# Bar 1-Visibility Representation of Optimal 1-Planar Graph

Mohammed Emtiaz Ahmed*, Asad Bin Yusuf , Md. Zahid Hasan Polin
Dept. of CSE, Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
*E-mail: a.emtiaz@yahoo.com

***Abstract*—In a visibility representation of a graph, the vertices map to objects in Euclidean space and the edges are determined by certain visibility relations. A bar visibility representation of a planar graph is a drawing where each vertex is drawn as a horizontal line segments called bars, each edge is drawn as a vertical line segment where the vertical line segment representing an edge must connect the horizontal line segments representing the end vertices. A graph is called a 1-planar graph if it can be drawn in the plane so that each its edge is crossed by at most one other edge. A 1-planar graph is said to be optimal if there are highest number of edges available. In this Research, we proposed an algorithm to numbering the optimal 1-planar graph and also bar 1-visibility representation of optimal 1-planar graph.**

***Keywords-1-planar graph,visibility representation,bar 1-visibility representation,optimal 1-planar graph.***

## I. INTRODUCTION

The problem of determining a visibility representation of a graph has been studied extensively in the literature due to the large number of applications (as in VLSI design, CASE tools, hidden-surface elimination problem, etc.) and, also, by the combinatorial properties of those graphs.

In a visibility representation of a graph, the vertices map to objects in Euclidean space and the edges are determined by certain visibility relations.

A graph is called a 1-planar graph if it can be drawn in the plane so that each its edge is crossed by at most one other edge. We can also say a 1-planar graph is a graph that has a1-planar drawing. For a 1-planar graph G, we have an inequality $m<=4n-8$ where n and m are its number of vertices and edges, respectively. A 1-planar graph G is said to be optimal if $m=4n-8$.

Since every planar graph has a 0-visibility representation, algorithms for finding 1-visibility representation of planar graphs are known. Thus the main idea is as follows: first obtain a planar graph by deleting some edges from the input non-planar graph, then obtain a visibility drawing of the planar graph and finally place the deleted edges which give bar 1-visibility representation.

The aim of this research is to study some properties of optimal 1-planar graph, developing such an algorithm that will produce bar 1-visibility representation for optimal 1-planar graph.

## II. APPLICATION OF VISIBILITY REPRESENTATION

The problem of computing a compact Visibility Representation is important not only in algorithmic graph theory, but also in practical applications such as VLSI layout [3] .From a visibility representation, a planar polyline drawing can be generated with O(1) bends per edge in linear time [16]. Visibility representations can also be used to generate planar orthogonal drawings. In this section, we present some of these applications in detail.

### *A. Planar Polyline Drawing*

We can construct a planar upward polyline drawing of a planar st-graph G using its visibility representation. We draw each vertex in an arbitrary point inside its vertex segment. We draw each edge (u, v) of G as a three segment polygonal chain.

### *B. VLSI Design*

Modules and their interconnections of a VLSI circuit are given as a graph where a vertex of the graph represents a module of the VLSI circuit and an edge represents an interconnection between two modules.

### *C. Motion Planing*

Motion planning is a problem of finding a path to move a robot from a start configuration to a goal configuration without colliding with obstacles.

### *D. Graph drawing*

Graph drawing is an area of mathematics and computer science combining methods from geometric graph theory and information visualization to derive two-dimensional depictions of graphs arising from applications such as social network analysis, cartography, and bioinformatics.

## III. HISTORY OF VISIBILITY REPRESENTATION

The problem of visibility representation has gained its own inherent interest in related topics and has significant researches based in its specific application and parameters. In this section, we give an outline of the results found in this area. Visibility representation has practical applications in VLSI layout [3] and several researchers concentrated their attention on visibility representations [3, 6]. Otten and VanWijk [17] have shown that every planar graph admits a

visibility representation and Tamassia & Tollis [21] have given a linear-time algorithm for constructing a visibility representation of a planar graph. Di Battista, Tamassia, R. and Tollis, I. G. have given Constrained Visibility Representation of Graphs [2].

Alice M. Dean et al. have introduced a generalization of visibility representation for a nonplanar graph which is called bar k-visibility representation [8] . While it is easy to see that all bar visibility graphs are planar, this is not true for bar k-visibility graphs, and no other immediate property provides an approach to their structure. Since all bar visibility graphs are planar, they seek measurements of closeness to planarity for bar k-visibility graphs. They have obtained an upper bound on the number of edges in a bar k-visibility graph. As a consequence, they have obtained an upper bound of 12 on the chromatic number of bar 1-visibility graphs, and a tight upper bound of 8 on the size of the largest complete bar 1-visibility graph. They also considered the thickness of bar *k*-visibility graphs, obtaining anupper bound of 4 when $k = 1$, and a bound that is quadratic in k for $k > 1$.

For the case $k = 1$, Dean et al. used the Four Color Theorem to show that their thickness is bounded by 4. They conjectured that no bar 1-visibility graph has thickness larger than 2.Recently, Fleshner and Massow have investigated some graph theoretic properties of bar 1-visibility graphs [10].They proved the tight upper bound on the thickness of bar 1-visibility graphs is 3. In recent years, several works are devoted to this field. Igor Fabrici,Tomas Madaras [14] study the existence of subgraphs of bounded degrees in 1-planar graphs which is also called bar 1-visibility graph. It is shown that each 1-planar graph contains a vertex of degree at most 7; they also prove that each 3-connected 1-planar graph contains an edge with both end vertices of degrees at most 20.

Peter Eades and Giuseppe Liotta study the relationship between RAC graphs and 1-planar graphs. Yusuke Suzuki studied the Optimal 1-planar graphs which triangulate other surfaces [24]. He proposed that A simple graph G is said to be 1-planar if it can be drawn on the sphere S2(or the plane) so that each of its edges crosses at most one other edge. For a 1-planar graph G, we have an inequality $m<=4n-8$ where n and m are its number of vertices and edges, respectively. (This inequality was proved in some papers. For example, see [14].) A 1-planar graph G is said to be optimal if $m=4n-8$. It had already proved in [23] that every optimal 1-planar graph G is obtained by adding a pair of crossing edges to each face of a 3-connected quadrangulation on the sphere; hence every vertex of G has even degree.

## IV. The Proposed Methodology for Bar 1-Visibility Representation of Optimal 1-Planar Graph

We first introduce our proposed algorithm based on Quadrangle Labeling of vertices of optimal 1-Planar graph. Later we would like to represent our proposed algorithm for finding the bar 1-visibility representation for Optimal 1-Planar graph

### *A. Quadrangle Labeling of Vertices of Optimal 1-Planar Graph*

A simple graph G is said to be 1-planar if it can be drawn on the sphere S2(or the plane) so that each of its edges crosses at most one other edge. The drawing is regarded as a continuous map f: G → S2 which may not be injective. To simplify our notation, we often consider that a given 1-planar graph G is already mapped on the sphere, and denote its image by G itself. An edge is said to be crossing if it crosses another edge in a 1-planar graph G, and to be non-crossing otherwise.

Quadrangle labeling is a way of numbering the vertices of a quad so that the highest and lowest numbered vertices are not in same diagonal.

From our proposed algorithm we have the following lemma:

**Lemma 1.** Every optimal 1-planar graph admits quadrangle labeling .

Proof : Let, G(V,E) be a input graph with maximum $4n-8$ edges where $n$= number of vertices. Now we will number its vertices. For this, at first we would select a vertex i=1 from the graph G which is outer from sphere surface as source. Then we will calculate n quadrangles which are connected to i vertex. Number all the vertices of quadrangles adjacent to the source, from left quadrangles to right quadrangles sequentially ensuring that lowest and highest numbered vertices of the quadrangle will not be in same diagonal and If Ep={Vi,Vj} is an edge which intersects with another edge, Eq={Vk,Vl} ; then assign number such that i<k<j<l. An example is shown below

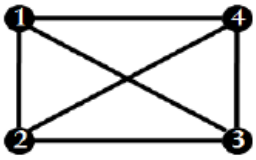



Figure 1: Quadrangle Labeling.

Same process will be repeated for i+1 to n. Then, for i=1 to n, Check if there is any diagonal edge consisting of lowest and highest numbered vertices of the quadrangle. If yes then swap the vertex number with any vertex number adjacent to it to ensure that lowest and highest numbered vertices of the quadrangle will not be in same diagonals . Then, go to the initial state, i=1 and repeat the step. If there is no diagonal edge consisting of lowest and highest numbered vertices of the quadrangle, increment i. Terminate when i=n.

The output of this algorithm would be such a numbering of vertices of G for which there will be no diagonal edge consisting of lowest and highest numbered vertices of the quadrangle. An example of this algorithm is shown below:

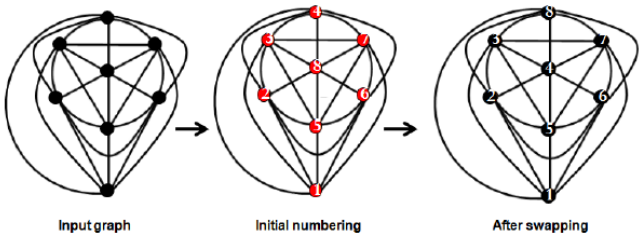



Figure 2: Labeling the vertices of an Optimal 1-Planar graph with n=8.

Algorithm:

Input : Optimal 1-planar graph.
Output : Quadrangle labeled optimal 1-planar graph.

- At first selected a vertex i=1 from the graph G which is outer from sphere surface as source.
- Calculate n quadrangles which are connected to i vertex.
- Number all the vertices of quadrangles adjacent to the source, from left quadrangles to right quadrangles sequentially ensuring that lowest and highest numbered vertices of the quadrangle will not be in same diagonals and If Ep={Vi,Vj} is an edge which intersects with another edge, Eq={Vk,Vl}; then assign number such that i<k<j<l.
- Repeat the process for i+1 to n.
- For i=1 to n Check if there is any diagonal edge consisting of lowest and highest numbered vertices of the quadrangle:
  - If yes then swap the vertex number with any vertex number adjacent to it to ensure that lowest and highest numbered vertices of the quadrangle will not be in same diagonals. Then i=1 and repeat the step.
  - If no then i=i+1.

### *B. Bar 1-Visibility Representation of Optimal 1-Planar Graph*

From quadrangle labeled optimal 1-planar graph we have the following Theorem:

**Theorem 1.** Every quadrangle labeled optimal 1-planar graph admits bar 1-visibility representation.

Proof : Let G is a quadrangle labeled optimal 1-planar graph. First, between the two intersecting edges of a quadrangle, we will remove the edge whose starting vertex number is higher than the starting vertex number of other edge in order to get a planar structure. This step is demonstrated in Figure 3(b).

Then, we convert the present graph into up-word planar graph as Figure 3(c).

Next, we will draw the dual graph of the upward planar graph as shown in Figure 3(d).

From the dual graph we got the Bar visibility representation using R. Tamassia, I. G.Tollis algorithm [1986]. Excepting the fact that ,starting and ending vertex u and v , draw the vertex -segment τ (u) at y-coordinate Y (u) and the vertex -segment τ (v) at y-coordinate Y (v) and between x-coordinates X(left(u)) and X(right(u)) and X(left(v)) and X(right(v)). This step is demonstrated in Figure 3(e).

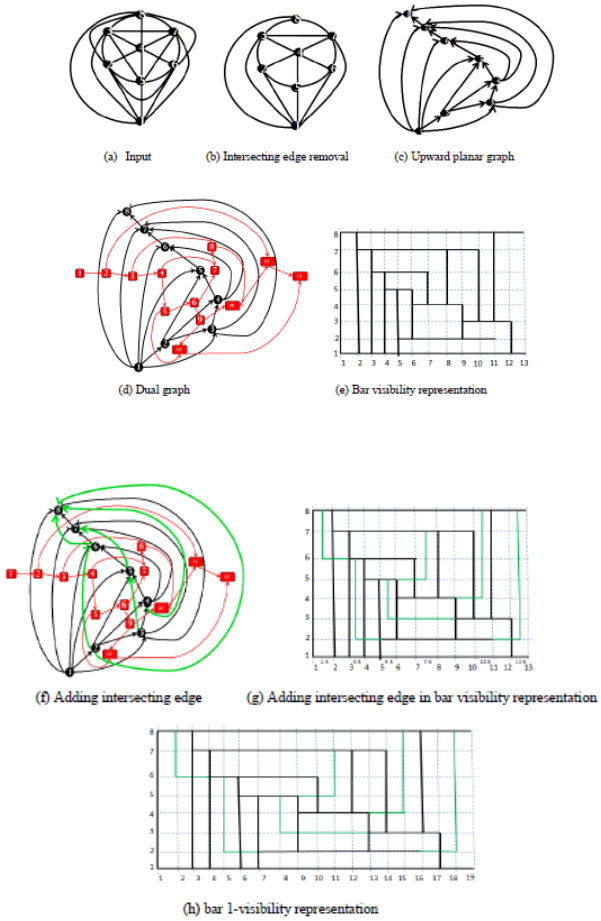



Figure 3: Bar 1-Visibility Representation of Quadrangle Labeled Optimal 1-Planar graph.

Then, added the intersecting edges which had been removed from the input graph to the up-word planar graph. The edge should be drawn on left or right of the point of the left region number, P. If there is more than one left region then choose the minimum one. We have to increase the intersecting bar either on Left or Right depending on the value of left region, P. If we want to increase the bar in left direction than the bar must have starting point at P-0.5 and inserted the edge at P-0.5. If we want to increase the bar in right direction than the bar must have ending point at P+0.5 and Insert the edge at

P+0.5. This step is demonstrated in Figure 3(f) and Figure 3(g).

Finally, if there is any edge drawn on a fraction point, draw it on the ceil value of the point and increase the sequential point by 1. Thus we get the final representation. This step is demonstrated in Figure 3(h).

Algorithm:

*Input*: Quadrangle labeled optimal 1-planar graph.
*Output*: Bar 1-visibility representation of optimal 1-planar graph.

- To make the graph planar; between the two intersecting edges of a quadrangle, remove the edge who's starting vertex number is higher than the other edge starting vertex number.
- Converting the present graph into up-word planar graph.
- Converting the up-word planar graph into its Dual graph.
- Bar visibility representation of the dual graph using R. Tamassia, I. G. Tollis algorithm[1986].Excepting the fact that ,starting and ending vertex u and v , draw the vertex -segment $\tau$ (u) at y-coordinate Y (u) and the vertex –segment $\tau$ (v) at y-coordinate Y (v) and between x-coordinates X(left(u)) and X(right(u)) and X(left(v)) and X(right(v)).
- Adding the intersecting edges which had been removed from the input graph to the up-word planar graph. The edge should be drawn on left or right of the point of the left region number, *P*. If there are more than one left region then choose the minimum one.
- We have to increase the intersecting bar either on Left or Right depending on the value of left region, *P*.
  1) If we want to increase the bar in left direction than the bar must have starting point at *P*-0.5.
  a) Insert the edge at *P*-0.5.
  2) If we want to increase the bar in right direction than the bar must have ending point at *P*+0.5.
  a) Insert the edge at *P*+0.5.
- If there is any edge drawn on a fraction point, draw it on the ceil value of the point and increase the sequential point by 1.

### C. Complexity of the Algorithm

The drawing we have proposed in this Algorithm to represent bar 1-visibility of an optimal 1-planar graph can be constructed in linear time. We have the following theorem:

**Theorem 2.** Let G be a quadrangle labeled optimal 1-planar graph. Then a Bar 1-visibility representation of G can be computed in linear time.

Proof: Let G be a quadrangle labeled optimal 1-planar graph. In step 1 and 2, we construct planar graph from optimal 1 planar graph by deleting one edge from each intersection and draw upward planar graph. Then we construct dual graph of the planar graph and compute topological numbering of the dual graph. The visibility representation of the planar graph can be constructed in linear time. Then shifting the vertex segment corresponding vertices at x coordinate, the deleted edges are placed in the representation in linear time. One edge intersects one vertex segment which maintains the properties of bar 1-visibility representation. Thus we can obtain a Bar 1-visibility representation of quadrangle labeled optimal 1-planar graph in linear time.

## V. CONCLUSIONS

In this paper, we have regarded the problem of computing algorithms on bar 1-Visibility Representations of non-planar graph such as optimal 1-planar graph. Given an undirected 1-planar graph G, this paper gives a drawing algorithm that gives bar 1-Visibility Representations. In this paper, we have addressed the problem of drawing Visibility Representation of non-planar graph. We have proved that some classes of non-planar graphs admit Bar 1-visibility representation. Consequently, we focused on those non planar graphs which contain at most one edge crossing. If given the directed graphs having diagonal labeling where each edge crossing bounded by a quadrangle and vertices of lowest number and highest number aren't on a diagonal then this paper also gives drawing algorithm for computing bar 1- visibility representation. So, finally we introduced a new algorithm for quadrangle labeling of optimal 1-planar graph. Based on this labeling we also represented a linear time algorithm for finding the bar 1-visibility representation of optimal 1-planar graph.

## REFERENCES


[1] S. Even A. Lempel and I. Cederbaum. An algorithm for planarity testing of graphs. P.Rosestiehl(ed.) Theory of Graphs: International Symposium July 1966, pages 215–232,1967.

[2] Di Battista, R. Tamassia, and I. G. Tollis.Constrained visibility representation of graphs. In Inform. Process. Letters, volume 41, pages 1–7,1992.

[3] G. Di Battista, P. Eades, R. Tamassia, and I. G. Tollis. Graph Drawing: Algorithms for the Visualization of Graphs. Prentice-Hall, Upper Saddle River, New Jersey, 1999.

[4] Prosenjit Bose, Alice M. Dean, Joan P. hutchinson, and Thomas Shermar. On rectangle visibility graphs. In Graph Drawing, pages 25–44, 1996.

[5] M. Chrobak and S. Nakano.Minimum-width grid drawings of plane graphs. International Journal of Computational Geometry and Applications, 11:29–54, 1998.

[6] F. J. Cobos, J. C. Dana, F. Hurtado, A.Marquez, and F. Mateos.On a Visibility Representation of Graphs, volume 1027 of Lecture Notes in Computer Science.Springer- Verlag, 1995.

[7] H. de Fraysseix, J. Pach, and R. Pollack.How to draw a planar graph on a grid.Combinatorica, 10:41–51, 1990.

[8] M. Dean, W. Evans, E. Gethner, J. D. Laison, and M. Safari. Bar k-visibility graphs. Journal of Graph Algorithms and Applications, 11(1):45–59, 2007.

[9] I. Fary. On straight line representations of planar graphs.Acta Sci. Math. Szeged, 11:229–233, 1948.

[10] S. FleshnerandM.Massow. Thickness of Bar 1-Visibility graphs, volume 4372 of Lecture Notes in Computer Science. Springer-Verlag, 2007.
[11] S. Fleshner and M. Massow.Parameters of bar k-visibility graphs. Journal of Graph Algorithms and Applications, 12(1):5–27, 2008.
[12] M. R. Garey and D. S. Johnson.Computers and Intractability: a Guide to the Theory of NP-completeness. W. H. Freeman and Company, New York, USA, 1979.
[13] Joan P. Hutchinson. Arc and circle visibility graphs. Australasian J. Combinatorics,25:241–262, 2002.
[14] T.Madaras I. Fabrici. The structure of 1-planar graphs. Journal on Discrete Mathematics, 307:854–865, 2007.
[15] T. Nishizeki and N. Chiba. Planar Graphs: Theory and Algorithms. North-Holland, Amsterdam, 1988.
[16] T. Nishizeki and M. S. Rahman.Planar Graph Drawing.World Scientific, Singapore, 2004.
[17] J. Otten and J. G. Van Wijk. Graph representation in interactive layout design. In the Proceedings of IEEE Intternational Symposium OnCircuits and Systems, pages 914–918, 1978.
[18] H. C. Purchase. Which aesthetic has the greatest effect on human understanding In the Proceedings of the 5th International Symposium on Graph Drawing (GD 1997), volume 2, pages 464–466, 1951.
[19] W. Schnyder. Embedding planar graphs on the grid. In Proc. of the First ACM-SIAM Symposium on Discrete Algorithms, pages 138–148, 1990.
[20] K. S. Stein. Convex maps. In Proc. of American Mathematical Society, volume 1353 of Lecture Notes in Computer Science, pages 248–261.Springer-Verlag, 1997.
[21] R. Tamassia and I. G Tollis.A unified approach to visibility representations of planar graphs. In Discrete and Computational geometry, volume 1, pages 321–341, 1986.
[22] D. B. West. Introduction to Graph Theory.Prentice-Hall, Upper Saddle River, New Jersey, USA, 2001.
[23] Y. Suzuki, Uniqueness of embeddings of optimal 1-planar graphs, preprint.
[24] Yusuke Suzuki .Optimal 1-planar graphs which triangulate other surfaces. Discrete Mathematics 310 (2010) 6_11. Tsuruoka National College of Technology, Tsuruoka, Yamagata 997-8511, Japan.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Automated Color Pencil Sketch Generation

AHM Mahfuzur Rahman[1,*], Tasmiha Salam [1]
[1] Samsung R&D Institute Bangladesh (SRBD), Bangladesh
*Email: mahfuz.rubel@samsung.com

*Abstract*— **We present a fully automated non-photorealistic rendering technique for color pencil sketch drawing. Because of a very simple stroke drawing technique, it is more efficient than previous state-of-the-art algorithms. In our method, we first segment the input image into foreground and background through an automated region segmentation procedure; then we draw hatching strokes all over the image and cross-hatching strokes only in the foreground region; the color to draw strokes are taken from the input image after mapping the color to one of the 12 colors of our color palette. Finally we overlay the stroke-drawn image with the input image to create an intermediate output and impose an edge detected image on that intermediate image to produce the final output. Through experimental results, we demonstrate the time efficiency and aesthetic pleasantness of our algorithm.**

***Keywords*— Non-photorealistic rendering, pencil sketch, color pencil sketch, edge detection, color mapping, overlay.**

## I. Introduction

Color Pencil Sketch drawing is a very pleasing artistic drawing technique; with only a limited amount of color pencil, artists are able to draw versatile types of pictures. In NPR (non-photorealistic rendering) research, creating color pencil (sketch) drawing can be a very intriguing challenge, but surprisingly meager amount of works has been done on automated or interactive color pencil sketch drawing to date.

Works on artistic image filtering, in general, is going on for several decades. Many researchers are working on image filters like WaterColor [1], Pastel Drawing [2], Color Pencil Sketch [3]–[8] but the results from existing techniques still lacks from visual quality and efficiency. Most of the works on pencil and color pencil drawing underpins on Linear Integral Convolution (LIC) [9]. A lot of works in this area, especially which focuses on color pencil sketch, also depend appreciably on the segmentation of image into different regions, which are then painted with different colors; the segmentation is not automated and depends considerably on user interactions for good performance.

### A. Problem Statement

Given an image (color or gray-scale), the task is to convert it to a color pencil sketch drawing. We have divided the task into several sub-tasks:

1) **Region Division:** Extraction of foreground and background regions from the image.
2) **Color Selection:** Selection of colors from color palette to draw strokes.
3) **Stroke Generation:** Generation of different type of strokes to draw onto the image.
4) **Output Generation:** Blending the stroke-drawn image and the original image and perform further post-processing.

### B. Related Works

There are not much works in literature which focuses on color pencil drawing or color pencil sketch drawing. We can divide the works that have been done so far mainly on two categories: works that depend on LIC, and those do not.

Of the few works that has been done on color pencil sketch, most of the works depend, in one way or other, on LIC [9]. With the help of a vector field and noise addition in image, LIC can be used to create pencil stroke like effect; applying this method on three channels (RGB) of a color image, it is also possible to create color strokes in images. In their papers, Yang and Min [3] and Yang et. al [4] used modified LIC for stroke generation. Yang and Min applied a modified LIC filter named swing bilateral LIC (SBL) filter and then produced monochrome pencil drawing images; Yang et. al, on the other hand, first find out feature lines and segment the image according to those lines; they then generate noises in CMY color space and implement two types of strokes: feature strokes (for subject) and hatching strokes (for background). In the work of Matsui et al. [7] which precedents the works of Yang and Min, they proposed a stroke-based algorithm in which they place strokes along the boundary of the curves of an image using 12 colored pencils. Another work regarding LIC is done by Yamato et al. [5], where they after segmenting the source image into several regions, choose two colors for each region, generate two noise images with those colors, apply LIC on both images and blend those images with Kubelka Munk (KM) Model [10] to create the final color sketch image.

Of the works that do not depend on LIC, Lu et al. [8] based their work on pencil stroke generation and tone drawing; they created a stroke image similar to edge detected image and transferred the tone of the image to give the output an abstracted look. The other and one of the most popular works on color pencil sketch generation, is covered by [6]; this attempt is a bit different from all others as it does not incorporate any kind of stroke to the image, rather segments the image into several regions, shrinks the image boundary and shifts the color of the image to match to their color database, and fills the regions chosen from some arbitrary colors.

(a)

(b)

Figure 1: (1a): Sketch of A natural scenario; used strokes - hatching (in different angles) and scumbling (in tress) (1b) : Sketch of apple; left: hatching, middle: cross-hatching, right: scumbling.


Figure 2: Common Stroke Types [11].

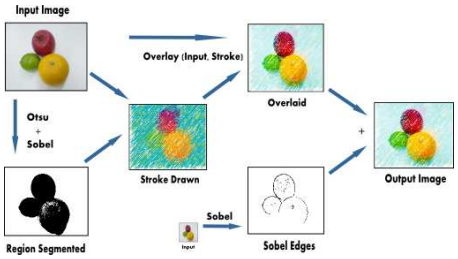



Figure 3: Process Overview.

### C. *Overview of the System*

In drawing color pencil sketches, the artists after making the rough outline of the objects, fill out different regions with different types of colors and strokes. The strokes used by color pencil artists include hatching(line drawing), cross-hatching, stippling, scumbling etc. With the employment of different colors and strokes one over another, the artists create different shades and complex colors from a limited number of color pencils. We have shown the basic strokes (Fig. 2) and their deployment in real color pencil sketch drawing in Fig. 1.

In our proposed system, we have tried to emulate the drawing styles of the artists. First we try to divide the image into two different regions: foreground and background; as there is no system of automatic segmentation of image into foreground and background, we have employed an approximated mechanism. We use Otsu thresholding [12] and Sobel edge detection [13] to produce two intermediate binary images and combine them to produce a segmented image. Depending on the foreground and background regions in segmented image, we draw two different kinds of strokes: hatching and cross-hatching. We cover the background region only with hatching and the foreground region is a mixture of cross-hatching over hatching. The color needed to draw the strokes are selected from the image color, which is converted to one of 12 colors from our color palette before drawing of strokes; this color mapping process gives us necessary color quantization of image that we observe in real color pencil sketch drawings. After the generation of stroke drawn image, input image and currently obtained stroke drawn image are overlaid [14] for intermediate output image. Finally we take the weighted average of the Sobel edge detected image and the intermediate output image to obtain the final output image. The whole process is described in Fig. 3.

The rest of the paper is organized as follows. In section II, we describe the schemes for automated foreground and background detection. Section III proposes the color selection method for drawing strokes. In section IV, we present methods for stroke generation. The main algorithm is described in section V which depends on the processes elaborated in section II, III and IV . Performance evaluation and comparison with previous works and conclusions are presented in section VI and section VII, respectively.

## II. Region Division

For any input image, we first convert it into a grayscale image. Then we deploy two algorithms on the input image:

1) **Thresholding**: Implement Otsu thresholding algorithm [12] to break the image into two regions.
2) **Edge Detection** : Edge detection using Sobel's Edge detection [13] operators.

The output of both of the procedures give us a binary image; we consider the black portion of the image from Otsu thresholding as the foreground of the image; the black (edge) regions from Sobel operator are considered the most prominent features (feature strokes) within the image. We then add both images using an OR operator assuming black region as 1 and white region as 0, which gives us an approximately separated foreground-background regions as final output: black as foreground and white as background; the algorithm used for region segmentation is given in algorithm 1.

We generate and use Sobel edge detected region at section V again, when we blend edge detected image

**Algorithm 1** $RegionSegmentation(I, I_{Otsu}, I_{Sobel})$

1: Create a binary image $I_{Otsu}$ using Otsu thresholding method.
2: Create a binary image $I_{Sobel}$ using Sobel Edge detection with threshold parameter 16.
3: Apply OR operation between $I_{Otsu}$ and $I_{Sobel}$ to generate final segmented image $I_{Seg}$.
4: **return** $I_{Seg}$

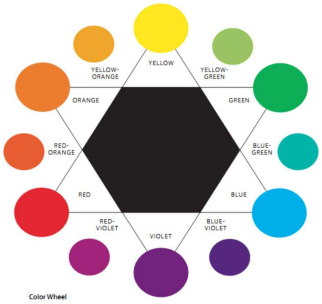



Figure 4: Most used basic colors by the artists.

with the overlaid image (between sketch and input). The parameter used as threshold for Sobel is different in those cases; in the first case, we want as much foreground as we can cover which leads us to assign a low threshold value of 16 to get a lot of edges (sometimes non-prominent edges also); in the second case though, we only want significantly prominent edges and therefore increase our threshold to a more restrictive 100; both of the values are determined empirically.

## III. Color Selection

Before drawing a stroke, we decide the color with which the stroke is going to be drawn; two things that we consider mostly are:

1) **Color Reduction:** To make a drawing, artists only have a limited number of colors at their disposal which motivated our decision to keep the number of colors restricted. In this paper, we have mapped almost all of the colors in an input image to the 12 basic colors artists use for drawing (Fig. 4) according to [15].
2) **Color Mapping:** To preserve color harmony between input and output image, in spite of choosing random colors for mapping, we take the closest color, considering color difference between the color of input image (at the center point) and the constituent colors of our intended color palette, taking each at a time.

### A. Color Mapping

This is an inevitable method to select colors before drawing any stroke. According to our color reduction consideration, we have already chosen 12 colors to map the whole image to. Now another important thing is how

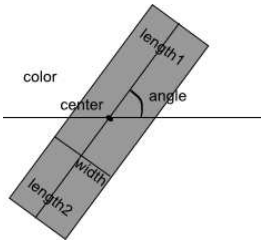



Figure 5: Each brush stroke is drawn with the properties shown in the image: an anchor point, suitably in middle of the length two length variables, width, angle and color. Color is chosen according to section III

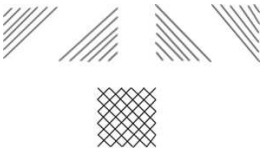


Figure 6: Top row shows four half squares drawn each time; combined together, they will give a cross-hatching output like the one shown in the bottom row.

should we map varieties of colors from the original image to only 12 colors. For that, We consider to measure color differences in three different ways:

1) Square root difference in RGB color space.
2) Color conversion to CIE Lab Color Space and finding difference using CIE 1974 difference [16].
3) Color conversion to CIE Lab Color Space and finding difference using CIE 1994 difference [17].

We have used the CIE 1994 color difference for our algorithm as the output from this one looks better than the other two difference mechanisms; it is less time-efficient than the other options though. To get a better visual quality, in spite of replacing the stroke color with the nearest color, we have taken 3 closest colors according to color differences and chosen one randomly from those three before choosing the final color to draw stroke. Also for color mixing, when we have to redraw a pixel, we are taking an average color between the previous color of that pixel and the new one presently obtained.

## IV. Stroke Generation

In original color pencil sketch drawing, the artists use different kind of strokes as shown in section I-C ; of those, the most commonly used strokes are hatching and cross-hatching. Therefore we have implemented hatching and cross-hatching strokes in our algorithm; the basic stroke generation process is shown in Fig. 5.

### A. Hatching

To implement hatching style strokes, we have to draw strokes of different lengths in parallel; for that, we take as parameter center (anchor) point, two lengths, width, color and angle of the stroke we are going to draw. A stroke of $width = 1$ is essentially a single line with the predefined angle, and a stroke of $width > 1$ is simply a

Figure 7: Sample Input Images (In-House images from SRBD)

width amount of lines drawn side by side. Fig. 5 illustrates the parameters necessary to draw a stroke and algorithm 2 gives the procedure to draw a stroke where, $SP$ = starting point, $L$ = length, $W$ = width (preferably odd, and $>= 3$ ), $C$ = color (RGB), $A$ = angle.

**Algorithm 2** $Hatching(I, SP, L, W, C, A)$

1: Taking the starting point as the middle element of width, we create width amount of starting points
2: **for** each starting point in starting points **do**
3: Draw a line with length $L$, angle $A$ and color $C$
4: **end for**

### B. Cross-Hatching

We draw cross-hatching strokes always within a square block whose size is determined beforehand. To generate cross-hatching within a block we need to draw four times, differently, each time half of a square as shown in Fig. 6; the process is also described in algorithm 3.

**Algorithm 3** $CrossHatching(I, Start, BlockSize)$

1: Divide the block into 4 sub-regions (4 half squares)
2: **for** each half square regions **do**
3: Generate a set of middle points for strokes, $SPoints$
4: Calculate length $L$ and width $W$ to draw hatching strokes.
5: **for** each starting point $SP$ in $SPoints$ **do**
6: find three closest colors to the color at $SP$ from the color Palette and randomly choose a color $C$.
7: $Hatching(SP, L, W, C, A)$
8: **end for**
9: **end for**

## V. Output Generation

Using the methods described in section II, III, IV, we are now going to devise the complete algorithm for color sketch generation. Before using these algorithms, we do some pre-processing steps as described in algorithm 4; then after going through the main algorithms, we perform some post-processing as described in algorithm 5; the entire process is described in algorithm 6.

Fig. 8 shows various color pencil drawing results for the input images provided in Fig. 7; we have chosen different images covering disparate subjects: natural scenarios, human, artificial objects, cartoon like drawing, real painting from artists, gray-scale images etc.

**Algorithm 4** $PreProcessing(I, I_{Seg})$

1: Prepare a color palette of 12 colors and convert those to Lab color space.
2: Generate two variables, NumOfHatchingPoints and NumOfCrossPoints, depending on the height and width of input image.
3: Generate a set of points $SP_{Hatching}$ of size NumOfHatchingPoints randomly from the image
4: Generate a set of points $SP_{Cross}$ of size NumOfCrossPoints randomly from the black (1) portion of $I_{Seg}$
5: **return** $SP_{Hatching}$, $SP_{Cross}$

**Algorithm 5** $PostProcessing(I, I_{Stroke})$

1: Overlay $I_{Stroke}$ and $I$ and keep the output in $I_{Out}$
2: Create Sobel edge detected image $I_{Sobel}$ using threshold 100
3: **for** each point $SP$ from the black pixels of $I_{Sobel}$ **do**
4: Set $I_{Out}$ as the average of color from $I_{Out}$ and $I_{Sobel}$.
5: **end for**
6: **return** $I_{Out}$

**Algorithm 6** $ColorSketchDrawing(I, L, W, BlockSize)$

1: Create images: $I_{Stroke}$, $I_{Otsu}$, $I_{Sobel}$, $I_{Seg}$ $I_{Out}$
2: $I_{Seg} = RegionSegmentation(I, I_{Otsu}, I_{Sobel})$
3: $SP_{Hatching}, SP_{Cross} = PreProcessing(I, I_{Seg})$

4: **for** each point $SP$ in $SP_{Hatching}$ **do**
5: find three closest colors from the color Palette to $SP$ and randomly choose a color $C$.
6: $Hatching(I_{Stroke}, SP, L, W, C, 45)$
7: **end for**

8: **for** each point $SP$ in $SP_{Cross}$ **do**
9: $CrossHatching(I_{Stroke}, Start, BlockSize)$
10: **end for**

11: $I_{Out} = PostProcessing(I, I_{Stroke})$
12: **return** $I_{Out}$

Figure 8: Output of Color Pencil Sketch

TABLE I.: Time Requirement of Our Algorithm

| Image | Image Size | Run time |
|---|---|---|
| Restaurant | 1664 * 2496 | 5 s |
| Girl | 900 * 1280 | 1 s |
| Two Boys | 831 * 1247 | 1 s |
| Fruits | 407 * 541 | 0.25 s |
| Barbie | 1350 * 1800 | 2.2 s |
| Hepbern | 1600 * 1265 | 2.1 s |

TABLE II.: Time Comparison

| Algorithm | Image Size | Time | |
|---|---|---|---|
| [3] | Pent. QuadCore | 872 * 847 | 33.69 s |
| [8] | Pent. Core i3 | 600 * 600 | 2 s |
| [6] | Pent. IV | 481 * 321 | 20 s |
| [7] | Pent. IV | 435 * 535 | 45 s |
| [5] | Pent. IV | 640 * 480 | 600 s |
| Our Algorithm | Pent. IV | 900 * 1280 | 1 s |

## VI. Performance Comparison

The performance of our algorithm for images of different resolutions are shown in Table I which was generated using Pentium IV processor and developed in C.

The output of our process is faster than previous algorithms suggested in the literature. A run-time comparison with other methods is given in Table II; in this table, we have compared the running time of [3], [8], [6], [7] and [5] with our proposed algorithm. As we don't have all the source codes available, we ran our algorithm with the lowest CPU configuration and on an image of higher resolution than all of the algorithms we compared with.

As an important goal of our work was to reduce the time complexity of color pencil sketch effect with insignificant reduction in output quality, we had to make some trade-offs. We used Sobel edge detection as it is faster than other popular edge detection methods, but if we give less concentration on time complexity, we can think of applying better edge detection techniques like Canny edge detection [18], Flow Based Difference of Gaussian (FDOG) filter [19] etc.

A subjective (visual) comparison of our algorithm and some of the previous state of the art algorithms are provided in Figure 9. Our method introduces cross-hatching technique first in color pencil sketch generation; the effect is clearly visible in the 3rd image (fruits) of Figure 9. Our color pencil sketch effect can be made even better in future, if strokes of different shapes and directions are considered. We may try to use other stroke types mentioned in section I-C. Also it is also possible to make the quality of generated strokes better by keeping library of strokes obtained from real artists as done in [20].

## VII. Conclusion

In this paper, we have proposed a fully automated color pencil sketch drawing system; our method includes region segmentation, two types of color stroke generation and mapping the colors of original image to 12 predefined colors; to keep the finer details, we have also kept the edges obtained from the original image.

Though our method is automated, with the help of user interaction, the segmentation could have been done almost realistic and the visual quality of the output can be made significantly better. Also, we can extend our work to provide color sketch effect in videos where we also need to consider temporal coherence across different frames of a video. In addition to color pencil sketch, with some changes in stroke drawing, the methods we have provided here can also be used to generate crayon, pastel etc. based drawing effects.

## References


[1] C. J. Curtis, S. E. Anderson, J. E. Seims, K. W. Fleischer, and D. H. Salesin, "Computer-generated watercolor," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 1997, pp. 421–430.

[2] K. Murakami and R. Tsuruno, "Pastel-like rendering considering the properties of pigments and support medium," in *ACM SIGGRAPH 2002 conference abstracts and applications*. ACM, 2002, pp. 227–227.

[3] H. Yang and K. Min, "Color pencil rendering of photographs," in *Convergence and Hybrid Information Technology*. Springer, 2012, pp. 302–308.

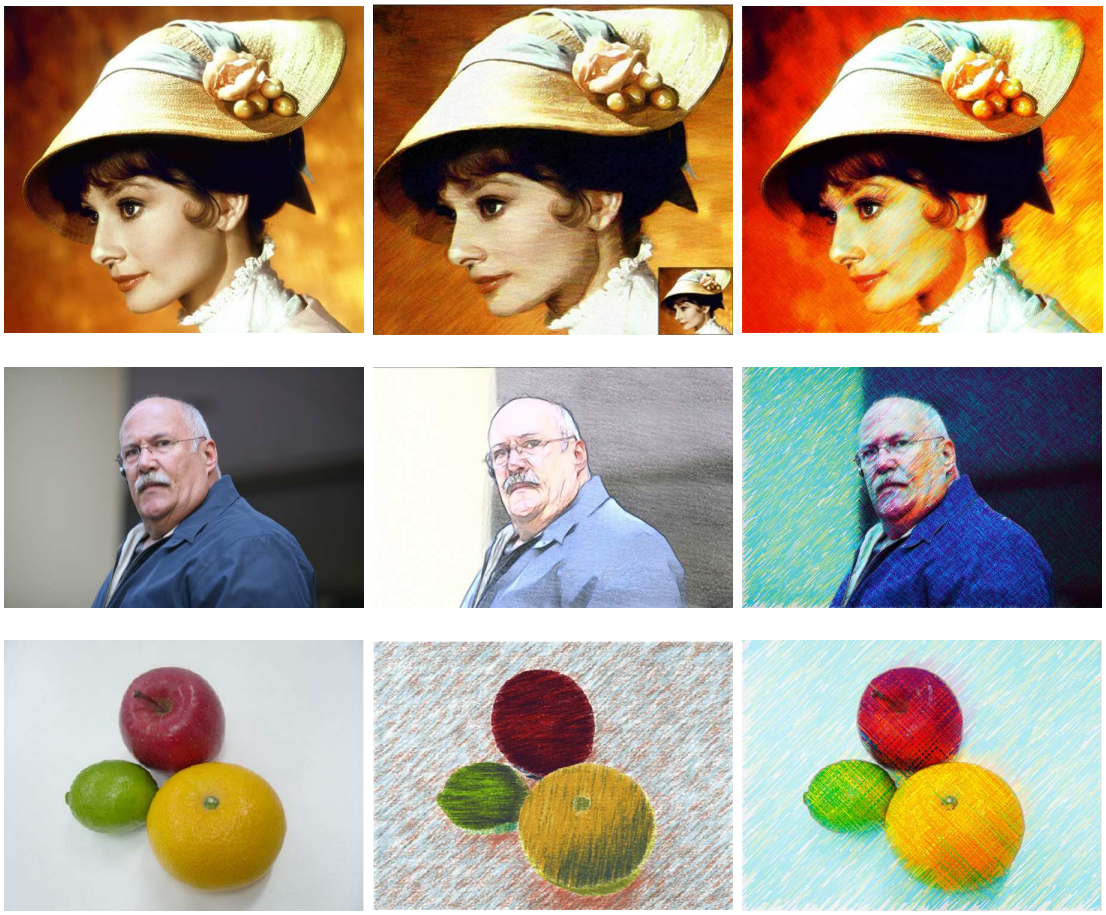


Figure 9: left: input images from [3], [8] and [5] respectively, middle: corresponding output images from the papers, right: output of our algorithm


[4] H. Yang, Y. Kwon, and K. Min, "A stylized approach for pencil drawing from photographs," in *Computer Graphics Forum*, vol. 31, no. 4. Wiley Online Library, 2012, pp. 1471–1480.

[5] S. Yamamoto, X. Mao, and A. Imamiya, "Colored pencil filter with custom colors," in *Computer Graphics and Applications, 2004. PG 2004. Proceedings. 12th Pacific Conference on*. IEEE, 2004, pp. 329–338.

[6] F. Wen, Q. Luan, L. Liang, Y.-Q. Xu, and H.-Y. Shum, "Color sketch generation," in *Proceedings of the 4th international symposium on Non-photorealistic animation and rendering*. ACM, 2006, pp. 47–54.

[7] H. Matsui, H. Johan, and T. Nishita, "Creating colored pencil style images by drawing strokes based on boundaries of regions," in *Computer Graphics International 2005*. IEEE, 2005, pp. 148–155.

[8] C. Lu, L. Xu, and J. Jia, "Combining sketch and tone for pencil drawing production," in *Proceedings of the Symposium on Non-Photorealistic Animation and Rendering*. Eurographics Association, 2012, pp. 65–73.

[9] B. Cabral and L. C. Leedom, "Imaging vector fields using line integral convolution," in *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*. ACM, 1993, pp. 263–270.

[10] C. S. Haase and G. W. Meyer, "Modeling pigmented materials for realistic image synthesis," *ACM Transactions on Graphics (TOG)*, vol. 11, no. 4, pp. 305–335, 1992.

[11] T. McArdle, "Colored pencil instruction). [online]. available: http://www.art-is-fun.com/colored-pencil-instruction.html," 2008.

[12] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[13] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.

[14] P. Software, "Overlay mode. [online]. available: http://www.pegtop.net/delphi/articles/blendmodes/overlay.htm," 2006.

[15] L. Hammond, *Acrylic Painting with Lee Hammond*. North Light Books, 2006.

[16] K. McLaren, "Xiiithe development of the cie 1976 (l* a* b*) uniform colour space and colour-difference formula," *Journal of the Society of Dyers and Colourists*, vol. 92, no. 9, pp. 338–341, 1976.

[17] B. Lindbloom, "Delta e (cie 1994). [online]. available: http://www.brucelindbloom.com/index.html," 2006.

[18] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6, pp. 679–698, 1986.

[19] H. Kang, S. Lee, and C. K. Chui, "Flow-based image abstraction," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 15, no. 1, pp. 62–76, 2009.

[20] K. Zeng, M. Zhao, C. Xiong, and S.-C. Zhu, "From image parsing to painterly rendering," *ACM Trans. Graph*, vol. 29, no. 1, p. 2, 2009.

# Variable Dependency Analysis of a Computer Program

Muhammad Sheikh Sadi, Linkan Halder and Seemanta Saha*
Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh.
* E-mail: seemantasaha@gmail.com

*Abstract*-**To keep pace with the advancement of technology, software products are overlooking the chances of soft errors in the program. Program analysis to check program execution flow is an effective way to detect soft errors. High level language computer program execution can be analyzed on the basis of the dependence of the variables used in the program. This paper illustrates a novel method to analyze variable dependencies of program based on automated generation of dependence graph. Dependence graph depicts the connectivity between the program variables where variables works as the vertices and dependence between variables perform as edges. The automated generated dependence graph also discovers the critical variables of a program and these critical variables perform as the key to detect the sequential execution of a program outperforming the existing methods of program analysis.**

***Keywords- Critical variables, variable dependence, dependence graph, program execution flow, soft error.***

## I. INTRODUCTION

Software programs need to be more reliable and error free with the advancement of technology. But, the speed at which technologies are being advanced, error rate is proportionally increasing too. So, there is no chance to avoid analyzing the code to check program execution flow and program data consistency. Developing a technology which can add checkers to the program steps early in the design process will be very much effective.

The method proposed in this paper is automated variable dependence analysis [11] [12] which can be used to analyze the flow of a computer program and work as an error checker. The method introduced in this paper adds a new dimension to the variable Dependence analysis [12] [13] by adding dynamic and automatic tracing of the critical variables [15] instead of all effecting variables.

High level language computer program consists of variables, functions, procedures, structures etc. Variables are the important component of a program which has the effect of total program flow. A program can be generalized using three main parts. These are input, output and the program execution based on an algorithm. Input, output is defined using variables. Variables are also used to maintain program processing. All the variables which used in the program are connected to each other directly or indirectly. Some variables can be also in the program which has no effect on any other variables for a particular algorithm. These variables can be termed as isolated variables. Dependence of variables can also be classified as data dependence and control dependence of variables. In this paper proposed analysis is based on the data Dependence of variables [12]. If the data value of variable x depends on data value of y then x has data dependence to y. A variable can be dependent on one or more than one variable. The variables which have a Dependence on other variables is called the precedence variable to the others. Those variables by which we can check the data value effect of the other variables of the total code (most precedence variables) as well as the isolated variables are summed as the critical variables [15]. The term 'Critical variable' in this paper denotes the variable which holds data dependency of all other variables in a program. Those critical variables can be used to detect soft error in a system with less time consumption as the effects of other variables are centralized to the critical variables. Critical variables will be traced out using the auto-generated dependence graph. As a result, the proposed method of variable Dependence analysis will be effective to the field of program analysis as program flow can be determined automatically.

The paper is structured as follows: Section II describes related works. Section III presents the proposed methodology to analyze variable dependence of computer programs. Section IV presents experimental analysis. Section V summarizes the contributions of this paper and discusses future possibilities.

## II. RELATED WORKS

A good number of works have been performed in the field of variable dependence analysis to analyze codes. All these works have introduced diversity of methods like program slicing, procedural approach, transformational approach, program rewriting etc. Among these methods, program slicing has been the most of interest for researchers. Program slicing is a technique which analyzes the variables and their dependence of a computer program written in high level language. Program slicing [1] [2] consists of identifying the parts of a program that can potentially affect the values of a

chosen set of variables. Program slicing also introduces various ways of methods like inter-procedural program slicing [3], dynamic program slicing [4]-[6], forward and backward slicing [7], program slicing based on dependence graph [8] [9]. Two types of slicing introduced previously are forward slicing and backward slicing. Both forward slicing and backward slicing consists of all program points that are affected by a given point in the program. According to the slicing techniques, a program point has been set by the programmer and then the dependence is analyzed. The programmers were responsible to find out the program point or precedence variables manually. Any mistake taken by programmers would change the output remarkably. Daniel Jackson et al. [10] introduced another method of program dependences for reverse engineering of software products. K. Muthukumar et al. [11] introduced abstract domain independent fix point algorithm using abstract interpretation to derive variable Dependence in compile time. M. Harman et al. [12] proposed to rewrite program code in an intermediate core language using a subset of the C programming language. They also introduced transformational approach of variable Dependence analysis [13] which also known as VADA where they used the relationship between variable Dependence and slicing. Alfred V. Aho mentioned some techniques based on compiler tools which has been very much useful in the field of code analysis [14]. The idea of critical code blocks and critical variables is also introduced [15] to mitigate soft errors in software programs. But, automation of detecting precedence and critical variables was never introduced for code analysis which can pave the way of dynamic detection of errors in program execution flow. The proposed methodology finds out the precedence variables automatically by generating dependence graph and detects the critical variables.

## III. PROPOSED METHODOLOGY

The proposed method illustrates the analysis of variable dependence which will be achieved by generating a dependence graph of the program variables. An algorithm is also mentioned later in this paper which will generate the dependence graph automatically based on the data dependence of variables. A computer program will be passed as input and the proposed method will generate the precedence graph as output with identifying the critical variables. And this process consists of several steps.

### *A. Parsing the Computer Program and Detecting Variables*

As mentioned earlier in this paper, a computer program consists of more than one variable. It is very easy to detect the variables and their dependence in a small computer program manually. But, automatic and dynamic detection of variables from a high level language computer program needs an efficient parser. So, a parser has been developed which detects all the variables from the program and saves in the memory. The first stage is to trace all variables and save those variables. Fig.1 portrays a very simple C program.

```
int a=4,c=1,p=6;
int x=10,q,w=40;
int y=x+c;
int z=y+a+1;
int d=5;
 q=w+z;
```

Figure 1. A Simple C Program

Here the variables are ‘a’, ‘c’, ‘p’, ‘x’, ‘q’, ‘w’, ‘y’, ‘z’ and ‘d’. These variables are filtered by parsing the simple C program of Fig.1. A generalized parser is developed which can trace out all the variables of any types of C program based on regular expression and C program language syntaxes.

### *B. Dependence Analysis*

The main target of this step is to find precedence and isolated variables and finally the critical variables. To achieve this, an algorithm is also proposed which analyzes dependence between variables and in consequence traces out the precedence variables. Fig.2 depicts the pseudo code to find the precedence variables.

```
[Definition:] variables_array=set of variables, right_variable= variable
right side of assignment operator, right=set of right_variable,
Variable_status=type of variable.
[initialize:] variables_array=NULL, variable,
right_variable=NULL,right=NULL,Variable_status=NULL.
Begin
        Begin
            Each statement of a program
            Assign variable to variables_array
            Assign right_variable to right
            End loop
        Begin each variable of variables_array
variable not in right and has no  Dependence
            Variable_status:=isolated
            Else Variable_status:=precedence
        End loop
End
```

Figure 2. Pseudo Code to Find Precedence

Dependence is the term defined as, if one or more variable’s data with numeric operation or without numeric operation is assigned to another variable, then the last variable is said to be dependent on previous variable(s). Mathematically, dependency of variables can be described using “Transitive Relation”. In mathematics, a relation ***R*** over a set ***X*** is transitive if whenever an element ***a*** is related to an element ***b***, and ***b*** is in turn related to an element ***c***, then ***a*** is

also related to ***c.*** In mathematical syntax:- $\forall \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \boldsymbol{X}: (\boldsymbol{aRb} \wedge \boldsymbol{bRc})$.

For our proposed method, if **b** is dependent on **c** and **a** is dependent on b then **a** is also dependent on **c**. Here, **b and c** are intermediate variables and **a** is precedence variable. The effect of **b** and **c** in the program can be detected from the variable **a**. From Fig.1 it is clear that variable 'y' is dependent on variable 'x' and 'c'. Variable 'z' is dependent on 'y' and 'a'. And so on.

### C. *Generating Dependence Graph*

This step is the most vital step which automatically generates a dependence graph. Data dependence of variables is used to generate this graph as described earlier in the paper. The variable which is in the left side of an assignment operator is dependent on the right side variable(s) of assignment operator. If any variable has numeric value on the right side of assignment operator or no other variable is dependent on this variable, the variable is defined as isolated variable. In general, an isolated variable has no dependence on other variable(s) and other variable(s) has no dependence on isolated variables. In Fig. 1, d and p are isolated variables. The vertex of dependence graph is the variables of the code segment. And the edge is directed from right sided variables of assignment operator to the left sided variable. As a result, step by step proceeding of the program execution builds a graph which finally takes a form of dependence graph. The algorithm to generate dependence graph is shown in Fig. 3.

[Definition:] ***variable_array***= set of all ***variables***, ***parent***=first ***variable*** of ***variable_array***, ***child***= right variable of assignment operator, ***new_parent***= a variable in ***variable_array***
Begin
  Draw ***parent***
  Until all ***variable*** of ***variable_array*** drawn
    If ***variable*** has child
      For each ***child*** of ***variable***
      Draw ***child***
      draw edge from ***child*** to ***parent***
      ***End loop***
      ***Parent***:=***child***
    Else if ***variable*** in ***child*** of ***new_parent***
      ***child***:=***variable***, ***parent***:=***new_parent***
      draw edge from ***child*** to ***parent***
    Else draw ***variable***
  End loop
End

Figure 3. Pseudo Code to Generate Dependence Graph

Now, applying the above mentioned algorithm of Fig.3 on the simple program of Fig.1, we will finally get a dependence graph which depicted Fig. 4.

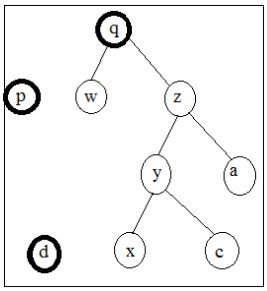



Figure 4. Dependence Graph

### D. *Precedence and Critical Variables*

Here, y is precedence to x and c as value of y is dependent on x and c, z is precedence to y and a as value of z is dependent on y and a, and q is precedence to w and z as value of q is dependent on w and z. As, q is the variable which precedes the variables x, c, z, y, a, w, it is detected as one of the critical variable. Variable d and p are also detected as critical as they are isolated. Rests of the variables are neglected for the method as their effect can be achieved from the critical variables of the program.

## IV. EXPERIMENTAL ANALYSIS

The proposed method for variable dependency analysis enlightens the idea of automatic detection of precedence and critical variables generating a dependence graph from a computer program.

### A. *Experimental Setup*

A software tool has been developed in C# language and using GDI in Microsoft Visual Studio 10.0. This tool takes input a C program and generates a precedence graph and finally detects the precedence and critical variables. The graph is built by a step by step process parsing the total C program line by line.

### B. *Preprocessing*

When a simple C program is input to the tool it filters out the variables and discards the unnecessary line to find the precedence variables. A lexical analyzer is introduced which detects all the variables in a C program. This is done by using building regular expression for C program syntax. So, the C program is either a 100 line program or a professional software code, the tool will filter out all the variables.

### C. *Classifying Variables*

After preprocessing the tool works only with the specific program lines containing variables. Fig. 5 depicts a program portion after doing the preprocessing of a simple C program.

After that, applying the algorithm of Fig. 2 on the program portion of Fig.5 variables will be classified to three

types of variables. These variables are precedence variables, isolated variables and intermediate variables. Table I shows the variable classification.

```
int a,b,c,x=5,y,z;    //line 1
b=10;                 //line 2
c=20;                 //line 3
a=b+c;                //line 4
y=x+a;                //line 5
z=y+30;               // line 6
int p=2,q,r=4;        // line 7
q=p*p;                // line 8
```

Figure 5. Program Portion after Preprocessing

TABLE I VARIABLE CLASSIFICATION

| Line number | variable | Dependent on | Status of variable |
|---|---|---|---|
| 2 | 'b' | -- | intermediate |
| 3 | 'c' | -- | intermediate |
| 4 | 'a' | 'b', 'c' | intermediate |
| 5 | 'y' | 'a', 'x' | intermediate |
| 6 | 'z' | 'y' | precedence |
| 7 | 'r' | -- | isolated |
| 8 | 'q' | 'p' | precedence |

### *D. Automatic Dependence Graph Generation*

After classifying the variables they are stored in data structure. An array holds all the variables and another array holds the index of the connectivity variables. When a variable is discovered it also shows the other variables if it has connectivity to those other variables. So, the precedence graph is gradually generated as the program parsing proceeds. When the program portion ends the precedence graph is completely generated. Those variables which have no connection with other variables are kept in another structure as isolated variables. Finally the most preceding variables and isolated variables results as critical variables. These critical variables define the effects of other variables in program flow execution as well as can be used to detect program steps and soft errors if there is any. Fig.6 illustrates the gradual building of precedence graph through line 2 to line 8. The final graph traces out the critical variables.

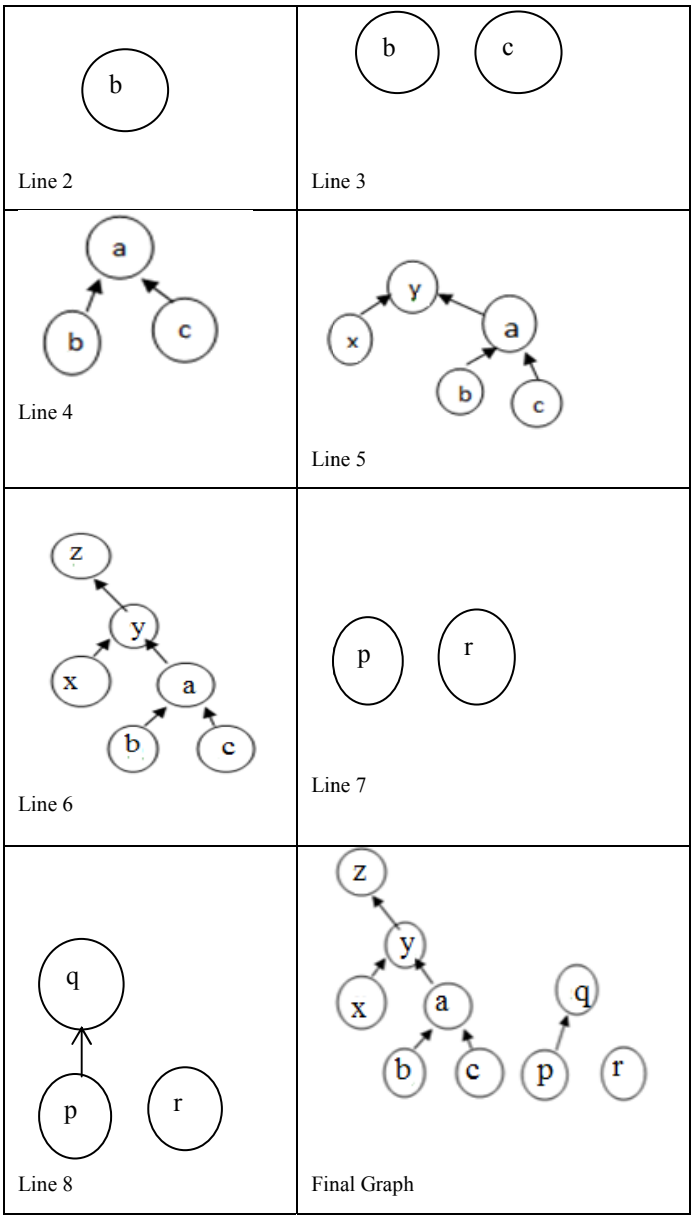



Figure 6. Automated Graph Generation (step by step)

### *E. Analysis Results*

The proposed method decreases the number of variables needs to check as well as covering all the variables effect in program.

Fig. 7 portrays a graphical view which shows the number of variables needs to check in proposed method is less by tracing out the critical variables using dependence graph. To show the difference, proposed method is applied to merge sort, quick sort and radix sort C program.

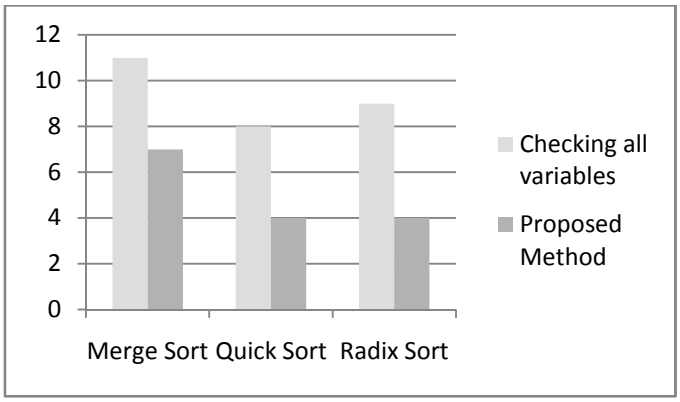



Figure 7. Number of Variables to Check to Detect Soft Errors

In VADA method [13], the idea of precedence of variables was mentioned somewhat but isolated variables inclusion to critical variables was overlooked. The proposed methodology developed in such a way that it can cover the total program effects as well as less number of variable to check by tracing out the critical variables automatically and dynamically from the program. These critical variables in accordance with critical blocks [15] will reduce the time complexity in large scale to detect soft errors in a software program.

Complexity analysis is negligible in the sense that generating the graph can be achieved at compilation time during building a high level computer program. When developing a software program if the variable analysis is done using automated precedence graph in the design phase, the efficiency of the software programs can be increased in large scale. The soft error rate of a software program can be decreased catastrophically too.

## V. CONCLUSIONS

Variable dependence is a major concern of a computer program especially when the code is used to detect soft error. The proposed methodology of this paper automatically detects all the isolated, intermediate and precedence variables using data dependence analysis. The automatically generated dependence graph depicts all the dependences among variables of a computer program. Programmers neither need to keep track of program steps nor need to trace out the dependence variables. Proposed methodology reduces the time to detect soft errors as works with only critical variables. In terms of a software product, proposed method ensures reliability too. Efficiency and accuracy rate increased dramatically as there is less chance of human error as there were in some existing works. Finally, the proposed method indicates a great future in the field of software fault tolerance and soft computing.

## REFERENCES

[1] Mark Weiser. "*Program slicing,"* IEEE Transactions on Software Engineering, 10(4):352–357, 1984.

[2] D. W. Binkley and K. B. Gallagher. *"Program slicing,"*. In M. Zelkowitz, editor, Advances in Computing, Volume 43, pp. 1–50. Academic Press, 1996.

[3] Susan Horwitz, Thomas Reps, and David Wendell Binkley. "*Interprocedural slicing using dependence graphs,"* ACM Transactions on Programming Languages and Systems, 12(1):26–61, 1990.

[4] H. Agrawal and J. R. Horgan "*Dynamic program slicing," In ACMSIGPLAN*" Conference on Programming Language Design and Implementation, pp. 246–256, New York, June 1990

[5] R. Gopal. "*Dynamic program slicing based on dependence graphs,"* In IEEE Conference on Software Maintenance, pp. 191–200, 1991.

[6] B. Korel and J. Laski. "*Dynamic program slicing,"* Information Processing Letters, 29(3):155–163, Oct. 1988.

[7] Dave Binkley,Sebastian Danicic,Tibor Gyim´ othy, Mark Harman, Akos Kiss Lahcen Ouarbya. "*Formalizing Executable Dynamic and Forward Slicing,"* Proceedings of the Fourth IEEE International Workshop on Source Code Analysis and Manipulation (SCAM'04).

[8] S. Horwitz, T. Reps, and D. W. Binkley. Interprocedural *slicing using dependence graphs*. In ACM SIGPLAN Conference on Programming Language Design and Implementation, pages 25–46, Atlanta, Georgia, June 1988. Proceedings in SIGPLAN Notices, 23(7), pp.35–46, 1988.

[9] S. Horwitz, T. Reps, and D. W. Binkley. "Interprocedural *slicing using dependence graphs*," ACM Transactions on Programming Languages and Systems, 12(1):26–61, 1990.

[10] Daniel Jackson and Eugene J. Rollins. "*A new model of program dependences for reverse engineerin,,"* In Proceedings of the ACM SIGSOFT '94 Symposium on the Foundations of Software Engineering, pp. 2–10, December 1994.

[11] K. Muthukumar, M. Hermenegildo, "*Compile-Time Derivation of Variable Dependence Using Abstract Interpretation,*" Journal of Logic Programming, Vol. 13, N. 2-3

[12] Chris Fox, Mark Harman, Youssef Hassoun. *Variable Dependence Analysis Technical Report: TR-10-0,* Elsevier 2010.

[13] M. Harman, C. Fox, R. Hierons, L. Hu, S. Danicic, and J. Wegener. "*VADA: A Transformation-based System for Variable Dependence Analysis,*" In Proceedings of the 2nd IEEE International Workshop on Source Code Analysis and Manipulation (SCAM2002), Montreal, Canada, pp. 55–64, 2002.

[14] Alfred V. Aho, Ravi Sethi, and Jeffrey D. Ullman. "*Compilers: Principles, techniques and tools,*" Addison Wesley,1986.

[15] Muhammad Sheikh Sadi, Md. Mizanur Rahman Khan, Md. Nazim Uddin and Jan Jürjens. "*An Efficient Approach towards Mitigating Soft Errors Risks,"* Signal & Image Processing: An International Journal (SIPIJ) Vol.2, No.3, September 2011.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# A Distributed Neighbor Discovery Based Approach for Cluster Head Selection in Wireless Sensor Networks

Mohammad Mamun Elahi[1,*], Mohammad Mahfuzul Islam[2]
[1]United International University, Bangladesh.
[2]Bangladesh University of Engineering and Technology, Bangladesh.
[*] E-mail: mmelahi@cse.uiu.ac.bd

*Abstract*—**Clustering is an important concept in the area of wireless sensor network research because of the large number of capability-limited sensors. Most of the clustering algorithms proposed so far in the literature needs the complete topology in the memory and faces the problem of exponential increase of execution time with the size of the topology. In this paper, we propose a distributed neighbor discovery based algorithm, which does not need the complete topology for the cluster head selection. The simulation results show that the proposed approach takes much less time and memory when compared to approaches considering the full topology.**

***Keywords—Wireless sensor network, clustering, cluster head, neighbor discovery.***

## I. INTRODUCTION

A wireless sensor network consists of a large number of capability-limited devices, which lack processing power, memory, and power consumption. There are lots of trade-offs to deal with since these sensors are deployed in remote places that are not easy to reach. The devices have a finite lifetime and must be recharged or replaced by new sensors. These limitations have shown an increasing interest from the scientific community to research in such area that would enhance the longevity and coverage of the resulting network. The main emphasis is on maximizing the life time of sensors and to use the limited resources efficiently by adopting mechanisms, algorithms and protocols that consider these limited resources as main priorities and challenges to produce efficient and reliable networks.

Clustering is one of the methods used to employ limited available capacity of these tiny sensors for extended lifetime. Lots of clustering algorithms have already been proposed in the literature, most of them need the complete topology of the sensor network. LEACH is one of the most well known energy efficient clustering algorithms for WSNs that forms node clusters based on the received signal strength and uses these local cluster heads as routers to the BS[1] – [3]. A number of variants to enhance LEACH are proposed to make clustering more efficient [4] – [7]. Other well-known hierarchical routing protocols are HEED, TEEN, PEGASIS [13]-[14] etc.

Most of the clustering algorithms proposed in the literature need to have the complete topology mapped in the memory to form the clusters. The main problem with those approaches is that the execution time exponentially increases with the increase in the graph size, i.e. number of nodes. We can reduce the execution time because of less number of comparisons if we can distribute the clustering process.

Neighbor discovery is an important part of the wireless sensor network operation as it is required for the construction of the topology as well as for the routing of the data correctly to the base station. A number of neighbor discovery algorithms have been proposed considering different parameters [10] – [12]. It can be initial phase of each round of clustering or can be continuous throughout the operation of the network, particularly when nodes are mobile. Maximum degree analysis is a method to select nodes as cluster heads in some applications [7] – [8]. In this paper, we have proposed a neighbor discovery based clustering algorithm, which is a distributed process and does not need complete topology for the cluster head selection. Each node will exchange the node degrees with other nodes and the node with the maximum degree will be selected as cluster head. The simulation results show that the proposed approach takes much less time and memory when compared to approaches considering the full topology.

The paper is organized as follows. Section II describes the proposed algorithm in details. Section III discusses the simulation environment and the simulation results. Finally, section IV provides the concluding remarks.

## II. PROPOSED ALGORITHM

The algorithm is divided into two phases. In the neighbor discovery phase, all nodes will discover its immediate neighbors, calculates the node degree, and exchange that information with neighbors. In the clustering phase, node with the maximum degree will be selected as the cluster head. The iterations will continue until no more nodes are left.

### A. *Neighbor Discovery phase*

First, each node will discover the immediate neighbors by exchanging HELLO messages among themselves. Then, nodes will calculate their own node degree and exchange with its neighbors.

### B. *Maximum Degree Clustering*

After receiving the node degrees from neighbors, each node will compare its own degree with each of its neighbor's node degree. Only when its own degree is greater than all its neighbors, it will select itself as the cluster head. Otherwise, it will add itself to the cluster of the node with the maximum degree.

Let us illustrate the procedure as an example. Consider the graph shown in the Figure 1. *n(i)* denotes number of neighbors of node i. At step 1, each node will exchange the *NoOfNeighbors()* to each other. Initially, node 1 calculates n(1) = 3, sends to node 2, 3, 4 (neighbors). Similarly, node 2 calculates n(2) = 2, sends to node 1, 3, 4 (neighbors), and so on. Nodes 1, 2, 3 will update node 4 as the cluster head since their n values are lower than node 4. So, node 4 is a prospective cluster head.

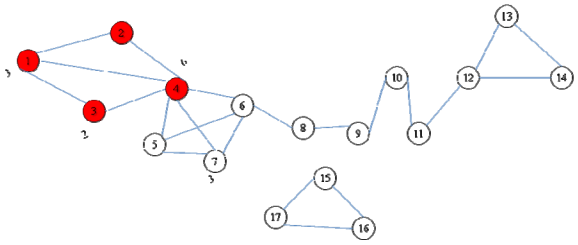


Figure 1. Nodes 1, 2, 3, 4 exchanging node degrees

Similarly, node 4, 5, 6, and 7 will exchange the *NoOfNeighbors()* to each other as shown in the Figure 2. Now, node 5, 6, and 7 will also update node 4 as the monitoring node since their n values are lower than node 4. Now, node 4 gets the n(i) from all of its neighbors i and it is confirmed that node 4 has the highest n value. So, node 4 will be selected as a cluster head as shown in the Figure 3.

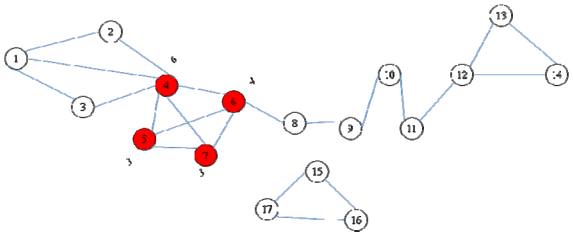


Figure 2. Nodes 4, 5, 6, 7 exchanging node degrees

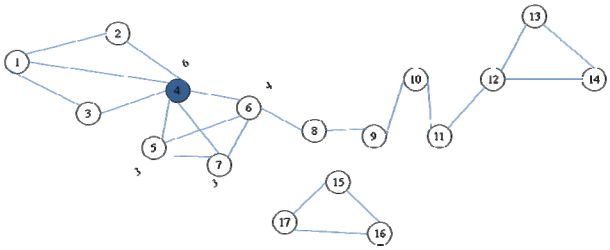


Figure 3. Nodes 4 selected as the cluster head

Similarly the algorithm will continue until all the nodes are processed and clustered. Final clustering scenario is shown in the Figure 4.

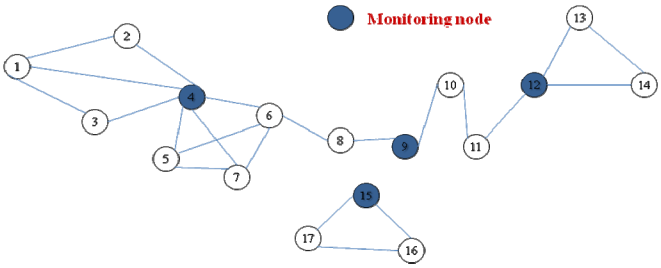



Figure 4. Final clustering scenario with cluster heads selected

The advantages of the proposed approach can be summarized as follows:

1. No need to have the complete topology of the network.
2. If a new node is added, only the neighbors will re-calculate and exchange, no global re-calculation is needed.

### C. *Mathematical Model*

We assume all the nodes have same initial energy. The amount of energy consumption greatly varies for different number of clusters. In our case, since we use the neighbor discovery mechanism for clustering, we cannot pre-define the expected number of clusters. It will be decided by the topology of the sensor network.

Suppose, we have N number of nodes. During the neighbor discovery phase, neighbor information will be exchanged among the nodes. A node i will be cluster head, if:

$$n(i) > \max\{n(j) \mid j \in neighbors(i)\} \quad (1)$$

Where, n(i) is the number of neighbors of node i. If $T_n$ is the time required for neighbor discovery phase, $T_s$ denotes the service time in the communication phase, then the required time in every round is $T_r$, which can be defined as:

$$T_r = T_n + T_s \quad (2)$$

The Cluster head of a cluster looses energy in the communication phase by receiving packets from non-cluster head nodes, aggregating them, and transmitting the aggregated information to the base station. Let, $C_H$ be the average number of packets transmitted by a cluster head, and $C_{NH}$ is the average number of packets of *l* bits transmitted by a non-cluster head node at time $T_r$. The estimated time for the communication round $T_s$ is:

$$T_s = \frac{l}{R_b} C_H \quad (3)$$

Where, $R_b$ is the available bit rate. For each subsequent round the node with second largest number of neighbors and so on until first node in the sensor network is dead.

## III. EXPERIMENTAL SETUP AND RESULTS

We have simulated the algorithm in MATLAB and to illustrate the efficacy of the proposed approach a maximum degree analysis algorithm, which needs to have complete topology in the memory and was used in [8] – [9]. The parameters used in the experiment are listed in Table 1.

TABLE I. SIMULATION PARAMETERS

| **Parameters** | ***Value*** |
|---|---|
| Number of Nodes | 50 nodes ~ 200 nodes |
| Range | 500 ~ 2000 units |
| Area | 1000 x 1000 ~ 4000 x 4000 units |

For the experiment, we have considered homogeneous wireless sensor network where each node has same sensing range and capabilities. Nodes have been placed randomly in the given area. For the performance parameter, we have used number of comparisons and the execution time of the algorithms.

### *A. Number of Comparisons vs. Number of Nodes*

Figure 5 shows the graph, which compares the number of comparisons needed for the centralized approach with the proposed approach. As shown in the figure, total number of comparisons exponentially increases with the number of nodes increased in the maximum degree analysis with complete topology in the memory, whereas number of comparisons increases only linearly with the increase of number of nodes in the neighbor discovery based approach. This is because in the neighbor discovery each node only considers the degrees of the neighboring nodes and does not need to consider other nodes.

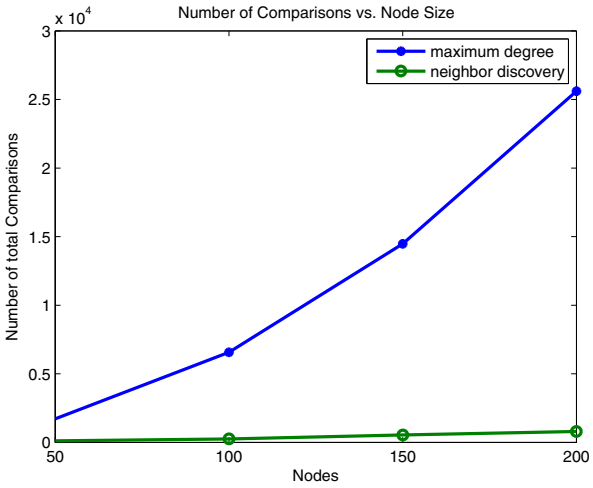



Figure 5. Graph showing total comparisons as a function of number of nodes

### *B. Time vs. Number of Nodes*

The graph in the Figure 6 compares the execution of the two different algorithms. This graph shows the same trend as figure 5, where increase in number of nodes exponentially increases the execution time in the centralized maximum degree analysis, whereas execution time only increases linearly with the increase in number of nodes in the neighbor discovery based approach.

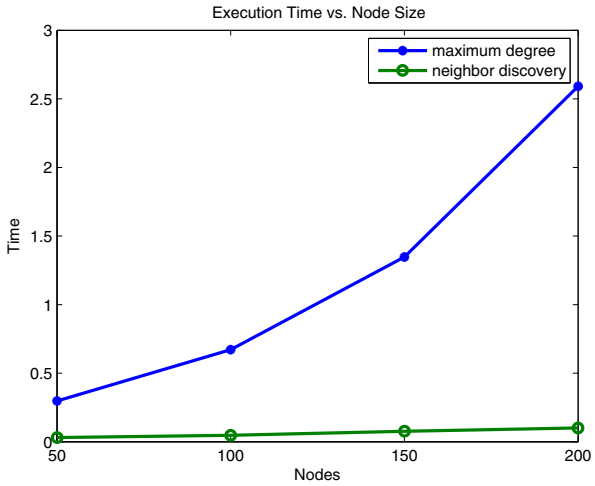



Figure 6. Graph showing total execution time as a function of number of nodes

## IV. CONCLUSIONS AND FUTURE WORKS

Hierarchical clustering approaches are used in wireless sensor networks to make the routing process more efficient. But most of the clustering protocols proposed in the literature require the knowledge of complete topology of the network and cost of processing increases exponentially with the increase in the graph size. On the other hand, neighbor discovery is an important part of the topology construction process. In this paper, we have proposed a distributed clustering method, which takes the advantage of neighbor discovery process and does not need complete topology of the network to be in memory. Experimental results show that the proposed algorithm performs better with the increase in the network size (number of nodes).

In our experiment we have assumed homogeneous sensors with same capacity and power. One limitation in our approach is cluster heads will lose more power compared to other nodes. So, after certain amount of time, cluster heads should be re-calculated. We will consider those issues in the future work.


## ACKNOWLEDGEMENT

The contribution of this paper is the part of the research work carried out for my PhD thesis in the Department of Computer Science and Engineering of Bangladesh University of Engineering and Technology (BUET).


## REFERENCES


[1] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks", *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, 2002, pp. 660-670.

[2] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-efficient routing protocols for wireless microsensor networks", in Proc. *33rd Hawaii Int. Conf. System Sciences (HICSS)*, 2000.

[3] J. Chen, J. Lu and Q. Wang, "Research and Improvement of Adaptive Topology Algorithm LEACH for Wireless Sensor Network", *IEEE Wireless Communication, Networking & Mobile Computing*, WiCOM, 2008, pp. 1-4.

[4] A. Azim, and M. M. Islam, "A Dynamic Round-Time Based Fixed Low Energy Adaptive Clustering Hierarchy for Wireless Sensor Networks", *IEEE 9th Malaysia International Conference on Communications (MICC 2009)*, December 14-17, 2009, Malaysia.

[5] R. Munjal and B. Malik, "Approach for Improvement in LEACH Protocol for Wireless Sensor Network", *International Conference on Advanced Computing & Communication Technologies (ACCT)*, 2012.

[6] A. Azim, and M. M. Islam, "Dynamic Service Policybased Clustered Wireless Sensor Networks", *The 6th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2010)*, October 11-13, 2010, Canada.

[7] A. Azim, M. M. Islam, and G. Sorwar, "Estimate the Initial Battery Power of Nodes and Optimal Number of Clusters To Achieve the Expected Life Span of Wireless Sensor Networks", *IEEE TENCON* 2010, Japan.

[8] F. T. Jaigirdar, M. M. Islam and S. R. Huq, "Grid approximation based inductive charger deployment technique in wireless sensor networks," *International Journal of Advanced Computer Science and applications (IJACSA),* vol. 2, pp.30-37, 2011.

[9] F. T. Jaigirdar, and M. M. Islam, "Assurance of the Maximum Destruction in Battlefield using Cost-Effective Approximation Techniques", *Journal of Networks*, Academy Publisher, 2012.

[10] S. Fang, S. M. Berber, and A. K. Swain, "Analysis of neighbor discovery protocols for energy distribution estimations in wireless sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 4386–4390, 2008.

[11] S. Vasudevan, D. Towsley, D. Goeckel, and R. Khalili, "Neighbor discovery in wireless networks and the coupon collector's problem," in *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking (MOBICOM '09)*, pp. 181–192, September 2009.

[12] S. Vasudevan, M. Adler, D. Goeckel, and D. Towsley, "Efficient algorithms for neighbor discovery in wireless networks", *IEEE/ACM Trans. Netw.* 21, 1 (February 2013), 69-83.

[13] W. R. Heinzelman, A. Chandrakasan, and H.Balakrishnan, "Energy-efficient communication Protocols for Wireless Microsensor Networks," *Proc of the Hawaii International Conference on Systems Sciences*, vol. 2,pp.10-19, Jan. 2000.

[14] O. Younis and S. Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed clustering approach for Ad Hoc sensor networks," *IEEE Transactions on Mobile Computing*, Vol. 3, No. 4, 2004, pp. 366–379.

# A Distributed Load Balancing Algorithm for Adaptive Cognitive Radio Network

Mohammad Mamun Elahi*, Shahrier Siddique
United International University, Bangladesh.
* E-mail: mmelahi@cse.uiu.ac.bd

*Abstract*—**Cognitive radio network, which deals with dynamic frequency assignment, is one of the most popular research topic nowadays in the context of cellular networks, WLANs and mesh networks as well. The growing interest in the cognitive radio technology and its capability to offer more efficient spectrum usage attracts researchers to find efficient algorithms to accommodate more users and to provide better service in wireless communication. In this paper, we present a distributed algorithm for selecting channels, in cognitive radio environment, so that the load is distributed over them for more efficient service. Simulations and experimental results show the efficacy of proposed approach.**

***Keywords—Cognitive radio, nash equilibrium, channel allocation, load balancing.***

## I. INTRODUCTION

Cognitive radio (CR) is one of the latest technologies in mobile networks. It offers dynamic and real-time network management. The cognitive radio was first proposed by Mitola [1]. Cognitive radio uses learning process and intelligence to provide the best-possible radio settings and user experience. Channel allocation in CR wireless networks has been an interesting research topic for long time. Many solutions have been proposed. There are solutions for centralized and distributed systems, addressing the frequency allocation problem in WLANs [2-4][15-16]. There are examples of using game theoretical approach as solutions [5-7]. Even there are several proposed frameworks for load balancing [15-17].

A distributed load balancing algorithm for adaptive channel allocation for cognitive radios was proposed in [13] by Fischer et. al., which attracted our attention because it can co-exist with any centralized and access point (AP) based load balancing system. And it can be embedded in end user devices or used as an add-on app for end devices to enrich user experience. But considering some of the drawbacks in their proposed algorithm, we intend to modify it and enhance it in this paper. Mentioned in the original paper, there are, two separate channel allocation problems related to cognitive radio networks:

- First and the obvious one is the much studied problem for the secondary users to detect spectrum opportunities that are opened by the non-transmitting primary users. In this case, the secondary users need to find out the spectrum opportunities and decide how to use them.
- Second, there is a problem how secondary users should choose between the available channels. This is, of course, a sort of load balancing or resource allocation problem, which can be quite difficult to solve if there is no central authority and the environment becomes very dynamic. In fact, Mahonen and Petrova have noted that both the first and second problem may be open to flash crowd effects in certain circumstances [8].

The authors in [13] have addressed the problem of frequency selection by secondary users in a CR environment, and provided a solution to minimize the interference level and maximize the throughput for each individual user without damaging the performance of the other users. But their solution adds some drawbacks like:

- Uncertainty of achieving optimum load distribution for choosing random sample agent.
- Might never take best decision due to choosing random sample agent.
- It also introduces more resource consumption from end devices.

In [14], the authors use a similar algorithm to compare and balance algorithm from [13]. It introduces a threshold value. But it also contains the same drawbacks. The goal of this paper is to ensure near optimum distribution of channel using minimum resources from end devices and achieve near optimal decision within minimum possible time. We present a simple algorithm to achieve load balancing among the channels for the secondary users in quite general framework. The algorithm is sequential sample based and reaches equilibrium at optimum speed. And this speed determines the adaptability in very fast dynamic environment.

The paper is organized as follows. Section II describes system and computational model. Section III illustrates the proposed load-balancing algorithm. Section IV discusses the experimental results. Finally, section V concludes the paper with findings.

## II. SYSTEM AND COMPUTATIONAL MODEL

Here we discuss the system model we have used in our work. We consider n balls are given (*CR enabled agents*) and m bins (usable channels or frequencies). To simplify the model, we assume all balls are of same size means all users introduce equal traffic load to the system and uses same communication technology, e.g., WLAN (802.11b/g). Each radio is assigned with a frequency, and $n_i$ denotes the *load* of channel i, i.e., the number of balls that select bin i ∈ [m]. Because our agents are indistinguishable, the system is described by the state vector $n = (n_i)_{i\in[m]}$. This solution will also work for uneven loads. Each bin *i* have an individual capacity bound δi which may or may not be elastic. But we are considering only systems like CDMA with elastic capacity. Fixed systems like TDMA are not considered here. Every ball should decide should it stay in same bin or change the bin to achieve the best possible performance. It should be considered that new balls (users) come into the system and try to use a channel. This means every time there will be fluctuations of the number of 'n' balls. Each bin i ∈ [m] is associated with value of $c_i(n_i)$ specifies the cost incurred to all balls choosing bin i. The value of $c_i(n_i)$ is the cost incurred to all balls for choosing bin i. As the cost function we can choose any of the utilities described below.

Considering that the users aim at minimizing these values, we're setting the scale such that the maximum value of any $c_i$ is 1 and the minimum value is 0.

$$C(n) = \sum_{i\in[m]} \frac{n_i}{n} c_i(n_i)$$

This denotes the average cost sustained by the agents. A ball has an incentive to migrate from its current bin to another bin if, it gains more utility by doing so. We can consider a state n to be stable, if no agent migrating. This can be defined by the concept of Nash equilibrium.

*Definition (Nash equilibrium):* A state n is at a *Nash equilibrium* if for all machines i and j with $n_i > 0$ it holds that $c_i(n_i) \leq c_j(n_j + 1)$.

Due to a potential function argument, pure Nash equilibrium will always exist in this model, even with different weights [9][10]. Though Nash equilibrium does not necessarily optimize the overall performance of the system, but it will utilize every channel load to ensure stability. We are assuming that there is no inter-channel interference in the system. This means that the effect of overlapping channels will not be considered. And the balls can sample the number of balls '$n_i$' from a certain bin. Utilities are defined bellow:

1) *Utility 1: Minimizing interference* - The goal is to minimize the number of balls per bin, means less load per channel. If inter-channel interference is considered we should minimize the number of ni in a single bin but also in the neigbouring (i − 1) and (i + 1) bins.
2) *Utility 2*: *Maximizing throughput* - Maximizing the throughput is very much related to minimizing interference. In case of a stochastic MAC less balls will lead both to a minimized interference and maximized throughput.
3) *Utility 3: Minimizing Latency* - This is valid for time-sensitive applications like streaming or voice calls. In this paper we are considering as the original paper [13], this utility function uses two different MAC protocols, slotted-ALOHA and CSMA/CA. The slotted ALOHA [11] where the users are allowed to transmit only at the beginning of the slot. The maximum throughput that can be achieved is $S = Ge^{-G}$, where G is an offered load. In order to consider CSMA/CA and to have a functional form for the utility, Bianchi's analytical model has been used for the saturation throughput of IEEE 802.11 DCF systems [12].

To handle real-world load in a decentralized manner, some properties need to be maintained:

- Algorithm is executed locally or in end devices. They do not need to rely on base station or any node which acts as a centralized manager.
- Algorithm may rely only on the information which they are able to gather by themselves.
- Algorithm should be simple that any node can join and leave the network dynamically.
- Algorithm should be most energy efficient.
- Each device should strive to maximize its own utility rather than the overall system utility. We are assuming that an agent using channel i can easily determine its own cost $c_i(n_i)$ (e. g., by measuring its throughput). CRs can estimate, the number of agents using each channel, mentioned earlier. To improvise it our algorithm is includes a function MEASURE LOAD(i) which returns the value of $n_i$. This value may be different under the influence of uncertainty.

To decide the migration of agents, we also need to know the value of utility, it will achieve by migrating. This can be achieved by decoding the headers of target channel. In our algorithm, this function is named MEASURE COST(i), which returns $c_i(n_i)$.

## III. PROPOSED LOAD BALANCING ALGORITHM

We assume that the MEASURE COST() can be applied to any channel. Consider an agent currently using channel i, at first this agent determines C by measuring its own Cost(), then the $n_{channel}$ by measuring its own channel load(). At intervals, $n_i$ is determined by measuring load of the next channel in sequence, where n is a real-time dynamic vector. The value of n is calculated only for that much range, where at least one suitable channel is available or there is no more available channel. If a suitable channel is found by comparing current through output and probability (ℙ) of switching, then agent i

migrates to the suitable channel and immediately the algorithm resets itself. Sequential sample selection confirms migration if any better channel is available, it also confirms the best possible load distribution and responses very fast to select available channel. It only uses that much resources required in finding a better channel ["break" in the if block ensures optimum resource is used].

Pseudo code:

```
for all balls in parallel do
    c ← Measure_Cost(channel)
    n_channel ← Measure_Load(channel)
    for all channels i ∈ [m] do
        n_i ← Measure_Load(i)
        n ← ∑n_i
    ℙ[channel] = n_channel /n
    ℙ[i] = n_i / n
    c' ← Measure_Cost(i)
    if c' < c then
        if ℙ[channel]> ℙ[i] then
            channel ← i
                break
        end if
    end if
end for
end for
```

Figure 1. Proposed algorithm pseudo code.

The expected load vector that results from one round starting at a Nash equilibrium should be a Nash equilibrium again in end. To see this, we consider a load vector n. Since any ball in bin j migrates to bin i with probability $c_j(n_j) \cdot n_i/n$, the expected load $n'_i$ of any channel i after one step is (q=1):

$$E_2[n_i'] = n_i - \sum_{s\in q} n_s.c_s(n_s) + \sum_{s\in q} n_s.c_s(n_s).\frac{n_i}{n}$$

$$= n_i - \sum_{s\in q} n_s.c_s(n_s) + n_i.C(n)$$

$$= n_i - \sum_{s\in q} n_s.c_s(n_s)$$ [The latter equality holds since at a Nash Equilibrium, cs(ns)=C(n) (or ns = 0)]

$$= n_i - n_s.c_s(n_s)$$

$$= n_i$$

## IV. Experimental Results

We have performed simulations over the classic compare and balance algorithm and the new enhanced compare and balance algorithm. Simulations are organized in 3 sections: construction, goal and result.

### *A. Graph-1: user migration graph.*

Figure 1 shows the user migration graph.

Construction-

- Data used 20 users with 20 available channels.
- Exhibits how users were distributed before and after the algorithms are performed.

Goal-

- Finding comparative quality of distribution.
- Analyze load and cost effect on user and channel.

Result-

- Chart-1(classic compare and balance) reaches Nash Equilibrium after 30 iterations where Chart-2(Enhanced compare and balance) reaches Nash Equilibrium after 4 iterations.
- Chart 2 shows better distribution than chart 1. The red line shows load is better distributed among channels.

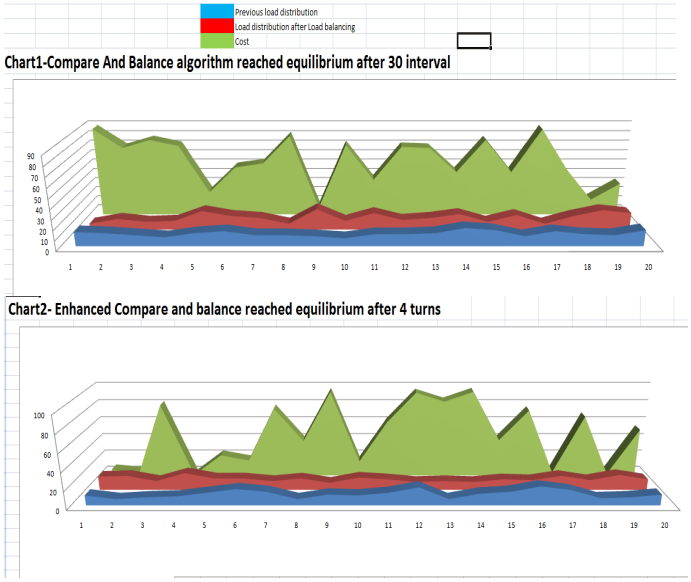



Figure 2. User Migration Graph.

### *B. Graph 2 - Chart 3.1: Average cost graph.*

Chart 1 in Figure 2 shows the average cost graph.

Construction-

- Data used: load= 1000 users on each channel, 1000 channel available to each user.
- Exhibits how time and resources are consumed to reach Nash equilibrium.

Goal-

- Comparing resource consumption.
- Estimation of channel switching (migration) occurred.

Result-

- Blue (classic compare and balance) reaches Nash Equilibrium after 1589 migrations or .15% of the total network load. Where red(Enhanced compare and balance) reaches Nash Equilibrium after 822 migrations or 0.0822% of the total network load and which requires approx. 49% less migrations than classic compare and balance.
- Blue (classic compare and balance) reaches Nash Equilibrium after 30.32 milliseconds, where red (Enhanced compare and balance) reaches Nash Equilibrium after 18.55 milliseconds. This is approx.

more 63% faster than the classic compare and balance algorithm.

- In general more migration requires more time to achieve network stability. Less migration provides better user experience and fewer over heads on network.

Chart 3.1

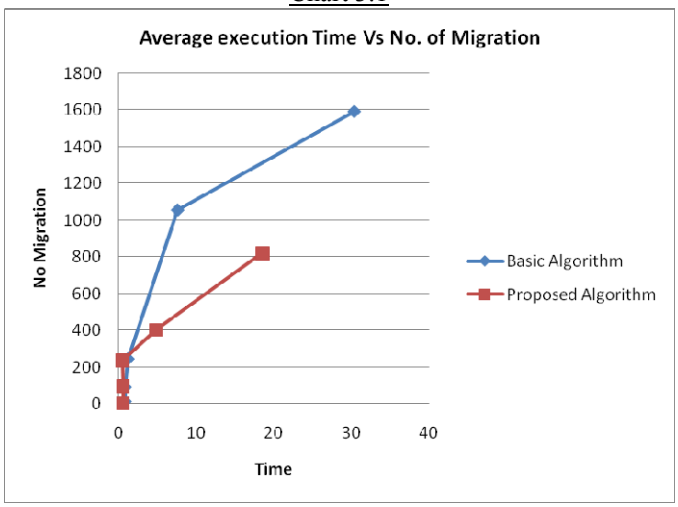



Chart 3.2

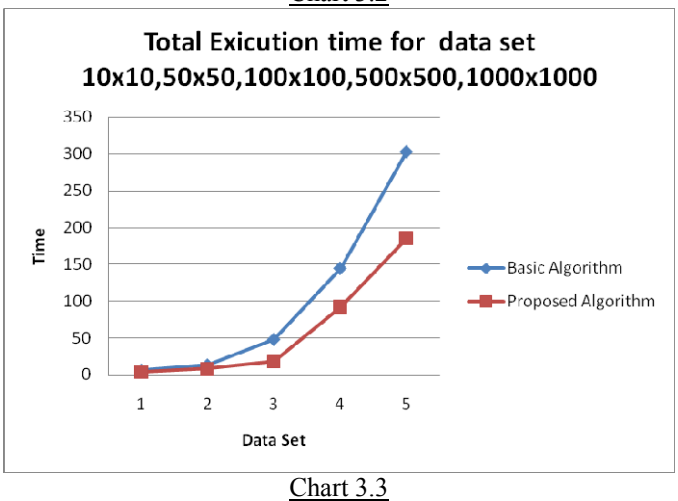



Chart 3.3

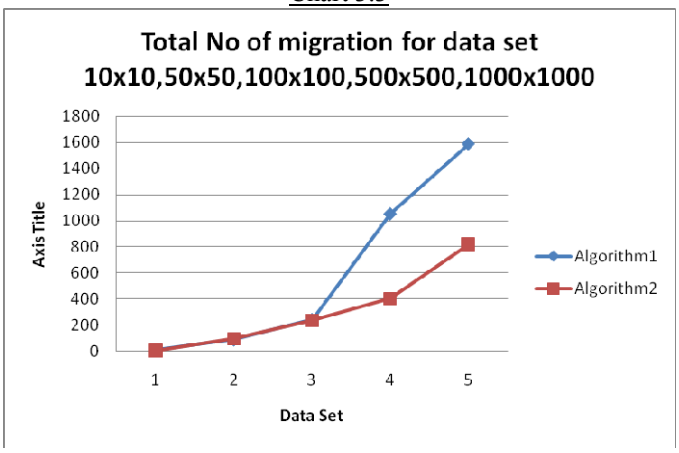



Figure 3. Resource Consumption Graph.

*C. Graph 2 - Chart 3.2, 3.3: Total time consumption graph.*

Chart 3.2 and 3.3 in Figure 2 shows resource consumption graphs.

Construction-

- Data : Set of (load= 10 users on each channel, 10 channel available to each user)
  Set of (load= 50 users on each channel, 50 channel available to each user)
  Set of (load= 100 users on each channel, 100 channel available to each user)
  Set of (load= 500 users on each channel, 500 channel available to each user)
  Set of (load= 1000 users on each channel, 1000 channel available to each user)
- Exhibits detail view of chart 3.1, in time dimension in respect to different data sets(chart 3.2) and migrations per data sets(chart 3.3)
- Chart 3.1, 3.2 and 3.3 uses same input data set and output data set with different dimensions.

Goal-

- Analyzing resource consumption behavior for different data sets.
- Analyze estimated migration required to reach stability in different network environment.

Result-

- Stability is inversely proportional to load of the channels in a network.
- Number of migration is inversely proportional to quality of distribution.
- Resources consumed in end device is proportional to available channels to that specific device for classic compare and balance, but for enhanced compare and balance Resources are consume in end device has no proportionality to available channels or any other metrics. It uses only that much resource required in finding a better channel.

The Compare and Balance algorithm mentioned in [13] and [14] has two key drawbacks, which we have improved in our system. First one is the random selection of channels for comparison, which introduce some uncertainty for optimization of load balancing. We removed that with sequential channel selection for comparison, which reflects its result on graph-1, chart-2. A better distribution of channels due to sequential selection rather than random selection, compared in graph-1, chart-1. And second one is the comparing the load and cost for every available channel. We have changed this to find a channel of lower load and if the cost is also lower then migrate to that channel. This limits the use of resources to find a channel with lower load instead of computing load and cost for the whole available channel. And the result is in Graph-1, Lower numbers of iterations are required to reach Nash equilibrium and it is about 86% less than the original algorithm. This means 86% reduction in use of resources. This may change over the dynamic situation of the network. But even in worst case enhanced compare and balance will use fewer resources than the original compare and balance, because of comparing only the load first instead of load and cost.

## V. CONCLUSIONS

We have studied in this paper the secondary user channel allocation subject to selfish load balancing, game theoretical approach, and have suggested a modified adaptive distributed algorithm it to reach an equilibrium solution faster than ever. We have shown that our proposed algorithm converges very fast and completely removes the uncertainty of the channel load. Furthermore, the results show that the convergence behavior of the algorithm is independent of tested utility function type. We consider the simulation results very encouraging for the implementation of a load balancing adaptive channel assignment solution and testing it to a more complex system model.

## REFERENCES


[1] J. Mitola, *Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio.* Ph.D. Thesis, KTH, 2000.

[2] J. Riihijarvi et al., "Performance evaluation of automatic channel assignment mechanism for IEEE 802.11 based on graph colouring," in *Proceedings of IEEE PIMRC*, Helsinki, September, 2006.

[3] K. Leung and B. Kim, "Frequency assignment for IEEE 802.11 wireless networks," *Proc. of IEEE 58th Vech. Tech. Conference Vehicular Technology Conference*, vol. 3, pp. 1422–1426.

[4] F. Gamba, J. F. Wagen, and D. Rossier, "A simple agent-based framework for adaptive wlan frequency management." *Proc. of MOBICOM*, September 2003.

[5] L. Berlemann et al., "Spectrum load smoothing for cognitive medium access in open spectrum," in *Proc. of IEEE PIMRC*, vol. 3, 2005, pp. 1951–1956.

[6] A. Laufner and A. Leshem, "Distributed coordination of spectrum and the prisoner's dilemma," in *Proc. IEEE DySPAN*, vol. 1, 2005, pp. 94–100.

[7] S. H. Wong and I. J. Wassell, "Application of game theory for distributed dynamic channel allocation," *Proc. of IEEE Trans. on Vehicular Technology Conference*, Birmingham, AL, USA, vol. 1, pp. 404–408, 2002.

[8] P. Mahonen and M. Petrova, "Flash crowds in cognitive radio environment: Beware an unpredictable mob," *PIMRC 2007* (available as a technical report).

[9] R. W. Rosenthal, "A class of games possessing pure-strategy Nash equilibria," *Int. Journal of Game Theory*, vol. 2, pp. 65–67, 1973.

[10] D. Fotakis et al., "The structure and complexity of Nash equilibria for a selfish routing game," in *Proc. 29th Int. EATCS Coll. on Automata, Languages and Programming (ICALP)*, Malaga, Spain, 2002, pp. 123–134.

[11] A. Tanenbaum, *Computer Networks*, 4th ed. Pearson, 2003.

[12] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, March 2000.

[13] S. Fischer, M. Petrova, P. Mahonen, and B. Vocking, "Distributed Load Balancing Algorithm for Adaptive Channel Allocation for Cognitive Radios", *RWTH Aachen University*, Computer Science 1, 52056 Aachen, Germany, 2007.

[14] S. Fischer, P. Mahonen, M. Schongens, and B. Vocking, "Load Balancing for Dynamic Spectrum Assignment with Local Information for Secondary Users", *RWTH Aachen University, Department of Wireless Networks*, Kackertstrasse 9, 52072 Aachen, Germany, 2008.

[15] Q. Chen, J. Chen and L. Tang, "Dynamic Spectrum Load Balancing for Cognitive Radio", *Key Lab of Mobile Communication Technology* Chongqing University of Posts and Telecommunications Chongqing, China, 2009.

[16] M. Bennai, J. Sydor, and M. Rahman, "Automatic Channel Selection for Cognitive Radio Systems", *Communications Research Centre*, 3701 Carling Avenue Ottawa, Ontario K2H 8S2, Canada, 2010.

[17] L. Wang, C. Wang, and F. Adachi, "Load-Balancing Spectrum Decision for Cognitive Radio Networks", *Selected Areas in Communications, IEEE Journal on* , vol.29, no.4, pp.757,769, April 2011.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Integration of Solar Heater and Solar Pond Technology for Enhancing the Thermal Efficiency of Water Heating System

M. Sabbir Rahman, Ismat Zerin, and A Z M Shahriar Muttalib
Electrical and Electronic Engineering Department, American International University-Bangladesh
Bonani, Dhaka, Bangladesh
E-mail: engsabbir.007@gmail.com, zerin.eee@gmail.com, A.Muttalib@brunel.ac.uk

*Abstract*— **Nowadays solar energy is used in different purposes, which is beneficial for our daily life. Solar heater and solar pond both are used identically for the collection and storage of solar energy at present. In this paper a new concept (integration of solar heater and solar pond technology) of increasing the efficiency of solar thermal energy is proposed. Solar heater technique is used generally to collect solar energy. And solar pond is used to collect and store solar energy simultaneously. But during night and cloudy weather it is unable to collect heat as energy from solar resources. In order to continuous and long period heat supply efficiently integration of solar pond and solar heater concept is proposed. This new concept will help to supply continuous heat at a time for a long period because of the storage thermal energy. The combination of solar pond and solar thermal energy will also decrease the heating time and increase the heat stored capacity of water as well.**

Keywords— ***Renewable energy, solar energy, solar pond, thermal efficiency***

## I. INTRODUCTION

The energy which comes from natural resources such as sunlight, wind, rain, tides, waves and geothermal heat can be denoted as renewable energy. Basically, all forms of energy in the world as we know it are solar energy. Because, oil, coal, natural gas and wood are originally produced by photosynthetic process, followed by complex chemical reactions in which decaying vegetation was subjected to very high temperature and pressure over a long period of time. The wind and tide energy have a solar energy since they are caused by difference in temperature in various region of the earth. The most important benefit of renewable energy system is the decrease of environmental pollution. Over the past century fossil fuels have provided most of our energy because these are much cheaper and more convenient than energy from alternative energy sources and until recently environmental pollution has been of little concern. There are many alternative energy sources which can be used instead of using fossil fuels. As compared to other forms of renewable energy solar energy [1, 2, 3] is the most preferable because it is clean and can be supplied without any environmental pollution [9].

Many technologies have been created to collect solar energy as heat from sun radiation. Solar heater technology is one of them. But it is unable to store heat for a long period of time. But for the future utilization solar energy storage is an important objective, because during night and cloudy weather solar energy is not available. To overcome this problem solar pond (salinity gradient) has been established. Solar pond is another technology which has an attractive cheap means of collecting and storing solar energy in the form of hot high density salt water. Through the use of solar pond solar energy can be trapped. This paper mainly deals with the concept of solar pond and solar heater technology [1, 4].

## II. AIMS AND OBJECTIVES

The aim of this research is to increase the storage of thermal efficiency of solar energy. The research will allow us

- The same volume of water compare to conventional solar heater with a higher temperature.
- The higher volume of water can be heated within short period.
- Cost effective way to heat up the water.
- To supply water at night by using efficient water storage.
- Enhancing the thermal efficiency of water heating system.

The main objective of this research is to take some necessary steps to increase the solar thermal efficiency which will be more efficient than using solar pond and solar water heater identically. The objectives to achieve the goal of the projects are following:

- To increase the overall storage temperature solar pond and solar heater technology is used together.
- As an insulator material wooden wall can be used in solar pond to protect heat transfer to the environment
- To increase the storage heating capacity black colour coating can be used in inner side of the solar pond for better absorption of solar heat.

- Greenhouse effect technology can also be implemented by using glass plate at the top of the solar pond.

Metal thin sheet can be preferable at the bottom of the solar pond for good heat conductor in solar heater purposes.

## III. BACKGORUND CONCEPT REVIEW

### A. *Solar Pond Technology*

Solar pond is mainly a body of processing water to trap solar energy. Because of normal water is not capable to storing the solar heat energy. If sun source is used to heat the water then the warm water will rise and expand because of less dense. It will lose its heat to the air through convection or evaporation and by the replacement of cold water through natural convective circulation that mixes the water and dissipates heat. The design of solar ponds reduces either convection or evaporation in order to store the heat collected by the pond [6].

### B. *Brief explanation of water layers* [7, 8]

A typical salinity-gradient solar pond consists of three main zones which are briefly described below:

- *Upper Convective Zone (UCZ)*

This is the first water layer which has a least cost, salinity (low salt concentration), temperature and is close to the ambient temperature. The thickness of this zone is typically 0 .3m or .5m. And it should be kept as thin as possible.

- *Non-Convective Zone (NCZ)*

This layer is also called gradient zone. It is located in the middle of the two layers, UCZ and LCZ. In this layer salt is not homogeneously distributed. Because In this zone salinity increases with the depth. Because of less salt content water or lighter water above the UCZ layer water in this gradient zone cannot rise. Similarly water in this layer cannot fall because higher salt content means heavier water placed below the LCZ layer.

- *Lower Convective Zone (LCZ)*

This layer has a high salinity at the bottom of the pond. It is also called a storage zone which acts as a transparent insulator. Because the permitting sunlight to be trapped in the bottom layer and become hot. As the LCZ's depth increases, the heat capacity increases, the capacity increases and temperature variation decreases. By this process heat will be stored. From which useful heat can be withdrawn [6, 4].

Figure 1 indicates three layers, position of layers and the depth ratio between the three water layers of solar pond also.

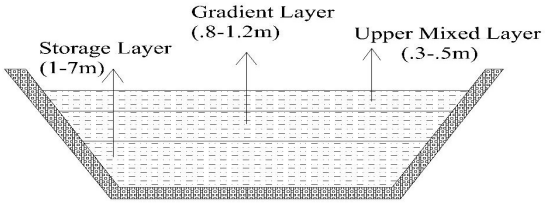



Figure 1. Water layer concept.

### C. *Solar Heater*

Flat plate solar-heat collector concept is used here. The working principal of solar thermal collectors is to capture and retain heat from the sun and use it to heat a liquid. In order to heat water using solar energy, a collector, often fastened to a roof or a wall facing the sun, heats working fluid that is either pumped or driven by natural convection through it.

Figure 2 shows the basic arrangement and the heat flow system of solar heater.

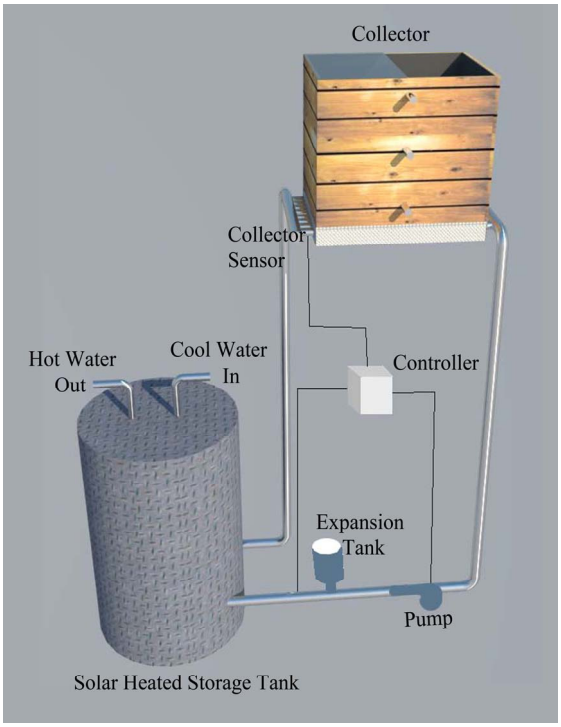



Figure 2. Heat flow arrangement of solar water heater.

### D. *Integration of solar pond and solar heater technology*

The whole system is based on two unique concepts one is solar heater and another is solar pond with gradient salinity water.

E. *Advantage of the proposed integration of solar pond and solar heater*

It has seen from the above description is that in spite of having many advantages of solar pond and solar heater identically there have some limitations. The proposed model will help to reduce those limitations properly. It will able to store the solar energy during night and cloudy weather. The large area will not require. Additionally, maintenance for these systems is generally simple and low-cost. This technology can be used in residential buildings where the demand for hot water has a large impact on energy bills. This generally means a situation with a large family, or a situation in which the hot water demand is excessive due to frequent laundry washing. Commercial applications include Laundromats, car washes, military laundry facilities and eating establishments. The technology can also be used for space heating if the building is located off-grid or if utility power is subject to frequent outages. Solar water heating systems are most likely to be cost effective for facilities with water heating systems that are expensive to operate, or with operations such as laundries or kitchens that require large quantities of hot water [10].

F. *Construction and necessary steps to increase the heat performance*

The new concept can be constructed by using a rectangular glass topped mini solar pond. This mini solar pond will contain salinity gradient water into a small water holder as like rectangular box. Salinity gradient of the water will be maintained by the diffuser process. The increase of thermal efficiency of the proposed model the following steps should be taken.

- *Wooden frame*

Wood will be used as frame or wall material due to heat insulation purposes as well as cost effective solution.

- *Black plastic sheet*

The inner surface of the wooden frame can be covered with a thin layer of black plastic sheet to protect the wooden frame from salty water. Black colour coating inside the frame will help in quick heat absorption process.

- *Glass plate*

To reduce the loss of heat energy from water into the air the top of the pond can be designed to be covered with the transparent glass. Transparent glass plate helps in greenhouse effect that trapped thermal energy inside the pond. The black bottom of the pond heats up when struck by sunlight. UCZ layer will collects the heat from sun light passes through glass plate. The internal heat energy is trapped inside the pond. As a result pond salinity gradient water is also heated up gradually.

- *Metal plate*

Due to transfer heat easily from lower convective zone to the bottom, metal plate is used. Metal sheet should not be too thick to pass the heat too thin to create imbalance condition [2, 5].

Figure 3 indicates the different fundamental parts (glass plate, black plastic sheet, wooden layer, metal body) of the proposed model.

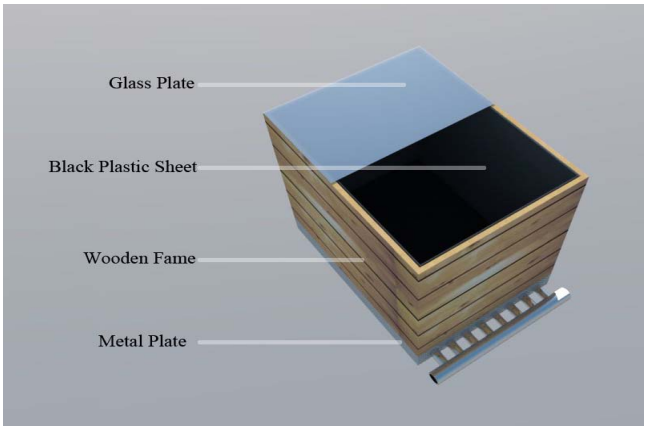



Figure 3. Proposed integration of solar heater and solar pond model.

## IV. HEAT FLOW ANALYSIS

A. *Sensible heating system*

Sensible heat is heat exchanged by a body or thermodynamic system that has as its sole effect a change of temperature. The term is used in contrast to a latent heat, which is the amount of heat exchanged that is hidden, meaning it occurs without change of temperature. For example, during a phase change such as the melting of ice, the temperature of the system containing the ice and the liquid is constant until all ice has melted. The sensible heat of a thermodynamic process may be calculated as the product of the body's mass (m) with its specific heat capacity (c) and the change in temperature ($\Delta T$) [11].

$$\Delta T = (T_H + T_L) \quad \text{.....(1)}$$

$$Q_{sensible} = mc\Delta T \quad \text{.....(2)}$$

$$Q_{air} = m_{air} C_p \Delta T \quad \text{.....(3)}$$

$$m_{air} = (D \times V)kg \quad \text{.....(4)}$$

$$D = \frac{P}{RT} \quad \text{.....(5)}$$

$$V_{air} = length \times height \times width \text{ t} \quad \text{.....(6)}$$

$$Q_{total} = Q_{metal} + Q_{air} \quad \text{.....(7)}$$

## V. MATLAB SIMULATION

In solar heater and solar pond heat transfer and heat absorbing through convective, conductive heat transfer system

in sensible heating process. The process is shown through simulation.

Figure 4 shows the Matlab Simulink model of the proposed model

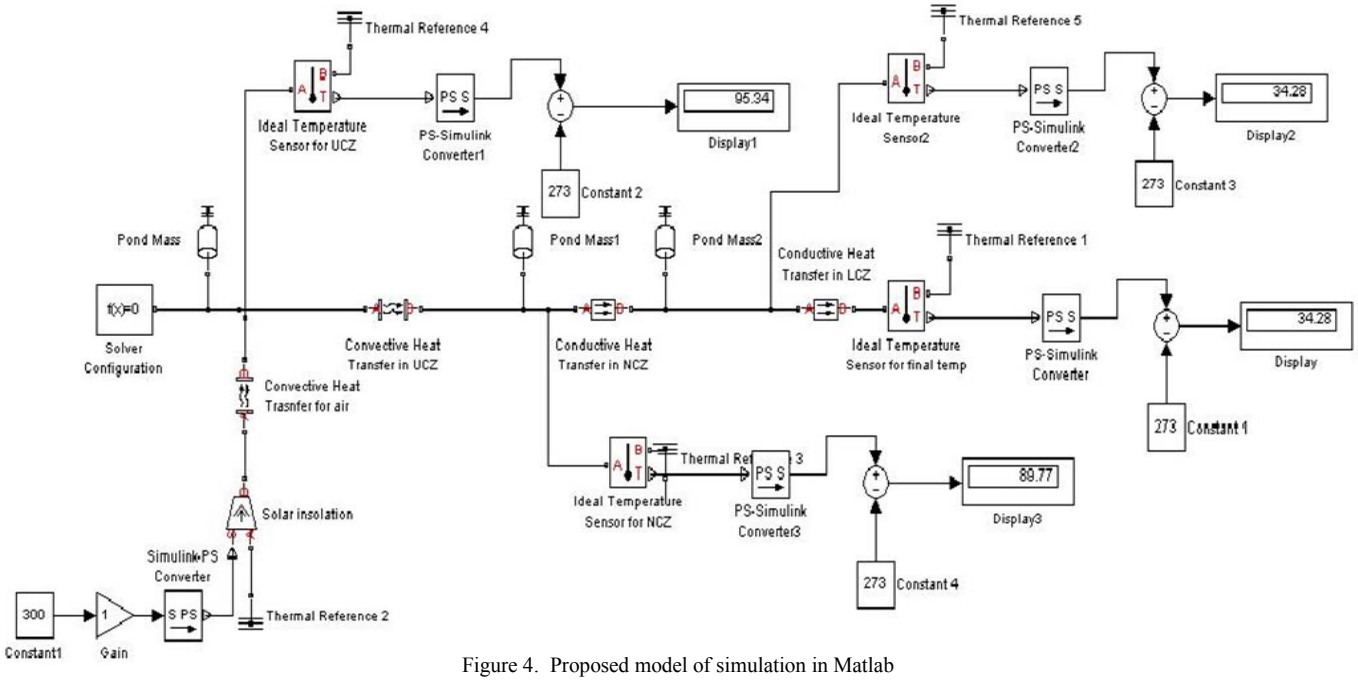



Figure 4. Proposed model of simulation in Matlab

### *A. Simulation output*

The line diagram is depicted below according to the different density of water (normal, gradient salinity and salinity) for a fixed time sixty seconds. As output, there are three heat performance graphs Blue, Red, Green for normal water, gradient salinity water and salinity water respectively are shown.

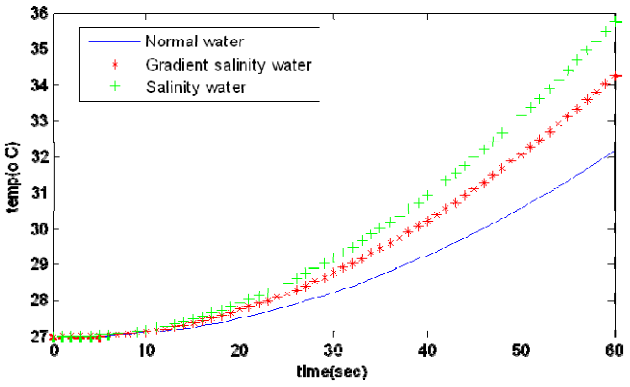



Figure 5. Temperature vs Time curve for normal, gradient and salinity water

For normal water, temperature peaked slightly above 32 degree Celsius which is lowest point. On the other hand for salinity water red graph line reached near to 36 degree Celsius sharply this is the highest peak point. And heat performance graph for gradient salinity water stays in between these two points above 34 degree Celsius. The graph emphasis that, the changes of salinity rate is responsible for changing the thermal performance of water.

## VI. CONCLUSIONS

The implementation of the given model will be efficient for enhancing the thermal efficiency of water heating system as well as storage of capacity of heat. In order to minimize the cost and increase the performance wooden box, metal plate, glass plate is considered. Based on numerical calculation it is shown that integration of solar heater and solar pond is suitable for collecting as well as supplying maximum heat to the liquid of solar heater. On the other hand in the Matlab simulation it is shown that high concentration brine salinity water gives better output means high temperature than normal water and gradient water for a fixed time. The increasing salinity of the water simultaneously increases the efficiency of the suggest model.

## VII. FUTURE WORK

The prototype of the project did not reach at the construction due to unavailabilty of suitable material as well as modern technology. So the main future intention of the project is to construct prototype and check the performace in practical basis.We aslo predict that this concept will help to minimize cost in power generation sector [3].

## REFERENCES


[1] Huanmin Lu, Andrew H. P. Swift, Herbert D. Hein and John C.Walton, "Advancements in Salinity Gradient Solar Pond Technology Based on Sixteen Years of Operational Experience", *J. Sol. Energy Eng.,* 126(2), pp. 759-767, 2004.

[2] Nuri O., Mehmet, K., Nalan, C.B., "A Solar Pond Model With Insulated and Glass Covered Surface, *Bulgarian Journal of Physics*, 27(4), pp. 67-70, 2000.

[3] Omer B., A. Ceylen S., "SOLAR PONDS," PHYS. *471 SOLAR ENERGY FALL* 2004.

[4] Celestino A., Erminia L., "A one-dimensional numerical study of the salt diffusion in a salinity-gradient solar pond", *International Journal of Heat and Mass Transfer*, Vol. 47, Issue 1, pp. 1-10, 2004.

[5] Hillel, R., Giorgio, A. B., "The Advanced Solar Pond (ASP): Basic Theoretical Aspects," *Solar Energy*, Volume 43, Issue 1, pp. 35-44, 1989.

[6] Singh, B.; Tan, L.; Date, A.; Akbarzadeh, A., "Power generation from salinity gradient solar pond using thermoelectric generators for renewable energy application," *Power and Energy (PECon), 2012 IEEE International Conference on* , vol., no., pp.89,92, 2-5 Dec. 2012

[7] U. Fisher, J. Weinberg, B. Doro, "Integration of Solar Pond with Water Desalination", RENEWABLE ENERGY SYSTEMS AND DESALINATION – Vol. II, Encyclopedia of Desalination and Water Resources (DESWARE) .

[8] Solarpond. http://edugreen.teri.res.in/explore/renew/pond.htm

[9] Soteris A. Kalogirou, "Solar thermal collectors and applications", *Progress in Energy and Combustion Science*, Volume 30, Issue 3, pp. 231-295, 2004.

[10] Hazim Al-Hussaini, Ibrahim A. Alenezi Adel O. Sharif, "New Method for Predicting the Performance of Solar Pond in any Sunny Part of the World", *World Renewble Energy Congress*, 2011.

[11] J. Srivasan, "Solar pond technology," *Sadhana*, Vol. 18, Part 1, pp. 39-55, 1993.

# Hybrid Energy Assisted Electric Auto Rickshaw Three-Wheeler

Necolus Shaha[1*], and Md. Bashir Uddin[2]
[1] Dhaka University of Engineering & Technology, Bangladesh.
[2] Dhaka University of Engineering & Technology, Bangladesh.
[*]E-mail: rinkushaha@gmail.com

*Abstract*— **Hybrid energy technology has become the latest goal in the automobile industry. This paper proposed an auto rickshaw which is based on hybrid energy storage system that operates in an environmentally friendly way. The main source of power of the proposed auto rickshaw is battery and the battery module is charged by the two ways- the on-board plug in battery charger and the solar panel. The battery is fully charged by the plug in battery charger and the PV panel is trying to keep the state of full charged battery module by continuous trickle charging. The charge sustaining capability enables the vehicle to accelerate faster, enhance the cruising capability, driving range, battery life-span, and reduce road accident. The proposed vehicle model is analyzed by the Advanced Vehicle Simulator (ADVISOR) software. Finally, the analysis showed that the use of hybrid energy technology can effectively improve the vehicle performance and more profitable than the prevailing auto rickshaw.**

***Keywords—Hybrid energy, energy status, ADVISOR.***

## I. INTRODUCTION

The first HEV was built in 1898, and there were several automotive companies who were selling HEVs in the early 1900s [1]. But the production of HEVs did not proceed due to the requirement for a smooth coordination between the engine and the motor, which was not possible due to the use of only mechanical controls and the poor efficiency compare to the combustion engine vehicle. Motorization was increasing day by day with the development of industrialization and urbanization.

The oil crisis of the 70s and the growing threat of global warming, excessive fossil fuel dependence, and increasing the fuel prices which have accelerated the development of Hybrid Electric Vehicles (HEV) into the political spotlight. Hybrids have been looked at as a possible solution to resolve consumption and pollution problems without having to reduce performance or range compared to a normal car. Therefore, serious research on hybrid cars began in the late 70s [2]. Electric drives are becoming very popular in the control of hybrid vehicles. The revolution in the design of automobile's electrical system creates a very large and diverse market for a practical 48V electrical system, new electrical functions, alternative electrical sources and necessity of power electronic controls and interfaces [3], [4]. Many automobile industries developed hybrid vehicles namely Honda Insight, Honda Civic, Toyota Prius, Volkswagen, Venturi Astrolab, Chevrolet Volt etc.

The hybrid vehicles still have the disadvantage of a too high price. Also this type of hybrid vehicles (a four seated middle class car) does not really fulfill the needs of the Bangladeshi society and does not meet the requirements of a typical Bangladeshi city either. Conventional auto rickshaws are well suited to the Bangladeshi society and environment. They are small and narrow, allowing maneuverability on congested roads and travel cost is also tolerable for middle class family. So, conventional auto rickshaws were adaptable to cover a small distance in the development country. The auto rickshaws are powered by the rechargeable batteries and they are becoming very popular day by day. Since Bangladesh faces acute power crisis and it is a matter of great regret that a lot of power is used daily to recharge the batteries. For this reason, the power crisis is increasing. In this situation, the hybrid energy is essential to recharge the batteries.

Bangladesh is situated between $20^0$34' and $26^0$38' North Latitude and $88^0$01' and $92^0$41' East Longitude and the climate is tropical [5]. It gets abundant sunlight year round. The monthly average solar insolation at different Locations of Bangladesh is given in Table I [6]. The Table I showed that the monthly solar insulation is the highest in Rajshahi and the lowest in Sylhet. The daily average of bright sunshine hours at Dhaka city is given in Table II and showed that the daily sunlight hour is around the range from 10 to 7 hours [7]. Maximum amount of radiation is available on the month of March-April and minimum on December-January [6].

The total solar energy reaching in Bangladesh is $180\times10^9$ MWH/year which is $10^5$ times the energy generated as electricity [5]. So, the solar resource in Bangladesh is rich and suitable form of renewable energy for urban region because of availability of plenty of sunshine. In effective operation, keeping the produced energy by this green source and the designed stand-alone system can operate economically. The vehicle is emission, noise and maintenance free. The oil prices all over the world are increasing day by day. In this system, our country does not depend on foreign oil. At night time charging of EVs will help to balance the load and improve the power plant efficiency.

TABLE I. MONTHLY SOLAR INSOLATION (KWH/M$^2$) AT DIFFERENT LOCATIONS OF BANGLADESH

| Month | Dhaka | Rajshahi | Sylhet | Bogra | Barisal | Jessore |
|---|---|---|---|---|---|---|
| Jan | 4.03 | 3.96 | 4.00 | 4.01 | 4.17 | 4.25 |
| Feb | 4.78 | 4.47 | 4.63 | 4.69 | 4.81 | 4.85 |
| Mar | 5.33 | 5.88 | 5.20 | 5.68 | 5.30 | 4.50 |
| Apr | 5.71 | 6.24 | 5.24 | 5.87 | 5.94 | 6.23 |
| May | 5.71 | 6.17 | 5.37 | 6.02 | 5.75 | 6.09 |
| Jun | 4.80 | 5.25 | 4.53 | 5.26 | 4.39 | 5.12 |
| Jul | 4.41 | 4.79 | 4.14 | 4.34 | 4.20 | 4.81 |
| Aug | 4.82 | 5.16 | 4.56 | 4.84 | 4.42 | 4.93 |
| Sep | 4.41 | 4.96 | 4.07 | 4.67 | 4.48 | 4.57 |
| Oct | 4.61 | 4.88 | 4.61 | 4.65 | 4.71 | 4.68 |
| Nov | 4.27 | 4.42 | 4.32 | 4.35 | 4.35 | 4.24 |
| Dec | 3.92 | 3.82 | 3.85 | 3.87 | 3.95 | 3.97 |
| Avg | 4.73 | 5.00 | 4.54 | 4.85 | 4.71 | 4.85 |

Source: Mondal, M. A. H., 2005, p.29

TABLE II. DAILY AVERAGE OF BRIGHT SUNSHINE HOURS AT DHAKA CITY

| Month | Daily Mean | Maximum | Minimum |
|---|---|---|---|
| January | 8.7 | 9.9 | 7.5 |
| February | 9.1 | 10.7 | 7.7 |
| March | 8.8 | 10.1 | 7.5 |
| April | 8.9 | 10.2 | 7.8 |
| May | 8.2 | 9.7 | 5.7 |
| June | 4.9 | 7.3 | 3.8 |
| July | 5.1 | 6.7 | 2.6 |
| August | 5.8 | 7.1 | 4.1 |
| September | 6 | 8.5 | 4.8 |
| October | 7.6 | 9.2 | 6.5 |
| November | 8.6 | 9.9 | 7 |
| December | 8.9 | 10.2 | 7.4 |
| Average | 7.55 | 9.13 | 6.03 |

Source: Bashar, REEIN, 2010a

With the factors of pollution, increased traffic in mind, drive range, charging cost, battery life and power crisis the best way to revamp the auto rickshaw is to develop a more efficient design that will be powered by a nonpolluting hybrid energy source, which can be achieved with an electric drive train since there are zero pollutants at the tailpipe. A hybrid energy source would make it a better solution compared to the prevailing alternative-fuel-powered rickshaws and auto rickshaw.

In this paper, hybrid energy for auto vehicles is presented and compared to the conventional auto vehicles. The main purpose of this study is to evaluate the hybrid energy based auto vehicles for zero environmental pollution and to propose an efficient model for increasing the driving range, driving speed, minimize the charging time, enhance battery life and reduce the charging cost. This paper also established the economic exposition of the proposed system. The proposed vehicle model is design and analyzed by the Advanced Vehicle Simulator (ADVISOR) software. For design and parameters analysis the data are collected from local markets in the Dhaka city.

## II. CONVENTIONAL AUTO RICKSHAW

Auto rickshaws are three-wheeled vehicles that are widely used in many Asian countries as taxis or Easy bike for people. The existing electric three-wheelers are popularly known as Borak, E-Bike, Easy Bike, City Bike, auto etc. are now being used in 32 districts in Bangladesh. It has only battery powered electric vehicle. At night, the battery is charged around 8 to 10 hours by the on-board battery charger. At new condition, it covered 140km to 150km per day in a single charge with the top speed approximates 40 to 45 km/ph. The vehicle speed depends on the battery charge. The slope and bad road condition discharge the battery state of charge rapidly and reduces the drive range. At night, the driver turn on the head light and it reduces the drive range. Because of this, most of the drivers drive the vehicle "turns off" the head light due to increase their drive range. For this reason, the prolong road accident has been happened. The battery efficiency is decreasing constantly and the improper operation of the vehicle the battery may be damaged less than one year. After one year operation of the vehicle, the battery module is discharged rapidly and it is covered around 60km to 70km. Again, the battery module is charged around 2 to 3 hours at launch time and then it is covered around 30km to 40km. After one and half year, the battery module has been damaged and replaced it with new one.

## III. PROPOSED VEHICLE MODEL CONFIGURATION

The electric three wheelers are characterized by its tin/iron body supported by three small wheels (one in front and two at the back), with a seat for the driver with a passenger in the front and two bench which are seating four or six passengers in the rear. It has an open design: no doors for the driver or passengers, allowing immediate pick-ups and drop-offs. The redesign auto vehicles do not want to change any part of the aforementioned characteristics. Only redesign its energy storage system. It is reflected on the auto rickshaw. The proposed vehicle will be well-suited with the old one and it is easily operated by the drivers and comfortable for the passengers. The proposed vehicle model is shown in fig.1.

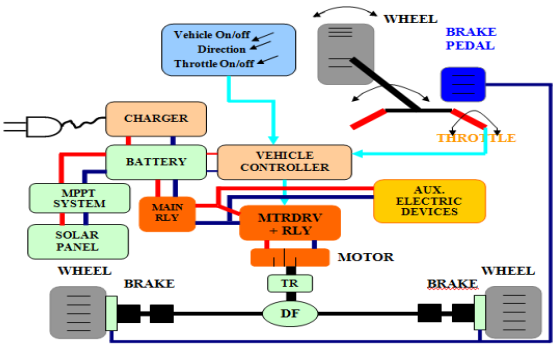



Figure 1. Proposed Auto Vehicle Model

The proposed vehicle is capable to use all over the places in Bangladeshi environment. This vehicle is also small, narrow and environmental friendly. The vehicle top speed is 45km/h and usually carries one to seven passengers with a driver. It covers a small distance 170 to 190 km per day in a single charge. The prototype system consists of a Solar panel, a Plug in Battery charger, a Maximum Power Point Tracking (M.P.P.T) system, rechargeable battery module, Power controller, Control circuitry, Instrumentation system and DC series excitation motor.

At night, the battery module is charged at 5 to 6 hours by the Plug in on-board battery charger system and at day, the PV panel with M.P.P.T system is continued trickle charge the state of full charge battery. In this way, the PV system is tried to keep the battery state of charge. So, the large current discharging of battery is avoided, enhance the cruising capability and the driving range. Thus, the battery life-span is extended. When the battery is charged by the plug in system, the control circuitry cut off the M.P.P.T circuit, which is sensed by the instrumentation system. The M.P.P.T control circuit not only takes maximum power from the solar panel but also monitor the condition of the battery.

The hybrid energy is stored in the Vehicle's rechargeable battery module. The battery module supplies power to the motor by the power controller. The power controller is a device, which controls the amount of power supplied to the electric drive motor(s) based on the position of the accelerator through the control circuitry. The electrical power supplied to the electric drive motor(s) is used to generate an electromotive force, which turns the shaft of the electric motor(s). This shaft is coupled to the wheels of the vehicle and causes movement either forward or reverse, depending on the direction the shaft is turning through a gearbox.

The dc motor actuator is driving the vehicle, a power controller for matching the various voltage and power levels of the battery according to the motor speed, a control circuitry control the driving speed and human interface. The display shows the battery performance through the instrumentation system. In this way, the electric energy is converted to mechanical energy that drives the vehicle. The proposed vehicles physical specification is given in the Table III.

TABLE III. PHYSICAL SPECIFICATION FOR THE PROPOSED VEHICLE

| Components | Value |
|---|---|
| Outline dimension(L*W*H) | 2650*1000*1650 mm |
| DC series excitation Motor | 60V--1000W |
| Top speed | 50km/h |
| Front wheel | 3.25-16 |
| Back wheel | 4.00-12 |
| Break distance | 25km/h less than 4m |
| Storage battery | 12V, 120Ah(5 set) |
| Charger voltage | 220v-(50Hz) |
| Solar panel | 280W, 45~47V |
| Solar panel size | 1955*982mm |
| Daily distance covered | 170km and 190km |
| Loading capacity | 450~500Kg |
| Vehicle weight | 300~400Kg |

## IV. ENERGY STATUS OF THE AUTO VEHICLES

The energy status is the heart part of the auto vehicles. The auto vehicles start at morning and stop at night. The vehicle operator covered more than10Km at 30 minutes. The conventional auto vehicles total running hour is 6 hour and it is covered 140Km per day in a single charge. The proposed auto vehicle total running hour is more than 8 hour and it is covered 190Km per day. In order to improve the vehicle efficiency, a hybrid energy storage scheme has been proposed. The battery module is charged by the on-board home charger and solar system in a convenient way. The hybrid energy storage system block diagram is shown in figure 2.

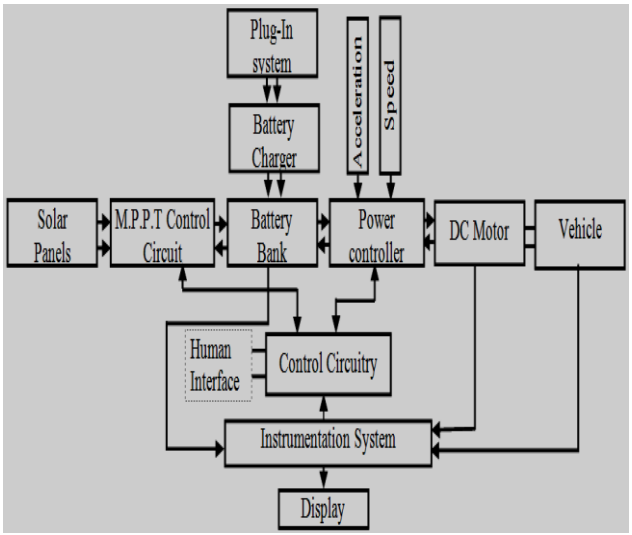



Figure 2. Hybrid Energy Storage System

### A. Home Charge

When the auto vehicle is parked at home at night, the vehicle on-broad-charger can be connected to a single phase AC plug for slow night time charging. The battery is recharged at the correct charge rates and the current is automatically switched off when the charge is completed. In Figure 2 shows block diagram of a typical controlled on-board-charger. Depending on the battery capacity and depth of discharge, the charging time takes about 5-7 hrs and charging current is usually limited to 12A to 18A. This on-board-charger should be light & inexpensive. As the electricity demand is relatively low at night, this home charging scheme can facilitate the low level control of power utilization.

### B. Solar Energy

In order to further enhance the vehicle drive time, batteries can also be charged by the solar panel embedded in the vehicle roof. The rickshaw has about 2.65 $m^2$ of space available on the roof alone to put solar panels. The output characteristic of PV cells could be expressed by volt-ampere (I-V) characteristic. The I-V characteristic is changed with the solar radiation intensity(S) and temperature (T). The MPPT systems can be designed in such a way to achieve great output even under changing atmospheric conditions, shading, or irradiance conditions, such as those that would inevitably occur on the roof of a vehicle and in addition, to prolong the life-span of

the battery, a three-stage charging method is controlled the battery charge [8, 9]. A fully charged battery needs only trickle charge. The solar panel specification is given in Table IV.

TABLE IV. 280W SOLAR PANEL SPECIFICATION

| Parameter | Value |
|---|---|
| Power Output | 280W |
| Module Efficiency | 14% |
| Cell Efficiency | 17% |
| Voltage at $P_{max}$,$V_{mpp}$ | 36V |
| Current at $P_{max}$,$I_{mpp}$ | 8.2A |
| Open circuit voltage, $V_{oc}$ | 45~47V |
| Short circuit current, $I_{sc}$ | 8.4A |

(Source: Rahimaafroz Bangladesh)

In Bangladesh, the monthly average solar radiation is around 4.5 to 5 kWh /$m^2$/day and the MPPT system track the maximum power from the panel. Considering only being able to capture about 5–10% of that energy due to inefficiencies of the panel, converters, dust, shadow and less-than-ideal tracking conditions, the actual energy recoverable per day may be more than 2kWh when using the entire surface. So, to find out the daily energy generation, multiply the (280*90%) 252W to monthly average sunshine duration. The calculated data is shown in Table V. From the Table V found that the monthly average energy generation from the solar system is 1902Wh per day.

TABLE V. DAILY AVERAGE ENERGY STORAGE FROM SOLAR PANEL

| Month | Sunshine duration(hr) | Daily energy(Wh) |
|---|---|---|
| January | 8.7 | 2192.4 |
| February | 9.1 | 2293.2 |
| March | 8.8 | 2217.6 |
| April | 8.9 | 2242.8 |
| May | 8.2 | 2066.4 |
| June | 4.9 | 1234.8 |
| July | 5.1 | 1285.2 |
| August | 5.8 | 1461.6 |
| September | 6 | 1512 |
| October | 7.6 | 1915.2 |
| November | 8.6 | 2167.2 |
| December | 8.9 | 2242.8 |
| Monthly Average | | 1902.6 |

## V. ECONOMICS EXPOSITION

The fixed cost is calculated by the survey of the local market. The wiring, installation and maintenance cost is considered approximately. The cost of these components may be varied depending on its brand, quality, place and quantity. The electrical components may be damaged in any time by miss operation. Here the component life time and cost is considered by the market survey upon the fifteen auto vehicles. The Table VI shows the components estimated cost with the life time.

The daily battery module charging cost of the vehicle is two hundred taka and the running cost is changeable in each year. After one year, the prevailing system is covered 120Km per day and the proposed system is covered more than 170Km per day. The minimum fare is five taka for each passenger. The night time fare is two taka more than day time. The Table VII and VIII show the daily income. The Table IX shows the yearly running cost and income analysis of the prevailing system and the proposed system. Comparing the two systems, the hybrid energy system is more profitable than the prevailing system.

TABLE VI. DAILY COST WITH LIFE TIME

| Sl. No. | Component Name | Quantity | Life Time | Cost in Taka |
|---|---|---|---|---|
| 01 | Solar Panel | 1 | 20 Year | 30000 |
| 02 | MPPT controller | 1 | 20 Year | 15000 |
| 03 | Motor | 1 | 2 Year | 7000 |
| 04 | Motor controller | 1 | 6~12 month | 1000 |
| 05 | Carbon brush | 2 | 2~3 month | 80~100 |
| 06 | Pick up | 1 | 6~12 month | 300~450 |
| 07 | Gear box | 1 | 1 year | 400~800 |
| 08 | Battery | 5 | 12~18 month | 60000 |
| 09 | Acid water | 5~8 L | 18 month | 150~180 |
| 10 | Battery Charger | 1 | >2 year | 3500~4000 |
| 11 | Contact Switch | 1 | | 300 |
| 12 | Wheel | 3 | 6~8 month | (2100~2300)*3=6300~6900 |
| 13 | Tube | 3 | 6~8 month | (300~320)*3=900~960 |
| 14 | Brake shoe | 2 | 2~3 month | 120~150 |
| 15 | Bearing | 6 | 3~6 month | (40~200)*6=240~1200 |
| 16 | Head Light | 1 | 10~12 month | 100~120 |
| 17 | Horn | 1 | 2~3 month | 70~150 |
| 18 | Contact Switch | 1 | | 300~350 |
| 19 | Body and others | | | 81340 |
| Total | | | | 2,07,300~2,10,000 |

TABLE VII. Daily Income in the Prevailing System

| Sl. No | No. of passenger | Fare in (Tk) | Total trip | Distance Cover | Income (Tk) |
|---|---|---|---|---|---|
| 1 | 7 | 10 | 14 | 140 | 980 |
| 2 | 7 | 10 | 12 | 120 | 840 |

TABLE VIII. Daily Income in the Proposed System

| Sl. No | No. of passenger | Fare in (Tk) | Total trip | Distance Cover | Income (Tk) |
|---|---|---|---|---|---|
| 1 | 7 | 10 | 14 | 190 | 1330 |
| 2 | 7 | 10 | 12 | 170 | 1190 |

TABLE IX. YEARLY COST ANALYSIS

| Year | Running Cost in taka | Income in prevailing system (Tk.) | Income in proposed system (Tk.) | Profit in prevailing system (Tk.) | Profit in proposed system (Tk.) |
|---|---|---|---|---|---|
| First | 99,000 | 3,57,700 | 4,78,800 | 258,700 | 379,800 |
| Second | 1,67,500 | 3,27,600 | 4,28,400 | 160,100 | 260,900 |
| Third | 1,60,000 | 3,27,600 | 4,28,400 | 167,600 | 268,400 |
| Four | 1,06,500 | 3,57,700 | 4,78,800 | 251,200 | 372,300 |
| Five | 1,60,500 | 3,27,600 | 4,28,400 | 167,100 | 267,900 |
| Six | 1,06,500 | 3,27,600 | 4,28,400 | 221,100 | 321,900 |

## VI. SIMULATION RESULT AND DISCUSSION

### *A. Conventional Vehicle Power Calculation*

Battery is the main power source for the prevailing auto vehicle. The vehicle uses five pieces of battery and each battery is 12V, 120Ah. Assume the depth of discharge (DOD) of a battery is 70%. The vehicle consume battery energy 5040Wh and it could cover 140Km in 6hour. So, the vehicle consumes 840Wh to covered 23.33Km in an hour. The power consumption calculation is given in below:

Battery capacity=120Ah
Battery voltage=12V
Total energy storage capacity= 120*12*5= 7200Wh
Battery Depth of Discharge=70%
Uses battery capacity=120Ah*70%=84Ah
Total usable vehicle energy= (84Ah*12V)*5=5040Wh
Total distance covered per day=140Km
Total driving time=6h
Vehicle Consume power per hour=5040/6=840W
Distance covered in an hour= 140Km/6=23.333Km

### **B.** *Proposed Auto vehicle Power Calculation:*

The proposed vehicle main source of energy is battery and solar system. The battery is charged by the on-board home charger in the night and it stored energy 5040Wh as like as the prevailing system. The power consumption calculation is given in below:

Now,
The vehicle total energy consume per day= 5040+2000=7040Wh
The vehicle consume power per hour=840W
Total running hour= 7040/840= 8.38h
Total distance covered= 23.333*8.38= 195.55Km

When the vehicle is started, the battery is discharged in the rate of 840Wh. But in the day time, the solar system is started to charge the batteries and it can be recovered 252Wh. The Solar system also acts as a pulse charging, and it reduces the battery charge-discharge current and prolongs the battery life span [16].So, the vehicle operating time is extended more than 2.3 hour and the driving range is extended more than 50Km.The vehicle is also continued operated at night with head light and avoid the prolong road accident.

The auto vehicle is designed and simulated by the ADVISOR software. ADVISOR is written in the MATLAB/Simulink and developed by the National Renewable Energy Laboratory. It is used to analyze performance, fuel economy, and emissions of conventional, electric, hybrid electric, and fuel-cell vehicles [17]. The actual components and parameters can be changed or added for these simulations and showed the vehicle performance without need to actually assemble a test vehicle. The first two ADVISOR input screens provide the interface to change the vehicle parameters and test the vehicle. By clicking "View Block Diagram," it may look at and change the Simulink blocks of the overall vehicle model. In this way, the solar panel model was chosen by subtracting a constant value from the required power of the power bus and also adding a corresponding weight to the vehicle, which can be done at the first input screen. The constant value is based on research for the power ratings of the feasible panel. This method will not show dynamics but will give the general results for the extended range of the vehicle and the effects on vehicle efficiency. To develop the standard driving cycle, some approaches were considered and some are similar to the Indian urban driving cycle [18]. From the ADVISOR output the vehicle efficiency curve, speed vs. time curve, battery energy discharging curve is shown in figure 5, 6 and 7, and it will meet the urban driving cycle.

The simulator results showed that the vehicle is covered more than 179.26 Km at an average speed 14.3 mph and maximum speed 56.7 mph. In the evening, the battery state of charge is falling rapidly. The hybrid energy storage system increases the vehicle performance and efficiency. The vehicle does not need to charge at day time and it reduces the charging cost.

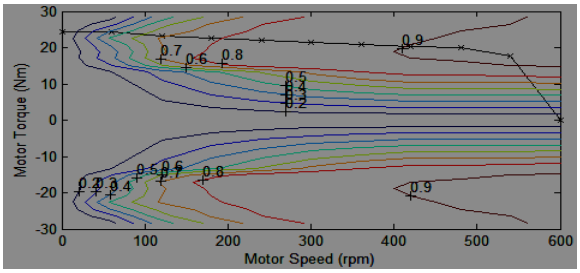



Figure 5. Motor Torque vs. Speed

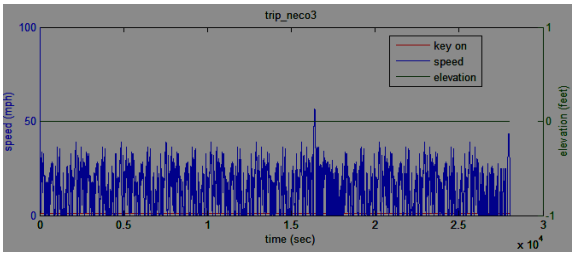



Figure 6. Speed vs. Time

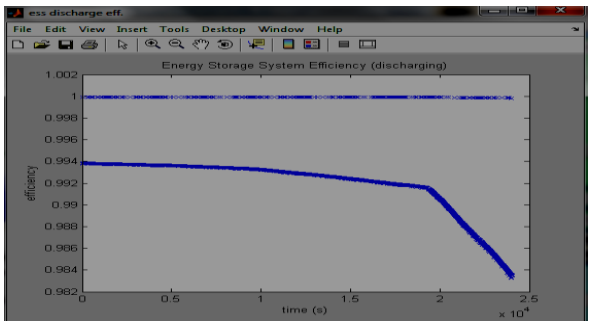



Figure 7. Energy Storage System Efficiency

## VII. CONCLUSION

The auto vehicle plays a fundamental role in the world. Moreover, this vehicle is very popular day by day for the developing country due to the low transportation cost with comfortable journey and zero pollution. Research showed that there are adequate renewable energies to support the infrastructure development of the auto vehicles. Thus, simulations have been performed on the electric vehicle supporting infrastructure and built in a prototype system. The system facilitate with such feature that reduces the energy dependency, eliminate the additional pressure of the grid in the day time, the drivers will not be worried about the battery charge, increase the driving range and avoid the prolong road accident at night. In Future research, the vehicle will be practically developed and analysis the result. To further increase the drive range more efficient motor with controllers, solar system and mechanical losses will be explored.

## REFERENCES

[1] S. Diego, E. W. P. John, "Hybrid electric vehicles: current concepts and future market trends," *Rama de estijdiantes del IEEE de barcelona,* Tech. Rep.5-30, 2006.

[2] L. Umanand, Y. Pradeep, N.V. Chalapathi Rao and K.Gopakumar, "Design and development of low pollution three wheeler, Project Report," *Centre for Electronics Design and Technology*, Bangalore, 2002.

[3] A. Graf, D. Vogel, J. Gantioler and F. Klotz, "Intelligent power semiconductors for future automotive electrical systems," *17th Meeting Electronikim Kraflfahrzeug*, VDA, Munich, June 1997.

[4] I. G. Kassakian, "Automotive electrical systems- The power electronics market of the future," *Proceedings of the 15th Applied Power Electronics Conference,* Vol.1 pp. 3-9, 2000.

[5] M. Zaman, M. A. Islam, and M. A. R. Sarkar., (2006). "Two phase heat transfer in solar water heater," *National Seminar on Renewable Energy: Biomass/Bio Energy,* Dhaka, 24thMarch 2006.

[6] M. A. H. Mondal, (2005). "Technical and Socio-economical Aspects of Selected Village Based Solar Home Systems in Gazipur District, Bangladesh," M.Sc.- Thesis, SESAM- Sustainable Energy Systems and Management, International Institute of Management, University of Flensburg, Germany, March 2005.

[7] S. A. Bashar, (2009). "Solar power as a prime energy source in Bangladesh," [Online]. Available: http://www.thedailystar.net

[8] K. H. Hussein, I. Muta, T. Hoshino, and M. Osakada., "Maximum photovoltaic power tracking: an algorithm for rapidly changing atmospheric conditions," *IEEE Proceedings, Generation, Transmission and Distribution,* vol. 142, no.1, pp.59-64, 1995.

[9] S. Zheng and L. Wang., "Research on charging control for battery in photovoltaic system," *Proceedings of the 6th IEEE Conference on Industrial Electronics and Applications*, pp. 2321~2325, 2011.

[10] B. S. Borowy, and Z. M. Salameh., "Methodology for optimally sizing the combination of a battery bank and PV array in a wind/PV hybrid system [J]," *IEEE Transactions on Energy Conversion*, 11(2), pp. 367~375, 1996.

[11] B.Ai, H.Yang , H.Shen, X.Liao., "Computer-aided design of PV/wind hybrid system[J]," *Renewable Energy,* pp. 1491~1512, 2003.

[12] T.A. Singo, A.Martinez, and S.Saadate., "Using Ultracapacitors to Optimize energy storage in a photovoltaic system," *SPEEDAM International Symposium on Power Electronics*, pp. 229 ~234, 2008.

[13] C. Binggang, Z. Chuanwei, B. Zhifeng., "Trend of development of technology for electric vehicles," *Journal of Xi'an Jiaotong University*, 38(1), pp. 1-5, 2004.

[14] Binggang Cao and Z. B., "Research on control for regenerative braking of electric vehicle," in *Proc. IEEE Int. Conf. Vehicular Electronics and Safety, Oct.* 2005, pp: 92-97.

[15] B. E. Dickson, T. R. Ialik, and D. G. Herrey., "Characterization of a Fuel cell/Battery Hybrid system for EV applications*," SAE Trans.*, pp. 1961-1969, 1993.

[16] C.C. Hua and M. Y. Lin, "A study of charging method control of lead acid battery for electric vehicles," *Proc. of the IEEE Ind. Electron.*, vol. 1, pp.135-140.

[17] K. D. Wipke, M. R. Cuddy, and S. D. Burch, "ADVISOR 2.1: A user friendly advanced power train simulation using a combined backward/forward approach," *IEEE Trans. Veh. Technol.,* vol. 48, no. 6, pp. 1751–1761, Nov. 1999.

[18] *Gandhi*, *K.K.*, *Zvonow*, *V.A.* and Harbans, S., "Development of a driving cycle for fuel economy in a developing country," *Transportation Research*, Vol. 17, No. 1, pp. 1-11, 1983.
.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Study on Strain in InGaN-based Multijunction Solar Cell

Md. Aminur Rahman* and Md. Rafiqul Islam
Department of Electrical and Electronic Engineering
Khulna University of Engineering & Technology (KUET)
*E-mail: eee.amin@yahoo.com

*Abstract*— **In case of MJSC residual strain is induced due to the change in lattice constants in different layers. So far our knowledge efficiency of multijunction solar cell (MJSC) has not been studied considering the strain effect on the energy bandgap of subcells. In this paper, we have analytically studied residual strain in different MJSC structures using multilayered strain model. Three structures are investigated to realize the structure-dependent state of strain in MJSC. The results obtained from the present study demonstrate that the strain induced in MJSC depends not only on the numbers of subcells but also on the position of subcells, window, tunnel, and BSF layers as well as their thicknesses. The influence of strain found more in the structure indicated by MJSC-3 compared to the structures MJSC-2, and MJSC-1.**

***Keywords— Multi-Junction Solar Cell (MJSC), Strain, Window layer, Tunnel junction, Back Surface Field (BSF)***

## I. INTRODUCTION

Solar energy is an environmental pollution free unlimited renewable source of energy and can be converted into electrical energy using solar cell. However, the conversion efficiency of solar cell is very poor, that is, mainly caused by losses associated with the carrier recombination and low and high energy photon absorptions. To be competitive with the conventional energy resources, the conversion efficiency of a solar cell must be improved.In previous studies [1], [2] several solar cell approaches have been proposed like multijunction (MJ) solar cell, concentrator solar cell, intermediate bandgap solar cell, quantum well, quantum dot solar cells etc. Among them MJ approach is very much attractive to resolve the main issues related with efficiency degradation [3].

$In_xG_{1-x}N$ (0.64 to 3.4 eV) is a promising material for multijunction solar cell (MJSC), because its bandgap energy is nicely matched to whole solar spectrum [4]. The InGaN-based MJSC structure is composed of number of cells in which the bandgaps of the cells are divided by tuning the composition. For efficient absorption of whole solar energies, the higher bandgap cell is fabricated at the top and others are below the top cell [5].

In case of MJSC structure, cells are fabricated layer by layer resulting in layered change in lattice constant. Since the lower bandgap cell is placed at the bottom and higher at the top in InGaN-based MJSC structure, the lattice constant increases in different layers from bottom to top that leads to induce compressive strain and cannot be determined using simple epilayer strain model [6], [7]. In order to determine strain from the MJSC structure multilayered strain model [8] can be used where strain is calculated by integrating strain induced in different layers due to change in lattice constant.

In previous studies [3], [4], [9] the efficiency of MJSC was investigated without taking into account of strain. It is well known that the bandgap of semiconductor materials is modified under the influence strain. Under application of compressive strain the bandgap of InGaN is increased and opposite is happened for tensile strain [10]. It is therefore very much important to investigate the actual efficiency of MJSC on account of strain.

In this work, study of strain is carried out on different MJSC structures using multilayered strain model. Three types of MJSC structures are considered in the present study. The state of strain and its magnitude are investigated as a function of cell parameters. Finally, the strains calculated with respect to the number of layers are compared for different MJSC structures.

## II. MJSC STRUCTURES

The MJSC structures studied hereare shown in Figs 1, 2,and 3 [3], [8], [9]. The subcells are placed from bottom to top with lower to higher bandgap by adjusting the composition. In MJSC-1, the window layer is used at the top to reduce surface recombination, on the other hand, back surface field (BSF) at the bottom to reduce carrier scattering [11].The function of high bandgap window is also to reduce the cell's series resistance. In addition to window and BSF layers tunnel junctions are used in MJSC-2 to provide low electrical resistance and optically low loss path between two subcells [9]. Each unit cell is comprised of a window, BSF and tunnel junction in MJSC-3 structure [3].The bandgaps used in 3, 5, and 7 subcells of different MJSC structures are listed in Table 1. The bandgap of the material used in the window, tunnel junction, and BSF layers must be higher than the bandgap of adjacent subcell.

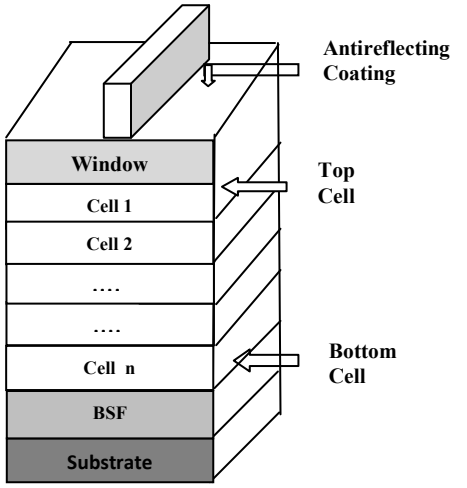



Fig. 1: The schematic illustration of proposed MJSC-1.

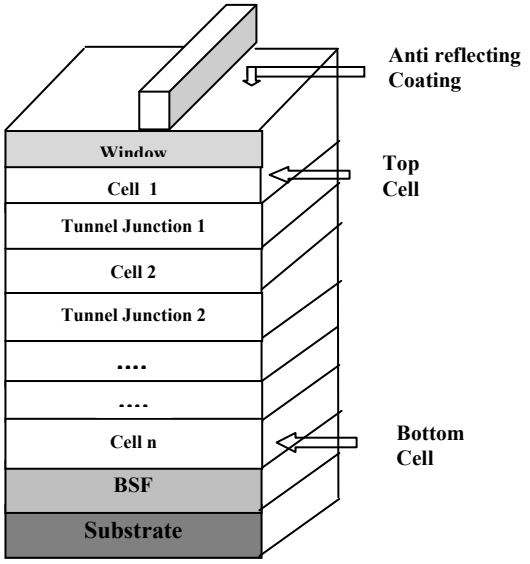



Fig. 2: The schematic illustration of proposed MJSC-2.

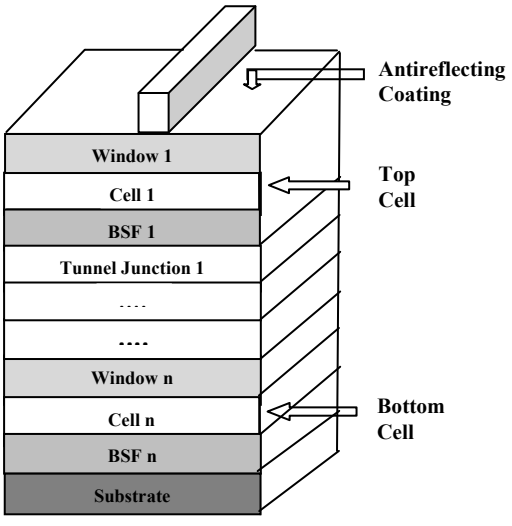



Fig. 3: The schematic illustration of proposed MJSC-3.

TABLE I
BANDGAP ENERGY DISTRIBUTION IN DIFFERENT SUBCELLS OF MJSC

| *No. of Stack* | *Values of Bandgap (eV)* |
|---|---|
| 3 | 0.7 1.37 2 |
| 5 | 0.53 0.95 1.4 1.93 2.68 |
| 7 | 0.47 0.82 1.191 1.56 2 2.5 3.21 |

## III. MULTILAYERED STRAIN MODEL

The multilayered strain model is developed [8] based on the multilayer structure shown in Fig. 4. The position and thickness of the layers are indicated by $y_n$ and $t_n$, respectively. A sacrificial layer is used to reduce lattice mismatch between the substrate and layer-1. Layer-1 to layer-n is grown with different lattice constants that lead to induce residual strain in the MJ structure.

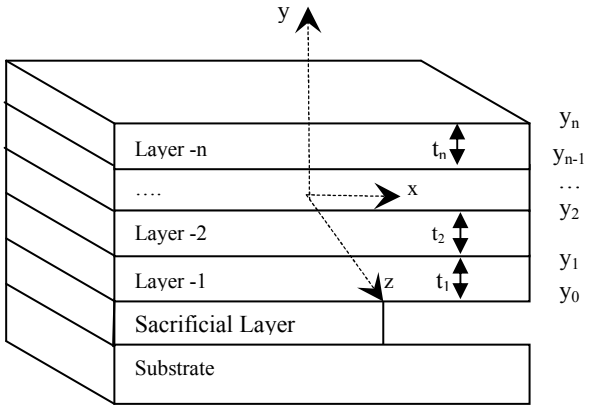



Fig. 4: Multilayer structure with n-layer

The plane strain in multilayered system can be given [8] by

$$\varepsilon = k + \frac{y - y_b}{G} \quad (1)$$

where $k$ is the uniform strain component and $G$is the radius of curvature. The parameter $y_b$,$G$, and $k$ can be represented by

$$y_b = \frac{\sum_{i=1}^{n} E_i' t_i (y_i + y_{i-1})}{\sum_{i=1}^{n} E_i' t_i} \quad (2)$$

$$G = \frac{-2\sum_{i=1}^{n} E_i' t_i \left[y_i^2 + y_i y_{i-1} + y_{i-1}^2 - 3y_b(y_i + y_{i-1} - y_b)\right]}{3\sum_{i=1}^{n} E_i' t_i (y_i + y_{i-1} - 2y_b(c - \eta_i \varepsilon_i^0 + v_i d)} \quad (3)$$

and

$$k = \frac{\sum_{i=1}^{n} E_i' t_i (\eta_i \varepsilon_i^0 - v_i d)}{\sum_{i=1}^{n} E_i' t_i} \quad (4)$$

where, thickness of each layer $t_i = y_i - y_{i-1}$. $E_i$ and $v_i$ are Young's modulus and Poisson's ratio. $E_i' = E_i/(1 - {v_i}^2)$ . Here $\eta_i = 1 + v_i$ and $\varepsilon_i^0$ is the initial strain. The parameter $d$ can be given by

$$d = \frac{(a_{12} - a_{11})a_{23}b_1 - (a_{21}b_1 - a_{11}b_2)}{(a_{11}a_{22} - a_{12}a_{21})a_{22} + (a_{12}^2 - a_{11}^2)a_{23}} \quad (5)$$

The details of Eqn. (5) are available in [8].

## V. RESULTS AND DISCUSSION

Using the multilayer strain model the strain induced due to change in lattice constants in different layers is calculated. Here we have studied three different MJSC structures shown in Figs. 1, 2 and 3 to understand structure-dependent state of strain. To calculate the strain the composition-dependent Young's modulus and Poisson's ratio for InGaN are determined by linear interpolation between the values available [12] for binaries GaN and InN. Initial strain plays an important role on the resultant strain, and primarily it depends on the difference of lattice constants between layer-1 and the layer below the layer-1. The strain are investigated for the subcell numbers 3, 5, and 7 as a function of subcells dimension while keeping the dimension of window, tunnel, and BSF layers are constant.

Figure 5 shows a comparison of cell position-dependent strains among the MJSC structures shown in Figs. 1-3 with subcells numbers 3 and cell thickness 100 nm. The cell position changes for different structures, because the number and placement of additional layers are not the same

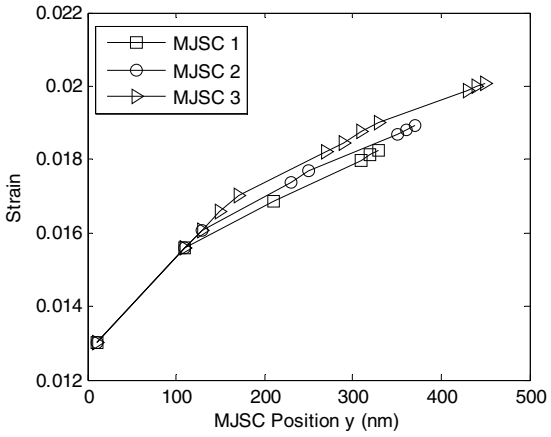



**Fig. 5:** Comparison of cell position-dependent strains in MJSC structures1, 2 , and 3. The results are obtained for the subcell numbers 3 having cell thickness 100nm.

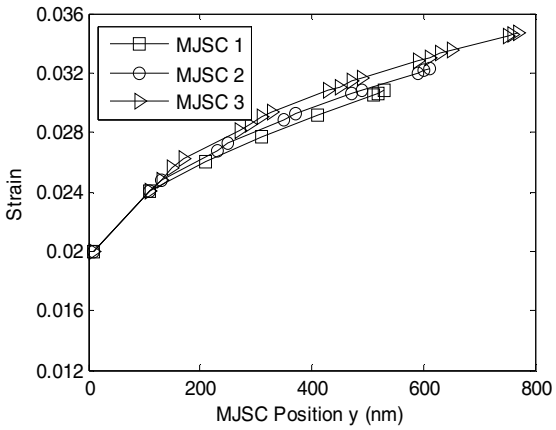



**Fig. 6:** Comparison of cell position-dependent strains in MJSC structures1, 2 , and 3. The results are obtained for the subcell numbers 5 having cell thickness 100nm.

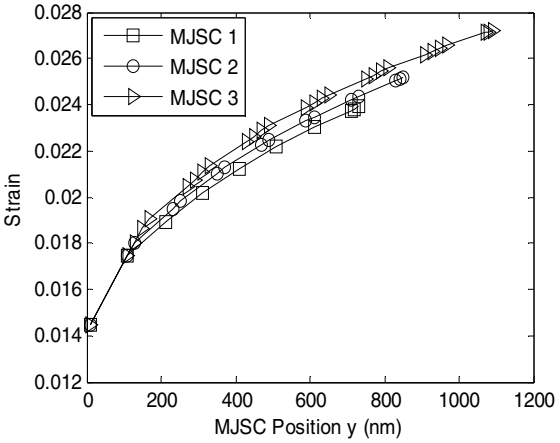



**Fig. 7:** Comparison of cell position-dependent strains in MJSC structures1, 2 , and 3. The results are obtained for the subcell numbers 7 having cell thickness 100nm.

in different structures. It is found that the magnitude of strain is the same up to the cell position 100 nm. When the cell position more than 100 nm the magnitude of strain changes for different MJ structures and is found high in MJSC-3 and low in MJSC-1. More number of layers in MJSC-3 may accumulate more strain. Similar tendency is found in Figs. 6 and 7 for the subcell numbers 5 and 7 except the initial strain. Because, the energy bandgap stepping of the subcells changes with the number of subcells as indicated in Table-1.Further the influence of strain of the subcells will not be same due to their asymmetrical placement in different MJSC structures. Since the bandgap of InGaN material is increased under the influence of compressive strain, the resultant bandgap of different subcells will be increased more for the MJSC-3 structure compared to MJSC-1.Also the increment of resultant bandgap will be more for more number of subcells for a particular MJSC structure. According to the results shown in Figs. 5, 6, and 7 the MJSC-3 structure will produce more open circuit voltage for the 7 subcells and thereby may be resulted more efficiency with respect to the efficiency obtained without taking into account of strain [5].

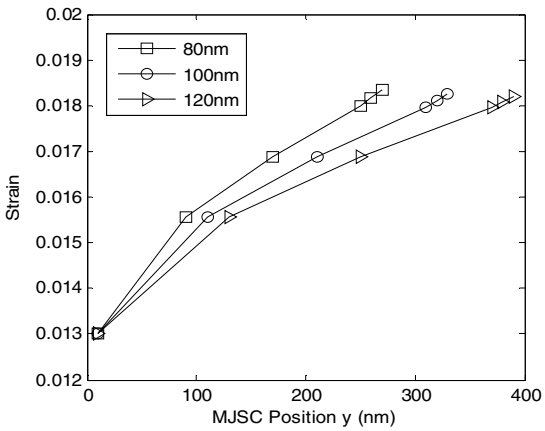



**Fig. 8:** Cell position dependent strain in MJSC-1 plotted for 3 subcells for the cell thickness of 80, 100 and 120nm.

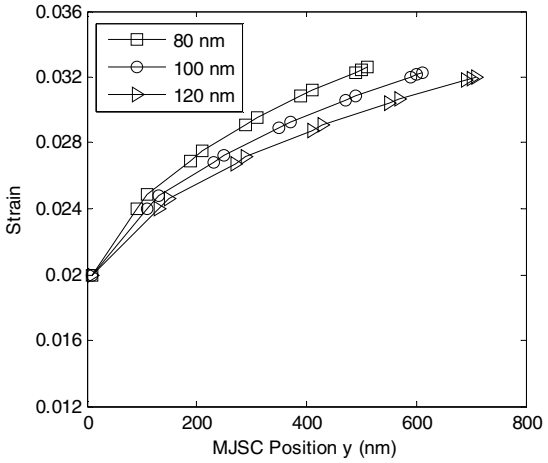



**Fig. 9:** Cell position dependent strain in MJSC-2 plotted for 5 subcells for the cell thickness of 80, 100 and 120nm.

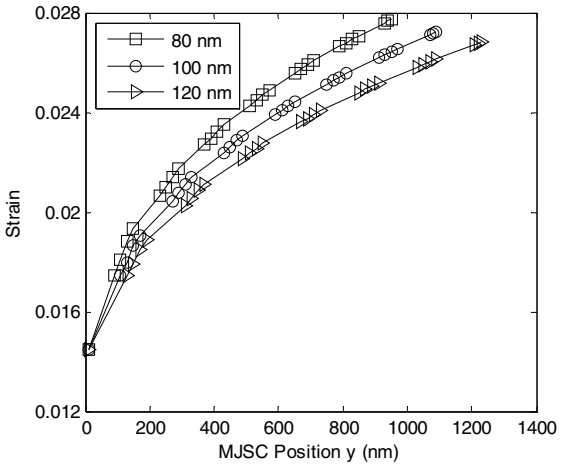



**Fig. 10:** Cell position dependent strain in MJSC-3 plotted for 7 subcells for the cell thickness of 80, 100 and 120nm.

In order to understand the subcell thickness-dependent state of strain, strains for MJSC-1 structure are plotted for the variation of cell thicknesses 80, 100, and120 nm and shown in Fig. 8. It is found that the magnitude of strain is reduced with increasing cell thickness. Similar results are also found for the MJSC-2 and MJSC-3 structures as shown in Figs. 9 and 10. It is well known that the strain is gradually relaxed with increasing thickness of epitaxial layer [12].

Since more efficiency is expected in strained subcells, it is expected that lower the thickness of subcells may give high efficiency. However, if the thickness of subcells becomes very narrow, electro-hole pair generation may decrease due to photon absorption loss. Therefore, there is an optimization of subcells thickness in order to have high efficient MJSC.

## VI. CONCLUSION

We have studied three types of MJSC structures composed of InGaN materials to investigate the residual strain induced due to change in lattice constant in different layers. The quantitative amount of strain is determined from the analytical approach developed for multilayered structure [8]. The results obtained from the present study demonstrate that the magnitude of strain depends on the difference in lattice constant among the adjacent subcells, particularly, on the initial strain which is induced due to the change in lattice constant between the bottom layer and the layer on which it is grown. Further, strain magnitude is strongly dependent on the number of layers and the thickness of subcells. The maximum strain is found for the MJSC-3 structure, and minimum in MJSC-1 compared to the structure MJSC-2. It is expected that the MJSC-3 structure may be more efficient than other structures due to more open-circuit voltage caused by strain-induced opening of bandgap under compressive strain.

## REFERENCE

[1] M. A. Green and K. Emery, "Solar cell efficiency tables (version 37)," *Progress in Photovoltaic,* vol. 19, no. 1, pp. 84–92, 2011.

[2] T. Trupke, M. A Green, and P. Wurfel, "Improving solar cell efficiencies by up-conversion of sub-band-gap light," *J. of Appl. Phys.,* vol. 92, no. 7, pp. 4117-22, 2002.

[3] M. Rafiqul Islam, M. T. Hasan, A. G. Bhuiyan, M. R. Islam and A. Yamamoto, "Design and Performance of $In_xGa_{1-x}N$ -based MJ Solar Cells," *IETECH Journal of Electrical Analysis*, vol. 2, no. 4, pp. 237-243, 2008.

[4] A. G. Bhuiyan, A. Hashimoto, and A. Yamamoto, "Indium nitride (InN): A review on growth, characterization, and properties," *J. Appl. Phys.,*vol. 94, no. 5, pp. 2779-2808, 2003.

[5] M. R. Islam, M. T. Hasan, M. A. Rayhan and Ashraful G. Bhuiyan, "High Efficiency $In_xGa_{1-x}N$-Based Multi-Junction Solar Cells with Concentrator," *Journal of Electrical Engineering, JEE (IEB)*, vol. 33, no. 1 and 2, pp. 48-52, Dec. 2006.

[6] J-H. Wong, B-R. Wu, and M-F Lin, "Strain Effect on the Electronic Properties of Single Layer and Bilayer Graphene," *J. Phys. Chem.*, vol. 116, pp. 8271−8277, 2012.

[7] G. P. Nikishkov, "Curvature estimation for multilayer hinged structures with initial strains," *J. Appl. Phys*., vol. 94, no. 8, pp.5333-5336, 2003.

[8] Y. Nishidate and G. P. Nikishkov, "Generalized plane strain deformation of multilayer structures with initial strains," *J. Appl. Phys*., vol. 100, no. 11, pp. 113518-4, 2006.

[9] X. Shen, S. Lin, F. Li, Y. Wei S. Zhong, H. Wan, J. Li, "Simulation of the InGaN-based tandem solar cells," *Proc. of SPIE*, vol. 2, pp. 7045-70450, 2008.

[10] J. Qi, X. Qian, "Strain-Engineering of Band Gaps in Piezoelectric Boron Nitride Nanoribbons," *Nano Lett.,* vol. 12, pp. 1224−1228, 2012.

[11] H. Salhi, H. Samet, M. Ben Amar, "Effect of BSF layer on the performance of a monocristalline solar cell" *International Renewable Energy Congress.,* Sousse, Tunisia, ID108, November 5-7, 2010.

[12] C. G. Van de Walle, M. D. McCluskey, C.P. Master, L.T. Romano, N.M. Johnson, "Large and composition-dependent band gap bowing in $In_xGa_{1-x}N$ alloys," *Materials Science and Engineering,* vol. 59, pp. 274–278, 1999.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# The Feasibility Study of Solar Irrigation: Economical Comparison between Diesel and Photovoltaic Water Pumping Systems for Different Crops

Md. Tareek-Al-Islam Khan, Suvasish Sarkar, Shakhawat Hossain, Ahsan Uddin Ahmed, and Bishwajit Banik Pathik

Dept. of Electrical and Electronic Engineering, American International University-Bangladesh (AIUB)
E-mail: tareek.islam@gmail.com

***Abstract*— This article investigates the feasibility of solar powered irrigation process in Bangladesh where photovoltaic technology could be used to gather solar energy for running a submersible pump and supply water for crop cultivation. It also depicts a comparative picture of irrigation costs for 27 Bangladeshi crops for diesel and photovoltaic irrigation systems. The researchers have collected data concerning required water height during farming of those crops and then have calculated water volume for 1 hectare of land. Subsequently, two commonly used pumps (solar, diesel) with same power ratings (5 hp) have been chosen. Specific area covered by these pumps for different crops are calculated furthermore from the attained water volumes. Finally, total irrigation costs (at present condition) of these types of irrigation choices for a period of ten years have been computed and analyzed. The study highlights that irrigation with solar energy for certain crops, namely potato, cotton, soybean, sunflower, strawberry, lentil, mustard etc. are very much lucrative compared to diesel powered irrigation.**

***Keywords*— Solar irrigation, renewable energy, green farming, photovoltaic pumping, solar for agriculture.***

## I. INTRODUCTION

Due to the fossil fuel resources decline and their great share in environmental pollution and issues, many countries and researchers are looking for green energy resources based on each region's potentials. So far many kind of renewable energy sources such as solar, wind, geothermal and others are utilized for power generation. In general, to meet electricity demand and to cope with environmental problems using green energies there are two steps: first, finding renewable energy resources in a special region, second, to utilize these energy resources economically and efficiently.

Being a tropical country, Bangladesh endowers with abandon supply of solar energy. The range of solar radiation is between 4 and 6.5 kWh/m2/day and the bright sunshine hours vary from 6 to 9 hours/day [1]. Being an agrarian economy, the agricultural sector alone accounts for 20% of GDP and provides employment for more than half of the labour force. Furthermore in Bangladesh, about 59% land is under irrigation system, based on diesel and grid electricity. However there remains vast area of cultivable land which is needed to be irrigated where grid connection is not available. Solar PV pump may be used for irrigating these lands for better crop production. This study presents the scenario of solar pump irrigation system in Bangladesh along with its economic feasibility for different crops [7].

The government of Bangladesh is planning to install close to 19,000 solar-powered irrigation pumps by 2016 [17], in a bid to expand the country's irrigated land area and boost food production, while limiting its reliance on fossil fuels. Once installed, the planned 18,750 solar-powered pumps will irrigate an additional 590,000 hectares (1.5 million acres) of land for cultivating rice and vegetables, without requiring any grid electricity or diesel fuel [3]. The government estimates that, once all the pumps are in place, their solar panels will save 675 MW hours of electricity per day, cut imports of diesel fuel by 47,000 tons per year, saving $45 million annually, and reduce carbon dioxide emissions by an annual 126,000 tons [19].

In this study the main objective was to do an economical evaluation of different cultivated crops in Bangladesh in a ten years time period using both diesel and photovoltaic water pumping systems and hence to find out which crops are viable to the newest and environment friendly method of irrigation process.

## II. RELATED STUDY

### *A. Photovoltaic(PV) Technology*

The Photovoltaic (PV) system is composed of a variety of components in addition to the photovoltaic modules, a balance-of-system that wired together to form the entire fully functional system capable of supplying electric power; and these system elements are:

- PV cells, represent the fundamental power conversion units. They are made from semiconductors and convert sunlight to electricity. To increase the power output of PV cells, they are connected together to form larger units called modules. Modules, in turn, are connected in parallel and series to form a larger unit called panel.
- A storage medium (battery bank), stores the electrical energy produced by the PV cells, and makes the energy available at night or on dark days (days of autonomy or no-sun-days).

☐ A voltage regulator (or charge/discharge controller), reverses current and prevents battery from getting overcharged and over discharged.

☐ An inverter, converts a low DC-voltage into usable AC-voltage; it may be a stand-alone installation or grid-connected installation.

☐ AC or DC loads, appliances and devices, which consume the power generated by the PV system.

Figure 1 shows the configuration of the stand-alone PV system with all the functional components.

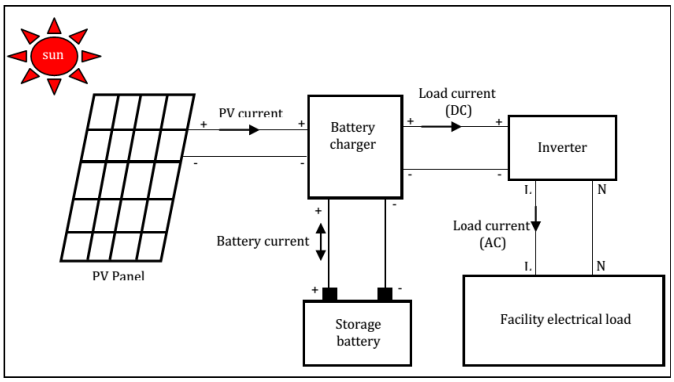



Fig. 1 Configuration of a stand-alone PV system

## B. Solar Pumping System

At remote places, far away from electricity and clean water, electrical pumps with solar panels (solar pumps) can provide in a clever solution. The solar pump system consists of solar panels on a mounting structure, a pump controller, an electric pump and an optional storage tank for water [13]. The big advantage of the solar pump is that it can be operated with or without the battery back-up for solar power. The pump is connected to solar panels, so water is pumped from low to high level in case the sun shines [2].

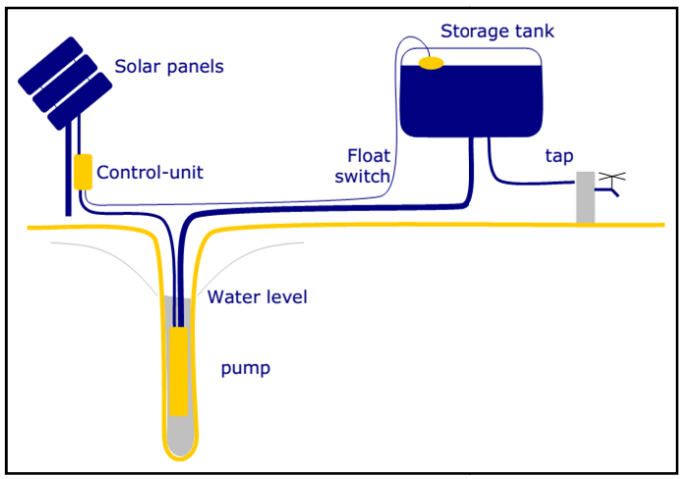



Fig 2: A general setup for solar pump system

To obtain a good match between solar panels and the pump, a pump controller is connected in between. The controller converts the direct current from the solar panels into alternating current with a frequency that depends to the irradiation. At low irradiation, e.g. in the morning at sunrise, the pump will be driven by a slowly rotating engine [6]. The speed of rotation will increase when the sun rises in the course of the day.

## C. Irrigation and Water Requirement of Crops

Irrigation makes agriculture possible in areas previously unsuitable for intensive crop production. Irrigation transports water to crops to increase yield, keep crops cool under excessive heat conditions and prevent freezing. The process of irrigation varies crop to crop and in a crop cycle it needs 3-5 times irrigation. The common processes of irrigation in Bangladesh are as follows [12]:

*(i) Furrow Irrigation:* Furrow irrigation is conducted by creating small parallel channels along the field length in the direction of predominant slope. Water is applied to the top end of each furrow and flows down the field under the influence of gravity.

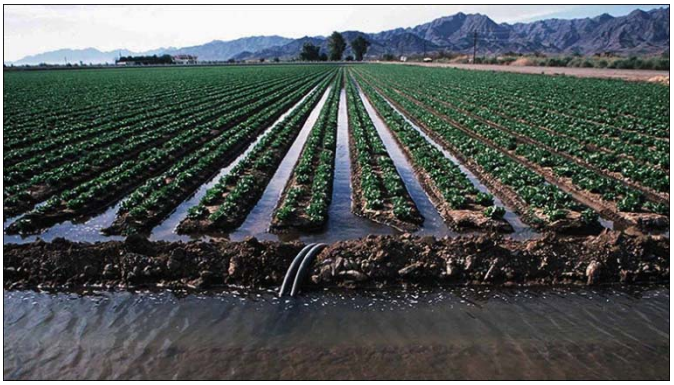


Fig. 3 Furrow irrigation system

*(ii) Basin Irrigation:* Level basin irrigation has historically been used in small areas having level surfaces that are surrounded by earth banks. The water is applied rapidly to the entire basin and is allowed to infiltrate.

Fig. 4 Basin irrigation system

The Water Requirement (WR) of crops depends upon retention and transmissivity of water in soil, absorption and transmission within plant, transpiration, effective rainfall, vapour pressure, and energy etc. WR is that quantity of water regardless of its sources required by a crop in a given period of time for its maturity. For example in case of paddy [11], it

is enough to impound 5 cm of water and recharge to the same level once in four days. Again, for maize irrigation frequency should be once in four days and six days in case of clay soils. Cotton and groundnut requires irrigation once in 10 days in red and 15 days in clay soils.

## III. Methodology

The information of this research was collected from many sources i.e. books, journal/conference papers, websites etc. In this study the system has not been considered to recharge the rechargeable battery whereas it has been considered that the pump is run directly from solar power.

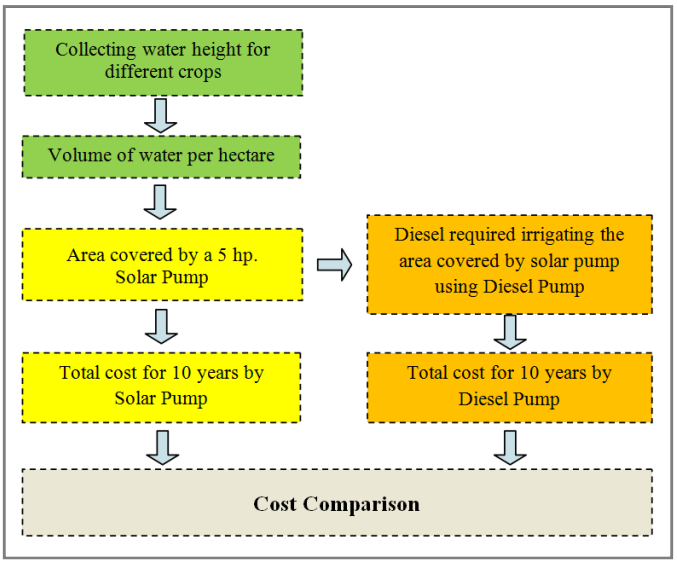



Fig. 5 Illustration of study process

To compare the total cost for irrigation in the period of ten years using PV technology and diesel pump, area covered by a 5hp pump has been calculated first. Then the required diesel has been evaluated to irrigate the same amount of land. It is noted that manufacturers of solar pump usually give a warranty of ten (10) years; hence the associated cost for solar pump is fixed for this span of time. However, diesel price increases every year, so in order to find out the total present value, 10% inflation rate has been considered.

## IV. Analysis of the Study

To draw comparative picture of irrigation costs of different Bangladeshi crops for diesel, grid electricity and solar power based irrigation systems, there are several steps of calculation needed to be considered.

*(A) Calculation of Water Volume* [12]

For *Basin* irrigation process,

Water Volume = Required water height × Area (1)

For *Furrow* irrigation process,

Water Volume = [{(Required water height × bed to bed distance × 100 × No. of canals) ÷ 100} ÷100] (2)

Where, No. of Canals = [{100 - bed to bed distance ÷ 100} ÷ {(width of bed + bed to bed distance) ÷ 100} + 1]

TABLE 1
CALCULATION OF REQUIRED WATER FOR DIFFERENT CROPS

| Name of Crop | Method of Irrigation | Width of Bed (cm) | Bed to Bed Distance (cm) | No. of Canals | Required Water Height (cm) | Water Volume (m3/hectare) |
|---|---|---|---|---|---|---|
| Rice | Basin | n/a | n/a | n/a | 8.5 | 850 |
| Wheat | Basin | n/a | n/a | n/a | 7 | 700 |
| Potato | Furrow | 25 | 60 | 117.9 | 4 | 283.1 |
| Maize | Furrow | 110 | 30 | 72.2 | 7 | 151.7 |
| Onion | Furrow | 100 | 30 | 77.7 | 2.5 | 58.3 |
| Tomato | Furrow | 100 | 30 | 77.7 | 3.5 | 81.6 |
| Sugarcane | Furrow | 120 | 30 | 67.5 | 12.5 | 235 |
| Cotton | Basin | n/a | n/a | n/a | 7 | 700 |
| Chill | Furrow | 100 | 30 | 77.7 | 5 | 116.5 |
| Carrot | Furrow | 100 | 30 | 77.7 | 5 | 116.5 |
| Soybean | Basin | n/a | n/a | n/a | 6 | 600 |
| Garlic | Furrow | 90 | 30 | 84.1 | 6.5 | 164 |
| Brinjal | Furrow | 100 | 30 | 77.7 | 3 | 69.9 |
| Gourd | Furrow | 560 | 40 | 17.6 | 4 | 28.2 |
| Sunflower | Basin | n/a | n/a | n/a | 8 | 800 |
| Ginger | Furrow | 100 | 30 | 77.7 | 3.5 | 81.6 |
| Strawberry | Basin | n/a | n/a | n/a | 6 | 600 |
| Turmeric | Furrow | 90 | 30 | 84.1 | 4.5 | 113.5 |
| Lentil | Basin | n/a | n/a | n/a | 5 | 500 |
| Pumpkin | Furrow | 330 | 30 | 28.7 | 6 | 51.7 |
| Cabbage | Furrow | 90 | 30 | 84.1 | 5 | 126.1 |
| Cauliflower | Furrow | 100 | 30 | 77.7 | 4 | 93.2 |
| Mustard | Basin | n/a | n/a | n/a | 3 | 300 |
| Banana | Furrow | 105 | 30 | 74.9 | 4 | 89.8 |
| Ladyfinger | Furrow | 100 | 30 | 77.7 | 3 | 69.9 |
| Papaya | Furrow | 200 | 30 | 44.3 | 5 | 66.5 |
| Groundnut | Furrow | 40 | 20 | 167.3 | 4 | 133.9 |

*(B) Calculation of Covered Area (Hectare):*

For the simplicity of the research calculation, the area of the field has been considered as 1 hectare which has the dimension of 100 m × 100 m. Again, from the personal communication of different agricultural scientists it has been ensured that water required for single irrigation could be supplied in about 5 days [9], [12].

So, Area Covered = Discharge rate ÷ Water volume (3)

For 5 days cycle, Total Cover Area = Area Covered × 5 (4)

Different crops need different amount of water at cultivation process with different irrigation process [15]. Calculation of required water per hectare of different crops as well as area covered by a 5hp PV pump have been tabulated at Table 1 and 2 consecutively.

*(C) Calculation for Solar Powered Irrigation*

In order to find out the total cost of a solar powered irrigation system in a span of 10 years, the future maintenance

TABLE 2

CALCULATION OF AREA COVERED BY A 5 HP SOLAR PUMP FOR DIFFERENT CROPS

| Name of Crop | Water Volume (m3/hectare) | Area Covered (hectare) | 5 Days Cycle (hectare) |
|---|---|---|---|
| Rice (*Oryza sativa*) | 650 | 0.4 | 1.9 |
| Wheat (*Triticum aestivum*) | 700 | 0.4 | 1.8 |
| Potato (*Solanum tuberosum*) | 283.1 | 0.9 | 4.4 |
| Maize (*Zea mays*) | 151.7 | 1.6 | 8.2 |
| Onion (*Allium cepa*) | 58.3 | 4.3 | 21.5 |
| Tomato (*Solanum lycopersicum*) | 81.6 | 3.1 | 15.3 |
| Sugarcane (*Saccharum officinarum L.*) | 235 | 1.1 | 5.3 |
| Cotton (*Gossypium spp.*) | 700 | 0.4 | 1.8 |
| Chill (*Capsicum annum L.*) | 116.5 | 2.1 | 10.7 |
| Carrot (*Daucus carota*) | 116.5 | 2.1 | 10.7 |
| Soybean (*Glycine max*) | 600 | 0.4 | 2.1 |
| Garlic (*Allium Sativum*) | 164 | 1.5 | 7.6 |
| Brinjal (*Solanum melongena*) | 69.9 | 3.6 | 17.9 |
| Gourd (*Lagenaria siceraria*) | 28.2 | 8.9 | 44.4 |
| Sunflower (*Helianthus annus*) | 800 | 0.3 | 1.6 |
| Ginger (*Zingiber officinale*) | 81.6 | 3.1 | 15.3 |
| Strawberry (*Fragaria ananassa*) | 600 | 0.4 | 2.1 |
| Turmeric (*Curcuma longa*) | 113.5 | 2.2 | 11 |
| Lentil (*Lens culinaris*) | 700 | 0.4 | 1.8 |
| Pumpkin (*Cururbita maxima*) | 51.7 | 4.8 | 24.2 |
| Cabbage (*Brassica oleracea*) | 126.1 | 2.0 | 9.9 |
| Cauliflower (*Brassica oleracea*) | 93.2 | 2.7 | 13.4 |
| Mustard (*Brassica juncea*) | 300 | 0.8 | 4.2 |
| Banana (*Musa paradisiac*) | 89.8 | 2.8 | 13.9 |
| Ladyfinger (*Abelmoschus esculentus*) | 69.9 | 3.6 | 17.9 |
| Papaya (*Carica papaya*) | 66.5 | 3.8 | 18.8 |
| Groundnut (*Arachis hypogaea*) | 133.9 | 1.9 | 9.3 |

costs of next 9 years should be added at the present time. It is a general concept of engineering that every year 5% maintenance cost should be considered for any system. So,

Cost of first year = Pump cost + 5% Maintenance Cost (5)

Now, to consider the future maintenance cost at present, *Present Value (PV)* is a formula that calculates the present day value of an amount that will be spent at a future date.

Present value = Future value / (1+i)^n (6)
Where, i = inflation rate, n = number of year.

TABLE 3
CALCULATION OF TOTAL COST A 5 HP SOLAR PUMP

| | | |
|---|---|---|
| Price of the solar module (panel + pump) | | 2,00,000.00 |
| Maintenance Cost at Present (taka) | Year 1 | 2,10,000.00 |
| | Year 2 | 9,090.90 |
| | Year 3 | 8,264.46 |
| | Year 4 | 7,513.14 |
| | Year 5 | 6,830.13 |
| | Year 6 | 6,209.21 |
| | Year 7 | 5,644.74 |
| | Year 8 | 5,131.58 |
| | Year 9 | 4,665.07 |
| | Year 10 | 4,240.97 |
| **Total Present Value (taka)** | | **2,67,590.24** |

Using solar energy, we calculated the cost in taka for total ten years. It will fix for all crops, because we implement one system. Our module is fixed for all cops.

*(D) Calculation for Diesel Powered Irrigation*

For Diesel powered irrigation,

Cost (first year) = Pump cost + (Diesel Requirement*number of irrigations in a crop cycle*Present diesel rate) *3 (7)

But for the second to next ten years (lifetime of diesel module) it depends in diesel price of that year, 5% maintenance cost, and number of irrigation in a crop cycle.

Cost for 2nd year and onwards = {(Diesel Requirement × number of irrigations in a crop cycle × Present Diesel rate × 3) + 5% maintenance cost} (8)

Now, if anyone wants to calculate the total cost for ten years period, present value of money for the future years should be calculated (Table 3).

## V. Findings and Conclusion

The analysis of this study discovered that irrigation with solar power of certain crops like potato, cotton, soybean, sunflower, strawberry, lentil, mustard are very much lucrative compared to diesel powered irrigation (fig. 6).

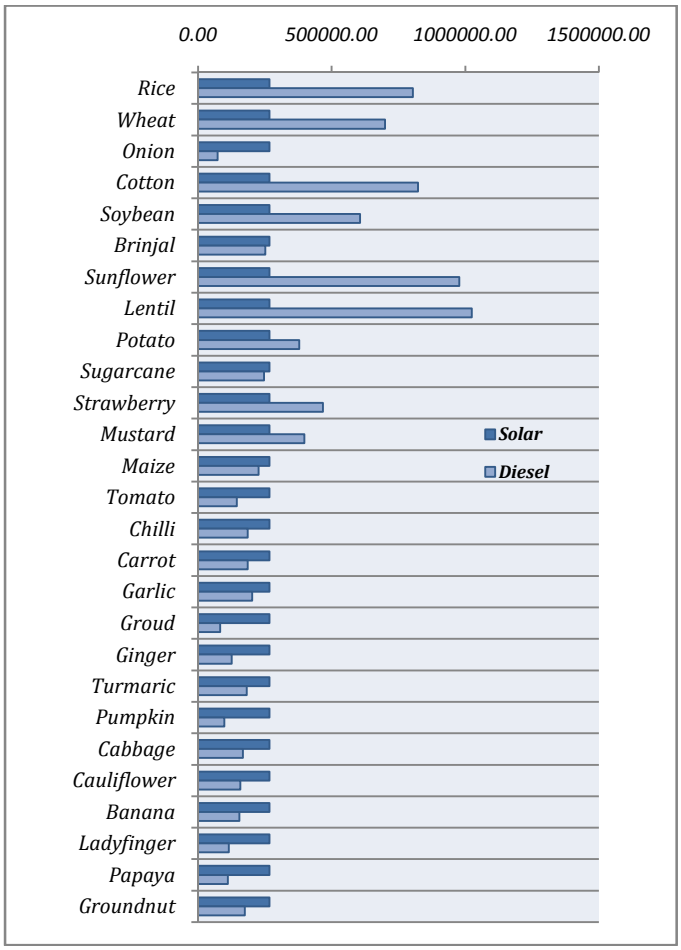



Fig. 6 Comparison between diesel and solar irrigation for different crops

TABLE 4
COST CALCULATION OF DIESEL POWERED IRRIGATION

| Crop | Water Volume ($m^3$) | Water Volume (L) | Diesel Requirement (L) | No. of Irrigation in a Crop Cycle | Present Value Calculation for 10 years (thousand taka) | | | | | | | | | | Total Present Value (in Thousand Taka) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | Year 6 | Year 7 | Year 8 | Year 9 | Year 10 | |
| Rice | 850.0 | 850000 | 96.59 | 5 | 137.5 | 104.8 | 109.3 | 114.0 | 119.0 | 124.3 | 129.7 | 135.5 | 141.5 | 147.8 | 1263.5 |
| Wheat | 700.0 | 700000 | 79.55 | 4 | 103.9 | 69.6 | 72.6 | 75.6 | 78.9 | 82.3 | 85.9 | 89.6 | 93.5 | 97.7 | 849.6 |
| Potato | 283.1 | 283060 | 32.17 | 5 | 71.8 | 36.1 | 37.5 | 39.0 | 40.5 | 42.2 | 43.9 | 45.8 | 47.7 | 49.8 | 454.3 |
| Maize | 151.7 | 151650 | 17.23 | 5 | 56.6 | 20.1 | 20.8 | 21.5 | 22.3 | 23.2 | 24.0 | 25.0 | 26.0 | 27.0 | 266.6 |
| Onion | 58.3 | 58270 | 1.28 | 6 | 40.6 | 3.4 | 3.3 | 3.3 | 3.2 | 3.2 | 3.1 | 3.1 | 3.1 | 3.2 | 69.5 |
| Tomato | 81.6 | 81580 | 9.27 | 5 | 48.5 | 11.7 | 11.9 | 12.3 | 12.6 | 13.0 | 13.4 | 13.9 | 14.4 | 14.9 | 166.7 |
| Sugarcane | 235.0 | 235000 | 26.70 | 4 | 39.0 | 24.6 | 25.4 | 26.4 | 27.4 | 28.4 | 29.5 | 30.7 | 32.0 | 33.3 | 296.7 |
| Cotton | 700.0 | 700000 | 79.55 | 5 | 81.1 | 86.6 | 90.3 | 94.2 | 98.3 | 102.5 | 107.0 | 111.8 | 116.7 | 121.9 | 1010.4 |
| Chill | 116.5 | 116540 | 13.24 | 5 | 52.5 | 15.9 | 16.4 | 16.9 | 17.5 | 18.1 | 18.7 | 19.4 | 20.2 | 21.0 | 216.5 |
| Carrot | 116.5 | 116540 | 13.24 | 5 | 52.5 | 15.9 | 16.4 | 16.9 | 17.5 | 18.1 | 18.7 | 19.4 | 20.2 | 21.0 | 216.5 |
| Soybean | 600.0 | 600000 | 68.18 | 4 | 94.6 | 59.9 | 62.4 | 65.0 | 67.8 | 70.7 | 73.7 | 76.9 | 80.3 | 83.8 | 735.3 |
| Garlic | 164.0 | 163960 | 18.63 | 4 | 54.2 | 17.7 | 18.2 | 18.8 | 19.5 | 20.2 | 20.9 | 21.8 | 22.6 | 23.5 | 237.4 |
| Brinjal | 69.9 | 69923 | 7.95 | 4 | 45.5 | 8.6 | 8.7 | 8.9 | 9.1 | 9.3 | 9.6 | 9.9 | 0.1 | 10.5 | 120.1 |
| Gourd | 28.2 | 28160 | 3.20 | 5 | 42.3 | 5.2 | 5.2 | 5.2 | 5.2 | 5.3 | 5.4 | 5.5 | 5.6 | 5.7 | 90.4 |
| Sunflower | 800.0 | 800000 | 90.91 | 5 | 131.7 | 98.7 | 103.0 | 107.4 | 112.1 | 117.0 | 122.2 | 127.6 | 133.2 | 139.2 | 1192.1 |
| Ginger | 81.6 | 81570 | 9.27 | 4 | 46.6 | 9.7 | 9.9 | 10.1 | 10.4 | 10.7 | 11.0 | 11.3 | 11.7 | 12.1 | 143.4 |
| Strawberry | 600.0 | 600000 | 68.18 | 3 | 80.7 | 45.4 | 47.2 | 49.1 | 51.2 | 53.3 | 55.6 | 58.0 | 60.5 | 63.1 | 564.1 |
| Turmeric | 113.5 | 113510 | 12.90 | 5 | 52.2 | 15.5 | 16.0 | 16.5 | 17.1 | 17.6 | 18.3 | 19.0 | 19.7 | 20.5 | 212.3 |
| Lentil | 500.0 | 500000 | 56.82 | 6 | 108.5 | 74.5 | 77.6 | 80.9 | 84.4 | 88.1 | 91.9 | 95.9 | 100.2 | 104.6 | 906.7 |
| Pumpkin | 51.7 | 51650 | 5.87 | 4 | 43.8 | 6.8 | 6.8 | 6.9 | 7.1 | 7.2 | 7.4 | 7.5 | 7.7 | 8.0 | 109.2 |
| Cabbage | 126.1 | 126125 | 14.33 | 4 | 50.7 | 14.0 | 14.4 | 14.8 | 15.3 | 15.8 | 16.4 | 17.0 | 17.6 | 18.3 | 194.2 |
| Cauliflower | 93.2 | 93230 | 10.59 | 5 | 49.8 | 13.1 | 13.4 | 13.8 | 14.2 | 14.7 | 15.2 | 15.7 | 16.3 | 16.9 | 183.2 |
| Mustard | 300.0 | 300000 | 34.09 | 5 | 73.8 | 38.1 | 39.6 | 41.2 | 42.9 | 44.6 | 46.5 | 48.5 | 50.5 | 52.7 | 478.4 |
| Banana | 89.8 | 89822 | 10.21 | 5 | 49.4 | 12.7 | 13.0 | 13.4 | 13.8 | 14.2 | 14.7 | 15.2 | 15.8 | 16.4 | 178.5 |
| Lady Finger | 69.9 | 69930 | 7.95 | 4 | 45.5 | 8.6 | 8.7 | 8.9 | 9.1 | 9.3 | 9.6 | 9.9 | 10.2 | 10.5 | 130.1 |
| Papaya | 66.5 | 66521 | 7.56 | 4 | 45.2 | 8.2 | 8.4 | 8.5 | 8.7 | 8.9 | 9.2 | 9.4 | 9.7 | 10.0 | 126.2 |
| Ground Nut | 133.9 | 133867 | 15.21 | 4 | 51.4 | 14.7 | 15.2 | 15.6 | 16.2 | 16.7 | 17.3 | 17.9 | 18.6 | 19.3 | 203.1 |

The findings of this study have led us to believe that the time is now ripe to advance towards a new phase implementing solar powered irrigation system [4]. At this time, financing and technical support for individual pilot initiatives is recommended. Therefore, the government should more facilitate the development of this renewable water pumping sector so that private entrepreneurs come forward to take the initiatives to invest in this sector.

## REFERENCES


[1] N. C. Bhowmik, *"Bangladesh Renewable Energy Report"*, APCTT-UNESCAP, 2012.

[2] A. Keyhani, M. N. Marwali, M. Dai, *"Integration of Green and Renewable Energy in Electric Power systems"*, A John Wiley and Sons, USA, 2010.

[3] E. F. Fuchs, M. A.S. Masoum, *"Power Conversion of Renewable Energy Systems"*, Springer – Verlag London Limited, UK, 2011.

[4] Dr. M. M. Qurashi, Engr. T. Hussain, *"Renwable Energy Technologies for Developing Countries Now and to 2023"*, Islamic Educational, Scientific and Cultural Organization, 2005.

[5] R. Foster, M. Ghassemi, A. Cota, *"Solar Energy Renewable Energy and the Environment"*, CRC Press Tylor & Francis Group, 2010.

[6] R. Pode, B. Diouf, *"Solar Lighting"*, Springer – Verlag London Limited, UK, 2011.

[7] V. Nelson, *"Solar Energy Renewable Energy and the Environment"*, CRC Press Tylor & Francis Group, 2009.

[8] M. R. Patel, *"Wind and Solar Power Systems"*, CRC Press, USA, 2000.

[9] R. Subarkah, Belyamin, *"Improved Efficiency Solar Cells with Forced Cooling Water"*, retrieved on 16/10/2011.

[10] R. Margolis and J. Zuboy, *"Nontechnical Barriers to Solar Energy Use: Review of Recent Literature"*, Colorado, 2006.

[11] M. Islam, *"Utilization of Renewable Energies in Bangladesh"*, 2002.

[12] M. Abdus Sattar, Dr. M.Motiur Rahman, *"Agriculture Technology Handbook"* 3rd ed., Joydebpur, Gazipur, 2004.

[13] (2013) The Solorb website. [Online]. Available: http:// www.solorb.com/elect/solarirc/scc3

[14] (2013) The Wikipedia website. [Online]. Available: http:// en.wikipedia.org/solarenergy

[15] (2013) The Rahimarooz Renewable Energy Ltd. website. [Online]. Available: http:/ /www.rre.com/solarpump

[16] (2013) Rahimafrooz homepage [Online]. Available:http://www.rahimafrooz.com/solarproject

[17] Bangladesh Prime-minister Office Website (2013) homepage [Online]. Available: http:// www.pmo.gov.bd

[18] Alibaba Office Website (2013) homepage [Online]. Available: http://www.alibaba.com/diesel_pump.php

[19] (2013) Bangladesh Government website. [Online]. Available: http:/ / www.bangabhaban.gov.bd

[20] Md. M. Biswas, Kamol K. Das, Ifat A. Baqee, Mohammad A. H. Sadi, Hossain M. S. Farhad, *"Prospects of Renewable Energy and Energy Storage Systems in Bangladesh and Developing Economics"*, Global Journal of researches in engineering , Volume 11, Issue 5 Version 1.0 July 2011.

[21] Khairul Anam, Husnain-Al-Bustam, *"Power Crisis & Its Solution through Renewable Energy in Bangladesh"*, Journal of Selected Areas in Renewable and Sustainable Energy *(JRSE)*, and September-2011.

[22] M. M. Hasan, M.F. Khan, "*A comparative study on installation of solar PV system for grid and non grid rural areas of Bangladesh*", Developments in Renewable Energy Technology (*ICDRET*), 2012.

# Power Converters and Control of Wind Energy Conversion Systems

Syed Naime Mohammad*, Nipu Kumar Das, Saikat Roy, and Arif Ahammad
Department of Electrical and Electronic Engineering, Chittagong University of Engineering and Technology, Bangladesh
*E-mail: syednaeem.08cuet@gmail.com

*Abstract*— **The rapid increase in global energy consumption and the impact of greenhouse gas emissions has accelerated the transition towards greener energy sources. The need for distributed generation employing renewable energy sources such as wind, solar and fuel cells has gained significant momentum. Advanced power electronic systems, affordable high performance devices, and smart energy management principles are deemed to be an integral part of renewable, green and efficient energy systems. This paper discuss the most emerging renewable energy source, wind energy, which by means of power electronics is changing from being a minor energy source to be acting as an important power source in the energy system and the standard power converter topologies from the simplest converters for starting up the turbine to advanced power converter topologies, where the whole power is flowing through the converter.**

***Keywords— Wind energy conversion, wind turbines, power electronics, control, reliability.***

## I. INTRODUCTION

The global energy consumption has been continually increasing over the last century. Official estimates indicate a 44 percent increase in global energy consumption during the period 2006 - 2030 [1]. It can be said that fossil fuels (liquid, coal and natural gas) have been the primary energy source for the present day world. Sustained urbanization, industrialization, and increased penetration of electricity have led to unprecedented dependency on fossil fuels. Presently, the most important concerns regarding fossil fuels are the green house gas emissions and the irreversible depletion of natural resources. Based on the official energy statistics from the US Government, the global carbon dioxide emissions will increase by 39 percent to reach 40.4 billion metric tons from 2006 to 2030 [1]. Green house gas emissions and the related threat of global warming and depleting fossil fuel reserves have placed a lot of importance on the role of alternative and greener energy sources. The wind turbine technology is one of the most emerging renewable technologies. Over the last ten years, the global wind energy capacity has increased rapidly and became the fastest developing renewable energy technology [2]. The controllability of the wind turbines becomes more and more important as the power level of the turbines increases. Power electronic, being the technology of efficiently converting electric power, plays an important role in wind power systems. It is an essential part for integrating the variable-speed wind power generation units to achieve high efficiency and high performance in power systems. Even in a fixed-speed wind turbine system where wind power generators are directly connected to the grid, thyristors are used as soft-starters [3]. The power electronic converters are used to match the characteristics of wind turbines with the requirements of grid connections, including frequency, voltage, control of active and reactive power, harmonics, etc [4]. This paper presents the requirements of the power electronic interface as applicable with respect to wind and qualitatively examines the existing power electronic topologies that can be employed. Section II and section III presents an overview of wind energy system and modern power electronics. Section IV presents power electronic converters used in wind turbine applications. Section V presents reliability of wind turbines and future of wind energy and finally, in section VI, the conclusion are drawn.

## II. WIND ENERGY SYSTEM

A wind energy conversion system is a structure that transforms the kinetic energy of the incoming air stream into electrical energy. This conversion takes place in two steps. At first the extraction device, named wind turbine rotor turns under the wind stream action, thus harvesting a mechanical power. Then the rotor drives a rotating electrical machine, the generator, which outputs electrical power. A block diagram of wind energy conversion system is shown in Fig. 1. The performance and efficiency of any wind energy conversion system (WECS) depends upon the characteristics of wind turbines. The mechanical power of the turbine is given by:

$$Pm = \frac{1}{2}\rho A u^3 Cp \qquad (1)$$

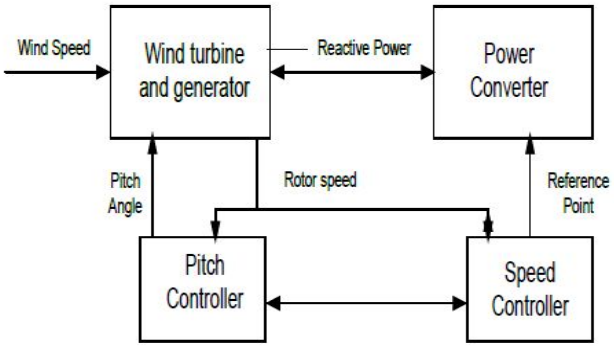



Figure 1. Block diagram of WECS

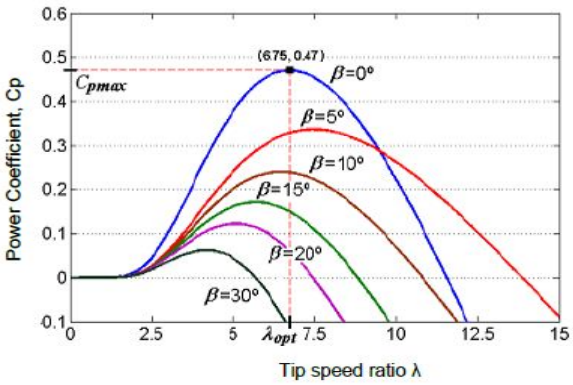



Figure 2. Cp- λ characteristics for different values of β

Where Pm is the power extracted from the airflow, ρ is the air density, A is the area covered by the rotor, u is the wind speed upstream of the rotor, and Cp is the performance coefficient or power coefficient. The power coefficient is a function of the pitch angle of rotor blades β and of the tip speed ratio λ, which is the ratio between blade tip speed and wind speed upstream of the rotor. The computation of the power coefficient requires the use of blade element theory and the knowledge of blade geometry. We consider the blade geometry using the numerical approximation developed in [5], assuming that the power coefficient is given by:

$$Cp = 0.73\lambda_i e^{\frac{-18.4}{\lambda_{ii}}} \tag{2}$$

Where $\lambda_i$ and $\lambda_{ii}$ are respectively given by:

$$\lambda_i = \frac{151}{\lambda_{ii}} - 0.58\beta - 0.002\beta^{2.14} - 13.2 \tag{3}$$

$$\lambda_{ii} = \frac{1}{\frac{1}{(\lambda_i - 0.02\beta.} - \frac{0.003}{(\beta^3 + 1)}} \tag{4}$$

The maximum power coefficient is given for a null pitch angle and is equal to:

$$Cpmax = 0.4412 \tag{5}$$

Where the optimum tip speed is equal to:

$$\lambda_{opt} = 7.057 \tag{6}$$

The power coefficient is illustrated in Fig. 2. as a function of the tip speed ratio. So, for a wind rotor with radius r, (1) can be rewritten as [6],[7]:

$$Pm = \frac{1}{2}\rho\pi r^2 u^3 Cp(\lambda,\beta) \tag{7}$$

The relation between wind speed and generated power is given by the power curve, as depicted in Fig. 3. When the wind speed is below the cut-in wind speed then there will be no generation. When the speed is above cut-in speed but below the rated speed the power optimization method or MPPT method is used to extract maximum power. In this region the speed is maintained at a constant value corresponding to optimum tip speed ratio λ. Above rated speed but below cut-out wind speed pitch controller is used to control the wind power at the wind turbine blade. At very high wind speed it means above cut out speed there will be again no generation of power.

## III. MODERN POWER ELECTRONICS AND SYSTEMS

Power electronics is the technology that links the two major traditional divisions of electrical engineering, namely, electric power and electronics. It has shown rapid development in recent times, primarily because of the development of semiconductor power devices that can efficiently switch currents at high voltages, and so can be used for the conversion and control of electrical energy at high power levels. The development of microprocessors/microcomputer technology has a great impact on the control and synthesizing the control strategy for the power semiconductor devices. Power electronic techniques are progressively replacing traditional methods of power conversion and control, causing what may be described as a technological revolution, in power areas such as regulated power supply systems, adjustable speed DC and AC electric motor drives, high voltage DC links between AC power networks, etc. The power electronic device technology is still undergoing important progress shown in Fig. 4., including some key self-commutated devices, such as

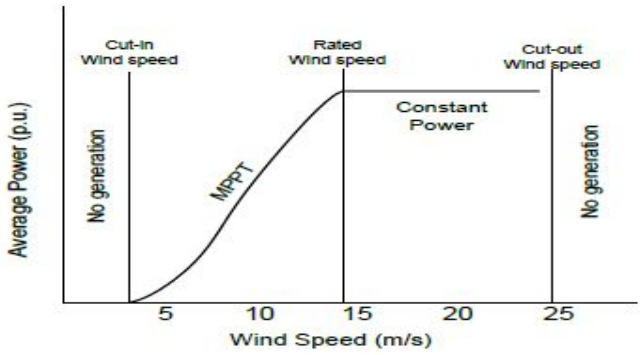



Figure 3. Typical power curve of a variable speed pitch controlled wind turbine [8].

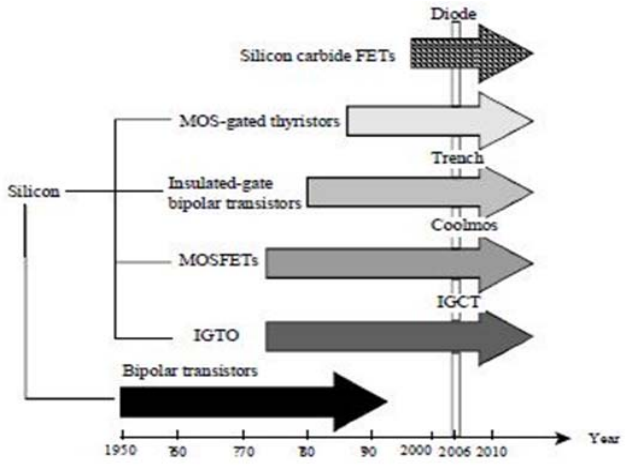



Figure 4. Development of power semiconductor devices in the past and in the future [9]

insulated gate bipolar transistor (IGBT), MOSFET, integrated gate commutated thyristor (IGCT), MOS-gate thyristors, and silicon carbide FETs. The breakdown voltage and/or current carrying capability of the components are also continuously increasing. Important research is going on to change the material from silicon to silicon carbide. This may dramatically increase the power density of the power converters. Power electronic converters are constructed by semiconductor devices, driving, protection, and control circuits to perform voltage magnitude and frequency conversion and control.

## IV. POWER ELECTRONIC CONVERTERS

Many different power converters can be used in wind turbine applications. In the case of using an induction generator, the power converter has to convert from a fixed voltage and frequency to a variable voltage and frequency. Other generator types can demand other complex protection. However, the most used topology so far is a soft-starter, which is used during start up in order to limit the in-rush current and thereby reduce the disturbances to the grid.

### *A. Soft Starter*

The "Danish concept" [10] of directly connecting a wind turbine to the grid is widely used in early wind turbine systems. The scheme consists of an SCIG, connected via a transformer to the grid and operating at an almost fixed speed. Connecting the induction generators to power system produces transients that are short duration with very high inrush currents, thus causing disturbances to both the grid and high torque spikes in the drive train of wind turbines with a directly connected induction generator. The soft starter is a power converter, which has been introduced to fixed speed wind turbines to reduce the transient current during connection or disconnection of the generator to the grid. When the generator speed exceeds the synchronous speed, the soft-starter is connected. Using firing angle control of the thyristors in the soft starter the generator is smoothly connected to the grid over a predefined number of grid periods. The connection diagram for the soft-starter with a generator is presented in Fig. 5 and Control characteristic for a fully controlled soft starter is shown in Fig. 6. When the generator is completely connected to the grid a contactor (Kbyp) bypass the soft-starter in order to reduce the losses during normal operation. The soft-starter is very cheap and it is a standard converter in many wind turbines.

### *B. Capacitor Bank*

For the power factor compensation of the reactive power in the generator, AC-capacitor banks are used, as shown in Fig. 7. The generators are normally compensated into whole power range. The switching of capacitors is done as a function of the average value of measured reactive power during a certain period. The capacitor banks are usually mounted in the bottom of the tower or in the nacelle. In order to reduce the current at connection/disconnection of capacitors a coil (L) can be connected in series. The capacitors may be heavy loaded and damaged in the case of over-voltages to the grid and thereby they may increase the maintenance cost.

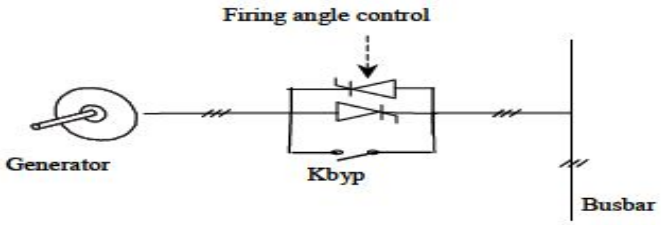



Figure 5. Connection diagram of soft starter with generators

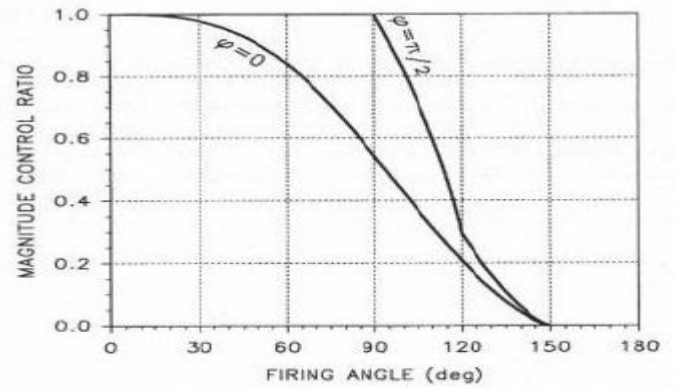



Figure 6. Control characteristic for a fully controlled soft starter

### C. *Diode Rectifier*

The diode rectifier is the most common used topology in power electronic applications. It is used as AC/DC converter. There is a DC link element, which can be a capacitor in Voltage Source Converter (VSC) or an inductor in a Current Source Converter (CSC) [11]-[12]. The applicability of this topology occurs because there is no need of external excitation. For a three-phase system it consists of six diodes. It is shown in Fig. 8. The diode rectifier can only be used in one quadrant, it is simple and it is not possible to control it. It could be used in some applications with a dc-bus.

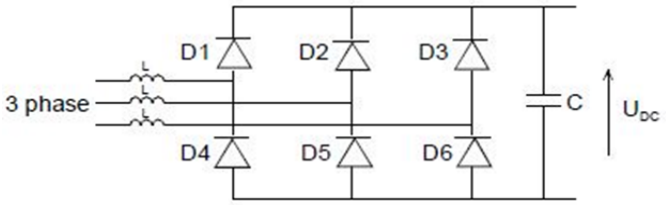



Figure 8. Diode rectifier for three-phase ac/dc conversion

### D. *Bidirectional Back-to-Back Two-Level Power Converter*

This topology is state-of-the-art especially in large DFIG based wind turbines [13]. The back-to-back PWM-VSI is a bi-directional power converter consisting of two conventional PWM-VSCs. The topology is shown in Fig. 9. To achieve full control of the grid current, the DC- link voltage must be boosted to a level higher than the amplitude of the grid line-line voltage. The power flow of the grid side converter is controlled in order to keep the DC-link voltage constant, while the control of the generator side is set to suit the magnetization demand and the reference speed. A technical advantage of the PWM-VSC is the capacitor decoupling between the grid inverter and the generator inverter. Besides affording some

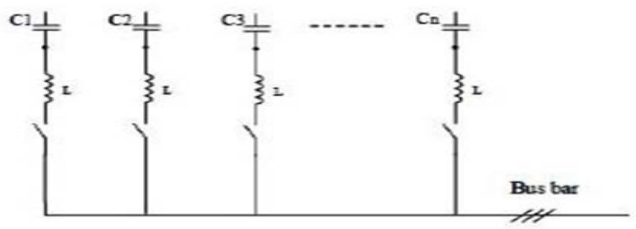



Figure 7. Capacitor bank configuration for power factor compensation in a wind turbine

protection, this decoupling offers separate control of the two inverters, allowing compensation of asymmetry both on the generator side and on the grid side, independently. The inclusion of a boost inductance in the DC-link circuit increases the component count, but a positive effect is that the boost inductance reduces the demands on the performance of the grid side harmonic filter, and offers some protection of the converter against abnormal conditions on the grid. However some disadvantages of the back-to-back PWM-VSI are reported in literature [14], [15] and [16]. In several papers concerning adjustable speed drives, the presence of the DC-link capacitor is mentioned as a drawback, since it is bulky and heavy, it increases the costs and maybe of most importance it reduces the overall lifetime of the system. Another important drawback of the back-to-back PWM-VSI is the switching losses. Every commutation in both the grid inverter and the generator inverter between the upper and lower DC-link branch is associated with a hard switching and a natural commutation. Since the back-to-back PWM-VSI consists of two inverters, the switching losses might be even more pronounced. The high switching speed to the grid may also require extra EMI-filters. To prevent high stresses on the generator insulation and to avoid bearing current problems [17] the voltage gradient may have to be limited by applying an output filter.

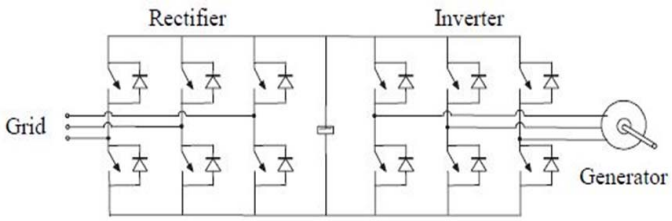



Figure 9. The back-to-back PWM-VSI converter topology.

### E. *Multilevel Converter*

These converters are a good choice in application where high voltage rating is necessary [18], due to voltage level of the converters. The general idea behind the multilevel converter

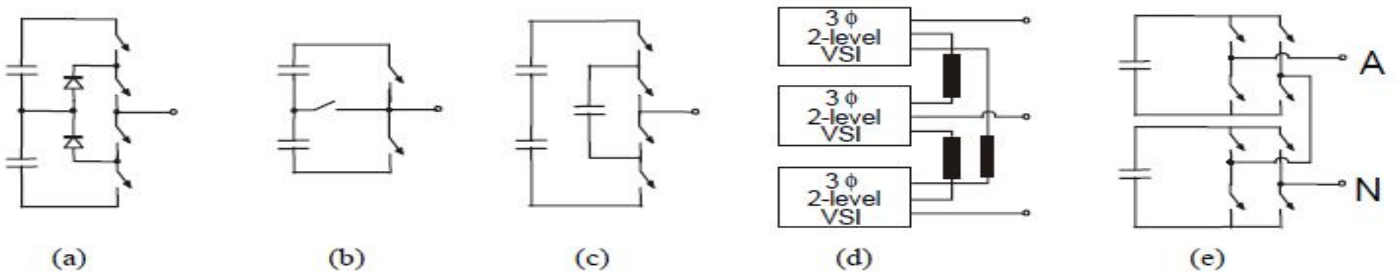



Figure 10. Multilevel topologies. a) One inverter leg of a three-level diode clamped multilevel converter. b) One inverter leg of a three-level multilevel converter with bidirectional switch interconnection. c) One inverter leg of a three level flying capacitor multilevel converter. d) Schematic presentation of a three-level converter consisting of three three-phase inverters. e) One inverter leg of a three-level converter consisting of H-bridge inverters.

technology is to create a sinusoidal voltage from several levels of voltages, typically obtained from capacitor voltage sources. The different proposed multilevel converter topologies can be classified in the following five categories [19]-[21]:

- Multilevel configurations with diode clamps
- Multilevel configurations with bi-directional switch interconnection
- Multilevel configurations with flying capacitors
- Multilevel configurations with multiple three-phase inverters
- Multilevel configurations with cascaded single phase H-bridge inverters

A common feature of the five different multilevel converter concepts is, that in theory, all the topologies may be constructed to have an arbitrary number of voltage levels, although in practice some topologies are easier to realize than others. The principle of the five topologies is illustrated in Fig. 10. Initially, the main purpose of the multilevel converter was to achieve a higher voltage capability of the converters. As the ratings of the components increases and the switching- and conducting properties improve, the secondary effects of applying multilevel converters become more and more advantageous. The switching losses of the multilevel converter are another feature, which is often accentuated. From the topologies in Fig. 10, it is evident that the number of semiconductors in the conducting path is higher than for the other converters treated in this paper, this might increase the conduction losses of the converter.

### *F. Matrix converter*

The basic idea of the matrix converter is that a desired input current (to/from the supply), a desired output voltage and a desired output frequency may be obtained by properly connecting the output terminals of the converter to the input terminals of the converter. In order to protect the converter, the following two control rules must be complied with: Two (or three) switches in an output leg are never allowed to be on at the same time. All of the three output phases must be connected to an input phase at any instant of time. The actual combination of the switches depends on the modulation strategy. Ideally, the matrix converter should be an all silicon solution with no passive components in the power circuit. The ideal conventional matrix converter topology is shown in Fig. 11. It provides a smaller converter promoting more reliability in comparisons to others. But due to its complex control it has not been accepted in industrial applications [22].

## V. RELIABILTY OF WIND TURBINES

The reliability of a wind turbine is critical to extracting the maximum amount of energy from the wind. Different techniques, methodologies and algorithms have been developed to monitor the performance of wind turbine as well as for early fault detection to prevent catastrophic failures. Implementation of condition monitoring system (CMS) and fault detection system (FDS) is essential for achieving high availability of the system. All wind turbines should be equipped with an integrated condition monitoring system for planned preventive maintenance.

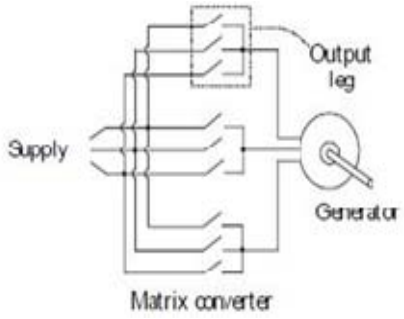



Figure 11. The conventional matrix converter topology.

This will also reduce the probability of catastrophic failures and consequent damage to components. Two of the more common condition-based maintenance (CBM) techniques are vibration-based monitoring and fluid-based monitoring. A typical-vibration monitoring system includes multiple types of sensors with multiple monitoring algorithms, which can be complex and costly. However, fluid-based monitoring provides only limited information related to the condition of wind turbine components [23].

## VI. CONCLUSIONS

Wind energy technology is a renewable, available, and environmentally clean resource that has reached a degree of technological maturity to be an acceptable utility generation technology. The increased wind power penetration in power systems networks leads to new technical challenges, implying research towards more realistic and physical models for wind energy systems. This paper presents a more realistic modeling of power electronics converter used in wind power systems. There is a continuing effort to make converter and control schemes more efficient and cost effective in hopes of an economically viable solution to increasing environmental issues. Wind power generation has grown at a significant rate in the past decade and will continue to do so as power electronic technology continues to advance.

## REFERENCES


[1] International energy outlook (2013) .[online]. Available: www.eia.doe.gov/oiaf/ieo/highlights.html

[2] B. W. Roberts, D. H. Shepard, K. Caldeira, M. E. Cannon, and J. F. Freidin, "Harnessing highaltitude wind power," *IEEE Trans. Energy Convers.*, vol. 22, no. 1, pp. 136–144, Mar. 2007.

[3] P. Singh, S.M. Tripathi and N. Mishra, "Recent trends in power converters for wind energy conversion system with grid integration impact," *International Journal of Research in Engineering & Applied Sciences*, vol. 02, issue: 02, pp. 1660-1661, Feb. 2012.

[4] N.R. Bhasme, and W.Z. Gandhare, "Power converters for grid integration of wind power system," *International Journal of Engineering Research and Applications*, vol. 02, issue: 04, pp. 2181-2185, Aug. 2012.

[5] R. Erickson, S. Angkititrakul, and K. Almazeedi, "A new family of multilevel matrix converters for wind power applications: Final report," pp. 15-25, Dec. 2006.

[6] I. S. M. Kulam, P.Aravindan, and M. Y. Sanavullah, "Mppt based permanent magnet synchronous generator for wind energy conversion system," *International Journal of Communications and Engineering,* vol. 04, issue: 02, pp. 421-434, Mar. 2012.

[7] G. Ofualagba, "Wind energy conversion system- Wind turbine modeling," *in Proc. EPE* , Jul. 2008, pp. 20-24.

[8] V. K. Dwivedi, M. Bajaj, and A. Kumar, "Power electronics converters for a wind energy conversion system: Review," *International Journal of Scientific & Engineering Research*, vol. 04, issue: 06, pp. 2686-2693, Jun. 2013.

[9] B.J. Baliga, "Power IC's in the saddle*", IEEE Trans. Power Electron.,* vol. 54, no. 6, pp. 2930-2945, Dec. 2007.

[10] F. Blaabjerg, Z. Chen, and S. B. Kjaer, "Power electronics as efficient interface in dispersed power generation systems," *IEEE Trans. Power Electron.,* vol. 19, no. 5, pp. 1184–1194, Sep. 2004.

[11] B. Ni, and C. Sourkounis, "Influence of wind-energy-converter control methods on the output frequency components." *IEEE Trans. on Industry Applications*, vol. 45, no. 6, pp. 2116-2121, Nov. 2009.

[12] D.S. Oliveira, M.M. Reis, C.E.A. Silva, and B. L. Soares, "A three-phase high-frequency semicontrolled rectifier for PM WECS," *IEEE Trans. on Power Electronics*, vol. 25, no. 3, pp 677-685, Mar. 2010.

[13] G. Ramtharan, N. Jenkins, O. Anaya-Lara, and E. Bossanyi, "Influence of rotor structural dynamics representations on the electrical transient performance of FSIG and DFIG wind turbines." *in Proc. ICEM*, Dec 2007, pp 293-301.

[14] T. Tafticht, K. Agbossou, A. Cheriti, and M. L. Doumbia, "Output power maximization of a permanent magnet synchronous generator based standalone wind turbine," *in Proc. IEEE ISIE*, Aug. 2006, pp. 2412–2416.

[15] J.A. Sanchez, C. Veganzones, S. Martinez, F. Blazquez, N. Herrero, and J. Wilhelmi, "Dynamic model of wind energy conversion systems with variable speed synchronous generator and full-size power converter for large-scale power system stability studies." *Renew Energy*, vol. 32, pp. 1186-98, Dec. 2008.

[16] S. Jia, X. Wang, K.J. Tseng, "Matrix converters for wind energy systems." *In Proc. IEEE Int. Conf. Ind. Electron. Appl*., Jun. 2007, pp. 488-94.

[17] A. M. Knight, and G. E. Peters, "Simple wind energy controller for an expanded operating range," *IEEE Trans. Energy Convers*., vol. 20, no. 2, pp. 459–466, Jun. 2005.

[18] H. N. Chong, M.A. Parker, L. Ran, P.J. Tavner, and J.R. Bumby, "A multilevel modular converter for a large, light weight wind turbine generator." *IEEE Trans. Power Elect*., vol. 23, no. 3, pp. 1062-1074, May 2008.

[19] M. Malinowsk, "A Survey on Cascaded Multilevel Inverters,"Industrial Electronics*," IEEE Trans. Power Elect*., vol. 3, no. 5 pp. 1-10, Sep. 2009.

[20] Ruderman, and B. Reznikov, "Three-level H-bridge flying capacitor converter voltage balance dynamics analysis," *in proc. European Conf. Power Elect. and App*., jun. 2009, pp. 1-10.

[21] J. Rodriguez., "A Survey on neutral point clamped inverters," *IEEE Trans. Industrial Elect*., vol. 22, no. 2, pp. 1002-1016, Sep. 2009.

[22] G. Yang, and H. Li, "Application of a matrix converter for PMSG wind turbine generation system," *in Proc. Int. Conf. Clean Electrical Power*, Aug. 2009, pp. 619-623.

[23] L. David and M. Emil, "Method and apparatus for condition-based monitoring of wind turbine components," US Patent 7322794, Jan. 2008.

# Experimental Study of Breakdown Voltage for Different Types of Vegetable Oils Available in Bangladesh

Md. Abdul Goffar Khan[1], Md. Khaled Hossain[2] and Md. Fazle Arosh[3] .
[1,2,3] Department of Electrical and Electronic Engineering
Rajshahi University of Engineering and Technology, Rajshahi-6204, Bangladesh.
[1]E-mail:agmagk@gmail.com

*Abstract*---**The increasing crisis of petroleum oil that is currently leading to uncertainty in its sustainability has forced the researchers worldwide to search alternative insulating liquid to mineral oil. Researchers are looking for suitable vegetable oils as alternatives. These oils are biodegradable, environmental friendly and available in different countries including Bangladesh. They may require some processing and modification to improve some of their properties to ascertain their safe use in power and distribution transformers as well as in high voltage equipment. This paper provides a comparative assessment of breakdown property through experimental investigation of several vegetable oils with respect to mineral oil. It is found that the vegetable oils are suitable for low voltage application (less than 11KV) to use as dielectric fluid.**

***Keywords---Breakdown voltage, vegetable oils, dielectric strength, mineral oil, electric field intensity, partial discharge.***

## I. INTRODUCTION

For more than a century, petroleum based mineral oil have been used in liquid filled transformer. Several billions of liters of transformer oil are used in transformer worldwide. The mineral oils were extracted from petroleum which is going run out in the future [1]. The popularity of mineral transformer oil is due to its availability, low cost, excellent as dielectric, cooling medium and easily inflammable. Despite the obvious popularity over time these oils have been found to be lacking in property requirement and most importantly is their negative environmental impact. The disadvantages of mineral oil are non-biodegradable, contaminate soil and water, disturb the plantation and other lives, and harm the environment, low fire point [2]. The natural vegetable oil, on investigation have shown impressive properties and stand as alternative or total replacement to mineral and petroleum oils. It is important to find alternative oil sources that have similar dielectric characteristics with the existing one and probably can increase the performance of related equipments. Vegetable oils have already been applied successfully to small transformers in the United States [3]. The alternative fluids such as synthetic esters, natural esters are already in common use at voltage levels up to 40 KV [4].

John Luksich identified that natural esters (Vegetable oil) is finding growing acceptance and application in electrical equipment using liquid insulation. It is fire safety, interaction with insulation and environmental characteristics make it an excellent choice for many applications [5].

K. Sirikulrat and N. Sirikulrat (2008) investigated the dielectric properties of different maturity of soybean oil was found to increase with the increase in the heating time and was well correlated with acid value, density and viscosity [6].

D. Martin et,al (2006) presented some of the findings to ascertain the suitability of vegetable oil based dielectrics as alternative to mineral oil in large power transformers. Their aim was to investigate whether ester oils are suitable replacements for mineral oil in large transformers above 132KV voltage levels. They compared esters to mineral oil and concluded in the following way:

*Esters are more biodegradable than mineral oil.
* Esters are non-toxic.
* Natural esters are from renewable sources.
* Esters have higher flash points and fire points than mineral oil making esters better suited to transformers [7].

R. Eberhardt et, al (2010) pointed out that different companies are offering alternative insulation liquids which are already used in distribution transformers. On the other hand, there is little experience with those fluids in large power transformers. For the confirmation of the usability of new insulating liquids comparison methods must be tested to find the differences between the alternative insulating fluids and commonly used mineral oil. They studied the partial discharge behavior of alternative insulation liquids such as synthetic and natural esters compared to commonly used mineral oil [8].

In order to settle down the flammability and environmental issues, many researchers started to look for alternative sources for mineral oil. This research has been carried out to search alternatives for mineral oil which are available in Bangladesh. The breakdown voltage and electric field intensity of commercially available vegetable oil such as soybean oil,

palm oil , mustard oil and transformer oil have been carried out in a laboratory based setup and comparison have been made to find out best substitute for mineral oils used in high voltage equipments.

## II. EXPERIMENTAL SETUP

In order to measure breakdown voltage, High Voltage Oil Tester Set, model OTS-E Series (Semi-automatic) was used which is available in RUET HV Laboratory as shown in fig.1. The oil tester was equipped with adjustable 36mm diameter mushroom electrodes. An AC voltage was applied using a 60KV step up transformer for the breakdown test. The maximum output voltage capacity of 60KV was possible for this instrument. Breakdown voltage was determined by applying 50 Hz AC voltage to the electrodes. During each experiment the applied voltage was increased manually from zero at a rate of approximately 0.4KV/Sec. until the breakdown occurred. A transformer control unit (TCU) monitored the cell current and interrupted the supply voltage to the step up transformer when breakdown occurred in the sample test cell as shown in fig.2.


Figure1. Liquid dielectric test set, model OTS-E Series for breakdown voltage measurement.

The dimension of oil test cup was approximately 15 cm in length, 10 cm in width and 12 cm in height. The cell was designed to take small sample volumes with a gap distance of 2.5mm, 5mm, 7.5mm and 10mm. For each of the vegetable oil samples four breakdown measurements were carried out at room temperature of 32$^{o}$ C. The samples were purchased from the local market and were used without processing and filtration. For each sample, PD inception voltages were recorded and breakdown voltages were measured. At the same time video recording was performed to observe PD inception and breakdown phenomena. After finishing each experiment the oil sample was taken out from the test fixture and then the fixture was disassembled, cleaned and dried at room temperature.

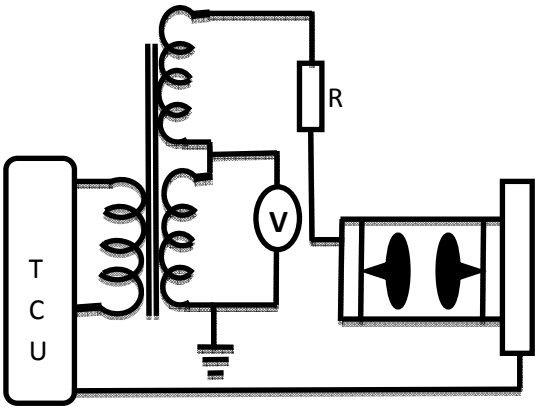



Figure2. Schematic of electric breakdown test setup.

## III. RESULTS

Based on the above experiment the data were plotted to make comparison and observation. Fig.3 shows that for a gap distance (2.5mm-5mm) of 2.5mm, the increasing rate of breakdown voltage is 4KV which increase up to 7KV for a gap distance (5mm-7.5mm) of 2.5mm and 14KV for the next gap distance (7.5mm-10mm) of 2.5mm. Fig.4 shows the breakdown characteristics of Palm oil. It starts from 6.5KV and goes up to 19KV for gap distances of 2.5mm, 5mm, 7.5mm and 10mm with more or less similar increasing rate. From Fig.5 , It is seen that for a gap distance of 2.5mm the breakdown voltage is 8KV and goes up to 21KV for the gap distances mentioned in the figure. Fig.6 shows the breakdown characteristics of mineral Transformer oil.For a gap distance of (2.5mm-7.5mm) the breakdown voltage increasing rate is the same and after that the increasing rate of the breakdown voltage is higher compared to other vegetable oils. Fig.7 shows the comparison of different electric breakdown voltages of several vegetable oils with respect to the breakdown properties of transformer oil. Breakdown property has been used as the base line as it is widely used for the selection of dielectric for power and distribution transformers. Fig.8 and Fig.9 shows the comparison of Partial Discharge (PD) inception voltages and electric field intensities respectively.

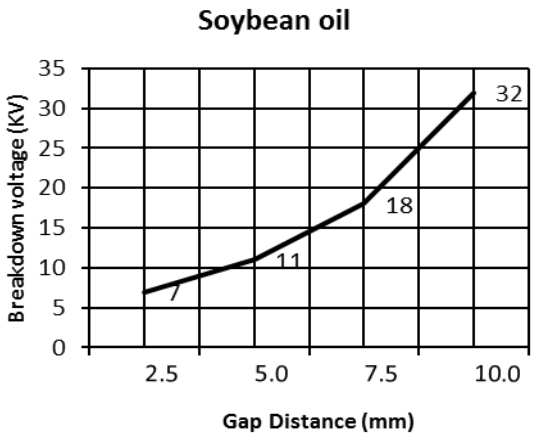



Figure3. Breakdown voltage versus gap distance between electrodes for Soybean oil.

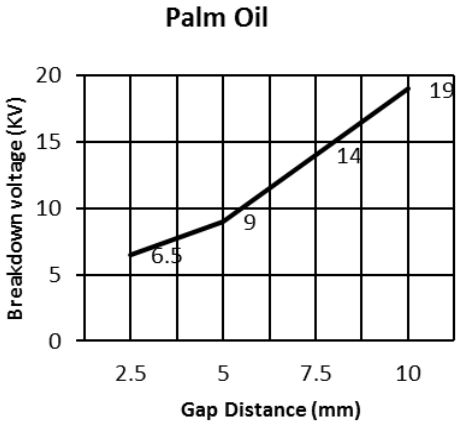



Figure 4. Breakdown voltage versus gap distance between electrodes for Palm oil.

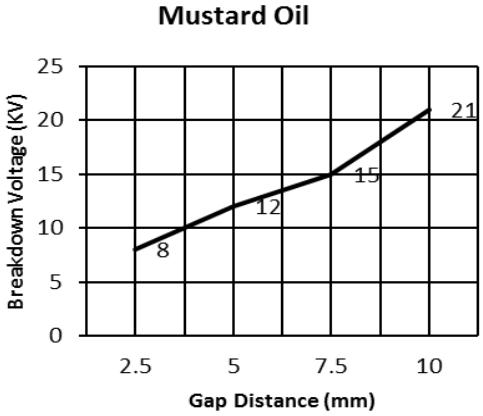



Figure 5. Breakdown voltage versus gap distance between electrodes for Mustard oil.

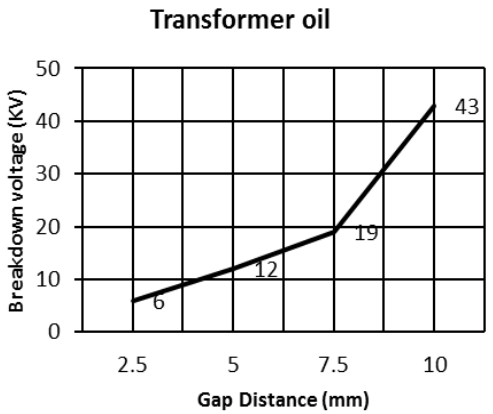



Figure 6. Breakdown voltage versus gap distance between electrodes for Transformer oil.

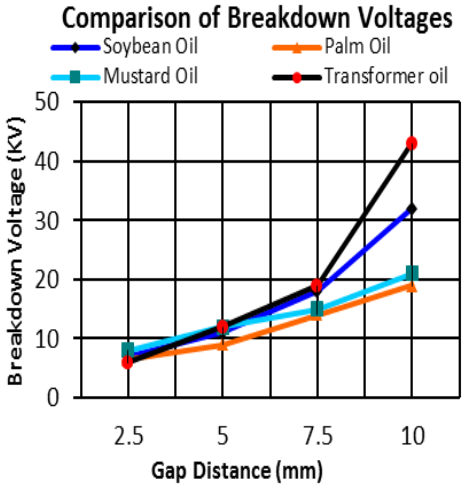



Figure 7. Comparison of breakdown voltages for samples.

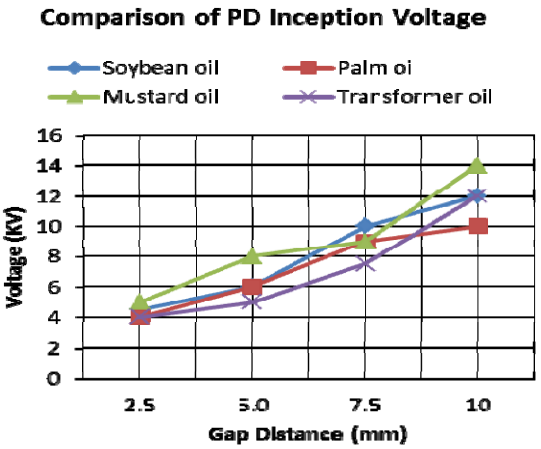



Figure 8. Comparison of PD inception voltages for samples.

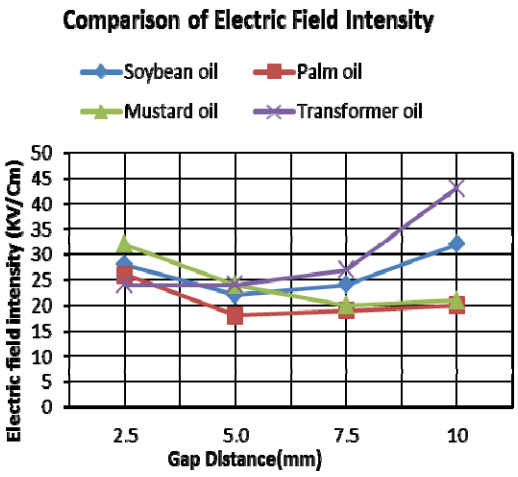



Figure9. Comparison of electric field intensities between electrodes for samples.

## IV. DISCUSSIONS

A. From "Fig. 7", it is identified that for a gap distance of 10mm between electrodes, the breakdown voltage of transformer oil is higher compared to other vegetable oils.

B. The breakdown voltage of transformer oil is rapidly increasing compared to vegetable oils with the increasing gap distance.

C. Mustard oil has the higher breakdown voltage compared to other oils for the use of less than 12KV.

D. For lower gap distance (2.5mm-5mm), the breakdown voltage of all types of oils are approximately equal.

E. For the shortest gap distance (2.5mm), the breakdown voltage of mustard oil is comparatively higher than other oils.

F. It is seen that mustard oil has the higher breakdown voltage compared to other oils for working voltage level within the range of 5-10KV.

G. Palm oil shows the less significant behavior as dielectric liquid.

H. From "Fig. 8", it is seen that, Palm oil and transformer oil has shown the partial discharge inception at comparatively lower voltage.

I. Mustard oil shows the highest voltage for partial discharge inception at higher gap distance of 10mm.

J. Palm oil shows the lowest voltage for partial discharge inception at a gap distance of 10mm.

K. From "Fig.9", it is observed that electric field intensity is more in all the oils for smaller gap distance and decreases with the increased gap distance but after 7.5mm gap distance transformer oil and soybean oil show typical anomalous behavior. This is because of the fact that dielectric strength of mentioned oils increase with the increase in gap distance between the electrodes.

L. The electric field intensity in transformer oil is less compared to other vegetable oils for gap distance of 2.5mm.

## CONCLUSIONS

Breakdown voltages of soybean oil, palm oil, mustard oil and mineral oil have been investigated and compared. As a result of the above discussions, it is concluded that the vegetable oils are suitable for low voltage application (less than 11KV) to use as dielectric fluid. Mustard oil could be considered as a potential dielectric. Among the samples, palm oil is inferior which needs modification to use. Further investigation can be done with pre-processing of the samples and even higher electrode gap with higher applied voltage and also with different electrode configuration.

## REFERENCES

[1] T.V. Oommen, "Vegetable oils for liquid-filled transformers", *IEEE, Electrical Insulation Magazine*, vol.18, no.1, pp.7-11, 2002.

[2] G. J. Pukel, R.Schwarz, F. Schatzl, F. Baumann, A. Gerstl, "Environmental friendly insulating liquids – A challenge for power transformers", *6th Southern Africa Regional conference*, pp.1-8, 2009.

[3] T.V. Oommen, *IEEE, Electrical Insulation Magazine*, vol.18, pp.6-11, 2002.

[4] G. J. Pukel, R. Eberhardt, H. M. Muhr, F. Baumann, W. Lick, "Large power transformers for alternative insulating fluids", *Proc. of the 16th Int. Symposium on High Voltage Engineering, Johannesburg*, paper F-27, 2009.

[5] John Luksich, "Evaluating new and in-service vegetable oil dielectric fluids", Cooper Power systems, Waukesha, WI.

[6] K. Sirikulrat and N. Sirikulrat, "Dielectric properties of different maturity soybean", KMITL Sci.J., vol.8, No.2, pp12-18,2008.

[7] D. Martin, I. Khan, J. Dai, and Z. D. Wang, "An overview of the suitability of vegetable oil dielectrics for use in large power transformers", *Euro.Tech.Con*. pp. 4-23, 2006.

[8] R. Eberhardt, H. M. Muhr, W. Lick, B. Weiser, R. Schwarz and G. Pukel, "Partial discharge behavior of an alternative insulating liquid compared to mineral oil", *IEEE*, 2010.

[9] R. Arora and W. Mosch, *High Voltage Insulation Engineering*, New Age International Publishers Limited, Wiley Eastern Limited, New Delhi, 1995.

# Efficient Power Generation System Design with High Rated KVA Generators for Medium Load Industry

Mohammad Nazmul Houque[1], Md. Shahid Iqbal[2, *], Md. Moniruzzaman[3], Ruchira Shikdar[4], Md. Janibul Alam Soeb[5], Md. Monjurul Islam[6], Ataullah Mishkat[7]
[1, 2, 3, 4, 5, 7] Department of EEE, Mymensingh Engineering College, Bangladesh
[6] Department of CSE, Mymensingh Engineering College, Bangladesh
[*] E-mail: shahidiqbal_05@yahoo.com

*Abstract*—**Power crisis is increasing day by day. By designing efficient power generation model it is possible to meet up some extent of power required by the industry especially for the medium load industry. This paper represents such a model to design an efficient power generation system in medium load (range up to 6000 amp, 440 volts) industry. Previous Paper represents the system with lower rating up to 1250 KVA but the proposed system is high rated up to 2500 KVA. The value of neutral current is also minimized (up to 8 amps). This model used 4 generators in parallel and sufficient cooling arrangement has been built. The final result shows that the proposed model is stable and optimum power is generated.**

***Keywords—Efficient, medium load industry, different loan condition.***

## I. INTRODUCTION

Power system is a complex system and the demand is increased day by day rapidly. The whole world especially the third world is hungry for power and they show a vast eagerness to get it. Many industries in Bangladesh are in land storage and the power crisis is observed almost every day [1]. So it is necessary to make efficient power generation system. Industries are the main consumers of power and this is highly appears that the want of power can only be minimized by efficient power generation system. Actually for an efficient power generation system, at first the design of that system is necessary. And during the designing period it is essential to know the amount of power consumption in KW is needed.

The climate change has a great impact on power system. The last few years many natural calamities took place in the third world country like Bangladesh, India, and Pakistan. These natural calamities affect the electrical energy production. And this change is affecting power systems at all levels, including: distribution transmission and loads [2]. Many methods are proposed to develop electrical energy generation and some solution is obtained with fabulous efficiency. This represents a great change in power system designing, traditionally based on large concentrated power plants. In the case of renewable energy sources, a problem is arisen that the sources of them are not available for all the time and this is an auxiliary support for power system. And the integration problem is also arisen in renewable energy. The rapid depletion (hence, increase of cost) of fossil fuels, rising demand for electricity power, government policy on reduction of greenhouse gas emission, are the key factors leading to a growing in the use of distributed generation (DG) units, including both renewable and non-renewable sources such as photovoltaic (PV), wind turbines, wave generators, fuel cells and gas/steam powered Combined Heat and Power (CHP) stations [1]. So this is also a problem for power system and as well as unit commitment. But like other countries Bangladesh shows its eagerness in developing the power system by using renewable energy. But to full-fill the power demand they should be installed with capabilities which mitigate the oscillation.

On the other hand, the demand of electricity is rising and the existing power system needs to be modified. This paper shows a reliable and efficient power generation system that can be minimized the demand and as well as it tries to generate power with minimal cost.

## II. LIMITATIONS OF EXISTING POWER SYTEM

The main drawbacks of the existing power systems can be summarized below as [1&3]

### *A. Inefficiency:*

Almost 8% of the total power is lost along transmission lines while only one-fifth of its generation capacity exists to meet the peak demand.

### *B. Domino-effect failures:*

It is a strictly hierarchical system where power plants at the top of the chain ensure power delivery to customers' loads at the bottom of the chain. In other words, the power flows in only one direction, which will leads to large-scaled blackout triggered by power plants intermittence or even transmission lines problems. The most well-known failure occurred in August 2003, when 50 million customers in the USA and Canada lost power for up to two days due to cascading events.

### *C. Instability:*

The unprecedented fluctuation of demand for electrical Power, coupled with increasing penetration level of DGs and lagging investments in the electrical power infrastructure, has decreased system stability.

Hence, a design is adapted for efficient power system to meet up the required power demand of the consumers [4].

In the following paragraphs, we describe more in detail the model of efficient power generation system and then we focus on the challenges of designing this system. For this purpose we take a reference model which is done by another person. This model is given below,

## III. System Architecture

The proposed model has been built in the medium load industry with high rated generators in the city of Dhaka, Bangladesh. The present system has four generators which are rated 550, 660, 1250 and 2500 KVA respectively. The generators are connected in parallel. The data taken for 2500 KVA generator with following specification:

Serial: FGWOSPECPSOA02019
Year of manufacture: 2011
Rated Power: 2500 KVA
Rated Current: 3609 A
Rated voltage: 400/230
Rated rpm: 1500
Maximum ambient temperature: 30° C
Excitation voltage: 47 volt
Excitation current: 5A
The system design shown below:

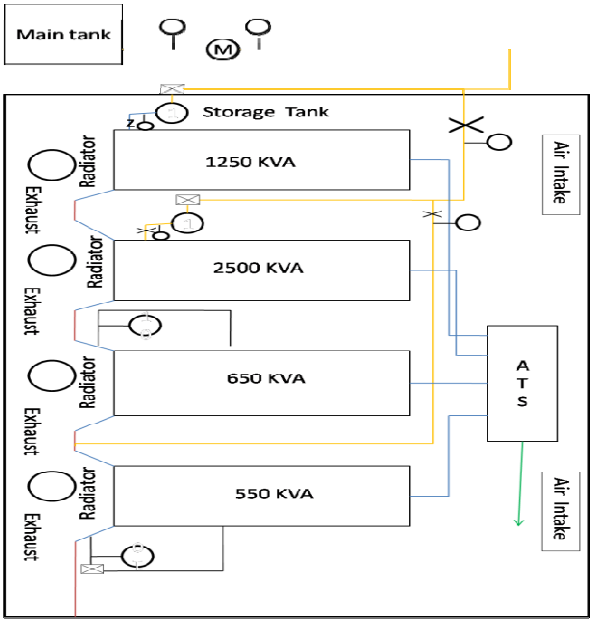



Figure 1: Proposed Model for Power Generation

The proposed model has been designed to get the optimum use of land and materials for maximum and efficient power generation system for the medium load industry which needs 4 or 5MW. This model has some special specifications which are given below:

### A. Compact Space

In medium load industry, there is space shortage. So, idea about the compact space rule is required to design the system within a limited area. A compact space system is a crucial thing to build a system which results in the reduction of its land size, less cost and efficient power system.

### B. Reduced distance from generation to load

From distance rule of resistance drop, it is well know that the resistance drops as power losses increases as the distance between generating point and load point increases. In this proposed system, the generation and distribution systems are very closed. So the power loss is reduced in great extent. So the maximum and efficient power generation and distribution is possible.

### C. Reduced Conductor Cost

As the generation point is very near to the load point, so it reduces the conductor cost and insulation cost which is economical for power generation.

### D. Noise and Heat reduction

For about 3.96 MW power generating station, comfortable heat and noise margin level has to be considered in this model for human safety. Silencer box is used to keep noise level within 85-95 db. Heat reduction is made by the exhaust pipe.

### E. Proper air intake and exhaust

For 4 generators the requirement of total air is above 3,40,000 $cm^3$. Centrifugal blower arrangement has been mounted at the front side of the room to fulfill the demand of the existing air. To meet the air requirement high pressurized blowers are used in this model. The system has 4 exhaust pipe and 3 blower at the back side for exhaust air and free hot air in the outside of the room respectively.

### F. Compact diesel tank

Inlet temperature of generators should be below the 40 degree centigrade. Normally the big Diesel tank (More than 1600 L) is used to fulfill the requirement. But in this model heat exchanger has been used. It is connected with the return of the fuel (200L) line of generator to cool the fuel as its temperature rises above 60°C. So the water coming from chilled water line absorb the heat and fuel mixes with the existing diesel makes the temperature of the total fuel below 40°C. Then finally fuel from main tank comes in the storage tank to keep the fuel in the required level if diesel level low down

## IV. Model Outline

According to the proposed model the outline of the system is given below:

### A. Generator Arrangement

Parallel connected generators has been installed 10 inch high from ground level each and the spacing provided for four

different generators is 3 feet, 6.5 feet, 13.5 feet. During the running time generator creates vibration. For this vibration any kind of coincidence may be occurred in the system. To remove this occurrence a concrete at directly on the top of the floor slab and mount the generator set on this pad. The purpose of this pad is for cleaning around the generator set and provide a strong base level.

### B. Air intake/outlet

The high pressurized air keeps the temperature at stable limit. So in this model the air intake pressure for the first two generators is 5700 CFM and the rest two consecutive generators have 9500 CFM intake pressure. Air is released in environment after circulation. For this purpose 6 exhaust fans are used in outlet to keep environment stable..

### C. Louver System

At the right side of the generator the arrangement of this system is made, so air can be easily flowed by itself. It has 16 feet height and 20 feet long. The generator room and environment is cooled and also reduced the noise of the generator room by this system.

### D. Arrangement of Diesel Tank

The main tank of diesel supply is installed outside the generator room and it is connected to the respective storage tanks of each generator. The main tank consists of four individual tanks and the net capacity is 90000 L. These four tanks are internally connected. One storage tank is at the right side of the 1250 KVA and another one is at the left side of the 2500 KVA generator. And the other two are situated at the down of the 550 KVA and 660 KVA generator. The capacity of the storage tanks for the 2500 KVA, 1250 KVA, 660 KVA and 550 KVA generators are 5000L, 2000L, 1000L and 800L respectively. Here motor is used to pull the diesel to the storage tank and the surrounding pipes. At the side of the pipe beside the generator there is two accessories, gate valve and pressure gauge. And at the side of the motor there is a gate valve, a pressure gauge and a meter. The meter is used to see the flow of the fuel through the pipe, gate valve is used to control the flow of diesel and pressure gauge is for monitoring pressure inside the tank and also in the pipe. Ball valve is also used for some maintenance and flexible use. The unused diesel comes back to the storage tank through a pipe which is connected with the heat exchanger cools the oil. Two thermometers have been used in the two side of the storage tank to see the temperature.

### E. Exhaust System

It is well known that the diesel engine work in four cycles. After a complete cycle of diesel engine each generator produces hot exhaust gas. For this exhaust system has been built. The purpose of this system is to safety discharge the engine combustion product called exhaust gas into the atmosphere outside of the building with minimum hazard to people and environment. Piping arrangement has been built at the roof of the generator room above 20ft from the ground level. It comes from silencer pipe. For four generators four flexible pipes have been used. Exhaust pipe size is adequate to prevent back pressure. In the outside of the room exhaust pipe is horizontal. Exhaust piping components are insulated as necessary to prevent operator burns and reduce pipe radiant heat losses.

### F. Radiators

Radiators gave been built to release the produced heat in engine to the environment. The radiator is not directly mounted at the engine. It is mounted close proximately to the generator set. The close radiator is mounted vertically. The radiator consists of electric fans to provide cooling air.

### G. Fire fighting equipment and safety system

Firefighting equipments have been provided to the generator room. There are five cylinders of $CO_2$ extinguishers are used. Fire proof materials are also used where exhaust pipe passes through building. Also personnel are provided with ear plugs, ear muffs, and maintenance tools.

### H. Noise and heat reduction

Each generator has silencer to reduce the noise in the engine. It is situated in the upper side of the engine. Compliance with local noise codes is always maintained here. For heat elimination dry type insulation has been used across the silencer and exhaust pipe. To reduce the heat in the generator room sufficient air flow is maintained to remove the heat. Cooling arrangement and exhaust fans and also used for this purpose.

### I. Utility Power

During the periods when generators are not running utility power must be provided to power critical component. Battery charger, turbo charger, space heaters and other devices have been used to maintain the generators set components and allow fast easy starting.

## V. EXPERIMENTAL DATA

We know for power, the equation is,

$$P = (\sqrt{3})\ VI \cos\varphi. \quad (1)$$

Here, P = Power; V = Voltage; I = Current; and cosφ = Power factor. This equation is used here and for the simulation the following tables are used,

TABLE I. EXPERIMENTEL DATA OF THE SYSTEM(9.00-10.00AM)

| START TIME | END TIME | TOTAL HOURS | AVG. VOLTS | AVG. AMPS | KW |
|---|---|---|---|---|---|
| 9.00 AM | 10.00 AM | 1 | 400 | 1385.64 | 960 |
| | | | 400 | 1400 | 970 |
| | | | 401 | 1403.78 | 975 |
| | | | 402 | 1421.83 | 990 |
| | | | 403 | 1495.77 | 1044 |
| | | | 403 | 1489.94 | 1040 |
| | | | 403 | 1504.26 | 1050 |
| | | | 404 | 1594.86 | 1116 |
| | | | 404 | 1600 | 1120 |
| | | | 404 | 1629.16 | 1140 |

TABLE II. EXPERIMENTEL DATA OF THE SYSTEM(1.00-2.00PM)

| START TIME | END TIME | TOTAL HOURS | AVG. VOLTS | AVG. AMPS | KW |
|---|---|---|---|---|---|
| 1.00 PM | 2.00 PM | 1 | 402 | 998.16 | 695 |
| | | | 400 | 1154.70 | 800 |
| | | | 403 | 1415.44 | 988 |
| | | | 403 | 1431.19 | 999 |
| | | | 400 | 1472.24 | 1020 |
| | | | 402 | 1522.37 | 1060 |
| | | | 403 | 1540.08 | 1075 |
| | | | 404 | 1564.85 | 1095 |
| | | | 403 | 1575.89 | 1100 |
| | | | 401 | 1655.74 | 1150 |
| | | | 400 | 1703.18 | 1180 |
| | | | 403 | 1719.16 | 1200 |
| | | | 402 | 1737.8 | 1210 |

TABLE III. EXPERIMENTEL DATA OF THE SYSTEM(3.00-4.00PM)

| START TIME | END TIME | TOTAL HOURS | AVG. VOLTS | AVG. AMPS | KW |
|---|---|---|---|---|---|
| 3.00 PM | 4.00 PM | 1 | 404 | 1371.92 | 960 |
| | | | 403 | 1396.82 | 975 |
| | | | 402 | 1407.47 | 980 |
| | | | 400 | 1424.61 | 987 |
| | | | 403 | 1425.47 | 995 |
| | | | 401 | 1497.37 | 1040 |
| | | | 404 | 1514.83 | 1060 |
| | | | 401 | 1547.76 | 1075 |
| | | | 400 | 1580.5 | 1095 |
| | | | 403 | 1604.55 | 1120 |

TABLE IV. OPERATION STATUS AT DIFFERENT POINT

| N.C | C.T | D.T | E.T | LOP |
|---|---|---|---|---|
| 7 | 61 | 36 | 31 | 8.1 |
| 7.1 | 61.1 | 36.1 | 31 | 8.1 |
| 7 | 61 | 36 | 31 | 8.1 |
| 7.1 | 61.2 | 36.2 | 31 | 7.9 |
| 7 | 61 | 36 | 31 | 7.9 |
| 7.1 | 61.1 | 36.1 | 31 | 8.1 |
| 7 | 61 | 36 | 31 | 8.1 |
| 7.1 | 61.2 | 36.2 | 31 | 8.1 |
| 7 | 61 | 36 | 31 | 8.1 |
| 7 | 61 | 36 | 31 | 7.8 |
| 7.2 | 61 | 37 | 32 | 7.8 |
| 7 | 61 | 37 | 32 | 8.1 |
| 7 | 62 | 38 | 32 | 8.1 |
| 7 | 61 | 37 | 32 | 8.1 |
| 7 | 61 | 37 | 32 | 8.1 |
| 7 | 61 | 37 | 32 | 8.1 |
| 7.2 | 61 | 37 | 32 | 8.1 |
| 7.2 | 61 | 37 | 32 | 8.1 |
| 7 | 62 | 38 | 32 | 8.1 |
| 7 | 62 | 38.1 | 32 | 8.01 |
| 7 | 62 | 38.1 | 32 | 8.01 |
| 7 | 62 | 38.1 | 32 | 8.01 |
| 7 | 62 | 37 | 32 | 8.01 |
| 8.1 | 61.2 | 38.2 | 34 | 8.01 |
| 8 | 61 | 38 | 34 | 8.01 |
| 8.2 | 61.2 | 38.2 | 34 | 8.01 |
| 8 | 61 | 38 | 34 | 8.01 |
| 8.1 | 61.3 | 38.3 | 34 | 8.1 |
| 8 | 61 | 38 | 34 | 8.1 |
| 8.1 | 61.4 | 38.4 | 34 | 8.1 |
| 8 | 61 | 38 | 34 | 8.01 |
| 8.1 | 61.5 | 38.6 | 34 | 8.1 |
| 8 | 61 | 38 | 34 | 8.1 |

## VI. EXPERIMENTAL RESULTS

The experimental result shows the variation of coolant temperature and diesel temperature with respect to Load (KW). Here the figures are sectioned by two sections. In first section has three figures which describe the response of variation of CT and DT with Load and second section has three figures which describe the variation of N.C with Load. The figures are shown in below,

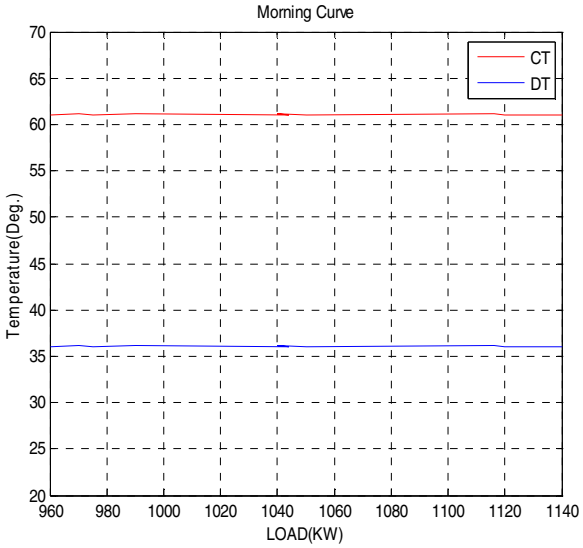



Figure 2: Variation of CT and DT with Load(from 9:00 am to 10:00 am).

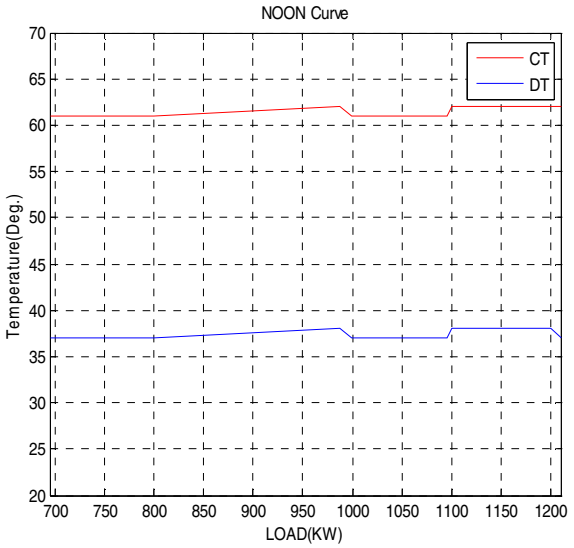



Figure 3: Variation of CT and DT with Load (from 1:00 pm to 2:00 pm).

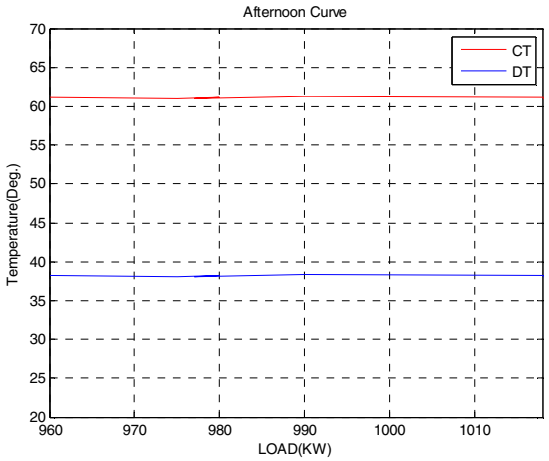



Figure 4: Variation of CT and DT with Load(from 3:00 pm to 4:00 pm).

The variation of neutral current with Load is given below,

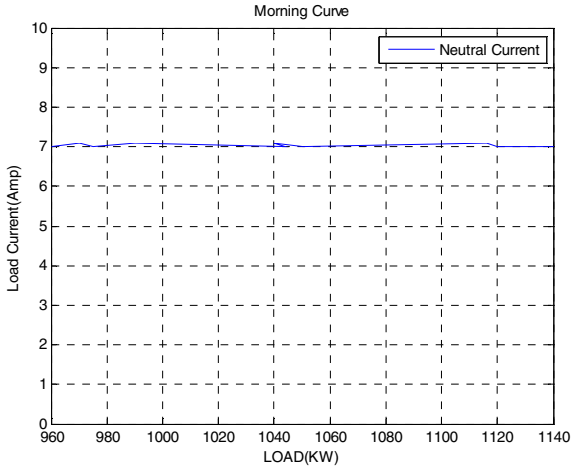



Figure 5: Variation of N.C with Load(from 9:00 am to 10:00 am).

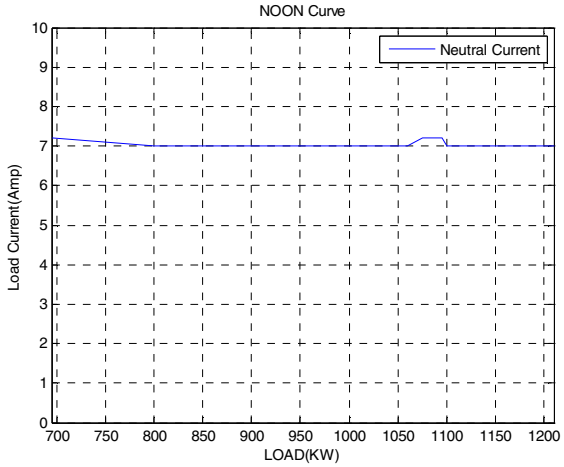



Figure 6: Variation of N.C with Load(from 1:00 pm to 2:00 pm).

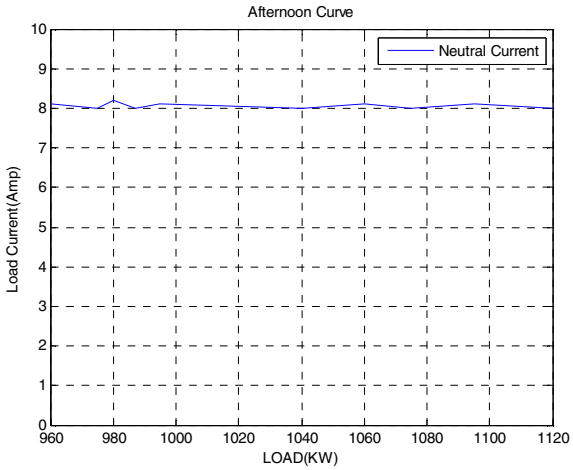



Figure 7: Variation of N.C with Load(from 3:00 pm to 4:00 pm).

## VII. CONCLUTION

This paper shows that by the designing efficient power generation model, it is possible to meet up some extent of power. And this paper presents efficient power production of around 3 MW in industry. In future there will be a provision for altering the present 03 generators with the new 2 MW generator each which can add more 3 MW than the present arrangement. Moreover, the present system will comply with the future upcoming generators easily. Proper diesel fuel supply and air intake should be kept in mind while installing.

## REFERENCES


[1] M. N. Hoque, A. F. Mitul, F. H. M. Rafi and H. M. Hasnaine, "Design of Efficient Power Generation System in Medium Load Industry", *International conference on Mechanical, Industrial and Energy engineering 2012.*

[2] M. Bauer, W. Plappert, C. Wang and K. Dostert, "Packet- oriented communication protocols for Smart Grid Services over low-speed PLC", *Proceeding of ISPLC, Dresden, Germany,* pp.89-94, 2009.

[3] G. N. SrinivasaPrasanna, A. Lakshmi, S. Sumanth, V. Simha, J. Bapat, G. Koomullil, "Data communication over the Smart Grid", *Proc. ISPLC, Dresden, Germany, 2009, pp.273-279.*

[4] A. Ipakchi and F. Albuyeh, "Grid of the future", *IEEE Power Energy Magazine,* vol.7 no. 2, pp. 52-62, Mar/Apr. 2009.

[5] F. Lobo, A. Cabello, A. Lopez, D. Mora and R. Mora, "Distribution Network as communication system: Smart Grids Distribution", CIRED Seminar, Frankfurt, Germany, pp.1-4, 2008.

[6] M. McGranagaghan, D. von Dollen, P. Myrda and E. Gunther "Utility experience with developing a smart grid roadmap", in Proc. Power Energy Soc.Gen.Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, Pittsburg, PA, 2008, pp. 1-5.

[7] S. E. Collier, "Ten steps to smarter grid", Proc. REPC, Fort Collins, CO, 2009, pp. B2-1-B2-7.

[8] A. Mahmood, M. Aamir and M. I. Anis, "Design and Implementation of AMR Smart Grid System", *Proc. EPEC*, Vancouver, BC, Canada, 2008, pp.1-6.

[9] C. Bennett and D. Highfill, "Networking AMI smart meters", in *Proc. Energy 2030 conf.(ENERGY*), Atlanta, GA,pp.1-8, 2008.

[10] A. Sarafi, G. Tsiropoulos and P. Cottis, "Hybrid Wireless-broadband over power lines: A promising broadband solution in rural areas", *IEE Communication*, vol. 47, no. 11, pp.140-147, Nov. 2009.

# Negative Sequence Protection of Generators against Unbalanced Loads using VAMP 210 Generator Protection Relay

Ahmadullah Siddiq[1], Syed Enam Reza[2]
[1]Ahsanullah University of Science & Technology, Bangladesh
[2]Dhaka International University, Bangladesh
[1*] E-mail: ahmadullahsiddiq@gmail.com

*Abstract*— **In rotating machines, an unbalanced voltage or current in one phase winding due to electrical faults or uneven load cause negative sequence current. It is very well known that this current could cause rotor damage, and that damage is highly detrimental to rotating machines such as motors and generators. This paper focuses on the negative sequence components; its effects and protection of an alternator against it using the microprocessor based relay VAMP 210.**

***Keywords— Unbalanced loading, symmetrical components, Inverse delay, VAMP 210 generator protection Relay.***

## I. INTRODUCTION

There are a number of system conditions that can cause unbalanced three-phase currents in a generator [1]. Any sudden disturbance or fault causes an unbalance condition in the network which results in unbalanced phase currents. These currents create negative sequence current in generator stator. Negative sequence current interacts with normal positive sequence current i.e. this component rotates at synchronous speed in a direction opposite to the direction of rotation of rotor to induce a double frequency current (120 Hz) in the rotor. These rotor currents can cause an excessive temperature rise in a very short time and damage the rotor [3]. Unbalanced stator currents also cause severe vibrations to the stator. Under normal conditions, the electric and magnetic fields are circumferentially distributed periodically and evenly across the air gap. However, during electrical disturbances, these fields change abruptly and become distorted. The unsymmetrical magnetic flux densities in the air gap create an unbalanced pull on the rotor [4, 5]. Healthy or loaded phase windings will have higher flux densities than the open-circuited phase or phases. This condition is most likely to exist as a result of unbalanced loading. The worst-case effect is when the rotor rubs on the stator coils.

## II. SYMMETRICAL COMPONENTS

The theory of symmetrical components was developed by Charles Fortesque in 1918. Symmetrical components allow us to break down the problem of solving unbalanced fault conditions in the power system. He showed that any unbalanced system of 3 phase currents or voltages may be regarded as being composed of three separate sets of balanced vectors such as positive phase sequence components, negative phase sequence components and zero sequence components. These three sets of phasors are given the name symmetrical components of the original unbalanced system.

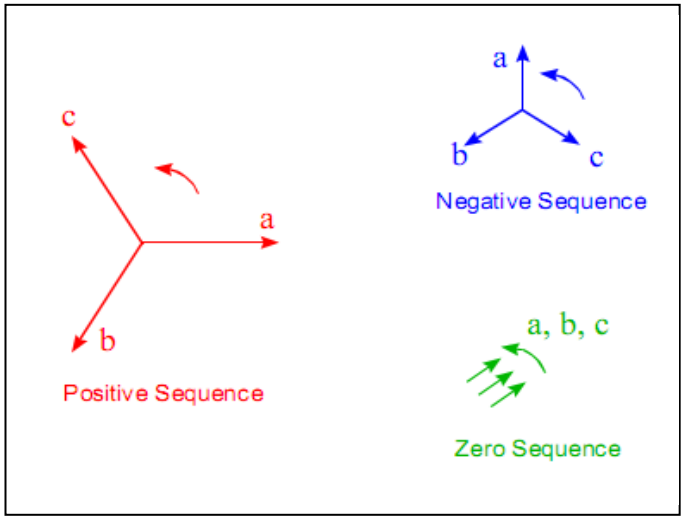



Figure 1: Phasor Diagram (Three phase voltage or current)

### A. *Negative Sequence Current in a Cylindrical Machine Rotor*

These rotor currents can cause an excessive temperature rise in a very short time. Figure shows the general flow of negative sequence current in a cylindrical machine rotor. The current flows across the metal-to-metal contact of the retaining rings to the rotor forging wedges. Due to skin effect, only a small portion of this high-frequency current flows in the field windings. Excessive negative sequence heating beyond rotor thermal limits results in failure of the generator [2]. The negative heating follows the resistance law so it is proportional to the square of the current. The heating time constant usually depend upon the cooling system used and is equal to $K = I_2^2 t$ where $I_2$ is rms value of negative sequence current in pu, $t$ is the current duration in seconds and $K$ is the constant depending on the generator design and size usually

lies between 3 and 20 (The larger the generator the smaller the value of *K*).

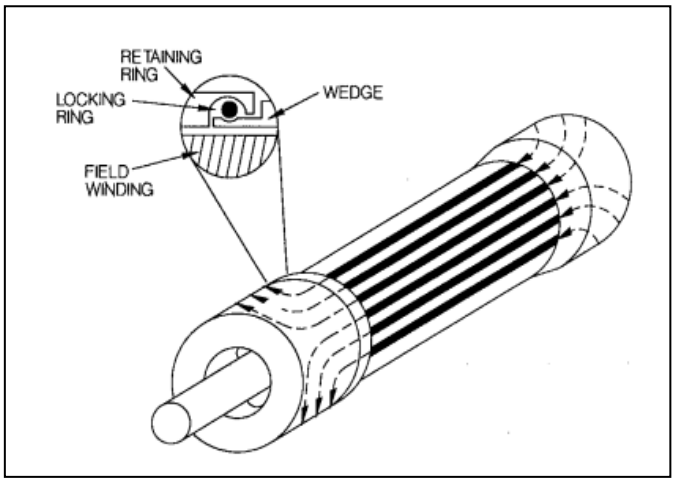



Figure 2: Currents in the rotor surface

### *B. Negative Sequence (Unbalanced Current Protection)*

Unbalanced loading can be protected by using negative sequence current filter with over current relay. It is a general practice to use negative sequence current relays which matches with the heating characteristics ($I^2K$) of the generator. In this type of protection three CTs are connected to three phases and the output from the secondaries of the CTs is fed to the coil of over current relay through negative sequence filter. Negative sequence circuit consists of the resistors and capacitors and these are connected in such way that negative sequence currents flows through the relay coil. The relay can be set to operate at any particular value of the unbalance currents or the negative sequence component current.

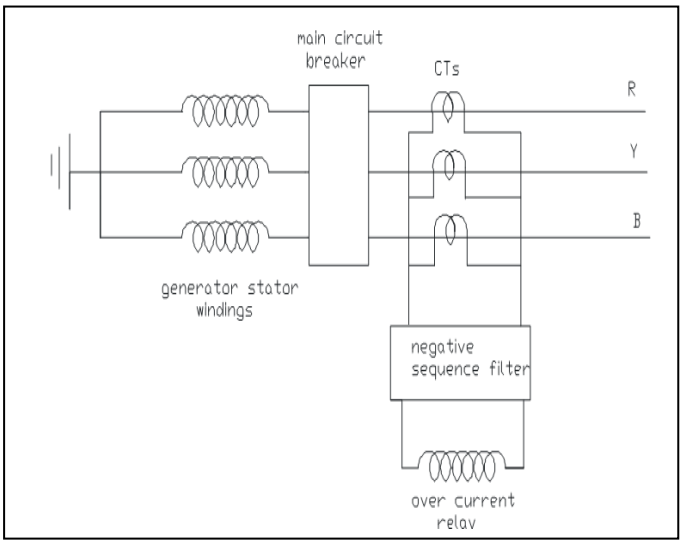



Figure 3: Protection against unbalanced loading

## III. INVERSE DELAY CHARACTERISTICS

The current unbalance protection is based on the negative sequence of the base frequency phase currents. In this relay (VAMP 210) both definite time and inverse time characters are available.

The inverse delay is based on the following equation:

$$t = \frac{K_1}{\left(\frac{I_2}{I_1}\right)^2 - K_2^{\,2}}$$

Where, $t$ is operation time, $K_1$ is delay multiplier, $I_2$ is measured and calculated negative sequence phase current of fundamental frequency, $I_{gn}$ is rated current of the generator, $K_2$ is pick-up setting $I_2>$ in pu. The maximum allowed degree of unbalance.

If more than one definite time delay stages are needed for current unbalanced protection, the freely programmable stages can be used in this relay.

### *A. Setting Group of VAMP 210*

There are two setting groups available. Switching between setting groups can be controlled by digital inputs, virtual inputs (mimic display, communication, logic) and manually.

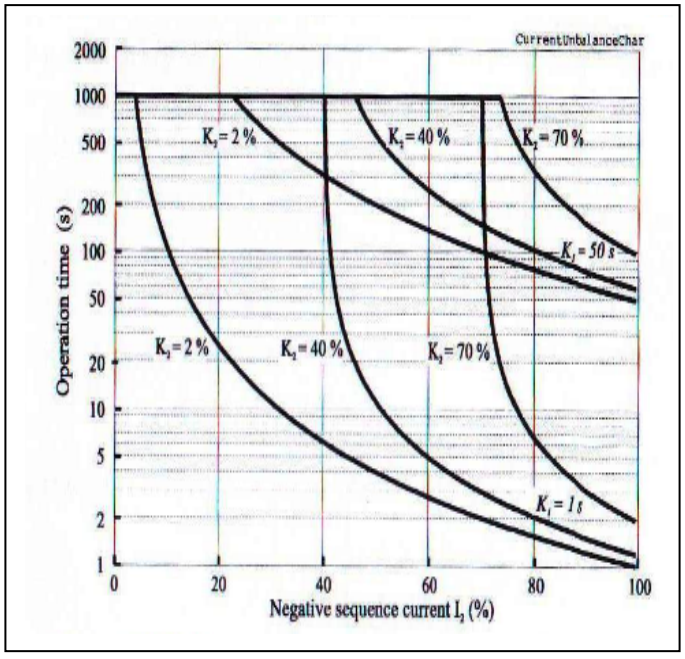



Figure 4: Inverse operation delay characteristics unbalance stage $I_2>$

TABLE I. PARAMETERS OF THE CURRENT UNBALANCED STAGE $I_2$> IN VAMP 210TABLE TYPE STYLES

| Parameter | Value | Unit | Description | Notes |
|---|---|---|---|---|
| Status | -<br>Blocked<br>Start<br>Trip | | Current Status of the Stage | <br><br>F<br>F |
| SCntr | | | Cumulative start counter | C |
| TCntr | | | Cumulative trip counter | C |
| SetGrp | 1 or 2 | | Active setting group | Set |
| SGrpDI | -<br>DIx<br>VIx<br>LEDx<br>VOx | | Digital signal to select the active setting group<br>None<br>Digital input<br>Virtual input<br>LED indicator signal<br>Virtual output | Set |
| Force | Off<br>On | | Force flag for status forcing for test purposes. This is a common flag for all stages and output relays, too. Automatically reset by a 5 minute timeout. | Set |
| $I_2/I_{gen}$ | | %Ign | The supervised value. | |
| $I_2$> | | %Ign | Pick-up setting | Set |
| t> | | s | Definite time operation (Type=DT) | Set |
| Type | DT<br>INV | | Definite time<br>Inverse time | Set |
| $K_1$ | | s | Delay multiplier (Type=INV) | Set |

## IV. MAIN FEATURES OF VAMP 210 RELAY

VAMP 210 relay utilizes the numerical protection technique which means all the signal filtering, protection and control functions are implemented through digital processing. An adapted Fast Fourier Transformation (FFT) algorithm is used in this regard. Synchronized sampling of the measured voltage and currents signals are used. The sampling rate is 32 samples/cycle within the frequency range of 45 Hz to 65 Hz.

The main features of VAMP 210 are:

- Easy adaptability of the relay to the power plant automation system or SCADA systems using the wide range of available communication protocols.
- Flexible control and blocking possibilities due to digital signal control inputs (DI) and outputs (DO).
- Freely programmable interlocking schemes with basic logic functions and timer.
- Recording of time stamped events and faults values with built-in disturbance recorder for evaluating all the analogue and digital signals.
- Arc protection is available as option.
- Fully digital signal handling with a powerful 16-bit microprocessor.

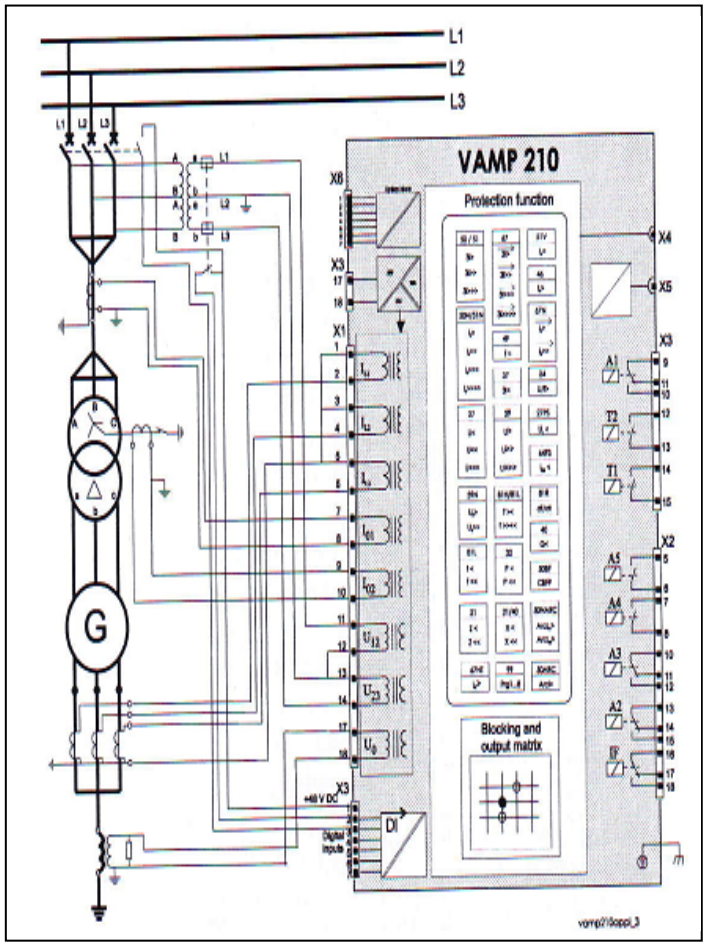



Figure 5: Functional diagram of the VAMP 210 generator protection relay

## V. BLOCK DIAGRAM

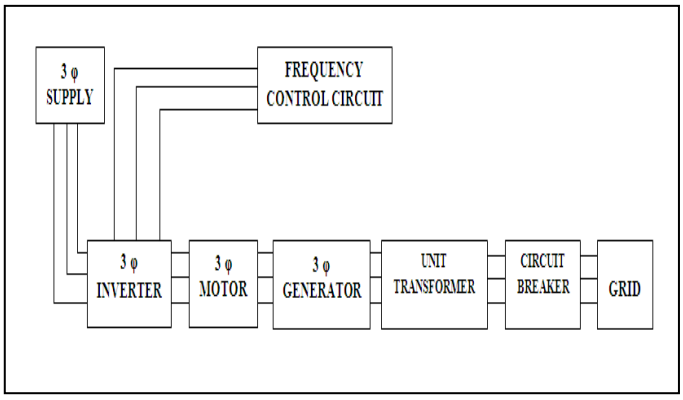



Figure 6: The block diagram of the generator protection scheme designed in our project

The block diagram above shows the entire generator protection scheme installed in our project. According to the diagram a three phase supply is directly fed to the three phase inverter. The inverter is connected with the generator through a three phase motor. In this diagram the inverter has used to vary the frequency of the motor by using the frequency control

circuit. By varying the motor frequency, generator frequency can be varied. The generator shaft has been rotated by coupling it with the output of the motor. The next stage is a unit transformer which is used to provide an electrical isolation between the generator and the grid so as to protect the generator from getting damaged by the overcurrent from the grid. The transformer unit is connected to the grid through a circuit breaker.

Figure 7: The complete schematic diagram of the generator protection project using VAMP 210 relay. First picture shows the Generator-Transformer unit, Second picture indicates the control panel for the fault clearing along with the entire block diagram of the system, 3 (right) denotes side view of generator along with the motor below which rotates the shaft of generator, 4 (left) Shows the Inverter system which Is used to control the speed of the generator.

## V. SIMULATION & RESULTS

In our project we have modeled different abnormal conditions in order to observe the performance of the VAMP 210 relay. As we know Negative sequence relay protects generator from excessive heating in the rotor due to unbalanced stator currents. During unbalanced condition Rotor temperature rise in proportion to $I_2^2t$ and relays provide settings for this relationship in the form of a constant, $k = I_2^2t$. In our project the relay setting was 2.5% of the unbalanced Current and with a operation delay of 5 second was provided. We took data for the variation of load. Initially we had 480W resistive load and then gradually decreased the load.

Ahsanullah power station

UNBALANCE STAGE I2> 46

| | |
|---|---|
| Enable for I2> | On |
| Unbalance I2/Igen | 2.5 %Ign |
| Status | - |
| Start counter | 0 |
| Trip counter | 0 |
| Set group DI control | - |
| Group | 1 |

| | Group 1 | Group 2 |
|---|---|---|
| Pick-up setting K2 | 5 % | 20 % |
| Delay type | DT | DT |
| Operation delay | 5.0 s | 10.0 s |

**(a)**

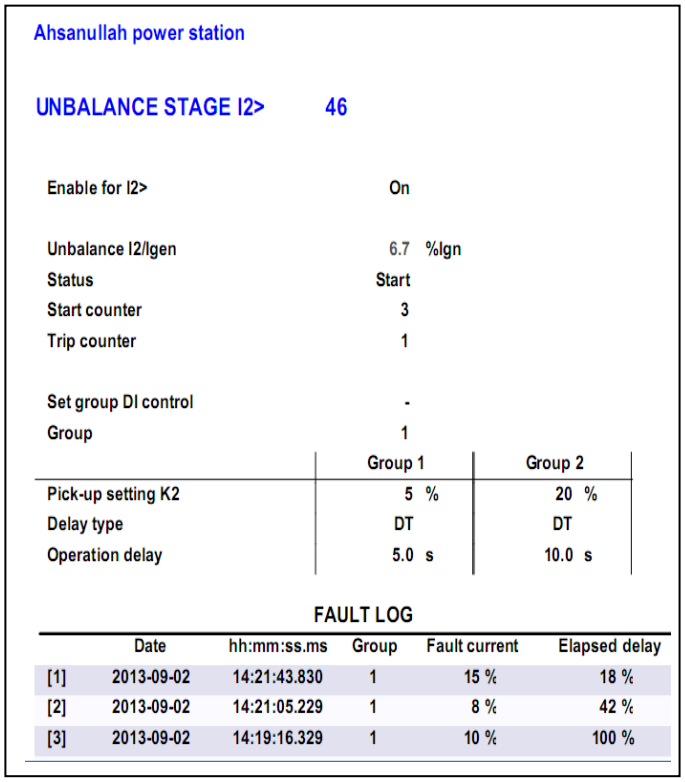
Ahsanullah power station

UNBALANCE STAGE I2> 46

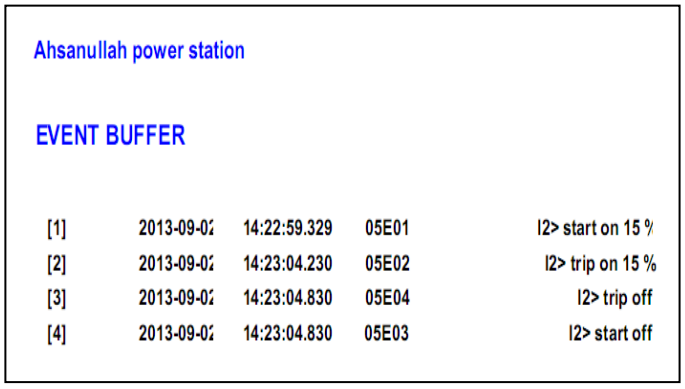
| | |
|---|---|
| Enable for I2> | Off |
| Unbalance I2/Igen | 4.1 %Ign |
| Status | - |
| Start counter | 0 |
| Trip counter | 0 |
| Set group DI control | - |
| Group | 1 |

| | Group 1 | Group 2 |
|---|---|---|
| Pick-up setting K2 | 5 % | 20 % |
| Delay type | DT | DT |
| Operation delay | 5.0 s | 10.0 s |

**(b)**




**(c)**




**(d)**

Figure 8: (a), (b) and (c) shows the simulation data incorporating VAMP 210 relay during the gradual variation of load to view the effects of unbalanced condition and (d) shows the history of relay operation.

## VI. CONCLUSION

In a rotating machine, the negative sequence current vector rotates in the same direction as the rotor. It is the magnetic flux produced by the negative sequence current that rotates in the reverse direction of the rotor. Thus, the rotor cuts through the flux at twice the synchronous speed, and the induced current in the rotor is twice the line frequency which is harmful for the machine. In this paper, the authors first review the concept of sequence components and its effects in the rotating machines. Then explains the features and functional operation of the microprocessor based relay VAMP 210 for the protection against negative sequence current. Moreover, necessary experiment has been performed to observe the effects of unbalanced condition.

This project has been performed using VAMP 210 relay which is comparative a new relay with a robust protection schemes. The necessary results have been shown in the result & simulation section. The relay was operated perfectly against the unbalanced fault during the performing this project. In future, faster operating digital relays can be introduced incorporating IEC standards.

## REFERENCES

[1] Steve Turner, “Considerations for Generator Protection during Black Start Conditions” Transmission and Distribution Conference and Exposition,Doi:10.1109/TDC.2010.5484200,19-22April 2010

[2] Charles J. Mozina, “Upgrading the Protection of Industrial-Sized Generators Using Digital Technology”, IEEE Transactions on Industry Applications, VOL. 33, NO. 4, p1117-1123, JULY/AUGUST 1997

[3] Charles J. Mozina, “Impact of Power System Instability on Generator Protection“, Industrial & Commercial Power Systems Technical Conference (I&CPS), Doi:10.1109/ICPS.2012.6229613, 20-24 May 2012

[4] S. Morii, N. Nagai, and K. Katayama, "Stability Analysis of Rotors Operating in Magnetic Field, Paper C500/047/96," presented at 6th International Conference on "Vibrations on Rotating Machines, 9-12 September 1996.

[5] T. J. E. Miller, "Faults and Unbalance Forces in the Switched Reluctance Machine," IEEE Transactions on Industry Applications, vol. 31, pp. 319 to 328, 1995.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Optimization of Power System Operation with Static Var Compensator applying ACO Algorithm

S. M. Rakiul Islam, Md. Alimul Ahsan* and Bashudeb Chandra Ghosh
Khulna University of Engineering and Technology, Bangladesh
* E-mail: kuet0803013@hotmail.com

*Abstract*— **Loss minimization, cost minimization and voltage stability are the major concerns for modern power system networks operating under highly stressed conditions due to continuously increased power demand. So, it is necessary to analyze the power system considering these factors. A Flexible AC Transmission System (FACTS) device such as Static Var Compensator (SVC) in a power system improves the voltage stability, reduces the power loss and also improves the load ability of the system. This study investigates the application of Ant Colony Optimization (ACO) technique to find optimal location and rated value of SVC thus optimizing total power loss, generation cost, FACTS device cost and improving voltage profile in the power system. Simulations are performed on IEEE 14 bus and IEEE 57 bus. It is observed from the results that total loss is optimized, total cost is minimized and voltage profile is improved.**

***Keywords— Power System Optimization, Ant Colony Optimization (ACO), Static Var Compensator (SVC).***

## I. INTRODUCTION

The operation of power system is becoming more and more challenging because of continuously increasing load demand which is leading to an augmented stress of the transmission lines, voltage instability, increase in loss and cost. To meet the ever increasing demand it is now essential to maximize the utilization of the existing transmission system.

In recent years, due to advancement in high power solid-state switches, transmission controllers have been developed which provides more flexibility and controllability. A new solution for controlling power flow known as FACTS was introduced in 1988 by Hingorani [1]. FACTS devices have made the power system operation more flexible and secure. They have the ability to control, in a fast and effective manner. FACTS controllers minimizes loss, enhance the voltage profile and the load ability of power systems [2].

FACTS devices include Thyristor Controlled Series Compensator (TCSC), Static VAR Compensator (SVC), Thyristor Controlled Phase Angle Regulator (TCPST), Static Compensator (STATCOM), Unified Power Flow Controller (UPFC), etc.

In this paper, SVC is used for several reasons. The most widely used shunt FACTS devices within power networks is the SVC due to its low cost and good performance in system enhancement. It is more conventional and available. SVC can control voltage with higher level of accuracy. It is a shunt-connected static Var generator or absorber whose output is adjusted to exchange capacitive or inductive current so as to provide voltage support and when installed in a proper location, it can also reduce power losses [3]. For these reasons, SVC is chosen over other FACTS devices in this paper.

Various FACTS controllers, their modelling and their impact on power systems have been reported in the literature [4]-[10]. But, most effective use of the FACTS devices largely depends on how these devices are placed in the power systems, i.e. on type, location and size [11]. An optimal location of FACTS devices allows controlling its power flows and thus enhances the reliability of the power systems [12]. Optimal location and rated value can be found by applying established search algorithms such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Improved Harmony Search (IHS), Ant Colony Optimization (ACO) etc.

ACO is a probabilistic technique for optimization initially proposed by Marco Dorigo [13][14].The inspiring source of ant colony optimization is the foraging behaviour of real ant colonies. This behaviour is exploited in artificial ant colonies for the search of approximate solutions to discrete optimization problems, to continuous optimization problems.

In the natural world, ants (initially) wander randomly, and upon finding food return to their colony while laying down pheromone trails. If other ants find such a path, they are likely not to keep travelling at random, but to instead follow the trail; returning and reinforcing it if they eventually find food.

Over time, however, the pheromone trail starts to evaporate, thus reducing its attractive strength. The more time it takes for an ant to travel down the path and back again, the more time the pheromones have to evaporate. A short path, by comparison, gets marched over more frequently, and thus the pheromone density becomes higher on shorter paths than longer ones. Pheromone evaporation also has the advantage of avoiding the convergence to a locally optimal solution. If there were no evaporation at all, the paths chosen by the first ants would tend to be excessively attractive to the following ones. In that case, the exploration of the solution space would be constrained.

Thus, when one ant finds a good (i.e., short) path from the colony to a food source, other ants are more likely to follow that path, and positive feedback eventually leads to all the ants' following a single path. The idea of the ant colony algorithm is to mimic this behaviour with "simulated ants" walking around the graph representing the problem to solve.

Various researches using different FACTS devices and applying GA, PSO have been reported [15]-[20]. This paper therefore proposes an ACO technique to determine optimal location and rated value of SVC in transmission network to optimize loss and voltage profile. The proposed method is verified for IEEE 14 bus and IEEE 57 bus.

## II. PROBLEM FORMULATION

### A. Mathematical Formulation

In this present research, optimal location and size of SVC has been found by formulating multi-objective optimal power flow. Certain Objectives subject to satisfying some network constraints have been minimized. Mathematically, the OPF problem can be written as follows:

Minimize,

$$F(x,u)=\begin{bmatrix} f_1(x,u) \\ . \\ . \\ f_k(x,u) \end{bmatrix} \quad (1)$$

Where, $k = 1, 2,\ldots\ldots$ number of objectives

Subject to

$$g(x,y)=0$$
$$h(x,u)\le 0$$

Here, $x$ and $u$ represent vector of dependent and control variables respectively. For example, the dependent variables include slack bus power, bus voltage angles, load bus voltage magnitudes etc. Whereas, control variables include PV bus voltage magnitude, generated power etc. The real and reactive power balance equations are denoted by *g(x,u)* and components operational limits are denoted by *h(x,u)*.

### B. Objective Functions

*1) Cost*

The total cost is calculated from fuel cost and SVC cost. The fuel cost can be determined from the following quadratic equation.

$$\min(F(P_i))=\sum_{i=1}^{NG}(a_iP_i^2+b_iP_i+c_i) \quad (2)$$

Where, $a_i$, $b_i$, $c_i$ are cost coefficients, $P_i$ is real power generation and $NG$ is the number of generation buses.

The SVC cost can be determined from the following equation.

$$\min(C_{SVC})=\min(0.0003S^2-0.305S+127.38) \quad (3)$$

Where, $C_{SVC}$ denotes Cost of SVC, $S$ denotes Operating range of SVC in MVAR.

The objective function is then formulated using total cost.

*2) Power Loss*

The objective function is formulated using total real power loss. The value of current depends on the bus voltages and transmission line parameters. After getting all branch data the loss is accumulated and finally the total loss is found.

*3) Voltage Deviation*

To have a good voltage performance, the voltage deviation at each load bus must be made as small as possible. The minimum voltage deviation (VD) is determined by the following equation:

$$\min(VD)=\min(\sum_{i=1}^{n}(|V_i-1|)^2) \quad (4)$$

Where, $V_i$ denotes voltage magnitude at load bus i.

### C. Constraints

*1) Facts Device Constraints*

The FACTS device i.e. SVC injects extra reactive power to the PQ bus. So, the search domain for FACTS device constraint is location and value. The location domain is only PQ buses. Due to capacity limit, the FACTS device rated value limit is represented by following equation.

$$Q^{\min}\le Q\le Q^{\max} \quad (5)$$

Where, $Q$ is the FACTS device rated value in MVAR.

*2) Power Loss Constraints*

The system should have minimum loss. But, the FACTS devices have a limit up to which the loss can be minimized. So, the loss constraint is

$$P_{Loss}^{\min}\le P_{Loss}\le P_{Loss}^{\max} \quad (6)$$

Where, $P_{Loss}$ denotes Real power loss in MW

*3) Cost Constraints*

Due to decrease in fuel cost the total cost per unit decreases. But, the per unit cost increases due to FACTS device cost. It is also not possible to make the loss zero. So, the power system should be optimized in such a way that the ultimate per unit cost must be less than per unit cost without using FACTS device. Hence, the cost constraint is as follow:

$$F^{\min}\le F\le F^{\max} \quad (7)$$

Where, $F$ denotes Total cost.

*4) Voltage Profile Constraints*

The desired value of voltage for each bus is unity in per unit scale. But, the reality is the voltage of buses never to be equal to unity. So, a range of voltage is considered as bus voltage profile constraints in the time of optimization.

$$|V_i^{\min}|\le|V_i|\le|V_i^{\max}| \quad (8)$$

Where, $V_i$ denotes Voltage of $i^{th}$ bus in per unit.

*5) Power Balance Constraints*

Total generated power of the power system must be equal to the sum of total power demand and total power loss. These criteria can be expressed by the following equations.

$$\sum P_G = \sum P_D + \sum P_L \quad (9)$$

$$\sum Q_G = \sum Q_D + \sum Q_L \quad (10)$$

Where, $P_G$= Generated real power in MW
$P_D$= Real power demand in MW
$P_L$= Real power loss in MW
$Q_G$= Generated reactive power in MVAR
$Q_D$= Reactive power demand in MVAR
$Q_L$= Reactive power loss in MVAR

Power demand and power loss are variables. But, power generation is limited from minimum to maximum output due to economy and capacity. Hence, power generation constraints can be represented as

$$\left|P_{Gi}^{\min}\right| \le \left|P_{Gi}\right| \le \left|P_{Gi}^{\max}\right| \quad (11)$$

$$\left|Q_{Gi}^{\min}\right| \le \left|Q_{Gi}\right| \le \left|Q_{Gi}^{\max}\right| \quad (12)$$

Where, $P_{Gi}$=Real power generation at $i^{th}$ bus

$Q_{Gi}$= Reactive power generation at $i^{th}$ bus.

## III. IMPLEMENTATION OF ACO

### *A. Initialization of Nodes*

The procedure starts with creation of a population of a fixed number of initial nodes from where the search process is initiated. Each node is an n-dimensional vector given by $(x_1,x_2,x_3,\ldots\ldots,x_n)$. Each coordinate of this vector is a randomly generated real number between the ranges of search space.

### *B. Fitness of Nodes*

The effectiveness of generated nodes is tested using a fitness function. For the node $(x_1,x_2,x_3,\ldots\ldots,x_n)$, it is given by

$$f(x) = 1 - \frac{1}{1+F(x)} \quad (13)$$

Here, *F(x)* is the objective function. A node whose fitness value, which is always taken as positive, equals to zero is taken as a solution. The value of fitness gives an idea about how far the node is from the solution of the equation. The guiding spirit behind the search process is to find nodes, which minimizes the value of this fitness function through movement of ants.

### *C. Creation of Neighbours*

The neighbours of a node are chosen using a special strategy. As per this, for each node, a fixed number of neighbours are generated within an admissible range. The coordinates of each of these neighbours are altered by adding a real random number between ranges to the value of the coordinate of the node. If the updated value of the coordinate is within the admissible range of the nodes, the value is retained. If this number crosses the range, the remainder is taken on dividing the number by "*p*". Similarly, other coordinates of the neighbour are constructed. That means, the neighbours of $(x_1,x_2,x_3,\ldots\ldots,x_n)$, are the collection of nodes $(x_{p1},x_{p2},x_{p3},\ldots.x_{pn})$ where each $x_{pi}$ can be at a maximum distance of modulo "*p*" from the coordinate $x_i$.

### *D. Pheromone Distribution*

An ant deposits a pheromone amount equivalent to its effectiveness of the node where it resides. Based on the strength of the pheromone deposit at the node, the surrounding ants could be attracted to choose the node with higher pheromone content. The quantity of pheromone deposited by an ant at a visited node is given by

$$Phero(x) = \frac{1}{f(x)} \quad (14)$$

Here, *f(x)* is the fitness of the node *x* where the ant is located now.

This simple and straight forward depiction of pheromone is derived from the fact that the present problem is a minimisation problem to find a node with fitness value zero. Thus, nodes with smaller fitness values have been given greater pheromone content as they are close to the solutions. This way, more ants are encouraged to traverse through such nodes.

Once each of the nodes in the topological neighbourhood of a node has been allocated the pheromone, the ant chooses a node using the probability

$$prob(x) = \frac{Phero(x)}{\sum Phero(x)} \quad (15)$$

Here the summation applies to all nodes, which are in the topological neighbourhood of the given node. In practice, this is implemented by generating a uniform random number by the system and selecting the node satisfying the probability condition

### *E. Pheromone Update*

The nodes which have already been encountered in the search process have been given additional weightage in the search by offering an additional amount of pheromone. The pheromone update used in the work is given by

$$PheroUpdate(x) = 0.01 * Phero(x) \quad (16)$$

This additional pheromone content is offered in addition to the existing pheromone for the nodes, which have already been visited by another ant.

### *F. Pheromone Evaporation*

The unchecked and continuous update of pheromone at the nodes, which have already been visited by other ants, creates a possibility of premature convergence, the phenomenon in

which most of the ants are forced to visit the nodes with an accumulated pheromone deposit. The pheromone content is reduced gradually when an ant visits a node, which has already been visited more than once by other ants. The procedure adopts the following formula for the pheromone evaporation.

$$PheroEvaporation(x) = \frac{Visits \times Phero(x)}{100} \quad (17)$$

After each iteration, ants are distributed through nodes according to the new pheromone level determined by adding pheromone update and pheromone evaporation.

## IV. Simulation Results

The proposed ACO technique is simulated to validate. The simulation has been completed by developing and incorporating codes in MATLAB. MATPOWER4.1 (a package of code) has been used as associates of main code to get and format system data. The simulation technique has been tested in IEEE 14 bus system and IEEE 57 bus system. Both the solutions with and without using SVC have been obtained by ACO. Parameters of ACO are given below:

TABLE I. ACO Parameters

| Parameters | Values |
|---|---|
| Nodes | 10 |
| Pheromone evaporation rate | 0.99 |
| Pheromone update rate | 0.01 |

### A. IEEE 14 Bus System

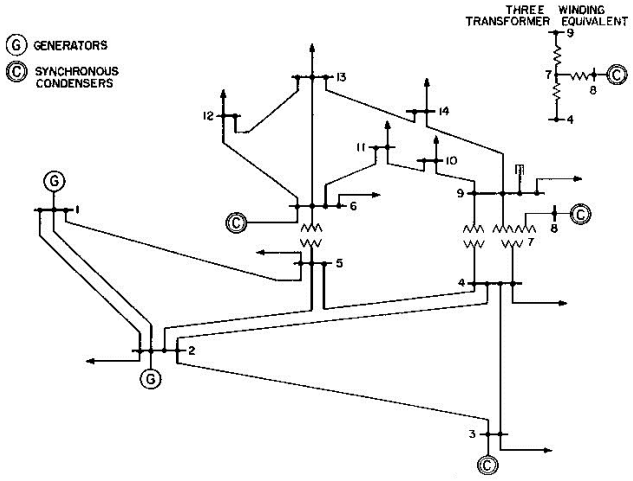



Figure 1. IEEE 14 bus system.

The IEEE 14 bus system is shown in Fig.1. The optimized system data for IEEE 14 bus system both using and without using SVC are shown in table II. The optimal location of SVC is found at bus 5 and the optimal value is 26.8016 MVAR. Using SVC, loss is reduced by 0.74%. The cost is reduced by 0.0028 unit per megawatt. The voltage profile of IEEE 14 bus system optimized without using SVC and using SVC are shown in Fig. 2 and Fig. 3 respectively.

TABLE II. Optimized IEEE 14 Bus System Data applying ACO

| Terms | Optimization using SVC | Optimization without using SVC |
|---|---|---|
| Buses | 14 | 14 |
| Generators | 5 | 5 |
| Loads | 11 | 11 |
| Transformers | 3 | 3 |
| Total Generation Capacity(MW) | 772.4 | 772.4 |
| Total generated power (MW) | 272.3 | 272.4 |
| Power demand(MW) | 259.0 | 259.0 |
| Power loss(MW) | 13.3 | 13.4 |
| Cost ($/MW) | 29.9968 | 29.9996 |

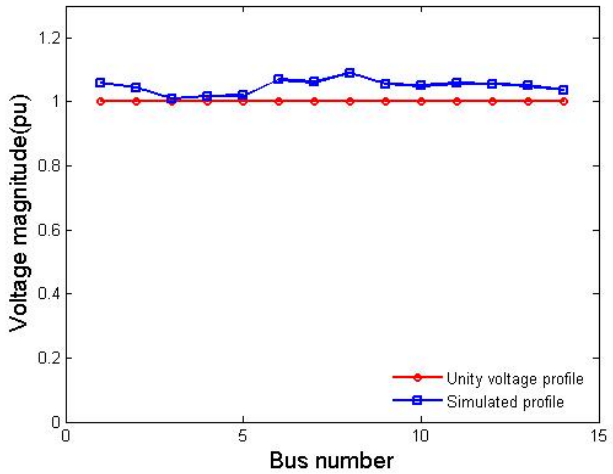



Figure 2. Voltage profile of IEEE 14 bus system optimized without using SVC.

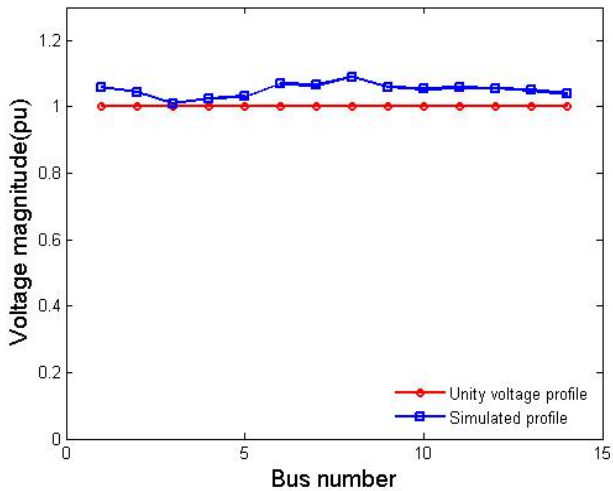



Figure 3. Voltage profile of IEEE 14 bus system optimized using SVC.

## B. IEEE 57 Bus System

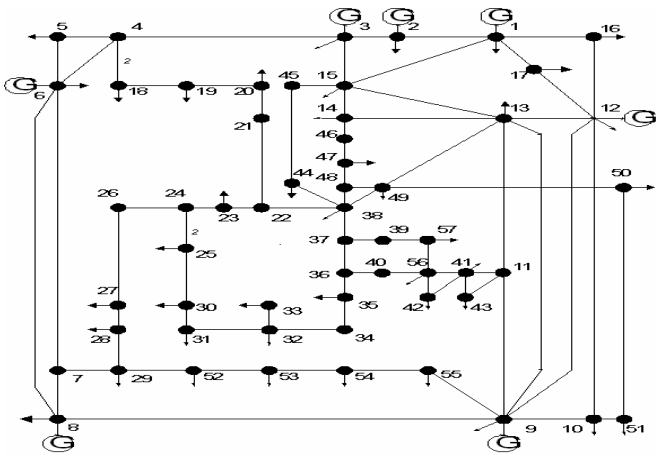


Figure 4. IEEE 57 bus system.

IEEE 57 bus system is shown in Fig. 4. The optimized system data for IEEE 57 bus system both using and without using SVC are shown in table III. The optimal location of SVC is found at bus 38 and the optimal value is 42.1807 MVAR. Using SVC, loss is reduced by 2.59%. The cost is reduced by 0.0305 unit per megawatt. The voltage profile of IEEE 57 bus system optimized without using SVC and using SVC are shown in Fig. 5 and Fig. 6 respectively.

TABLE III. OPTIMIZED IEEE 57 BUS SYSTEM DATA APPLYING ACO

| Terms | Using SVC | Without using SVC |
|---|---|---|
| Buses | 57 | 57 |
| Generators | 7 | 7 |
| Loads | 42 | 42 |
| Transformers | 17 | 17 |
| Total Generation Capacity(MW) | 1975.9 | 1975.9 |
| Total generated power (MW) | 1277.9 | 1278.7 |
| Power demand(MW) | 1250.8 | 1250.8 |
| Power loss(MW) | 27.141 | 27.864 |
| Cost ($/MW) | 40.1271 | 40.1576 |

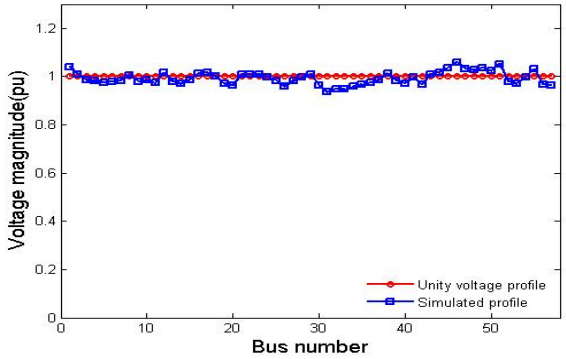



Figure 5. Voltage profile of IEEE 57 bus system optimized without using SVC.

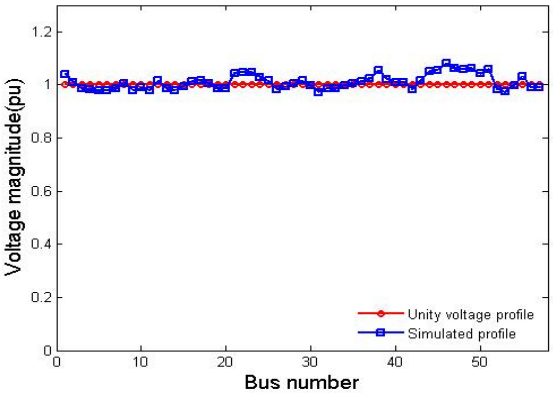



Figure 6. Voltage profile of IEEE 57 bus system optimized using SVC.

## V. CONCLUSION

In this paper, ACO has been used to find the optimal location and size of SVC device for the purpose of minimizing loss, cost and improving voltage. Other popular optimization techniques such as GA, PSO have already been shown in different papers. So, they have not been repeated here for the sake of the length of the paper. Simulations have been performed on IEEE 14 bus and IEEE 57 bus systems using MATLAB. From the obtained results, it is obvious that loss and cost have been reduced and voltage profile has been improved to a great extent using SVC. In fact, the improvement is clearer in IEEE 57 bus system. The proposed ACO technique has shown superior features including high quality solution, stable convergence characteristics and good computational efficiency.

## REFERENCES


[1] N. G. Hingorani, "Power electronics in electrical utilities: role of power electronics in future power systems", *Proceedings of the IEEE,* Vol. 76 No. 4, pp.481-482, Apr. 1988.

[2] Sode-Yome, A., N. Mithulananthan and K.Y. Lee, 2005. "Static voltage stability margin enhancement using STATCOM, TCSC and SSSC.", *Proceedings of the IEEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific, Aug.M18-18*, IEEE Xplore Press, Dalian, pp: 1-6. DOI: 10.1109/TDC.2005.1547141.

[3] Reza Sirjani, Azah Mohamed, Hussain Shareef, "Optimal placement and sizing of Static Var Compensators in power systems using Improved Harmony Search Algorithm", *PRZEGLĄD ELEKTROTECHNICZNY (Electrical Review),* ISSN 0033-2097, R. 87 NR 7/2011

[4] M.Noroozian, L. Angquist, M. ghandhari, and G.Andersson, " Use of UPFC for optimal power flow control", *IEEE Trans. Power Deliv.,* vol. 12, no. 4, pp. 1629-1634, Oct. 1997.

[5] Noroozian, M. Angquist, L. , Ghandhari, M. ; Andersson, G, "Improving power system dynamics by series-connected FACTS devices", *Power Delivery, IEEE Transactions on* (Volume:12 , Issue: 4 ), pp. 1635-1641, Oct 1997

[6] C.R. Fuietre-Esquivel, E. Acha, and H. Ambriz-Perez, "A thyristor controlled series compensator model for the power flow solution of practical power networks", *IEEE Trans. Power Syst.,* vol 15, no. 1, pp. 58-64, Feb.2000

[7] E. Acha, C.R. Fuietre-Esquivel, H. Ambriz-Perez, *FACTS modelling and simulation in power networks*, England: John Wiley & Sons Ltd, 2004, pp. 200-215.

[8] A.M. Vural and M. Tumay, " Mathematical modelling and analysis of a unified power flow controller: A comparison of two approaches in power flow studies and effect of UPFC location", *Electr. Power Energy Syst*., vol.29, pp. 617-629, 2007.

[9] Bhowmick S., Das B. and Narendra Kumar, "An indirect UPFC model to enhance reusability of Newton power-flow codes," *IEEE Trans. On Power Delivery,* vol. 23, no. 4, pp.2079-2088, Oct. 2008.

[10] Rafael Mihalic and Uros Gabrijel, "Transient stability assessment of system comprising phase-shifting FACTS devices by direct methods", *Elect. Power and Energy Syst.,* vol.26, pp.26, pp.445-453,2004.

[11] S. Gerbex, R. Cherkaoui, and A.J. Germond, "Optimal location of multi-type FACTS devices in a power system by means of genetic algorithms," *IEEE Trans. Power Syst.,* vol.16, no.3, pp. 537-544, Aug. 2001.

[12] D.J. Gotham and G.T. Heydt, "Power flow control and power flow studies for systems with FACTS devices", *IEEE Trans. Power Syst*., vol. 13, no. 1, pp. 60-65, Feb. 1998.

[13] A. Colorni, M. Dorigo et V. Maniezzo, "Distributed Optimization by Ant Colonies", *actes de la première conférence européenne sur la vie artificielle*, Paris, France, Elsevier Publishing, 134-142, 1991

[14] M. Dorigo, "Optimization, Learning and Natural Algorithms", PhD thesis, Politecnico di Milano, Italie, 1992.

[15] Y. Lu and A.Abur "Static Security enhancement via optimal utilization of thyristor-controlled series capacitors", *IEEE Trans. Power Syst.,* vol.17, no.2, pp. 324-329, May 2000.

[16] L. Ippolito and P. Sinao, "Selection of optimal number and location of thyristor-controlled phase shifters using genetic based algorithms", *Proc. Inst. Elect. Eng. Gen. Transm. Distrib*, vol.151, no. 5, pp. 630-637, Sep. 2004.

[17] S. Rahimzadeh, M. Tavakoli Bina, and A.H. Viki, "Simultaneous application of multi-type FACTS devices to the restructured environment: achieving both optimal number and location", *IET Gener. Transm. Distrib.,* vol. 4, no. 3, pp. 349-362, 2010.

[18] Tiwari, R., Niazi, K.R., Gupta, V., "Optimal Location of FACTS Devices for Improving Performance of the Power Systems", *Power and Energy Society General Meeting,* 2012 IEEE, pp. 1-8, 22-26 July 2012.

[19] Malakar, T., Sinha, N., Goswami, S.K., Saikia, L.C., "Optimal location and size determination of FACTS devices by using multiobjective optimal power flow", TENCON 2010 - 2010 *IEEE Region 10 Conference,* pp. 474-478, 21-24 Nov. 2010.

[20] S. Panda, and N.P. Padhy, "Optimal location and controller design of STATCOM for power system stability improvement using PSO", *Journal of Frank Inst.,* vol. 345, pp. 166-181, 2008.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Restricted Earth Fault Protection with Superconducting Fault Current Limiter for 100% Stator Winding

Md. Elias Khan, Apurbo Biswas, and Md.Rafiqul Islam
Khulna University of Engineering & Technology (KUET), Bangladesh
Email: rafiq043@yahoo.com, eliaskhaneee@hotmail.com, apurbo.eee2k8@gmail.com

*Abstract*— **Conventional Restricted earth fault protection provides protection about 80 to 90% of generator stator winding against earth fault when neutral of generator is earthed through a resistance. In this paper, superconducting fault current limiter an innovative device is used along with restricted earth fault protection to provide 100% stator earth fault protection. A specific generator is considered which is provided with restricted earth fault protection to protect 90% of its winding against earth fault. Then a resistive SFCL is applied in the protective scheme replacing the neutral earthed resistance to a lower value than the previous one. It is found that excellent operational characteristics of SFCL facilitate the restricted earth fault relay to operate from very beginning to the end of the generator winding and provide 100% stator earth fault protection.**

***Keywords— earth fault current, restricted earth fault protection, superconducting fault current limiter, neutral earthed resistanc.***

## I. INTRODUCTION

Generator and transformer are the most important and expensive essentials of power system. Their protection is inevitable and it must be adequate. Biased differential protection is extensively used for the over current protection of both generator and transformer. For generator, although differential protection provides a complete protection of stator winding against phase to phase fault, but doesn't in the case of phase to ground fault. The differential relay may be tripped in the absence of any earth fault [1]-[2]. This protection system is also less sensitive for earth faults close to the neutral point in an impedance earthed generator [3]-[4]. So, conventional earth fault protection using over current elements fails to provide adequate protection for generator or transformer windings. The degree of protection is very much improved by the application of restricted earth fault protection (or REFP).

Restricted earth fault protection by differential system is increasingly used for the protection of generator stator windings against earth fault, due to the outstanding selectivity and excellent sensitivity. But the operation of restricted earth fault relay is influenced by the magnitude of earth fault current. The earth fault current depends on the value of earthed impedance and position of fault in the stator winding. When earth fault occurs near to neutral, the restricted earth fault relay remains inoperative and this portion of stator winding remains unprotected. This continues up to 10-20% of the generator winding. As a result, REFP provides protection about 80-90% of generator winding against earth fault [5]. In recent years, study has been carried out for the 100% protection of generator winding against earth fault. This protection scheme is a combination of low frequency injection method and restricted earth fault protection [6]. Low frequency injection provides protection about 5-20% of generator winding and the remaining 80% winding is protected by restricted earth fault protection. But low frequency injection uses coupling transformer, low frequency generator and measuring system which makes the protection scheme more complex. 100% stator earth fault protection can be accompanied simply using superconducting fault current limiter (SFCL) along with REFP.

Superconducting fault current limiter (SFCL) is the most attractive device which operates based on the superconductor properties [7]. It has many advantages including having no effect on the system below critical current of superconductor, limiting fault current significantly and responding automatically without any external trigger. Many studies have been moved forward for the practical application of SFCL in power system protection [8]-[9].

In this paper, a particular generator is considered which gets protection about 90% of its stator winding by conventional REFP. Then a resistive type SFCL is implemented in the neutral of stator and the overall performance is studied.

## II. OPERATIONAL CHARACTERISTICS OF RESISTIVE SFCL

The operation of SFCL depends on the current flowing through it. The resistance of SFCL is zero when passing

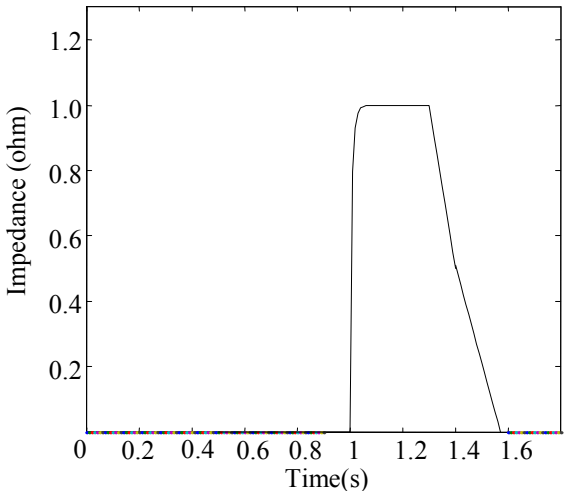


Figure 1. Quench and Recovery characteristics of SFCL

current remains below the critical current of superconductor and SFCL have no impact on the system. But when current exceeds the critical current of superconductor then the superconductor goes to resistive state and inserts resistance to the system without external triggering. This transition is known as quench. As SFCL is connected in series with the system, this additional resistance is added with the system impedance and limits fault current. Quench and recovery characteristics are designed on the basis of [10] is shown in Fig. 1. In normal condition impedance of SFCL is zero. Quenching process of SFCL starts at t=1s due to occurrence of fault and the exceeding of fault current above the critical current of superconductor causing SFCL's impedance rises to its maximum value. Impedance again becomes zero after the fault clears. More details about SFCL have been discussed in our previous paper [11].

## III. METHODOLOGY

### *A. Limitations in Designing 100% Stator Earth Fault Protection by Conventional REFP*

Restricted earth fault protection for a generator stator winding is shown in Fig. 2. The neutral is earthed through resistance to limit earth fault current. During earth fault, the fault current $I_f$ flows through a part of the generator winding and neutral to ground circuit. The corresponding secondary current of current transformer flows through the differential relay and restricted earth fault relay. The restricted earth fault relay is selected independent of the setting of the differential relay to operate in case of earth fault.

If earth fault occurs at point 'f' of generator winding $V_{af}$ is available to drive enough earth fault current $I_f$ for the operation of restricted earth fault relay though high resistance connected in neutral. If earth fault occurs at point 'a' then $V_{af}$ is small and $I_f$ is also small due to high resistance connected in neutral. In this case the restricted earth fault relay remains inoperative and some portion of generator winding remains unprotected against earth fault.

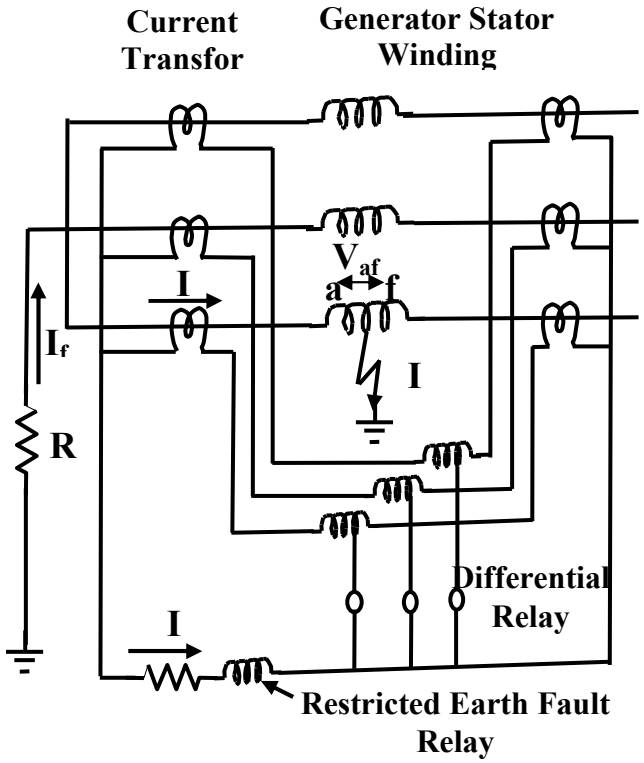


Figure 2. Restriced earth fault protection for generator

To get 100% protection from these protection scheme two conventional techniques can be applied. But they doesn't work successfully because

1. If relay current setting may keep too sensitive to sense the earth fault current of small magnitude, the relay may respond during through faults of other faults due to inaccuracies of CT's, saturation of CT's etc. This is not expected from restricted earth fault relay. And

2. If the value of earthed resistance may keep low, it will cause high fault current when earth fault occurs in the winding far from neutral. This high fault current may exceed the highest rating of the winding which causes huge damage to the winding.

So, 100% protection of generator winding by REFP is not possible by using conventional techniques. But this can be done efficiently by using SFCL along with REFP.

### *B. Design of 100% Stator Earth Fault Protection by REFP with SFCL*

Normally, REFP system with a high resistance connected in neutral can give 80-90% protection of stator winding against earth fault. The 10-20% winding remains unprotected against earth fault. This can be overcome by adding SFCL with the neutral of REFP system which is shown in Fig. 3. Now, the value of earth resistance is kept small. It is described previously; SFCL does not show any resistance below critical current of superconductor. This critical current is also known as triggering current of SFCL. So, when earth fault occurs in the stator winding near to neutral , voltage is small as well.

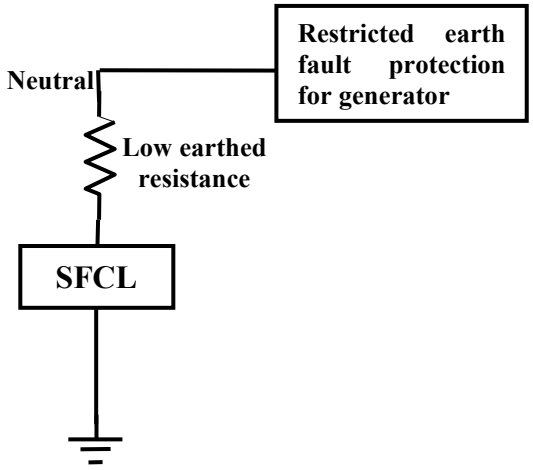



Figure 3. Restricted earth fault protection with SFCL in neutral.

But due to the low resistance connected to neutral, the fault current $I_f$ is large enough to operate the earth fault relay and the small percentage of winding get protected against earth fault. This process continues up to 10-20% of the stator winding. Fault current will be large enough beyond 10-20% percentage to damage the winding due to the small resistance. To get rid of this, SFCL starts triggering when fault current exceeds the level of the critical current of the superconductor.

Now, if a fault occurs beyond 10-20% of winding, quench of SFCL will start and it will insert a resistance to the neutral line. This resistance of SFCL is added up with the low resistance and gives a high resistance which limits the fault current to a considerable value that cannot cause any damage to the winding. This current is also sufficient to operate earth fault relay and this portion of winding get protected against earth fault.

So, 10%-20% of generator winding get protected against earth fault without the trigger of SFCL and remaining 80-90% get protected with the trigger of SFCL. It is also said that the remaining 80-90% winding get protected in the same way as done in conventional process. In this way 100% protection of stator winding of a generator against earth fault by REFP system is possible with the presence of SFCL.

*C. Practical Consideration*

Generator specifications for Conventional REFP system [5]:

Generator rating = 3phase, 11kV, 5000kVA , Full load current= 200A, Reactance of winding =15%

REFP by differential protection is set to operate on earth fault current of more than 200A.

90% protection of stator winding against earth fault can be given by a neutral resistance.

**Calculation of resistance**

Let, Resistance in neutral in ohms = R, Reactance per phase in ohms = X, Reactance of the winding at fault point = x, Voltage at fault point of the winding = v

Here, Rated current I = 200A, Phase voltage V =6350 volt Reactance of winding %X = 15%

We know,

$$\%X = \frac{X \times I}{V} \times 100 \qquad (1)$$

Fault current,

$$I_f = \frac{v}{\sqrt{R^2 + x^2}} \qquad (2)$$

From (1) we get, $X = 4.75\ \Omega$

When fault occurred at 10% of the generator winding then,

Fault current, $I_f$ =200A

$x = 4.75 \times 0.1 = 0.475\ \Omega$ , $v = 6350 \times 0.1 = 635$ volt

and from (2), $R = 3.1475\ \Omega$

So, this resistance 3.14 is necessary to give 90% protection of stator winding against earth fault by conventional REFP system.

**Calculation of low resistance**

Now, the resistance (R) is replaced by a resistance (r) which is lower than before.

From (2) we get,

Maximum possible current of stator winding = 1116 A and the low resistance, $r = 0.31\ \Omega$

**Calculation of SFCL resistance**

SFCL give a resistance so that the series combination of low resistance(r) and SFCL resistance are equal to the value of resistance (R).

$$\text{Value of SFCL résistance, } r_1 = 3.14 - 0.31$$

$$= 2.83\Omega$$

The value of fault current from very beginning to 100 % winding of the above generator is shown in Table. 1. The calculation is done using (2) at various portion of winding considering the presence of resistance(R), low resistance(r) and SFCL.

TABLE I. CURRENT IN VARIOUS PORTION OF GENERATING WINDING

| %of winding | Fault Current(A) | | |
|---|---|---|---|
| | *Considering resistance (R)* | *Considering low resistance ( r)* | *Considering SFCL operation* |
| 1 | 20.2 | 202.5 | 202.5 |
| 2 | 40.4 | 391.7 | 391.7 |
| 3 | 60.6 | 558.3 | 558.3 |
| 4 | 80.7 | 698.5 | 698.5 |
| 5 | 100.8 | 813.1 | 813.1 |
| 6 | 120.8 | 904.8 | 904.8 |
| 7 | 140.7 | 977.8 | 977.8 |
| 8 | 160.6 | 1035.8 | 1035.8 |
| 9 | 180.3 | 1082.2 | 1082.2 |
| 10 | 200 | 1116 | 200 (SFCL is triggered) |
| 11 | 219.4 | 1149.5 | 219.4 |
| 12 | 238.7 | 1174.4 | 238.7 |
| 20 | 387.1 | 1270.8 | 387.1 |
| 30 | 552.4 | 1306.3 | 552.4 |
| 40 | 692 | 1328.3 | 692 |
| 50 | 807 | 1329.3 | 807 |
| 60 | 898 | 1331.2 | 898 |
| 70 | 971.9 | 1332.1 | 971.9 |
| 80 | 1030 | 1333.2 | 1030 |
| 90 | 1077 | 1334.4 | 1077 |
| 100 | 1116 | 1336 | 1116 |

## IV. RESULT AND DISCUSSION

The specified generator winding get 90% protection against earth fault by conventional REFP system with a neutral resistance 3.14 ohm. For the specified generator, REFP by differential protection is set to operate on earth fault current of more than 200A. From Fig. 4, it is seen that fault current remains below 200A up to 10% of generator winding. So, the earth fault relay does not operate and this 10% winding remains unprotected. This occurs due to a resistance (3.14 ohm) connected in neutral.

If no resistance is connected to neutral, then the output curve is found as shown in Fig. 5. From Fig. 5 it is seen that fault current is constant at 1336A from very beginning to 100% of the winding. This current is larger than 200A which is available to operate the earth fault relay. As the current has exceeded the maximum rating 1116A this could damage the winding. So, for a solidly grounded neutral, it is possible to protect complete generator winding against earth fault. But due to the possibility of getting damaged the generator winding, this is not a reliable method.

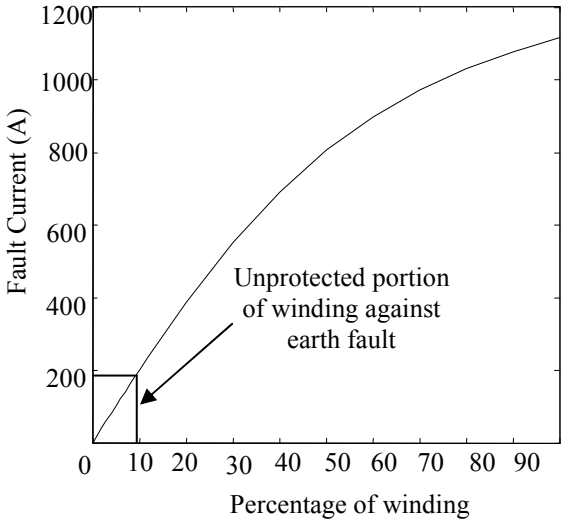



Figure 4.Output curves of 90% winding protection of a generator against earth fault by REFP system

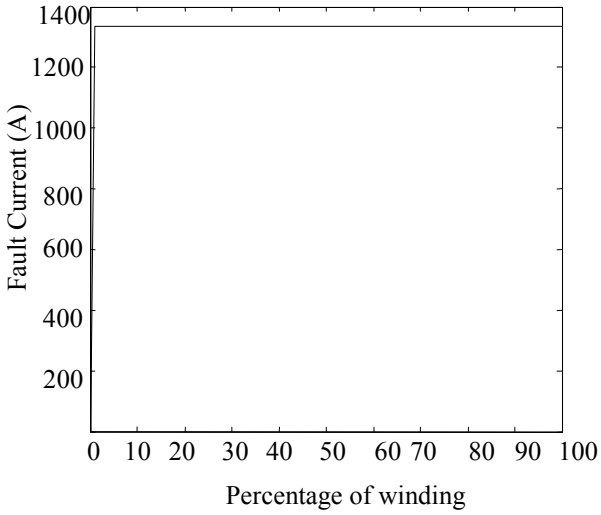



Figure 5. Output curves of a generator winding protection against earth fault by REFP system without neutral resistance

To get 100% protection of generator stator winding against earth fault reliably, SFCL is added in the neutral. Now, the neutral resistance is kept at a value of 0.31 ohm which is lower than the resistance of 3.14 ohm. This technique gives the following output curves which are shown in Fig. 6 and Fig. 7. From Fig. 6 it is seen that when earth fault occurs at the very beginning of winding, the value of fault current is 200A. So, the earth fault relay operates and protects the winding. The SFCL remains at normal operating condition up to 10% of the winding against earth fault because fault current is less than triggering value of SFCL. The triggering current of SFCL is considered just below the maximum value of winding current. When earth fault occurs beyond 10% of the winding, the fault current exceeds the maximum current of the winding. SFCL starts self-triggering and limits the fault current to a considerable value which is sufficient to operate the earth fault relay.

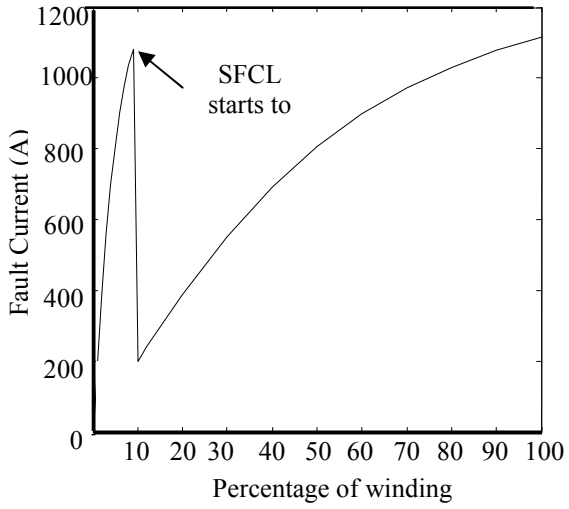



Figure 6. Output curves of 100% winding protection of a generator against earth fault by REFP system using SFCL.

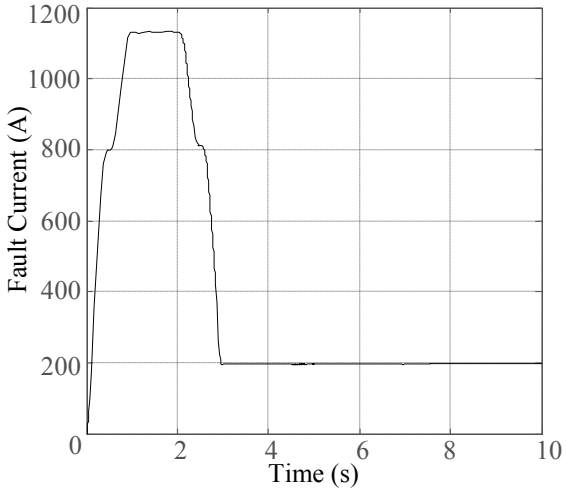



Figure 7. Fault current limitations by SFCL at 10% of the generator winding.

For this instance, the fault current beyond 10% of the winding is same as fault current shown in Fig. 5. This has been possible because SFCL inserts resistance of 2.83 ohm which is added up with the low resistance and makes the resistance equal to 3.14 ohm which is same as the previous resistance.

The current limiting behavior of SFCL at the 10% of the winding is shown in Fig. 7. The current reaches to 1116A which is the maximum value of current of the winding and limited by SFCL to 200A. In the same way, the fault current limiting behavior of SFCL can be shown anywhere beyond 10% of the winding.

Thus, it is said that up to 10% of generator winding get protected against earth fault without the trigger of SFCL and remaining 90% get protected with the trigger of SFCL. In this way the 100% winding of generator get protected against earth fault by REFP system using SFCL.

## V. CONCLUSION

In this paper, 90% stator winding protection against earth fault of a specific generator given by conventional restricted earth fault protection is upgraded to 100% using superconducting fault current limiter along with REFP. It is clear from the results that the exceptional operational characteristics of SFCL enable REFP to provide 100% stator earth fault protection. The value of neutral earthed resistance should be kept low in this concern. This work will create a new era in the generator and transformer winding protection. Coordination between the SFCL and protective device is a major concern for the practical implementation of this work.

## REFERENCES

[1] D. Jones, M.Sc., C. Eng., M.I.E.E., *Analysis and Protection of Electrical Power Systems*, Pitman Publishing, 1971.

[2] Denis Robertson, "*Power system protection", reference manual, reyolle protection*, first published in 1982 by Oriel Press Ltd., pp. 104.

[3] GEC Alsthom Measurements Limited, *Protective Relays Application Guide*, 3rd edition, 1990.

[4] D. Robertson, ed. Power System Protection Reference Manual, Reyrolle Protection, Chapter 6, Stockfield: Oriel Press, 1982.

[5] Sunil S.Rao, *"Protection of generators," in Switchgear protection and power system*, 12th ed. Delhi-110006, India, Khanna Publishers, 2011-2012, ch.33, sec. 33.4, pp. 623-627.

[6] U.A Bakshi, M.V Bakshi, *"Generator Protection," in Protection and Switchgear*, 4th ed. Pune-411030, India, Technical Publications Pune, 2009, ch.5, sec. 5.9, pp. 5-18

[7] Steven M. Blair, Campbell D. Booth, Nand K. Singh, and Graeme M. Burt, "Analysis of Energy Dissipation in Resistive Superconducting Fault-Current Limiters for Optimal Power System Performance," *IEEE Trans. Appl. Supercond*, vol. 21, no. 4, Aug. 2011.

[8] Jin-Seok Kim, Sung-Hun Lim, and Jae-Chul Kim, "Study on Application Method of Superconducting Fault Current Limiter for Protection Coordination of Protective Devices in a Power Distribution System," *IEEE Trans. Appl. Supercond*, vol. 22, no. 3, Jun. 2012.

[9] B. C. Sung, D. K. Park, J.-W. Park, and T. K. Ko, "Study on optimal location of a resistive SFCL applied to an electric power grid," *IEEETrans. Appl. Supercond*, vol. 19, no. 3, pp. 2048–2052, Jun. 2009.

[10] Jong-Fil Moon, Sung-Hun Lim, Jae-Chul Kim, and Sang-Yun Yun, "Assessment of the Impact of SFCL on Voltage Sags in Power Distribution System," *IEEE Trans. Appl. Supercond*, vol. 21, no. 3, Jun. 2011.

[11] Apurbo Biswas, Md. Elias Khan and Md. Rafiqul Islam, "Enhancement of Power System Capacity with Existing Switchgear using Superconducting Fault Current Limiter," *2nd International Conference on Informatics, Electronics & Vision (ICIEV-2013)*, Dhaka, Bangladesh, May 17-18, 2013.

.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Designing a Mobile Satellite Communication Antenna and Link Budget Optimization

Anindya Kumar Kundu, MD. Tofael Hossain Khan*, Wahida Sharmin, Md. Osman Goni, Kazi Abul Barket[2]
Dept. of Electronics & Communication Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh.
[2] Physics Discipline, Khulna University, Khulna-9208, Bangladesh.
*tofael18@gmail.com

*Abstract*—**With the rapid growth of modern mobile satellite communication technology, the development of very small size, low-cost, low-profile, high gain and high directivity antennas is a must. The design of a mobile satellite communication antenna named as Parabolic Helix and link budget optimization of the proposed antenna system are illustrated in this paper. The proposed antenna has a gain of 15.95dB, directivity of 19.36dBi; return loss value of -28.0573dB, Voltage Standing Wave Ratio (VSWR) of 1.08, bandwidth of 68.3MHz, antenna efficiency of 45.6% and 3dB angular beamwidth i.e., the Half Power Beamwidth (HPBW) of 29.5deg. The resonant frequency of the antenna array is 5.982 GHz. The proposed antenna system can be used for C-band applications like satellite communications transmissions, Wi-Fi, cordless telephones, weather radar systems and other wireless systems. The antenna system is designed and simulated in the CST Microwave Studio. Link budgets are performed in order to analyze the critical factors in the transmission chain and to optimize the performance characteristics, such as transmission power, system noise temperature, figure of merit and so on. The link budget determines what size antenna is to use, power requirements and in general, the overall customer satisfaction. This paper deals with the rudiments of a satellite link design with some simulation results.**

***Keywords—C-band, gain, link budget analysis, transmission power, satellite communications.***

## I. INTRODUCTION

With the drastic development of modern mobile satellite communication technology, the use of mobile antennas have increased due to their very small size, low-cost, low-profile, high gain and high directivity. Antennas for mobile satellite communications are widely presented in books and papers in the last decade as presented in [1-3]. Global Mobile Satellite Communications: For Maritime, Land and Aeronautical Applications are illustrated in [4]. A 12 GHz Planar Array Antenna for Satellite Communication is studied in [5]. Satellite microwave transmission systems are similar to terrestrial microwave systems except that the signal travels from a ground station on Earth to a satellite and back to another ground station on Earth, thus achieving much greater distances than Earth-bound line-of-sight transmissions. Since three equidistant satellites in the geosynchronous orbit (having 120 degrees apart) can effectively cover over almost the entire earth surface with some overlapping except for the polar region, the need for multiple retransmissions is removed [6]. The geostationary satellite can cover almost 38% of the surface of the earth with a 5 degrees minimum elevation angle of the earth station antenna [7]. The factors required for determining the quality of performance can be divided into two broad categories; the conduit factors and the content factors. The conduit factors include effect on signal propagation, quality of earth station equipment, uplink and downlink etc. On the other hand, for suitability of transmission over a microwave medium, the content factors deal with the transmitted message type and the devices involved in its transformation [8]. G/T (the ratio of antenna gain to system noise temperature or a figure of merit), effective isotropic radiated power (EIRP) and C/N (the ratio of carrier power to noise power density) are the parameters that characterize the performance of the three segments- namely, earth station, gateway and satellite. In satellite communication, two usually used concepts are G/T and EIRP and they indicate the transmitting and receiving capabilities, respectively, of satellite, a gateway earth station and a mobile terminal. The quality of the communication channel is insured by the C/N ratio [9]. Points on earth beyond about 80 degrees latitude are not visible which the disadvantage of a geostationary satellite. On the other hand visibility to the higher northern and southern latitudes can be provided by the inclined orbits. This often necessitates an acquisition operation and sometimes involves handover from an orbiting satellite leaving the area to a new satellite entering the area. In addition, in order to provide continuous coverage inclined orbits usually require multiple satellites to be spaced along the orbit [10].

## II. ANTENNA ARCHITECTURE

The Copper (annealed) lossy metal is used as substrate for the Helix of the proposed antenna and the relative permeability ($\mu_r$) and electrical conductivity of the substrate material are 1.0 and 5.8e+007 (S/m) respectively. The helix is 30 mm long and 10 mm of diameter. The diameter of the wire used for building helix is 1mm. Aluminum is used as substrate for the reflector of the proposed antenna and the relative permeability ($\mu_r$) and electrical conductivity of the substrate material are 1.0 and 3.56e+007 (S/m) respectively. The reflector diameter is 150 mm with 0.2 mm thickness and focal

distance of 75 mm. Base of the helix antenna is kept on the reflector surface without any electrical conduction. The excitation signal is applied below the base to a coaxial cable connected with the helix through the base. Fig. 1 shows the structure of the proposed Parabolic-Helix antenna-

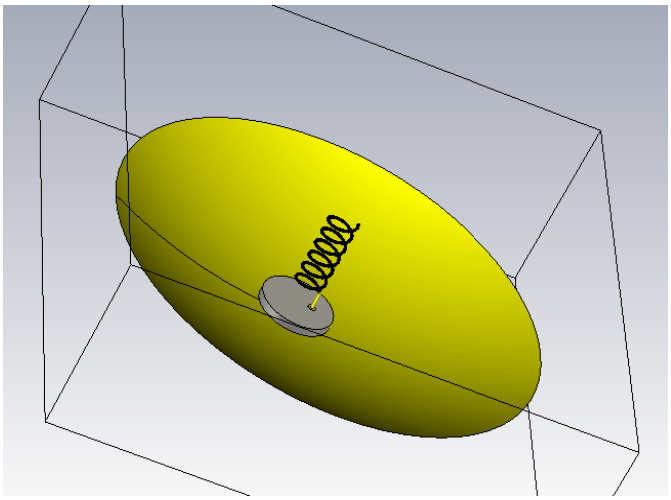


Figure 1. Structure of the proposed Parabolic-Helix antenna.

## III. SIMULATION RESULTS

### *A. Electric Field Distribution*

Electric field distribution of the Parabolic–Helix antenna is given below –

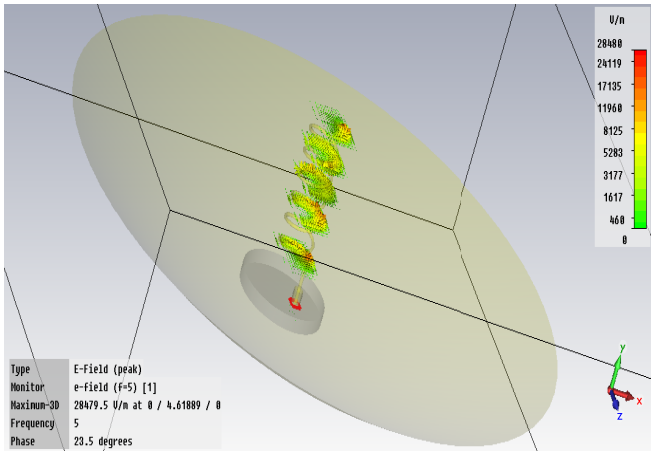



Figure 2. Electric field distribution of Parabolic-Helix antenna.

### *B. Voltage Standing Wave Ratio (VSWR) and Return Loss*

Fig. 3 shows the graph of VSWR vs. Frequency. From this figure it can be seen that the VSWR value is very near to the unity which is mostly expected. The VSWR value of the antenna is 1.08235 at 5.982 GHz.

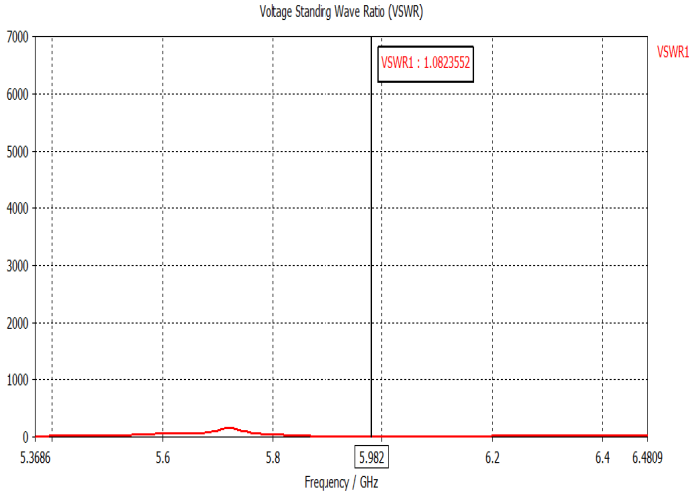



Figure 3. VSWR vs. Frequency curve of the parabolic-helix antenna.

Fig. 4 shows the graph of return loss vs. frequency. From this figure we can observe the bandwidth of the antenna. The bandwidth of the antenna is 68.3 MHz which operates in C band (5.9475 to 6.0158 GHz) at 5.982 GHz. The value of return loss is -28.057276 dB at 5.982 GHz.

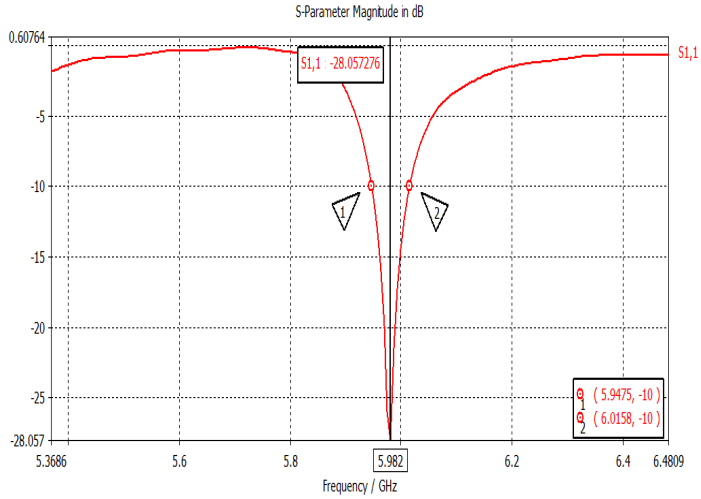



Figure 4. Return loss of proposed Parabolic-Helix antenna.

### *C. Farfield Radiation Pattern (3D), Gain Pattern and Antenna Efficiency*

Fig. 5 shows the farfield radiation pattern of the Parabolic-Helix antenna. The radiation pattern of the antenna is very much pointed which is one of the most important requirements for satellite communication antennas for ground base earth station. From this figure it can be seen that the directivity and gain of the antenna is about 19.36dBi and 15.95dB respectively. The radiation efficiency is 0.8239dB.

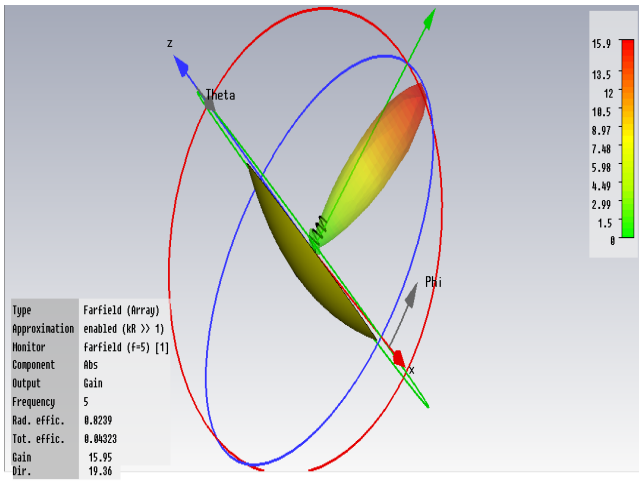



Figure 5. Farfield radiation pattern.

Fig. 6 shows the farfield gain pattern of the antenna. It can be seen that there is no side lobe and only one main lobe in the gain pattern. The main lobe magnitude is 18.8 dB, main lobe direction is 98.0 deg., and 3 dB angular beamwidth is 29.5 deg. and side lobe level is -11.4 dB.

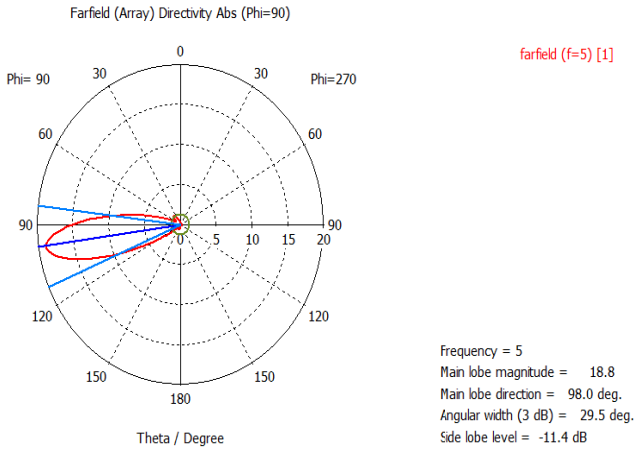



Figure 6. Farfield gain pattern.

## IV. Result Analysis

VSWR, return loss, farfield radiation pattern, antenna gain, directivity and antenna efficiency of the proposed antenna system show reasonable characteristics. The performance of the antenna is quite good. The analysis has been done depending on gain. This antenna can be used in the C-band applications for its effective performance.

TABLE I. Proposed Parabolic-Helix Antenna Parameters and Their Values at A Glance

| **Designed Parabolic Helix Antenna Parameters** | **Simulation Results** |
|---|---|
| VSWR | 1.08235 |
| Return Loss (in dB) | -28.0573 dB |
| Gain (in dB) | 15.95 dB |
| Directivity (in dBi) | 19.36 dBi |
| Half Power Beamwidth | 29.5 deg. |
| Antenna Efficiency (%) | 45.6% |
| Bandwidth (MHz) | C-band → 68.3 MHz |

The results in the Table I reveal that the proposed Parabolic-Helix antenna system is useful for C-band applications like satellite communications transmissions, Wi-Fi, cordless telephones, weather radar systems and other wireless systems.

## V. Link Budget Analysis

The satellite link is much like the terrestrial microwave radio relay link with the advantage of not requiring as many re-transmitters as are required in the terrestrial link. Transmission of signals over a satellite communication link requires Line-of-Sight (LoS) communication. Link analysis basically relates the transmit power and the receive power. Basic transmission parameters are- flux density, received power, antenna gain, noise power, figure of merit etc.

A link consists of three parts namely transmitter, receiver and media. The two main items that are associated with transmitters are flux density and EIRP. A measure of the amount of energy that is received at a distance $r$ from a transmitter of gain $G_t$ and transmitter power $P_t$ watts is the flux density which is given by-

$$\phi = \frac{P_t G_t}{4\pi R^2} \ [W/m^2] \qquad (1)$$

where, $P_tG_t$ is called the Effective Isotropic Radiated Power or EIRP which is closely associated with a radiating source or a transmitter and is a subset of flux density.

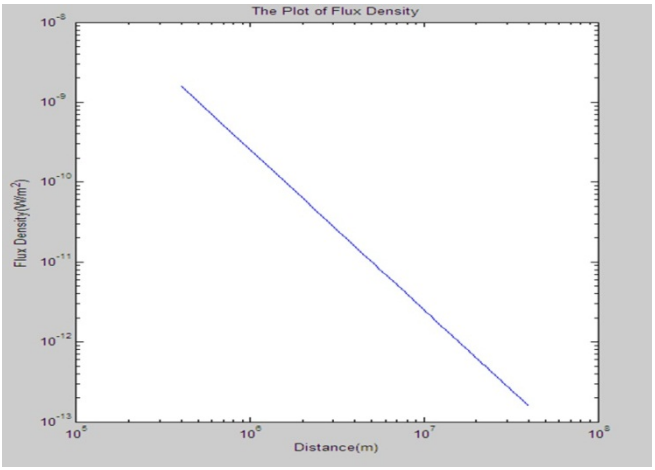



Figure 7. Graphical representation of flux density.

The above figure shows that flux density decreases as square of the distance. The plot has been made between 400 to 40000 km for distance and the corresponding flux density lies between near about $10^{-9}$ to $10^{-13}$ W/ $m^2$ . The constant parameters are Pt = 20 W, Gt = 22 dB.

For an ideal receiver antenna of aperture area A, the total received power at the receiver is given by-

$$P_r = \phi \times A = \frac{P_t G_t A}{4\pi R^2} \text{ [W]} \quad (2)$$

A practical antenna with physical aperture A will not deliver this power as some energy will be reflected and some will be absorbed by lossy elements. Thus the actual power received will be-

$$P_r = \eta A \times \varphi \text{ [W]} \quad (3)$$

where, □ is the antenna efficiency and □A is referred to as the effective collecting area of the antenna. The antenna efficiency □ accounts for all losses between the incident wavefront and the antenna output port.

An antenna of maximum gain Gr is related to its effective area by the following equation-

$$G_r = \eta \frac{4\pi A}{\lambda^2} \quad (4)$$

where, ג is the wavelength of the received signal.

Rearranging equation (4) and substituting in (3) we get-

$$P_r = \frac{P_t G_t G_r}{(4\pi R/\lambda)^2} \text{ [W]} \quad (5)$$

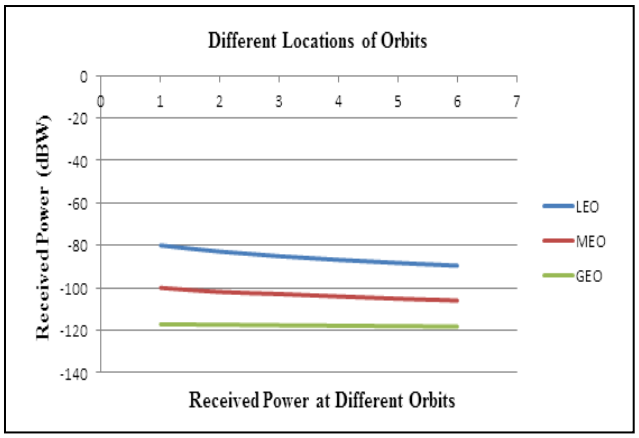



Figure 8. Received power at different orbits.

From the above graph it is seen that received power remains almost constant for GEO due to its longer distance which is above 35786 km above the earth surface. Here, the ranges of different orbits are taken as- LEO→500-1500km , MEO→5000-10000km, GEO→36000-41000km while the values of constant parameters of equation (5) are $P_t = 20$ W, $G_t = 22$dB, $G_r = 52.3$ dB, f = 11 GHz.

For a parabolic antenna of diameter D, equation (4) can be rewritten as-

$$G_r = \eta \left(\frac{\pi^2 D^2}{\lambda^2}\right) \quad (6)$$

The variation in antenna gain for a range of transmission frequencies that are employed in satellite communications is shown below assuming an efficiency of 60%.

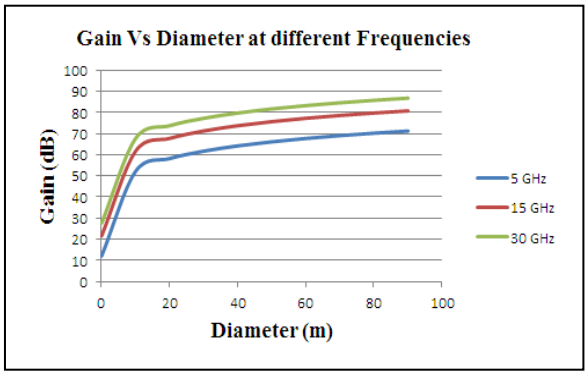



Figure 9. Illustration of gain Vs diameter at different frequencies.

The above figure resembles that highest frequency shows maximum gain while the lowest one illustrates the lowest gain as the gain is directly proportional to the square of the frequency.

## VI. SYSTEM NOISE

### A. *Noise Power*

The thermal noise power Pn delivered to the optimum load by the thermal noise source of resistance R at temperature T is given by -

$$P_n = kT_pB_n \quad (7)$$

where, k – Boltzmann constant = 1.38 x$10^{-23}$ J/K = – 228.6 dBW/K/Hz; $T_p$– Noise temperature in Kelvin; $B_n$ – Noise bandwidth in Hz.

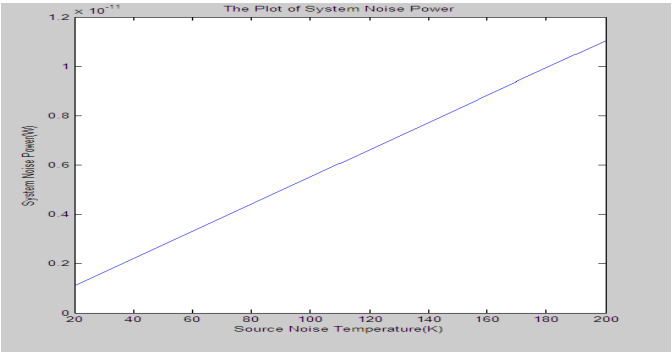



Figure 10. Demonstration of system noise power.

Since the system noise power is directly proportional to the source noise temperature, so the noise power will increase as the temperature increases. Here, $B_n$ = 4 GHz.

### B. *Figure of merit (G/T)*

The figure of merit (G/T) has been introduced to describe the capability of an earth station or a satellite to receive a signal. Since the C/N ratio is the ratio of signal power to noise power, we have that-

$$C/N = \frac{P_r}{P_n} = \frac{P_t G_t G_r}{(4\pi R/\lambda)^2} / (kT_sB_s) \quad (8)$$

$$\text{i.e., } C/N = f(G_r / T_s)$$

The ratio $G_r/T_s$ (or simply G/T) is known as the Figure of Merit which indicates the quality of a receiving satellite earth system and it is measured in [dB/K].

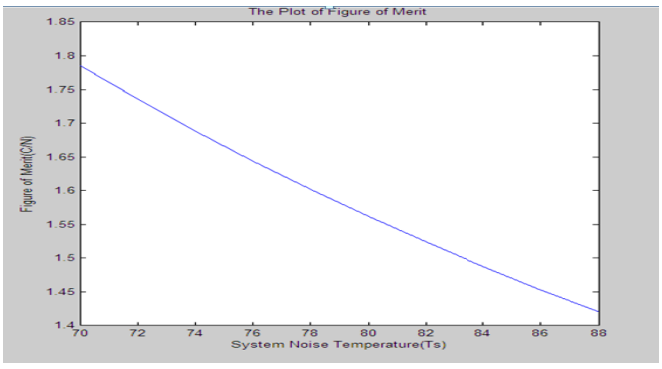



Figure 11. Realization of figure of merit.

Since the figure of merit is inversely proportional to the system noise temperature, so the figure of merit will decrease as the system noise temperature increases. Here, $P_t$ = 20 W, $G_t$ = 22 dB, R = 39,000 km, f = 4.15 GHz, B = 4 GHz, $G_r$ = 10,000-100,000.

## VII. LOSSES EXPETIENCED IN LOS LINKS

### A. *Free Space Path Loss*

The term $(4\pi R/\lambda)^2$ is called the free space path loss.

The link equation expressed in equation (5) may be read as

$$\text{Power received} = \frac{\text{EIRP} \times \text{Receive antenna gain}}{\text{Path loss}} \text{ [W]} \quad (9)$$

Expressing the above equation in terms of dB we get-

$$P_r = \text{EIRP} + Gr - L_p \text{ [dBW]} \quad (10)$$

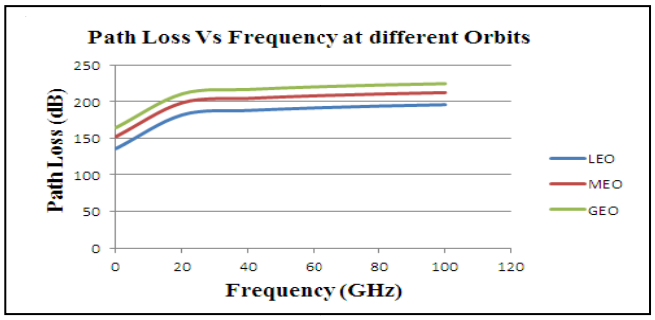



Figure 12. Representation of path loss Vs frequency at different orbits.

The above figure resembles that highest path loss is obtained in case of GEO and the lowest path loss is achieved in LEO.

### B. *Rain Loss*

Signal attenuation due to rain is the second most significant after free space loss. It is particularly significant for frequencies in the Ku and Ka bands. It becomes severe at above 10 GHz. Rain loss can be calculated from the chart of rain rate as shown below-

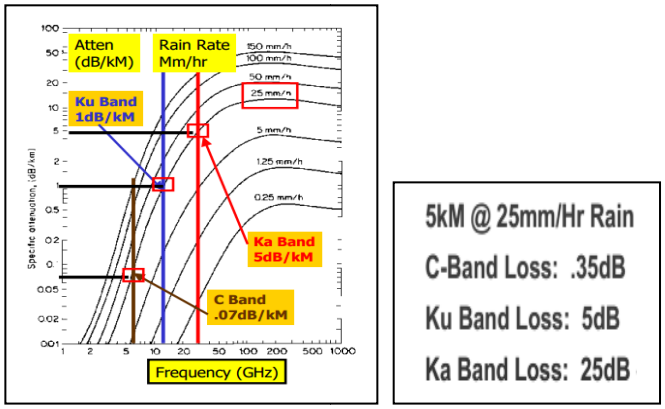



Figure 13. Calculation of rain loss from rain rate. [11]

## VIII. LINK BUDGET

The results obtained from the link budget calculator are shown below [12]-

### A. Uplink Budget

| | |
|---|---|
| Uplink frequency GHz | 6 |
| Uplink antenna diameter m | .15 |
| Uplink antenna aperture efficiency e.g. 0.65 | .456 |
| Uplink antenna transmit gain dBi | 16.07506 |
| Uplink antenna, power at the feed W | 850 |
| Uplink EIRP dBW | 45.36925 |
| Range (35778 - 41679) km | 38500.0 |
| Uplink path loss dB | 199.7222 |
| Uplink pfd at satellite dBW/m^2 | -117.340 |
| Bandwidth Hz | 68300000 |
| Satellite uplink G/T dB/K | 40.938 |
| Uplink C/N dB | 36.84080 |

### B. Downlink Budget

| | |
|---|---|
| Downlink frequency GHz | 4 |
| Downlink receive antenna diameter m | .15 |
| Downlink receive antenna aperture efficiency e.g. 0.65 | .456 |
| Downlink system noise temperature (antenna+LNA) K | 120 |
| Downlink receive antenna gain dBi | 12.55323 |
| Downlink receive antenna G/T dB/K | -8.23857 |
| Downlink satellite EIRP dBW | 35.5 |
| Downlink path loss dB | 196.2004 |
| Downlink C/N dB | -18.6831 |

The basic assumptions of this calculator are uplink antenna power at the feed, range, satellite uplink G/T, downlink system noise temperature and downlink satellite EIRP. The results obtained from the link budget calculator shows that our proposed antenna system is realizable by showing similarity between the gain (the main reasoning factor) obtained by simulation and the gain obtained by the calculator which is approximately equal to 16 dB. Simulation results are obtained from the CST Microwave Studio software while we get the calculated values from a database specified for standard C using our proposed antenna dimension. The white colored values are the input values of this calculator.

## IX. COST CALCULATION

Large antennas are expensive to construct and install, with costs exceeding $1M for 30m diameter fully steerable antennas [12]. The cost of large fully steerable antennas has been quoted as [13]-

$$\text{Cost} = \$\ y(D)^{2.7} \tag{11}$$

where, D is the diameter of the antenna aperture in feet. The constant y in equation (11) depends on the currency used and inflation, but might typically be around five U.S. dollars in the early 1980s.

The diameter of the proposed Parabolic Helix antenna is 15cm i.e.; 0.492125984 feet which in turn gives-

$$\text{Cost} = \$\ y(D)^{2.7} = \$\ \{5 \times (0.492125984)^{2.7}\} = \$\ 0.74$$

## X. CONCLUSION

The design of a Parabolic-Helix antenna and its performance analysis has been demonstrated in this paper. The simulation results of the proposed Parabolic Helix antenna resemble very good performance. The results obtained from the link budget calculator shows that our proposed antenna systems are realizable by showing similarity between the gain obtained by simulation and the gain obtained by the calculator. Also the cost calculation reveals that our proposed antenna system is cost effective. There are a lot of factors that should be kept in consideration for designing a satellite link. This paper has also illustrated the most important factors among all and has shown their interrelation by plotting curves. If one considers all the discussed factors for link budget design and his/her system parameter curves follow the characteristics of the curves given in this paper, the designed link will be a robust one for satellite communication.

## XI. FUTURE WORK

We will try to design an Ultra-Wide band mobile satellite communication antenna system and to develop practical infrastructure of the proposed antenna system.

## REFERENCES


[1] Basari, "Development of simple antenna system for land mobile satellite communications," Doctoral Dissertation, Feb. 2011.

[2] S. Ohmori, H. Wakana, and S. Kawase, "Mobile satellite communications," *Artech House*, Norwood, MA, USA, 1998.

[3] J.H. Lodge, "Mobile satellite communication systems: toward global personal communications," *IEEE Commun. Mag.*, vol. 29, no. 11, pp. 24–30, Nov. 1991.

[4] S.D. Ilcev, "Global mobile satellite communications: For Maritime, Land and Aeronautical Applications," *Springer*, Dordrecht, the Netherlands, 2005.

[5] Adel M. Abdin, "12 GHz Planar Array Antenna for Satellite Communication," *PIERS Online*, Vol. 4, No. 4, 2008.

[6] Monojit Mitra, *Satellite Communication*, ISBN: 978-81-203-2786-3.

[7] Tri T. Ha., *Digital Satellite Communications*, 2nd edition, McGraw Hill, Copyright©1990, ISBN-13: 978-0-07-007752-2.

[8] Aderemi A. Atayero, Matthew K. Luka and Adeyemi A. Alatishe, "Satellite link design: A Tutorial," *International Journal of Electrical and Computer Sciences*, Vol: 11 No: 04, Aug.2011.

[9] Pratt et al., *Satellite Communications*, ISBN: 9971-513-22-6.

[10] Wikipedia English (2010). Geosynchronous Satellite. [Online]. Available: http://en.wikipedia.org/wiki/Geosynchronous_satellite.

[11] Howard Hausman, "Fundamentals of Satellite Communications, Part-2," Hauppauge, NY 11788, Jan. 22, 2009.

[12] Satellite Link Budget Calculator. [Online]. Available: http://www.satsig.net/ linkbugt.htm/

[13] N.E. Feldman, "The link from a communications satellite to a small ground terminal," *Microwave journal*, 7, 39-44, 1964.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Handover Priority Based on Adaptive Channel Reservation in Wireless Networks

Tahsin Ahmed Chowdhury, Rahul Bhattacharjee, and Mostafa Zaman Chowdhury
Department of Electrical and Electronic Engineering
Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh
Email: tawrinkle@gmail.com, rahul0803036@gmail.com, mzceee@gmail.com

*Abstract*— **The process, handover was introduced to facilitate users' mobility and quality of service to the users in wireless networks. Forced termination of a continuing call, which is undesirable to the users, is the result of handover failure. So, it is a requirement to keep the handover request in the higher priority than the new originating call. Several schemes were introduced to give priority to the handover calls which include fixed guard channel reservation scheme, handover queuing scheme, channel transferred scheme, etc. where simultaneous minimization of handover call dropping probability and new call blocking probability did not come out with successful results. Again, a non-priority scheme results in higher handover call dropping probability with the highest channel utilization. In our proposed idea, at first we study about the different schemes for handover call priority, analyze their performances in wireless networks and discuss their merits and demerits. Then, we propose our scheme to overcome the problems arose in the previous schemes. In our proposed scheme, we combine both the features of minimizing handover call dropping probability and new call blocking probability. This results in improved channel utilization keeping the handover call dropping probability within an acceptable range based on variable guard channels which has been made adaptive on the basis of channel occupancy.**

**Keywords—Handover, adaptive channel reservation, handover call dropping probability, new call blocking probability, channel utilization.**

## I. INTRODUCTION

The coverage area of a wireless network is divided into some small service areas, called cells, where each cell is covered by a base station (BS), serving a number of mobile stations (MSs) [1]. Handover is the process of changing channels from one cell to another cell while a call is in progress [2]-[4]. The process of handover depends on some parameters such as: cell size, incoming or outgoing traffic intensity, users' mobility, and the direction to that mobility. As, the resources are limited and demand of the services is high, the process of channel reuse has been accomplished with increased capacity and decrease in cell size, from macrocells, to microcells, to femtocells and to picocells [5], [6]. Reduction in cell size along with users' motion result in frequent handovers in wireless communication systems.

According to the users' point of view, forced termination of continuing calls due to handover failure is more aggravating than the blocking of new calls. Various schemes were proposed earlier for givving priority to handover calls which can be categorized into fixed guard channel reservation Scheme (CRS) [1], handover queuing scheme (HQS) [1], [3], channel transferred scheme [1], [7], etc. In CRS [8], small portion of total allocated channels are reserved (known as guard channel) for prioritization of handover calls. HQS allows either the handover to be queued or both the originating calls and handover requests to be queued [9], [10]. The HQS schemes [9] give priority to handover attempts by permitting them to be queued, instead of denying the access in the prospective new BS if it is busy. The vital issue of channel transferred scheme is that, if there are no available channels to accommodate a handover call request, a channel from a neighboring cell may be transferred [1]. Again, a non-priority scheme was also introduced earlier [3], which gives no priority to any type of calls. An adaptive bandwidth allocation scheme [5] was proposed where the bandwidth was made adaptive to give priority to handover calls over new calls. Several metrics such as: new call blocking probability, handover call dropping probability, handover probability, call dropping or forced termination or interrupted call probability, delay, and channel utilization are used to evaluate the performance of a handover scheme [1], [7]. New call blocking probability, handover call dropping probability, and channel utilization are the prime concerns of the research field and in this paper, we focus on these three performance metrics.

Non-priority scheme shows better channel utilization but higher handover call dropping probability. Priority scheme (fixed guard channel reservation scheme) gives poor channel utilization, higher new call blocking probability but reduced handover call dropping probability. The prime concerns of our paper are: the performance analysis of the non-priority scheme, fixed guard channel reservation scheme and to overcome the drawbacks of these schemes with our proposed adaptive channel reservation scheme for handover call priority, which reduces the handover call dropping probability within an acceptable range, reduces new call blocking probability and improves the channel utilization, so that the overall performance of the proposed scheme in wireless networks is optimized. The guard channels are made adaptive in our scheme based on channel occupancy. The concept of variable guard channel has been introduced in our proposed scheme to make an optimized use of channels which remain unused when handover call arrival rate is low for fixed guard channel reservation scheme. So, in our proposed scheme we can make better utilization of our resources and

simultaneously, giving priority to the handover calls, which were not possible at a time both in non-priority scheme and fixed guard channel reservation scheme.

The rest of the paper is organized as follows. Section II represents our proposed idea with mathematical modeling and queuing analysis. The performance analysis of the scheme and the comparison with the other schemes have been shown in section III. Finally, section IV gives a conclusion to our study.

## II. ADAPTIVE CHANNEL RESERVATION SCHEME

In our proposal, adaptive channel reservation for handover priority has been introduced based on fixed guard channel reservation scheme (or CRS), with the additional feature of variable guard channels; so that, when one or more guard channels are not in used or released, it can be used for new call (or both for new call and handover call) instead of being reserved for handover calls only. This has the benefit of both minimized handover call dropping and new call blocking probability along with better channel utilization.

We have introduced a new factor by which the number of reserved guard channels is varied. The factor is based on the percentage of channels used in the system. The increasing quantity of channel usage means more traffics are accepted into the cell and so, possibility of more handover calls in the system. Thus, in this strategy the factor varies depending on the number of channels occupied in the system.

Let, the total number of channels be $S$, number of priority-based fixed guard channels be $g$, new (originating) call arrival rate be $\lambda_n$, handover call arrival rate be $\lambda_h$, average call life time be $1/\mu_c$, and average dwell time be $1/\eta$.

There is a relation among channel holding time, dwell time and call life time, which can be expressed as [11]:

$$\frac{1}{\mu} = \frac{1}{\mu_c + \eta} \tag{1}$$

Probability of handover calls is also related to dwell time and call life time [11] by:

$$P_h = \frac{\eta}{\eta + \mu_c} \tag{2}$$

We find the relation between the handover call arrival rate and the new call arrival rate [5], [11] as:

$$\lambda_h = \frac{P_h(1-P_B)}{1-P_h(1-P_D)}\lambda_n \tag{3}$$

Fixed channels both for new calls and handover calls for fixed guard channel reservation scheme is:

$$S_C = S - g \tag{4}$$

Suppose,

$$U = \frac{channels\ occupied}{Total\ number\ of\ channels} \tag{5}$$

Now, we propose a new factor $x$ where,

$$x = \begin{cases} U, & U > x_{\min} \\ x_{\min}, & U \le x_{\min} \end{cases} \tag{6}$$

Here, $x_{min}$ is the minimum value of $x$. Now, the number of guard channels in our proposed scheme is $gx$ instead of $g$ which was fixed in fixed guard channel reservation scheme. The range of the factor $x$: $x_{min} \le x \le 1$.

Now, channels available both for new calls and handover calls are:

$$S_{Cn} = S - gx \tag{7}$$

As $x \le 1$, $S_{Cn} \ge S_C$.

Minimum guard channels for handover calls is,

$$C_h = gx_{\min} \tag{8}$$

The maximum value of $S_{Cn}$,

$$\begin{aligned} S_h &= S - C_h \\ &= S - gx_{\min} \end{aligned} \tag{9}$$

Thus, we can set a limit for $S_{Cn}$ by: $S_C \le S_{Cn} \le S_h$; which means, the number of available channels both for new calls and handover calls is not fixed at $S_C$, rather it can be varied up to $S_h$ in our proposed scheme.

The relevant state transition diagram is shown in Fig. 1, which has been modeled as an M/M/S/S queuing system.

The steady-state probability $P_i$ is easily found as follows:

$$P_i = \begin{cases} \dfrac{(\lambda_n+\lambda_h)^i}{i!\mu^i} P_0, & 0 \le i \le S_C \\ \dfrac{(\lambda_n+\lambda_h)^{S_C}(\lambda_n+\lambda_h)^{i-S_C}}{i!\mu^i} P_0, & S < i \le S_{Cn} \\ \dfrac{(\lambda_n+\lambda_h)^{S_C}(\lambda_n+\lambda_h)^{S_{Cn}-S_C}\lambda_h^{i-S_{Cn}}}{i!\mu^i} P_0, & S_{Cn} < i \le S_h \\ \dfrac{(\lambda_n+\lambda_h)^{S_C}(\lambda_n+\lambda_h)^{S_{Cn}-S_C}\lambda_h^{S_h-S_{Cn}}\lambda_h^{i-S_h}}{i!\mu^i} P_0, & S_h < i \le S \end{cases} \tag{10}$$

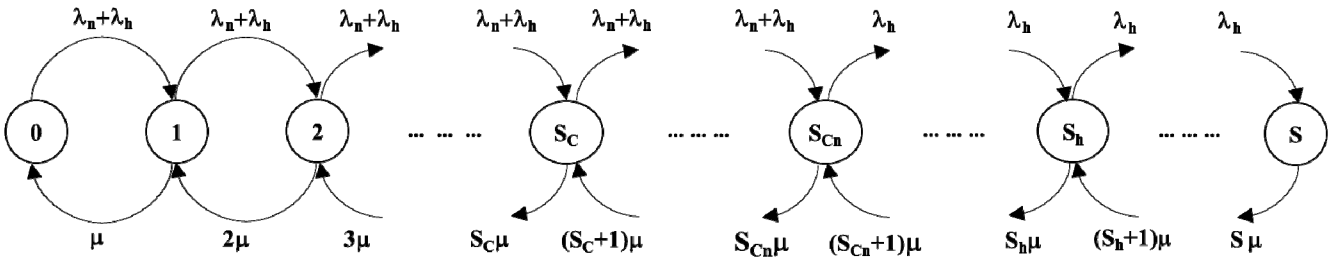



**Fig. 1:** State transition diagram for proposed adaptive channel reservation scheme

$$P_0 = \left[ \begin{array}{l} \sum_{i=0}^{S_c} \frac{(\lambda_n+\lambda_h)^i}{i!\mu^i} + \sum_{i=S_c+1}^{S_{cn}} \frac{(\lambda_n+\lambda_h)^{S_c}(\lambda_n+\lambda_h)^{i-S_c}}{i!\mu^i} \\ + \sum_{i=S_{cn}+1}^{S_h} \frac{(\lambda_n+\lambda_h)^{S_c}(\lambda_n+\lambda_h)^{S_{cn}-S_c}\lambda_h^{\,i-S_{cn}}}{i!\mu^i} \\ + \sum_{i=S_h+1}^{S} \frac{(\lambda_n+\lambda_h)^{S_c}(\lambda_n+\lambda_h)^{S_{cn}-S_c}\lambda_h^{\,S_h-S_{cn}}\lambda_h^{\,i-S_h}}{i!\mu^i} \end{array} \right]^{-1} \quad (11)$$

The blocking probability $P_B$ for an originating new call is:

$$P_B = \sum_{i=S_{cn}}^{S} P_i = \sum_{i=S_{cn}}^{S} \frac{(\lambda_n+\lambda_h)^{S_c}(\lambda_n+\lambda_h)^{S_{cn}-S_c}\lambda_h^{\,i-S_{cn}}}{i!\mu^i} P_0 \quad (12)$$

The dropping probability $P_D$ of handover request is:

$$P_D = P_S = \frac{(\lambda_n+\lambda_h)^{S_c}(\lambda_n+\lambda_h)^{S_{cn}-S_c}\lambda_h^{\,S_h-S_{cn}}\lambda_h^{\,S-S_h}}{S!\mu^S} P_0 \quad (13)$$

The equation for channel utilization is given below:

$$\%\, channel\ utilization = \frac{\left[\{(1-P_B)\lambda_n\}+\{(1-P_D)\lambda_h\}\right]}{\mu S} \times 100 \quad (14)$$

So, in our scheme, we have proposed a new factor $x$, the value of which has been limited to a minimum value ( $x_{\min}$ ) to 1. This limit is adapted based on channel occupancy. This factor is multiplied with the total number of guard channels reserved for handover priority. Thus, the number of guard channels varies from a minimum value to its maximum value which will be adapted with the change of factor $x$, determined by the channel occupancy. It is to be noted that, for prioritizing handover calls, it is required to keep a minimum number from the total number of guard channels fixed and let the rest of the guard channels vary, so that, we can find a promising outcome in our proposed scheme. Again, we can use a fixed and limited number of allocated channels for better quality of services. We do not need to borrow channels from other cells or, we do not require to split one channels into two or more by splitting their bandwidth that makes the number of channels allocated per cell variable.

The overall procedure of the proposed scheme can be viewed through the flowchart given in Fig.2. The algorithm of the proposed scheme starts with the incoming of either a new call or a handover call, which is accepted if there are channels available within $S_C$. If no, then system calculates the factor $x$ and sets the parameters. Then the system checks for available channels within $S_{Cn}$ both for new calls and handover calls.

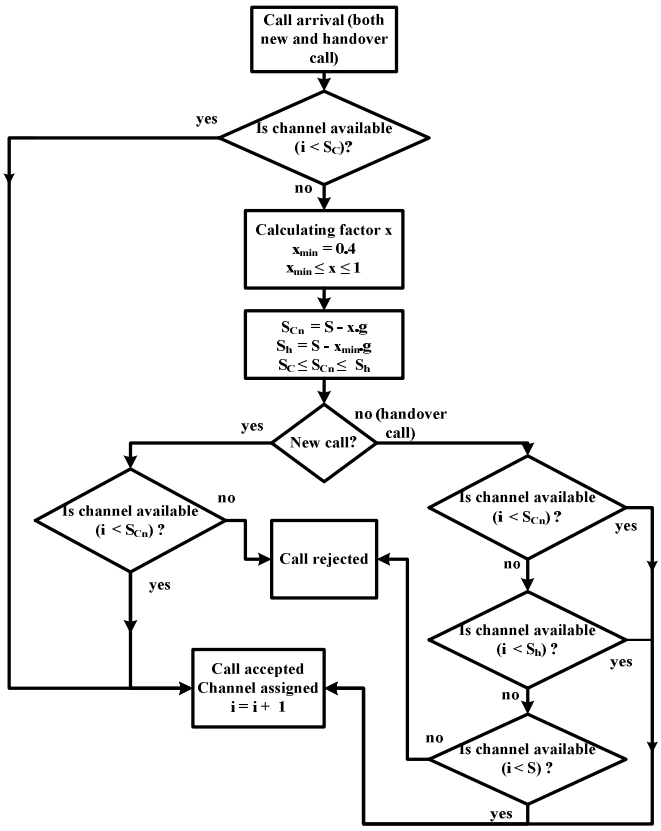



**Fig. 2:** Channel reservation policy of the proposed scheme

From $S_{Cn}$ to S, the system accepts handover calls only. The factor $x$ adapts the number channels which should be reserved only for handover calls.

The algorithm of non-priority scheme is very simple and not promising to be used in wireless networks. Since, it does not give any priority to handover calls, the system does not check for reserved channels for handover calls. The total number of channels can be used both by the handover calls and new calls.

Fixed guard channel reservation scheme follows an algorithm where, the system checks for fixed amount of reserved channels for prioritize handover calls. So, up to a fixed states, the new calls are accepted, even if there is no handover calls, the guard channels remain unused, which is referred as the misuse of limited resources.

However, it is required to use this limited resources to the maximum extent to provide better quality of service in wireless networks. The proposed idea of adaptive channel reservation scheme has been designed to use limited resources by using unused guard channels for new calls when there is no or less handover calls. Thus, the algorithm of the proposed idea always checks for the available channels both for new calls and handover calls and uses the unused guard channels which were not possible in fixed guard channel reservation scheme.

## III. PERFORMANCE ANALYSIS

In this section, we have performed the performance analysis of the proposed scheme. The overall performance has been simulated in MATLAB environment. Every data has been analyzed to get a proper outcome in our proposed scheme. The performance of the scheme was analyzed with different values of the input parameters to obtain better optimized result.

We have compared the performance of the proposed scheme with *non-priority scheme* and *fixed guard channel reservation scheme*. The main performance metrics considered here are: handover call dropping probability, new call blocking probability, and channel utilization. As, it has been stated earlier that, fixed guard channel reservation scheme shows minimum handover call dropping probability. In fact it is possible to achieve handover dropping probability tends to zero in the fixed guard channel reservation scheme. However, lower channel utilization is achieved in this case. Again, non-priority scheme shows the highest channel utilization, but, it does not give any priority to handover calls which cannot be accepted either. So, we have compared our scheme with those schemes and achieved promising performance. In our proposed scheme, we have shown that it gives better channel utilization keeping reduced handover dropping probability. Thus, it overcomes the drawbacks of both the non-priority scheme and fixed guard channel reservation scheme.

For the analysis of the proposed scheme and the other schemes, we have selected 4% guard channel of the total number of channels, which optimize both the handover priority and channel utilization for a better performance. Minimum value of factor $x$ has been set to 0.4. All the parameters are under continuous observation in the system by the operator.

Fig 3 shows that our proposed adaptive scheme can reduce the handover call dropping probability within a good acceptable range. The proposed scheme shows handover dropping probability around $10^{-3}$ which is a promising outcome.

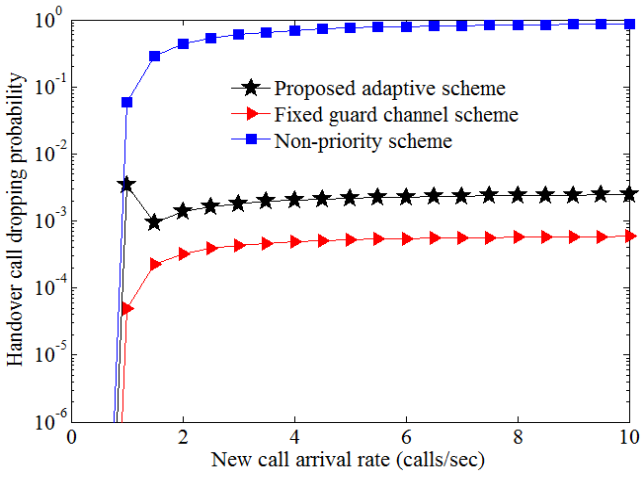



**Fig. 3:** Comparison of handover call dropping probability with 4% guard channel

The curve for the proposed scheme in fig. 3 shows much better response than the non-priority scheme.

Fig. 4 presents the performance of the wireless system in the case of new call blocking probability. The proposed adaptive scheme has made a decrease in new call blocking probability. Here, we also have to sacrifice the new call acceptance to get better performance for handover priority but with slightly reduced new call blocking probability, which is an improved performance compared to fixed guard channel reservation scheme.

Fig 5 represents the comparison of channel utilization. The proposed scheme shows better channel utilization than the fixed guard channel scheme. The non-priority scheme always shows the highest channel utilization as it sacrifices the priority of the handover calls. Fixed guard channel scheme allows handover priority with reduced channel utilization. The proposed scheme gives both reduced handover call dropping probability and improved channel utilization, which is an optimum performance in wireless networks. About 97% channel utilization has been obtained in the proposed adaptive scheme.

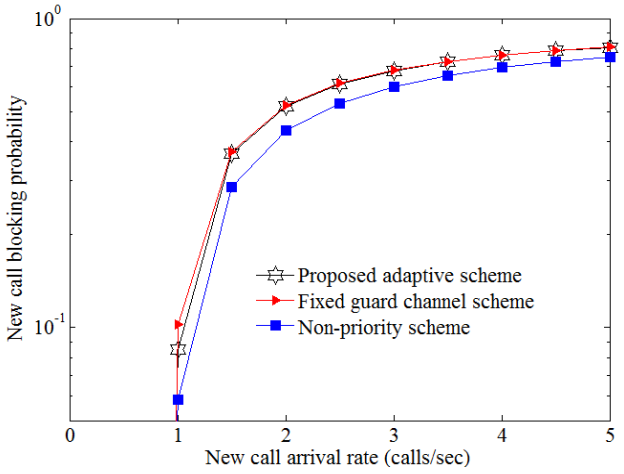



**Fig. 4:** Comparison of new call blocking probability with 4% guard channel

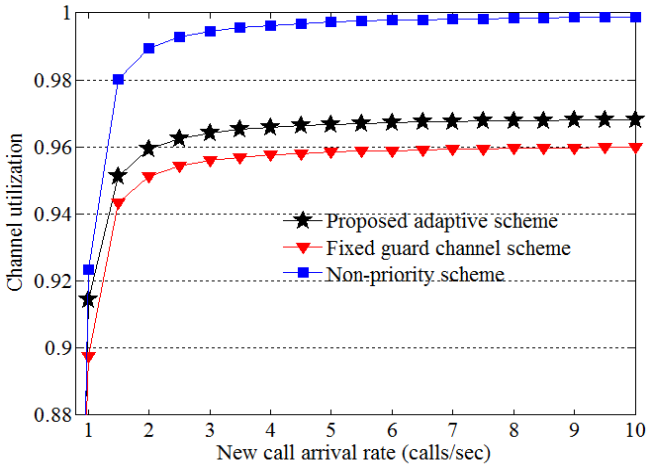



**Fig. 5:** Comparison of channel utilization with 4% guard channel

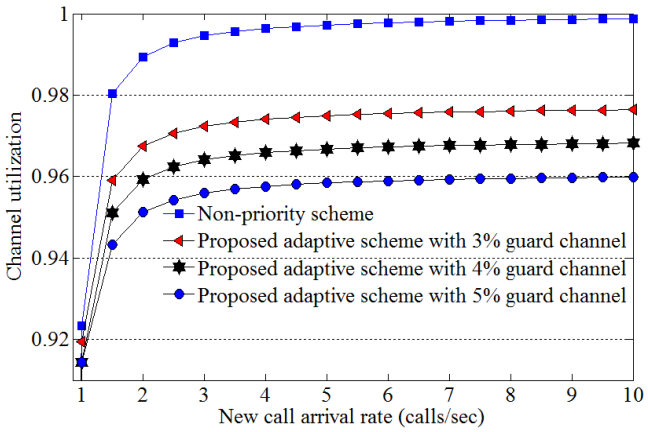



**Fig. 6:** Channel utilization of the proposed adaptive scheme at different percentage of guard channels

We have also performed the analysis of the proposed scheme at 3%, 4%, and 5% guard channels. Here, the operator has the flexibility to change quantity of guard channels according to the requirement of the system. Fig 6 shows that, the increasing number of reserved guard channels gives reduced channel utilization but still gives better performance than the fixed guard channel scheme. Though 5% guard channel gives better performance for handover priority, it degrades the channel utilization. Again, 3% guard surely gives better channel utilization but degrades the performance of handover priority. Hence, the operator can vary the percentage of guard channels to get optimized performance based on the network characteristics.

## IV. CONCLUSION

Prioritizing handover calls is a vital process in wireless communication. By comparing with the other schemes, we found that though our proposed scheme has a utilization slightly less than the non-priority scheme, it has a sharp decrease in the handover dropping probability and our proposed scheme has a better utilization than the channel reservation scheme. In addition, our proposed scheme has been made more user friendly so that operator can use this scheme more efficiently by varying the different parameters such as - channel holding time, dwell time, and number of guard channels according to the network characteristics.

## REFERENCES

[1] A. Sgora and D. D. Vergados, "Handoff Prioritization and Decision Schemes in Wireless Cellular Networks: A Survey," *IEEE Commun.Surveys & Tutorials*, vol 11, no. 4, pp. 57-77, 2009.

[2] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd edition, Prentice Hall, 2002.

[3] Stojmenovic, *Handbook of Wireless Networks and Mobile Computing*, John Wiley & Sons, Inc, 2002.

[4] W. C. Y. Lee, *Wireless and Cellular Communications*, McGraw-Hill, 2005.

[5] M. Z. Chowdhury, Y. M. Jang and Z. J. Haas, "Call Admission Control Based on Adaptive Bandwidth Allocation for Wireless Networks," *Journal of Communications and Networks (JCN)*, vol. 15, no. 1, pp. 15-24, February 2013.

[6] M. Z. Chowdhury, Y. M. Jang and Z. J. Haas, "Cost-Effective Frequency Planning for Capacity Enhancement of Femtocellular Networks," *Wireless Personal Communications,* vol. 60, no. 1, pp. 83-104, September 2011.

[7] D. D. Vergados, "Simulation and Modeling Bandwidth Control in Wireless Healthcare Information Systems," *SIMULATION*, vol. 83, no. 4, pp. 347-364, April 2007.

[8] Z. Ye, L. K. Law, S. V. Krishnamurthy, Z. Xu, S. Dhirakaosal, S. K. Tripathi, M. Molle, "Predictive Channel Reservation for Handoff Prioritization in Wireless Cellular Networks," *Computer Networks,* vol. 51, no. 3, pp. 798-822, February 2007.

[9] J. Diederich and M. Zitterbart, "Handoff Prioritization Schemes Using Early Blocking," *IEEE Communications Surveys & Tutorials*, vol. 7, no. 2, pp. 26-45, 2005.

[10] G. Sharma and G. N. Purohit, "Queue Modeling of Handoff Calls with Sub Rating in Wireless Cellular Networks," *Applied Mathematical Sciences*, vol. 5, no. 21, pp. 1017-1024, 2011.

[11] M. Schwartz, *Mobile Wireless Communications, Cambridge University Press*, Cambridge, 2005.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Design and Performance Analysis of a Dual Band Waveguide Dipole Feed Cassegrain Antenna and Link Budget Optimization

MD. Tofael Hossain Khan*, Wahida Sharmin, Anindya Kumar Kundu, Md. Osman Goni, Kazi Abul Barket[2]
Dept. of Electronics & Communication Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh.
[2] Physics Discipline, Khulna University, Khulna-9208, Bangladesh.
*tofael18@gmail.com

*Abstract*—**This paper deals with the design and performance analysis of a very small size, low-cost, low-profile, high gain and high directivity Dual Band Waveguide Dipole Feed Cassegrain antenna considering the link budget optimization. The proposed antenna system has a gain of 28.49dB, directivity of 28.87dBi, return loss of -18.837447dB, Voltage Standing Wave Ratio (VSWR) of 1.2582, bandwidth 433.7 MHz (6.4354 to 6.8691 GHz) at C band and at X band it operates in two regions with 195.2 MHz (8.1748 to 8.37GHz) and 253 MHz (11.241 to 11.494 GHz), antenna efficiency of 91.56% and 3dB angular beamwidth i.e., the Half Power Beamwidth (HPBW) of 2.9deg. The resonant frequencies of the antenna are 6.678, 8.2778 and 11.386 GHz respectively. The proposed antenna system can be used for both C-band and X-band applications like satellite communications transmissions, Wi-Fi, cordless telephones, weather radar systems, medical applications and other wireless systems. The antenna system is designed and simulated in the CST Microwave Studio. Link budget optimization is performed in order to analyze the critical factors in the transmission chain and to optimize the performance characteristics. The link budget determines what size antenna is to use, power requirements and in general, the overall customer satisfaction.**

***Keywords—Cassegrain antenna, dual band, gain, link budget optimization, satellite communications.***

## I. INTRODUCTION

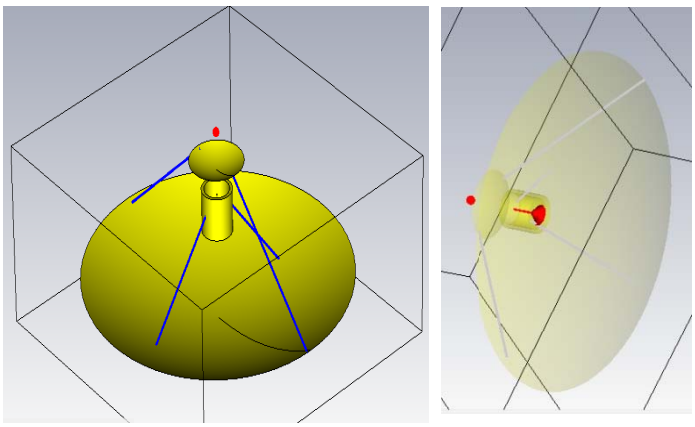
With the change of era new technologies are getting familiarity in satellite communications. Due to the drastic growth of modern mobile satellite communication technology, the use of mobile antennas have increased due to their very small size, low-cost, low-profile, high gain and high directivity. Antennas for mobile satellite communications are widely presented in books and papers in the last decade as presented in [1-3]. Global Mobile Satellite Communications: For Maritime, Land and Aeronautical Applications are illustrated in [4]. Small Cassegrain antenna for passive remote sensing at L-band is studied in [5]. A compact dual-band dipole antenna fed by a coplanar waveguide for wireless communication is presented in [6]. S/X band feed development for 12m Cassegrain antenna is described in [7]. A C/X/Ku-band dual polarized Cassegrain antenna system is demonstrated in [8]. The satellite link is much like the terrestrial microwave radio relay link with the advantage of not requiring as many re-transmitters as are required in the terrestrial link. Transmission of signals over a satellite communication link requires Line-of-Sight (LoS) communication. Link analysis basically relates the transmit power and the receive power [9]. The communication link between a satellite and the Earth Station (ES) is exposed to a lot of impairments such as free space path loss, rain loss, pointing loss and atmospheric attenuations etc. [10].

The organization of this paper is as follows- Section II conducts the antenna architecture. Section III illustrates the simulation results. Section IV reveals the result analysis. Section V conducts link budget optimization. Section VI resembles the cost calculation of the proposed antenna system. Finally, Section VII provides some concluding remarks.

## II. ANTENNA ARCHITECTURE

The Aluminum lossy metal is used as substrate for the main reflector, sub-reflector, dipole antenna and waveguide. The relative permeability ($\mu_r$) and electrical conductivity of the substrate material are 1.0 and 3.56e+007(S/m) respectively. The main reflector diameter is 50 cm with 0.2 cm thickness and focal distance of 25 cm. The sub reflector diameter is 10 cm with 0.2 cm thickness and focal distance of 5 cm. The distance between the two reflectors is 20 cm which ensures that the focal of main reflector and virtual focal of sub reflector is at same point. A dipole antenna of 4 cm long and 0.1 cm diameter is placed between two reflectors facing toward the sub reflector convex surface. The distance of dipole antenna is 15 cm from the main reflector surface and 5 cm from the sub reflector surface (the focal point of the sub reflector) without any electrical conduction. A one side open waveguide with a length of 5 cm and 0.165 cm thickness is used around the dipole to direct the E and H field of the dipole to the sub reflector convex surface. Two bars are used as a stand to keep the antenna above the reflector which has no effect on electric or magnetic field. Another two bars are used to keep the sub reflector stable. Both the stands connected

with main reflector body. The excitation signal is applied at the end of the dipole at the closed end of waveguide. Fig. 1 shows the structure of the Waveguide Dipole Feed (WDF) Cassegrain antenna-



(a) (b)

Figure 1. Structure of the waveguide dipole feed (WDF) Cassegrain antenna (a) Solid view (b) Transparent view

## III. SIMULATION RESULTS

### A. *Electric Field Distribution*

Electric field distribution of the WDF Cassegrain antenna is given below –

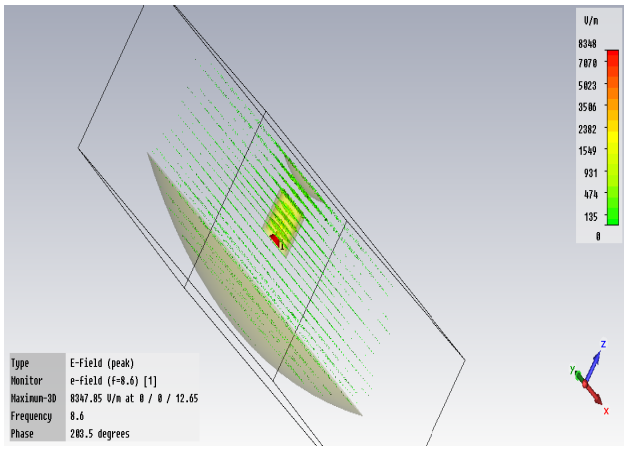



Figure 2. Electric field distribution of WDF Cassegrain antenna.

### B. *Voltage Standing Wave Ratio (VSWR) and Return Loss*

Fig. 3 shows the graph of VSWR vs. Frequency. From this figure it can be seen that the VSWR value is very near to the unity which is mostly expected. The VSWR value of the antenna is 1.2582 at 6.678 GHz, 1.7977 at 8.2778 GHz and 1.6405 at 11.386 GHz.

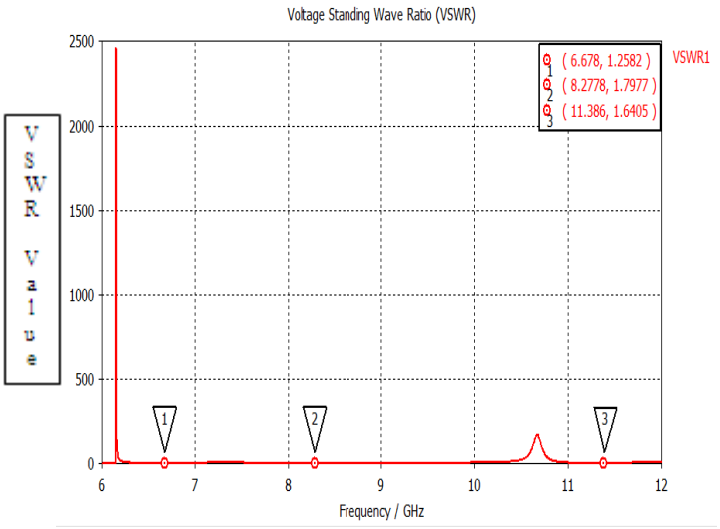



Figure 3. VSWR vs. Frequency curve of the WDF Cassegrain antenna.

Fig. 4 shows the graph of return loss vs. frequency. From this figure we can observe the bandwidth of the antenna. The bandwidth of the antenna is 433.7 MHz which operates in C band (6.4354 to 6.8691 GHz) at 6.678 GHz. The antenna also operates in X band in two regions with a bandwidth of 195.2 MHz (8.1748 to 8.37GHz) at 8.2778 GHz and 253 MHz (11.241 to 11.494 GHz) at 11.386 GHz. The value of the return loss is -18.837447dB at 6.678GHz.

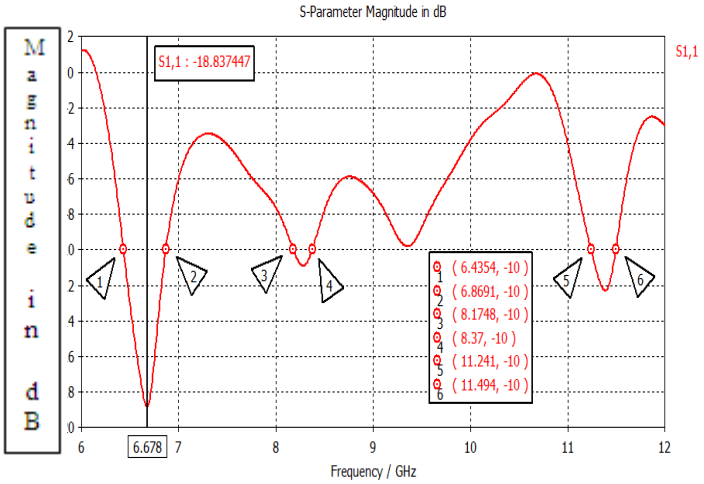



Figure 4. Return loss of proposed WDF Cassegrain antenna.

### C. *Farfield Radiation Pattern (3D), Gain Pattern and Antenna Efficiency*

Fig. 5 shows the farfield radiation pattern of the Waveguide Dipole Feed Cassegrain antenna. The radiation pattern of the antenna is very much pointed and beamwidth is narrow. From this figure it can be seen that the directivity and gain of the antenna is about 28.87 dBi and 28.49 dB

respectively. The radiation efficiency is -0.3830dB. The antenna efficiency is 91.56%.

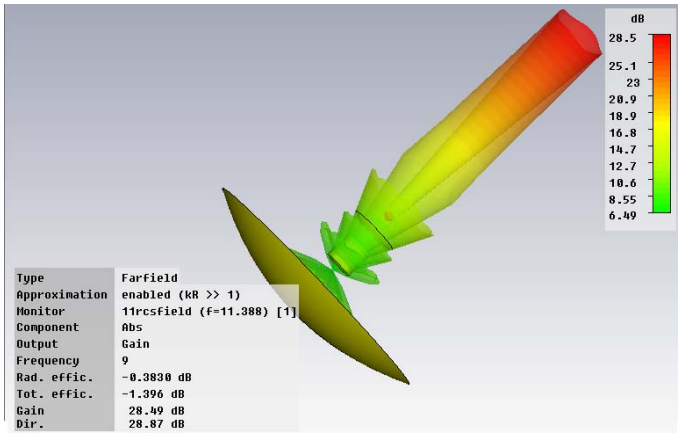



Figure 5. Farfield radiation pattern.

Fig. 6 shows the farfield gain pattern of the antenna. It can be seen that there is negligible side lobes and only one main lobe in the gain pattern. The main lobe magnitude is 33.5 dB, main lobe direction is 3.0 deg., 3 dB angular beamwidth is 2.9 deg. and side lobe level is -9.1 dB.

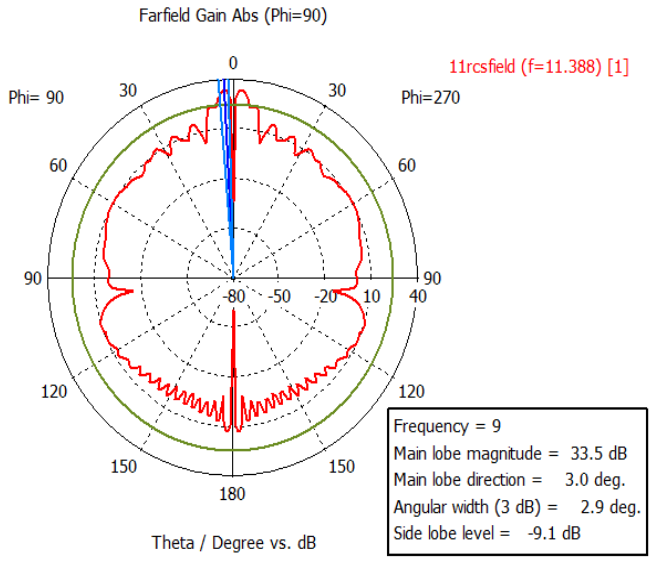



Figure 6. Farfield gain pattern.

## IV. Result Analysis

VSWR, return loss, farfield radiation pattern, antenna gain, directivity and antenna efficiency of the proposed antenna system show reasonable characteristics. The performance of the antenna is quite good. This antenna can be used in both C and X-band applications for its effective performance.

TABLE I. PROPOSED DUAL BAND WDF CASSEGRAIN ANTENNA PARAMETERS AND THEIR VALUES AT A GLANCE

| Designed Dual Band WDF Cassegrain Antenna Parameters | Simulation Results |
|---|---|
| VSWR | 1.2582 |
| Return Loss (in dB) | -18.837447dB |
| Gain (in dB) | 28.49dB |
| Directivity (in dBi) | 28.87dBi |
| Half Power Beamwidth | 2.9 deg. |
| Antenna Efficiency (%) | 91.56% |
| Bandwidth (MHz) | C-band → 433.7 MHz<br>X-band →253 MHz and 195.2 MHz |

The results in the Table I reveal that the proposed Dual Band WDF Cassegrain antenna system is useful for both C and X-band applications like satellite communications transmissions, Wi-Fi, cordless telephones, weather radar systems, medical applications and other wireless systems.

## V. Link Budget Optimization

The results obtained from the link budget calculator are shown below [11]-

### *A. Uplink Budget*

| | |
|---|---|
| Uplink frequency GHz | 6 |
| Uplink antenna diameter m | .5 |
| Uplink antenna aperture efficiency e.g. 0.65 | .9156 |
| Uplink antenna transmit gain dBi | 29.56004 |
| Uplink antenna, power at the feed W | 850 |
| Uplink EIRP dBW | 58.85423 |
| Range (35778 - 41679) km | 38500.0 |
| Uplink path loss dB | 199.7222 |
| Uplink pfd at satellite dBW/m^2 | -103.855 |
| Bandwidth Hz | 433700000 |
| Satellite uplink G/T dB/K | 40.938 |
| Uplink C/N dB | 42.29810 |

*B. Downlink Budget*

| | |
|---|---|
| Downlink frequency GHz | 4 |
| Downlink receive antenna diameter m | .5 |
| Downlink receive antenna aperture efficiency e.g. 0.65 | .9156 |
| Downlink system noise temperature (antenna+LNA) K | 120 |
| Downlink receive antenna gain dBi | 26.03822 |
| Downlink receive antenna G/T dB/K | 5.246410 |
| Downlink satellite EIRP dBW | 35.5 |
| Downlink path loss dB | 196.2004 |
| Downlink C/N dB | -13.2258 |

The results obtained from the link budget calculator shows that our proposed antenna system is practically realizable by showing similarity between the gain obtained by simulation and the gain obtained by the calculator which is approximately equal to 29 dB.

## VI. Cost Calculation

Large antennas are expensive to construct and install, with costs exceeding $1M for 30-m diameter fully steerable antennas [12]. The cost of large fully steerable antennas has been quoted as [13]-

$$\text{Cost} = \$\ y(D)^{2.7} \quad (1)$$

where, D is the diameter of the antenna aperture in feet. The constant y in equation (1) depends on the currency used and inflation, but might typically be around five U.S. dollars in the early 1980s.

The diameter of the proposed Dual Band WDF Cassegrain antenna aperture is 50 cm i.e.; 1.640419948 feet which in turn gives-

$$\text{Cost} = \$\ y(D)^{2.7} = \$\ \{5 \times (1.64041998)^{2.7}\} = \$\ 19.02605$$

## VII. Conclusion

The design of a Dual Band Waveguide Dipole Feed Cassegrain antenna and its performance analysis for both C and X-band along with the link budget optimization has been demonstrated in this paper. The simulation results of the proposed antenna system resemble very good performance. The results obtained from the link budget calculator show that our proposed antenna system is practically realizable by showing similarity between the gain obtained by simulation and the gain obtained by the calculator. Also the cost calculation reveals that our proposed antenna system is cost effective. In this paper, CST Microwave Studio software has been used for all the simulations which provide effective and satisfactory results. The proposed antenna system provides high gain, directivity, half power beamwidth, efficiency and bandwidth. It also shows very low value in case of VSWR which is near about unity that satisfies the antenna specification.

## VIII. Future Work

We will try to design an Ultra-Wide band satellite antenna system and to develop practical infrastructure of the above mentioned antenna system.

## References


[1] Basari, "Development of simple antenna system for land mobile satellite communications," Doctoral Dissertation, Feb. 2011.

[2] S. Ohmori, H. Wakana, and S. Kawase, "Mobile satellite communications," *Artech House*, Norwood, MA, USA, 1998.

[3] J.H. Lodge, "Mobile satellite communication systems: toward global personal communications," *IEEE Commun. Mag.*, vol. 29, no. 11, pp. 24–30, Nov. 1991.

[4] S.D. Ilcev, "Global mobile satellite communications: For Maritime, Land and Aeronautical Applications," *Springer*, Dordrecht, the Netherlands, 2005.

[5] W. Wasylkiwskyj, R. Lang, M. Ogut, "Small cassegrain antenna for passive remote sensing at L-band," *US National Committee of URSI National, Radio Science Meeting (USNC-URSI NRSM)*, 9-12 Jan. 2013, pp. 1.

[6] Liping Han, Runbo Ma, Xinwei Chen, Wenmei Zhang, "A compact dual-band dipole antenna fed by a coplanar waveguide for wireless communication," *International Conference on Microwave and Millimeter Wave Technology (ICMMT)*, (Volume: 1) 5-8 May 2012, pp. 1-3.

[7] C.S. Kim, N. Moldovan, N. Hanchett, "S/X band feed development for 12m Cassegrain antenna," *Antennas and Propagation Society International Symposium*, APSURSI '09. IEEE, 1-5 June 2009, pp. 1-4.

[8] M. Beadle, R. Chugh, D. Geen, S. Patel, R. Schwerdtferger, S. Casey, "A C/X/Ku-band dual polarized Cassegrain antenna system," *Antennas and Propagation Society International Symposium*, IEEE, (Volume: 1), 11-16 July 1999, pp. 692-695.

[9] Aderemi A. Atayero, Matthew K. Luka and Adeyemi A. Alatishe, "Satellite link design: A Tutorial," *International Journal of Electrical and Computer Sciences*, Vol: 11 No: 04, Aug.2011.

[10] Charan Langton. (2002) Intuitive Guide to Principles of Communications, copyright 1998 and 2002. [Online]. *Available:* http://www.Complextoreal.com/

[11] Satellite Link Budget Calculator. [Online]. Available: http://www.satsig.net/ linkbugt.htm/

[12] Timothy Pratt et al., *Satellite Communications*, ISBN: 9971-513-226.

[13] N.E. Feldman, "The link from a communications satellite to a small ground terminal," *Microwave journal*, 7, 39-44, 1964.

# Performance of Relay Assisted Multiuser Uplink MIMO Wireless Communication Using Walsh Hadamard Sequences

M.M. Kamruzzaman
Key Lab of Information Coding & Transmission,
Southwest Jiaotong University, China
E-mail: m.m.kamruzzaman@gmail.com

*Abstract*— **This paper compares the performance of single user and multiuser with Walsh Hadamard spreading sequences for relay assisted wireless link as well as direct link. Space Time Block Coding (STBC) is used for encoding the information at user's handset and relay. Input information are mapped using Walsh Hadamard spreading sequences and send to a QPSK or 16 QAM or 64 QAM modulator. Modulated symbols are encoded by STBC encoder and split into n streams which are simultaneously transmitted using n transmit antennas. Relay receives the rayleigh fading effected signals where it decodes and forwards the signals for destination after further encoding using STBC. It is observed that there is around 9-12 dB coding gain for using Walsh Hadamard spreading sequences in multiuser system and there is around 11 dB coding gain for using relay in the MIMO system.**

***Keywords—Multiuser, Walsh Hadamard Spreading Sequences, STBC, MIMO, Relay.***

## I. INTRODUCTION

Direct Sequence Spread Spectrum (DSSS) is an important technology for multiuser applications and there has been considerable research on it over the past several decades [1-3]. In DSSS system, the information signal is modulated by a spreading sequence prior to transmission, and the spreading sequence is typically known to the desired user's receiver, where it used to perform the dispreading the signal and recover the transmitted information. Walsh-Hadamard and Gold spreading codes are widely used for DSSS. In our research, we have used Walsh Hadamard spreading sequences. Walsh Hadamard sequences are orthogonal, so it is very convenient to use for multiple users.

On the other hand, Space Time Block Coding (STBC) has received huge attention for supporting much greater data rate and higher reliability than the single Input Single Output (SISO) system [4-12]. So, information of every user is encoded by STBC before transmission to get higher data rate and better BER performance. We have further included relay for decoding and forwarding the transmitted information to provide long distance wireless communication as well as further improving the capacity and transmission reliability [13-16].

Relay assisted multiuser communications have been widely studied [17-39] but mostly precoding techniquies have been used [33-38]. In [39] non orthogonal spreading codes are used but it is not MIMO system. In this paper, we have studied a relay assisted multiuser MIMO wireless communication with orthogonal Walsh Hadamard sequences and our scheme performs better than earlier proposed schemes with respect to BER.

## II. SYSTEM MODEL

Consider an uplink wireless communication with $u$ active users. Each user and relay are equipped with two transmit antennas ( $n = 2$ ), whereas the relay and destination are equipped with 2/3/4 receive antennas ( $m = 2/3/4$ ) as shown in fig.1 and fig. 2. In the following, we will first describe the structure of transmitter of user followed by the channel model with decoding and reencoding the signal at relay and the base station receiver design.

### *A. Encoding at Transmitter of User*

The information bit sequences $b_u$ for each user u, u=1,…,k, are mapped using spreading code. Different rows of Walsh Hadamard matrix ( $w_p$ ) are used for different user as spreading code. Walsh Hadamard sequences belong to the binary sample space, and therefore, they have fixed power. Walsh Hadamard sequences are available for lengths that are powers of 2 which can be write as:

$$W_2 = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix}$$

For many users we use the larger matrices by using the recursion as:

$$W_{2p} = \begin{bmatrix} W_p & W_p \\ W_p & -W_p \end{bmatrix}$$

Suppose four users are communication to base station, so we can write the matrix for four user as :

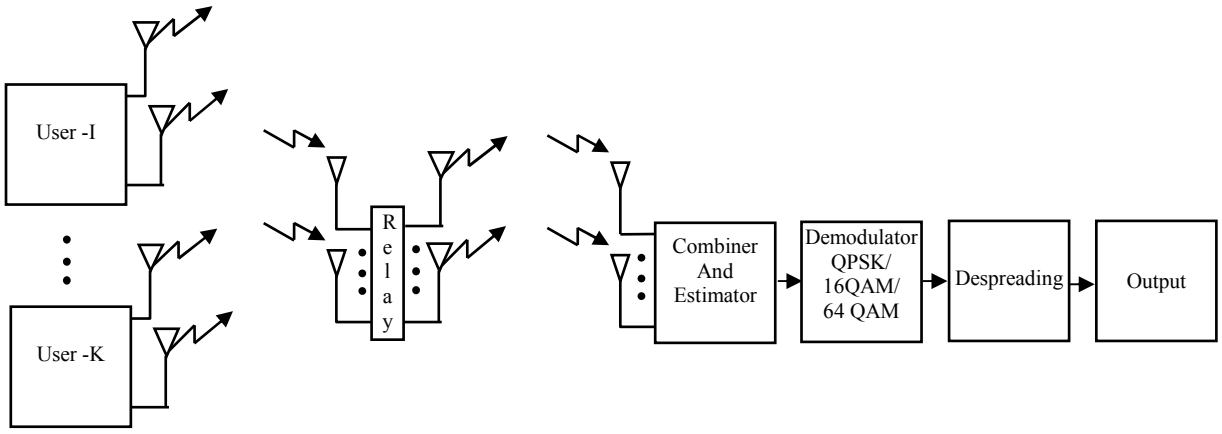



Fig. 1. System Block diagram

$$W_4 = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix}$$

If the each row of Walsh Hadamard sequences is represented by $r_u$, the mapped bits of each user are $m_u = b_u r_u$. . A QPSK or 16 QAM or 64 QAM modulator modulates the mapped bits $m_u$. Modulated symbols are encoded using STBC encoder before transmission according to table I .

*B. Received Signal at Relay*

If the signals $s_t^i$ , $i$ =1, 2, ... $n$ are transmitted simultaneously at each time slot $t$ using 2 transmit antennas of each user's handset $u$, then the signal received at antenna $j$ of relay can be written as:

$$\left(r_t^j\right)_u = \sum_{i=1}^{2} \left(p_{i,j}^{SR}\right)_u \left(h_{i,j}\right)_u \left(s_t^i\right)_u + \left(n_t^j\right)_u \qquad (1)$$

where $\left(r_t^j\right)_u$ is the received symbol on the $j^{th}$ receiver antenna of relay at time slot $t$ of user $u$

$\left(p_{i,j}^{SR}\right)_u$ is path loss from transmit antenna i of source to receive antenna j of relay of user $u$ and $p_{i,j}^{SR} \alpha \frac{1}{d_{SR}^2}$

TABLE I. THE ENCODING AND TRANSMISSION SEQUENCE FOR TWO TRANSMISSION ANTENNAS OF USER'S HANDSET

| | **Antenna-I** | **Antenna-II** |
|---|---|---|
| **Time slot-I** | $(s_1)_u$ | $(s_2)_u$ |
| **Time slot-II** | $(-s_2^*)_u$ | $(s_1^*)_u$ |

$\left(h_{i,j}\right)_u$ is the path gain from transmit antenna $i$ to receive antenna $j$ of user $u$. The path gains are modeled as samples of independent complex Gaussian random variables with variance 0.5 per real dimension. The wireless channel is assumed to be quasi-static so that the path gains are constant over a frame of length *l* and vary from one frame to another.

$\left(s\right)_u$ is the transmitted symbol at time slot $t$ of user $u$ and

$\left(n_t^j\right)_u$ is the noise on $j^{th}$ receive antenna at time slot $t$ of user $u$. It is assumed that the noise on each receive antenna at each time slot is independent from the noise on the other receive antennas.

It is considered that $H_{ij} = p_{i,j} h_{i,j}$, then (1) can be rewritten as:

$$\left(r_t^j\right)_u = \sum_{i=1}^{2} \left(H_{i,j}\right)_u \left(s_t^i\right)_u + \left(n_t^j\right)_u \qquad (2)$$

*C. Decoding at relay*

For detecting symbols $\left(s_1\right)_u$ and $\left(s_2\right)_u$ of two transmit antennas of every user, (3) and (4) decision metrics have been used [4-5]

$$\left| \left[ \sum_{j=1}^{m} \left( \left(r_1^j\right)_u \left(H_{1,j}^*\right)_u + \left(r_2^j\right)_u^* \left(H_{2,j}\right)_u \right) \right] - \left(s_1\right)_u \right|^2 + \left( -1 + \sum_{j=1}^{m} \sum_{i=1}^{2} \left| \left(H_{i,j}\right)_u \right|^2 \right) \left| \left(s_1\right)_u \right|^2 \qquad (3)$$

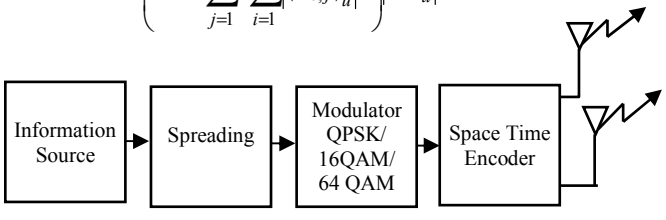



Fig. 2. Block diagram of transmitter of every user

TABLE II. THE ENCODING AND TRANSMISSION SEQUENCE FOR TWO TRANSMISSION ANTENNAS OF RELAY

| | Antenna-I | Antenna-II |
|---|---|---|
| **Time slot-I** | $(\hat{s}_1)_u$ | $(\hat{s}_2)_u$ |
| **Time slot-II** | $(-\hat{s}_2^*)_u$ | $(\hat{s}_1^*)_u$ |

$$\left\| \left[ \sum_{j=1}^{m} \left( \left(r_1^j\right)_u \left(H_{2,j}^*\right)_u - \left(r_2^j\right)_u^* \left(H_{1,j}\right)_u \right) \right] - (s_2)_u \right\|^2 + \left( -1 + \sum_{j=1}^{m} \sum_{i=1}^{2} \left| \left(H_{i,j}\right)_u \right|^2 \right) \left| (s_2)_u \right|^2 \quad (4)$$

we denote the detected symbols of two tansmit antennas as $(\hat{s}_1)_u$ and $(\hat{s}_2)_u$ .

*D. Encoding at relay:*

STBC encoder of relay encodes the decoded symbols $(\hat{s}_1)_u$ and $(\hat{s}_2)_u$ according to the Table II and then at each time slot *t*, signals $\left(\hat{s}_t^i\right)_u$ , *i* =1, 2, … *n* are transmitted simultaneously using two transmit antennas.

*E. Received signal at destination:*

At time *t* the signal $y_t$ , received at destination, is given by

$$\left(y_t^j\right)_u = \sum_{i=1}^{2} \left(P_{i,j}^{RD}\right)_u \left(a_{i,j}\right)_u \left(\hat{s}_t^i\right)_u + \left(\eta_t^j\right)_u \quad (5)$$

where $\left(y_t^j\right)_u$ is the received symbol on the $j^{th}$ receiver antenna of destination at time slot *t* of user *u*

$\left(p_{i,j}^{RD}\right)_u$ is path loss from transmit antenna i of source to receive antenna j of destination of user *u* and $p_{i,j}^{RD} \alpha \frac{1}{d_{RD}^2}$

$\left(a_{i,j}\right)_u$ is the path gain from transmit antenna *i* to receive antenna *j* of user *u* .

$\left(\hat{s}_t^i\right)_u$ is the transmitted symbol from transmit antenna *i* of user *u* at each time slot *t*.

$\left(\eta_t^j\right)_u$ is the noise on $j^{th}$ receive antenna at time slot *t* of user *u* .

It is considered that $A_i = P_i^{RD} a_i$ , then (8) can be rewritten as:

$$\left(y_t^j\right)_u = \sum_{i=1}^{2} \left(A_{i,j}\right)_u \left(\hat{s}_t^i\right)_u + \left(\eta_t^j\right)_u \quad (6)$$

*F. Decodingl at Destination*

The combiner combines received signals of destination which are then sent to the maximum likelihood detector. For detecting symbols $\hat{s}_1$ and $\hat{s}_2$ of two transmit antennas, (7) and (8) decision metrics have been used.

$$\left\| \left[ \sum_{j=1}^{m} \left( \left(y_1^j\right)_u \left(A_{1,j}^*\right)_u + \left(y_2^j\right)_u^* \left(A_{2,j}\right)_u \right) \right] - (\hat{s}_1)_u \right\|^2 + \left( -1 + \sum_{j=1}^{m} \sum_{i=1}^{2} \left| \left(A_{i,j}\right)_u \right|^2 \right) \left| (\hat{s}_1)_u \right|^2 \quad (7)$$

$$\left\| \left[ \sum_{j=1}^{m} \left( \left(y_1^j\right)_u \left(A_{2,j}^*\right)_u - \left(y_2^j\right)_u^* \left(A_{1,j}\right)_u \right) \right] - (\hat{s}_2)_u \right\|^2 + \left( -1 + \sum_{j=1}^{m} \sum_{i=1}^{2} \left| \left(A_{i,j}\right)_u \right|^2 \right) \left| (\hat{s}_2)_u \right|^2 \quad (8)$$

The detected symbols are demodulated by a QPSK or 16 QAM or 64 QAM demodulator to extract the $\hat{m}_u$ . A bank of match filter is used to determine the coding sequence used by the transmitter where the received signal is matched to a different coding waveform to get the output $\hat{b}_u$ of user *u* using the following :

$$\hat{b}_u = \hat{m}_u \, r_u^T \quad (9)$$

where $r_u^T$ is the transpose of the Walsh Hadamard sequences for user *u*

## III. SIMULATION RESULTS

In this section, computer simulation is carried out to show the BER performance of the proposed system. The results are evaluated for several combinations of $T_x$ and $R_x$ antennas with and without relay. 64 QAM is used for simulation. It is considered that relay is placed at the middle of source and destination. We used two terms in fig.5-fig .7: Direct Link (DL) and Via Relay Link (VRL). DL means that information pass from source to destination without relay. On the other hand, VRL means that information pass from source to relay and then from relay to destination. 10 users is considered to simulate the system with Walsh Hadamard sequences.

We present the BERs to compare the performance of single user and multiuser with and without Walsh Hadamard

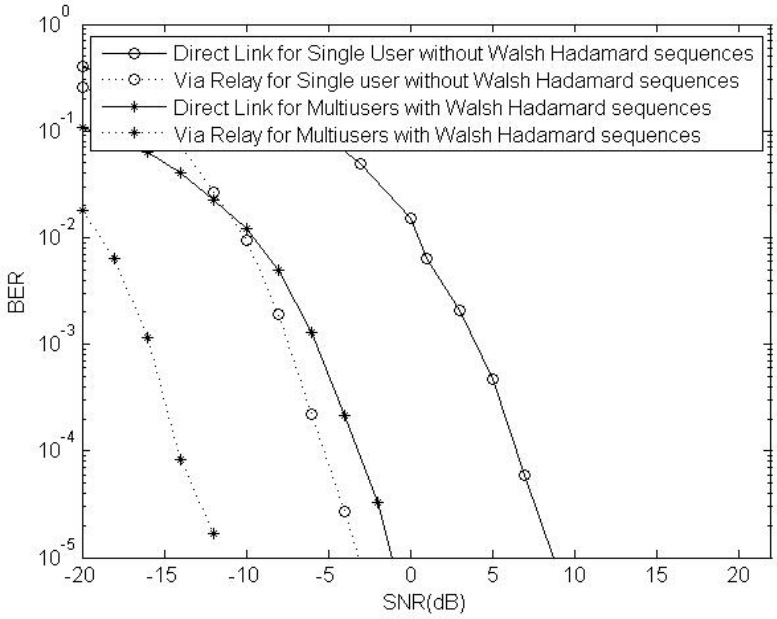



Fig. 4. BER performance comparison of direct line and via relay link with $2T_x$ & $3R_x$ for single user and multiuser with and without Hadamard sequences

sequences in fig.3. It also compares the performance of relay assisted link with direct link where user's handset is equipped with 2 $T_x$, relay is equipped with 2 $T_x$ and 2 Rx and base station is equipped with 2 Rx. It is observed that multiuser system with Walsh Hadamard sequences provides 12 dB coding compared to single user without Walsh Hadamard sequences at $10^{-5}$ and is around 11 dB gain for using relay with the same configuration at $10^{-5}$.

Fig. 4 shows the performance of of multiuser and single user with and without Walsh Hadamard sequences where user's handset is equipped with 2 $T_x$, relay is equipped with 2 $T_x$ and 3 Rx and base station is equipped with 3 Rx. It also compares the performance of relay assisted link with direct link. It is observed that multiuser system with Walsh Hadamard sequences provides 10 dB coding compared single user without Walsh Hadamard sequences at $10^{-5}$ and there is around 11 dB gain for using relay with the same configuration at $10^{-5}$.

Fig. 5 shows the performance of of multiuser and single user with and without Walsh Hadamard sequences

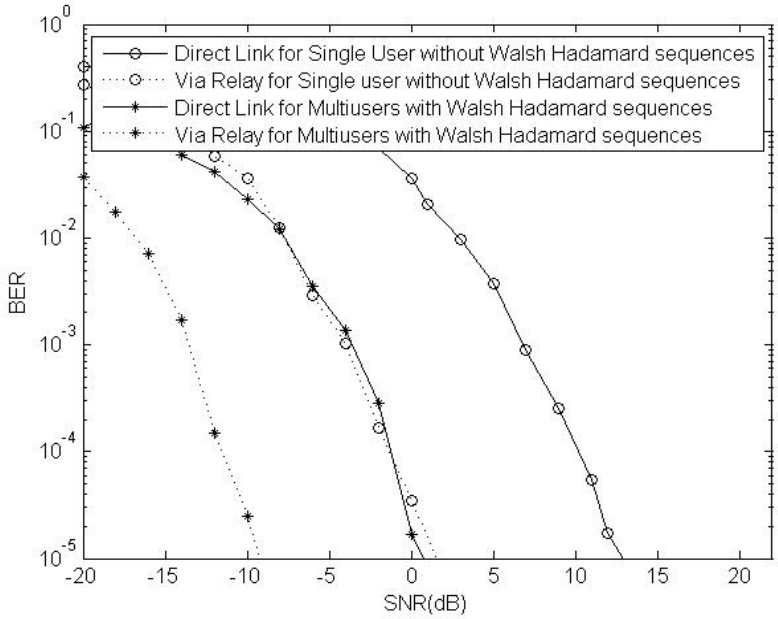



Fig. 3. BER performance comparison of direct line and via relay link with $2T_x$ & $2R_x$ for single user and multiuser with and without Hadamard sequences

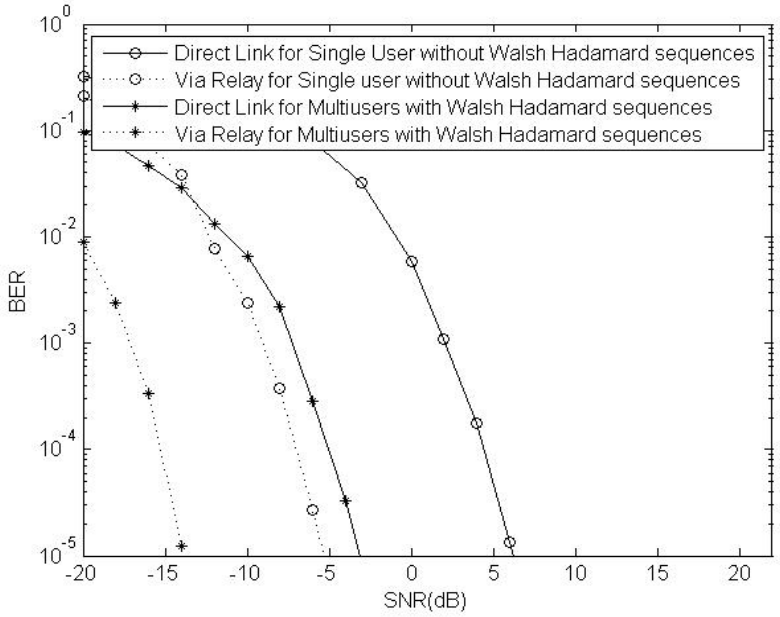



Fig. 5. BER performance comparison of direct line and via relay link with $2T_x$ & $3R_x$ for single user and multiuser with and without Hadamard sequences

where user's handset is equipped with 2 $T_x$, relay is equipped with 2 $T_x$ and 4 Rx and base station is equipped with 4 Rx. It also compares the performance of relay assisted link with direct link. It is observed that multiuser system with Walsh Hadamard sequences provides 9 dB coding compared to single user without Walsh Hadamard sequences at $10^{-5}$. There is around 11 dB gain for using relay with the same configuration at $10^{-5}$.

## IV. CONCLUSION

From the simulations results, it is observed that Walsh Hadamard sequences make a significant difference in performance with respect to BER. It is possible to get 9-12 dB gain by using Walsh Hadamard sequences for multiuser system with different combination of transmit and receive antennas at a BER of $10^{-5}$. And the performance can be further improved by employing relay between source and destination. It is possible to get around 11 dB gains by using relay in the system.

## ACKNOWLEDGMENT

The authors would like to thank the reviewers for the suggestions which help to improve the quality of this paper. In addition, the authors are also very thankful to Key Lab of Information Coding & Transmission, Southwest Jiaotong University, Chengdu, Sichuan, China for providing resources.

## REFERENCES

[1] Pickholtz, R.L.; Schilling, D.L.; Milstein, L.B., "Revisions to "Theory of Spread-Spectrum Communications - A Tutorial"," Communications, IEEE Transactions on , vol.32, no.2, pp.211,212, February 1984

[2] Rappaport Stephen, S.; Grieco, Donald M., "Spread-spectrum signal acquisition: Methods and technology," Communications Magazine, IEEE , vol.22, no.6, pp.6,21, June 1984

[3] A. J. Viterbi, CDMA: Principles of Spread Spectrum Communication. Reading, MA: Addison-Wesley, 1995.

[4] S.M. Alamouti, "A simple transmit diversity scheme for wireless communications," IEEE J. Selected. Areas Commu., vol 16, no. 8, pp. 1451-1458, Oct. 1998.

[5] Vahid Tarokh, Hamid Jafarkhani and A. Robert Calderbank, "Space time block coding for wireless communication: performance result," IEEE J. Select Areas Commun., vol. 17, pp. 451-460, Mar. 1999.

[6] KAMRUZZAMAN, M., LI, H.. Performance of Turbo coded OFDM wireless link for SISO, SIMO, MISO and MIMO system. Journal of Communications, North America, 7, nov. 2012.

[7] ZAMAN, M., HAO, L.. Performance of Turbo-SISO, Turbo-SIMO, Turbo-MISO and Turbo-MIMO system using STBC. Journal of Communications, North America, 6, nov. 2011

[8] Kamruzzaman, M.M., "Performance of Turbo coded wireless link for SISO-OFDM, SIMO-OFDM, MISO-OFDM and MIMO-OFDM system," Computer and Information Technology (ICCIT), 2011 14th International Conference on, vol., no., pp.185-190, 22.

[9] Kamruzzaman, M.M., "Performance of Turbo Coded Wireless Link for SISO, SIMO, MISO and MIMO System," Multimedia Information Networking and Security (MINES), 2012 Fourth International Conference on , vol., no., pp.187,190, 2-4 Nov. 2012.

[10] Kamruzzaman, M.M., "Performance of Turbo Coded Wireless Link for SIMO Using SC EGC and MRC," Multimedia Information Networking and Security (MINES), 2012 Fourth International Conference on , vol., no., pp.191,194, 2-4 Nov. 2012.

[11] M.M.Kamruzzaman, Dr. Mir Mohammad Azad, "Single Input Multiple Output (SIMO) Wireless Link with Turbo Coding", International Journal of Advanced Computer Science and Applications, Vol. 1, No. 5, pp. 69-73, November 2010.

[12] Kamruzzaman, M. M.; Hao, Li, "Performance of Turbo-VBLAST Coded Wireless Link for Multiple-Input Multiple-Output-Orthogonal Frequency Division Multiplexing System" Sensor Letters, Volume 10, Number 8, December 2012 , pp. 1911-1917(7)

[13] P. A. Anghel and M. Kaveh, "Exact symbol error probability of a cooperative network in a Rayleigh-fading environment," IEEE Trans. Wireless Commun., vol. 3, pp 1416–1421, Sept. 2004.

[14] Ribeiro, X. Cai, and G. B. Giannakis, "Symbol error probabilities for general cooperative links," IEEE Trans. Wireless Commun., vol. 4, pp. 1264–1273, May 2005.

[15] Y.-W. Peter Hong, Wan-Jen Huang And C.-C. Jay Kuo, Cooperative Communications and Networking technologies and system design, Springer . 2010.

[16] Kamruzzaman, M.M., " Performance of decode and forward MIMO relaying for wireless uplink," Computer and Information Technology (ICCIT), 2012 15th International Conference on , vol., no., pp.321,325, 22-24 Dec. 2012.

[17] Min Chen; Yener, A., "Multiuser two-way relaying: detection and interference management strategies," Wireless Communications, IEEE Transactions on , vol.8, no.8, pp.4296,4305, August 2009

[18] Nan Yang; Elkashlan, M.; Yeoh, P.L.; Jinhong Yuan, "Multiuser MIMO Relay Networks in Nakagami-m Fading Channels," Communications, IEEE Transactions on , vol.60, no.11, pp.3298,3310, November 2012

[19] Zhu Han; Xin Zhang; Poor, H.V., "High performance cooperative transmission protocols based on multiuser detection and network coding," Wireless Communications, IEEE Transactions on , vol.8, no.5, pp.2352,2361, May 2009

[20] Umehara, D.; Huang, C.; Denno, S.; Morikura, M.; Sugiyama, T., "Enhancement of IEEE 802.11 and network coding for single-relay multi-user wireless networks," Signal Processing and Communication Systems (ICSPCS), 2010 4th International Conference on , vol., no., pp.1,9, 13-15 Dec. 2010

[21] Chao Wang; Ming Xiao; Skoglund, M., "Efficient Multiple Access Protocols for Coded Multi-Source Multi-Relay Networks," Communications (ICC), 2011 IEEE International Conference on , vol., no., pp.1,6, 5-9 June 2011

[22] de Melo, Marco Antonio Beserra; da Costa, Daniel Benevides, "Downlink performance of multiuser multi-relay cooperative networks: A low-complexity relay-destination selection scheme," European Wireless, 2012. EW. 18th European Wireless Conference , vol., no., pp.1,7, 18-20 April 2012

[23] Jian Tang; Mumey, B.; Zhubayev, K.; Wolff, R.S., "On Exploiting Cooperative, Channel and Multiuser Diversities in Wireless Relay Networks," Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE , vol., no., pp.1,6, 6-10 Dec. 2010

[24] Liang Yang; Weiping Liu, "On the Throughput of MIMO Relay Wireless Network with Receive Antenna Selection," Communications Letters, IEEE , vol.15, no.6, pp.626,628, June 2011

[25] Guo, W.; Liu, J.; Zheng, L.; Xu, H.; Liu, Y.; Li, Y., "Capacity study of multiuser cooperative networks with fixed-gain relays," Networks, IET , vol.1, no.3, pp.101,107, Sept. 2012

[26] Liangbin Li; Yindi Jing; Jafarkhani, Hamid, "Interference Cancellation at the Relay in Two User Wireless Relay Networks," Wireless Communications and Networking Conference (WCNC), 2010 IEEE , vol., no., pp.1,6, 18-21 April 2010

[27] Di Wang; Xin Wang; Xiaodong Cai, "Optimal Power Control for Multi-User Relay Networks over Fading Channels," Wireless Communications, IEEE Transactions on , vol.10, no.1, pp.199,207, January 2011

[28] Xing Zhang; Chao Yang; Shuping Chen; Yong Li; Wenbo Wang, "Performance Analysis of Multiuser Diversity in Multiuser Two-Hop Cooperative Relay Wireless Networks," Communications Workshops, 2008. ICC Workshops '08. IEEE International Conference on , vol., no., pp.337,341, 19-23 May 2008

[29] Xiaoyi Liu; Xin Zhang; Dacheng Yang, "Outage Probability Analysis of Multiuser Amplify-and-Forward Relay Network with the Source-to-Destination Links," Communications Letters, IEEE , vol.15, no.2, pp.202,204, February 2011

[30] Nan Yang; Elkashlan, M.; Jinhong Yuan, "Outage Probability of Multiuser Relay Networks in Nakagami Fading Channels," Vehicular Technology, IEEE Transactions on , vol.59, no.5, pp.2120,2132, Jun 2010

[31] Shuping Chen; Wenbo Wang; Xing Zhang, "Performance analysis of multiuser diversity in cooperative multi-relay networks under rayleigh-fading channels," Wireless Communications, IEEE Transactions on , vol.8, no.7, pp.3415,3419, July 2009

[32] Guangping Li; Blostein, S.D.; Jiuchao Feng, "Performance Analysis of OSTBC Transmission in Multiuser Multiantenna Relay Networks," Vehicular Technology, IEEE Transactions on , vol.62, no.1, pp.421,427, Jan. 2013

[33] Aboutorab, N.; Hardjawana, W.; Vucetic, B., "A New Resource Allocation Technique in MU-MIMO Relay Networks," Vehicular Technology, IEEE Transactions on , vol.60, no.7, pp.3485,3490, Sept. 2011

[34] Jongyeol Ryu; Wan Choi, "A simple linear multiuser precoding technique in cellular relay networks," Communications Letters, IEEE ,vol.14, no.1, pp.12,14, January 2010

[35] Li, Jianing; Jianhua Zhang; Zhang, Yu; Ping Zhang, "A comparison of TDMA, dirty paper coding, and beamforming for multiuser MIMO relay networks," Communications and Networks, Journal of , vol.10, no.2, pp.186,193, June 2008

[36] Hyun Jong Yang; Youngchol Choi; Namyoon Lee; Paulraj, A., "Achievable Sum-Rate of MU-MIMO Cellular Two-Way Relay Channels: Lattice Code-Aided Linear Precoding," Selected Areas in Communications, IEEE Journal on , vol.30, no.8, pp.1304,1318, September 2012

[37] Chang-Kyung Sung; Collings, I.B., "Multiuser Cooperative Multiplexing with Interference Suppression in Wireless Relay Networks," Wireless Communications, IEEE Transactions on , vol.9, no.8, pp.2528,2538, August 2010

[38] Jian Tang; Mumey, B.; Zhubayev, K.; Wolff, R.S., "Leveraging Cooperative, Channel and Multiuser Diversities for Efficient Resource Allocation in Wireless Relay Networks," Selected Areas in Communications, IEEE Journal on , vol.30, no.9, pp.1789,1797, October 2012

[39] Venturino, L.; Xiaodong Wang; Lops, M., "Multiuser detection for cooperative networks and performance analysis," Signal Processing, IEEE Transactions on , vol.54, no.9, pp.3315,3329, Sept. 2006

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Effect on Performance of Wireless Uplink for Placing Decode and Forward MIMO Relay At Different Position Between Source and Destination

M.M.Kamruzzaman
Key Lab of Information Coding & Transmission,
Southwest Jiaotong University, Chengdu, Sichuan, China.
E-mail: m.m.kamruzzaman@gmail.com

*Abstract*— **This paper investigates the effect on performance for placing a decode and forward MIMO relay at different position between source and destination for wireless uplink in the presence of rayleigh fading where source is equipped with single transmit antenna, relay is equipped with multiple transmit and receive antennas, and destination has multiple receive antennas. It is observed that relay at 0.4 and/or 0.5 and/or 0.6 provides best performance. On the other hand, relay at 0.2 and/or 0.8 shows the worst performance. That is, the system shows the best performance placing relay at the middle of the source and the destination, or slightly closer to middle. And the system shows the worse performance placing relay close to the source or the destination. It is also observed that relay at 0.4 or 0.5 or 0.6 provides 5 dB to 12 dB gains at $10^{-5}$ compared to relay at 0.2 or 0.8.**

***Keywords—Space Time Block Code; MIMO; Relay; Decode and Forward; Uplink; Wireless Communication.***

## I. INTRODUCTION

Wireless relaying techniques have been receiving a considerable attention for providing significant improvements on coverage and performance. In wireless relay networks, signals are transmitted from source to destination through a number of relay. Relay accepts the signal from source or/and from other relays, improves the qualities of signal and retransmits it to the destination. There are mainly two types of relays: Amplify and Forward (AF) and Decode and Forward (DF). AF simply amplifies the incoming signal and forwards it to the destination without any attempt to decode it. AF relay is easy to implement but cannot achieve high performance gain. On the other hand, DF decodes the incoming signal, re-encodes it, and then retransmits it to the destination. Although the complexity of DF is high but can obtain high performance gain [1]. So we have used DF to show the performance of our system.

On the other hand, multiple transmit and/or receive antennas promises very high data rates on a scattering-rich wireless channel, especially when propagation environment or channel is known at the receiver. Two types of transmission techniques are used in the MIMO system. One is transmit diversity in which different duplicates of the same transmission sequence are transmitted through different transmission antennas. One of the typical transmit diversity technique is Space Time Block Coding (STBC) [2-7]. The other one is transmitting multiplexing in which different transmission sequences are transmitted through different transmission antennas. The typical transmitting multiplexing technique is Bell Labs Layered Space-Time (BLAST)[8,9]. In this paper we have used STBC to show the performance of our system.

Relay assisted wireless communication has been widely studied [10-42]. [10-13] show the performance of relay using amplifying and forwarding. [14-16] show the performance of relay having difference time slot to transmit information. [17-25] show the performance of relay having single antenna at source, relay and destination. [26-28] show the performance of relay having single antenna at source and destination but multiple antennas at relay. [29-42] show the performance of relay having multiple antennas at source, relay and destination. Our previous paper investigates the performance of relay assisted uplink wireless communication by placing relay only at middle of source and destination where source is equipped with single transmit antenna, relay is equipped with multiple transmit and receive antennas, and destination is equipped with multiple receive antennas [43]. This paper investigates the effect on performance for placing relay at different position with the same configuration of our previous paper because the transmission power saving can be optimized by placing relay at different location.

## II. SYSTEM MODEL

It is considered that the system consists of source, relay and destination where source is equipped with single transmitte antenna and relay is equipped with multiple transmit antennas and multiple receive antennas, and destination is equipped with multiple receiving antennas. Data are modulated by a QPSK or 16 QAM or 64 QAM modulator before transmitting as shown in Fig.1.

### *A. Received Signal at Relay*

If the symbol transmitted by transmit antenna is $s$, then the signal received by the receiving antennas at relay can be written as:

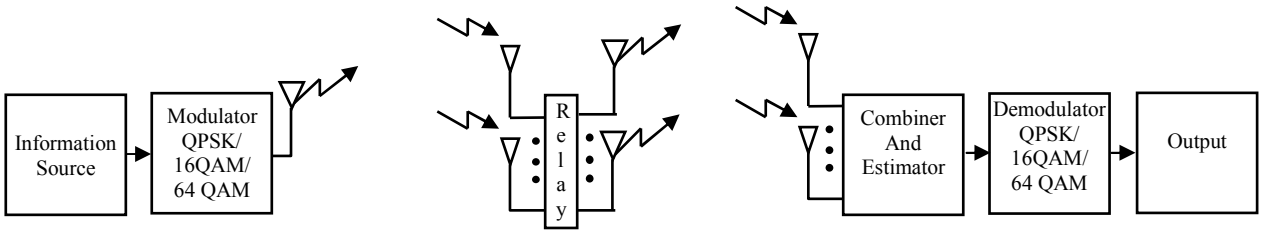



Fig. 1. System Block diagram

$$r_i = h_i s + n_i \quad (1)$$

where $r_i$ is the received symbol on the ith receiver antenna of relay.

$h_i$ is the channel on the i[th] receive antenna. It is assumed that the channel experienced by each receive antenna is independent from the channel experienced by other receive antennas. And the channel $h_i$ is know at the receiver.

$s$ is the transmitted symbol and

$n_i$ is the noise on i[th] receive antenna. It is assumed that the noise on each receive antenna is independent from the noise on the other receive antennas.

The combiner combines received signals which are then sent to the maximum likelihood detector. The combiner generates the following signals [44-45]:

$$\tilde{s} = h_1^* r_1 + h_2 r_2^* + h_3^* r_3 + h_4 r_4^* \text{ ------ } h_{n-1}^* r_{n-1} + h_n r_n^* \quad (2)$$

Maximum likelihood decoding of combined signal $\tilde{s}$ can be achieved using the decision metric:

$$\sum_{i=2}^{n} \left( \left| r_{i-1} - h_{i-1} s \right|^2 + \left| r_i - h_i s^* \right|^2 \right) \quad (3)$$

over all possible values of $s$ .

We expand the above equation and delete the terms that are independent of the code words. So the above can be rewrite:

$$-\sum_{i=2}^{n} \left[ r_{i-1} h_{i-1}^* s^* + r_i^* h_i s \right] + \left| s \right|^2 \sum_{i=1}^{n} \left| h_i \right|^2 \quad (4)$$

For detecting $s$ at relay, we use the following decision metric.

TABLE I. THE ENCODING AND TRANSMISSION SEQUENCE FOR TWO TRANSMISSION ANTENNAS OF ALAMOUTI WITH CODE RATE ONE

| | **Antenna-I** | **Antenna-II** |
|---|---|---|
| **Time slot-I** | $x_1$ | $x_2$ |
| **Time slot-II** | $-x_2^*$ | $x_1^*$ |

$$\left\| \left[ \sum_{i=2}^{n} \left( r_{i-1} h_{i-1}^* + r_i^* h_i \right) \right] - s \right|^2 + \left( -1 + \sum_{i=1}^{n} \left| h_i \right|^2 \right) \left| s \right|^2 \quad (5)$$

*B.* Encoding using STBC and retransmitting from relay:

STBC encoder encodes the decoded symbols of relay according to the Table I and then at each time slot $t$, signals $S_t^i$ , $i$ =1, 2, … $n$ are transmitted simultaneously using $n$ transmit antennas.

*C.* Received signal at destination:

At time $t$ the signal $y_t^j$ , received at antenna $j$ of destination, is given by

$$y_t^j = \sum_{i=1}^{n} \alpha_{i,j} S_t^i + \eta_t^j \quad (6)$$

where $y_t^j$ is the received symbol on the jth receiver antenna of destination at time t.

$\alpha_{i,j}$ is the channel from transmit antenna $i$ to receive antenna $j$.

$S_t^i$ is the transmitted symbol from transmit antenna $i$ at each time slot $t$.

$\eta_t^j$ is the noise from transmit antenna $i$ to receive antenna $j$.

The combiner combines received signals of destination which are then sent to the maximum likelihood detector. For detecting symbols $s_1$ and $s_2$ , (7) and (8) decision metrics have been used [2, 3]:

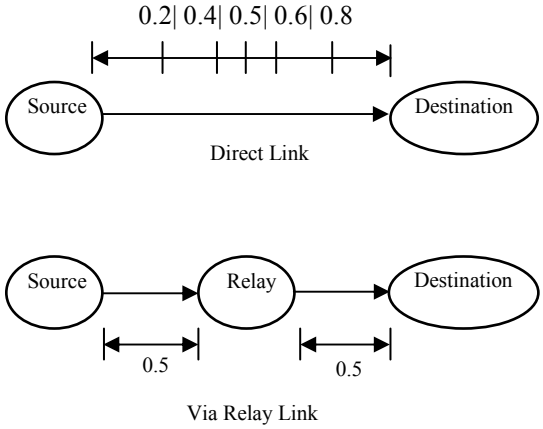



Fig. 2. Direct Link and Via Relay Link

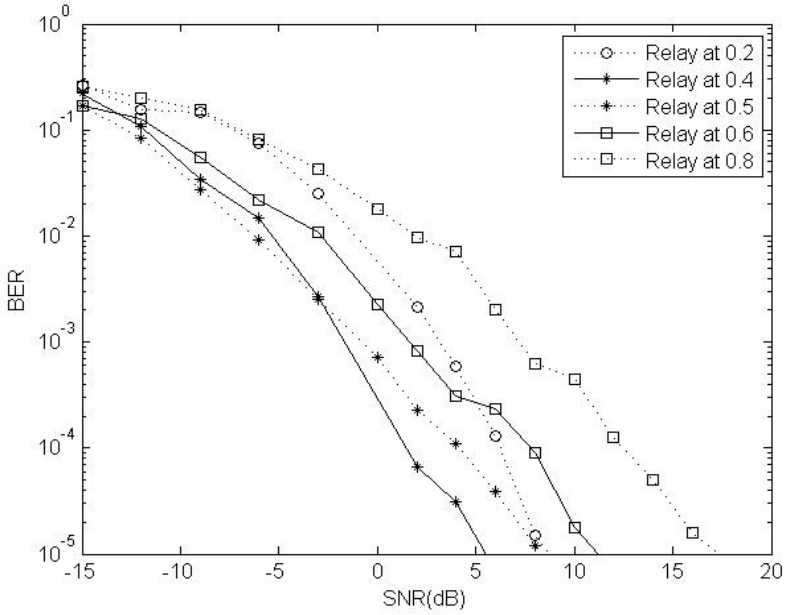



Fig. 3. BER performance comparisons of wireless uplink for placing relay at different position where relay is equipped with $2T_x$ & $2R_x$ and destination is equipped with 2 Rx

$$\left[\left[\sum_{j=1}^{m}\left(r_1^j\alpha_{1,j}^{*}+\left(r_2^j\right)^{*}\alpha_{2,j}\right)\right]-s_1\right|^2 \qquad (7)$$
$$+\left(-1+\sum_{j=1}^{m}\sum_{i=1}^{2}\left|\alpha_{i,j}\right|^2\right)\left|s_1\right|^2$$

$$\left[\left[\sum_{j=1}^{m}\left(r_1^j\alpha_{2,j}^{*}-\left(r_2^j\right)^{*}\alpha_{1,j}\right)\right]-s_2\right|^2 \qquad (8)$$
$$+\left(-1+\sum_{j=1}^{m}\sum_{i=1}^{2}\left|\alpha_{i,j}\right|^2\right)\left|s_1\right|^2$$

The detected symbols are demodulated by 64QAM demodulator to get the output.

## III. Simulation Results

In this section, computer simulation is carried out to show the BER performance of the proposed system. The results are evaluated for several combinations of $T_x$ and $R_x$ antennas with relay placing at different position (at 0.2, 0.4, 0.5, 0.6 and 0.8) as shown in fig. 2. 64 QAM is used for simulation.

Fig. 3 shows the performance of wireless uplink for placing relay at different position (at 0.2, 0.4, 0.5, 0.6 and 0.8) where relay is equipped with 2Tx & 2Rx and destination is equipped with 2 Rx. It is observed that relay at 0.8 shows the

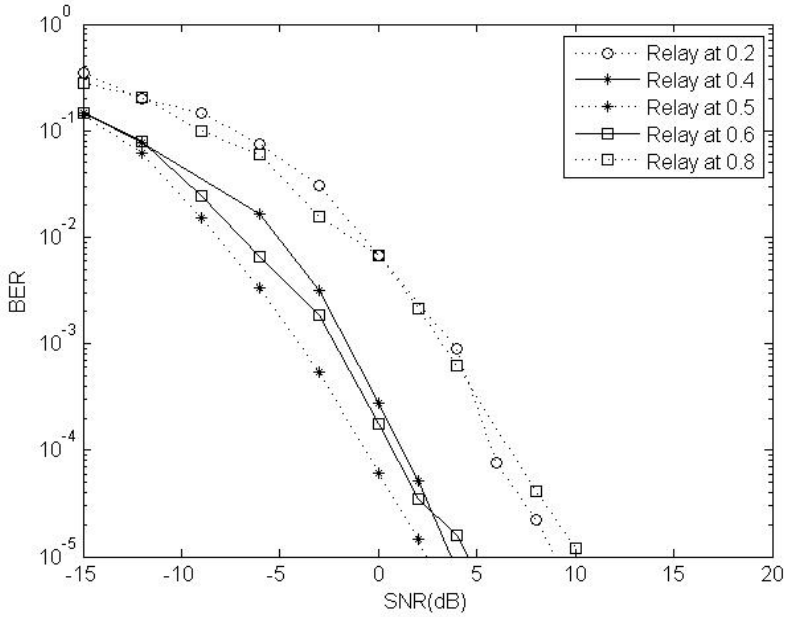



Fig. 4. BER performance comparisons of wireless uplink for placing relay at different position where relay is equipped with $2T_x$ & $3R_x$ and destination is equipped with 2 Rx

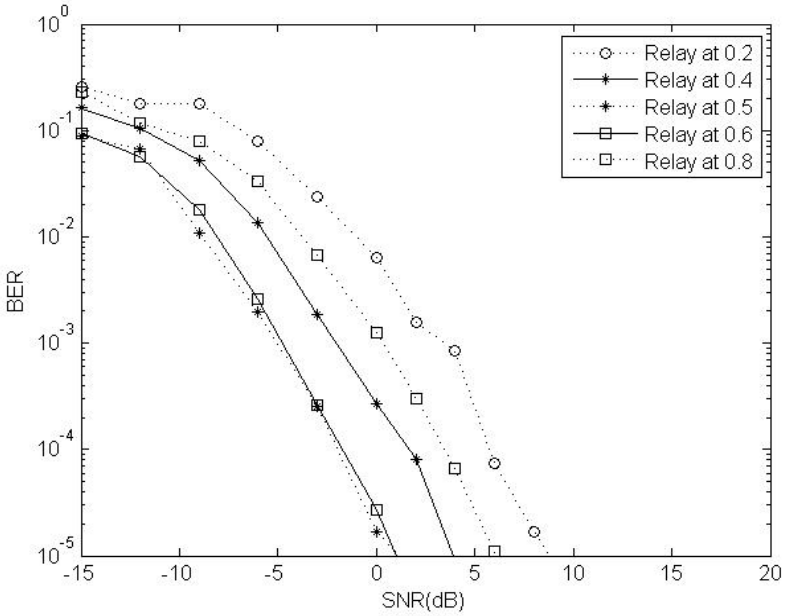



Fig. 5. BER performance comparisons of wireless uplink for placing relay at different position where relay is equipped with $2T_x$ & $4R_x$ and destination is equipped with 2 Rx

worst performance. Relay at 0.6, 0.5, 0.4, 0.2 provides 7 dB, 10 dB, 12 dB and 6 dB gain at $10^{-5}$ respectively compared to relay at 0.8. And relay at 0.4 shows the best performance.

Fig. 4 shows the performance of wireless uplink for placing relay at different position (at 0.2, 0.4, 0.5, 0.6 and 0.8) where relay is equipped with 2Tx & 3Rx and destination is equipped with 2 Rx .It is observed that relay at 0.8 and 0.2 shows the worst performance. Relay at 0.6, 0.5 and 0.4 provides 5 db,7 dB, 6 dB gain at $10^{-5}$ respectively compared to relay at 0.8 or 0.2. And relay at 0.5 shows the best performance.

Fig. 5 shows the performance of wireless uplink for placing relay at different position (at 0.2, 0.4, 0.5, 0.6 and 0.8) where relay is equipped with 2Tx & 4Rx and destination is equipped with 2 Rx .It is observed that relay at 0.2 shows the worst performance. Relay at 0.8, 0.6, 0.5 and 0.4 provides 2 dB, 7 dB, 7 dB and 5 dB gain at $10^{-5}$ respectively compared to relay at 0.2. And relay at 0.5 and 0.6 shows the best performance.

Fig. 6 shows the performance of wireless uplink for placing relay at different position (at 0.2, 0.4, 0.5, 0.6 and 0.8) where relay is equipped with 2Tx & 5Rx and destination is equipped with 2 Rx .It is observed that relay at 0.2 shows the worst performance. Relay at 0.8, 0.6, 0.5 and 0.4 provides 3 dB, 8 dB, 8 dB and 5 dB gains at $10^{-5}$ respectively compared to relay at 0.2. And relay at 0.5 and 0.6 shows the best performance.

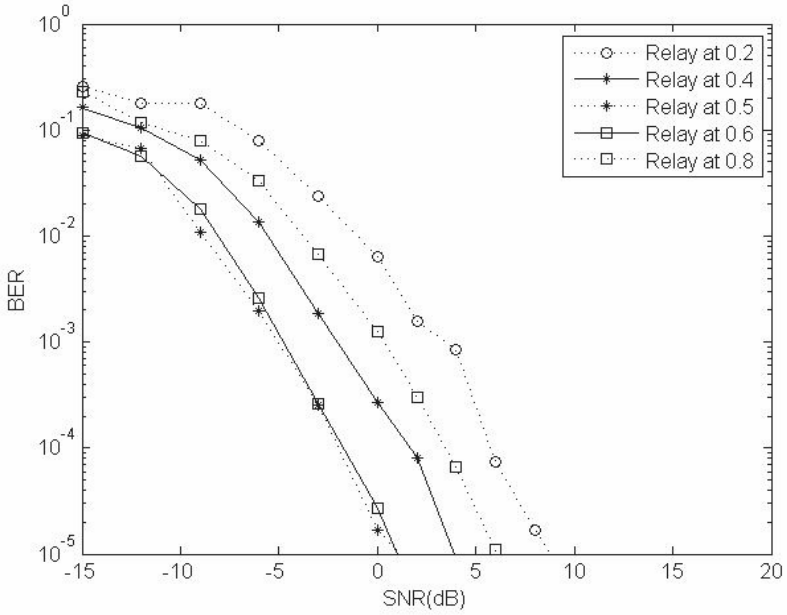



Fig. 6. BER performance comparisons of wireless uplink for placing relay at different position where relay is equipped with $2T_x$ & $5R_x$ and destination is equipped with 2 Rx

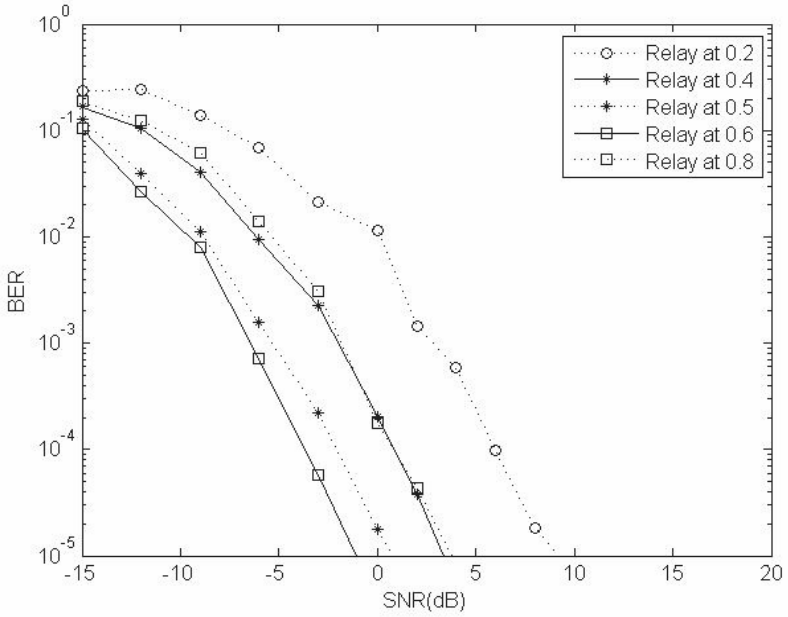



Fig. 7. BER performance comparisons of wireless uplink for placing relay at different position where relay is equipped with $2T_x$ & $6R_x$ and destination is equipped with 2 Rx

Fig. 7 shows the performance of wireless uplink for placing relay at different position (at 0.2, 0.4, 0.5, 0.6 and 0.8) where relay is equipped with 2Tx & 6Rx and destination is equipped with 2 Rx .It is observed that relay at 0.2 shows the worst performance. Relay at 0.8, 0.6, 0.5 and 0.4 provides 5 dB, 10 dB, 8 dB and 6 dB gains at $10^{-5}$ respectively compared to relay at 0.2. And relay at 0.6 shows the best performance.

## IV. CONCLUSION

From the simulations results, it is observed that if the relay is placed at the middle of the source and the destination, or slightly closer to middle, the system shows the best performance. On the other hand, if the relay is near to source or near to destination, the system shows the worst performance. So the preferred position of the relay is the middle between the source and destination or slightly closer to middle.

## ACKNOWLEDGMENT

The authors would like to thank the reviewers for the suggestions which help to improve the quality of this paper. In addition, the authors are also very thankful to Key Lab of Information Coding & Transmission, Southwest Jiaotong University, Chengdu, Sichuan, China for providing resources.

## REFERENCES

[1] Y.-W. Peter Hong, Wan-Jen Huang And C.-C. Jay Kuo, Cooperative Communications and Networking technologies and system design, Springer . 2010.

[2] S.M. Alamouti, " A simple transmit diversity scheme for wireless communications," IEEE J. Selected. Areas Commu., vol 16, no. 8, pp. 1451-1458, Oct. 1998.

[3] Vahid Tarokh, Hamid Jafarkhani and A. Robert Calderbank, " Space time block coding for wireless communication: performance result," IEEE J. Select Areas Commun., vol. 17, pp. 451-460, Mar. 1999.

[4] Kamruzzaman, M.M., "Performance of Turbo Coded Wireless Link for SISO, SIMO, MISO and MIMO System," Multimedia Information Networking and Security (MINES), 2012 Fourth International Conference on , vol., no., pp.187,190, 2-4 Nov. 2012

[5] ZAMAN, M., HAO, L.. Performance of Turbo-SISO, Turbo-SIMO, Turbo-MISO and Turbo-MIMO system using STBC. Journal of Communications, North America, 6, nov. 2011 .

[6] KAMRUZZAMAN, M., LI, H.. Performance of Turbo coded OFDM wireless link for SISO, SIMO, MISO and MIMO system. Journal of Communications, North America, 7, nov. 2012.

[7] Kamruzzaman, M.M., "Performance of Turbo coded wireless link for SISO-OFDM, SIMO-OFDM, MISO-OFDM and MIMO-OFDM system," Computer and Information Technology (ICCIT), 2011 14th International Conference on, vol., no., pp.185-190, 22.

[8] Foschini, Gerard J. "Layered space time architecture for wireless communication in a fading environment when using multi‐element antennas."Bell labs technical journal 1.2 (1996): 41-59.

[9] Kamruzzaman, M. M.; Hao, Li, "Performance of Turbo-VBLAST Coded Wireless Link for Multiple-Input Multiple-Output-Orthogonal Frequency Division Multiplexing System" Sensor Letters, Volume 10, Number 8, December 2012 , pp. 1911-1917(7)

[10] J. N. Laneman, D. N. Tse and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behaviour," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[11] Canpolat, O..; Uysal, M..; Fareed, M.M.; , "Analysis and Design of Distributed Space–Time Trellis Codes With Amplify-and-Forward Relaying," *Vehicular Technology, IEEE Transactions on* , vol.56, no.4, pp.1649-1660, July 2007

[12] Berger, S.; Kuhn, M.; Wittneben, A.; Unger, T.; Klein, A.; , "Recent advances in amplify-and-forward two-hop relaying," *Communications Magazine, IEEE* , vol.47, no.7, pp.50-56, July 2009

[13] Abdaoui, A.; Ikki, S.S.; Ahmed, M.H.; , "Performance Analysis of MIMO Cooperative Relaying System Based on Alamouti STBC and Amplify-and-Forward Schemes," *Communications (ICC), 2010 IEEE International Conference on* , vol., no., pp.1-6, 23-27 May 2010

[14] Vien, N.H.; Nguyen, H.H.; Le-Ngoc, T.; , "Diversity analysis of smart relaying over Nakagami and Hoyt generalised fading channels," *Communications, IET* , vol.3, no.11, pp.1778-1789, November 2009

[15] Heesun Park; Joohwan Chun; , "Alternate Transmission Relaying Schemes for MIMO Wireless Networks," *Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE* , vol., no., pp.1073-1078, March 31 2008-April 3 2008

[16] Ho Van Khuong; Tho Le-Ngoc; , "A bandwidth-efficient cooperative relaying scheme with space-time block coding and iterative decoding," *Communications and Electronics, 2008. ICCE 2008.*

*Second International Conference on* , vol., no., pp.262-267, 4-6 June 2008

[17] Janani, M.; Hedayat, A.; Hunter, T.E.; Nosratinia, A.; , "Coded cooperation in wireless communications: space-time transmission and iterative decoding," *Signal Processing, IEEE Transactions on* , vol.52, no.2, pp. 362- 371, Feb. 2004

[18] Abouei, J.; Bagheri, H.; Khandani, A.; , "An efficient adaptive distributed space-time coding scheme for cooperative relaying," *Wireless Communications, IEEE Transactions on* , vol.8, no.10, pp.4957-4962, October 2009

[19] Zhang, C.; Zhang, J.; Yin, H.; Wei, G.; , "Selective relaying schemes for distributed space-time coded regenerative relay networks," *Communications, IET* , vol.4, no.8, pp.967-979, May 21 2010

[20] Tourki, K.; Alouini, M.-S.; Deneire, L.; , "Blind Cooperative Diversity Using Distributed Space-Time Coding in Block Fading Channels," Communications, 2008. ICC '08. IEEE International Conference on , vol., no., pp.4596-4600, 19-23 May 2008

[21] Duong, T.Q.; Alexandropoulos, G.C.; Zepernick, H.; Tsiftsis, T.A.; , "Orthogonal Space–Time Block Codes With CSI-Assisted Amplify-and-Forward Relaying in Correlated Nakagami- m Fading Channels," Vehicular Technology, IEEE Transactions on , vol.60, no.3, pp.882-889, March 2011

[22] Feng Tian; Wei Zhang; Wing-Kin Ma; Ching, P.C.; , "Distributed Space-Time Coding for Two-Path Successive Relaying," Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE , vol., no., pp.1-5, 6-10 Dec. 2010

[23] Torabi, M.; Haccoun, D.; , "Performance analysis of cooperative diversity systems with opportunistic relaying and adaptive transmission," *Communications, IET* , vol.5, no.3, pp.264-273, Feb. 11 2011

[24] Le Quang Vinh Tran; Berder, O.; Sentieys, O.; , "Non-regenerative full distributed space-time codes in cooperative relaying networks," Wireless Communications and Networking Conference (WCNC), 2011 IEEE , vol., no., pp.1529-1533, 28-31 March 2011

[25] Mimura, T.; Kuwabara, A.; Murata, H.; Yamamoto, K.; Yoshida, S.; , "Packet Transmission Experiments of STBC-Based Multi-Hop Cooperative Relaying," *Communications (ICC), 2011 IEEE International Conference on* , vol., no., pp.1-5, 5-9 June 2011

[26] Sharma, G.V.V.; Ganwani, V.; Desai, U.B.; Merchant, S.N.; , "Performance Analysis of Maximum Likelihood Detection for Decode and Forward MIMO Relay Channels in Rayleigh Fading," *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE* , vol., no., pp.1-6, 5-8 April 2009

[27] Miyano, T.; Murata, H.; Araki, K.; , "Cooperative relaying scheme with space time code for multihop communications among single antenna terminals," Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE , vol.6, no., pp. 3763- 3767 Vol.6, 29 Nov.-3 Dec. 2004

[28] Mheidat, H.; Uysal, M.; , "Space-Time Coded Cooperative Diversity with Multiple-Antenna Nodes," *Information Theory, 2007. CWIT '07. 10th Canadian Workshop on* , vol., no., pp.17-20, 6-8 June 2007

[29] Fan, Y.; Thompson, J.; , "MIMO Configurations for Relay Channels: Theory and Practice," Wireless Communications, IEEE Transactions on , vol.6, no.5, pp.1774-1786, May 2007

[30] Yang, Q.; Kwak, K.S.; , "Outage performance of cooperative relaying with dissimilar Nakagami-m interferers in Nakagami-m fading," Communications, IET , vol.3, no.7, pp.1179-1185, July 2009

[31] In-Ho Lee; Dongwoo Kim; , "Achieving maximum spatial diversity with decouple-and-forward relaying in dual-hop OSTBC transmissions," Wireless Communications, IEEE Transactions on , vol.9, no.3, pp.921-925, March 2010

[32] Bastami, A.H.; Olfat, A.; , "Optimal SNR-based selection relaying scheme in multi-relay cooperative networks with distributed space-time coding," Communications, IET , vol.4, no.6, pp.619-630, April 16 2010

[33] Maham, B.; Hjorungnes, A.; , "Opportunistic relaying for space-time coded cooperation with multiple antennas terminals," Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on , vol., no., pp.251-255, 13-16 Sept. 2009

[34] Abdaoui, A.; Ikki, S.S.; Ahmed, M.H.; Cha□telet, E.; , "On the Performance Analysis of a MIMO-Relaying Scheme With Space–Time Block Codes," Vehicular Technology, IEEE Transactions on , vol.59, no.7, pp.3604-3609, Sept. 2010

[35] Van Khuong, H.; Le-Ngoc, T.; , "Performance analysis of a decode-and-forward cooperative relaying scheme for MIMO systems," Communications (QBSC), 2010 25th Biennial Symposium on , vol., no., pp.400-403, 12-14 May 2010

[36] Dharmawansa, Prathapasinghe; McKay, Matthew R.; Mallik, Ranjan K.; , "Analytical Performance of Amplify-and-Forward MIMO Relaying with Orthogonal Space-Time Block Codes," *Communications, IEEE Transactions on* , vol.58, no.7, pp.2147-2158, July 2010

[37] Liping Su; Dong Chen; Tong Wu; Jing Huang; , "Efficient power allocation schemes for MIMO cooperative relaying systems," *Computer Application and System Modeling (ICCASM), 2010 International Conference on*, vol.14, no., pp.V14-179-V14-183, 22-24 Oct. 2010

[38] Chao Zhao; Champagne, B.; , "Non-regenerative MIMO relaying strategies — from single to multiple cooperative relays," *Wireless Communications and Signal Processing (WCSP), 2010 International Conference on* , vol., no., pp.1-6, 21-23 Oct. 2010

[39] Zaman, N.I.; Al Mamun, K.M.A.; Uddin, A.M.M.S.; Kanakis, T.; , "Performance Analysis of MIMO Relaying Channels' Cooperative Diversity for Mobile Networks," *Data, Privacy and E-Commerce (ISDPE), 2010 Second International Symposium on* , vol., no., pp.95-99, 13-14 Sept. 2010

[40] Nguyen Thi Be Tam; Thanh Tran-Thien; Tuan Do-Hong; Vo Nguyen Quoc Bao; , "Performance analysis of decode-and-forward relaying for multi-hop Alamouti transmission over Rayleigh fading channels," *Advanced Technologies for Communications (ATC), 2010 International Conference on* , vol., no., pp.195-200, 20-22 Oct. 2010

[41] Yang, C.; Wang, W.; Chen, S.; Peng, M.; , "Outage performance of orthogonal space-time block codes transmission in opportunistic decode-and-forward cooperative networks with incremental relaying," *Communications, IET* , vol.5, no.1, pp.61-70, January 2011

[42] Lei Cao; Li Chen; Xin Zhang; Dacheng Yang; , "Cooperative diversity with OSTBC transmission and adaptive-gain amplify-and-forward MIMO relaying," *GLOBECOM Workshops (GC Wkshps), 2010 IEEE* , vol., no., pp.115-119, 6-10 Dec. 2010.

[43] Kamruzzaman, M.M., " Performance of decode and forward MIMO relaying for wireless uplink," Computer and Information Technology (ICCIT), 2012 15th International Conference on , vol., no., pp.321,325, 22-24 Dec. 2012.

[44] M.M.Kamruzzaman, Dr. Mir Mohammad Azad, "Single Input Multiple Output (SIMO) Wireless Link with Turbo Coding", International Journal of Advanced Computer Science and Applications, Vol. 1, No. 5, pp. 69-73, November 2010.

[45] Kamruzzaman, M.M., "Performance of Turbo Coded Wireless Link for SIMO Using SC EGC and MRC," Multimedia Information Networking and Security (MINES), 2012 Fourth International Conference on , vol., no., pp.191,194, 2-4 Nov. 2012

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# An Integrated Online Courseware: A Step to Globalization

G. M. M. Bashir [1], A. S. M. L. Hoque[2], S. Majumder[1], B. Bepary [1], and K. Rani[1]
[1]Patuakhali Science and Technology University, Bangladesh
[2]Bangladesh University of Engineering and Technology, Bangladesh
murad98csekuet@yahoo.com, asmlatifulhoque@cse.buet.ac.bd, {sbmajumdercse07, bipashabepary, kakalicse}@gmail.com

*Abstract*—**With the rapid development of technology, e-learning pretends to be more useful and popular in education. E-learning can be used in or out of the classroom. Courseware is becoming a well known form of e-learning in present days. It helps to convert the conventional process of learning and teaching towards e-learning. Different educational organizations are providing their resources to all through Courseware. In this paper, authors attempt to give an overview of some of the renowned online courseware. The authors have performed a comparative analysis among those courseware and then summarize them in the form of table and tree so that their overall structure can be easily approximated. After studying the similarities and dissimilarities among various contents provided by various courseware, the authors have proposed a design approach for an integrated global online courseware which will combine all the facilities of all courseware into a single one. A methodology for the fulfillment of the proposed idea has also been proposed with proper demonstration by Venn diagram. If the integrated courseware can be developed, then both of the instructors and learners will be benefitted. Because of integrating all relevant contents to relevant courses, it will be more beneficial for the third world learners also. The completion of the proposed idea requires the permission from the concerned authority. But if the idea can draw the attention of the authority, then it will surely be one of the steps to globalization of learning especially for the self learners.**

***Keywords— Courseware, E-learning, Globalization, Venn Diagram*.**

## I. INTRODUCTION

New and innovative ways of learning and teaching are now being emerged through the improvements in web technologies and the increased influence of the World Wide Web. E-learning means the use of these ways in learning and teaching. Present e-learning system technologies and services enable actions that allow users to be self learners, actively participating in the on-line learning process. There are several forms of e-learning such as multimedia based f2f learning, computer-based training, web-based training, online learning etc. Several institutions use e-learning in the form of courseware. Courseware is an additional educational software material that is used as kits for teachers or trainers and as tutorials for students, for use with a self-learning or coach assisted program. There are several papers based on e-learning and courseware. L. Vicent *et al*. [1] conducted a survey upon the effects of multimedia lecture and traditional lecture. The result shows that multimedia processes for learning is better than traditional. But, bad functional technology disturbs multimedia process sometimes. F. Buendia *et al.* [2] proposed a system that allows students to send their system program using the web forms, execute them in a native operating system and receive their results and feedback information through web-browsers. L. Zhang *et al*. [3] propose an assessment model based on Bayesian networks, which assess learning status by knowledge map after absorbing and analyzing test results. A survey is conducted based on academic results and a questionnaire (MCQ) which shows that 1% student's attitude is negative that is, they are not satisfied. The authors have no way to know their mentality and what they actually require. N. Hoic-Bozic *et al.* [4] describes a blended learning approach to course design and implementation using a LMS named AHyCo, applied to the senior students in the undergraduate program in a Mathematics and Information Science major at the department of Information Science. But, it is appropriate only where student/ instructor ratio is lower. J. R. Galvão *et al.* [5] analyzed various courseware features that are available on the internet and propose a new model that is based on some of the technologies, such as computers and Telecommunication.

The paper is structured as follows. In Section II, definition and types of courseware are described with the present status of overall courseware. Their tree and table representations are described in Section III with a brief overview of our analyzed courseware. In this section, a comparative analysis is made according to dissimilarity. In Section IV, our proposed methodology for course design and development with application is explained. Benefits of the proposed idea and challenges to do this are given in Section V and VI respectively. Finally, the conclusions are summarized in Section VIII.

## II. CONCEPT OF COURSEWARE?

All human beings have the right to learn, improve, and progress. Educational opportunity is the way by which they can fulfill that right. In this century, most of the persons related with education require delivering and acquiring knowledge with the use of technology. Technology is one of the biggest reasons for courseware. A courseware can be: (1) Visual – "hands on" type of person; Need pictures, graphs, etc. to learn and apply material, (2) Non-Visual - can apply learned material from just reading it. We can also characterize a courseware in the following sense: (1) On CD-ROMs – visual and customizable, (2) Books and Text – non-visual and very standardized, (3) On the internet – visual and accessible [6]. We are about to integrate the third type of courseware. There are different types of courseware user such as students,

educator, self-learner etc. The MIT open courseware conducts a survey based on who use their courseware and represents the result in a pie chart [7] where the chart shows that the maximum number (43%) of courseware users is self learners whereas the percentage of using courseware by educators, students and others are 9%, 42% and 6% respectively. So, we are concerned about designing an integrated courseware which will help all of the self learners. For a knowledge seeker, it will be time consuming to search all the individual courseware to acquire vast knowledge about a desired topic. But, if all the contents can be supplied by only a single courseware, then all the users will be benefitted.

We have analyzed several online courseware. Most of them are for undergraduate students. There are only a few for the post graduate students. Courseware for primary level students is even less. Khan academy [8] is one of this. No courseware is present for disable students. Different courseware have their own structure (not following any general structure). Courseware of different universities uses different books of different writers for providing contents to a course. There is an absence of content clustering technique for course material. For example, a course named 'Database System' may be taught from the book written by Korth or from the book written by Jeffrey Ullman or so on. Several books may contain several new topics and new methods for making a topic understandable on the same subject. In this case, existing courseware do not provide all topics from all books. But content clustering is very much necessary for self-learners. Though, there exist some attempts to integrate courseware, such as edx [9], open courseware consortium [10] etc. But, edx, combining different courseware, provides only some special courses. They also do not combine all courseware components, which may also be needed for a learner. They have combined courseware from several renowned universities such as MIT, Havard University, Berkely (University of California), The University of Texas System. So, here we do not get all courses in a single one. On the other hand, in open courseware consortium, they also provide some link of courseware for different courses. So, as usual we need to visit all courseware sites for acquiring knowledge.

## III. A Comparative Analysis

We have analyzed 9 courseware provided by different reputed institutions. Different institutions present their courseware in different styles. All are designed to be appropriate in their own perspectives. MIT OpenCourseWare is a web-based publication of virtually all MIT courses (more than 2,000) content [11]. Open Yale Courses provide free and open access to a selection of introductory courses taught by distinguished teachers and scholars at Yale University [12]. Notre Dame OCW is a free and open educational resource provided by the Notre Dame University for faculty, students, and self-learners throughout the world [13]. Connexions is a place to view and share educational material made of small knowledge chunks called modules that can be organized as courses, books, reports, *etc.* [14]. Utah State OpenCourseWare is a collection of educational material used in the formal campus courses of Utah State University, and seeks to provide people around the world with an opportunity to access high quality learning opportunities [15]. Open Michigan, free learning from

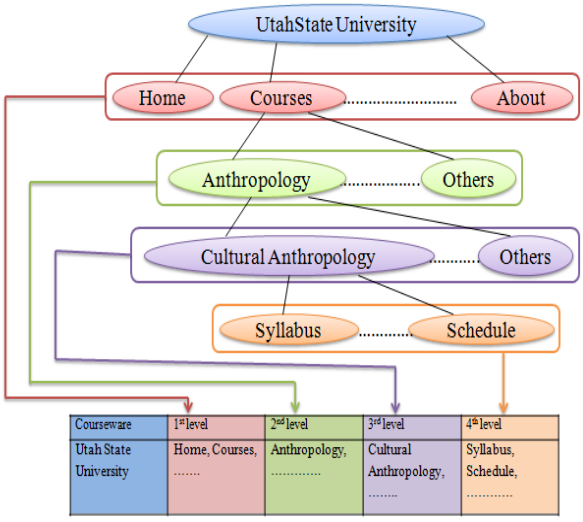



Figure 1. Overview of Courseware Provided by UtahState University

University of Michigan helps people find, use, and create openly licensed content and provides a space to share their educational content [16].

### *A. Representation*

After analyzing all courseware, we make a tree representation of all of them to make them structured. A tree helps to represent hierarchical data. So, the tree representation established by us will help to easily understand a courseware. A tree shows the navigation through a courseware at a glance. At first a tree representation of each courseware is made. When the tree is constructed, we can convert it to an equivalent table. The overview of courseware provided by UtahState University [15] as a table got from tree representation is represented in fig. 1. Table representation is important, because it will help us to compare different courseware on the basis of some common criteria. We have found that about 50% of the courseware don't fulfill sufficient contents for a specific course. Some courseware provides unnecessary level or specification for getting a course material. Though so much specification may be time consuming and boring, nevertheless it helps to find a course material quickly. MIT is a vast courseware. But here, if the users can search course by department, then the 'find by topic' portion seems to be not required especially for self learners (Table I). In Notre Dame and Utah State University, there is no easy option of searching, like --search by dept. But, Notre Dame University provides bibliography for each course which can be useful for a knowledge seeker. Open Yale University provides support for buying books. Connexions has two important search options –search by Author & search by Keywords. Some institutions provide online courses only for some specific field. As for example, JHSPH [17] provides courses related only health. University of Washington [18] and CS50 [19] provide courses related only computer science. Now, if one equivalent table can be obtained from one tree, then the table I can be found from observing other courseware.

In this table, similar components are kept in the same level, even though they are not the member of that level. For example, courseware from Open Michigan has *Find* in the first level, but it has been kept in the 2nd level in table I. In this table '…….' represents that there are more contents, which is not shown due to space congestion.

### *B. Percentage of Dissimilarity*

We make a mathematical analysis for extracting the dissimilarities among all courseware and express them in a percentage form. Every level of each courseware is compared for dissimilarity with every level of all other courseware. In the present context, the number of comparison will be (9!), because here 9 different courseware have been analyzed. To make this comparison, we first assign a weight for each level of the courseware. Weights of levels are assigned according to their importance. Last level is the most important level, we want to consider. Since, it provides the desired courses and course materials. We primarily work for first seven levels. Assigned weights are shown in table I. For example, 1st level=5 in the third column, here 5 represents the weight of 1st level. Obtained percentage on the basis of dissimilarity among courseware is given in the table II. For all diagonal elements, similarities will be 100%, because comparison occurs itself. So, dissimilarity is 0%. If we consider for ID (1, 2) that 1st level of ID 1 contains 6 elements and 1st level of ID 2 contains 5 elements. There are 2 elements which are common in both ID 1 and ID 2. So, Similarity = (2*5)/9=1.11. Here, 2 is the number of common elements, 5 is the assigned weight and 9(6+5-2) is the number of elements which is found by applying union operation between the 1st level elements of ID 1 and ID 2. If all the elements of these two IDs in this level were matched, then their matching was assigned as 5. If a less matching is found, then a matching<5 is provided. For 2nd, 3rd and 4th levels, ID 1 contains value, but ID 2 not. So we can say that for these three levels, similarity=0% and dissimilarity=100%. Similarly, the values for other levels are measured. Now, for ID (1, 2) total similar weight = 1.11+0 +0+0+0.1525+0.2376+3=4.5001. The values of other levels are also assumed. Now, percentage of weight= (4.5001*100)/56=8.04%. Here, 56 is the total weight (5+6+7+8+9+10+11=56). Now, percentage of dissimilarity is found by subtracting the obtained result from 100. The percentage of this example is found manually. This technique has been applied in the program with real data and we find the table II. After getting the results, we manually check some of the percentages. We would never claim that these percentages are correct, because no synonym table is used. The mathematical calculation is correct, but when we are concerned with the terms like true positive, false positive, true negative and false negative, the result is too much inaccurate. True positive is very weak, because without a synonym table, the program is unable to recognize that the course 'Basic Database' and 'Database System' are the same. So, the possibility of false negative is very high. Again, there may also be some words with same name but different meaning. For example, the spelling of a 'table' in database and a furniture 'table' are same, but their meanings are never the same. So, there may also have some false positive. Even though we have used a program to compute the percentage, many works have been done manually.

TABLE I. TABULAR REPRESENTATION OF DIFFERENT COURSEWARE

| ID | Courseware Origin | 1st level=5 | 2nd level=6 | 3rd level=7 | 4th level=8 | 5th level=9 | 6th level=10 | 7th level=11 | ……. nth level=n |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Massachusetts institute of technology | Courses ….. | Find course by | Topic, MIT courseNo, Dept | School of science, ----- | Biology-------,(6) | Genetics ----- | syllabus, lecture note, assignment, exam,------ | |
| 2 | Utah state University | Courses …… | | | | Anthropology,------- | Cultural anthropology | Syllabus, schedule, about professor,-- | |
| 3 | Notre dame University | Courses …… | Areas of study | | | Aerospace and mechanical engg., …… | Thermo-dynamics, ------- | lesson, bibliography, additional resources,--- | |
| 4 | Open Yale University | Courses …… | View all courses | Dept, course title, course no, ----- | | | Environmental politics and law, …… | Session, survey, buy books, …… | |
| 5 | Connexions | Content …….. | Search for content, Browse content | Sub, title, keyword, all collection ……. | Arts, Science ……. | | Arts and culture, …….. | E-mail, PDF, …… | |
| 6 | CS50 | | Search | Course name, .. | | | Mobile software engineering,---- | Problem set, quizzes,---- | |
| 7 | Open Michigan | Connect …….. | Find | | Engg, …….. | Chemical engg, …… | Process dynamics and controls , ……. | Overview, highlights, ……. | |
| 8 | JHSPH OCW | Course, topics, ..... | | | | Adolescent Health,… ……. | Child health & development, …….. | Syllabus, readings , ……. | |
| 9 | University of Washington | Courses | | | | | Computer programming …… | Lecture , homework , ---------- | |

TABLE II. PERCENTAGE OF DISSIMILARITIES

| ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 95.01 | 95.48 | 95.56 | 98.15 | 99.44 | 97.1 | 97.04 | 98.4 |
| 2 | 95.01 | 0 | 91.29 | 96.54 | 98.02 | 100 | 96.65 | 96.26 | 99.11 |
| 3 | 95.48 | 91.29 | 0 | 94.2 | 98.02 | 100 | 97.6 | 97.25 | 97.69 |
| 4 | 95.56 | 96.54 | 94.2 | 0 | 99.11 | 99.4 | 96.23 | 97.1 | 98.84 |
| 5 | 98.15 | 98.02 | 98.02 | 99.11 | 0 | 100 | 98.56 | 100 | 100 |
| 6 | 99.44 | 100 | 100 | 99.4 | 100 | 0 | 99.4 | 100 | 100 |
| 7 | 97.1 | 96.65 | 97.6 | 96.23 | 98.56 | 99.44 | 0 | 98.36 | 98.91 |
| 8 | 87.04 | 96.26 | 97.25 | 97.1 | 100 | 100 | 98.36 | 0 | 98.84 |
| 9 | 98.4 | 99.11 | 97.69 | 98.84 | 100 | 100 | 98.91 | 98.84 | 0 |

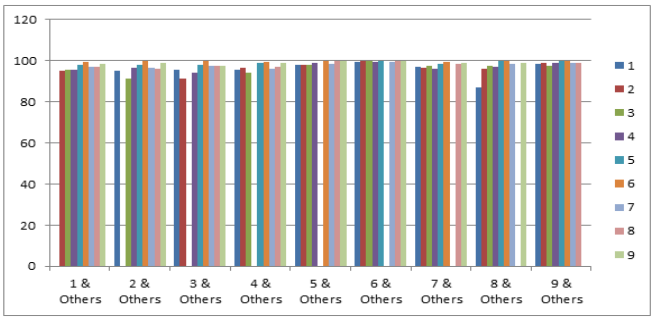



Figure 2. Percentage of Dissimilarity Chart

The values in table II are represented by the column chart given in fig. 2. After analyzing the percentage of similarities, we have found that there exists high dissimilarity in most of the courseware. Every courseware has at least a new idea. Some contents are important, but we see that one courseware have it, but others do not have. In this case, integration will combine all of the new ideas from different courseware.

## IV. COURSE DESIGN & DEVELOPMENT

The present era is an era of globalization. Now-a-days every conscious economic, financial, trade related institutes want to communicate with other similar organizations to get the advantages of globalization. There are different attempts to do this. BDREN [20] is one such attempt, which is an exclusive super highway communication network, linking education, research and innovation organizations in Bangladesh, and across the world. If we could combine the facilities of all of the courseware into a single one, then it can also be a step to globalization. We have already divided the courseware facilities into several levels through trees. Now if we can merge all the fields corresponding to a specific level of all courseware sites into a level of a single courseware so that it contains all the respective fields, then it can help different Universities to communicate with each other. It can also help a knowledge seeker to find the combined facilities into one.

### *A. Proposed Methodology*

To design an integrated courseware, we have to compare each courseware with all other courseware. At first, we have to

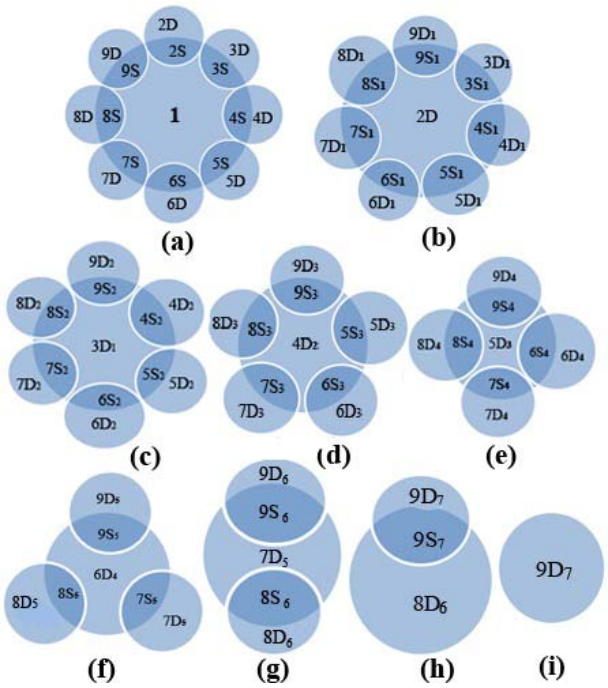



Figure 3. Representation of Methodology for Integrating Content

separate the similar and dissimilar contents of each level. Then we have to select which contents are to be kept in the integrated courseware. Courseware contents can be combined by using the union operation. But, here we have proposed a different methodology to achieve the integration. We have started by considering one courseware as static, whose value will be kept as it is. To represent this methodology, we have used Venn diagram. In the fig. 3(a), we have kept the contents of MIT courseware as static. Here, the MIT courseware is represented in the middle circle by it's ID 1, given in the table I. Other courseware are also represented by their ID. In the figure, the overlapped region represents similarity. In this case, we have considered only the individual similarities of each courseware with ID 1. Here, we have not considered the similarities among the courseware instead of ID 1. It may seem wrong in the first view, but we have considered them in the next parts of this methodology. After separating the similarities and dissimilarities, we have represented them as S and D respectively. For example, in the fig. 3(a), 2D represents the dissimilar contents among ID 1 and ID 2. Accordingly, 2S represents the similar contents among them. Size of the circles is not given according to the volume of contents of a courseware. They are drawn arbitrarily. The same procedure will be applied repeatedly for the next until we find only one circle. In fig. 3(a), we will find all the contents of 1. If 2S contains a content that is same with a content of 3S or…. 9S, remove from the group of these. It is also applicable for other steps. In fig. 3(b), among 2D, 3D, 4D, 5D, 6D, 7D, 8D, and 9D, we will take 2D as static. Then, with respect to this, we will find the following figure according to the same methods of the first step. From this step, we will find 2D, the inner circle. It will not be used for the next steps. In fig. 3(c), among $3D_1$, $4D_1$, $5D_1$, $6D_1$, $7D_1$, $8D_1$, and $9D_1$, we will take $3D_1$ as static. Applying the same procedures stated

above, we will have the following figure. From this step, we will find $3D_1$, the inner circle. It will not be used for the next steps. In fig. 3(d), among $4D_2$, $5D_2$, $6D_2$, $7D_2$, $8D_2$, and $9D_2$, $4D_2$ is taken as static. From this step, we will find $4D_2$ the inner circle. It will not be used for the next steps. In the next fig. 3(e), among $6D_3$, $7D_3$, $8D_3$, and $9D_3$ with respect to $3D_1$, $5D_3$ is taken as static. From this step, we will find $5D_3$, the inner circle. It will not be used for the next steps. In the fig. 3(f), among, $6D_4$, $7D_4$, $8D_4$, and $9D_4$, $6D_4$ is taken as static. From this step, we will find $6D_4$, the inner circle. It will not be used for the next steps. In the next fig. 3(g), among, $7D_5$, $8D_5$, and $9D_5$, $6D_4$ is taken as static. From this step, we will find $6D_4$, the inner circle. It will not be used for the next steps. In the fig. 3(h), among, $8D_6$, and $9D_6$, $8D_6$ is taken as static. From this step, we will find $8D_6$. From the last fig. 3(i), we will find the last total unique component $9D_7$. Now, we have reached at the end of this recursion. Now, the required unique contents for the analyzed courseware in an integrated form can be found by

$$1 + 2D + 3D_1 + 4D_2 + 5D_3 + 6D_4 + 7D_5 + 8D_6 + 9D_7$$

Now, we can generalize this methodology according to the above steps in the following form:

$$\text{Integrated Content} = 1 + \sum_{n=2}^{n=m} (nD_{n-2})$$

Where, 1=contents of first Id; m = total number of courseware; n =2,.....,m; D = dissimilarity

This operation will be conducted for each level separately. For each component found by union of one level, same operation will be used to find the next level for that component. For example, from the above calculation, we have found Home, Courses, About, Feedback etc. in the first level. Now, for each of these components same operation will be used to find the next important level. Among the dissimilar contents, there can have some synonyms. To make them computer understandable, we have to use a synonym table. When developing a synonym table, we face many problems. For example, in MIT, the course *Introduction to Psychology* is taught in Fall 2011 and Fall 2004 with distinct topics. For the simplicity of synonym table, we compress these two courses into one. It has been done manually. An efficient synonym table is needed for an efficient integration of unique contents.

### *B. Application of Methodology*

If we apply the methodology for each level stated in section IV (A), the result will be all dissimilar contents with the similar contents only once. For example, we consider that only 'Courses' is found by applying the methodology in the 1st level. Now, MIT has Topic, Dept., and Course No as the next important level got from 'Courses' level and another courseware from Open Yale University provides Course No, Dept., and Course Title as the next important level of 'Courses' and so on for other courseware. Now, the result of applying methodology will be Topic, Course No., Dept., Course Title etc. as the next important level for the integrated courseware. Here, the important level means the level which we want to keep in our courseware as a level on the basis of their importance. For this reason, we have omitted the 2nd level

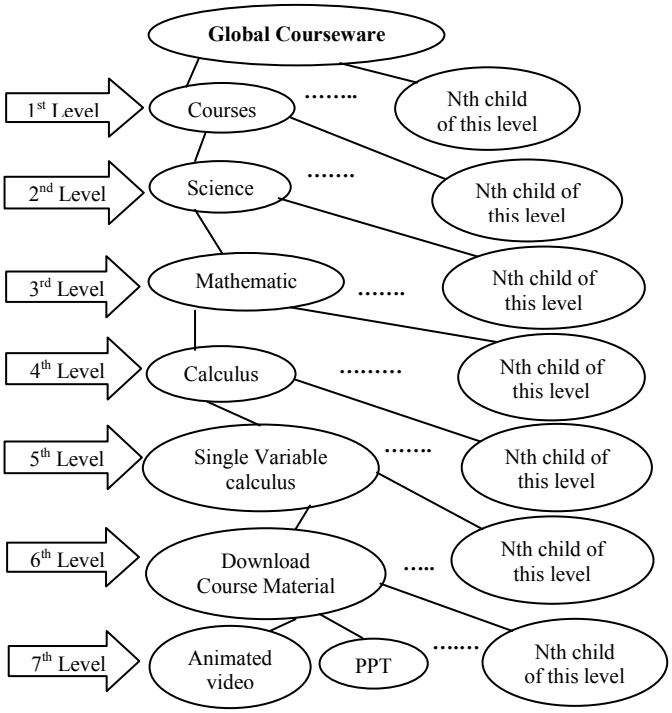



Figure 4. Tree Representation of Proposed Courseware

of table I. 2nd level is not so important for the integrated courseware to consider, because instead of using the contents of this level like 'Find Course By', we can go directly to the next level, which will be more easy and comfortable. Now, for all the components of the 2nd level, the same operation is done. Accordingly, after manual analysis of each level, we have made some summarization or component reduction such as Course No. It is university specific. So, in our proposed courseware, we do not want to keep this component. If a component seems to be not important from our perspective, then all elements under that component will be removed. From the collected data from different courseware and through the analysis of them in different manner, we have proposed an integrated global online courseware design approach for all kinds of learners. In fig. 4, we have made a tree representation according to our idea. Level wise components description for fig. 4 are given as: In the 1st level, we just want to keep the Union of the 1st level of all courseware like Home, Courses, About, Help, Feedback etc. From the components of 1st level, we have chosen Courses to represent our intention. In the 2nd level, we want to keep search option for courses according to school i.e. Science, Humanities, Business, Engineering etc. We have expanded the tree from Courses to the next level in the following figure on the basis of school. The 2nd level is expanded according to department of corresponding school. For the school of science, the department can be Chemistry, Physics, Mathematics etc. Other schools will also be expanded in the same manner. The 3rd level is expanded according to subjects under respective departments. For the department of Mathematics, the subjects can be Calculus, Differentiation, Digital Signal etc. The 4th level is expanded according to distinct sections under relevant subjects. For the subject Calculus, the distinct sections can be Single Variable Calculus, Calculus with Applications etc. 6th level will contain

section wise components such as Syllabus, Download Course Material etc. required to get course related materials and information. The 7th level, we will provide the entire course materials in the form of pdf, video, audio, animations etc as a child of Download Course Material. It is the most important level we are to consider. Learners will get all of their desired contents in integrated form from this level. We do not claim that only these levels are to be considered. Levels and components can be added or removed according to demand.

## V. BENEFITS

Different Universities have distinct courseware which are helpful for students of corresponding university. Different courseware have different good course material. In this case, searching for content requires searching so many addresses to meet the desired need for a self learner. It is very much time consuming and difficult for him. But, if all of the accumulation of knowledge of various instructors from different universities can be found in a single one, then it will reduce the time and difficulties to search for contents. Also an instructor will be benefitted from this. He will find new material for self learning and teaching his students. For example, some courseware may provide only pdf for a course named 'Algorithm', another can provide the video file containing the lecture conducted by that university for that specific course, another can provide animation to make understandable different complex algorithms. If the integrated global online courseware can be developed, then all of these course materials will be found into a single page. So, we can claim that our proposed idea will provide greater benefits for modern age and it will surely be a step to globalization.

## VI. CHALLENGES

To design an integrated courseware, we have faced some questions, such as: (1) Different university developed different contents for their courses. In many cases, the contents of a course vary sometimes on teacher, on level of universities, on prerequisite course. They have their own syllabus. Now question is 'how to integrate this?' (2) Another question arises for 'who will manage these contents globally?' (3) Third question is 'how to get permission from concerned authority?' We think that combining all courses and their contents is so tough through programming and manually, but not impossible. Managing contents globally is difficult. We assume that, any International organization (educational or others) can handle this. If the answers of the above questions can be given with proper solutions, then our proposed global online courseware will be the future necessity.

## VII. CONCLUSION

The purpose of the research is to develop an integrated courseware which will contain all courses of all courseware. If it can be completed, it will play an important role in e-learning. Instead of visiting all courseware, one can see all courses of every courseware by visiting only one site and this saves one's time. Getting permission for merging contents of courseware from all universities will be a tough matter. But, today or tomorrow, through the gradual evolution of globalization, it will be one of the necessities of the world. In different courseware, same course can be represented by different name. For this reason, all the calculations are not fully correct. To solve this problem, a proper synonym table can be used. To create a synonym table, authors have to do much work manually. Even after using this table, there exists some false positive which has to be checked manually. The more contents a synonym table will contain, the less false positive will appear. The removal of the number of these false positives totally depends on the intelligence of the synonym table. The proposed design approach is a continuous process. In order to make relevant changes in the functionality, comprehensive evaluation is needed.

## REFERENCES


[1] L. Vicent, X. Àvila, J. Anguera, D. Badia, and J. A. Montero, "Do multimedia contents increase the effectiveness of learning?," 36th ASEE/IEEE Frontiers in Education Conference, October 28 – 31, 2006, San Diego, CA, pp. 12-17.

[2] F. Buendía and J. C. Cano, "WebgeneOS: A generative and web-based learning architecture to teach operating systems in undergraduate courses," IEEE Transactions on Education, vol. 49, no. 4, pp. 464-473 November 2006.

[3] L. Zhang, Y. Zhuang, Z. Yuan, G. Zhan, "Auto diagnosing: an intelligent assessment system based on bayesian networks," 37th ASEE/IEEE Frontiers in Education Conference, October 10 – 13, 2007, Milwaukee, WI, pp. 7-10.

[4] N. Hoic-Bozic, V. Mornar, and I. Boticki, "A blended learning approach to course design and implementation," IEEE Transactions on Education , vol. 52 , no. 1, pp. 19-30 , February 2009.

[5] J. R. Galvão and A. M. Barreto, Polytechnic Institute of Leiria, Leiria, Portugal,"What is courseware? A comparative analysis," World Transactions on Engineering and Technology Educatio Vol.4, No.2, 2005.

[6] (2013) Types of Courseware on FreeQuality. [Online]. Available: www.freequality.org/documents/training/**Courseware**%5B1%5D.ppt

[7] (2013) Site statistics of visitors in MIT Open Courseware. [Online]. Available: http://ocw.mit.edu/about/site-statistics/

[8] (2013) The Khan Academy website. [Online]. Available: https://www.khanacademy.org/

[9] (2013) The EDX website. [Online]. Available: http://www.edx.org/

[10] (2013) The Open Courseware Consortium website. [Online]. Available: http://www.ocwconsortium.org/

[11] (2013) The MIT Open Courseware website. [Online]. Available: http://www.ocw.mit.edu

[12] (2013) The Open Yale Courses website. [Online]. Available: http://oyc.yale.edu/

[13] (2013) The University of Notre Dame Open Courseware website. [Online]. Available: http://ocw.nd.edu/

[14] (2013) The Connexions website. [Online]. Available: http://cnx.org/

[15] (2013) The Utah State OpenCourseWare website. [Online]. Available: http://ocw.usu.edu/

[16] (2013) The Open.Michigan website. [Online]. Available: https://open.umich.edu/

[17] (2013) The JHSPH Open Courseware website. [Online]. Available: http://ocw.jhsph.edu/

[18] (2013) The Computer Science and Engineering website on University of Washington. [Online]. Available: http://www.cs.washington.edu/

[19] (2013) The CS50.tv website. [Online]. Available: http://cs50.tv

[20] (2013) The Bangladesh Research and Education Network website. [Online]. Available: http://www.bdren.net.bd/

# Chunking Implementation of Extendible Array to Handle Address Space Overflow for Large Multidimensional Data Sets

K. M. Azharul Hasan *, Mehnuma Tabassum Omar* , S. M. Masudul. Ahsan, and Nazmin Nahar
Department of Computer Science and Engineering
Khulna University of Engineering & Technology (KUET)
Khulna 9203, Bangladesh.
*E-mail: misty2409@gmail.com

*Abstract*— **Multidimensional array is commonly used in data warehouse or in many scientific computation of OLAP for dealing with application dataset. Different models of multidimensional array have been introduced to do this job like Traditional Multidimensional Array, Extendible array, Extended Karnaugh Map Representation etc. But Application that requires operation on large amount of data cannot be handled by these models. One common approach is to store data in chunked multidimensional arrays. Indices into an array are obtained by a normalization process that maps attribute values to integers. We have introduced an approach of storing huge amount of data contiguously in disk by applying the chunking approach to an extendible multidimensional array. This implemented approach is effective in managing memory than any other existing approach such as traditional multidimensional array. In our system data are stored in memory according to some predefined amount of memory space that will help the memory to store large amount of data through small unit of chunk where the size of chunk will never overflow the address space limit of a computer**.

***Keywords—Array Chunking, Subarray, Extendible Array, Addressing Function, Chunk Subscript.***

## I. INTRODUCTION

Scientific and engineering data sets that are used in large volume computations are generally modeled as multidimensional array or matrices as their basic data structure [1][2]. Storage of multidimensional array has many advantages but still some issues to be resolved. This includes the followings:

i. The array cannot be extended dynamically i.e. if an array needs to be extended dynamically it needs total reorganization of the array. Reorganization will be time consuming and will not store data if new data along with reorganization is not possible[3][4].

ii. In general the array is very sparse that causes huge amount of nonempty array cells in the array.

iii. In many cases an array will need to be loaded from data that is not in array format (e.g., from a relational table or from an external load file).

iv. It is difficult to fit large size array in the memory, because memory is fixed but data grows rapidly. Some method should be introduced which will efficiently use this fixed memory to hold data as large as possible[5].

v. The array overflows very fast of its address space if the number of dimension and length of dimension is large [6].

To overcome reallocation in case of dynamic extension (problem (i) above) the popular data structure extendible array model [7][8] has been adopted. An extendible array can be extended dynamically during runtime without reallocating the data already stored. Such advantages make it possible to apply extendible array to apply in many database applications namely MOLAP [9], multidimensional database [8][10] and parallel database [11]. But still there is a problem to deal data which is big in size as we have problem (iv) and (v). Extendible array performs well compared with others but it also overflows quickly [5]. One solution to resolve problem (ii) is to break the large n dimensional original array into chunks [12]. For an n-dimensional array $\boldsymbol{A}$, whose dimensionality is $|d_1| \times |d_2| \times \ldots \times |d_n|$ , the chunks can be formed by breaking each $d_i$ into several ranges. Within $\boldsymbol{A}$, two positions are in the same chunk if and only if, in every dimension, they fall within the same range. In memory or disk, values within a chunk are stored consecutively [5]. The existing multidimensional array structures cannot perform well in case of handling a large amount of data because it overflows quickly the available address space of memory even for 64 bit computers when the number of dimension and length of dimension is large. The chunking of array delays the overflow situation of the array [2] as the length of dimension is broken into number of chunks. In this paper we are going to propose an extendible array model that solves the problems above. The model is based on index array or extendible array [7] and we incorporate chunking to the array. In the rest of the paper, we named the data structure as Chunked Extendible Array, CEA and traditional multidimensional array as TMA.

We calculate the chunk size in such a way that the chunk size becomes small and within limit of address space. This data structure provides all the facilities of normal extendible array as well as treats the memory like a small unit called chunk to utilize memory for better performance.

## II. THE EXTENDIBLE ARRAY

Extendible array is a multidimensional storage array which is extendable in all direction. Here reorganization of the existing array is totally omitted which reduces the cost of extension compared with conventional multidimensional array. Extendible array stores not only a single array but also all of its potential extensions. It contains ***n*** dimensions of rectangular array with joining a block of ***n-1*** dimensional sub array. The following are the steps that can describe an extendible array more elaborately.

- Extendible array for ***n*** dimensions is a combination of ***n-1*** dimensional sub array
- This structure consists of three auxiliary tables such as address table, coefficient vector table, history table for each dimension to provide the system with fast access of element.
- Address table contains the first address of each constructed sub array, history table contains the history extension of each dimension and coefficient contains the coefficient vector for ***n-2*** dimensions
- There has a variable for holding the history counter of each construction of sub array.
- Extendible array can be extended in any dimension by only the cost of these three auxiliary tables.

In the Fig. 1 a simple example of extendible array has given where coefficient table is void (as only 2 dimension). By comparing the history value of <3,2> the desired sub array can be defined and from the first address of the sub array the element 11 can be accessed through the calculation of addressing function by using the first address of the address table.

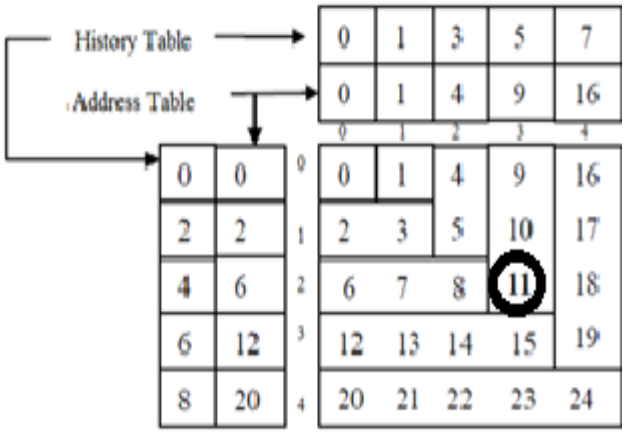



Figure 1: A 2D 5x5 Extendible array.

## III. RELATED WORKS

Traditional Multidimensional Array (TMA) [13], [14] is a well understood and widely used data structure. But TMA cannot be extended dynamically and one problem is that TMA overflows quickly. The Extended Karnaugh Map Representation (EKMR) [15],[16] is a good structure for matrix computation but EKMR is similar to TMA with a different organization. Hence EKMR is not dynamically extendible and it also overflows quickly. The Extendible multidimensional Array [7],[17] is dynamically extendible but it also overflows soon because the subarrays are n-1 dimensional if the extendible array is *n* dimensional. The extendible array is employed in [17] to extend the array and it only treats an organization scheme of the history tables. The subarrays are n dimensional hence it will be difficult to apply in actual implementation when address space overflow is concerned. An extendible array model [18],[19] where there is record for each dimension called axis vector but the overflow situation is the same as in [7]. For round robin expansion, [18][19] is similar to extendible array [7]. [2], [20] presents the chunking of multidimensional arrays. In this scheme the large multidimensional arrays are broken into chunks for storage and processing. All the chunks are n dimensional with smaller length than the original array. But dynamic extension is not possible for this model and it will overflow quickly because the chunks are ***n*** dimensional. In [21] there is an array system based on extendible array [7] presented that can insert a row in the middle of the dimension which deals with overflow situation.

## IV. CHUNKING OF EXTENDIBLE ARRAY

Logically the chunk is an ***n*** dimensional array for an ***n*** dimensional CEA. The offset of a two dimensional array can be realized by rows or by columns and in general ***n***-dimensional there are ***n***! possible linearization orders according to the possible ordering of the dimensions. Hence we use the multidimensional chunk using the following addressing function for linearization.

$$f(i_n, i_{n-1}, i_{n-2}, \ldots\ldots, i_2, i_1) = s_1 s_2 \ldots\ldots s_{n-1} i_n + s_1 s_2 s_3 \ldots\ldots s_{n-2} i_{n-1} + s_1 i_2 + i_1 \quad \ldots\ldots (1)$$

Here $s_i$ is the length of dimension i.

### A. The Basic Data Structure for Chunked Subarray

Each of the chunked subarrays requires a number of chunks. So these chunks need some correspondence to be interconnected to each other for working as a single unit. The number of chunks required for a subarray is determined as follows.

$$nc = \prod_{i-1}^{n} L_i (i \neq j) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

where ,
*nc* = Number of chunk required for a sub array;

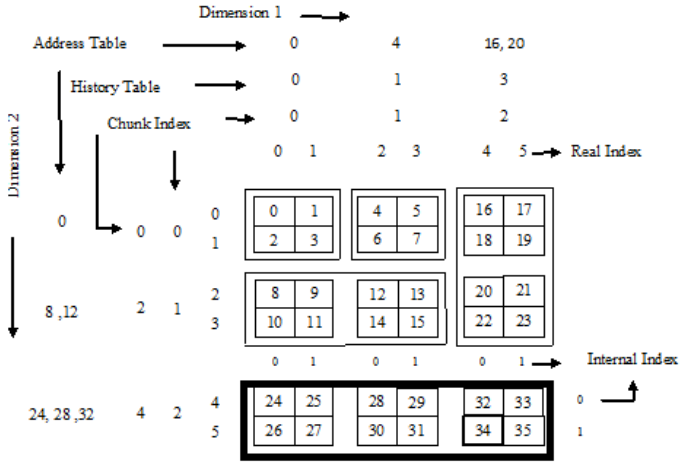



Figure 2: An extension on dimension two

$L_i$ = the length of dimension $i$ of $CEA\,(1 \le i \le n)$ before extension;

$j$= The dimension in which direction (or dimension) of the CEA is extended or the dimension where the subarray is constructed or inserted;

Fig. 2 shows a 2 dimensional *CEA* with its expansion history values. The subarray having history value 4 has ***nc***=3. This is because the length of dimension 1 (i.e. $L_1$) is 3.

The operations on *CEA* are completely different from that of the conventional multidimensional array. The following auxiliary tables play the vital role for the proposed CEA.

*a) History table: It consists of the extension history of the CEA. The history number is a consecutive. When an extension occurs the history number is incremented by 1.*

*b) Coefficient table: The length of dimension of CEA is determined by the number of chunks in that dimension. The length is used to calculate the addressing function. Hence the coefficient table consists of the coefficient vector of the addressing function.*

*c) Address table: The address table that stores each of the first addresses of the chunks of the subarrays. This address is used later for accessing data. It stores all the first addresses of the chunks that exist in a subarray.*

*B. Choice of Chunk Size*

The chunk size choice is very important for *CEA*. Let $\boldsymbol{L_i}$ $(1 \le i \le n)$ be the length of each dimension of a chunk. Then chunk size (***cz***) is determined as

$$cz = \prod_{i=1}^{n} L_i \text{ ..... ..... ..... .....} (3)$$

The length $\boldsymbol{L_i}$ is chosen in such a way that ***cz*** never overflows for a specific system. Let the data type for a specific machine is ***k*** bit. Then the range of data type is $0\text{-}2^k\text{-}1$. Hence the largest number that can be stored is $2^k\text{-}1$. Therefore the $\boldsymbol{L_i}$ is chosen in such a way that

$$cz = \prod_{i=1}^{n} L_i \le 2^k - 1 \text{.....} (4)$$

Chunks are the basic storage area that stores the array element. Each chunk is $n$ dimensional and each of the dimensions of the chunk is predefined. Hence the total size of the chunk is also predetermined. Thus the over flow condition does not occur in CEA scheme.

*C. Accessing an Arbitary Element*

The chunked implementation of extendible array (CEA) scheme considers an $n$ dimensional array as partitioned into blocks having a subarray of n-1 dimensions. Hence the CEA scheme is not a single stream of array elements rather it is a combination of adjoining blocks called subarrays. Moreover the subarrays are not the conventional array they are combination of single arrays. Fig.3 depicts the scheme layout for a two dimensional *CEA*. Each of the dimensions has an index set called subscripts. The subscript maps to a unique chunk which itself a multidimensional array maps to a unique cell. Accessing any element in the multidimensional CEA has major two parts:

*1) Determining the subarray inside which the element exists.*

*2) Determining the corresponding chunk and hence the element inside the chunk.*

*a)* ***Determining the candidate subarray:*** There are three types of subscipt in CEA namely real subscipt, chunk subscipt and internal subscipt. See Fig. 3. The real subscipt that corresponds to the real co-ordinate, which is the continuation from the starting of the corresponding dimension. This real subscipt is denoted as $< x_1, x_2, ..., x_n >$. The chunk subscipt is the chunk number inside the subarray which is denoted by $< xc_1, xc_2, ..., xc_n >$ and the third subscipt is the subscipt inside the chunk, which is termed as internal subscipt and denoted as $< xi_1, xi_2, ..., xi_n >$.

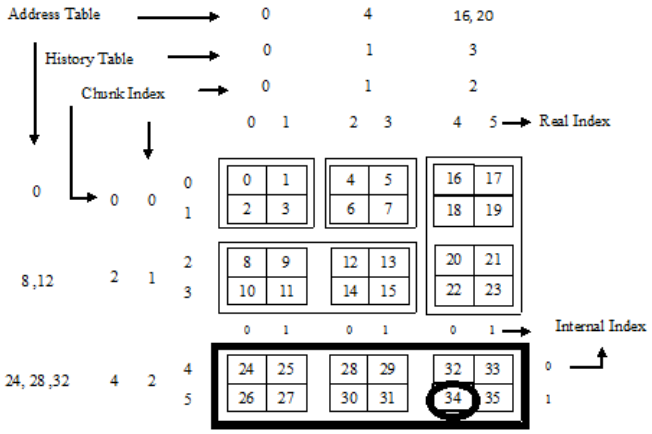



Figure 3: A Complete CEA example

The chunk subscipt $< xc_1, xc_2, ..., xc_n >$ is determined by the following function

$$xc_i = x_i / L_n \quad \ldots\ldots(5)$$

The role of the chunk subscript is to correspond to the specific chunk in the subarray.

Example 1: Consider the element <4, 5>, here $x_1$=4 and $x_2$=5. Hence $xc_1 = 4/2 = 2$ and $xc_2 = 5/2 = 2$ . So, our candidate chunk subscript is $< xc_1, xc_2 > \approx < 2,2 >$.

Once the chunk subscript is determined the candidate subarray is determined as follows.

Let the history tables are $H_1$ and $H_2$ of the corresponding dimension. Now compare $H_1$ *[xc1]* and $H_2$ *[xc2]*. Then if $H_1$ *[xc1]*> $H_2$ *[xc2]*, it can be proved that the corresponding element is mapped to the subarray having the history value $H_1$ *[xc1]*. This is because the larger subarray history value was constructed later than the smaller one. Hence the corresponding coordinate was mapped to the subarray later. So the subarray of the maximum history value becomes the candidate subarray. To determine the subarray compare $H_2$ *[2]* and $H_1$ *[2]* in Fig 3.

*b)* ***Determining the chunk and the element:*** Two types of co-efficient vectors are identified and stored; one is for determining the corresponding chunk and another is the co-efficient vector for fixed size chunk that determines the corresponding cell in the chunk. That is these two co-efficient vectors determine two offset values one corresponds to the specific chunk and another corresponds to a specific cell in the chunk for which internal subscript is needed.

Example 2: As in example 1 consider the element <4, 5> the subarray is determined as the one having history value 4.The coefficient vector is stored in co-efficient table. Since the coefficient vector is of order n-2, the CEA in example 1 is two-dimensional and hence the co-efficient table is void. For this reason the offset is 2 and the chunk number is determined as 2 as shown in Fig. 3.

The internal subscipt $< xi_1, xi_2, ..., xi_n >$ is determined by the following function

$$xi_i = x_i \% L_n \quad \ldots\ldots(6)$$

Example 3: For the element in example 1 $xi_1 = 4\%2 = 2$ and $xi_2 = 5\%2 = 1$ . So, our candidate chunk subscript is $< xi_1, xi_2 > \approx < 0,1 >$.

The fixed size co-efficient vector for the chunk is stored only once since the chunk size is fixed in every dimension. The internal index is determined as $< xi_1, xi_2 > \approx < 0,1 >$ in example 1 and the offset inside the chunk is determined as

$$s_1 x_2 + x_1 = 2x_1 + 0 = 2 \quad \ldots\ldots(7)$$

So the element <4, 5> is determined as 34 that is shown in Fig.3.

## V. EXPERIMENTAL RESULTS

The experiment that we have done on our system is measuring the condition of memory when it overflows. The main purpose of the paper is to limit the occurrences of overflow as far as possible. In our analysis we have used the machine which has Intel processor (2.20 GHz), RAM (1 GB physical memory with virtual memory 2.48 GB), OS (Windows XP professional).

Fig. 4 shows the memory management performed by CEA and TMA. The total memory managed by CEA is better than that of TMA. The total data that we can store in secondary disk with TMA is approximately 990 MB while with CEA we have become successful to store data about 3.99 GB approximately. The reason is that the system maximum size is 1 GB (RAM) and the previous storage uses 990MB and if new insertions cross 1 GB then it will not store any data beyond 990MB. It will soon overflows with storing small amount of data and wasting some cell space. So the performance of TMA array almost same for various dimensions. For CEA it is possible to reduce overflow of memory by introducing the concept of chunked subarray. In every case it will only allocate a small chunk instead of a large array with reallocation. For the system as mentioned for TMA where it can't cross the limit beyond 990MB, CEA will run till only one chunk can't find space for allocation. That means it will use as large space as it can till the system maximum size and thus will store large amount of data. Hence it can store highest data and will overflows later. Here another issue is that in

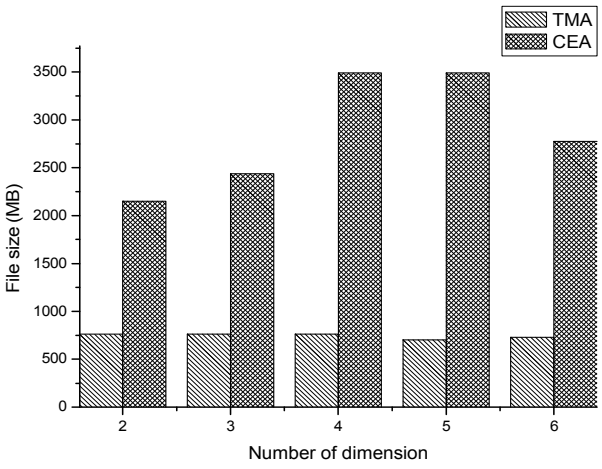



Figure 4: Memory management of CEA and TMA.

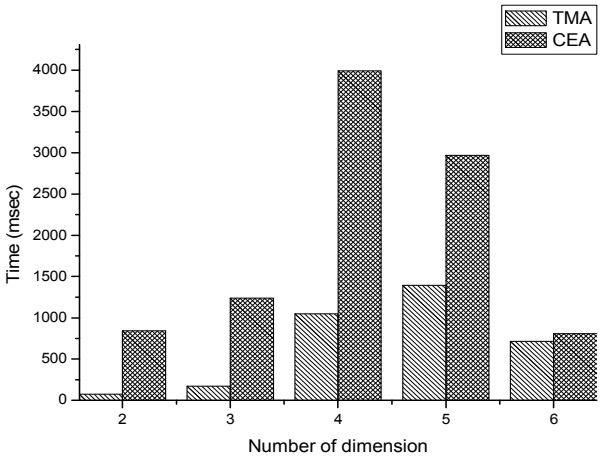



Figure 5: Average Construction Time of CEA and TMA.

some cases CEA can use full virtual memory for allocation but TMA cannot. That means if virtual memory is considered then the storage will be also high. But as using virtual memory slowdowns the performance hence virtual memory is not considered here.

Fig. 5 shows the comparison of construction time among CEA and TMA. To construct a TMA first of all the amount of data needed has to be calculated. After then the size of space to allocate this amount of data is calculated and if allocation is possible then data is stored. Here the allocation is done once. But to construct a CIEA for same amount of data as TMA does, it has to perform ***nc*** (no. of chunk**)** times allocations. The amount of allocation is not only larger but also it has to maintain some auxiliary table. The role of this auxiliary tables have mentioned before. Therefore time for constructing TMA is larger than CIEA.

Fig. 6 shows the comparison of extension time among CEA and TMA. For CEA an extension along any direction causes to increase length of dimension of an array by the length of dimension of the chunk. But for TMA if possible, an extension along any direction causes to increase length of dimension by one. So extension memory is always larger for CEA that means an extension for CEA will handle large data than TMA. To handle same amount of data as CEA, TMA has to extend chunk times length of dimension. During extension in any direction TMA has to reallocate its previous data. That means it has to keep its previous data position same. And the allocation for memory is always previous allocation including new allocation. So the time for TMA is larger. For CEA to extend in any direction it does not need any reallocation. It only requires number of chunks for an extended subarray and allocates them individually. So time needs smaller.

Here, one issue is that in every case the amount of output decreases as dimension increases, for example

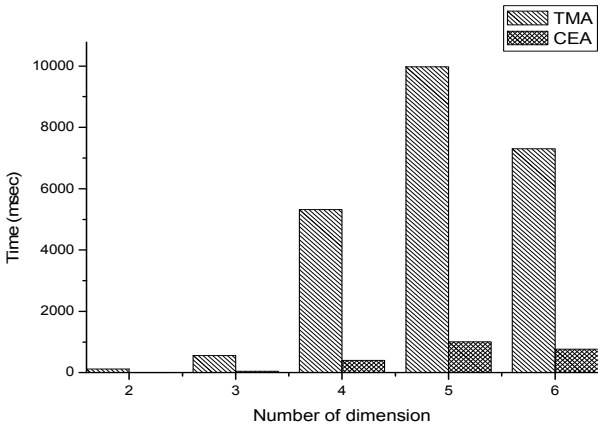



Figure 6: Average Extension Time of CEA and TMA.

extension time for 6 dimensions. This happens because along with increasing dimensions, the number of data to be stored increases, hence overflows.

So, it can be said that CEA is applicable for the situation where large data is to be handled.

## I. CONCLUSION

The model presented here is the extension of extendible array by including chunk concept in it. Our proposed approach CEA for ***n*** dimensions is superior for managing memory efficiently. With CEA we can utilize same amount of memory more effectively than TMA by making the system able to store as many values as the whole capacity of the hard disk. Though we have been able to write approximately 4 GB by using 1 GB RAM but it should not stop writing until there have spaces in hard disk. Because of our limitation of programming system we cannot write in the whole available disk. One important future application is to apply the system in parallel environment. Most of the operations described here are independent to each other. Hence it will be very effective to apply the system in parallel environment. CEA is a basic data structure. The system can be applied to database and data warehousing applications.

## REFERENCES

[1] Ekow J. Otoo and Doron Rotem ” A Storage Scheme for Multi-dimensional Databases Using Extendible Array Files”, Proceedings of the thirdWorkshop on STDBM Seoul, Korea, September 11, PP. 67-76, 2006.

[2] Yihong Zhao, Prasad M. Deshpande, and Jeffrey F. Naughton, **“**An Array-Based Algorithm for Simultaneous Multidimensional Aggregates” Proceeding SIGMOD conference on Management of data,pp 159 – 170, 1997.

[3] Sk Md Masudul Ahsan, K. M. Azharul Hasan (2011) “An Implementation Scheme for Multidimensional Extendable Array Operations and Its Evaluation”, In: ICIEIS 2011, Part III, CCIS 253, pp. 136–150, 2011.

[4] K. M. Azharul Hasan, Kamrul Islam, Mojahidul Islam, Tatsuo Tsuji, "An Extendible Data Structure for Handling Large Multidimensional Data Sets", In: Proceedings of the 12th International Conference on Computer and Information

Technology (ICCIT 2009), Dhaka, Bangladesh, pp. 669-674,2009.

[5] K. M. Azharul Hasan, "Compression Schemes of High Dimensional Data for MOLAP", In the Edited Book, "Evolving Application Domains of Data Warehousing and Mining: Trends and Solutions" Chapter IV, Information Science Reference, pp. 64-81, September 2009.

[6] Sk Md Masudul Ahsan, K M Azharul Hasan (2011) "A Solution of Address Space Overflow for Large Multidimensional Arrays" In: Proceedings of the 14th International Conference on Computer and Information Technology (ICCIT 2011), Dhaka, Bangladesh.

[7] Otoo E. J and T. H. Merrett, "A storage scheme for extendible arrays". Computing, Vol.31, pp.19, 1983.

[8] Hasan, K. M. A., Kuroda, M., Azuma, N., Tsuji, T., & Higuchi, K. (2005). An extendible array based implementation of relational tables for multidimensional databases. In Proceedings of 7th International Conference on Data Warehousing and Knowledge Discovery (DaWaK'05) (pp. 233-242). Heidelberg: Springer-Verlag.

[9] Masayuki Kuroda, Naoki Azuma, K. M. Azharul Hasan, Tatsuo Tsuji, Ken Higuchi (2005) "An Implementation Scheme of Relational Tables", In: Proceedings of 21st IEEE Conference on Data Engineering (ICDE'05), Tokyo, Japan, pp 100-104, IEEE CS, 2005.

[10] K. M. Azharul Hasan, T. Tsuji, and K. Higuchi, **"**An Efficient Implementation for MOLAP Basic Data Structure and Its Evaluation", Proceedings of DASFAA, LNCS 4443, pp. 288 – 299, 2007.

[11] Hasan K. M. A, Tsuji T, Higuchi K., A Parallel Implementation Scheme of Relational Tables Based on Multidimensional Extendible Array, International Journal of Data Warehousing and Mining, 2(4), 66-85. 2006.

[12] S. Sarawagi, and M. Stonebraker, "Efficient organization of large multidimensional arrays", Proc. of ICDE, pp. 328-336, 1994.

[13] Seamons, K.E., Winslett, M.: Physical Schemas for Large Multidimensional Arrays in Scientific Computing Applications. In: 7th International Conference on Scientific and Statistical Database Management (SSDBM), pp. 218--227, IEEE CS, Washington, DC, USA, 1994.

[14] Sarawagi, S., Stonebraker, M.: Efficient Organization of Large multidimensional Arrays. In: 10th International Conference on Data Engineering, pp. 328--336, Houston, TX , USA, 1994.

[15] Chun, Y.L., Yeh, C.C., Jen, S.L.: Efficient Representation Scheme for Multidimensional Array Operations. IEEE Transactions on Computers, 51(3), 327--345, 2002.

[16] Chun, Y.L., Yeh, C.C., Jen, S.L.: Efficient Data Parallel Algorithms for Multidimensional Array Operations Based on the EKMR Scheme for Distributed Memory Multicomputer. IEEE Transactions on Parallel and Distributed systems, 14(7), 625--639, 2003.

[17] Rotem, D., Zhao, J.L.: Extendible Arrays for Statistical Databases and OLAP Applications. In: 8th International Conference on SSDBM, pp. 108--117, Stockholm , Sweden, 1996.

[18] Otoo, E. J., Rotem, D.: A Storage Scheme for Multi-dimensional Databases Using Extendible Array Files. In: 3rd Workshop on STDBM, pp. 67--76, Seoul, Korea, 2006.

[19] Otoo, E. J., Rotem, D.: Efficient Storage Allocation of Large-Scale Extendible Multi-dimensional Scientific Datasets. In: 18th International Conference on SSDBM, pp. 179--183, Vienna, Austria, 2006.

[20] Rotem, D., Otoo, E. J., Seshadri, S.: Optimal Chunking of Large Multidimensional Arrays for Data Warehousing. Lawrence Berkeley National Laboratory, University of California, LBNL-63230, 2007.

[21] Kumakiri, M., Bei, L., Tsuji, T., Higuchi, K.: Flexibly Resizable Multidimensional Arrays. In: 22nd International Conference on Data Engineering Workshops, pp. 83--88, Atlanta, GA, USA, 2006.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# A Novel Steganographic Scheme using Sudoku

Arnab Kumar Maji [1,*], Rajat Kumar Pal [2], and Sudipta Roy [3]
[1] Department of Information Technology, North Eastern Hill University, Shillong 793 022, Meghalaya, India
[2] Department of Computer Science and Engineering, University of Calcutta, Kolkata 700 009, West Bengal, India
[3] Department of Information Technology, Assam University, Silchar 788 011, Assam, India
E-mail*: arnab.maji@gmail.com

***Abstract*—The word 'Steganography' is derived from Greek words 'stegos' meaning 'cover' and 'graphia' meaning 'writing'. Steganography conceptually implies that the message to be transmitted is not visible to the normal eye. It is an art of hiding information inside a medium. The main objective of Steganography is mainly concerned with the protection of contents of the hidden information. A message in cipher text may arouse suspicion while an invisible message is not. A digital image is a flexible medium used to carry a secret message because the slight modification of a cover image is hard to distinguish by human eyes. In this paper, we have proposed a Steganographic scheme using Sudoku puzzle. The novelties of the work lies that nobody could be able to modify the cover image during transmission as well as the message is embedded using a Sudoku reference matrix.**

***Keyword: Sudoku, Steganography, Pixel, Reference matrix, Embedding, Extraction.***

## I. INTRODUCTION

Steganography is the art and science of writing hidden messages in such a way that no one apart from the intended recipient knows the existence of the message. The following formula provides a very generic description of the pieces of the Steganographic process:

*cover_medium* + *hidden_data* + *stego_key* = *stego_medium*

In this context, the *cover_medium* is the file in which we hide the *hidden_data*, which may also be encrypted using the *stego_key*. The resultant file is the *stego_medium* (which is, of course, the same type of file as the *cover_medium*). The *cover_medium* and thus, the *stego_medium* are typically image or audio files. In this paper, we focus on image files and therefore, refer to the *cover_image* and *stego_image*.

Before discussing how information is hidden in an image file, it is worth a fast review of how images are stored. An image file is merely a binary file containing a binary representation of the colour or light intensity of each picture element (pixel) comprising the image.

Images typically use either 8-bit or 24-bit colour. When using 8-bit colour, there is a definition of up to 256 colours forming a palette for this image, each colour denoted by an 8-bit value. A 24-bit colour scheme, as the term suggests, uses 24 bits per pixel and provides a much better set of colours. In this case, each pixel is represented by three bytes, each byte representing the intensity of the three primary colours red, green, and blue (RGB), respectively.

The simplest approach in hiding data within an image file is called least significant bit (LSB) insertion [4]. In this method, we can take the binary representation of the *hidden_data* and overwrite the LSB of each byte within the *cover_image*. If we are using 24-bit colour, the amount of change is minimal and indiscernible to the human eye. As an example, suppose that we have three adjacent pixels (nine bytes) with the following RGB encoding:

10010101 00001101 11001001

10010110 00001111 11001010

10011111 00010000 11001011

Now suppose we want to *hide* the following nine bits of data (the hidden data is usually compressed prior to being hidden): 101101101. If we overlay these nine bits over the LSB of the nine bytes above, we get the following (where bits in bold have been changed):

10010101 00001100 11001001

10010111 00001110 11001011

10011111 00010000 11001011

One of the potential problems of this type of method is that any intruder can easily modify the *cover_image*. Then it is impossible for the receiver to extract the hidden message in the *cover_image*. So, detecting whether a *cover_image* is modified during transmission is very much essential.

'Sudoku' is a popular Japanese puzzle game. It is usually a 9×9 grid based puzzle problem which is subdivided into nine 3×3 minigrids, wherein some clues are given and the objective of the problem is to fill it up for the remaining blank positions. Furthermore, the objective of this problem is to compute a solution where the numbers 1 through 9 occur exactly once in each row, exactly once in each column, and exactly once in each minigrid independently obeying the given clues. Besides the standard 9×9 grid, variants of Sudoku puzzles include the following:

- 4×4 grid with 2×2 minigrids,
- 5×5 grid with *pentomino* regions published under the name *Logi-5* [1]; a *pentomino* is composed of five congruent squares, connected orthogonally; *pentomino* is seen in playing the game *Tetris* [2],
- 8×8 grid with 2×4 minigrids [6],
- 16×16 grid (super Sudoku) [5],
- 25×25 grid (Sudoku, the Giant) [5], etc.

In our proposed scheme, we have embedded an 8×8 Sudoku puzzle into the cover image to prevent and detect the modification in the *cover_image*, if any. Along with it, we have used an 18×18 Sudoku reference matrix as a key for message embedding into the *cover_image*.

## II. Literature Survey

Initial work on Steganography using Sudoku puzzle was done by Chang, Chou, and Kieu [3]. The basic idea in the method is to use a Sudoku puzzle to generate a *reference matrix* (M) and alter the values at selected pixels in cover image according to values represented in reference matrix. For an 8-bit cover image, the size of reference matrix is 256×256.

*Chang et al.'s method [3]:*

1. Create a tile matrix (T) by using Sudoku puzzle solution and subtracting 1 from all the fields in the puzzle as shown in Figure 1. The initial puzzle contained data values ranging from 1 to 9. After subtracting 1 from each of the values in the solution of the puzzle, we obtain the solution that contains values ranging between 0–8.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 6 | | | | | 9 | 8 |
| 2 | 5 | | | | | | 1 | |
| | | | | 7 | | | | 5 |
| | | 5 | 6 | | 3 | | | 4 |
| 1 | | | 8 | | 5 | 2 | | 9 |
| 9 | | | | 6 | | | | |
| | 2 | | | | | | 4 | 7 |
| 5 | 4 | | | | 2 | 9 | | |

(a)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 3 | 7 | 6 | 2 | 5 | 1 | 4 | 9 | 8 |
| 2 | 5 | 9 | 4 | 8 | 6 | 7 | 1 | 3 |
| 8 | 1 | 4 | 3 | 7 | 9 | 6 | 2 | 5 |
| 7 | 9 | 5 | 6 | 2 | 3 | 1 | 8 | 4 |
| 4 | 8 | 2 | 9 | 1 | 7 | 5 | 3 | 6 |
| 1 | 6 | 3 | 8 | 4 | 5 | 2 | 7 | 9 |
| 9 | 3 | 7 | 1 | 6 | 4 | 8 | 5 | 2 |
| 6 | 2 | 1 | 5 | 9 | 8 | 3 | 4 | 7 |
| 5 | 4 | 8 | 7 | 3 | 2 | 9 | 6 | 1 |

(b)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2 | 6 | 5 | 1 | 4 | 0 | 3 | 8 | 7 |
| 1 | 4 | 8 | 3 | 7 | 5 | 6 | 0 | 2 |
| 7 | 0 | 3 | 2 | 6 | 8 | 5 | 1 | 4 |
| 6 | 8 | 4 | 5 | 1 | 2 | 0 | 7 | 3 |
| 3 | 7 | 1 | 8 | 0 | 6 | 4 | 2 | 5 |
| 0 | 5 | 2 | 7 | 3 | 4 | 1 | 6 | 8 |
| 8 | 2 | 6 | 0 | 5 | 3 | 7 | 4 | 1 |
| 5 | 1 | 0 | 4 | 8 | 7 | 2 | 3 | 6 |
| 4 | 3 | 7 | 6 | 2 | 1 | 8 | 5 | 0 |

(c)

Figure 1: (a) An instance of 9×9 Sudoku problem. (b) A solution of the Sudoku instance shown in Figure 1(a), where a digit / symbol occurs exactly once in each row, column, and minigrid. (c) The tile matrix (T) for creating the reference matrix (M).

2. Generate the reference matrix (R): Replicate the tile matrix on both axes to create a matrix of size 256×256. The reference matrix, M is then consisting of an m×m tiling of copies of T, where m = floor(256/9) + 1. The overflowing fields are ignored and M is truncated to 256×256 matrix.

3. Data embedding: Convert the cipher text to base-9 numeral system. So that the cipher text is converted as: $S = S_1 S_2 S_3 \ldots S_n$, where n is the number of converted secret digits and $S_k \in [0,8]$, $1 \le k \le n$. The converted cipher text can now be pointed into the reference matrix as the values in this matrix also lie in the same range (i.e., 0–8).

4. The cover image is partitioned into R nonoverlapping blocks of size 1×2 as shown in Figure 2**.** Each tuple contains one pair of values, in which each value ranges from 0 to 255. This value pair is treated as X and Y coordinates for the reference matrix. This is used to traverse the reference matrix for data embedding.

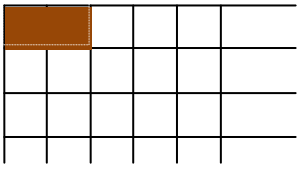


Figure 2: Pixel tuple from carrier image.

5. Find the value in the reference matrix at that position pointed by the pixel tuple values using pixel pair (X,Y) as x and y coordinates in 256×256 reference matrix. The base-9 value to conceal in image is selected and is looked up at the closest position pointed in reference matrix the coordinates of this block (X′,Y′) are picked.

6. These new coordinates are updated in place of original tuple. This provides the location of concealed data in reference matrix with a deviation of maximum two bits in each value of tuple. To obtain data back, receiver needs to generate reference matrix again and selects pixel pairs sequentially to map values in it.

But there are several limitations or drawbacks of this method.

(i) The works performed only on greyscale images.

(ii) With increase in pixel size, the size of the reference matrix multiplies. Hence, it is infeasible for RGB images.

This method is further modified by Shetty *et al.* [7]. They have used a 27×27 reference matrix for the same purpose. This method is slightly advantageous than the scheme proposed by Chang *et al.*, when the size of the matrix is less. But unfortunately none of the methods is able to detect, whether the *cover_image* is modified or not. If the cover image is changed, the message embedded inside the image is also changed. They have used Sudoku reference matrix of size 256×256 and 27×27, which is higher in size and computation is more.

## III. The Proposed Method

As discussed earlier in the previous sections, the detection and prevention of modification in *cover_image* is very much essential. The size of the reference matrix, used as a key for the hidden message, is also very much higher in all earlier methods. In our proposed approach, we have used one 8×8

Sudoku for detecting modifications in the cover image and an 18×18 Sudoku reference matrix is used as key, which is lesser in size than that of each of the previous methods.

The image used in our proposed method is a 24-bit colored image. Initially, the cover picture is taken and an 8×8 Sudoku is embedded onto it using the LSB embedding technique.

*A. Embedding an 8×8 Sudoku matrix in cover_image:*

*(i) Block preparation:*

Initially the *cover_image* is divided into 64 blocks. In each *block*, we have created several groups of three pixels each.

Then, extract the values of B (*Blue*) components of each pixel, which is represented by eight bits.

Now we take an 8×8 Sudoku puzzle. Subtract 1 from each cell. Then the value in each cell be ranging from 0–7, each of which can be represented by three bits. An instance of 8×8 Sudoku puzzle and its possible solution are shown in Figure 3.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 6 | | 3 | 4 | | | 1 |
| 4 | | | | 5 | 3 | | 2 |
| | | | 6 | | | | |
| | 1 | | | | 7 | | 5 |
| | | 8 | 2 | | | | |
| | 7 | | | | | 8 | 4 |
| | | | 4 | | 1 | | |
| | 5 | 2 | 1 | | 6 | | 8 |

(a)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2 | 6 | 5 | 3 | 4 | 8 | 7 | 1 |
| 4 | 8 | 1 | 7 | 5 | 3 | 6 | 2 |
| 5 | 2 | 7 | 6 | 8 | 4 | 1 | 3 |
| 3 | 1 | 4 | 8 | 6 | 7 | 2 | 5 |
| 1 | 4 | 8 | 2 | 7 | 5 | 3 | 6 |
| 6 | 7 | 3 | 5 | 1 | 2 | 8 | 4 |
| 8 | 3 | 6 | 4 | 2 | 1 | 5 | 7 |
| 7 | 5 | 2 | 1 | 3 | 6 | 4 | 8 |

(b)

Figure 3: (a) An instance of 8×8 Sudoku puzzle. (b) A solution of the Sudoku instance.

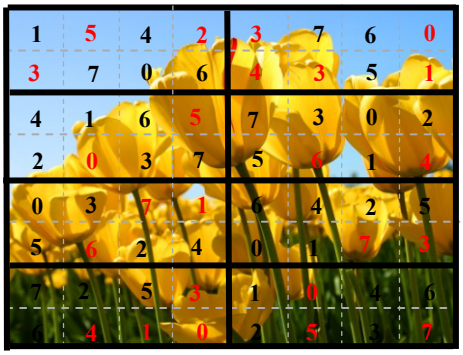


Figure 4: Embedding of the 8×8 Sudoku solution into the *cover_image*.

Now, in each *block* embed the values present in each cell. The embedding of these values is done in the following manner:

For each three pixels group of a *block*, insert the values in the LSB of B (*Blue*) components. Thus only one bit is changed. The embedding procedure of the given puzzle is shown in Figure 4.

After that, an 18×18 reference matrix is generated from a 9×9 Sudoku puzzle. The reference matrix is created by replicating 9×9 Sudoku solution in a square form. The values of the reference matrix lie between 0 through 8 as 1 is subtracted from each value. The reference matrix behaves as a reference look-up table for embedding. A sample 18×18 Sudoku reference matrix is generated, as shown in Figure 5, using the Sudoku instance in Figure 1.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 6 | 5 | 1 | 4 | 0 | 3 | 8 | 7 | 2 | 6 | 5 | 1 | 4 | 0 | 3 | 8 | 7 |
| 1 | 1 | 4 | 8 | 3 | 7 | 5 | 6 | 0 | 2 | 1 | 4 | 8 | 3 | 7 | 5 | 6 | 0 | 2 |
| 2 | 7 | 0 | 3 | 2 | 6 | 8 | 5 | 1 | 4 | 7 | 0 | 3 | 2 | 6 | 8 | 5 | 1 | 4 |
| 3 | 6 | 8 | 4 | 5 | 1 | 2 | 0 | 7 | 3 | 6 | 8 | 4 | 5 | 1 | 2 | 0 | 7 | 3 |
| 4 | 3 | 7 | 1 | 8 | 0 | 6 | 4 | 2 | 5 | 3 | 7 | 1 | 8 | 0 | 6 | 4 | 2 | 5 |
| 5 | 0 | 5 | 2 | 7 | 3 | 4 | 1 | 6 | 8 | 0 | 5 | 2 | 7 | 3 | 4 | 1 | 6 | 8 |
| 6 | 8 | 2 | 6 | 0 | 5 | 3 | 7 | 4 | 1 | 8 | 2 | 6 | 0 | 5 | 3 | 7 | 4 | 1 |
| 7 | 5 | 1 | 0 | 4 | 8 | 7 | 2 | 3 | 6 | 5 | 1 | 0 | 4 | 8 | 7 | 2 | 3 | 6 |
| 8 | 4 | 3 | 7 | 6 | 2 | 1 | 8 | 5 | 0 | 4 | 3 | 7 | 6 | 2 | 1 | 8 | 5 | 0 |
| 9 | 2 | 6 | 5 | 1 | 4 | 0 | 3 | 8 | 7 | 2 | 6 | 5 | 1 | 4 | 0 | 3 | 8 | 7 |
| 10 | 1 | 4 | 8 | 3 | 7 | 5 | 6 | 0 | 2 | 1 | 4 | 8 | 3 | 7 | 5 | 6 | 0 | 2 |
| 11 | 7 | 0 | 3 | 2 | 6 | 8 | 5 | 1 | 4 | 7 | 0 | 3 | 2 | 6 | 8 | 5 | 1 | 4 |
| 12 | 6 | 8 | 4 | 5 | 1 | 2 | 0 | 7 | 3 | 6 | 8 | 4 | 5 | 1 | 2 | 0 | 7 | 3 |
| 13 | 3 | 7 | 1 | 8 | 0 | 6 | 4 | 2 | 5 | 3 | 7 | 1 | 8 | 0 | 6 | 4 | 2 | 5 |
| 14 | 0 | 5 | 2 | 7 | 3 | 4 | 1 | 6 | 8 | 0 | 5 | 2 | 7 | 3 | 4 | 1 | 6 | 8 |
| 15 | 8 | 2 | 6 | 0 | 5 | 3 | 7 | 4 | 1 | 8 | 2 | 6 | 0 | 5 | 3 | 7 | 4 | 1 |
| 16 | 5 | 1 | 0 | 4 | 8 | 7 | 2 | 3 | 6 | 5 | 1 | 0 | 4 | 8 | 7 | 2 | 3 | 6 |
| 17 | 4 | 3 | 7 | 6 | 2 | 1 | 8 | 5 | 0 | 4 | 3 | 7 | 6 | 2 | 1 | 8 | 5 | 0 |

Figure 5: An 18×18 reference matrix (M) is created using tile matrix (T), for the Sudoku instance in Figure 1(c).

*B. Embedding hidden message:*

*Cover_image* is divided 64 equal sized blocks and in each block, each three pixels are grouped. Each group carries a character of the secret message and a value of the 8×8 Sudoku that corresponds to the block. The R and G components hold the secret character and the B component holds the value of the Sudoku. The benefit of the 8×8 Sudoku is that if anyhow the secret message or the picture is tampered, then the 8×8 Sudoku will not be solved at the receiver's side, and hence, the attack is detected.

*(i) Conversion of secret message into base-9:*

All the characters of the secret message are converted to base-9 having three digits so that the digits of the characters fall in the range 0–8 and can be referred to the 18×18 reference matrix.

Read character one-by-one from the message.

Convert the character to ASCII code (decimal).

Convert the decimal (ASCII code) to base-9.

Add padding bits (0's) in front of the base-9 number, if necessary, so that the number is a 3-digit number.

*(ii) Finding out the base-9 values in reference matrix:*

For each digit of the three-digit base-9 character code

Read pixel (h,f) /*h and f are random location of pixel*/

$X = (R\%6) + 6,$

$Y = (G\%6) + 6.$

Locate the cell (X,Y) in the reference matrix. X can be considered as row number and Y as column number.

Now consider nine cells in the same row, keeping the cell (X,Y) in the middle of it and store them in $C_r$, i.e., we like to consider only four left-most and four right-most cells, surrounding (X,Y).

Similarly, choose nine cells in the same column, keeping the cell (X,Y) in the middle of them, and store them in $C_c$. Select the minigrid, where the cell (X,Y) belongs to and store them in $C_m$.

If $(X_i,Y_i)$ = base-9 digit, then

Locate $(X_r,Y_r)$, $(X_c,Y_c)$, and $(X_m,Y_m)$ in $C_r$, $C_c$, and $C_m$ of the reference matrix, respectively.

Calculate deviation of $(X_r,Y_r)$, $(X_c,Y_c)$, $(X_m,Y_m)$ from (X,Y).

Select the cell from the above three candidate elements which has minimum deviation from (X,Y).

*(iii) Updating the R and G values of each pixel:*

Suppose, $(X_c,Y_c)$ has minimum deviation from (X,Y), then

The pixel at (h,f) has new data-embedded R-G-B values as:

$R = R - (X - X_c)$

$G = G - (Y - Y_c)$

B holds the 8×8 Sudoku value embedded previously.

Let us take an example, say after conversion from R-G values, using formula, say for a particular pixel we get X = 12 and Y = 7. The corresponding value in position (12,7) is 7, marked by a circle. Say we want to encrypt a character 'A'. The ASCII equivalent of 'A' is 65. If we convert it into base-9, it will be 027. So, 0 needs to be embedded in the first pixel, 2 in the second pixel, and 7 in the third pixel amongst the 3-pixel group. Let us try to embed 0 in the first pixel. We have first found out the elements of $C_r$, $C_c$, and $C_m$. The elements of $C_r$, $C_c$, and $C_m$ are marked with green, red, and blue dotted lines in Figure 5. We can easily mark the presence of 0 in $C_r$, $C_c$, and $C_m$. We find the least deviation in $C_r$. Then we can update the R and G values accordingly. We can easily see that the value of $(X - X_C)$ or $(Y - Y_c)$ is at most four. So there is a very negligible amount of distortion in the cover image.

*(iv) Checking the integrity of the cover_image:*

For checking whether the *cover image* has been modified or not, we first need to send an 8×8 Sudoku instance to the receiver. Receiver solves this puzzle and stores it. Then the receiver divides the *cover_image* into 64 blocks and in each of the blocks, a group of three pixels be prepared. Then the last bit of each 3-pixel group is taken and the equivalent decimal values are kept in the 8×8 matrix. If this matrix match with the already solved Sudoku puzzle, then the integrity has been maintained; otherwise, the image has been modified.

*(v) Hidden message extraction:*

For extracting the hidden message from the *stego_image*, we first need to transfer the instance of 9×9 Sudoku puzzle. Then after receiving it the receiver solves this puzzle and an 18×18 reference matrix is generated.

Then,

For each bit in a three-pixel group, we compute the following:

$X = (R\%6) + 6,$

$Y = (R\%6) + 6.$

Find the value present at (X,Y) from the reference matrix.

End for

Concatenate the three values from the three pixels.

Convert the result to base-10. This is an ASCII value of the hidden character message.

Get the character equivalent to the decrypted ASCII code.

*C. Sending of Sudoku instances:*

The sender of the secret message needs to send an 8×8 and a 9×9 Sudoku instance to the receiver. Sender encrypts the instances and sends them to the receiver.

The whole process (i.e., the proposed Steganographic scheme) is shown in Figure 6.

## IV. Results Obtained

The quality of *stego_image* is evaluated using histogram comparison in MAT Lab and embedding capacity in terms of characters. A sample result has been shown as follows.

After performing histogram analysis, it can easily be found that there are very less distortion in the *cover_image*.

The embedding capacity for the image shown in Figure 7 is calculated in the following manner:

Sample image size = 229 KB.

Number of pixels in the image = 78089.

Number of pixels required for hiding one character = 3.

Therefore, 78089/3=26029 characters.

Thus, this method can hide approximately 25 KB secret messages in the image shown in Figure 7.

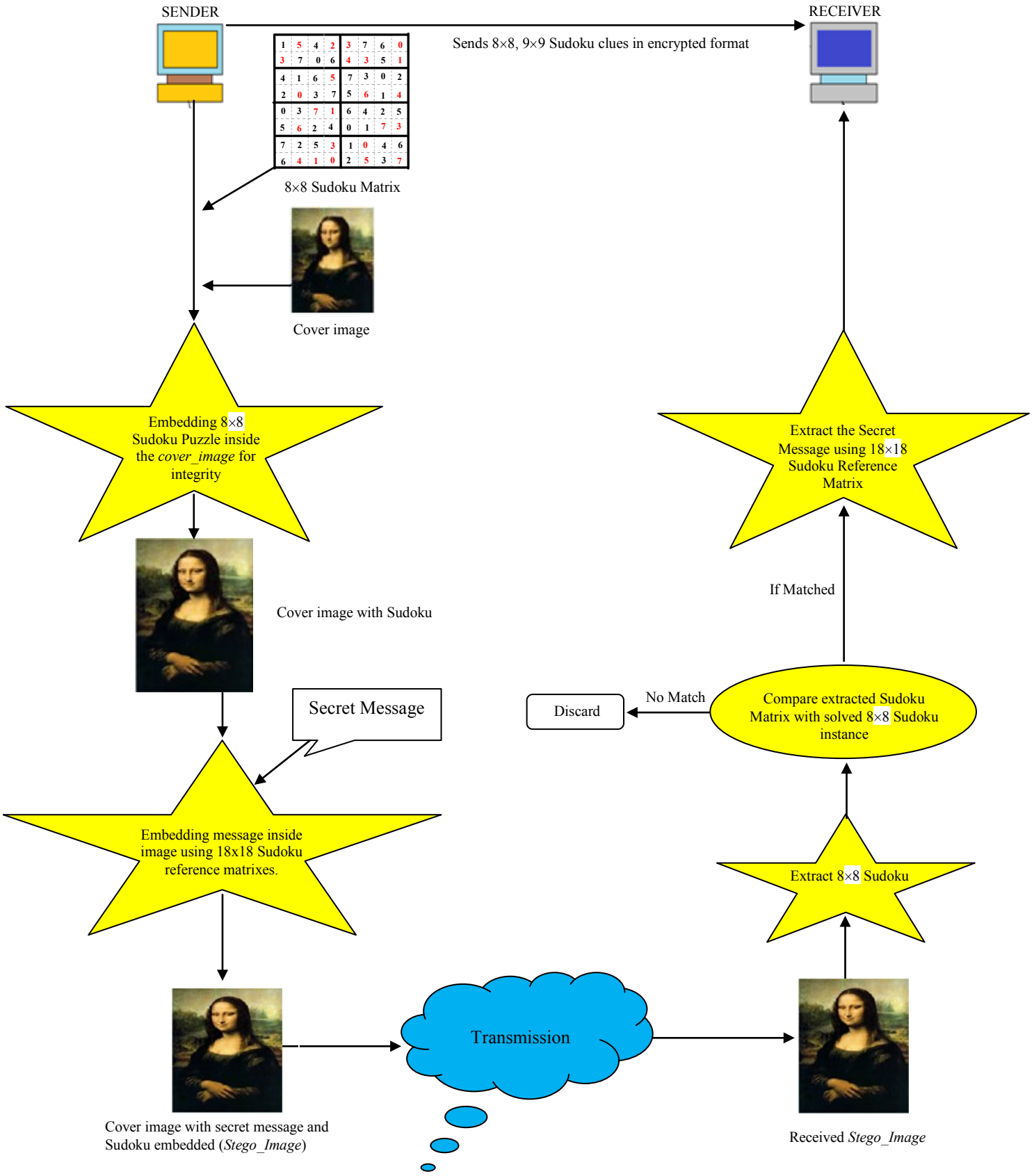



Figure 6: The proposed Steganographic scheme.

We have also performed steganalysis of the *stego_image*; we can find that the proposed scheme in this paper is safe against geometrical attack, noise based attack, and so on.

Figure 7: A sample *cover_image*.

Figure 8: *Stego image* after embedding the sample *cover_image* shown in Figure 7, in an 8×8 Sudoku and secret message.

Figure 9: Histogram of the *cover_image*, shown in Figure 7.

Figure 10: Histogram of the *stego_image*, shown in Figure 8.

*A. A brief comparison with existing method:*

The Chang *et al.*'s method [3] has used a reference matrix of size 256×256, whereas in the Shetty *et al.*'s method [7], a 27×27 reference matrix has been used. In this respect the reference matrix that we have used is of size 18×18, which is even much smaller in size. Use of smaller reference matrix also reduces the computation time. Moreover, we have embedded a separate 8×8 Sudoku matrix into the blue components of each pixel in each block, which provides an additional layer of security.

## V. Conclusion

In this paper we have proposed a Steganographic scheme, where we are using an 18×18 Sudoku reference matrix for message embedding and we are also embedding an 8×8 Sudoku, for checking whether the cover image has been modified (or not). If somehow the cover image gets modified, it can easily be detected as we have already embedded the 8×8 Sudoku matrix inside it. It can also prevent any modification, as each Sudoku puzzle matrix should have values 0 through 8 only once in the same row, in the same column, and in a minigrid. An 18×18 Sudoku reference matrix is used for hiding the secret message, whereas the entire earlier existing methods used 256×256 or 27×27 reference matrix. That is why, less computation is involved in our method. It can be claimed that our proposed scheme is more robust with less computation.

## References


[1] http://www.en.wikipedia.org/wiki/Pentomino

[2] http://www.en.wikipedia.org/wiki/Tetris

[3] C. C. Chang, Y. C. Chou, and T. D. Kieu, "An Information Hiding Scheme Using Sudoku", in *Proc. Third International Conference on Innovative Computing, Information and Control (ICICIC2008)*, June 2008.

[4] E. Cole, *Hiding in Plain Sight - Steganography and the Art of Covert Communication*, Wiley Publishing Inc., 2003.

[5] N. Jussien, *A-Z of Sudoku*, ISTE Ltd., USA, 2007.

[6] W.-M. Lee, *Programming Sudoku*, Apress, USA, 2006.

[7] B. R. Shetty, J. Rohith, V. Mukund, and R. Honwade, "Steganography using Sudoku Puzzle", in *Proc. International Conference on Advances in Recent Technologies in Communication and Computing*, pp. 623-626, 2009.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Native Language Identification using Probabilistic Graphical Models

Garrett Nicolai[1], Md Asadul Islam[1*], and Russ Greiner[1]
[1]Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada
[*] E-mail: mdasadul @ ualberta.ca

*Abstract*— **Native Language Identification (NLI) is the task of identifying the native language of an author of a text written in a second language. Support Vector Machines and Maximum Entropy Learners are the most common methods used to solve this problem, but we consider it from the point-of-view of probabilistic graphical models. We hypothesize that graphical models are well-suited to this task, as they can capture feature inter-dependencies that cannot be exploited by SVMs. Using progressively more connected graphical models, we show that these models out-perform SVMs on reduced feature sets. Furthermore, on full feature sets, even naïve Bayes increases accuracy from 82.06% to 83.41% over SVMs on a 5-language classification task.**

***Keywords— NLI, Machine Learning, SVM, Bayesian Methods, TAN***

## I. Introduction

Year-by-year, it is becoming more important to know one of a small number of languages if one wishes to succeed professionally. In particular, the number of people who speak English as a second language is growing faster than any other language. As more people devote their time and effort to learning the language, it is also becoming more important to be able to identify difficulties that may arise in the learning of English.

Native Language Identification (NLI) is the task of identifying the native language (L1) of a writer of an essay written in a second language (L2). The task has seen interest mainly in the past ten years, with a concentration of Machine Learning algorithms such as Support Vector Machines (SVMs) being applied to the problem. Probabilistic Graphical Models (PGMs) provide an alternative method of tackling the issue, and provide certain benefits over SVMs. Where SVMs can predict the L1 of the writer, PGMs can also infer likely feature groupings, given observed evidence, allowing second language teachers to adapt curricula to a student's particular needs.

Our paper is organised as follows: Section II describes the important work related to the NLI task; Section III briefly describes the data used to learn and test our models; Section IV describes the features that have been chosen to learn our models; Section V provides the methods that we use to build our graphical models; Section VI gives the results of our experiments, and Section VII provides some conclusions and areas for future work.

## II. Related Work

Koppel *et al.* [9] are one of the first groups to seriously consider the problem of NLI. Using a feature set that includes function words, character n-grams, errors, and rare part-of-speech (POS) bigrams, the authors train linear SVMs to classify English essays into five different L1s, obtaining 80.2% accuracy. Tsur and Rappoport [13] further analyse the importance of character bigrams, obtaining 65.6% across the same five L1s. Our approach takes inspiration from these papers, but we approach the problem differently. Aside from using PGMs instead of SVMs, we use a slightly different set of features. As in these papers, we originally considered only 200 features, but sort them by information gain, rather than frequency. Furthermore, we gradually increase the number of accepted features, and allow content words, as well as function words. We also use the most informative POS bigrams, instead of rare ones, and have our own implementation of spelling errors.

Wong and Dras [14] abandon SVMs in favour of a maximum entropy learner to classify essays across seven different L1s. Furthermore, the authors are more concerned with syntactic than lexical features, introducing rules generated by syntax trees as features to their model. By adding syntax rules to the feature set of previous work, they obtain around 80% classification accuracy. Although our features are mostly lexical, Wong and Dras consider using more than just 200 features for some of their feature sets. They also demonstrated that the rare bigrams of Koppel *et al*. [9] did not contribute to a classifier's accuracy, and motivated our decision to use more common and informative part-of-speech bigrams.

Bergsma *et al.* [1] tackle a slightly different task than the one that we are investigating. Using SVMs, the authors analyse scientific papers, and classify them according to three criteria: whether or not the first author's L1 is English, the gender of the first author, and whether the paper is a journal or workshop paper. This task is different from ours, but uses many of the same features that we consider, such as word n-grams and POS n-grams, as well as other features such as syntax tree rules.

Tomokiyo and Jones [12] also investigate the not-quite-NLI-task of native speaker identification, but rather than using SVMs, the authors use a naïve Bayes classifier. They consider both words and POS as features to their classifier for two types of transcribed speech: spontaneous speech and read text. We also consider words and POS, but on written text, which can be significantly different from transcribed speech, and our task is different from the task of identifying whether or not a speaker

is a native speaker. However, this work encourages us as we also use PGMs, albeit for the NLI task.

Graphical models have found uses in an area similar to NLI: text classification. Like NLI, text classification aims to use features of the text to separate documents into several classes, based on their content. However, unlike NLI, text classification is often able to look for a small set of keywords that are highly indicative of their class: business documents rarely discuss baseball, unless it is about the business side of the sport. Essays can be about various topics, yet still have the same L1.

Lam and Low [10] use Bayesian networks to perform text classification, learning the structure of their network from their data, and obtain an F-measure of 0.53. Although this task is different from ours, we use the same method as the authors of representing our documents: binary indicators of the presence or absence of features. Khor and Ting [7] also use Bayesian networks for the task of text classification, discovering that Bayesian classifiers perform admirably at the task of separating conference papers by topic, obtaining 90% accuracy when compared with human experts performing the same task.

## III. Data

Our data set is the TOEFL 2011 Corpus of Non-Native English [2], which contains 9900 essays evenly distributed across eleven languages and eight essay prompts. However, to make the task more manageable, we have reduced the data set to five languages: Chinese, French, German, Japanese, and Turkish. We thus have 4500 essays across these fives languages.

These languages were chosen to represent a set of languages that contained both linguistically related languages, such as French and German, culturally tied languages such as Chinese and Japanese, and Turkish: a language that is not culturally or linguistically related to any of the other languages. The data set was split into 90% for training and 10% for testing.

## IV. Features

Our features were selected from a set determined in previous work [11]. These features have been shown to improve classifier accuracy for the NLI task when using SVMs, and have thus been chosen as features for PGMs. All features are binary; if a document contains the feature, the value is set to 1, otherwise it is 0. The features are chosen from the training set only, and any features encountered in the test set that were not in the training set are ignored.

### A. Word Unigrams

If asked for the smallest unit of meaning in languages, many people would suggest the word. Long works are constructed of paragraphs, which are made up of sentences, which in turn are composed of words. While there are smaller units of meaning, words do contain much information.

Furthermore, writers from different linguistic backgrounds may prefer words that match ones in their first language, or make common spelling mistakes that can be captured at the word level. Unfortunately, the set of all words across all documents is rather large, and document vectors composed of words are very sparse. We perform some feature reduction, which we describe further in Section V.

### B. Character Bigrams

Even smaller than the word, characters can provide information about the tendencies of an author. Certain linguistic backgrounds may prefer certain spellings, either more reflective of the spelling of their L1, or more representative of the sound of the word. English is a very orthographically dense language, that is, one letter may be associated with many different sounds, and one sound is not necessarily tied to a single letter. Consider the sound 'ə', which is the first sound in “about” in fast speech. In this case, it is represented with an “a”, but in “burn”, it is represented by a “u”, in “woman”, by an “o”, and so on. Similarly, “a” can represent different sounds, such as the ones in “cat” and “arm”.

If a language usually uses one character for one sound, a native speaker may default to that letter if he is unsure of the spelling of a word. Character unigrams are not particularly informative: the letters in English have a given distribution amongst words, and changing the frequencies slightly contains little information. Character bigrams, on the other hand, provide context for when the letters are incorrectly used. If a document contains the bigrams “hi” and “in” for the spelling “thin” when the writer meant “then”, it provides more information than if we know that “i” was used instead of “e”.

### C. Part-of-Speech Bigrams

More information is contained in a word than just its meaning. Every word also carries grammatical functionality. Using annotated files generated by the Stanford Parser [8] for previous work, we were able to construct POS bigrams from the essays. As with characters, POS can be very informative of the writer’s L1. Languages vary considerably from one to another. Closely related languages tend to have similar grammatical structures, but unrelated languages have different syntactic rules that can be captured by POS. Consider the following example from German: *Obwohl es warm ist, will er nicht draußen gehen*, which literally translates as “Although it warm is, wants he not outside to go”, and means “Although it is warm, he does not want to go outside to play”. We can see that after the comma, we have a verb followed by a pronoun, which is a very awkward construction in English, but is perfectly fine in German. This feature may be indicative of German as L1.

### D. Spelling Errors

Although some aspects of spelling mistakes are captured by the character bigrams and word unigrams, we also have a feature that explicitly checks for spelling errors. Essays are evaluated by *ASpell*[1], and the spell-checker gives a list of recommended correct spellings. We choose the top

1 http://aspell.net

recommendation as the correct spelling, and compare it with the incorrect spelling. The two spellings are aligned using M2MAligner [5]. Consider the example misspelling of "computer" as "kompyuta". The aligner would provide the following alignment:

```
c o m p u  t e r
| | | | /\ | \/
k o m p y u t a
```

From this alignment, we get three error alignments: c-k, u-yu, and er-a. We use these alignments as features for our classifier.

## V. Methodology

The following section describes our methods for learning graphical models to classify essays by the first language of their writer. First, we describe methods used to reduce the feature set to a manageable number of input features. Secondly, we describe various structure learning methods that were applied to determine the best graph.

### A. Feature Reduction

As mentioned in Section IV, our original feature set was very large, containing approximately 50,000 features. Learning a graph on all 50,000 features would have been prohibitive, so we tried to reduce the dimensionality of our feature set.

There exist in the NLI literature several methods of reducing the number of features. Koppel *et al*. [9] and others ignore content words in essays, and only consider function words. Function words can be considered the "building blocks" of languages. They have little meaning themselves, but are necessary to build grammatically correct sentences. Function words belong to a small number of POS classes, such as prepositions, pronouns, determiners, and conjunctions. Some examples of these words include "for", "I", "the", and "and".

By reducing our word set to a list of 295 function words, and only considering features that appear in these words, we reduce our feature set from 50,000 to approximately 5000 features.

Furthermore, we use information gain on our feature sets to further reduce the number of features. Previous work [11] showed that when you only want to use a subset of features, information gain can be helpful in selecting useful features from a larger set. The equation used for information gain is provided in equation 1, where $H$ is the Entropy, defined by equation 2.

$$IG = \sum_{i=1}^{2} H(word = i) - \sum_{c=1}^{5} \sum_{i=1}^{2} H(word = i | class = c) \quad (1)$$

$$H = -P(x) \times \log(P(x)) \quad (2)$$

### B. Structure Learning

Even with a reduction to 1000 features and one class variable, the number of potential Directed Acyclic Graphs (DAG) that can be constructed is super-exponential, and we do not know which is the correct one.

We decided to build several different types of graphs in an attempt to maximize our chances of finding one that approximates the true I-Map for the data, which are described in subsections 1) through 4).

#### 1) Naïve Bayes

The very simplest method of constructing a graph is to assume that all features are independent of each other. This graph gives us no dependencies, but the next simplest method, to assume that features are independent if the class variable is known, at least provides a starting point. As we have 50,000 features, naïve Bayes was attractive, with its minimalist approach.

Khor and Ting [7] also used the naïve Bayes classifier as a baseline for text classification. While our task is not text classification, the challenges are similar. We use naïve Bayes as our simplest classifier, and as a baseline against which more connected graphs can be compared.

#### 2) Tree Augmented Naïve Bayes

Following the work of Chow and Liu, [3], we investigate the Tree-Augmented naïve Bayes (TAN) classifier. The TAN tree starts as a Naïve Bayes classifier, where all nodes are assumed independent of each other, given the class variable. A spanning tree that maximizes mutual information between nodes is then constructed.

The number of arcs in a TAN network is twice as many as in a naïve Bayes, but is still manageable. All of our features are binary, and thus, the number of parameters in the network is also approximately doubled. A TAN is able to encode some dependencies in the data. Furthermore, TAN networks scale well with the number of features, and should be able to be constructed in a reasonable amount of time. A network with more dependencies may also be more prone to over-fitting the training data.

#### 3) K2 Hill-Climbing

Although a TAN network models some dependencies in the data, it is still a graph with very few arcs, and may be missing key relationships between nodes. Consider the graph in Fig. 1. The character bigram "ck" is likely related not just to one word, but to many, such as *black, kicked*, as well as others that are not shown. Likewise, each of those words is related to $n$-1 character bigrams, where $n$ is the number of letters in the word.

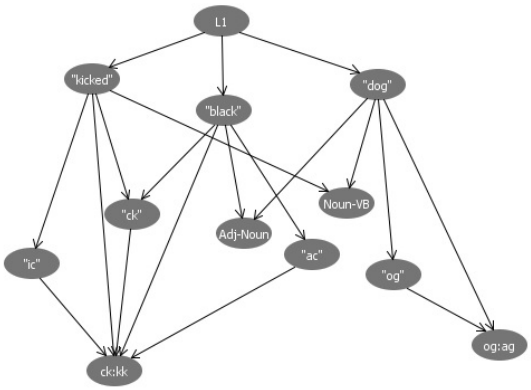



Figure 1. One of many potential graphs for a small number of features

Every word also has a likelihood of being tagged as a particular part-of-speech, and is then likely to be associated with any POS bigram that contains that POS. For example, in the sentence "The man kicked the black dog", *black* is an adjective, and is thus tied to the *Adj-NN* POS bigram. Similarly, *dog* is a noun, and would be tied to all POS bigrams that contain a noun. Likewise, words can be associated with several spelling errors. It is easy to see that one node may have relationships and dependencies with many others.

For this reason, we consider the K2 Hill-Climbing algorithm, as provided by Weka[2]. K2 uses an ordering of the variables to determine the best assignment of parents to nodes, using a particular scoring function. We have chosen the Bayes score, represented in equation 3, where $N_{ij}$ is the number of documents where a node's $i^{th}$ parent takes its $j^{th}$ value, and $N_{ijk}$ is the number of documents where a node's $i^{th}$ parent takes its $j^{th}$ value, and the node takes its $k^{th}$ value. $\Gamma(x)$ is the gamma function, which extends the factorial (in $\Gamma(x) = (x-1)!$ for an integer $x$). Given standard assumptions, this function represents the probability of a structure (marginalized with respect to the parameter values).

$$Q_{Bayes(BS,D)} = \prod_{i=0}^{n} \prod_{j=1}^{q_i} \frac{\Gamma(N'_{ij})}{\Gamma(N'_{ij} + N_{ij})} \prod_{k=1}^{r_i} \Gamma \frac{(N'_{ijk} + N_{ijk})}{\Gamma(N'_{ijk})} \qquad (3)$$

The K2 algorithm iteratively adds parents from the previous nodes in its ordering, up to the maximum specified. It chooses the parents based on which ones lead to the greatest increase in the score of the network. As we are unsure of the true ordering of the variables, we start the K2 algorithm with a random ordering, with the exception that the class variable is always the first variable in the ordering. Furthermore, we also require that each variable is dependent on the class.

Weka ensures that all variables lie within the Markov blanket of the class variable. The network is initialized as a naïve Bayes network, and edges are added to the network.

*4) Support Vector Machines*

As an alternative to graphical models, we consider SVMs. First proposed by Cortes and Vapnik [4], SVMs compute a high-dimensional hyper-plane that separates the data, using a small number of instances, known as Support Vectors, to determine the location and orientation of the hyperplane. We use a package for multi-class classification that is extended from the *SVM-Light* Package [6]. We use a linear kernel.

## VI. Results

Our data set contained 4500 essays that were divided into a 90% training set and 10% test set. We learned naïve Bayes networks, TAN networks, and more-fully connected Bayesian networks on the training data, as well as SVMs, and performed inference on the test data. We compared to the Bayesian networks. Beyond testing the accuracy of our learners, we also run an ablation study on one of our networks, and compare the results to a similar ablation study using an SVM. We wanted to determine whether the same features were contributing to the network as to the SVM.

### A. Using Function Words for Feature Reduction

Section V describes how we use function words in an attempt to reduce the dimensionality of our feature set. Table I shows results for networks built using only function words, compared with their equivalent networks using all potential words. All of these networks use 295 word features, and 200 of each of the other three feature sets, as selected by information gain.

It is quickly apparent that systems using only function words do not perform as well as those that allow any words. Naïve Bayes gains 5.6% accuracy when all words are considered, TAN gains 6.96% accuracy, the 5-parent K2 network gains 8.3%, and the SVM gains 2.69% accuracy. Using function words, we see that naïve Bayes out-performs the other two Bayesian networks, but when all words are considered, naïve Bayes is outperformed by TAN, which in turn is out-performed by a network that allows even more potential parents. For the next set of experiments, we only consider networks allowing all words from the word feature set, as the trend of this set out-performing the function word classifiers continued as the networks grew larger.

### B. Growing the Networks

In the previous subsection, all of the networks contained 895 nodes. However, there may be information in the other features that were not included in the network, and thus, we increase the size of the network in increments. First, we double the number of nodes allowed from each feature set; that is we choose the 400 most informative word unigrams, as well as 400 character bigrams, 400 POS bigrams, and 400 error alignments. We then double the size twice more, allowing 800 of each feature type, and then 1600. However, this tops out for some feature sets -- the character bigrams feature set only contains 1098 items, while the POS set only contains 1418 items.

2 http://www.cs.waikato.ac.nz/ml/weka/

*TABLE I.* ACCURACY OVER 895 FEATURES

| Classifier | Function Words | All Words |
|---|---|---|
| Naïve Bayes | 55.16 | 60.76 |
| TAN | 54.48 | 61.44 |
| K2 (Max 5 Parents) | 53.36 | 61.66 |
| SVM | 55.16 | 57.85 |

Thus, while the classifier that allows 800 of each feature type has 3200 total nodes, the classifier that allows 1600 of each will not have 6400, but rather 5716. The results are presented in Fig. 2.

Generally, we see that as the number of features increases, the accuracy of the classifier also increases. As can be expected, the increases in accuracy slow down as more features are introduced. However, the TAN network still appears to be out-performing the SVM, as the increases are leveling off. If this trend continues, it is expected that a TAN would still out-perform the SVM for feature sets for which it is very computationally expensive to build the TAN. Furthermore, as more features are included, the SVM eventually passes the naïve Bayes classifier, but the TAN classifier continues to perform better than the SVM. Somewhat surprisingly, the classifier constructed with K2 does not seem to benefit much from an increase in features, but this may be partially due to the restriction on the number of parents.

Although K2 adds the parents that most increase the score of the classifier, K2 is restricted in the nodes that it can choose as parents, while TAN is not. The parents of a node for K2 must come from the previous nodes in the ordering supplied to the algorithm, and if the ordering is less than optimal, so, too, will be the network. One final experiment was considered, where no feature reduction was performed. Due to time constraints, only the naïve Bayes model and the SVM were able to be compared. The naïve Bayes model achieved 83.41% accuracy, while the SVM obtained 82.06%. Following the results of our other experiments, it is not inconceivable that the TAN model would improve upon the naïve Bayes, further improving upon the SVMs.

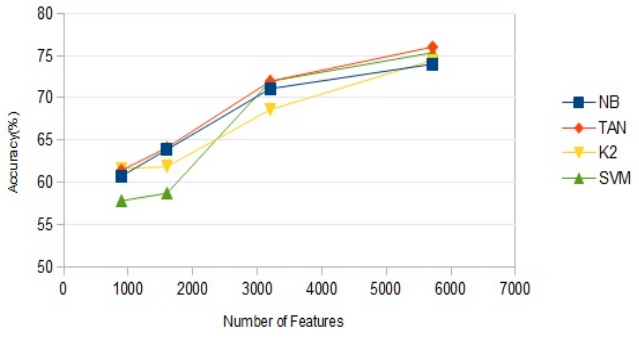



Figure 2. 10-fold cross validation accuracy, as number of features increase

### C. *Ablation Study*

To help understand which factors most contribute most to our classifiers, and whether this is constant across the various learners, we conduct an ablation study on the models created with 800 of each feature type. The results are presented in Fig. 3.

While it appears that the feature sets influence the models in a similar way, with word unigrams having the most impact, followed by POS bigrams, character bigrams, and errors, there are a few differences. The accuracy of the PGMs seems to be more dependent upon the word unigrams, while SVMs show similar falls in accuracy for POS bigrams, with larger drops for the other features than the drops observed for the PGMs.

## VII. DISCUSSION AND FUTURE WORK

We have learned probabilistic graphical models for classifying essays by the native language of their writer. Typically, this has been a task for machine learning methods such as SVMs or maximum entropy learners, but graphical models have certain innate advantages over the machine learning methods. As well as being able to take advantages in dependencies in the data that cannot be expressed by linear SVMs, graphical models can perform inference tasks that are beyond SVMs. If a user makes a set of mistakes, a graphical model can infer the likelihood that he will make other mistakes; it can also predict errors that differ between L1s . An SVM can make no such inference. This ability is of great importance to second-language teachers, who often need to adapt their teaching to suit their students. While this ability of the graphical models is important, we were more concerned with the construction of the models. We used four feature sets: word unigrams, character bigrams, part-of-speech bigrams, and errors alignments to learn four models: naïve Bayes, TAN, K2, and linear SVM. Using feature selection to allow graphs to be built in reasonable time, we found that the naïve Bayes and TAN models consistently out-perform SVMs when given the same features to learn a model.

Our models only consider five different L1s, but the literature has models that differentiate between as many as

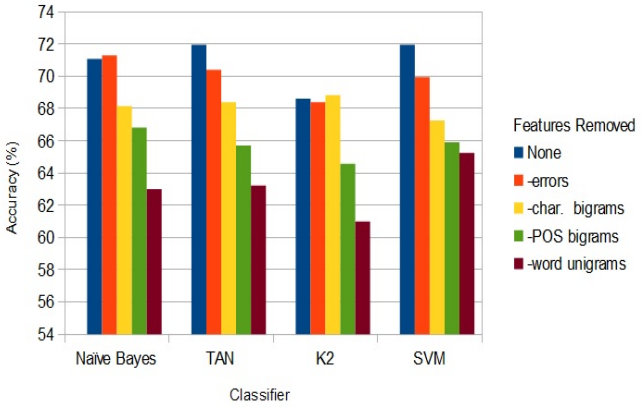



Figure 3. Ablation study for models with 3200 features

eleven. Future work can consider these other languages, and the increase in complexity that accompanies them. The feature sets that we used are only a subset of those commonly used to tackle the NLI task. It is possible that other features will further improve the accuracy of the classifier, and will do so when only a small number are considered. Furthermore, other methods of feature selection need to be considered. One disadvantage of the use of graphical models is the amount of time and resources required to build them. If we are limited in the number of features that can be used, we need to be sure that we are using the best ones available.

Acknowledgments:

The authors would like to thank Dr. Grzegorz Kondrak and Sheehan Khan for their advice and support on the research conducted in this paper. This research was partially funded by a scholarship from the National Science and Engineering Research Council of Canada (NSERC), as well as a scholarship from Alberta Innovates Technical Futures (AITF).

[1] Shane Bergsma, Matt Post, and David Yarowsky. Stylometric analysis of scientific articles. In Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 327–337. Association for Computational Linguistics, 2012.

[2] Daniel Blanchard, Joel Tetreault, Derrick Higgins, Aoife Cahill, and Martin Chodorow. Toefl11: A corpus of non-native english. Technical report, Educational Testing Service, 2013.

[3] C Chow and C Liu. Approximating discrete probability distributions with dependence trees. Information Theory, IEEE Transactions on, 14(3):462–467, 1968.

[4] Corinna Cortes and Vladimir Vapnik. Support-vector networks. Machine learning, 20(3):273–297, 1995.

[5] Sittichai Jiampojamarn, Grzegorz Kondrak, and Tarek Sherif. Applying many-to-many alignments and hidden markov models to letter-to-phoneme conversion. In Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference, pages 372–379, 2007.

[6] Thorsten Joachims. Making large scale svm learning practical. 1999.

[7] Kok-Chin Khor and Choo-Yee Ting. A bayesian approach to classify conference papers. In MICAI 2006: Advances in Artificial Intelligence, pages 1027–1036. Springer, 2006.

[8] Dan Klein and Christopher D Manning. Accurate unlexicalized parsing. In Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1, pages 423–430. Association for Computational Linguistics, 2003.

[9] Moshe Koppel, Jonathan Schler, and Kfir Zigdon. Determining an author's native language by mining a text for errors. In Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pages 624–628. ACM, 2005.

[10] Wai Lam and Kon-Fan Low. Automatic document classification based on probabilistic reasoning: Model and performance analysis. In Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on, volume 3, pages 2719–2723. IEEE, 1997.

[11] Garrett Nicolai, Bradley Hauer, Mohammad Salameh, Lei Yao, and Grzegorz Kondrak. Cognate and misspelling features for natural language identification. In The 8th Workshop on Innovative Use of NLP for Building Educational Applications, 2013.

[12] Laura Mayfield Tomokiyo and Rosie Jones. You're not from'round here, are you?: naïve bayes detection of non-native utterance text. In Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies, pages 1–8. Association for Computational Linguistics, 2001.

[13] Oren Tsur and Ari Rappoport. Using classifier features for studying the effect of native language on the choice of written second language words. In Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition, pages 9–16. Association for Computational Linguistics, 2007.

[14] Sze-Meng Jojo Wong and Mark Dras. Exploiting parse structures for native language identification. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, pages 1600–1610. Association for Computational Linguistics, 2011.

# Application of Neural Networks in Talent Management

Sajjad Waheed[1*], A Halim Zaim[2], Halil Zaim[3], Ahmet Sertbas[4], Selim Akyokus[5]
[1]Mawlana Bhashani Science and Technology University, Bangladesh
[2]Istanbul Commerce University, Turkey
[3]Fatih University, Turkey
[4]Istanbul University, Turkey
[5]Dogus University, Turkey
[*] E-mail: swaheed.iu@gmail.com

*Abstract*— **Study of talent management is getting more attentions in the recent years. It was found that there are no easy classification methods for verifying talents. This paper discusses the application of neural networks for a talent matrix based talent classification process. The proposed method is easy to implement, and free from biasing and nepotism.**

***Keywords—Talent, talent management, classification, neural networks.***

## I. INTRODUCTION

Neural networks have found many applications in various science and technological as well as in business fields, like stock market prediction [1]. Similarly, [2] is a good example of using neural networks for predicting global solar radiation (GSR). These applications were nonetheless used the neural networks' strength for predicting and classification of sample data to generate the predicted assumptions.

One of the strength of the neural networks is that it can calculate and predict almost like a human brain, as it was developed using the techniques based on the human brain. Out of various neural network models, multilayered perceptrons (MLP), backpropagation and radial basis function (RBF) networks are most prominent based on their usages in various fields. [3] discusses the basic differences between the RBF and MLP. Similarly, one can find the differences between the RBF and backprogpagation networks in [4].

One of the usages of the neural networks is calculation of correctness of data [3]. Thus, neural networks are mostly used for the error evaluation and classifications works. Despite their differences in basic structures, all these networks possess few unique characteristics. There are many hidden layers in the MLP, thus it can attempt to minimize the error rate while processing the data. The theme of a backpropagation network is that the errors from hidden layer units are determined through back propagating the errors of the immediate output layers.

The basic structure of the neural networks is almost similar. All the neural networks layers are generally input-hidden-output layers. For example, RBF has got only one hidden layer despite MLP and backpropagaition networks has got many layers, as shown in the figure 1.

In general, the neural networks are being used in different types of model prediction, speech-hand writing recognition, control for high dimensional and highly nonlinear systems, image processing, series analysis, and medical diagnosis [3]. Use of stock market prediction is another important area of neural network applications [1].

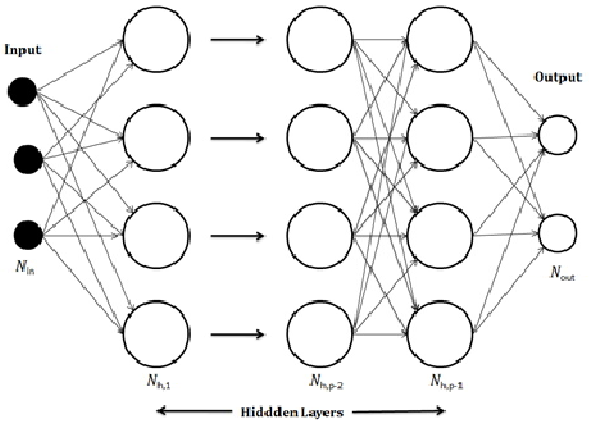



Figure 1: Generalized structure of neural networks

## II. PREVIOUS WORKS

We have examined two distinct approaches for literature reviews: first one was related to talent management and the second one was to find computer application in the talent management. Applications of various advanced techniques were found in the literature those mostly used Analytic Hierarchy Process (AHP) methods [5], [6], [7], [8]. In these studies, AHP produced serialized results of the talented employees. One of the drawbacks of the AHP was that performance and qualification of the employees were merged arbitrarily with random weighing values. Thus major capabilities of the employees – which may distinguish one employee from another – are diminished through random weights. Major talent management thoughts, however, requires distinguishable performance and qualification aspects to be measured properly [9].

Other approaches [10], [11] were limited in discovering the various characteristics of talents. Authors of [10] reported that

only the performance factors of employees were considered for defining the talent aspects. A detailed discussion was found in the [12], where the authors tried to find out sub-categories using decision tree induction under the categories Background, Previous performance evaluation, Knowledge and skill, Management skill, and Individual Quality to measure the talents amongst the academics in a university. One work [11] reported the support vector machine (SVM) based criteria evaluation of talent. The criteria were school, academic degree, job tenure, highest title, highest position, number of posts, work experience, language proficiency and degree. The result showed which criteria was more important in selecting talented people.

All these works did not yield good results for the talent management in discovering the "talent" for a talent management system. One reason could be marked as the drawback that these studies did not make any difference between the performance and qualification. Rather, in most of the cases, they used these two terms interchangeably. Thus, talents were not discovered, only few criteria were found for selecting appropriate candidates based on few selected criteria. The present study was conducted in such a way that if the detailed criteria were accounted for and measured appropriately, a better talent discovery could be made.

## III. The Present Study

As discussed in the section II, few of the talent management [5] criteria could be found as "Create good group spirit and working atmosphere", "Understanding the ability, characteristics, and needs of subordinates", "Encourage employees", and "Coaching skill and Fair and subjective evaluation of employees' performance" as Leadership Skill. Another aspect was the Interpersonal Skill that included "Understanding the meaning of other people", "Negotiation skill", "Listening skill", "Conversation skill", and "Conflict management skill". It was reported [6] that Project management, Cost management, Quality management, Purchase and machinery and equipment management are also important criteria. It was found that the performance and qualification criteria are not clearly defined or discussed in any of these articles. This paper intended to develop a mathematical model for talent management based on "performance" and "qualification". These two criterions were measured through survey questions gathered through supervisor evaluation, self-evaluation and previous experience of employees.

### A. *Talent criteria*

Analyzing [5, [6], [7], [8], [9], [10], [11], [12], this paper proposes Work quality, Work ethics, Problem solving skill, Communication skill, Leadership skill, Learning skill, and Planning and organization skill as the basic evaluation factors for supervisor evaluation and self-evaluation surveys. On the other hand, Previous Education, Test Scores, Received Training, Experiences gathered till date, Job tenure, and Expertise were the basic measuring factors for previous experience survey. The survey questions and scoring is discussed in section III (C).

### B. *Measurement through Likert scale*

The talent management criterions may include those, as mentioned in II (A), but are not strictly limited to these criteria. The criterions were clearly distinguishable as performance and qualification. This was achieved through survey questions format, with sub criteria under each of the major criterions mentioned in II (A). Each sub criteria was a question. These questions were scored using Likert scale [14]. According to [15], the Likert scale typically have five categories of response from 1 = Disagree Strongly to 5 = Agree Strongly. For a better measurement and accurate result in the present survey, an additional zero score was added so that the employees can answer for any criteria as "never achieved" or "never attained".

### C. *The survey questions and scores*

The survey questions were made in three categories: Supervisor evaluation, self-evaluation and previous experiences. Out of these, supervisor evaluations and self-evaluations were merged to generate performance scores of the employees. The previous experience was considered as qualification of the employees.

Since supervisors are the persons who directly observe the employees and know them well, their evaluations were considered with greater valuation. We proposed "performance" to be calculated by taking ninety percent of the supervisor evaluation and ten percent from the self-evaluation. In both supervisor and self-evaluations, a similar set of questioners was used. These questioners could later be used to analyze the employee's weakness, and appropriate measures could be used to improve their performance. There were a total of 34 questions in the supervisor and self-evaluation forms. Previous experience had 21 questions in the survey, and was taken to be the "qualification" of the employees.

### D. *The Talent Matrix*

Talent management was first introduced in the "The War for Talent" in a McKinsey & Company report, reports [16] and [17]. The major steps of the talent management were discovery, training and development, and retention of the talented employees within the organization. Based on the discussions in III (B) and (C), four distinguishable classes in the talent management could be arranged, as shown in the figure 2.

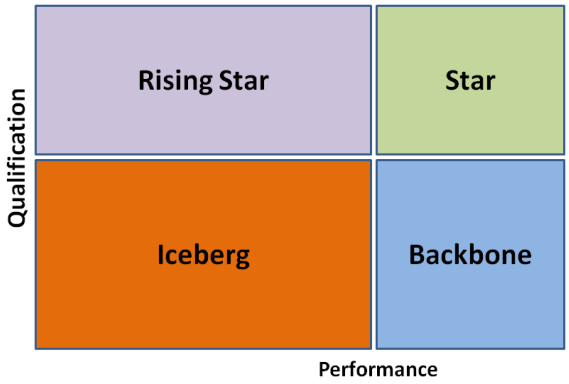



Figure 2: Talent Matrix

This classification is not necessarily to be strict along the dividing lines shown in the figure 1. Rather, based on the individual data sets, the results are mainly comparative and perceptional. One major advantage is that the talents are found to be in the star segment of the figure 2. Rest of the employees has scope to develop and enter into the star segment, which is also known as *talent pool*.

### *E. Classification methods*

The performance and qualification data was primarily classified using three statistical classification methods: mean based, z-score based [18] and Min-Max normalization [19] based classifications. The following argument was used for finding out four segments as shown in the figure 2.

Table I. The basic argument for Talent Matrix

| | ***Performance*** | ***Qualification*** |
|---|---|---|
| **Iceberg** | Low | Low |
| **Backbone** | High | Low |
| **Rising Star** | Low | High |
| **Star** | High | High |

Mean based classification was implemented along with the mean of each data column. The z-score was measured using the standard deviation and mean of each of the data columns. Lastly, the Min-Max normalization was used to scale the data for the scale of -1 and 1. Thus Talent Matrix depicted in the figure 2 was mathematically achieved.

## IV. CALCULATION AND RESULTS

The data used in this paper was collected from a reputed Turkish bank during November – December, 2012 through direct interviews between the bank employees and their supervisors. The employees also filled their previous experience forms along with the self-evaluation forms. All these three categories of data were then processed to develop the basic data set for the calculations. The neural network tools from Matlab™ were used for the calculation. Testing environment consisted of Matlab 2010a, installed on Windows 7 (64-bit) operating system.

### *A. Data preparation and classification*

All the collected data was measured on a 6-scale input. For ease of computing, numerical values were normalized into the percentiles. This normalization standardized numerical values for all data entries. Then the performance and qualification values were calculated and *n*-by-2 data set was prepared.

After data was ready according to the requirement of the system, star, rising star, backbone and iceberg classes were constructed, as mentioned in the table 1 and shown in figure 2. The data was classified using all the three classification methods, as mentioned in the section III (E).

### *B. Neural networks and Error Calculation*

As already discussed in the section I, neural networks are capable of making decisions and verifying classifications of some given data. By now, all three classification methods were used and three different classification results were obtained. These results were used as inputs of three neural networks, as mentioned in the section I.

The neural networks were configured using few common parameters:

Table II. Assigned Network Parameters

(a) Numerical parameters

| | |
|---|---|
| **Maximum Epochs** | 100 |
| **Minimum gradient** | 0.00000001 |
| **Goal** | 0.01 |
| **Learning rate** | 0.3 |

(b) Different functions

| | |
|---|---|
| **Training Function** | trainscg |
| **Adapt Function** | adaptwb |
| **Divide Function** | dividerand |
| **Performance Function** | mse |

Neural networks can efficiently measure the error rates from given data. Precisely, coefficient of determination ($R^2$), Root Mean Square Error (RMSE) and simple error rate was determined [2]. For optimal results, maximum values of $R^2$ and minimum values of RMSE and error were considered.

Table III. Results derived from neural network calculations

| | **$R^2$** | **RMSE** | **Error** |
|---|---|---|---|
| **Mean-based** | | | |
| MLP | 0.7439 | 0.2191 | 0.1312 |
| Backpropagation | 0.4466 | 0.3221 | 0.355 |
| RBF | 0.946 | 0.1006 | 0.04 |
| **z-Score based** | | | |
| MLP | 0.4618 | 0.3177 | 0.1396 |
| Backpropagation | 0.32 | 0.4537 | 0.3601 |
| RBF | 0.9589 | 0.0877 | 0.0315 |
| **Min-Max Normalization based** | | | |
| MLP | 0 | 0.441 | 0.2365 |
| Backpropagation | 0.4626 | 0.3174 | 0.3761 |
| RBF | 0.945 | 0.1016 | 0.0388 |

In all of the results shown in the table 3, the RBF network yielded the best result matching the pre-set requirement: maximum $R^2$ and minimum of RMSE and Error. Now taking out the best three classifications results and compare again among them, for the above condition.

Table IV. Best fit model

| | $R^2$ | RMSE | Error |
|---|---|---|---|
| Mean-based | 0.946 | 0.1006 | 0.04 |
| z-Score based | 0.9589 | 0.0877 | 0.0315 |
| Min-Max based | 0.945 | 0.1016 | 0.0388 |

In the table 4, maximum of $R^2$ and minimum of RMSE and Error can be found for the z-Score based classification. Thus, z-Score based classification result would be taken as the optimum result for this data set. The numbers in the braces, as shown in the table 5 below, denotes the number of employees in that class. For example, according to the table 5, there are 4 employees in the backbone class.

Table V. Final Classification of data

| Iceberg | Backbone | Rising Star | Star |
|---|---|---|---|
| B01, B04, B06, B16, B19, B20, B21, B22, B23, B24, B26, B27, B28, B29 | B07, B08, B12, B15 | B02, B03, B05, B10, B14, B25 | B09, B11, B13, B17, B18 |
| (14) | (4) | (6) | (5) |

## V. DISCUSSIONS

Talent management is a comparative, dynamic and perceptional way of judging people. This paper developed and discussed the use of neural networks techniques for the talent management process. Through this model, talents can be discovered without any bias and nepotism. $R^2$, RMSE and error rates were determined through neural networks with better accuracy. This model can be used in many areas where authorities want to find out talented people for positioning in the critical positions. The example used in the section IV yielded a very good result for the employer and the employees.


## ACKNOWLEDGMENT

The authors like to thank the Scientific Research Fund of the University of Istanbul, for supporting the PhD thesis, titled "Talent Management and Career Planning System Design", through a generous "PhD Thesis Grant", IU-BAP-21618.


## REFERENCES


[1] Elif Erdogan and Hamide Ozyurek (2012), *Estimating Stock Prices By Neural Networks*, SOSYAL ve BEŞERİ BİLİMLER DERGİSİ, Vol 4, No 1, 2012

[2] Maitha H. Al Shamisi, Ali H. Assi and Hassan A. N. Hejase (2011), *Chapter 9: Using MATLAB to Develop Artificial Neural Network Models for Predicting Global Solar Radiation in Al Ain City – UAE, Engineering Education and Research Using MATLAB*, edited by Ali H. Assi, ISBN 978-953-307-656-0, Published: October 10, 2011, page 220-238

[3] John A. Bullinaria (2004), *Radial Basis Function Networks: Applications*, Lecture slide 14, www.cs.bham.ac.uk/~jxb/INC/l13.pdf, visited 10 December, 2012

[4] Tiantian Xie, Hao Yu, Bogdan Wilamowski (2011), *Comparison between Traditional Neural Networks and Radial Basis Function Networks*, IEEE, website visited on February 2013, web link: eng.auburn.edu/~wilambm/pap/2011/Comparison%20between%20Traditional%20NN.pdf

[5] Ji-bin Ma, Hong-wei Yu, Chun-lei Gao (2010), *Study of the Issue of Personnel Promotion Based on Fuzzy Comprehensive Evaluation*, 2010 International Conference of Information Science and Management Engineering in collaboration with IEEE, pages 511 – 513

[6] Zhao Liqin, Guo Yuexian, Cui Wenming (2009), *The Application of Fuzzy Comprehensive Evaluation Methods in the Selection of a Project Manager*, 2009 Fourth International Conference on Computer Sciences and Convergence Information Technology in collaboration with IEEE, pages 1387 – 91

[7] Limin Xu and Yonggang Zhao (2009), *Fuzzy Evaluation on Administrators in Colleges and Universities Based on Improved Algorithm*, 2009 Second Asia-Pacific Conference on Computational Intelligence and Industrial Applications (PACIIA, 2009) in collaboration with IEEE, pages 185 – 188

[8] Wei Fulei (2010), *The Research on College Teacher Performance Evaluation Based on Fuzzy-AHP Method*, 2010 Second International Workshop on Education Technology and Computer Science with collaboration IEEE, IEEE 2010, DOI 10.1109/ETCS.2010.484, pages 561 – 564

[9] Clarke R & Winkler V (2006), *Reflections on talent management*, London, CIPD

[10] S.Yasodha and P. S.Prakash, *Data Mining Classification Technique for Talent Management using SVM*, 2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET), Nagercoil, Tamil Nadu, India, 21 Mar - 22 Mar 2012

[11] Hua HU, Jing Ye, Chunlai Chai (2009), *A Talent Classification Method Based on SVM*, 2009 International Symposium on Intelligent Ubiquitous Computing and Education, 15-16 May 2009, Chengdu, China

[12] Hamidah Jantan and Abdul Razak Hamdan, Zulaiha Ali Othman (2008), *Classification for Talent Management Using Decision Tree Induction Techniques*, 2nd Conference on Data Mining and Optimization 27-28 October 2009, Selangor, Malaysia, pages 15 – 20

[13] Rob Johns (2010), Likert Items and Scales, University of Strathclyde, *Survey Question Bank: Methods Fact Sheet 1*, March 2010

[14] Susan Jamieson (2004), *Likert Scales: How to (ab)use them*, Blackwell Publishing Medical Education, 38: 1212 – 1218

[15] Chambers, E.G., Foulton, M., Handfield-Jones, H., Hankin, S.M. and Michaels, E.G. (1998), *The war for talent*, The McKinsey Quarterly 3, 44–57

[16] Ed Michaels, Helen Handfield-Jones, Beth Axelrod (2001), *The War for Talent*, Harvard Business School Press Boston, Massachusetts

[17] Geoffrey R. Norman and David L. Streiner (2003), *PDQ Statistics*, 3rd edition, BC Decker Inc, Hamilton, London, UK, ISBN 1-55009-207-3

[18] T. Jayalakshmi, Dr. A. Santhakumaran (2011), *Statistical Normalization and Back Propagation for Classification*, International Journal of Computer Theory and Engineering, Vol.3, No.1, February, 2011

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Proposal of Possible OTEC Sites in Bangladesh

Shifur Rahman Shakil*,Md.Safwat Hossain and Nirjhor Tahmidur Rouf
BRAC University, Bangladesh
*E-mail: shifur.shakil@gmail.com

*Abstract*—**Bangladesh, a new name among the fast developing country, has a huge population of more than 160 Million people and a large demand for electricity. Although Bangladesh is still struggling to provide a stable energy source to the growing industrial sectors and general mass, the geographical location of this country is very advantageous for the development of renewable energy. The site of Bangladesh is naturally gifted area to establish alternate energy collection sources like Ocean Thermal Energy Conversion (OTEC). Since Bangladesh lies just beneath the tropic of cancer and on the shore of the Bay of Bengal, the two vital elements: constant sunlight and large littoral areas needed for OTEC can easily be found in this region. With huge water bodies throughout, this zone is very much suitable for developing multiple OTEC plant size.**

***Keywords— Environmentally sustainable, clean energy, possible OTEC sites, significant impact***

## I. INTRODUCTION

Ocean thermal energy conversion (OTEC) is a promising renewable energy technology which exploits the thermal gradient between warm, surface seawater and deep, cold seawater to generate electricity [1].OTEC has the least environment impact and is capable enough to provide thousands of megawatt which are urgently needed in developing countries[2].As OTEC does not emit any carbon dioxides so it has the least negative effect on the environment and this is what makes OTEC a clean, sustainable energy. A developing country like Bangladesh where power disruption become an acute problem, OTEC-seems like a likely solution to alleviate this situation as there are many sites for OTEC plants in Bangladesh which can sum up to a considerable amount of energy. Only 5% of total energy comes from renewable energy in Bangladesh [3]. Figure 1 shows a pie chart depicting the contribution of renewable energy sources in the energy sector of Bangladesh [4]. This should be mentioned that as the oceans cover around 70 percent of the total surface area of this project. OTEC has a future prospective all over the world in the field of renewable energy. At the same time the oceans can be thought of the world's largest solar energy collector and energy storage system. On an average day, approximately 60 million square kilometers of tropical seas absorb an amount of solar radiation which equals the heat content of about 250 billion barrels of oil [5].

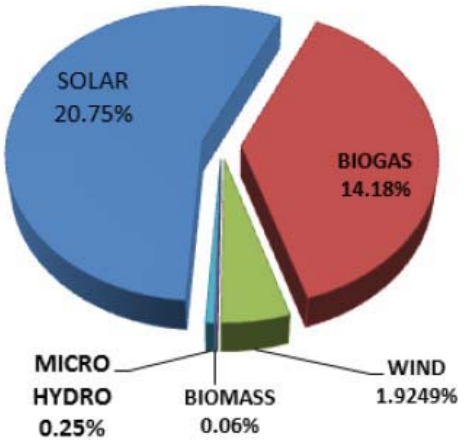



Figure 1. Contribution of different implemented renewable sources in Bangladesh

Again, properly designed OTEC plants produce negligible amount of greenhouse gas (GHG) or other pollutants. OTEC systems can produce fresh water besides electricity than can contribute a significant role in island areas where fresh water is scarce. Bangladesh is perfectly suitable for establishing OTEC power plants as many locations here fulfill the requirements.

## II. OCEAN THERMAL ENERGY CONVERSION

OTEC system generates electricity indirectly by harnessing the solar energy. It basically operates by using the difference in temperature between surface water and the deep sea water. The system actually pumps up the deep cold sea-water on its surface and then using the temperature difference between the pumped cold deep-sea water and the hot ocean surface water it runs a thermal engine which then in turns produces electricity by running a Rankine cycle to produce electricity in this system. The main advantage of OTEC system is that it not only produces electricity but as a by-product it also produces fresh water and for this reason it can be a viable choice to be used in islands where there limited fresh water. OTEC is much more advantageous compared to other renewable energy systems because the system itself is a base load source which can be operated both day and night continuously. The difference between an OTEC system and Solar PV is illustrated in Table 1. Moreover the feasibility of this system is higher because it uses the solar energy absorbed

by the ocean which is 4000 times more than the current solar technologies can harvest.

Table 1: Differences between OTEC & solar PV

| Solar PV Technologies | Ocean Thermal Electro Conversion Technology |
|---|---|
| Direct and continuous sunlight is essential to run the system with full efficiency and smoothly. | Direct/Continues sunlight is not necessary as the temperature difference between the sea surface and depth doest change rapidly. |
| Huge free area is necessary to collect sufficient amount of solar energy also additional area is necessary for battery placement and controllers. | Comparatively small operational area or project site is necessary to supply equivalent amount of electric energy. |
| Continuous maintenance and panel alignment changing is required in multiple solar energy colleting panels. | Single project site so maintenance is easy. |
| Uses solar light energy. | Uses solar heat energy. |
| Requires huge operational ground area with multiple sections containing battery storage and controllers. | Confined to a single large operational section consisting of mainly a generator. |
| Distribution is often difficult as energy is first stored in battery and then re-supplied. | Distribution is easy as output is directly connected to the main power grid. |
| Stores energy in batteries so completely battery dependent | Supplies directly to the main power grid. |

Initially the capital cost to establish an OTEC is system comparatively higher because of the large pipelines and heat exchangers but in the long run the cost is minimal because the whole system runs on the renewable solar energy. Also, the small amount of electricity that the system itself consumes is also provided from its own production. Significant approaches are still being made to lessen the initial cost.

## III. PROCESS AND TECHNOLOGY STATUS(SUITABLE FOR BANGLADESH)

To set up an OTEC system operational data is needed to earn the support required from the financial community and developers. A small overview is presented: - for a small 4-module system, a 1/5-scaled version of a 25MW module is proposed as an appropriate size. It should be noted that for typical Small Island Developing States (SIDS), a 5 MW plant can be used[6]. The OTEC process starts by collecting warm surface water using pumper then using the warmth, a heat exchanger is used to vaporize working fluids like propane and ammonia which results the vapor to expand and rotate a turbine connected to a generator. In the second phase the deep sea cold water is pumped and taken through a condenser (second heat exchanger). The coldness of the water is then used to cool down the vaporized propane/ammonia to the liquid state again. This process is done using multiple pumpers which run on a small amount of electricity and which is collected from the turbine itself. Aside from this minimum loss, a major portion of the remaining power can be acquired quite easily.

There are three types of OTEC design:

1. Closed Cycle
2. Open Cycle
3. Hybrid Cycle

The closed-cycle system uses a working fluid, such as ammonia; pumped around a closed loop. This design has three components: a pump, turbine and heat exchanger (evaporatorand condenser). Warm surface seawater is pumped through a heat exchanger that vaporizes the fluid with a low boiling point(e.g.,ammonia). Then the working fluid turns into vapour and runs the turbo generator as shown in figure 2. Later the steam is condensed using the cold from pumped deep sea water causing the liquefaction again. Thus the cycle goes round like this and keeps moving the turbine.

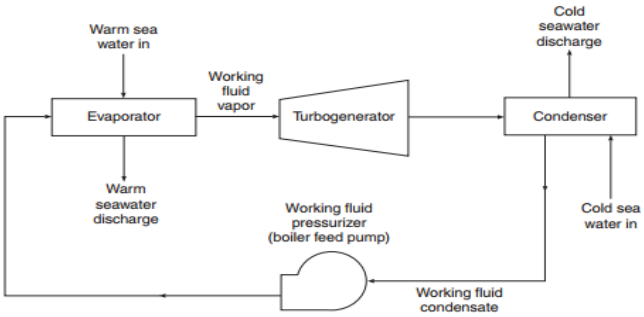



Figure 2. Closed cycle loop OTEC

The open-cycle system is quite similar to the closed-cycle system and uses the same basic components. The Open cycle OTEC uses warm surface water directly as working fluid. This system places the warm sea water in a low pressure container causeing it to boil (flash evaporator).In some schemes the expanding steam drives a low pressure turbine attached to a generator. The steam is then condensed into liquid in a heat exchanger by exposing it to cold water pumped from down below. This process is shown at figure.3.

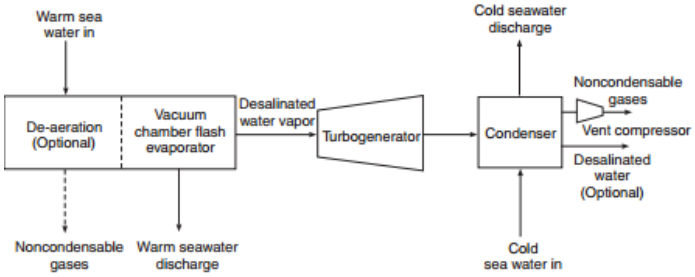



Figure 3. Open cycle loop OTEC

A Hybrid cycle system combines the features of the closed and open cycle systems. In a hybrid system, warm seawater enters a vacuum chamber, where it is flash-evaporated into steam (similar to the open-cycle evaporation process). The heat from the steam vaporizes ammonia in a separate container, and the vaporized ammonia drives a turbine to produce electricity (similar to closed-cycle process).Vaporization of the seawater removes its salt and other impurities. When the steam condenses in the heat exchanger, it emerges as fresh, pure water.

## IV. APPROACH FOR REDUCING COST OF OTEC PLANT

Researchers have been working on reducing the overall cost of OTEC systems. Keeping the cost and overall efficiency in mind Rankine cycle, Open or Claude cycle, Mist-life cycle, Kalina cycle and Uhera cycle were some concepts which were introduced in OTEC[7]. Amongst them Rankine cycle is much more accepted among the engineering community. As for the working fluids, ammonia remains the fluid of choice for closed-cycle OTEC system along with propylene. Here we have confined ourselves to a closed system using Rankine cycle and ammonia.

At OTEC conditions, there are considerable amount of experimental data available for heat exchangers using ammonia and seawater, both for boiling and condensation [08]. The reason for choosing this combination is for its practicality and the low risk factor with high yield from the seawater resources in the form of staging the Rankine cycle. Staging allows maximum potential extraction of heat and power from a set of given resources.

A Temperature-Entropy (T-S) diagram is given in figure 4. This diagram represents the characteristics of staging cycle which shows both single-stage and two-stage Carnot cycle[09]. Here Carnot's cycle is considered because it best explains the advantages of staging without any loss of the overall applicability. Also the diagram includes both cooling and heating lines from the warm and cold sea water. The single stage working fluid state points are indicated by points ABCD. The area of ABCD cycle shows that the amount of power it can be produced through by the cycle. The temperature approach at points A and C are dictated by the minimal internal temperature approach (MITA) on the evaporator and condenser.

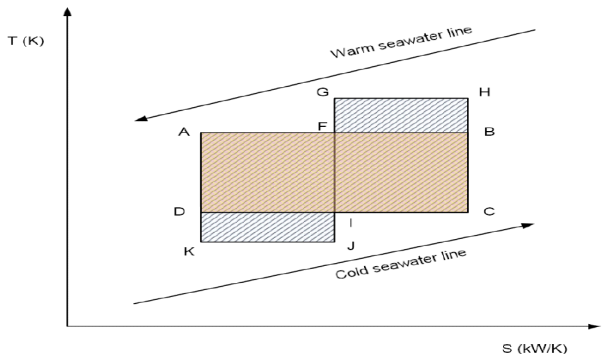



Figure 4. Cycle temperature-entropy (T-S) diagram for a Single and two-staged Carnot cycles (not to scale)

## V. ENVIRONMENTAL IMPACT

The water being from the deeper region of the ocean, it contains nutrient and comprises of a different salinity level. OTEC system involves of bringing up highly mineral enriched water from ocean depth. This mineral rich water,when dropped on the ocean surface helps in promoting the growth of photosynthetic phytoplankton. However as there is a question of maintaining the ocean mixed layer biota and maintenance of the natural surface temperature anomalies so the release of the water is conducted in a precise and controlled manner to avoid disrupting the natural balance. In case of any other off shore operations, there is a possibility of natural disruption of unintended fish or seabirds attraction, noise created from the plant might interfere with animal communication. Also, there is a fear of lubricants and anti-biofouling chemicals entering the ocean. Concerning the OTEC system, the release and redistribution of a huge amount of water periodically with cause some natural changes in geological stratification, ocean salinity, oxygen and nutrient levels near the site. It might affect the natural life habitat of some organisms. Some organisms base their behaviors on certain temperature or salinity gradients[10], while others may be affected by increased nutrient levels [11].

## VI. VIABLE OTEC SITES

The best locations for OTEC are in places where the temperature gradient of the ocean water is significantly observable. This is because the efficiency of the system depends on the difference between the higher and lower temperatures involved. Temperature difference between the ocean surface and the water at a depth of 1000 meters varies from less than 18 degree to more than 24 degree celsius. OTEC can be sited, in principle, almost anywhere in the tropical ocean-generally between Tropic of Cancer and Tropic of Capricorn. A temperature difference between the surface and 1000 meter depth more than 22 degree Celsius is usually available here[12].

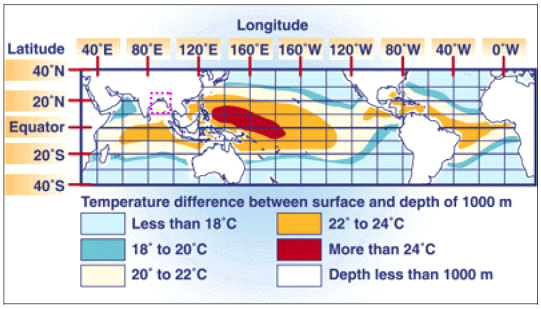



Figure 5. Suitable location for OTEC plant

From the figure 5, it can be seen that for Bay of Bengal the temperature difference between surface and sub surface (1000m) sea water is between 20 degree to 22 degree Celsius. So, OTEC project is quite to be feasible in this region.

Considering the OTEC site selection criteria as mentioned in multiple papers by L.A. Vega, who is one of the key

researchers and an internationally figure on this subject matter ,deliberating his mentioned factors like geographical location, surface and undersea temperature, land area, population, electric energy sources, electricity demand ,current availability and future electricity demand are quite important[13].

A number of possible OTEC plant site has been selected and these are: Cox's Bazaar Beach, St. Martin Island, Patenga Beach,Parki Beach,Nijhum Island,Kuakata.

### *A. Cox's Bazaar :*

Cox'z Bazaar is the world's longest unbroken 125 km natural sandy sea-beach, covering an area of 2491.86 sq.km with a population near to 51,918.It is a major fishing port and tourist spot of the country with immense need for electricity[14]. Currently there is an overall demand of 45 MW electricity where the national grid's contributes only 9 MW, the rest is fulfilled by individual diesel generators [15]. Being located on the foot of the country and directly facing the ocean this area is the most viable location to establish OTEC plant in Bangladesh.

### *B. St. Martin Island:*

This is small 8 sq.km coral island with a local population of 3,700 and is a major tourist spot of the country [16]. Each year thousands of tourists come here. In this little island even without stable electric supply form the main power grid there are 29 hotels and restaurants,4 educational institutes and cyclone shelters, hospitals, light house, government offices and a Military Naval Base[17]. Although having a perfect ocean surrounding landmass with a year round sun-light, this place has no electricity generation plant rather electricity is supplied only by private diesel generators making the electricity expensive and hazarding the environment. It is assumed that there is approximately a yearly demand of 1131.5 MW electricity throughout the island.

### *C. Patenga & Parki beach :*

Patenga and Parki beach are both located in Chittagong district and can be classified as the two major ocean facing landmasses of Bangladesh. Chittagong is a huge city covering 168 sq.km area with an approximate population of 2,579,107[18].A major sea-port and a major contributor in the national industrial activity, this city handles almost 90% of national export-import and has a daily electricity deficiency of 250-150 Megawatt daily [19]. The two sea-beaches Patenga and Parki are two viable positions here for OTEC plant location.

### *D. Nijhum Island :*

Located on the Noakhali district of Bangladesh,this small 56. sq.km island is an uprising tourist destination. This island is situated near Noakhali district where there is a calculated consumer of around 200,000 under rural electrification board ,yet each day almost all of the consumers has to face load-shedding for shortage of electricity supply[20]. Located near this huge district is Nijhum island,which can contribute to the local district grid if the natural sun-light and ocean water is used to generate electricity from renewable energy sources. OTEC is an efficient solution for this area.

### *E. Kuakata :*

Kuakata is sea-beach in the southern part of the country, located in Patuakhali district where jute-mills, ice-mills and fish farms which have a daily demand of 65 MW electricity but faces a daily shortage of nearly 40 MW [21]. Even after having a good source of procuring renewable energy, this part of the country faces a daily interruption of electricity supply. Kuakata beach facing the Bay of Bengal and getting constant sun-light throughout the year, is a very good location for building OTEC plant.

## VII. EFFICIENCY IN OTEC PLANT SITES

We can use equation 1 (Carnot's equation) to get an idea about the efficiency of OTEC [22]:

$$W = \frac{T - T_o}{T}.Q \qquad (1)$$

Here,

W = Work obtained (energy)

T = Surface water temperature (K)

T○ = The deep water temperature (K)

Q = Thermal Value

An example for a typical tropical climate is shown:

Let us take typical temperatures in the tropical area be,

T = 27 °C

T○ = 4 °C

$$W = \frac{(273+27) - (273+4)}{(273+27)}.Q = \frac{23}{300}.Q = 7.6\%.Q$$

Q being of the order of 0.5 for practical considerations in the conversation, the efficiency comes to slightly more than 3%.This may seem a low efficiency, but this may not make OTEC less competitive than other renewable energy sources. For example, it needs no more than a hydroelectric plant and the availability of resources of OTEC is vast; the capacity of OTEC is 300 times more than mankind's current total power usage [23]

## VIII. CONCLUSION

In order to tackle the rising demand of electricity in Bangladesh, potential energy sources are needs to be investigated to alleviate forthcoming energy crisis. Bangladesh is still a developing country and so keeping that in mind the new energy sources must be made inexpensive and easily procurable .Solar power is a good choice as a renewable

energy source but it requires huge funding, which would strain the economy .OTEC might have a huge initial cost but in the long run it pays off the investment. According to calculation, within five years we can profit from the system and we get clean drinkable ocean water as a byproduct. This paper looks into the prospect of OTEC in Bangladesh and the possible locations suitable for this technology. OTEC power systems satisfy the criteria for suitable energy systems for Bangladesh, and if the suggestions introduced in this paper are implemented, it would be a giant leap toward solving the present energy crisis of the country.

## REFERENCES


[1] L.A. Vega, "Ocean Thermal Energy Conversion Primer," *Marine Technology Society J.*, vol. 6, no. 4, pp. 25–35, Winter 2002/2003.

[2] A. Crump, *Dictionary of Environment and Development: People, Places, Ideas, and Organizations*,The MIT Press; 1st MIT Press Ed edition (March 2, 1993)

[3] D.M. Rahman, N.B. Sakhawat & R.Amin, "A Study on Renewable Energy as a Sustainable Alternative for Ensuring Energy Security in Bangladesh and Related Socio-Economic Aspects," *Eng. J.*, vol. 16, no. 2, Apr. 2012

[4] (2013) The Power Division website of Ministry of Power, Energy and Mineral Resources of Bangladesh. [Online]. Available: http://www.powerdivision.gov.bd/

[5] D.C. Abel & R.L. McConell, *Environmental Oceanography: Topics and Analysis.*, Massachusetts: Jones and Bartlett, 2009

[6] K. Koakutsu, N. Okubo, K. Takahashi, N. Torii and A. Fukui, "CDM Reform 2011 Veri_cation of the progress and the way forward", Institute for Global Environmental Strategies (IGES), Kanagawa, Japan,2011

[7] K. Koakutsu, N. Okubo, K. Takahashi, N. Torii and A. Fukui, "CDM Reform 2011 Veri_cation of the progress and the way forward", Institute for Global Environmental Strategies (IGES), Kanagawa, Japan,2011

[8] C. B.Panchal , D. L. Hillis, J. J. Lorenz, D. T. Young," OTEC Performance Tests of the Trane Plate-Fin Heat Exchanger," *Argonne National Laboratory Report,* ANL/OTEC-PS-8,1981

[9] T.N. Veziroğlu, *Thermal sciences 16: proceedings of the 16th Southeastern Seminar.*, Miami: University of Miami, 1986

[10] L.A. Meyer-reil & M. Koster, "Eutrophication of Marine Waters: Effects on Benthic Microbial Communities," *Elsevier*, vol. 41, no. 1-6, p. 255–263, 2000.

[11] G.W. Boehlert & B.C. Munday, "Vertical and onshore-offshore distributional patterns of tuna larvae in relation to physical habitat features," *Marine Ecology Progress Series*, vol. 107, no. 1-13, Apr. 1994.

[12] A.Lewis, S. Estefen, J. Huckerby, W. Musial, T. Pontes, J. Torres-Martinez, 2011: Ocean Energy. In IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation O. Edenhofer, R. Pichs- Madruga, Y. Sokona, K. Seyboth, P. Matschoss, S. Kadner, T. Zwickel, P. Eickemeier, G. Hansen, S. Schlömer, C. von Stechow (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[13] L.A.Vega,*Ocean Thermal Energy Conversion(OTEC)*,1999

[14] R. Ibrahim and G.M.Kamal (June 2012).Community Report Cox's BazarZila.[online] Available http://www.bbs.gov.bd/WebTestApplication/userfiles/Image/Census2011/Chittagong/Cox's%20Bazar/Cox's%20Bazar%20at%20a%20glance.pdf

[15] A. Hussain (2012) *Power outages vex Ctg dwellers*. [online]. Available : http://newagebd.com/detail.php?date=2012-03-01&nid=2422#.UaJhKthrxjE

[16] Administrator (2013) *St. Martin - Bangladesh Tourism Board*. [online]. Available : http://www.tourismboard.gov.bd/2012-03-10-05-09-35/st-martin-island.html

[17] R. Hasan (2012) *FAQs about traveling St. Martin Island in Bangladesh*. [online]. Available : http://www.bangladesh-travel-assistance.com/st-martin-island-faqs/

[18] A. Haliuc & A. Frantiuc, "A study case of Baranca drainage basin flash-floods using the hydrological model of Hec-Ras," *Georeview*, vol. 21, Jan 2013.

[19] Priyodesk (2012) *Power outage plagues Ctg*. [online]. Available: http://news.priyo.com/national/2012/03/24/power-outage-plagues-ctg-48712.html

[20] R.Shaw,F.Mallick&A.Islam,Climate Change Adaptation Action in Bangladesh,Japan,*Springer,* 2013

[21] (2013)*Patuakhali-Municipality*.[online].Available: http://www.patpou.net/

[22] M.R. Islam, *Journal of Nature Science and Sustainable Technology.*, Nova Publishers, 2008

[23] J. Baird (2013) OTEC and Energy Innovation: The Willie Sutton Approach[Online].Available*: http://theenergycollective.com/jim-baird/221801/energy-willie-suttonwill-rogers-approach*

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Technical and Economic Assessment of Biogas Based Electricity Generation Plant

Niloy Talukder[1,*], Anik Talukder[2], Debangshu Barua[3], and Anindya Das[3]
[1]Bangladesh University of Engineering & Technology (BUET), Bangladesh
[2]Khulna University of Engineering & Technology (KUET), Bangladesh
[3]Chittagong University of Engineering & Technology (CUET), Bangladesh
[*] E-mail: niloyeee147@gmail.com

*Abstract*—**To meet the ever-increasing demand of energy, renewable energy can open up new possibilities in developing countries. Renewable energy sources, such as biogas can be used for off-grid electricity generation. Although biogas is not known to be an efficient source of electricity, due to the abundance of raw materials it is still used in rural areas. Our study shows that mixing different raw materials in a proper ratio, also known as co-digestion can make a great difference in efficiency. We present in the paper, a model for biogas based low-cost off-grid power plant for rural areas. In our design, we take into account the generator size, and some maintenance tasks, such as $H_2S$ removal unit etc. We also make economic assessments of small power plants. Our model and assessments indicate that electricity generation from biogas is feasible even for relatively small application in the range of 10 – 50 kW. Potential of biogas in other applications is also mentioned here.**

***Keywords—Biogas, co-digestion, load calculation, payback time***

## I. INTRODUCTION

In the present era of human civilization, we are facing an increasing demand of power consumption and dwindling of fossil fuel reserves. Burning of fossil fuels leaves harmful effects on environment. Therefore, in order to prevent environmental deterioration and promote new sources of energy technologies especially in the developing countries, biogas can be adopted as an alternative to fossil-fuel. Recently, a report by Grameen Shakti [1] emphasizes that the biogas technology is very ideal for rural Bangladesh because it is simple to build and raw materials are available in rural household. People use animal manure to produce biogas for cooking and digested bio slurry in place of chemical fertilizer for harvesting. Domestic sized biogas plant (2-5$m^3$) is already popular in rural areas but most are reportedly very inefficient [2]. Although different implementing agencies in Bangladesh are active in promoting waste reuse technology, oftentimes the agencies do not pay attention to the design that would be more cost effective in a local setting where there is an abundance of certain type of raw materials. Furthermore, it is also required to bridge the gap between the researchers who are trying to come up with effective designs and the implementing authorities [3]. Another potential application of biogas is electricity generation which is not fully explored. Some attempts have already been taken to produce electricity from poultry waste [4]. In rural and remote areas most of the families do not get electricity from main power grid lines. Co-digestion of raw materials with appropriate proportions can yield optimum biogas production with high methane content. Methane content in biogas is almost similar to natural gas, therefore biogas can produce considerable heat during combustion. But it has relatively high amount of $CO_2$ and $H_2S$ content when compared to natural gas. High $CO_2$ content can lead to various problems when biogas is used in combustion engines; problems with $H_2S$ content are its toxicity and corrosiveness [4]. However, these harmful contents can be effectively removed and biogas can be used as alternative fuel for gas generator in small power plants to meet power demand of nearby houses.

## II. BIOGAS

Nowadays, biogas is one of the most popular renewable energy which can produce combustible gas through biological degradation of organic compounds known as anaerobic digestion. The composition of biogas is mainly methane (60-70%), carbon dioxide (30-40%), nitrogen (1-5%), hydrogen (0-3%), hydrogen sulfide (0.1-0.5%) and traces of oxygen [5]. It is a colorless, odorless gas and burns in similar way as natural gas. Raw materials for biogas are cow dung, poultry manure, agricultural residues, municipal waste, fish waste, water hyacinth etc. which are high in organic content.

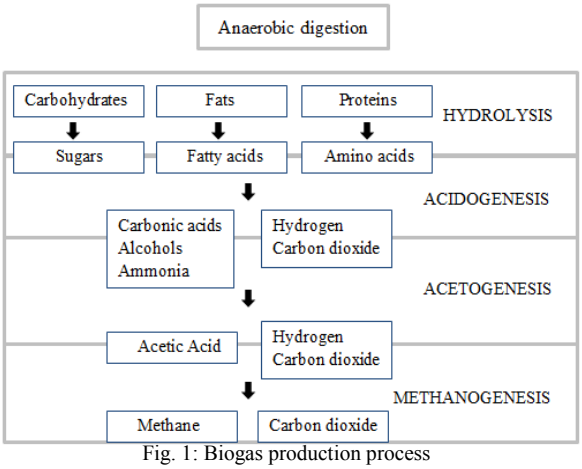



Fig. 1: Biogas production process

The whole anaerobic digestion process can be divided into three steps: hydrolysis, acidification, and methane formation.

Mainly three types of bacteria are involved in the whole process. Figure 1 summarizes the anaerobic process.

There are two types of anaerobic digestion: (1) Mesophilic digestion (35 - 40°C) and (2) Thermophilic digestion (55 - 60°C). Mesophilic digestion tends to be more robust and tolerant than the thermophilic process but gas production is less. A successful pH range for anaerobic digestion is 6.0 - 8.0 and carbon-nitrogen ratio (C/N) close to 30:1 for achieving an optimum rate of digestion. Hydraulic Retention Time (HRT) which is the average time spent by the input slurry inside digester before it comes out is 2 to 4 weeks.

### *A. Co-digestion*

Co-digestion is the simultaneous digestion of more than one type of waste in the same unit. Co-digestion of food waste and human excreta with 4:1 ratios has resulted increased biogas production than control material with 58% methane content [6]. For co-digestion of fish waste and cow dung highest gas yield was 2 L/kg when mixed with 1:1.2 ratios while only fish waste yielded 150 mL/kg [7]. Co-digestion substantially increased biogas yields by 24-47% over the control (organic kitchen waste and dairy manure only). Highest methane yield is obtained with 75% organic kitchen waste and 25% cattle manure [8]. The use of co-substrates usually improves the biogas yields from anaerobic digester due to positive synergisms established in the digestion medium and the supply of missing nutrients by the co-substrates [9].

## III. MATERIALS AND METHODS

Various advanced techniques have been developed to increase biogas yield from raw materials. But rural masses cannot afford these techniques as they are costly. They have to rely on cost effective, naturally available and easy to operate feed substrate such as poultry manure and cattle dung. Poultry manure contains high level of organic nitrogen and during anaerobic digestion ammonia nitrogen rises considerably. Part of ammonium ions are being utilized by some anaerobic bacteria but excess of ammonia can inhibit bacteria operation. To offset ammonia toxicity, dilution is required but it leads to a large increase in volume of waste and also digester volume which is non-viable in terms of cost. Another problem associated with poultry waste is that, the increase in organic load leads to reduction in methane content along with decrease in digester's performance efficiency. Hence, it is preferable to have a co-digestion of poultry manure with another substrate. Cattle dung with poultry droppings amendment may complement each other with supplying nitrogen to nitrogen deficient cattle dung and cattle dung supplying methane bacteria for a congenial digestion.

For a total 200 grams mixture of cattle dung and poultry manure, it is observed that control chicken manure produced more gas per unit weight as compared to the control cow dung but low in methane content. When poultry manure is added to cow dung up to 20%, the increase in gas production was significant and later it increased but methane volume started decreasing. It is clear from the studies that beyond 30% addition of poultry droppings may gradually lead to the digester failure due to ammonia toxicity. Table 1 gives a brief idea of this percentage [10] -

Table 1: Percentages of methane and carbon dioxide

| Digester (Cow dung : Poultry manure) | % $CH_4$ | %$CO_2$ |
|---|---|---|
| Cow dung only (control) | 55 | 45 |
| 90:10 | 58 | 42 |
| 80:20 | 66 | 34 |
| 70:30 | 52 | 48 |

So, from 200 grams cow dung and poultry manure mixture with 4:1 ratio, we get average 4300 mL good quality biogas. We get 0.022 $m^3$ biogas from one kg mixture. Results from 15 observations [10] are shown in figure 2.

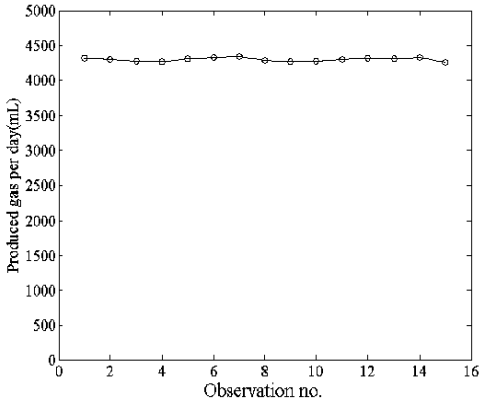



Fig. 2: Daily gas production of cattle dung digesters addition with poultry droppings.

Produced biogas which constitutes 66% methane is well compatible with natural gas which contains 65-75% methane. Another study [11] also reports optimum yield of biogas production maintaining 4:1 ratio of mixture but some less gas production.

## IV. DIGESTER DESIGN

A model plant is designed for 20 $m^3$ of total gas yield from a mixture of cow dung and poultry manure with ratio 4:1. A biogas plant has different parts: Inlet, Digester, Outlet, Gas purification unit etc. Here we only show calculation for the digester design which is the main part of the biogas plant.

To produce 20 $m^3$ biogas total raw material needed for digester is 20/0.022 = 910 kg. To maintain 4:1 ratio, 728 kg cow dung and 182 kg poultry manure is needed. A cow gives 11.5 kg cow dung and a hen gives 180 gram manure per day [12]. So, number of cows needed is 728/11.5=63.4≈64 and number of hens required is 182/0.18=1011.

Total solid content of cow dung is 17% and poultry manure is 25% [13]. So, we consider total solid content of mixture is 20%. Thus amount of total solid of 910 kg slurry is 910 x 0.2 = 182 kg. Most favorable total solid of raw material is 8%. Therefore, the quantity of input mixed material is 182/0.08 = 2275 kg. The amount of water needed to make the mixture dilute is 2275 - 910 = 1365 kg. The approximate volume of 2275 kg slurry is 2275 liters. Thumb rule of maximum volume of the slurry in the digester is about 80% to that of the total

volume of the digester. This gives total volume of digester is 2275/0.8= 2850 liters.

There is an optimum relation between the diameter D of the digester and the total volume V as D=1.3078 x $V^{1/3}$, where $f_1/D$ = 1/5, $f_2/D$ = 1/8 and H = D/2.5 [12]. This gives D = 18.54 m, H = 7.5 m, $f_1$ = 3.71 m, $f_2$ = 2.32 m. The chamber dimensions are shown in figure 3.

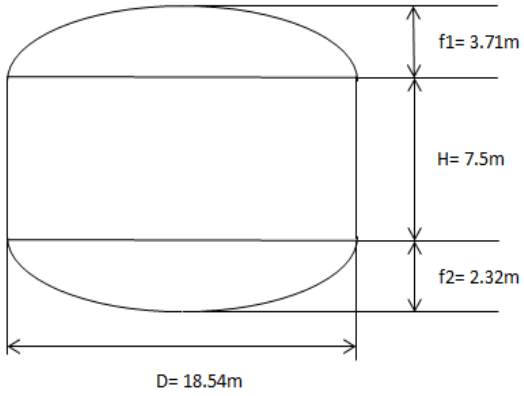



Fig. 3: Digester dimensions

## V. Power Generation

In practice, generators may run on gasoline, diesel, natural gas, propane, bio-diesel, water, sewage gas or hydrogen. Engine-driven generators fueled on natural gas fuel often form the heart of small-scale (less than 1,000 kW) combined heat and power installations [14]. In this case, biomass can be combusted, gasified, biologically digested or fermented to run a generator as an alternate source of natural gas. Difference between biomass and biofuel is that biomass can be used as fuel and biofuel is a form that can be readily used as fuel. Biomass is actually different types of waste, animal manure etc. In that sense, biofuel is a type of biomass. Several research institutions and international agencies rate biomass as one of the cheapest available renewable energy sources for power generation [15].

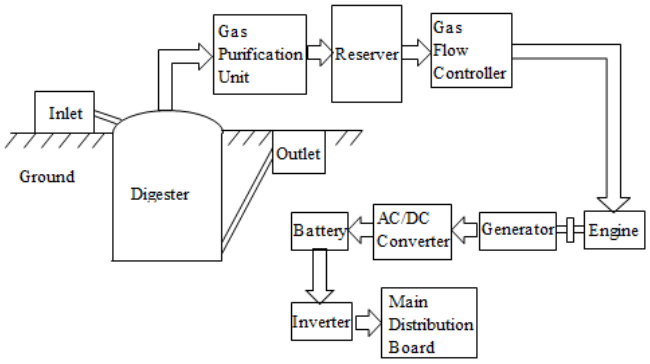



Fig. 4 : Block diagram of a typical biogas based power plant

Furthermore, conversion from biomass to electricity is a low-carbon process as the resulting $CO_2$ is captured by plant regrowth [16]. Figure 4 shows the block diagram of a biogas based small power plant. Here, biogas is used for combustion which can produce required force for the piston. However, $H_2S$ in biogas can cause corrosion of the engine and metal parts by emitting $SO_2$ from combustion, especially when engine is not in operation. As a result, $H_2S$ reduces the lifetime of the engine. Again, it pollutes workplace also.

## VI. Load Calculation

In a power system, a load curve or load profile is a chart that illustrates the variation in demand/electrical load over a specific time. Calculating load demand is very useful for a generator. Figure 5 shows a typical daily load curve for 3 houses at Unasattarpara, Raojan during summer season.

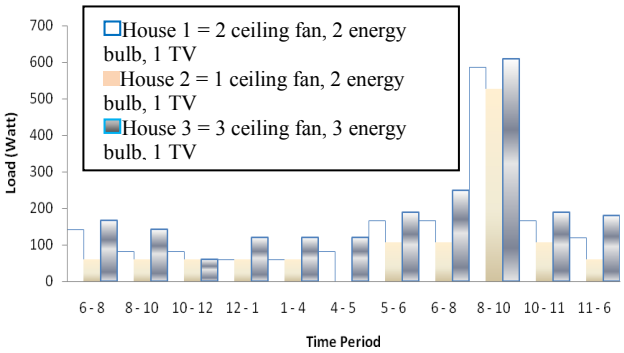



Fig. 5 : Typical daily load curve for 3 houses at Unasattarpara, Raojan

From figure 5 we see that average load per day is 3.6 kWh/house. Let us consider the average load per day is 4 kWh/day/house. In practice we see that the fuel consumption per hour is around 0.8 $m^3$ gas for producing 1 kWh in a gas generator. A digester of D= 18.54 m, H= 7.5 m can produce around 20 $m^3$ gas per day for the mixture of cow dung and poultry manure. So, it can produce daily electricity of 20/0.8=25 kWh. Then the number of houses can be covered per day is= 25/4 = 6.25 ≈ 6.

## VII. Cost Assessment

Since the system capacity is too small and there is a risk of unstable production of biogas, the produced electricity cannot be sold to national grid. But it can be used within the farm or community.

For a D= 18.54 m, H= 7.5 m digester, it is shown that the electricity generation per day is around 25 kWh. Thus the power plant capacity is 1kW. Let us consider, per unit cost of electricity for selling is 8 tk/kWh. Due to the small plant size, only maintenance cost is required. The total cost is around 1500–2000 tk/month. Total 3-4 times overhauling is required per year for a biogas based 1 kW power plant. In this case, total cost of overhauling per year = 10,000 tk. Overhauling is the procedure of cleaning up the whole plant. So, net profit per year is= (25 x 8 x 30 – 2000) x 12 – 10,000 = 38,000 tk.

The energy payback time of a power generating system is the time required to generate as much energy as was consumed during production of the system. For a 1 kW biogas based power plant, we see that cost of a 1.25 kVA generator = 24,000 tk, construction cost = 20,00,000 tk results the total installation

and commissioning cost including all utilities and accessories is around 21,00,000 tk. Hence, the payback time will be 21,00,000/38,000 ≈ 55 years. But, if we increase the plant capacity, the payback period will reduce. For a 10 kW power plant the payback period will be 48 years. Again, with the increase of plant capacity digester size will also be increased. For a 50 kW plant, the payback period will be 27 years.

For the low capital cost and easy maintenance, biogas plant can be a good replacement of IPS. Again, it can also be used as a main source of electricity where BPDB (Bangladesh Power Development Board) haven't reached yet or when load shedding will occur. Per unit cost of electricity for a diesel generator is around 15-18 tk which makes biogas a cheap alternative. But the major challenge in biogas plant is the removal of $H_2S$. If we use biogas without removing $H_2S$, it will increase engine corrosion rate, environmental pollution etc. Again, removing $H_2S$ will increase per unit cost of electricity, installation cost, payback time of the investment etc.

For this reason, $H_2S$ removal units are used only in large scale biogas plants. For small scale biogas plants there are almost zero $H_2S$ removal units in practice. This is mainly due to the lack of knowledge about $H_2S$ toxicity. In this case, 2% potassium iodide (KI) impregnated active carbon absorption technique can be used for low cost and good removal performance [17]. The activated carbon can be produced from coconut shells which is available at every rural area of Bangladesh. Then the activated carbon is impregnated with 2% KI.

The greenhouse gas effect is also low for biogas plant. Table 2 shows a comparison of greenhouse-gas production per unit of electrical power from different sources [18].

Table 2: Greenhouse-gas production per unit of electrical power

| Source of electricity | Total $CO_2$/MWh |
|---|---|
| Coal | 0.89 |
| Oil | 0.72 |
| Natural gas | 0.48 |
| Wind | 0 |
| Biogas (no burning of land fill gas in ''do-nothing'' scenario) | -2.18 |
| Biogas (burning of land fill gas in ''do-nothing'' scenario) | -0.0989 |

## VIII. Other Applications

### A. Cooking

In Bangladesh, about 44 million tons of fuel wood is used in rural areas as cooking fuel in each year [19]. These destroy our forest and put negative impact on weather and environment. They emit greenhouse gas which is causing hole in ozone layer. Again when biomasses are burnt for cooking purpose they cannot help as natural fertilizer as part of the cycle that keeps balance in the ecological system. Biogas gives easy, clean and hygienic cooking environment. People do not need to collect log as fuel for cooking. So, potential time is saved and work load is reduced. One cubic meter of biogas is equivalent to about cooking three meals for a family of five or six members [20].

### B. Bio-fertilizer

Digested bio slurry gives better results as a fertilizer than fresh cow dung and poultry litter. A comparative study of nutrient status of fresh and digested cow dung and poultry litter used in the field experiments is shown below in table 3 [21].

Table 3: Nutrient status of fresh and digested cow dung and poultry litter used in the field experiments

| Parameter | Cow dung | | Poultry litter | |
|---|---|---|---|---|
| | Fresh | Digested | Fresh | Digested |
| pH | 7.39 | 7.48 | 7.61 | 7.69 |
| Moisture | 17.90 % | 18.52 % | 16.69 % | 18.44 % |
| Nitrogen | 1.16 % | 1.55 % | 1.22 % | 1.95 % |
| Phosphorus | 0.86 % | 1.89 % | 2.01 % | 3.14 % |
| Potassium | 0.89 % | 1.04 % | 0.88 % | 1.12 % |
| Sulfur | 0.51 % | 0.81 % | 0.75 % | 1.01 % |
| Calcium | 0.74 % | 0.98 % | 2.25 % | 3.45 % |
| Magnesium | 0.28 % | 0.33 % | 0.75 % | 0.88 % |
| Iron | 0.089 % | 0.105 % | 0.125 % | 0.195 % |
| Manganese | 445 $mgkg^{-1}$ | 835 $mgkg^{-1}$ | 345 $mgkg^{-1}$ | 697 $mgkg^{-1}$ |

So, digested bio slurry can easily replace chemical fertilizer for greater food production.

## IX. Conclusion

The findings presented here are based mainly on different biogas plants and socio-economic condition in Bangladesh. From our study we can suggest that the availability of raw materials, simple setup conditions and multiple use of biogas can make it a promising technology in the developing countries like Bangladesh. It is inspiring to know from SNV report [22] that biogas plant scenario in Asia (upto end of 2010): 368,000 installations, 2 million people involved and operation rate is more than 90%. In Bangladesh there is potentiality of four million of biogas plant which can meet the household energy need of about 20% of the total families. To explore the untouched potential of biogas source following actions are recommended:

- To attract the investment in this sector government may give loan at small interest.
- Incentive can be given to private sectors for investing in electricity from biogas.
- The technologies used in the countries for successfully adopting waste to electricity system should be low in price and easily available at everywhere.
- Media can be used properly to develop consciousness among people.

## X. Future Work

In our work, co-digestion of cattle dung and poultry manure is considered. In Bangladesh which has a large population, human excreta can be a huge source of raw material for biogas production. Synergy between different raw materials needs to be determined through rigorous experiments for optimum

output. Co-digestion with other available raw materials also needs to be experimented. In urban areas municipality waste is left mostly untreated. Buriganga River, which was once the vein of Dhaka city, is now no more than a noxious drain due to indiscriminate dumping of municipality waste. But these wastes can be a powerful source of biogas production which can not only support our depleting energy sources but also avoid environment pollution. Government needs to encourage research in municipality waste treatment and take necessary steps to implement them.

## REFERENCES


[1] M. S. Islam, A. M. H. R. Khan, S. Nasreen, F. Rabbi & M. R. Islam, "Renewable Energy: The Key to Achieving Sustainable Development of Rural Bangladesh". Journal of Chemical Engineering, IEB, Vol. ChE. 26, No. 1, December 2011

[2] Rahman, Khondokar Mizanur, Elizabeth Manzanares, Ryan Woodard, and Marie K. Harder. "An evaluation of the Potential Anaerobic Digestion Feedstock, Biogas yield size and its Management Impact on Rural Society in Bangladesh."

[3] M.H. Rahman, "Energy Conversion of Organic Wastes: Prospects and Problems in Bangladesh". ISES 2001 World Solar Congress

[4] Sheikh Ashraf Uz Zaman."The Potential of Electricity Generation from Poultry Waste in Bangladesh. A Case Study of Gazipur District" Master of Science (M.Sc.) International Institute of Management University of Flensburg, Germany. August 2007

[5] Gasbook volume 1, Information and Advisory Service on Appropriate Technology.

[6] S. O. Dahunsiand, U. S. Oranusi. "Co-digestion of Food Waste and Human Excreta for Biogas Production", British Biotechnology Journal, 3(4): 485-499, 2013

[7] B. Salam, M. Islam and M. T. Rahman, "Biogas from anaerobic digestion of fish waste", ICME09-RT-06

[8] Tamrat Aragaw, Mebeaselassie Andargie and Amare Gessesse, "Co-digestion of cattle manure with organic kitchen waste to increase biogas production using rumen fluid as inoculums", Vol. 8(11), pp. 443-450, 23 March, 2013

[9] Mata-Alvarez, J.; S. Mace; P. Llabres. 2000. "Anaerobic digestion of organic solid wastes. An overview of research achievements and perspectives". Bioresource Technology 74: 3-16

[10] Shanta Satyanarayan, Rama Kant, "Co digestion of cattle dung and poultry droppings in an anaerobic bio digester". Institute of Gandhi Studies, Wardha (unpublished)

[11] Nnabuchi, M. N., Akubuko, F. O, Augustine, C. & G. Z. Ugwu. Assessment of the Effect of Co-Digestion of Chicken Dropping and Cow Dung on Biogas GenerationVolume 12 Issue 7 Version 1.0 Year 2012

[12] M.H. Rahman, M.A. Mottalib and M.H. Bhuiyan, Bangladesh, "A study on biogas technology in Bangladesh". 22nd WEDC Conference New Delhi, India, 1996

[13] Md. Forhad Ibne Al Imam, M. Z. H. Khan, M. A. R. Sarkar, S. M. Ali, "Development of Biogas Processing from Cow dung, Poultry waste, and Water Hyacinth", International Journal of Natural and Applied Science 2013; 2(1): 13-17

[14] http://en.wikipedia.org/wiki/Engine-generator

[15] http://www.giz.de/Themen/en/dokumente/giz2011-en-small-scale-electricity-generation-from-biomass-part-I.pdf

[16] Elmar Dimpl, Small-scale electricity generation from biomass

[17] Suneerat Pipatmanomai, Sommas Kaewluan, Tharapong Vitidsant, Economic assessment of biogas-to-electricity generation system with $H_2S$ removal by activated carbon in small pig farm, Applied Energy 86 (2009) 669–674

[18] J. D. Murphy, E. McKeogh, G. Kiely, Technical/ economic/ environmental analysis of biogas utilisation

[19] M S Islam and M R Islam (July 2011), Reducing biomass usage in rural areas of Bangladesh, South Asia Energy Network, Practical Action, Colombo, p. 4-5

[20] Stuckey, D. C. (1983), "The Interrogated Use of Anaerobic Digestion (Biogas) in Developing Countries: A State of the Art Review", International Reference Centre for Waste Disposal, Uberland Strasse, 133, Dubendorf, Switzerland

[21] M. S. Shahariar, M. Moniruzzaman, B. Saha, G. Chakraborty, M. Islam and S.Tahsin, "Effects of fresh and digested cowdung and poultry litter on the growth and yield of cabbage (Brassica oleracea)", Bangladesh J. Sci. Ind. Res. 48(1), 1-6, 2013

[22] Prakash C. Ghimire, "SNV supported domestic biogas programmes in Asia and Africa", Climate Summit for a Living Himalayas, Bhutan SNV side-event, November 14 2011

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Financial Feasibility Analysis of a Micro-controller Based Solar Powered Rickshaw

A. R. M. Siddique[1], A. A. Khondokar[2], M. N. H. Patoary[3], M. S. Kaiser[4] and A. Imam[5]
[1,2,3,5]Dept. of Electrical, Electronic and Communication Engineering,
Military Institute of Science and Technology, Dhaka, Bangladesh
[4]Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh
*Email:raihansiddique@ieee.org

***Abstract*—This paper presents financial feasibility analysis of a micro-controller based solar powered tricycle or rickshaw for Bangladesh. The main components of the Rickshaw are: rickshaw structure, a photovoltaic array, a battery bank, a battery charge controller, PIC micro-controllers and a DC motor. Here PIC micro-controllers are used to control the duty cycle and speed of the DC motor. The financial feasibility analysis have been done using standard methods of engineering econometrics such as Net Present Worth (NPW), Benefit-to-Cost Ratio (BCR) and Payback Periods (PP) methods. From the numerical analysis, it has been found that the proposed micro-controller based solar power rickshaw is financially feasible and Life cycle unit cost of the system is lowest compared to grid powered battery driven auto rickshaw.**

***Keywords*—Solar energy, battery operated rickshaw, economic analysis, microcontroller.***

## I. Introduction

Human energy driven tricycle or rickshaw has long been a very significant mode of transport in Bangladesh. About $38.3\%$ of the total population in Dhaka prefers to travel using rickshaw [1]. They have gained their popularity especially among the middle class and lower middle class people of the country because they offer lower travel cost and moreover these are environmental friendly. It has also been a good source of income to the poor people.

A traditional rickshaw converts human labor into rotational energy that will produce translational motion. The mean values of energy expenditure of pulling a cycle rickshaw varied from $23.5 \pm 2.66$ to $25.35 \pm 1.51$ kJ/min. Relative cardiac strain and cardiac cost indicated that the job is 'heavy' to 'very heavy' [2]. Rickshaw puller are exposed to air pollutants and they are engaged in very high level of physical activity more than 7 to 9 hours a day. Air pollutants and the reactive oxygen species causes generated by the physical activity both potentially can damage DNA [3].

With an attempt to eliminate this strenuous physical labor, some recent modifications to these rickshaws have been made where they have been equipped with a dc motor. The stored electrical energy from a battery bank runs the motor. The battery bank needs to be charged by supply lines of national grid where Bangladesh Government is giving several hundred million taka as subsidy.

A survey of High Beam research in January $2012$ shows that there were around fifty thousand battery operated rickshaw in Dhaka city [4]. This figure becomes hundred twenty-five thousand in the following six month. They also predicted that this figure will be more that three million in December 2013. If this rate continues, large amount of power is being used up charging the battery banks of these rickshaws. Several survey shows that battery operated rickshaws consume at least $300MW$ of electricity everyday to recharge their batteries [4].

In this work, we aim to eliminate this additional power demand by generating required energy through solar panel. Solar Photovoltaic (PV) is an alternative source of energy. In contrast to fossil fuel, these are renewable and environment friendly. But the initial establishment cost of the system is a bit high. Moreover, the solar radiation receives at our gourd surface is relatively high. The bright sunshine hours are vary from $4$ to $10$ hours daily and global radiation varies from $2.8$ to $6.1$ $kW/h/m^2/day$. Daily solar radiation received in Bangladesh lies between $4$ and $6.5kWh/m^2$ [6], [7].

The first solar powered auto rickshaw was introduced in England in $2000$ by Collinada Ltd, a British company along with Indian sponsor [8]. After that during $2003$ they commercially launched this solar powered rickshaw on the market but it was not well designed and energy efficient. London based research and design company Solar Lab worked on a solar powered rickshaw that would be the first human powered hybrid vehicle in $2008$. $80\%$ percent of the total power is fed by solar energy and rest of the energy is provided by the driven pedal power. After a long time, the first successfully solar powered auto rickshaw was introduced in India, Chandirarh in $2008$ which was designed by the Council of Science and Industrial Research (CSIR), named it Soleckshaw [9]. In $2012$, it was commercially launched. The Solecshaw is run by a brush less direct current (BLDC) motor. In Bangladesh, Beeva Tech Ltd launched solar powered rickshaw in January, $2012$ [5]. It is operated by $48$ volt battery, charge controller, speedometer and motor and named it Pedicab Rickshaw.

In this work, we have proposed a micro-controller based roof-top solar powered rickshaw. The financial analysis has been done using Net Present Worth (NPW), Benefit-to-Cost Ratio (BCR) and Payback Periods (PP) methods. The life cycle cost (LCC) is also analyzed for the proposed rickshaw.

The rest of the sections are: section 2 introduces solar powered rickshaw model, section 3 discusses system's cost assumption, section 4 depicts the life cycle cost analysis, section 5 showes energy scenario of Bangladesh, section 6 discuss investment evaluation analysis, section 7 includes the different evaluation methods, section 8 shows the economic

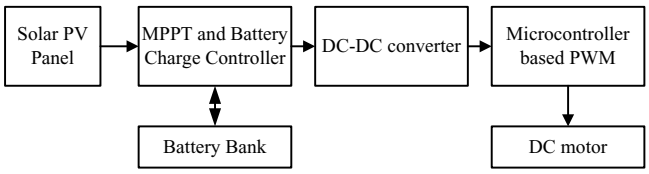



Fig. 1. Microcontroller based solar powered rickshaw

analysis of different methods and finally the work is concluded in section 9.

## II. Propped Solar Rickshaw Model

Figure 1 shows the proposed solar powered rickshaw model. The main components are: solar panel, battery charge controller, battery bank, microcontroller based motor driving circuit and dc motor.

Solar panel is a package, consists of photovoltaic cells, which converts solar radiation into electrical energy. Energy from sunlight is stored in lead acid battery bank, which is controlled by a micro-controller based charge controller. The charge controller track the maximum power point (MPP). Lead acid battery is considered because of its low cost, reliable, tolerate to overcharge, low initial impedance, deliver very high current and world most recycled product. Backup system is necessary as PV system only generates electricity when the sun is shining. Battery charge controller also protects the batteries from overcharge charging and excessive discharge. Micro-controller based motor driving circuit controls the duty cycle of the motor. There are two ways to move a vehicle whether to use a DC motor or a stepper motor. DC motors are always preferred over stepper motor when it comes to speed, size and cost. Our proposed solar powered rickshaw model is given in Figure 2.


Fig. 2. solar powered rickshaw model

In the existing solar powered rickshaw, there is no speed controller. As a result it causes unbalancing at high speed and consumes same power continuously. The existing old version of charge controller is not appropriate for this type of rickshaw. In our proposed model, we design microcontroller based charge controller and microcontroller based dc motor driving circuit. We also change some of the constructional design so that it makes the passenger comfortable as well as more mechanically efficient to drive.

## III. System Cost Assumption

The following assumptions have been made on the basis of present market prices to determine the life cycle cost (LCC) and cost of unit energy of the proposed solar powered rickshaw:

- The operation and maintenance cost is considered $10\%$ and for solar system it is $5\%$ of the total capital cost [10].
- The installation cost is considered on the basis of an average labor cost of Bangladeshi taka (Tk) $250/day$.
- Unit cost of the solar panel is Tk 60 (found in local market survey).
- Considering that the inflation rate, denoted by $\epsilon$, of conventional electricity is $5\%$.
- The LCC nonrecurring General Escalation (GE) of $3\%$ (typically the value is $3-8\%$), discount rate, denoted by$\delta$, of $7\%$ (typically the annual value is $7-15\%$), LCC nonrecurring cost factor of $0.565$ are considered [11].
- The interest rate, denoted by $\iota$, is $3\%$ [11].
- The period of analysis, denoted by $n$, is 20 years which is equal to the assumed physical and economic of the PV system [11].

## IV. Life Cycle Cost Analysis

For the LCC analysis, we consider the following five different cases

- Case A: Human driven rickshaw
- Case B: Existing battery operated rickshaw charged by the national grid.
- Case C: Solar power battery operated rickshaw
- Case D: Existing battery operated rickshaw charged by the national grid. But government is not giving subsidy.
- Case E: Microcontroller based proposed rickshaw

Table I presents the estimated costs for different cases of driving rickshaw in (Tk) [1 USD=78 Tk].

where $LCC_c$ is the Life cycle capital cost, $LCC_{NC}$ is the Life cycle non-curring cost and it includes battery replacement cost, $LCC_{O\&M}$ is the Life cycle operation and maintenance cost.

TABLE I. LCC/KWH FOR THE DIFFERENT CASES

| Cost Item | Case A | Case B | Case C | Case D | Case E |
|---|---|---|---|---|---|
| Rickshaw Structure | 16000 | 25000 | 25000 | 25000 | 25000 |
| 120W PV Array | – | – | 7200 | – | 7200 |
| Solar Charge Controller | – | – | 900 | – | 900 |
| Battery Bank | – | 18500 | 18500 | 18500 | 18500 |
| DC to DC converter | – | 3500 | 3500 | 3500 | 3500 |
| Installation Cost | – | 2000 | 2500 | 2000 | 2500 |
| $LCC_c$ | 16000 | 56400 | 65000 | 56400 | 66000 |
| License cost for 20 yrs | 6000 | 24000 | 24000 | 24000 | 24000 |
| $O\&M$ Cost in 20 yrs | 32000 | 112800 | 65000 | 112800 | 66000 |
| $LCC_{O\&M}$ | 38000 | 136800 | 89000 | 136800 | 90000 |
| $LCC_{NC}$ | – | 63382.5 | 63382.5 | 63382.5 | 63382.5 |
| Cost of Electricity | – | 226008 | – | 628905 | – |
| Total 20 yrs LCC | 54000 | 482590.5 | 217382.5 | 882484.5 | 219382.5 |
| Total Energy (kWh) | – | 43800 | 43800 | 43800 | 52560 |
| LCC/kWh (Tk/kWh) | – | 11.02 | 4.96 | 20.12 | 4.17 |

## V. ENERGY SCENARIO OF BANGLADESH

Power sector plays a key role in the development of a country. For a third world developing country like Bangladesh, this sector is still under the development process. The average electricity demand in Bangladesh was $6041MW$ and average generation was $5486MW$ on July $01, 2013$. At the same time average load shedding was $555MW$. The above statistics gives a picture of the fragile power and electricity conditions of the country. The per capita energy consumption in Bangladesh is one of the lowest ($265kWh$) in the world. In 2012, around $53\%$ of Bangladeshi population had an access to electricity [12].

Bangladesh largely depends on natural gas and hydro power stations to generate electricity. Although the country is rich in natural gas, it does not have the necessary infrastructure to exploit this natural gas to generate enough electricity to meet the total demand. Moreover the demand for power is increasing at a rapid rate as Bangladesh has a very fast population increasing rate. Use of renewable energy can help us to meet the demand. Renewable energy is the energy obtained from different natural renewable recourses mainly from sunlight, wind, tides, biofuel and geo thermal heat. The country's energy demand is actually growing annually $12\%$ and not as per government's estimation of $7\%$ [13].

Renewable energy as solar energy is a great source for solving power crises in Bangladesh. Some studies suggest that if solar energy is adopted, as much as $10,000MW$ daily of solar electricity can be generated in short and medium-runs which is equivalent to almost twice the total amount of electricity produced and supplied on the national grid.

Bangladesh is situated between $20.30^O$ and $26.38^O$ north latitude and $88.04^O$ and $92.44^O$. At this position the amount of hours of sunlight each day throughout a year is shown in Figure 3. The average solar radiation in Bangladesh varies from $5.05kWh/m^2$ day in winter to $8.03kWh/m^2$ day in summer[7].

## VI. ASSUMPTION FOR THE INVESTMENT EVALUATION ANALYSIS

The following are the assumptions considered for investment evaluation analysis:

- In Bangladesh, Electric Power is generated by gas, hydrro, steam- turbine and diesel power plants. All the generation units are inter connected through a national grid. As a result, the cost of per unit electricity is not uniform. The average per unit ($kWh$) electricity generation cost is around $14.29$Tk and the average sales rate is $3.16$Tk and for small business, it is $5.16$Tk which is considered in our proposed model [13].
- GHG emission reduction cost is calculated using RET Screen software and it is found to be.
- Incidental operating cost has been included in operation and maintenance cost.

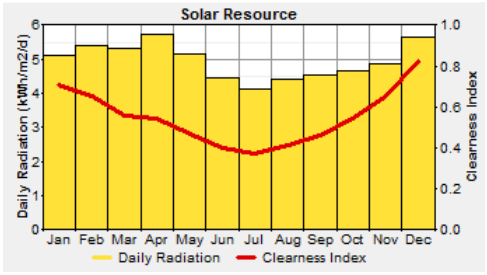



Fig. 3. Solar radiation variation for the different months of a year

## VII. METHODS OF INVESTMENT EVALUATION

Using three investment evaluation models, the investment in a solar run rickshaw can be evaluated. All these three models serve a common purpose, to determine whether an investment is economically feasible or not fulfilling the required criteria for meeting the user demands.

### A. Net Present Worth (NPW) Method

The net present worth (NPW) can be defined as the difference between the present worth of all cash inflows and outflows of a project. The mathematical expression for present worth factor (PWF) and thereby net present worth (NPW) or net present value (NPV) can be written as

$$PWF(\epsilon, \iota, n) = \left(\frac{1+\epsilon}{\iota-\epsilon}\right) \times \left[1 - \left(\frac{1+\epsilon}{1+\iota}\right)^n\right] \quad (1)$$

$$NPW = A_o \times PWF - C_c \quad (2)$$

where $A_o = A_s + GHG - (A_c + A_{NC} + A_{0\&M})$, $C_c$ is the capital cost, $A_s$ is the annual saving, $A_{NC}$ is the annual non-recurring cost, $A_{0\&M}$ is the annual operation and maintenance cost and other cost, and $GHG$ is the green house gas cost.

### B. Payback Period (PP) Method

This method calculates the number of years needed for an investor to recover the investment. This payback period is compared with maximum acceptable payback period determined by the investor. If the payback period exceeds maximum limit, then the project is unacceptable. The payback period can be expressed as

TABLE II. NPW FOR FOUR DIFFERENT CASES

| | Case B | Case C | Case D | Case E |
|---|---|---|---|---|
| $\epsilon$ | 0.05 | 0.05 | 0.05 | 0.05 |
| $\iota$ | 0.03 | 0.03 | 0.03 | 0.03 |
| $I$ | 56400 | 65000 | 56400 | 66000 |
| $A_s$ | 219000 | 237250 | 219000 | 255500 |
| $\Upsilon$ | 11300.4 | 0 | 31295.1 | 0 |
| $A_{NC}$ | 3170 | 3170 | 3170 | 3170 |
| GHG | 0 | 14234 | 0 | 14234 |
| $A_{O\&M}$ | 6840 | 4450 | 6840 | 4500 |
| $A_o$ | 197689.6 | 243864 | 177694.9 | 262064 |
| PWF | 24.63 | 24.63 | 24.63 | 24.63 |
| NPW | 5162429.9 | 5940394.1 | 4676040.5 | 6387587.3 |

$$PP = \frac{C_c}{A_o} \tag{3}$$

$$= \frac{\log\left[\frac{C_c}{A_o \times (\frac{1+\epsilon}{\epsilon-\iota})+1}\right]}{\log\left[\frac{1+\epsilon}{1+\iota}\right]} \tag{4}$$

### C. Benefit to Cost Ratio (BCR) Method

Public projects are evaluated using this investment evaluation method. The BCR is defined as

$$BCR = \frac{A_s + GHG}{(CR + A_{0\&M} + C_E + C_{NC})}. \tag{5}$$

If

$$BCR \begin{cases} \geq 1 & \text{project is feasible} \\ < 1 & \text{project is not feasible} \end{cases}$$

CR (capital recovery) is the equivalent annual capital cost $C_c$ or annualized life cycle cost (LCC), denoted by $\Xi$ and can be calculated by

$$\Xi = I \times \left[\frac{(\epsilon - \iota)}{(1+\iota)} \times \frac{(1+\iota)^n}{(1+\epsilon)^n - (1+\iota)^n}\right] \tag{6}$$

where $I$ is the Investment.

## VIII. ECONOMIC ANALYSIS

### A. Net resent Worth (NPW) Method

Table II shows the NPW of the different cases of rickshaw as mentioned in Section IV. Figure 4 and Figure 5 shows variation the net present worth of the investment with interest rate and variation the net present worth with the inflation rate for different cases respectively.

### B. Payback period (PP) Method

In this method, it is determined that the number of years required to recover the capital investment of the different cases. Table III shows the payback period of the different cases of solar power rickshaw

where $\Upsilon$ is the total annual cost of electricity

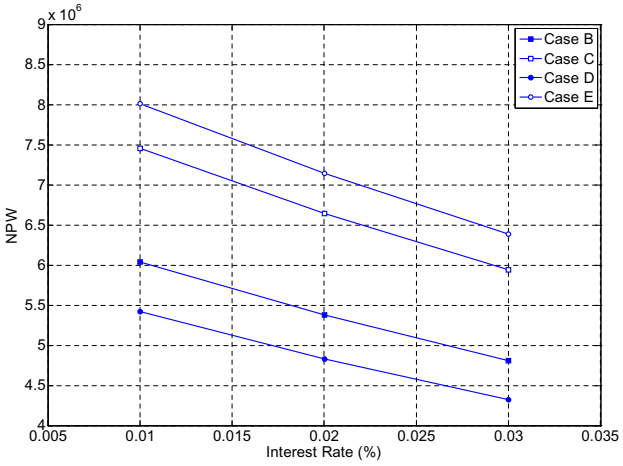



Fig. 4. Variation the net present worth of the investment with the interest rate

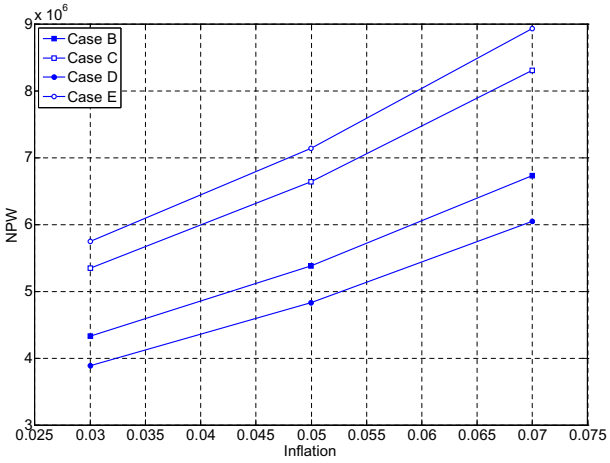



Fig. 5. Variation the net present worth of the investment with the inflation rate

TABLE III. PAYBACK PERIOD FOR FOUR DIFFERENT CASES

| | Case B | Case C | Case D | Case E |
|---|---|---|---|---|
| $\epsilon$ | 0.05 | 0.05 | 0.05 | 0.05 |
| $\iota$ | 0.03 | 0.03 | 0.03 | 0.03 |
| $I$ | 56400 | 65000 | 56400 | 66000 |
| $A_s$ | 219000 | 237250 | 219000 | 255500 |
| $\Upsilon$ | 11300.4 | 0 | 31295.1 | 0 |
| $A_{NC}$ | 3170 | 3170 | 3170 | 3170 |
| GHG | 0 | 14234 | 0 | 14234 |
| $A_{O\&M}$ | 6840 | 4450 | 6840 | 4500 |
| $A_o$ | 197689.6 | 243864 | 177694.9 | 262064 |
| PP | 0.29 | 0.27 | 0.32 | 0.25 |

### C. Benefit to Cost Ratio (BCR) Method

A BCR method is used most commonly to evaluate the feasibility study of a defined project. Table IV shows the BCR of different cases of the solar rickshaw project. Figure 6 shows the variation of BCR with the different interest rates for four cases.

TABLE IV. BCR FOR DIFFERENT CASES

| | Case B | Case C | Case D | Case E |
|---|---|---|---|---|
| $\epsilon$ | 0.05 | 0.05 | 0.05 | 0.05 |
| $\iota$ | 0.03 | 0.03 | 0.03 | 0.03 |
| $I$ | 56400 | 65000 | 56400 | 66000 |
| $A_s$ | 219000 | 237250 | 219000 | 255500 |
| $\Upsilon$ | 11300.4 | 0 | 31295.1 | 0 |
| $A_{NC}$ | 3170 | 3170 | 3170 | 3170 |
| GHG | 0 | 14234 | 0 | 14234 |
| Annualized LCC | 2334.73 | 2690.74 | 2334.73 | 2732.14 |
| BCR | 9.3 | 24.39 | 5.02 | 25.93 |

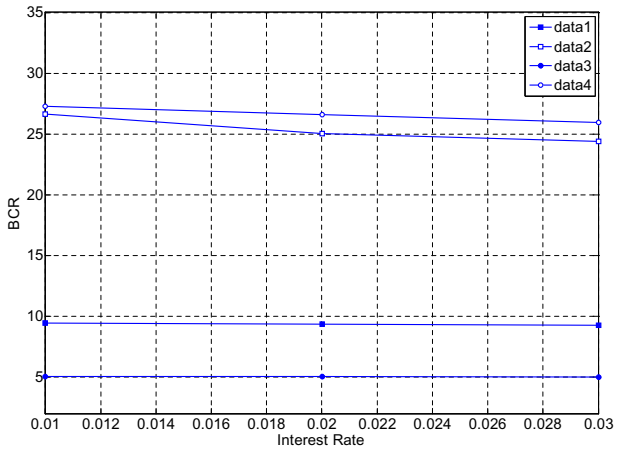



Fig. 6. Variation the BCR of different cases with the interest rate

## IX. CONCLUSION

It is clear from Table IV that the BCR of $Case - E$ is the highest among the four different cases. Thus the results presented in this paper shows that microcontroller based solar powered battery operated rickshaw is more financially feasible than others. The main advantages of such rickshaw are that these are environmental friendly and human labor is not required. Thus such rickshaw can be the good alternative than existing battery operated rickshaw charged by national grid.

## ACKNOWLEDGEMENT

This project was done by the research grant of Ministry of Science and Technology, Government of the People Republic of Bangladesh, Bangladesh. Authors would like to thank all anonymous reviewers who have evaluated this project.

## REFERENCES

[1] Ministry of Environment and Forests & Ministry of Communication, "Development of EST in Bangladesh", *5th Regional EST Forum*, Thailand , 23-25 August, 2010, http://www.uncrd.or.jp/env/5th-regional-est-forum/.

[2] A. R. Chowdhury, "Evaluation Of Occupational Health Problems Of Cycle Rickshaw Pullers And Redesign Of Cycle Rickshaw On Ergonomic Principles". *Indian Journal of RE*, Vol 23, 112-134.

[3] Alok Kumar Pandey et al, "DNA damage in lymphocytes of rickshaw pullers as measured by the alkaline comet assay", *Environmental and Molecular Mutagenesis*, Vol 47 Issue 1.

[4] Report on Battery Run Rickshaw, High Beam Research, *http://www.highbeam.com/doc/1P3-2550862211.html*, March 15, 2013..

[5] Report on Bangladesh Electric Vehicle Importers Association, The Daly Star news, *http://www.thedailystar.net*, January 15,2012.

[6] S. K. Nandi, M. N. Hoque, H. R. Ghosh, and R. Chowdhury, "Assesment of Wind and Solar Energy Resources in Bangladesh" *Arabian Journal for Science and Engineering*, December 2012.

[7] S. K. Nandi, M. N. Hoque, H. R. Ghosh, and S. K. Roy, "Potential of Wind and Solar Electricity Generation in Bangladesh" *ISRN Renewable Energy Vol. 2012*, Article ID 401761, doi:10.5402/2012/401761

[8] World Automotive Achievements- Solar Powered Rickshaw, *www.speedace.info*, February 11, 2013 .

[9] Eco-Friendly Dual-Powered Rickshaw Launched, Press Information Bureau Government of India. *http://pib.nic.in/newsite/erelease.aspx?relid=43313*, March 2010.

[10] Jain. G. C, "Development and manufacture of photovoltaic system in developing countries," *National Physical Lab* RIT Digital Media Library, March 1981.

[11] M. S. Kaiser, A. Anwar, S. K. Adyita, and R. K. Mazumder, "Financial Analysis of a Roof-top grid-connected PV system in Dhaka," *Journal of Science*, University of Dhaka, Vol. 54(2), pp:157-162, July 2006, ISSN: 1022-2502.

[12] Daily Generation Data sheet, "Bangladesh Power Development Board(BPDB)," *www.bpdb.gov.bd*, July 02, 2013

[13] Power cell of Bangladesh , "Power Division, Ministry of Power, Energy & Mineral Resources, " *http://www.powercell.gov.bd*, May 30, 2013.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Hypothetical Discussion on Windmill on Train

Jamil Sarwar Ahad[1*] and Ahnaf Shakil[1]
[1]Begum Rokeya University, Rangpur, Bangladesh
[*] E-mail: jamil.yahoo@gmail.com

*Abstract*— **Since 1887 wind has been a reliable source of power in order to address the global challenges of clean energy, climate change and sustainable development. Availability of wind has been always a matter of concern. And economically viable windmills depend most strongly on the quality of wind resource. This wind resource is available from nature. It can be achieved by many other ways. Moving bodies always get enough wind. Vehicles especially trains face much wind flow, which is more than enough to run small wind turbines. Using train compatible wind turbine likely small size, electricity can be generated nicely. We estimated that annually 27299.69378 KW power generations can be available from a particular train which is very much potential in modern energy-crisis era. We showed that, for train, using the electricity to the sub grid, the vehicles can be facilitated more efficiently.**

***Keywords—hypothetical, discussion, windmill, train, renewable energy***

## I. INTRODUCTION

Renewable energy is a key technology in the effort to reduce $CO_2$ emissions [1]. Wind power consumes no fuel and no water [2] for continuing operation, and has no emissions directly related to electricity production. The UK Energy Research Centre (UKERC) study of intermittency states that wind energy can displace fossil fuel-based generation, reducing both fuel use and carbon dioxide emissions [3].

The interest in renewable energy has been revived over last few years, especially after global awareness regarding the ill effects of fossil fuel burning [4]. Energy is the source of growth and the mover for economic and social development of a nation and its people. No matter how we cry about development or poverty alleviation- it is not going to come until lights are provided to our people for seeing, reading and working [4].

Wind is the movement of air across the surface of the Earth, affected by areas of high pressure and of low pressure [5]. Wind power, as an alternative to fossil fuels, is plentiful, renewable, widely distributed, clean, produces no greenhouse gas emissions during operation and uses little land [6]. Wind energy is one of the fastest growing forms of electric power generation. The search for pollution-free, sustainable energy has promoted this industry worldwide [7]. Wind power energy is getting more shares in the total energy production every year, with wind turbines growing bigger and bigger at the rhythm of technology does [8]. Usage of wind energy has been increased in recent times especially because of its availability and low cost rate [9]. In Bangladesh wind energy is recent concept and yet several place research is going on to measure the wind speed [9]. The total energy in Bangladesh major comes from oil and coal 56% where solar, wind and geothermal take places 5% of the total energy [9].

As of 2011, Denmark is generating more than a quarter of its electricity from wind and 83 countries around the world are using wind power on a commercial basis [10]. In 2010 wind energy production was over 2.5% of total worldwide electricity usage, and growing rapidly at more than 25% per annum. The monetary cost per unit of energy produced is similar to the cost for new coal and natural gas installations [11].

With the recent surge in fossil fuels prices, demands for cleaner energy sources, and government funding incentives, wind turbines are becoming a more viable technology for electrical power generation [12]. Fortunately there is an abundance of wind energy to be harnessed. Currently, horizontal axis wind turbines (HAWT) dominate commercially over vertical axis wind turbines (VAWT). However, VAWT do have some advantages over HAWT [12]. We encourage the maximum and efficient use of small scale wind turbines so that more energy can be achieved. A Carbon Trust study into the potential of small-scale wind energy in the UK, published in 2010, found that small wind turbines could provide up to 1.5 terawatt hours (TWh) per year of electricity (0.4% of total UK electricity consumption), saving 0.6 million tons of carbon dioxide (Mt CO2) emission savings. This is based on the assumption that 10% of households would install turbines at costs competitive with grid electricity, around 12 pence (US 19 cents) a kWh [13]. The American Wind Energy Association (AWEA) has released several studies on the small wind turbine market in the U.S. and abroad, showing that the U.S. continues to dominate the industry [14]. According to World Wind Energy Association (WWEA), it is difficult to assess the total number or capacity of small-scaled wind turbines, but in China alone, there are roughly 300,000 small-scale wind turbines generating electricity [15].

Development of wind energy use in urban environments is of growing interest to industry and local governments as an alternative to utility-based and non-renewable forms of electric production [8].

Modern world faces much difficulty to meet up the increasing demand of energy. And day by day energy becomes expensive

too. We already entered to a world in which every point of view to meet energy demand must be considered. Energy passes through (wind) and we just need to grab it. Wind energy is a vital source of electricity. And it is high time to use the most of that. Wind energy faced by the vehicles like train might be a countable and significant source of energy. As well as the roof bodies of the trains are very much compatible to small scale wind turbine which is simply unavoidable. In this paper we estimated that using train compatible small scale wind turbine, potential electricity can be generated. Annually 27299.69378 KW power can be achieved placing small turbine **(**1.3m diameter vertical axis wind turbine**)** on running vehicle like train which is very much potential in modern energy-crisis era. Consequently we recommend that efficient small VAWT wind turbine can be used to convert the massive wind power faced by train.

## II. Research Methodology

In order to fulfill the aim & objectives of the study procedure were structure as a review based-desk study. The secondary information was collected through literature review on "Hypothetical Discussion on Windmill on Train" related published & unpublished reports, scientific research publication, and newspaper & magazine articles. Therefore, most of the literatures are based on international publications. Reviewed data were analyzed interpreted and discussed to support the study objective.

## III. Windmill on Train

In energy crisis era it's smart to grab the energy passes through our environment especially the wind energy faced by running vehicles. Trains specially face potential wind energy. So the concept heating point that wind energy faced by vehicles can be a potential source of energy. As the figure: 01 describes the concept. Placing train compatible small wind turbine and generating electricity the preliminary demand of electricity of the train compartment can be fulfilled.

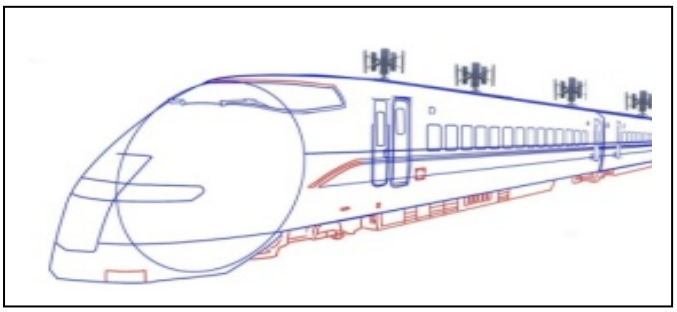


Figure: 01: Small vertical axis wind turbine on train

## IV. Choosing Vertical Axis wind Turbines (VAWTs)

Vertical-axis wind turbines (VAWTs) are a type of wind turbine where the main rotor shaft is set vertically and the main components are located at the base of the turbine. Among the advantages of this arrangement are that generators and gearboxes can be placed close to the ground, which makes these components easier to service and repair, and that VAWTs do not need to be pointed into the wind [16].

- VAWTs offer a number of advantages over traditional horizontal-axis wind turbines (HAWTs). They can be packed closer together in wind farms, allowing more in a given space. This is not because they are smaller, but rather due to the slowing effect on the air that HAWTs have, forcing designers to separate them by ten times their width. [17], [18].
- Not affected by the direction of the wind, which is useful in train areas where the wind changes direction frequently and quickly
- Unlike traditional horizontal axis wind turbines, a yaw mechanism is not needed to turn the wind turbine towards the wind
- Because of this, VAWTs outperform horizontal axis turbines in areas where a tall tower isn't feasible.
- Research at Cal Tech has also shown that carefully designing wind farms using VAWTs can result in power output ten times as great as a HAWT wind farm the same size [19].
- Ideal for both rural and urban applications, including roof top installations
- Depending on the shape of the roof, the wind flow over the roof can be concentrated, leading to an increased energy output
- Simple to install and maintain
- Quiet operation
- Pleasant appearance.

## V. Specification of VAWTs

For train compatible wind turbine we specify the parameters and issues related with it. We choose 1.3m diameter vertical axis wind turbine for the train "Rangpur Express" in Bangladesh. The roof width of "Rangpur Express" train is measured 2.73m .Here we have specification from *Beijing Asia Alliance Yike Technology Co., Ltd.* That is 500W Small Vertical Axis Wind Turbine with 1.3m Rotor Diameter and Electromagnetic Brake Method. Technical parameters of VAWTs

- Generator type: Direct Drive Permanent Magnetic VAWT, 3 phase AC
- Power: 500W
- Rated Wind Speed: 9 to 11m/s (32.4 to 39.6 km/hour)
- Rotor Diameter: 1.3m
- Start-Up Wind Speed: 2.2m/second
- Height of Blade: 1.6m
- Range Of Working Wind Speed: 2.2 to 25m/second
- Material of Blades and Numbers: Glass Fiber Reinforced Plastic/5
- Survival Wind Speed: 50m/second
- Rated Rotor Speed (m/s): 150r/min
- Output Voltage: 24 to 48V AC
- Speed Regulation Method: Electromagnetic Control

- Stop Method: Electromagnetic Brake
- Wind Energy Utilized Coefficient: 30 to 40%
- Sweep Wind Area: 1.32665m
- Range of Working Temperature: -45°C to 60°C
- Suggested Supporting Battery: 2 to 4 pieces/12V, 80Ah
- Top Weight: 50kg
- Tower Type and Height: Guy Cable Tower/ 0.37 m

## VI. How much power is in the wind?

Wind energy is the kinetic energy of air in motion, also called wind. Total wind energy flowing through an imaginary area *A* during the time *t* is:

$$E = \tfrac{1}{2} mv^2 = \tfrac{1}{2}(Avt\rho)\, v^2 = \tfrac{1}{2} At\rho v^3 \text{ [20]} \qquad (1)$$

Where $\rho$ is the density of air; $v$ is the wind speed; $Avt$ is the volume of air passing through A (which is considered perpendicular to the direction of the wind); $Av\rho$ is therefore the mass $m$ passing per unit time. Note that $\tfrac{1}{2}\rho v^2$ is the kinetic energy of the moving air per unit volume.

Power is energy per unit time, so the wind power incident on *A* (rotor area of a wind turbine) is:

$$P = E/T = \tfrac{1}{2} A\rho v^3 \text{ [20]} \qquad (2)$$

The power available in the wind, **P**, can be found from the following equation $P=\tfrac{1}{2} A\rho V^3$

Here we consider the specifications which are compatible to our selected train. Where capture area, $\mathbf{A} = \pi (D/2)^2 = 1.327326$ m [Diameter **D** = 1.3 m], wind speed **V** varies, Density of the air $\rho = 1.225$ kg/m³ at sea level, Taking Minimum Coefficient of Performance C = **0.3 [**Wind energy utilized coefficient: 30 to 40%**],** Generator Efficiencies (96%), $\eta = 0.96$. Wind power in an open air stream is thus *proportional* to the *third power* of the wind speed; the available power increases eightfold when the wind speed doubles. Wind turbines for grid electricity therefore need to be especially efficient at greater wind speeds. The maximum mechanical power that can be got from our wind turbine depends on both the rotational speed and on the undisturbed wind speed, as shown in the picture below. Vertical axis wind turbine (VAWT) spins like a top with an axis of rotation perpendicular to the wind [21].

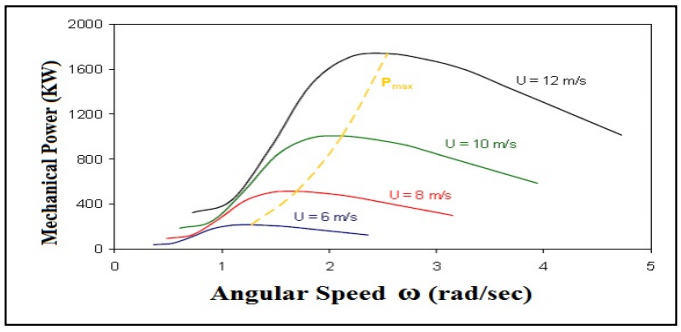



Figure: 02. Mechanical power and rotational speed for different wind speed [22].

The power available in the wind can be expressed as

$$P=\tfrac{1}{2} C_p \eta A \rho v^3 \qquad (3)$$

Where **A** is the capture area, **ρ** is the density of the air, and **V** is the wind speed. η is the generator efficiency. The power actually captured by the wind turbine rotor, $\mathbf{P^r}$, is some fraction of the available power, defined by the coefficient of performance, $\mathbf{C_p}$, which is essentially a type of power conversion efficiency:

$$C_p = P^r / P \qquad (4)$$

$$\text{So, } P=\tfrac{1}{2} C_p A \rho v^3 \qquad (5)$$

Where $\mathbf{C_p}$ is the power coefficient, **ρ** is the density of the air, **A** is the swept area of the turbine, and **v** is the wind speed [23].

The maximum theoretical value of the coefficient of performance is **0.593**, a value determined by a fluid mechanics constraint known as the **Betz limit**. According to Betz's law, no turbine can capture more than 16/27 (59.3 percent) of the kinetic energy in wind. The factor 16/27 (0.593) is known as Betz's coefficient. Practical utility-scale wind turbines achieve at peak 75% to 80% of the Betz limit. [24], [25]. Actual coefficients of performance are less than this limit due to various aerodynamic and mechanical losses. Here calculated value of the coefficient of performance is 0.49. Actually for a given turbine design, $\mathbf{C_p}$ is a function of Tip Speed Ratio (TSR). As shown in the curves in Figure: 03, there is a tip speed ratio for which the power capture is maximum. Comparisons of the various wind turbine types in Figure: 03 show how inefficient the drag-based Savonius turbine is compared to the lift-based turbines. The Darrieus turbines and the HAWT have similar values of the maximum coefficient of performance, but the HAWT can operate at much higher tips speed ratios (faster rotation speeds or lower wind speeds) [26].

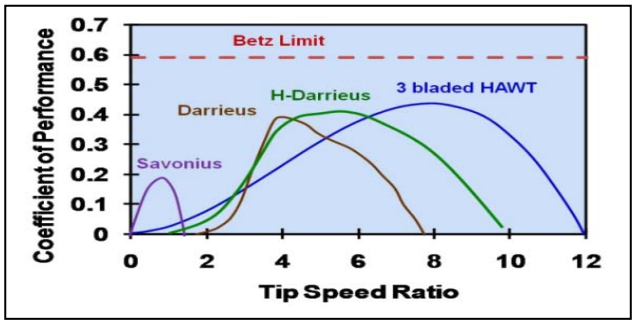



Figure: 03. Coefficient of power vs tip speed ration for different wind turbines [26].

Again an increase in rotor radius leads to greater maximum power coefficients, but they are achieved at greater tip speed ratios, so the little the radius, the less tip speed ratio is necessary to work at maximum power coefficient.

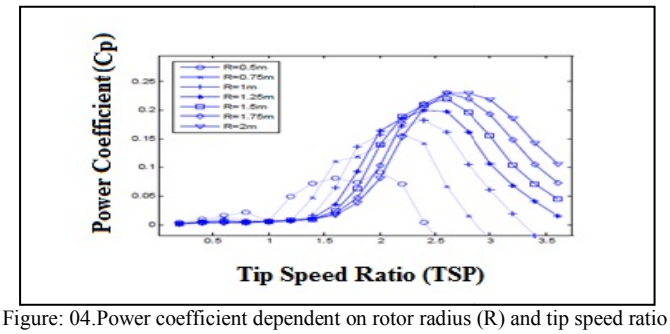



Figure: 04.Power coefficient dependent on rotor radius (R) and tip speed ratio (TSR) maintaining constant swept area [8].

The electrical power output from the generator is less than the power captured by the rotor, due to losses in both the gear train and generator:

$$P^T = C_p \eta_g \eta_b (\tfrac{1}{2} A V^3) \quad (6)$$

Where $\eta_g$ and $\eta_b$ are efficiencies (power output over power input) for the generator and the gearbox. Gearbox efficiencies are typically 90- 95%, while generator efficiencies range from 50% (for a car alternator) to better than 80% for a high quality, grid-connected model.

## VII. Windmill Setup

It is to set wind turbine on the roof top of the train containing a safety case to keep windmill safe. It may seem like extra portion of the train but it is a part of the train. Strong set up with safety case is the base of windmill. In Bangladesh train faces 15m/s to 25m/s wind force which is enough for our small vertical axis wind turbine with 1.3m rotor diameter and electromagnetic brake method. However, based on current turbine technology, for wind energy to be economically viable it has to deliver to a wind turbine an average annual wind speed of at least 5.36 m/s and above [27].

We choose "**Rangpur Express**" for our measurement. This train uses Rangpur-Dhaka-Rangpur route. Theoretically measuring wind velocity faced by "**Rangpur Express**" we became sure that on running train with small wind turbine can produce potential amount of electricity. The energy produced by this rotations is given to the main shaft (or to a gear box, if it is present) and from here to the electrical generator, that provide the electricity to the grid [22].

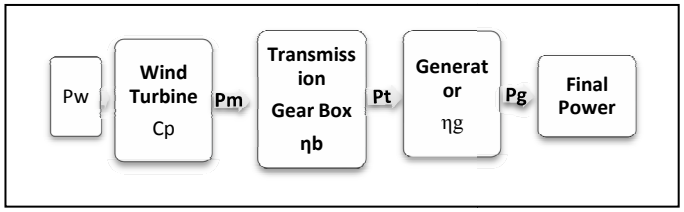



Figure: 05. Power generation

## VIII. Power Output

Theoretically our specified wind turbine generates energy as stated in the TABLE I. (taking wind energy utilized coefficient: 30%, which is minimum)

TABLE I. Hourly Energy Output

| Wind Speed (m/s) | Hourly Energy Output(Wh) |
|---|---|
| 0 | 0 |
| 2 | 1.87312 |
| 3 | 6.32178 |
| 4 | 14.98496 |
| 5 | 29.2675 |
| 6 | 50.57424 |
| 7 | 80.31002 |
| 8 | 119.87968 |
| 9 | 170.68806 |
| 10 | 234.14 |
| 11 | 311.64034 |
| 12 | 311.64034 |

Power curve show the relation between the rotary speed of the wind turbine and the produced power, for a range of wind speeds [28]. The power curve for a wind turbine shows this net power output as a function of wind speed.

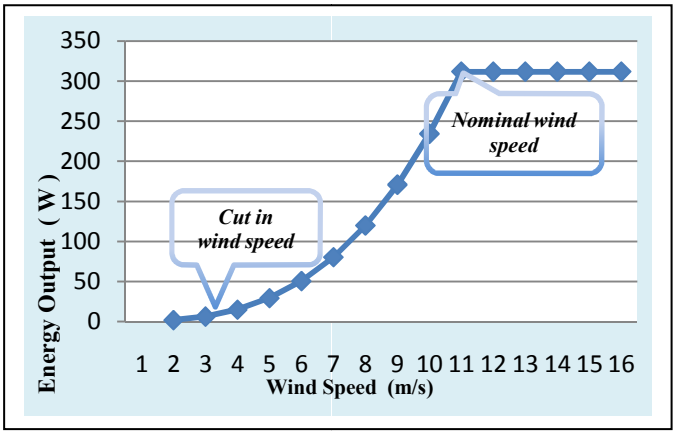



Figure: 06 Hourly wind turbine power curve

Starting from the 2m/s to 11m/s, the curve goes maintaining straight line which reflects the range of the turbine.

1. Cut in wind speed [6]: This is the wind speed at which the wind turbine will start generating power. Here typical cut-in wind speeds are 2 to 4 m/s.
2. Nominal wind speed [6]: This is the lowest speed at which the wind turbine reaches its nominal power output. Above this speed, higher power outputs are possible, but the rotor is

controlled to maintain a constant power to limit loads and stresses on the blades.

3. Cut-out wind speed [6]: This is the highest wind speed which the turbine will operate at. Above this speed, the turbine is stopped to prevent damage to the blades. Here typical cut-out wind speeds are 11m/s to 12 m/s

At rated speed annually we can achieve 2729.969378 KW (speed per hour*24*365) from just one wind turbine.

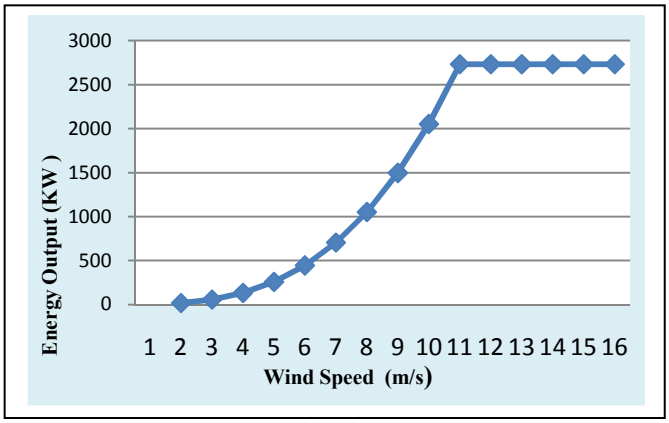



Figure: 07 Annual wind turbine power curve

## IX. Annual Energy Generation and Capacity Factor

The power curve combined with the annual wind speed distribution can be used to estimate how much energy a wind turbine could generate in typical year. Specifically, the power at each wind speed is multiplied by the number of hours per year that the wind blows at that speed to estimate how much energy is generated at each wind speed. This is then summed to get the annual energy generated. At 12m/s wind speed we consider 2729.969378 KWh is the annual calculated annual energy output for single wind turbine.

Since wind speed is not constant, a wind farm's annual energy production is never as much as the sum of the generator nameplate ratings multiplied by the total hours in a year. The ratio of actual productivity in a year to this theoretical maximum is called the capacity factor. Typical capacity factors are 15–50%, with values at the upper end of the range in favorable sites and are due to wind turbine improvements. [29], [30][1]. The capacity factor of a wind turbine is the total annual energy generated divided by the energy that could be generated if it were running continuously at rated capacity 24 hrs a day for 365 days a year. Capacity factor for the rated speed (12m/s) is

(2729.969378KW) / (1.0367KW * 8760 hr) = 30.06%

Studies have indicated that 20% of the total annual electrical energy consumption may be incorporated with minimal difficulty [31]. These studies have been for locations with geographically dispersed wind farms, some degree of dispatchable energy or hydropower with storage capacity, demand management, and interconnected to a large grid area enabling the export of electricity when needed. Beyond the 20% level, there are few technical limits, but the economic implications become more significant. Electrical utilities continue to study the effects of large scale penetration of wind generation on system stability and economics. [32], [33]–[35].

[1] For example, a 1MW turbine with a capacity factor of 35% will not produce 8,760 MWh in a year (1×24×365), but only 1×0.35×24×365 = 3,066 MWh, averaging to 0.35 MW.

## X. Amount of Electricity Produce

Thus, while the output from a single turbine can vary greatly and rapidly as local wind speeds vary, as more turbines are connected over larger and larger areas the average power output becomes less variable and more predictable [33].

TABLE II. Calculated Annual Energy Output

| Wind Speed (m/s) | Energy Output *Single Wind Turbine* (KW) | Energy Output *10 Wind Turbine* (KW) |
|---|---|---|
| 4 | 131.2682496 | 1312.682496 |
| 7 | 703.5157752 | 7035.157752 |
| 12 | 2729.969378 | 27299.69378 |

Table II shows that at 12m/s rated wind speed a single turbine can generate 2729.969378 KW energy output.

## XI. Recommendation

Lights, fans, water filter; mobile-laptop charging plugs are the basic for the train passengers. It is smart to use ten pieces train compatible wind turbine which can generate annually at least 27299.69378Wh (1.3m rotor diameter VAWT) at speed 12m/s. It requires more consideration if better amonut of electricity generates, which may facilitate the train even more.

While train stops, it faces low wind speed then an energy storage system can supply power, getting researching from windmill at first. So the installment of effective energy storage system can enhance the use of power outcome from windmill.

We think this is high time to consider seriously about this massive wind power. We also recommend taking initiatives and better consideration of this as an industry. Especially those countries use trains as mass transport, they can think about it.

## XII. Conclusion

The objective of the study was to hypothetically describe the concept "*Windmill on Train*". That is, the heavy wind force faced by speedy train cannot be ignored easily. It's discussed that the strong wind force faced by train can be converted as potential electricity which is very much significant for the internal use of the vehicle.

Question is how much economically extractable power in the world we have? The total amount of economically extractable power available from the wind is considerably more than present human power use from all sources [36]. The study says that heavy wind forces faced by speedy trains are fully economically extractable.

## REFERENCES


[1] Shin-Ichi Inage, *Prospects for Large-Scale Energy Storage in Decarbonised Power Grids*, International Energy Agency Working Paper Series, 2009.

[2] Gross, Robert; Heptonstall, Philip; Anderson, Dennis; Green, Tim; Leach, Matthew; Skea, Jim, *The Costs and Impacts of Intermittency*, March 2006.

[3] UK Energy Research Council. p. iii. ISBN 1-903144-04-3. , [Online]. Available:http://www.ukerc.ac.uk/Downloads/PDF/06/0604Intermittency/0604IntermittencyReport.pdf. Retrieved 2010-07-27.

[4] Nasima Akter, , *Alternative Energy Situation In Bangladesh Country Review*, Paper presented at the Regional Training Orientation Course on Alternative Energy Technologies, APPROTECH ASIA Philippine Social Development Center, Philippines, July 1997, PP. 11.

[5] Anon (2010). "What is wind?" *Renewable UK: Education and careers*. Renewable UK. Retrieved 9 April 2012.

[6] Fthenakis, V.; Kim, H. C., "Land use and electricity generation: A life-cycle analysis". *Renewable and Sustainable Energy Reviews* **13**(6–7): 1465. doi:10.1016/j.rser.2008.09.017, 2009

[7] EFMR Monitoring Group, Inc.4100 Hillsdale Road Harrisburg, *Wind Energy Lesson Plans and Resource Guide*, PA 17112, [Online]. Available: http://www.efmr.org

[8] Javier Castillo, *Small-Scale Vertical Axis Wind Turbine Design* Bachelor's Thesis Degree program in Aeronautical Engineering, Tampere University of Applied Sciences, December 2011.

[9] Mohammed Shahed Hossain, Minhaz Rahman, A.K.M. Nazrul Islam and Md. Fuhad Hassan, *Strategy for Promotions and Development of Wind energy in Bangladesh,* in Proc. National Seminar on Renewable Energy 2011, Dhaka, Bangladesh. April 6-8, 2011.

[10] REN21, "Renewables 2011: Global Status Report". p. 11. , 2011.

[11] "International Energy Outlook". Energy Information Administration. p. 66, 2006.

[12] Jon DeCoste, Denise McKay, Brian Robinson, Shaun Whitehead, Stephen Wright, *Vertical Axis Wind Turbine*, MECH 4010 Design Project, Group- 2, Department of Mechanical Engineering, Dalhousie University, December 5, 2005.

[13] "Smale scale wind energy". [Online]. Available: http://carbontrust.com.

[14] Hugh Piggott (6 January 2007), *Wind speed Measurement In The City*. [Online]. Available: http://Scoraigwind.com..

[15] World Wind Energy Association Statistics, [Online]. Available: http://www.wwindea.org/webimages/Half-year_report_2012.pdf

[16] Sandra Eriksson, Hans Bernhoff, Mats Leijon, *Evaluation of different turbine concepts for wind power*, Renewable and Sustainable Energy Reviews 12 (5): 1419–1434, doi:10.1016/j.rser.2006.05.017., ISSN 1364-0321, retrieved 2010-04-26, June 2008.

[17] Fish hold the key to better wind farms, [Online]. Available: http://www.mnn.com/earth-matters/energy/stories/fish-hold-the-key-to-better-wind-farms

[18] Kathy Svitil. (2012) *Wind-turbine placement produces tenfold power increase, researchers say*, retrieved 2012-07-31, [Online]. Available: http://phys.org/news/2011-07-wind-turbine-placement-tenfold-power.html

[19] Mielke, Erik, *Water Consumption of Energy Resource Extraction, Processing, and Conversion*, Harvard Kennedy School, October 2010.

[20] Grogg, K. (2005), "*Harvesting the Wind: The Physics of Wind Turbines"*, *Physics and Astronomy Department Carleton College* (Comp Paper),http://apps.carleton.edu/campus/library/digitalcommons/assets/pacp_7.pdf

[21] Kathryn E. Johnson, "*Adaptive Torque Control of Variable Speed Wind Turbines"*, National Renewable Energy Laboratory., Tech. Rep. (August 2004), http://www.osti.gov/bridge

[22] Mohammed Shahed Hossain, Minhaz Rahman, A.K.M. Nazrul Islam and Md. Fuhad Hassan, *Strategy for Promotions and Development of Wind energy in Bangladesh,* Proceeding of the National Seminar on Renewable Energy-2011 Dhaka, Bangladesh. April 6-8, 2011.

[23] Chiras, D., *Wind Power Basics: A Green Energy Guide*. Gabriola Island, BC, Canada: New Society Pub, 2010.

[24] Tony Burton David Sharpe, Nick Jenkins, Ervin Bossanyi, *Wind Energy Handbook* , John Wiley and Sons 2001 ISBN 0 471 48997 2 page 65

[25] Tony Burton David Sharpe, Nick Jenkins, Ervin Bossanyi, *Wind Energy Handbook* , John Wiley and Sons 2001 ISBN 0 471 48997 2, chapter 5 *Shrouded Turbines*

[26] *Wind Turbines (2008),*Mechanical Engineering Factsheet, Boston University

[27] IEC (ND). Iowa Wind Energy Data, Iowa Energy Centre Publications, USA.

[28] R. Bravo, S. Tullis, S. Ziada*, Performance Testing of a Small Vertical-Axis Wind Turbine,* Mechanical Engineering Department, McMaster University,

[29] Wind Power: Capacity Factor, Intermittency, And What Happens When The Wind Doesn't Blow?, http://www.ceere.org/rerl/about_wind/RERL_Fact_Sheet_2a_Capacity_Factor.pdf

[30] Shahan, Zachary (27 July 2012). *"Wind Turbine Net Capacity Factor – 50% the New Normal?"* [Online]. Available: http://cleantechnica.com.

[31] "Tackling Climate Change in the U.S". American Solar Energy Society. January 2007.

[32] The UK System Operator, *National Grid*. 2008. Archived from the original on 25 March 2009.

[33] Final Report – 2006, Minnesota Wind Integration Study, The Minnesota Public Utilities Commission. 2006.

[34] "Impact of Wind Power Generation in Ireland on the Operation of Conventional Plant and the Economic Implications", ESB National Grid. February 2004. p. 36.

[35] Sinclair Merz, *Growth Scenarios for UK Renewables Generation and Implications for Future Developments and Operation of Electricity Network,s* BERR Publication URN 08/1021 June 2008.

[36] Hurley, Brian. *How Much Wind Energy is there?* Wind Site Evaluation Ltd". Claverton Group. Retrieved Mar 2009.

# Design and Cost benefit Analyses of a Proposed Solar Recharging Station in Bangladesh

Sikder Sunbeam Islam [1,*], Mohammad Mahmud Hasan[2]
[1, 2]Department of EEE, International Islamic University Chittagong, Bangladesh
* E-mail: sikder_islam@yahoo.co.uk

**Abstract— In Bangladesh the electric vehicles powered by rechargeable batteries are becoming popular day by day. But it is a matter of great regret that a big amount of power is being used daily to recharge their batteries at the recharging stations. As Bangladesh faces acute power crisis this has become a big challenge. Renewable energy resources can play an important role in this regard. This paper has proposed a PV based recharging station for electric vehicles and made a detail analysis on it.**

***Keywords— Solar System, Recharging Station, Power System.***

## I. INTRODUCTION

An enormous amount of energy is extracted, distributed, converted and consumed daily in the present world. Today's global energy production is high, in fact 83% [1]. The energy demand in Bangladesh is increasing day by day. At present the power demand in Bangladesh is about 6500 MW which was 6000 MW in the last year (2012) and the generation capacity is around 5500MW but peak demand is estimated to exceed 5,500 MW [2]. Moreover, our average growth demand is around 7% which indicates that if we do not add 10% every year we are going to face a big electricity crisis problem in the near future [12]. On the other hand now about 0.8% renewable energy is being shared to the total generation. Moreover, the Government is committed to facilitate both public and private sector investment in renewable energy projects according to the Renewable Energy Policy approved in 2008 [2]. So, it is the right time to invest in this sector especially in solar energy. The monthly global solar insolation and daily average bright sunshine hour in Dhaka city are presented in Fig. 1.1 and 1.2 respectively [3].

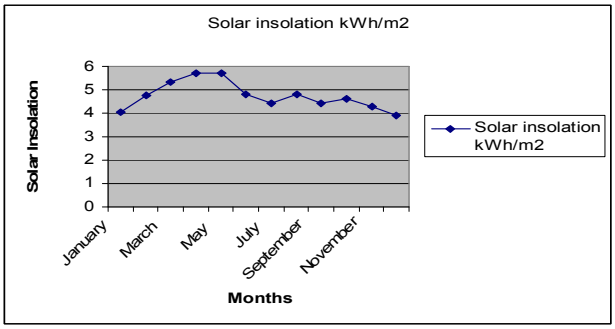



Figure 1. Monthly solar insolation at Dhaka

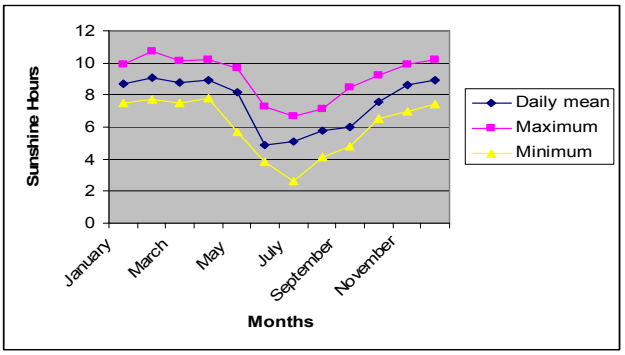



Figure 2. Daily average bright sunshine hours at Dhaka City

From the Figure 1, it is found that the average solar insolation is 4.73 for Dhaka and from Fig.2, the daily average bright sunshine is 7.55 hours for Dhaka. The solar resource is abundant in Bangladesh as the tropic of cancer passes closely through the geographical centre of Bangladesh. So, the average Global Horizontal Irradiance is 625 W/m2 [4].

*Irradiance = (Average Insolation)/Average daily bright sunshine hours),*

For our system we have to consider Solar insolation = 4.5 kWh/m$^2$

Electric Vehicles also known as Easy Bikes [5] in Bangladesh are rechargeable battery based energy efficient and environment friendly popular vehicles. As it is rechargeable battery based where a lot of power is needed every day. As a result the power crisis is increasing. But there is a lot of fuel recharging stations in Bangladesh where there is a great opportunity to implement solar based recharging station. Previously very few relevant works have been done [6] they were not too much practical for our country, whereas our works is more practical and implementable in Bangladesh.

For designing and cost analysis data are collected from local and international markets and the electricity production's data is calculated manually with respect to Dhaka city. The rest of the paper organized as follows. In section II, the design of a PV system has been explained including location planning and accommodation. In section III, the overall system cost analyses has been done including battery bank calculation. In section IV, monthly income has been calculated. In section V, pay back analyses has been done. In section VI, opportunities and benefits have been discussed and finally the conclusion has been drawn in section VII.

## II. PV System Design

### A. Location

Dhaka is the capital of Bangladesh. Dhaka's geographical coordinates are 23° 43' 23" North, 90° 24' 31" East [7].

### B. Module Accommodation on Roof Area

The PV modules are accommodated on the roof of the recharging stations. Generally the roof is plain as a result there is no problem to set up but for maximum efficiency the PV modules are tilted with an angle which depends on the location of the installation. For this system let us consider the length of the roof is around 50 feet and the breadth is around 28 feet. So, Total area of the roof will be 1400 $ft^2$. If we considered the main building of (30x20 sq.ft) 600 $ft^2$ then the shade will be 800 $ft^2$. The layout is given in Figure 3. The number of components is calculated with the help of this consideration. So, considering this area (1400 $ft^2$) of roof we can accommodate maximum 92 solar modules of 245 watt each easily.

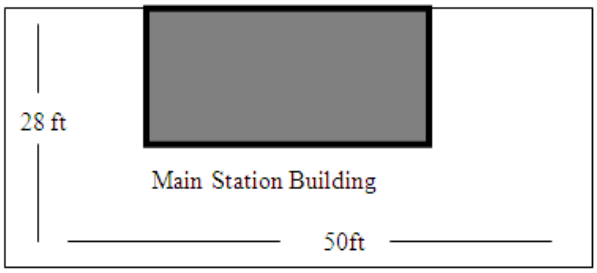



Figure 3. The considered layout

### C. System Sesign

As a solar based system the system can be designed with some PV modules, Charge Controller boxes, control circuit etc in the system and the representation of block diagram is given in Fig.4.

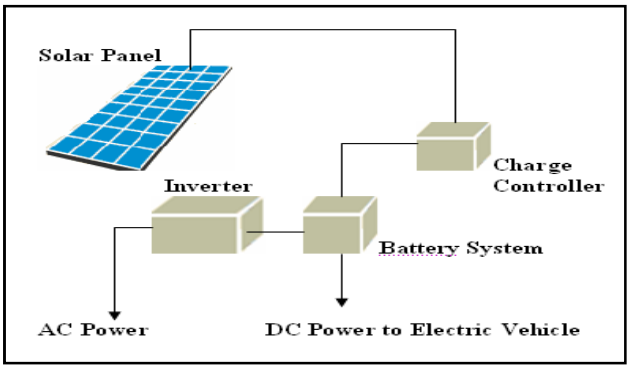



Figure 4. Total block diagram for the System

## III. Cost Analysis

### D. Load calculation and solar energy

The total load calculation for system is given in the below table.

TABLE I. LOAD CALCULATION OF THE STATION

| Load | Watt/unit | Number of Units | Total(watt) | Hours/Day | KWh/day |
|---|---|---|---|---|---|
| Easy Bike | 400 | 15 | 6000 | 8 | 48 |
| Bulb (CFL) | 7 | 4 | 28 | 12 | 0.34 |
| Fan | 40 | 2 | 80 | 8 | 0.64 |
| Net book | 15 | 1 | 15 | 5 | 0.07 |
| **Total:** | | | **9123** | | **48** |

For Bangladesh we have considered Solar Insolation = 4.5 kwh/$m^2$ /day. We have considered system efficiency 55%. Now, Energy to load (kwh/day) = $kw_p$ * Insolation * System Efficiency

$$kw_p = 48/(4.5*0.55) = 20k = 20{,}000\ w_p$$

Solar panels are rated in watt peak ($W_p$) i.e. the amount of power the panel can produce in a standard condition. So, total number of solar module required=20000/245 ≈ 82 module of 245 watt each although the system roof area able to contain 92 modules. For our system if we consider Sharp-ND-Q245 [8][9] Solar panel (each of 15 $ft^2$ surface), then total roof area needed for the 82 module is, 82*15= 1230 sq.ft.

### E. Battery Bank calculation

We are considering the days of autonomy or backup days i.e number of days when sunlight is not present is 2. Also, we are considering depth of discharge (DOD) of battery 70% as to preserve the battery the DOD is better to keep no more than 30%-50%. The life span of battery varies from 1-5 years. This life span can be prolonged by limiting the DOD of the battery i.e the amount of power drawn by the battery each day.

Backup days*Energy to Load = Purchase capacity of Battery * DOD

Purchase capacity of Battery=2*48000/0.7 =137143 watt-hour of battery

So, Amp-hour capacity* terminal voltage= 137143 watt-hour

Suppose, the voltage rating of the battery for the battery bank = 24 Volt and capacity of each battery =120 Amp–hour

Amp-hour capacity=137143/24=5714 Amp-hour

Normal PV standard battery rating = 120 Amp-hour

So, total battery required for the system= 5714/120 ≈ 47

In the same way it can be showed that, if we consider days of autonomy 1 day then we will need 47 battery of 12 volt each.

The different essential components with respective quantity and costs are given in Table-2.

TABLE II. COST COMPONENTS OF OUR PROPOSED SYSTEM

| Component | Description | Qty | Cost in Taka[#] |
|---|---|---|---|
| PV module | Sharp-ND-Q245 [8] | 82(Calculated) | 82*20,000 =16,40,000 |
| Inverter | 600Watt Inverter [10] | 7(Calculated) | 25,000*7= 1,75,000 |
| Charge controller | Local market | 12 | 1,700*12=20,400 |
| Wiring | Approximate | | 5,000 |
| Battery | 12Volt | 47 | 12,000*47 =5,64,000 |
| Control circuit, installation, maintenance and others | Approximate | | 20,000 |
| Total | | | 24,24,400 |

## IV. MONTHLY INCOME

In our proposed system the estimated generation capacity is around 20 KW. We can find the daily energy generation by multiplying the 20KW to the average sunshine hours/day of a month. Again for finding the monthly energy generation we have to multiply daily energy generation to number of days of each month respectively. From monthly energy generation we can calculate the monthly income and the summation of monthly income is referred to annual income which helps us to find payback time of our proposed system. In figure the monthly incomes are shown which is calculated in different consideration and the considerations are per kilowatt hour equal to 6, 8 and 10 respectively.

TABLE III. MONTHLY AND YEARLY INCOME OF OUR PROPOSED SYSTEM

| Months | Sunshine (hrs./day) | KWH/ Month | Income in Taka (1kwh=6tk.) | Income in Taka (1kwh=8tk.) | Income in Taka (1kwh=10tk) |
|---|---|---|---|---|---|
| January | 8.7 | 5220 | 31320 | 41760 | 52200 |
| February | 9.1 | 5460 | 32760 | 43680 | 54600 |
| March | 8.8 | 5280 | 31680 | 42240 | 52800 |
| April | 8.9 | 5340 | 32040 | 42720 | 53400 |
| May | 8.2 | 4920 | 29520 | 39360 | 49200 |
| June | 4.9 | 2940 | 17640 | 23520 | 29400 |
| July | 5.1 | 3060 | 18360 | 24480 | 30600 |
| August | 5.8 | 3480 | 20880 | 27840 | 34800 |
| September | 6 | 3600 | 21600 | 28800 | 36000 |
| October | 7.6 | 4560 | 27360 | 36480 | 45600 |
| November | 8.6 | 5160 | 30960 | 41280 | 51600 |
| December | 8.9 | 5340 | 32040 | 42720 | 53400 |
| Total Income in a year | | **54360** | **326160** | **434880** | **543600** |

## V. PAYBACK TIME ANALYSIS

The payback time means that the number of years required for the improvement to pay for itself and for cost benefit analysis of our proposed system it is needed.

*Simple payback time= (Cost of the system)/ (Annual savings), years*

**Considering 1kWh= 6 taka:**

Cost of the system = 24,24,400 taka

Annual income =326160 taka

Payback time (year) = (Cost of the system / Annual income) = 24,24,400 / 326160 ≈8 year

**Considering 1kWh= 8 taka:**

Cost of the system = 24,24,400 taka

Annual income =434880 taka

Payback time (year) = (Cost of the system / Annual income) = 24,24,400 / 434880≈ 6 year

**Considering 1kWh= 10 taka:**

Cost of the system = 24,24,400 taka

Annual income =543600 taka

Payback time (year) = (Cost of the system / Annual income) = 24,24,400 / 543600≈5 year

## VI. OPPORTUNITIES, BENEFITS AND CHALLENGES

### *A. Opportunities and Benefits*

The opportunities and benefits are given below:

1. The existing fuel stations can be used. No other site is needed to be selected.

2. The payback time is estimated and lesser than other system.
3. No extra labour cost and maintenance cost is needed.
4. Extra power can be used for the recharging station itself
5. This system becomes one of the factors to increase the income of the owners of fuel filling stations.
6. Proper utilization of solar energy.
7. Dependence on fossil fuels such as Diesel, Petrol and Gas etc is decreased.
8. It is a pollution free process.

### *B. Challenges and Limitations*

The challenges and limitations are given below:

1. Initial cost is high for the owner of the fuel filling stations.
2. The design must be perfect.
3. Lack of knowledge about proper maintenance of Solar based technologies.
4. During cloudy weather with the insufficient solar radiation the system can support maximum one day according to the considered 'days of autonomy' in this proposed system for the estimation.
5. Calculation has been done here for recharging only 15 vehicles.
6. For fully recharging the battery pack can take 4 to 8 hours. Even a "quick charge" to 80% capacity can take 30 min [11].

### *C. Comparative analysis with relevant works*

Previously in [6] there were some flaws such as, no load calculation was seen, roof area was uneconomical, no DOD and back up days considered for battery bank calculation which were impractical. Here we have cleared these problems in this work. In case of cost analysis in [6], even if we consider 6 taka/kwh it would take 24 years to recover but in our proposal we have brought it to approximately 8 years as recovery time.

## VII. CONCLUSION

Although the Electric vehicle has some limitations, its environment friendly nature has made it popular to all over the world. So, it is clear that our proposed system could be an effective one due to its necessity in the power crisis context. Though our proposed system is designed for Dhaka city and costs are calculated with respect to Dhaka city, this system is implemented anywhere in Bangladesh. And it is also said that this proposed system is economically and geographically feasible to Bangladesh.

## REFERENCES


[1] International Energy Agency: Share of total primary energy supply in 2003. Technical report, 2003.

[2] Bangladesh Power Development Board website. [Online]. Available: http://www.bpdb.gov.bd . Accessed on: February, 2013.

[3] REEI.[Online].Available: http://www.reein.org/solar/resource/index.htm.

[4] Final report of Solar and Wind Energy Resource Assessment (SWERA), Renewable Energy Research Centre, Dhaka, Bangladesh, February 2007.

[5] Bangladesh's First Internet Newspaper. [Online]. Available: http://biz.bdnews24.com/details.php?id=194924&cid=4

[6] Sikder Sunbeam Islam and M JashimUddin, "A Solar System to reduce the Power Crisis in Bangladesh through Electric Vehicle Recharging Station" IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE), Volume 5, Issue 3, April, 2013, India.

[7] Available at: http://www.maplandia.com/bangladesh/dhaka-div/dhaka-zila/dhaka/ dated:

[8] *Sharp NU-U235F1 datasheet;* Sharp Electronics Corporation, CA. [Online]Available at: www.mrsolar.com/pdf/**sharp**/**Sharp**235U1F.pdf

[9] *Midnite Solar MNPV3 datasheet,;* Midnight Solar, USA. [Online]. Available at: www.**midnitesolar**.com/**pdf**s/frontBack03.**pdf**

[10] Solarpac company Ltd. Data has been collected from local markets.

[11] US Department of Energy. [Online]. Available: http://www.fueleconomy.gov/feg/evtech.shtml

[12] HAQUE N.M. Ziaul; *Electricity problem of Bangladesh;* Wheel Buisness Magazine, Vol. 2 , Issue 1, January-March 2012.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Comparison of DTC and FOC for FSTP Inverter Fed IPMSM Drives

Tanvir Ahmed*, Anupam Das, and Kalyan Kumar Halder
Department of Electrical and Electronic Engineering
Khulna University of Engineering and Technology
Khulna-9203, Bangladesh
*Email: tanvir2362@gmail.com

***Abstract*—This paper presents a comparative study between field oriented control (FOC) and direct torque control (DTC), two most popular control strategies for inverter fed interior permanent magnet synchronous motor (IPMSM) drives. The comparison is done in four switch three phase (FSTP) inverter scheme instead of six switch three phase (SSTP) inverter scheme. The FSTP inverter scheme is better than SSTP inverter scheme because of the reduction in price, switching losses and the complexity of the control board. The comparison is based on various conditions such as normal operating condition, sudden change in load torque, speed reversal and change in stator resistance. To validate the effectiveness of the drive systems, simulation is carried out in MATLAB environment. The simulation results and comparative study have been found quite satisfactory.**

***Keywords*—Direct torque control, field oriented control, four switch three phase inverter, interior permanent magnet synchronous motor, and robustness.***

## I. Introduction

Permanent magnet synchronous motors (PMSMs) are widely used because of their advantages such as rugged construction, easy maintenance, high efficiency, high power factor, etc. Over the last few decades, several control systems have been proposed for the PMSMs [1]–[5]. These PMSMs are usually fed by SSTP inverters. However, they are relatively expensive because they need six switches and corresponding gate drive circuits. Hence, the schemes using FSTP inverters have been widely researched for cost reduction by the elimination of power switches and additional circuits such as gate drives [6]–[10].

In recent years, the inverter fed ac motor has largely taken over as the preferred solution for variable speed applications. Considering high-performance motion control, field oriented control (FOC), or more recently direct torque control (DTC) are used. The vector control philosophy started to be developed around 1970. Several types of vector control are possible: rotor-oriented, rotor-flux-oriented, stator-flux-oriented and magnetizing-flux-oriented. On the other hand, the publication of the DTC theory actually goes back to 1971. Recently, it is being considered as an industrial alternative to the FOC strategy. These control strategies are different on the operation principle but their objectives are the same.

Now, which one gives better performance for FSTP inverter fed IPMSM? This question has not been answered yet. Therefore, the objective of this paper is to present a comparative study in order to reach a verdict on the superiority between these two control strategies.

## II. Mathematical Model of IPMSM

A mathematical model of the IPMSM is required for proper simulation of the system. The dynamic model of the IPMSM motor in the synchronously rotating $d-q$ reference frame can be expressed as follows [11]:

$$v_d = Ri_d + pL_d i_d - \omega_r L_q i_q \tag{1}$$

$$v_q = Ri_q + pL_q i_q + \omega_r \psi_f + \omega_r L_d i_d \tag{2}$$

The developed electromagnetic torque is given as:

$$T_e = \frac{3P_p}{2}\left(\psi_f i_q + (L_d - L_q)\, i_d i_q\right) \tag{3}$$

The mechanical motion of the IPMSM can be expressed as:

$$T_e = T_l + J_m p\omega_m + B_m \omega_m \tag{4}$$

$$\omega_r = P_p \omega_m \tag{5}$$

Where,

$v_d$ and $v_q$ = the dq- axes stator voltages;

$i_d$ and $i_q$ = the dq- axes stator currents;

$L_d$ and $L_q$ = the dq- axes inductances;

$\psi_f$ = the permanent magnetic flux linkage;

$R$ = the stator resistance;

$\omega_r$ = the angular speed of rotor;

$\omega_m$ = the mechanical speed of rotor;

$T_e$ = the electromagnetic torque;

$J_m$ = the motor inertia;

$B_m$ = the motor friction coefficient;

$P_p$ = the number of pole pairs;

$p = \frac{d}{dt}$.

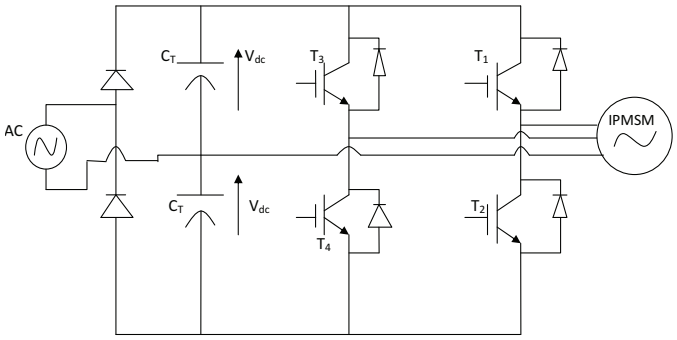



Fig. 1. FSTP inverter fed IPMSM drive.

## III. FOUR SWITCH THREE PHASE INVERTER MODEL

In the four switch inverter, as shown in Fig. 1, a three phase system is obtained by connecting the phase $c$ terminal of the stator windings directly to the centre tap of the dc link capacitors. The single phase ac supply is rectified by the front-end rectifier. The capacitors are used to level the output dc voltage. The three phase voltages to the IPMSM can be expressed as follows [12]:

$$V_a = \frac{V_{dc}}{3}\left[4S_a - 2S_b - 1\right] \tag{6}$$

$$V_b = \frac{V_{dc}}{3}\left[4S_b - 2S_a - 1\right] \tag{7}$$

$$V_c = \frac{V_{dc}}{3}\left[-S_a - S_b + 1\right] \tag{8}$$

where $V_{dc}$ is the maximum voltage across the dc link capacitors, $S_a$ and $S_b$ are the switching states (0 or 1) of upper switches in the legs of phases $a$ and $b$ respectively.

## IV. DIRECT TORQUE CONTROL SYSTEM

The complete block diagram of DTC based IPMSM drive system is shown in Fig. 2. The basic idea of DTC for IPMSM is to control the torque and flux linkage by selecting the voltage space vectors properly, which is based on the relationship between the lip frequency and torque. The voltage vector plane of a four switch inverter fed system is divided into four sectors as shown in Fig. 3. A voltage vector switching table for the four switch inverter fed DTC system is tabulated in Table I.

TABLE I. VOLTAGE VECTOR TABLE

| $\phi$ | $\tau$ | $S_1$ | $S_2$ | $S_3$ | $S_4$ |
|---|---|---|---|---|---|
| 0 | 0 | $V_4$ | $V_1$ | $V_2$ | $V_3$ |
| 0 | 1 | $V_3$ | $V_4$ | $V_1$ | $V_2$ |
| 1 | 0 | $V_1$ | $V_2$ | $V_3$ | $V_4$ |
| 1 | 1 | $V_2$ | $V_3$ | $V_4$ | $V_1$ |

The basic principle of the DTC is to select proper voltage using a pre-defined switching table. The selection is based on the hysteresis control of the stator flux linkage and the torque. In the basic form the stator flux linkage is estimated with:

$$\phi_s(t) = \int_0^t (V_s - R_s I_s)\, dt + \phi_{so} \tag{9}$$

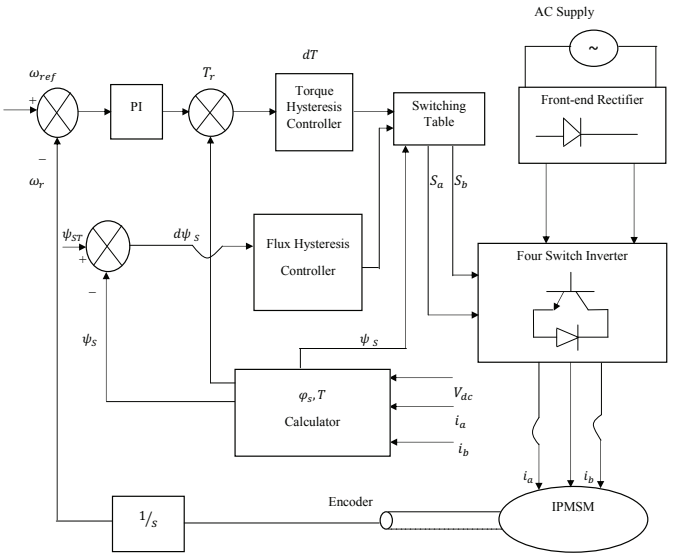



Fig. 2. Block diagram of the DTC scheme of IPMSM.

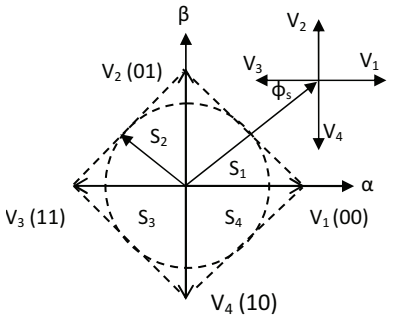



Fig. 3. Voltage vectors of a FSTP inverter fed system.

where $\phi_{so}$ is the initial value of the stator flux linkage. The composite $\alpha$ and $\beta$ components of vector $\phi_s$ can be obtained:

$$\phi_{s\alpha}(t) = \int_0^t (V_{s\alpha} - R_s I_{s\alpha})\, dt \tag{10}$$

$$\phi_{s\beta}(t) = \int_0^t (V_{s\beta} - R_s I_{s\beta})\, dt \tag{11}$$

Stator flux linkage can be written as:

$$\phi_s = \sqrt{(\phi_{s\alpha})^2 + (\phi_{s\beta})^2} \tag{12}$$

The angle is equal to:

$$\theta_s = \tan^{-1}\left(\frac{\phi_{s\beta}}{\phi_{s\alpha}}\right) \tag{13}$$

The torque can then be estimated with:

$$T_e = \frac{3}{2} P_p \left(\phi_{s\alpha} I_{s\beta} - I_{s\alpha}\phi_{s\beta}\right) \tag{14}$$

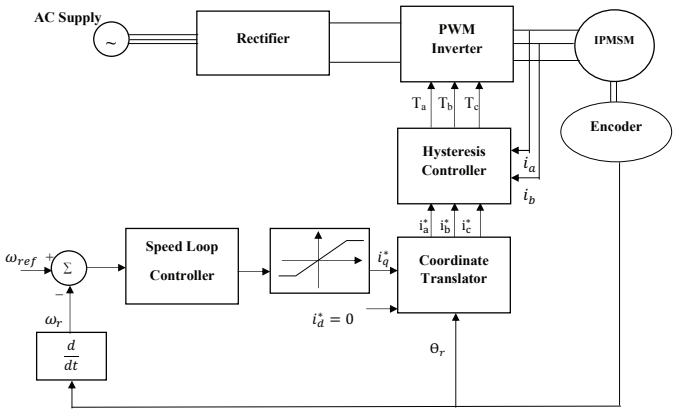



Fig. 4. Block diagram of the FOC scheme of IPMSM.

## V. FIELD ORIENTED CONTROL SYSTEM

The machine model of PMSM can be described in the rotor rotating reference frame as follows:

$$v_d = L_d \frac{d}{dt} i_d + R_s i_d - \omega_s L_q i_q \tag{15}$$

$$v_d = L_q \frac{d}{dt} i_q + R_s i_q - \omega_s L_d i_d \tag{16}$$

and

$$\lambda_d = \psi_f + L_d I_d \tag{17}$$

$$\lambda_d = L_q I_q \tag{18}$$

where $\lambda_d$ and $\lambda_q$ are the stator flux linkages.

The corresponding electromagnetic torque is:

$$T_e = \frac{3}{2} P_p \left[\psi_f i_q + (L_d - L_q)\, i_d i_q\right] \tag{19}$$

The associated electromechanical equations are as follows:

$$J_m \frac{d\omega_m}{dt} + B_m \omega_m = T_e - T_l \tag{20}$$

$$\frac{d\theta_m}{dt} = \omega_m \tag{21}$$

where $\theta_m$ is the rotor angular displacement.

The primary principle in controlling an IPMSM drive is based on field orientation. Fig. 4 shows the complete block diagram of FOC based IPMSM drive. Since the magnetic flux generated from the PM rotor is fixed in relation to the rotor shaft position, the flux position in the coordinates can be determines by the shaft position sensor. If $i_d = 0$, the $d$-axis flux linkage $\phi_d$ is fixed. Since $\phi_f$ is constant for an IPMSM, the electromagnetic torque is then proportional to $i_q$, which is determined by the closed loop control.

$$T_e = \frac{3}{2} P \phi_f i_q \tag{22}$$

The rotor flux is produced only in the $q$-axis while the current vector is generated in the axis in FOC. Since the generated motor torque is linearly proportional to the $q$-axis current, as the $d$-axis rotor flux is constant, the maximum torque per ampere can be achieved.

## VI. SIMULATION RESULTS

The performance comparison between DTC and FOC is made in terms of speed, torque and current response. The static and dynamic performances of the DTC and FOC schemes are obtained by simulation performed in MATLAB.

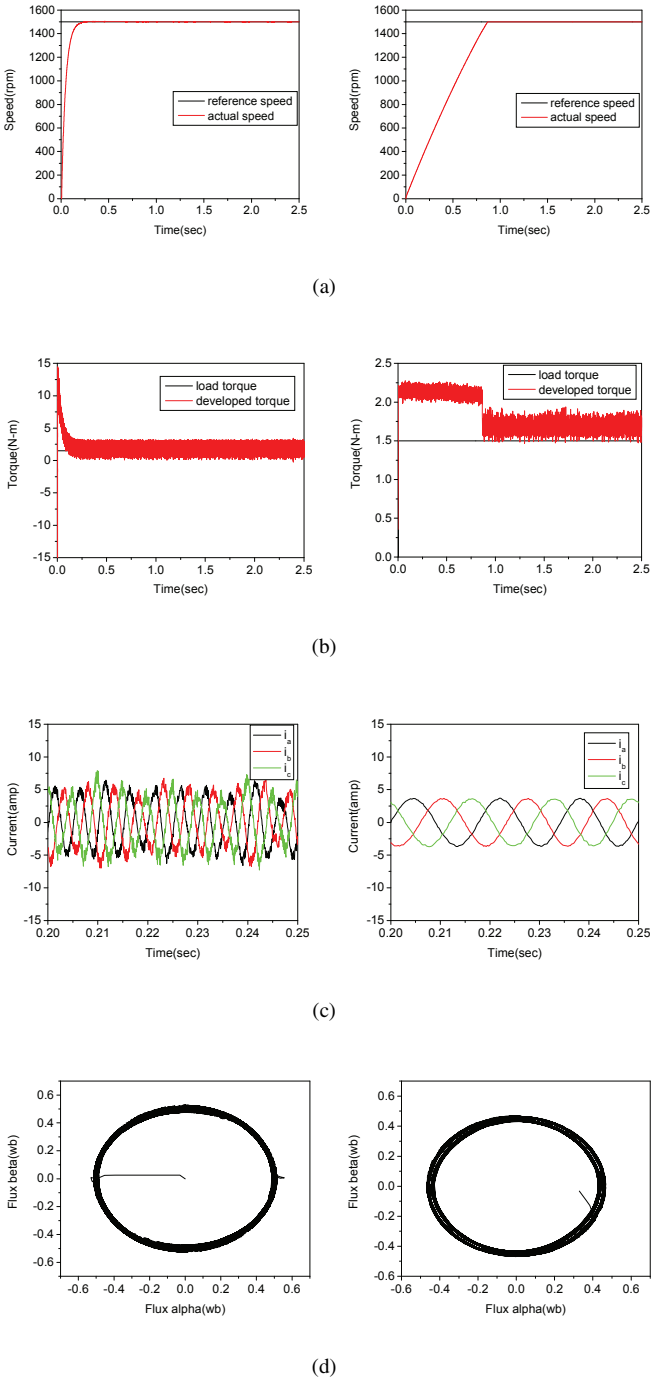



Fig. 5. Comparison of DTC and FOC in terms of (a) speed response (b) torque response, (c) three phase currents and (d) flux locus under normal operating condition. Left side figures represent the curve for DTC and right side figures represent curve for FOC.

### A. Normal Operating Condition

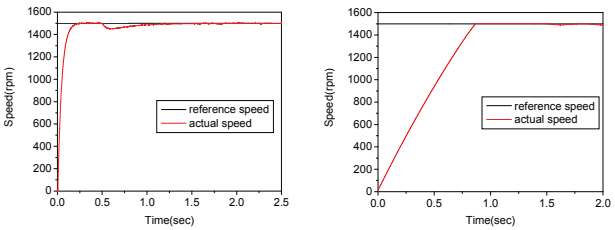
The motor starting performance at the rated speed of 1500 rpm is shown Fig. 5(a). Speed response for DTC is faster than FOC as it reaches the reference speed at 0.25 sec whereas that of FOC reaches at 0.8 sec. Fig. 5(b) shows the developed torque that oscillates around the load torque (=1.5 N-m) when the set speed is reached. It is noticed that higher electromagnetic torque is generated during the motor acceleration. From the figure it is clear that Torque response of DTC is better than that of FOC. Fig. 5(c) shows the actual 3-phase currents for both control schemes. Current response of DTC is quite acceptable for healthy motor operation though FOC gives better response than DTC. Fig. 5(d) shows the trajectory of the stator flux linkage. They are quite same for both DTC and FOC and as far our expectation.

### B. Load Change Condition

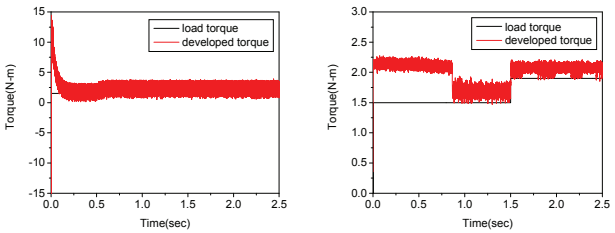
Fig. 6(a) and (b) show the speed and torque responses under load change condition, respectively. We investigated that when load increases speed falls down from reference speed and then automatically reaches the rated speed. We used 1.5 N-m as reference torque and changed it to 2 N-m for simulating load change condition. For DTC we changed the load at 0.5 sec and it reached the reference speed at 1.2 sec. On the other hand, for FOC we changed the torque at 1.5 sec and it reached the reference speed very quickly.

### C. Speed Reversal Condition

Fig. 7(a) and (b) show speed curve and the torque response under speed reversal condition respectively. By analyzing these figures we can see that DTC shows slightly better speed and torque response than FOC under speed reversal condition.

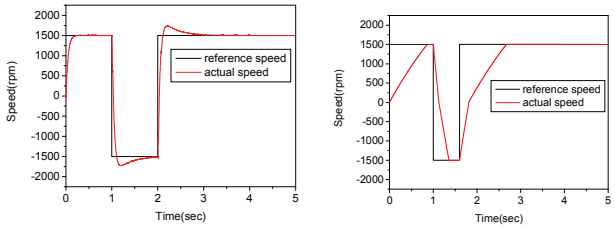



(a)

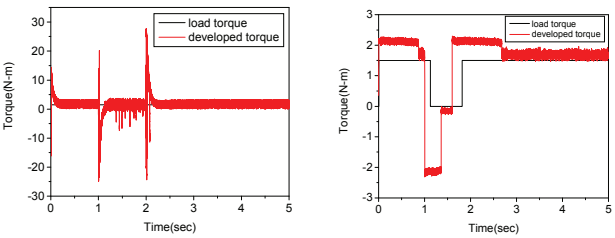



(b)

Fig. 6. Comparison of DTC and FOC at the variation of load torque in terms of (a) speed and (b) torque.




(a)

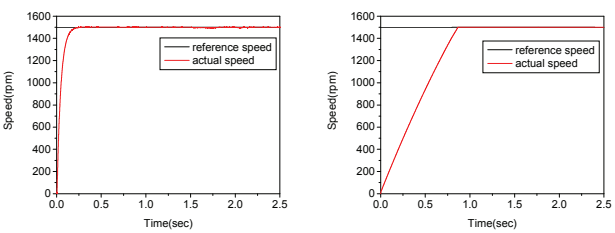



(b)

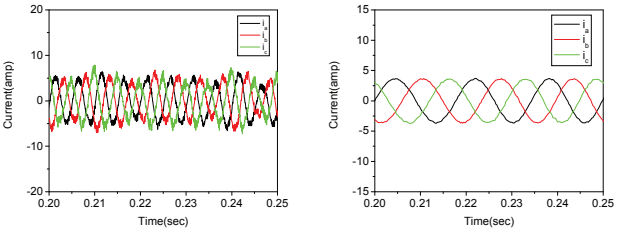
Fig. 7. Comparison of DTC and FOC at speed reversal condition in terms of (a) speed and (b) torque.

### D. Effect of Change of Stator Resistance

Fig. 8(a) and (b) show the speed response curve and three phase current response curve when the stator resistance is increasing, respectively. From the simulation figure we can see that there is no effect on speed or current response for an increase in stator resistance.




(a)




(b)

Fig. 8. Comparison of DTC and FOC for step change of stator resistance in terms of (a) speed and (b) current.

## VII. Conclusion

In this paper, a performance analysis of FOC and DTC schemes for FSTP inverter fed IPMSM drives has been investigated by simulation with a view to distinguishing their respective advantages and disadvantages. A detailed model for IPMSM drive system with DTC and FOC has been developed and operations have been studied using two current control schemes. A speed controller has been designed successfully for closed loop operation of the IPMSM drives system so that the motor runs at the reference speed. Robustness is present in both FOC and DTC schemes. Parameter sensitivity is large in FOC scheme, on the other hand it is average in DTC scheme. DTC reduces the speed response time, whereas FOC gives better current response. On different conditions, both the schemes showed very similar simulation output. So, we can conclude that it is hard to comprehend the superiority between DTC and FOC.

## VIII. APPENDIX

Motor parameters used in this simulation are given below:

| Parameter Name | Parameter Value |
|---|---|
| Number of pole pairs, $P_p$ | 2 |
| Stator Resistance, $R_s$ | 1.93 ohm |
| Permanent Magnet Flux Linkage, $\psi_f$ | 0.311 |
| Flux Density, $B_m$ | 0.001 Wb/m$^2$ |
| Moment of Inertia, $J_m$ | 0.003 |
| Rated Voltage, $V_{dc}$ | 300V |
| $d$-axis inductance, $L_d$ | 0.0244H |
| $q$-axis inductance, $L_q$ | 0.07957H |
| Rated speed, $\omega_m$ | 1500 rpm |

## References


[1] B. K. Bose, "A high-performance inverter-fed drive system of an interior permanent magnet synchronous machine," *IEEE Trans. on Industry Applications*, vol. 24, no. 6, pp. 987–997, Nov./Dec. 1988.

[2] Y. Utsumi, N. Hoshi, and K. Oguchi, "Comparison of FPGA-based direct torque controllers for permanent magnet synchronous motors," *Journal of Power Electronics*, vol. 6, no. 2, pp. 114–120, Apr. 2006.

[3] M. N. Uddin and M. A. Rahman, "Fuzzy logic based speed control of an IPM synchronous motor drive," *Journal of Advanced Computational Intelligence*, vol. 4, no. 2, pp. 212–219, Dec. 2000.

[4] K. K. Halder, N. K. Roy, and B. C. Ghosh, "Position sensorless control for an interior permanent magnet synchronous motor SVM drive with ANN based stator flux estimator," *International Journal of Computer and Electrical Engineering*, vol. 2, no. 3, pp. 475–480, Jun. 2010.

[5] F. F. M. El-Sonsy, "Robust tracking control based on intelligent sliding-mode model-following position controllers for PMSM servo drives," *Journal of Power Electronics*, vol. 7, no. 2, pp. 159–173, Apr. 2007.

[6] C. B. Jacobina, E. R. C. da Silva, A. M. N. Lima, and R. L. A. Ribeiro, "Vector and scalar control of a four switch three phase inverter," in *Industry Applications Conference, 1995. Thirtieth IAS Annual Meeting, IAS '95., Conference Record of the 1995 IEEE*, 1995, vol. 3, pp. 2422–2429.

[7] F. Blaabjerg, D. O. Neacsu, and J. K. Pedersen, "Adaptive SVM to compensate DC-link voltage ripple for four-switch, three-phase voltage source inverter," *IEEE Trans. Power Electronics*, vol. 14, no. 4, pp. 743–751, Jul. 1999.

[8] M. N. Uddin, T. S. Radwan, and M. A. Rahman, "Performance analysis of a 4-switch, 3-phase inverter based cost effective IPM motor drives," in *Canadian Conference on Electrical and Computer Engineering*, 2004.

[9] M. B. R. Correa, C. B. Jacobina, E. R. C. da Silva, and A. M. N. Lima, "A general PWM strategy for four-switch three-phase inverters," *IEEE Trans. Power Electronics*, vol. 21, no. 6, pp. 1618–1627, Nov. 2006.

[10] J. Klima, "Analytical investigation of an induction motor fed from four-switch VSI with a new space vector modulation strategy," *IEEE Trans. Energy Conversion*, vol. 21, no. 4, pp. 832–838, Dec. 2006.

[11] D. Y. Ohm, J.W. Brown, and V. B. Chava, "Modeling and parameter characterization of permanent magnet synchronous motors," in *Annual Symposium of Incremental Motion Control Systems and Devices*, 1995.

[12] K. K. Halder, M. J. Islam, M. A. Rafiq, and B. C. Ghosh, "Vector control of a cost effective FSTP inverter fed synchronous reluctance motor drive based on recurrent neural network," *Pacific Journal of Science and Technology*, vol. 12, no. 2, pp. 20–26, Nov. 2011.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Efficiency Improvement of Semi-Bridgeless Phase-Shifted Boost Converter with Power Factor Correction in Energy Storage System

Parvez Akter[1*], Muslem Uddin[1], Md. Mizanur Rahman[2], Monirul Islam[1] and Md. Rezaul Basher Bhuiyen[2]
[1]Department of Electrical Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
[2]Department of Electrical and Electronic Engineering, CUET, Chittagong, Bangladesh
[*] E-mail: suzon_cuet06@yahoo.com

*Abstract*—**Boost converter is an essential part of energy storage system for efficient charging of the static devices. This paper proposes a 7.5 kW semi-bridgeless phase-shifted boost converter to improve the efficiency and performance of the energy storage system. This investigation is carried out with the MATLAB Simulink to validate the feasibility of the proposed converter topology. The verification study depicted in this inquisition with the unity power factor and improved efficiency of the converter. The maximum efficiency of the proposed converter is determined as 98.3% for 7.52 kW power transfer at 200 V, 50 Hz input with 70 kHz switching frequency.**

***Keywords—Semi-Bridgeless Boost Converter, Power Factor Correction, Phase-Shift Modulation, Energy Storage System.***

## I. INTRODUCTION

Energy storage system is elaborately used in utility and transport applications as well as in the renewable energy research to ensure the power reliability. This energy storage system is used in further utility applications such as; active power control, load leveling and frequency control which have been investigated in [1-3]. Generally, in the energy storage system the used static storage devices are electric double layer capacitor, Li-ion battery, lead-acid battery and nickel metal-hydride battery [4]. These Static devices are also used immensely in pure, hybrid, plug-in hybrid and fuel-cell electric vehicles [5-7], elevators [8] and gantry crane systems [9] etc.

Therefore, several circuit topologies and control techniques of converters have been used in battery energy storage system. The two-stage cascaded AC-DC and DC-DC converter is commonly used in relatively high rated Li-ion storage system [10]. On the other hand, because of the low frequency output ripple current, the single-stage approach of boost converter is more suitable for lead-acid batteries. However, this boost converter faces some limitations with high switching losses in semiconductor switches and also high conduction losses in the diode bridge. Therefore, for reducing the conduction losses in the diode bridge, a bridgeless converter topology have been proposed in [11] and [12]. This topology eliminates the conduction losses in the diode bridge and also reduces the total amount of semiconductor devices from six to four, but it still faces some problems with the increase of EMI [13, 14]. This problem can be eliminated by using some powerful control algorithms, such as; fuzzy logic, neural network and predictive control etc. Some recent investigations are carried out with predictive control algorithm applied in the power converters for different inquisitions such as, current, torque, flux and unity power factor control [15, 16]. These investigations are validated the effectiveness of the control.

In this paper, a semi-bridgeless phase-shifted boost converter with power factor correction of the battery energy storage system has been proposed in order to improve the efficiency and performance of the same.

This paper is organized in the following: Section II presents the semi-bridgeless boost converter topology, Section III describes the working principle of the converter and Section IV depicts the results and analysis. The power factor correction is analyzed and the efficiency improvement is compared with interleaved PFC converter in Section V. Finally a fruitful conclusion is drawn in VI.

## II. SEMI-BRIDGELESS BOOST CONVERTER TOPOLOGY

Fig. 1 shows the semi-bridgeless boost converter topology which is controlled with phase-shift pulse width modulation method. This converter topology features improve efficiency and performance at low power rating compared with the conventional and bridgeless boost converters. Therefore, the charger size, component count, charger cost and EMI are reduced. Again, this converter also provides power factor correction capability for improving the power quality as well as the battery health of the energy storage system.

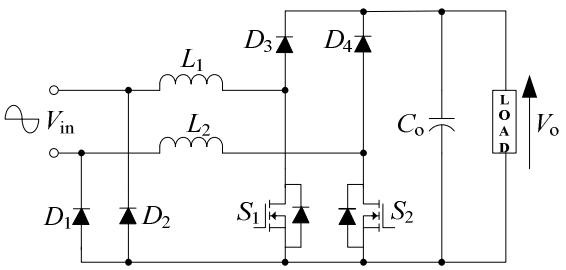



Figure 1. Semi-bridgeless boost converter

Two extra slow diodes ($D_1$ and $D_2$) have been configured in the proposed semi-bridgeless boost converter, compared with bridgeless converter topology for linking the ground with the input line. However, the conduction losses associated with the slow diodes remain at low level as the current does not always return through these diodes.

## III. PRINCIPLE OF OPERATION

The operating principle of semi-bridgeless boost converter is analyzed by separating the input AC power supply into the positive and negative half-cycles. A detailed circuit operation depends on the duty cycle is elaborately presented in the following sections.

### *A. Positive half-cycle*

Fig. 1 shows that, MOSFET switch $S_1$ turns ON during the positive half-cycle of the AC input voltage. Thus, current flows through the inductor $L_1$, switch $S_1$ and further continues through the switch $S_2$ inductor $L_2$, to return to supply when energy is stored in the inductors $L_1$ and $L_2$. When switch $S_1$ is OFF during this positive half-cycle, the stored energy of inductors ($L_1$ and $L_2$) is discharged with current flow through diode $D_3$, load and then further returns partially through slow diode $D_1$ and internal diode of MOSFET switch $S_2$.

The circuit operation of semi-bridgeless boost converter is widely depends on the duty cycle ($D$). In General, duty cycle of the converter is greater than 50%, while supply voltage is smaller than half with respect to output voltage during any half-cycle (positive or negative). On the other hand, in case of greater input voltage compared with half of output voltage, duty cycle is less than 50%. Detailed circuit operation of semi-bridgeless boost converter for duty cycle D > 50% and D < 50% during positive half-cycle of supply voltage are analyzed and presented in the following sections.

#### *1) Duty cycle D > 50%:*

Fig. 2 shows the operating principle of semi-bridgeless boost converter which can be described with three modes of operation for duty cycle greater than 50%.

*a) Mode I (time interval $t_0$ to $t_1$ and $t_2$ to $t_3$):*

Figs. 2(a) and 2(b) shows that, the MOSFET switches $S_1$ and $S_2$ are ON at mode I. At this time period, current increases linearly in the series inductances $L_1$ and $L_2$. Hence energy is stored into these passive elements. Further, stored energy in capacitor ($C_o$) is delivered to the load. Current ripple in $S_1$ and $S_2$ are similar to the current in $L_1$ and $L_2$. Therefore, the ripple current is:

$$\Delta i = \frac{1}{L_1 + L_2} v_{in}(D - 0.5)T_s \tag{1}$$

*b) Mode II (time interval $t_1$ to $t_2$):*

MOSFET switch $S_1$ remains ON and $S_2$ turns OFF in the mode II, as shown in Fig. 2(a) and 2(c). Hence, current increases linearly and stores energy in $L_1$ and $L_2$. During this time interval ($t_1$ to $t_2$), capacitor ($C_o$) provides the stored energy to load. Threfore, the current ripples in $S_1$ and internal diode of $S_2$ are as follows:

$$\Delta i = \frac{1}{L_1 + L_2} v_{in}(1 - D)T_s \tag{2}$$

*c) Mode III (time interval $t_3$ to $t_4$):*

At mode III of the semibridgeless boost converter, MOSFET switch $S_1$ is OFF and $S_2$ is ON, as shown in Fig. 2(a) and 2(d). During the time interval ($t_3$ to $t_4$) of mode III, the stored energy in series inductances $L_1$ and $L_2$ is discharged through $L_1$, $D_3$, $S_2$ and $L_2$ to load. Threfore, the current ripples in $D_3$ and $S_2$ are similar as in $L_1$ and $L_2$.

$$\Delta i = \frac{1}{L_1 + L_2} (v_{in} - V_o)(1 - D)T_s \tag{3}$$

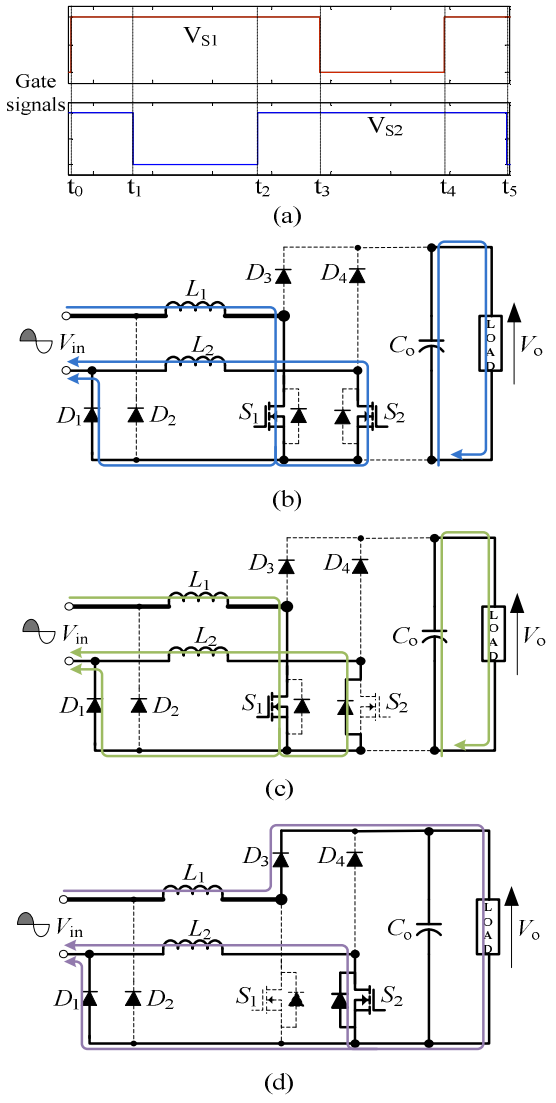



Figure 2. Operating principle of phase-shifted semi-bridgeless boost converter for duty cycle D > 50%. (a) gate signals $V_{S1}$ and $V_{S2}$ for two MOSFET switch $S_1$ and $S_2$ respectively, (b) switching state and current flow diagram for mode I, (c) mode II and (d) mode III.

*2) Duty cycle D < 50%:*

When duty cycle is D < 50%, the operating principle of semi-bridgeless boost converter can also be described with three mode of operation. A detailed explanation of operating principle is presented in Fig. 3 and elaborately discussed in the following sections.

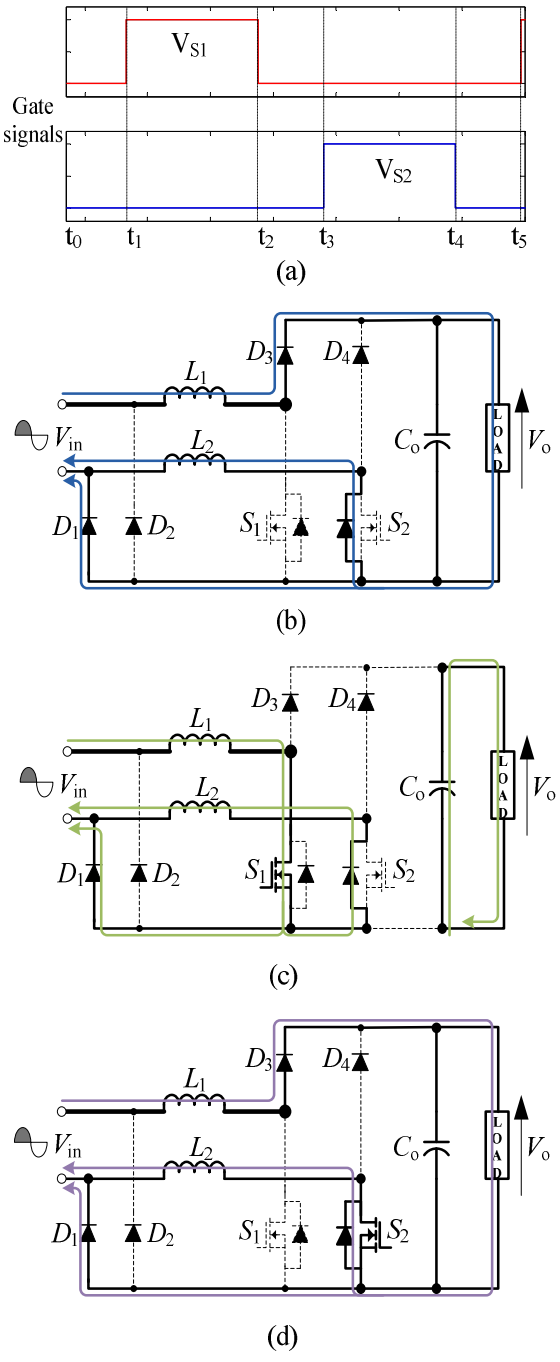



Figure 3. Operating principle of phase-shifted semi-bridgeless boost converter for duty cycle D < 0.5. (a) gate signals $V_{S1}$ and $V_{S2}$ for two MOSFET switch $S_1$ and $S_2$ respectively, (b) switching state and current flow diagram for mode I, (c) mode II and (d) mode III.

*a) Mode I (time innterval $t_0$ to $t_1$ and $t_2$ to $t_3$):*

At mode I of the semi-bridgeless boost converter for duty cycle D < 50%, switches $S_1$ and $S_2$ both are OFF, as mentioned in Figs. 3(a) and 3(b). The stored energy in $L_1$ and $L_2$ are released through the inductor $L_1$, diode $D_3$, internal diode of switch $S_2$ and inductance $L_2$. Therefore, current ripples in diode $D_3$ and internal diode of switch $S_2$ are in bellow:

$$\Delta i = \frac{1}{L_1 + L_2}(v_{in} - V_o)(0.5 - D)T_s \tag{4}$$

*b) Mode II (time interval $t_1$ to $t_2$):*

MOSFET switch $S_1$ turns ON and $S_2$ remains OFF in mode II, as presented in Figs. 3(a) and 3(c). In this mode, current ripples in $S_1$ and internal diode of $S_2$ are given by:

$$\Delta i = \frac{1}{L_1 + L_2} v_{in} D T_s \tag{5}$$

*c) Mode III (time interval t3 to t4):*

The MOSFET switch $S_1$ is OFF and $S_2$ is ON at mode III of the phase-shifted semibridgeless boost converter, as presented in Fig. 3(a) and 3(d). During the time interval ($t_3$ to $t_4$) of this mode III, energy stored in $L_1$ and $L_2$ are discharged through $L_1$, $D_3$, $S_2$ and $L_2$ to load. Threfore, current ripples in $D_3$ and $S_2$ are expressed as:

$$\Delta i = \frac{1}{L_1 + L_2}(v_{in} - V_o) D T_s \tag{6}$$

*B. Negetive Half-cycle*

During negative half-cycle of AC supply voltage, MOSFET switch $S_2$ turns ON, then the current flows through the inductor $L_1$ and switch $S_2$ which further continues through switch $S_1$ and returns through the inductor $L_1$ with storing in $L_2$ and $L_1$. After the turn OFF of the switch $S_2$, the stored energy in $L_2$ and $L_1$ is discharged with current flow through diode $D_4$ to the load. In this case, the current which returns to input is split between the internal diode of $S_1$ and slow diode $D_2$.

The operating principle of converter during the negative half-cycle of supply voltage is similar as the positive half-cycle for both the cases of duty cycle D > 50% and D < 50% but the direction is opposite.

## IV. Simulation Results

The proposed semi-bridgeless phase-shifted boost converter topology is simulated in MATLAB Simulink to validate the feasibility to improve the power factor as well as the efficiency. The proposed converter topology has been investigated for 7.5 kW power transfer. The system topology is analyzed with 240 V supply voltage, 70 kHz switching frequency and 8.0 kW load condition. The parameters used in simulation for the proposed inquisition are given in Table I.

Fig. 4 (a) shows the supply AC voltage and AC current which are exactly in phase with each other and maintain the sinusoidal wave shape. Therefore unity power factor is maintained in phase-shifted semi-bridgeless boost converter topology. Due to the pure sinusoidal wave form of the current, THD is in an acceptable limit. This is also one of the most important improvements of this investigation.

Fig. 4(b) shows the current wave shapes in $L_1$ and $L_2$. During positive period of supply voltage, the current in inductor $L_1$ is similar to supply AC current and during negative period it flows partially through the slow diodes $D_1$ and $D_2$. On the other hand, in case of positive period the current in inductor $L_2$ is partially flowing through the slow diodes $D_2$ and $D_1$ and remain same with supply current during negative interval.

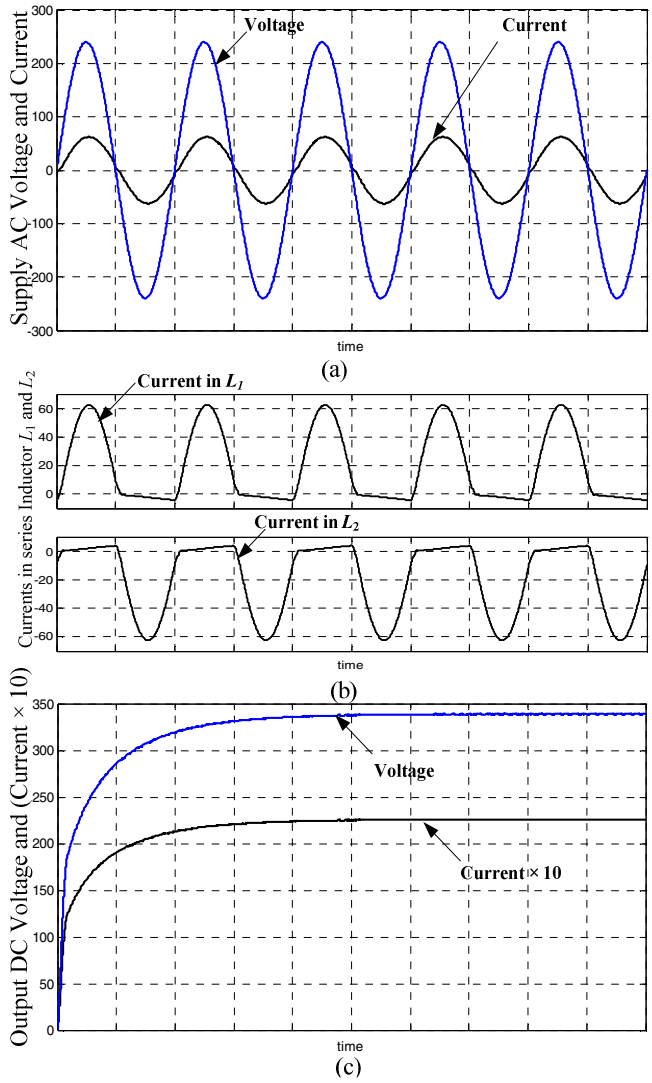



Figure 4. Supply AC voltage and AC current, output DC voltage and DC current for 7.5 kW power transfer of a semi-bridgeless boost converter. (a) Supply AC voltage and AC current. (b) Current wave shapes in upper and lower series inducotrs $L_1$ and $L_2$ respectively. (c) Output DC voltage and DC current.

TABLE I. SIMULATION PARAMETER

| SL. No | Parameters Values and Unit | | |
|---|---|---|---|
| | *Variables and parameters* | *Values* | *Unit* |
| 1. | Input Supply Voltage ($v_{in}$) | 240 | V |
| 2. | Supply frequency ($f_{in}$) | 50 | Hz |
| 3. | Upper Series Inductor ($L_1$) | 400 | µH |
| 4. | Lower Series Inductor ($L_2$) | 400 | µH |
| 5. | Capacitor Across Load ($C_o$) | 1500 | µF |
| 6. | Load Resistance ($R_o$) | 15 | Ω |
| 7. | MOSFET on State Resistance ($R_{ds(on)}$) | 0.5 | mΩ |
| 8. | Switching Frequency ($f_s$) | 70 | kHz |

The output DC voltage and DC current of the semi-bridgeless phase-shifted boost converter is presented in Fig. 4(c). The ripples of the DC output voltage is very low and maintains a smooth wave shape of results. Also, this slight ripple in output DC voltage is reciprocal to the value of capacitor connected with load which has a great impact on the power factor.

The efficiency and performance of the phase-shifted semi-bridgeless boost converter is measured with the input and output power from Figs. 4(a) and 4(c). From Fig. 4(a) the peak sinusoidal AC input voltage is 240 V and peak sinusoidal current is measured as 63.75 A. Hence, the input power is measured as 7.65 kW. On the other hand, in the Fig. 4(c), the output DC voltage is measured 340.20 V and DC current is measured 22.10 A. Hence, the output power of the converter is 7.52 kW. Therefore, the overall efficiency is measured as 98.3% for the proposed converter.

## V. POWER FACTOR CORRECTION AND EFFICIENCY IMPROVEMENT

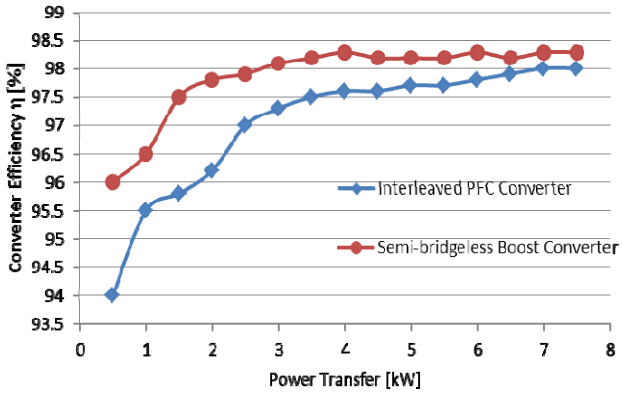



Figure 5. Measured efficiency of proposed semi-bridgeless boost converter and conventional interleaved PFC converter.

The power factor correction of phase-shifted semi-bridgeless boost converter topology can be confirmed from the Fig. 4(a). It shows that, the input voltage and current is pure sinusoidal in shape and also exactly in phase to each other. Hence, the power factor achieved by the converter is more than 0.99.

The efficiencies of proposed phase-shifted semi-bridgeless boost converter are measured between 0.5 kW to 7.5 kW power transfer range and compared with the same for interleaved PFC converter [17] shown in Fig. 5. The efficiencies associated with the proposed semi-bridgeless boost converter are higher compared to interleaved PFC converter. The peak efficiency of this semi-bridgeless boost converter is 98.3% at 7.5 kW power transfer condition.

## VI. CONCLUSIONS

A high efficient phase-shifted semi-bridgeless boost converter topology for the energy storage system is presented in this paper. This proposed converter topology exhibits high efficiency and performance for medium range power transfer with attractive features as small charger size, low cost, fast charging time and reduced EMI. The power factor of this proposed topology has been maintained unity during 50% to full load condition. So, the reactive power losses associated with this converter is very low. Therefore, the maximum efficiency of this proposed converter is achieved up to 98.3% at 240 V supply voltage with 50Hz supply frequency for 7.52 kW output power transfer. The results associated with the investigated converter topology are very much encouraging and will continue to play a strategic role in the improvement of modern high performance AC-DC power converters in energy storage system and will open a new era of power electronics research.

## REFERENCES


[1] S. Vazquez, S. M. Lukic, E. Galvan, L. G. Franquelo, and J. M. Carrasco, "Energy storage systems for transport and grid applications," *IEEE Trans. Ind. Electron.,* vol. 57, pp. 3881-3895, 2010.

[2] M. Bragard, N. Soltau, S. Thomas, and R. W. De Doncker, "The balance of renewable sources and user demands in grids: Power electronics for modular battery energy storage systems," *IEEE Trans. Power Electron.,* vol. 25, pp. 3049-3056, 2010.

[3] J. M. Carrasco, L. G. Franquelo, J. T. Bialasiewicz, E. Galván, R. P. Guisado, M. A. Prats*, et al.*, "Power-electronic systems for the grid integration of renewable energy sources: A survey," *IEEE Trans. Ind. Electron.,* vol. 53, pp. 1002-1016, 2006.

[4] N. M. Tan, T. Abe, and H. Akagi, "Topology and application of bidirectional isolated dc-dc converters," in *IEEE Int. Conf. Power Electron. and ECCE Asia (ICPE & ECCE), 2011*, 2011, pp. 1039-1046.

[5] C. Chan, "The state of the art of electric, hybrid, and fuel cell vehicles," *Proc. IEEE,* vol. 95, pp. 704-718, 2007.

[6] A. F. Burke, "Batteries and ultracapacitors for electric, hybrid, and fuel cell vehicles," *Proc. IEEE,* vol. 95, pp. 806-820, 2007.

[7] F. Musavi, W. Eberle, and W. G. Dunford, "A phase shifted semi-bridgeless boost power factor corrected converter for plug in hybrid electric vehicle battery chargers," in *IEEE Applied Power Electronics Conference and Exposition (APEC), 2011*, 2011, pp. 821-828.

[8] A. Rufer and P. Barrade, "A supercapacitor-based energy-storage system for elevators with soft commutated interface," *IEEE Trans. Ind. Appl.,* vol. 38, pp. 1151-1159, 2002.

[9] S.-M. Kim and S.-K. Sul, "Control of rubber tyred gantry crane with energy storage based on supercapacitor bank," *IEEE Trans. Power Elect.,* vol. 21, pp. 1420-1427, 2006.

[10] L. Petersen and M. Andersen, "Two-stage power factor corrected power supplies: The low component-stress approach," in *IEEE Appl. Power Electron. Conf. and Exposition, 2002. APEC 2002.* , 2002, pp. 1195-1201.

[11] P. N. Enjeti and R. Martínez, "A high performance single phase AC to DC rectifier with input power factor correction," in *Appl. Power Electron. Conf. and Exposition, 1993. APEC'93.*, 1993, pp. 190-195.

[12] Y. Jang and M. M. Jovanovic, "A bridgeless PFC boost rectifier with optimized magnetic utilization," *IEEE Trans. Power Electron.,* vol. 24, pp. 85-93, 2009.

[13] P. Kong, S. Wang, and F. C. Lee, "Common mode EMI noise suppression for bridgeless PFC converters," *IEEE Trans. Power Electron.,* vol. 23, pp. 291-297, 2008.

[14] T. Baur, M. Reddig, and M. Schlenk, "Line-conducted EMI-behaviour of a High Efficient PFC-stage without input rectification," *Infineon Technology–Application Note,* 2006.

[15] M. Rivera, J. Rodriguez, B. Wu, J. R. Espinoza, and C. A. Rojas, "Current control for an indirect matrix converter with filter resonance mitigation," *Industrial Electronics, IEEE Transactions on,* vol. 59, pp. 71-79, 2012.

[16] J. Rodriguez, J. Pontt, C. A. Silva, P. Correa, P. Lezana, P. Cortés*, et al.*, "Predictive current control of a voltage source inverter," *Industrial Electronics, IEEE Transactions on,* vol. 54, pp. 495-503, 2007.

[17] F. Musavi, W. Eberle, and W. G. Dunford, "Efficiency evaluation of single-phase solutions for AC-DC PFC boost converters for plug-in-Hybrid electric vehicle battery chargers," in *Vehicle Power and Propulsion Conference (VPPC), 2010 IEEE*, 2010, pp. 1-6.

# A NewTransformerless Inverter for Grid Connected Photovoltaic System With Low Leakage Current

Monirul Islam[1*], Mahamudul Hasan[2], Parvez Akter[1], Md. Mizanur Rahman[3]
[1]Department of Electrical Engineering, University of Malaya, Kuala Lumpur, 50603 Malaysia
[2]Department of Mechanical Engineering, University of Malaya, Kuala Lumpur, 50603 Malaysia
[3]Department of Electrical and Electronic Engineering, CUET, Chittagong, Bangladesh
*E-mail: monir04eee@yahoo.com

*Abstract*—**Transformer-less inverters are more attractive for grid-tied photovoltaic (PV) system due to its higher efficiency and lower cost. But unfortunately, a leakage current flows through the system. So transformer-less PV inverter have to be tackled carefully. In this paper, a new transformer-less PV inverter topology with low leakage current has been proposed. Proposed circuit structure and detail operation principle are presented in this paper. One additional switch with conventional full H- Bridge and a diode clamping branch make sure the disconnection of PV module from the grid at the freewheeling mode and clamp the short circuited output voltage at the half of DC input voltage. Therefore, the common mode (CM) leakage current is minimized. The relationship between H5 topology and proposed topology has been analyzed in this paper. The proposed inverter topology is simulated by MATLAB / Simulink software to validate the accuracy of the theoretical analysis. Finally, a 1kW prototype has been built and tested.**

***Keywords—Common mode voltage, Converter, Grid connected, Leakage current, Photovoltaic, Transformerless.***

## I. INTRODUCTION

Transformer-less grid-tied inverters have many advantages such as higher efficiency, lower cost, smaller size, and weight. However, there is a galvanic connection between the power grid and the PV module; this may cause fluctuation of the potential between the solar cell array and the ground. This CM voltage $V_{cm}$ shown in equation (1) may induce a leakage current, $i_{Leakage}$ flowing through the loop consisting of the parasitic capacitors ( $C_{pvg1}$ and $C_{pvg2}$), the filter inductors, the bridge, and the utility grid [1]. As a result, the grid current harmonics and losses are increased. A person, who touch the PV array and connected to the ground, may conduct by the capacitive current. At the same time, conducted and radiated interference will be brought in by the ground current [2]. The instantaneous CM voltage $V_{cm}$ in the full bridge topology shown in Figure 1 is defined as follows [3].

$$V_{CM} = \frac{V_{AN} + V_{BN}}{2} \quad (1)$$

In order to eliminate the leakage current, the CM voltage must be kept constant during all operation modes and many solutions have been proposed [2-10] as follows:

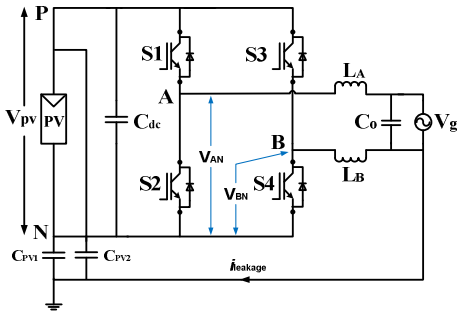



Figure 1. leakage current and parasitic capacitance in transformer-less grid connected PV system

*1. Bipolar sinusoidal pulse width modulated (SPWM) full bridge inverter:* The bipolar SPWM modulation can be used to solve the problem of the leakage current in the full bridge inverter. This inverter can keep constant CM voltage during the whole operation. However, the current ripples across the filter inductors and also the switching losses are large. Furthermore, it is critical to maintain a good synchronization among the gate signals of the bridge transistors [1].

*2. Unipolar sinusoidal pulse width modulated (SPWM) full bridge inverter:* The most common modulation used in this topology is unipolar SPWM, because it presents a lot of advantages in comparison to bipolar modulation such as higher dc voltage utilization, lower current ripple at high frequencies, better efficiency, lower electromagnetic interferences emission, etc. However, when unipolar SPWM is employed in the transformer-less conventional full H-bridge inverter; in active mode, the CM voltage $V_{cm}$ is equal to $0.5V_{pv}$ but in the freewheeling mode $V_{cm}$ is equal to $V_{pv}$ or zero depending on the leg midpoints (point *A* and *B* in Figure 1) connected to the positive or negative terminal of the input. Therefore, the CM voltage of conventional unipolar SPWM full-bridge inverter varies with high frequency, which leads to high leakage current [2].

To solve the aforementioned problem, the PV module should be separated from the grid during the freewheeling period and a lot of depth researches have been done to minimize the leakage current. Most of the topology has been proposed in the literature based on the structure of freewheeling path, where a new freewheeling path has been created [2-10].

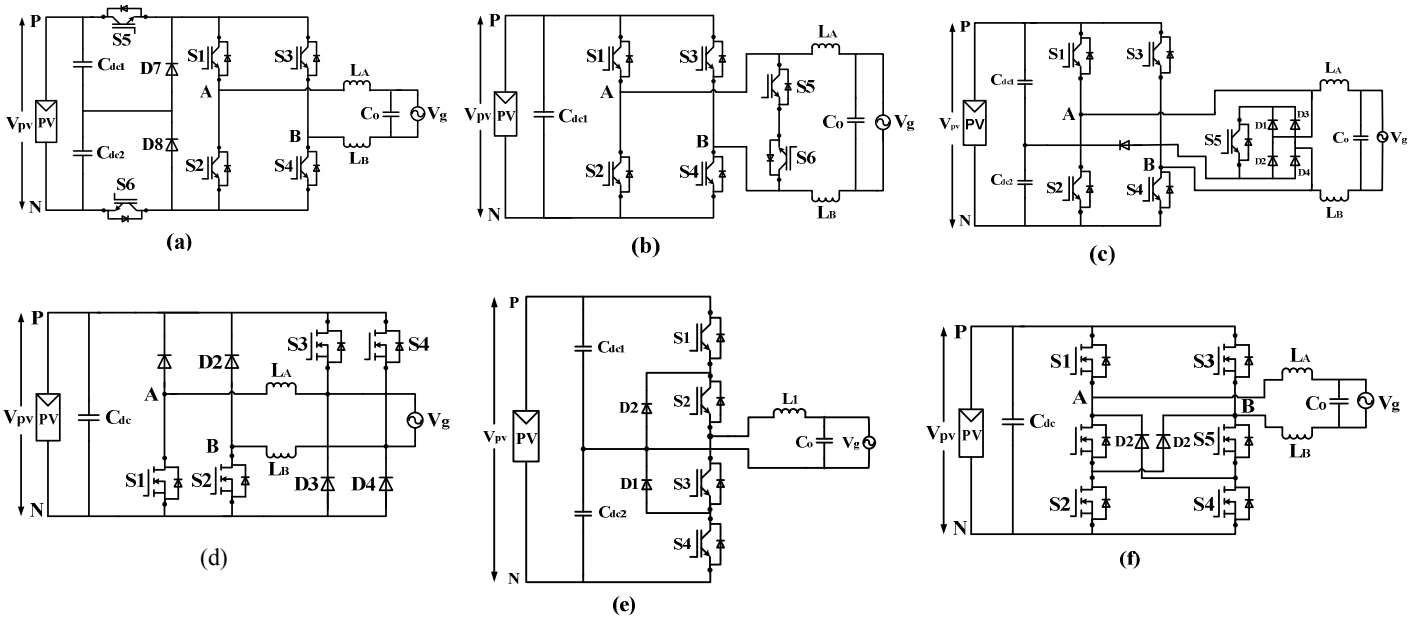



Figure 2. Some existing transformer-less topologies for grid-tied PV inverter (a) H6 topology proposed in [3] (b) HERIC topology proposed in [11] (c) HB-ZVR topology proposed in [7] (d) Dual-Paralleled-Buck inverter topology proposed in [12] (e) NPC topology proposed in [13] (f) H6 type topology proposed in [14]

In this paper, a new structure of transformer-less inverter has been proposed based on H5 topology. The proposed inverter ensures that the freewheeling voltage is clamped to the half of DC input voltage and the PV module is separated from the grid during the freewheeling period by adding an extra switch and a capacitor divider. As a result, the CM mode voltage kept constant during the whole operation period and the generated leakage current is almost zero. Also unipolar SPWM is applied to the proposed topology with three-level output voltage. The efficiency of the proposed inverter, H6 inverter and HERIC inverter are calculated and compared. Simulation and experimental results show that the proposed inverter operates with high efficiency and low leakage current.

This paper is prepared as follows: A review of single phase transformer-less topology is described in section II. CM voltage and leakage current of H5 topology is presented in section III. The proposed improved converter structure and operation principle with unipolar SPWM control scheme are investigated in section IV. Simulation and experimental results are presented in section V and section VI concludes the paper

## II. SINGLE PHASE TRANSFORMER-LESS TOPOLOGIES

### *A. H6 topology:*

Full-bridge inverter with dc bypass (FB-DCBP) topology has been proposed in [3] shown in Figure 2(a). Two extra switches are added into the conventional full H-bridge inverter in order to decouple dc bus from the grid during freewheeling period. Also a capacitive divider and two diodes are added to clamp the common mode voltage at half of the dc input voltage. The active voltage vector is achieved by switching S5 and S6 with high frequency. Switches S1 & S4 are switched with the grid frequency and in anti-parallel to S2 & S3, depending on whether the reference voltage is in the positive or negative half period. The output voltage has three levels. The main drawback of FB-DCBP topology is that it suffers more conduction losses from the inductor current flowing through four switches in the active mode [3].

### *B. HERIC Topology:*

Full-bridge inverter with ac bypass (FB-ACBP) topology has been proposed in [11] as shown in Figure 2(b), called Highly Efficient and Reliable Inverter Concept (HERIC). Two switches are added in the ac side to provide the path of freewheeling current. The PV module is decoupled from the grid during the freewheeling period because the switches S1, S2, S3, and S4 are turned-off. The output voltage of the inverter has three levels and the current ripple at output is very lower. The inverter generates constant CM voltage; therefore, the leakage current through the parasitic capacitance is minimized to safe level. Furthermore, the inverter efficiency is kept high because the load current is short circuited via S5 or S6 during the freewheeling period.

### *C. HB-ZVR Topology:*

Another full-bridge inverter topology with ac bypass is proposed in [7], where the midpoint of the dc link is clamped by means of a diode rectifier and one bidirectional switch. An extra diode is used to protect from short-circuiting the lower dc-link capacitor. The operational principle is same as HERIC topology. The zero voltage state is generated by short-circuiting the output of the inverter and clamping them to the midpoint of the dc-link. Three-level output voltage has been achieved by employing unipolar SPWM. The main disadvantage of this topology is the necessity of dead time which increases the distortion of output current [7].

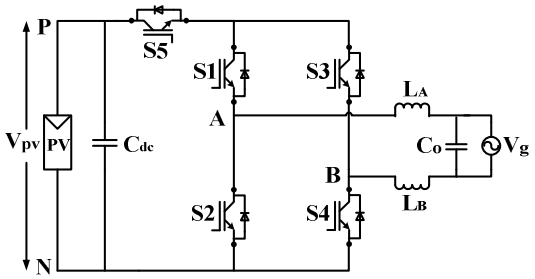



Figure 3. Power circuit structure of H5 topology

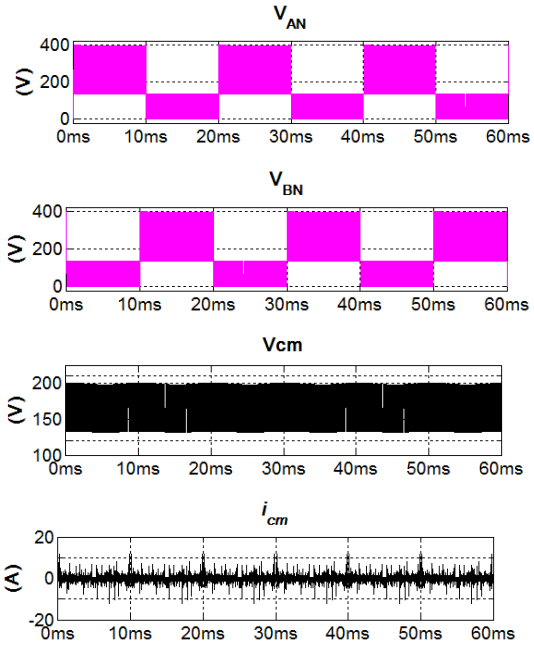



Figure 4. Simulated waveform of $V_{AN}$, $V_{BN}$, Vcm, and $i_{cm}$ of H5 topology

## III. H5 TRANSFORMER-LESS TOPOLOGY

### A. Circuit Configuration and Operation Principle:

An inverter topology proposed in [6] is called H5 topology which is shown in Figure 3. It is made up by adding an extra switch S5 with standard H-Bridge topology. $L_A$, $L_B$, and $C_o$ constructs the LCL type filter which is coupled to the grid. In this topology, unipolar SPWM is applied with three-level output voltage. This topology can meet the condition of eliminating CM leakage current. In the positive half cycle of grid current, switch S5 and S4 are commutates with switching frequency. During the zero voltage vectors, S5 and S4 are turned-off and the freewheeling current flows through S1 and the anti-parallel diode of S3. In the negative half cycle, S5 and S2 are commutates with switching frequency and the freewheeling current flows through S3 and the anti-parallel diode of S1.

### B. Common Mode Voltage and Leakage Current:

The CM voltage and the leakage current for H5 topology can be calculated using equation (1) and (2).

$$V_{CM} = \frac{V_{Ao} + V_{Bo}}{2} \tag{1}$$

$$i_{CM} = c_{PV} \frac{dV_{CM}}{dt} \tag{2}$$

where, $V_{AN}$ and $V_{BN}$ are the voltages of the full-bridge inverter from mid-point A and B of the bridge leg to the reference terminal N. Figure 4 shows the waveform of CM voltage and leakage current of H5 topology. We can see that a small variation of CM voltage is existed and as a result, a non-negligible leakage current is generated.

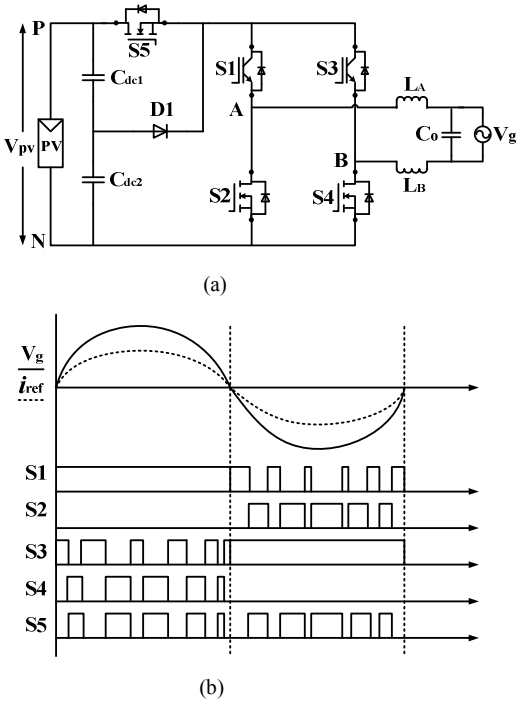



Figure 5. Proposed transformers-less grid-tied PV inverter (a) Proposed circuit configuration (b) Switching signals with unity power factor

## IV. PROPOSED TOPOLOGY

### A. Circuit Structure:

In order to de-couple the converter from the grid in the freewheeling mode, an extra MOSFET switch is added into the conventional full H-Bridge topology and the two lower high frequency IGBT switches of two phase legs are replaced by MOSFET switches in this paper which is shown in Figure 5(a). Also a diode and a capacitor divider are added to clamp the short circuited output voltage at the mid-point of DC link voltage. $L_A$, $L_B$, and $C_o$ constructs the LCL type filter coupled to the grid. This topology can achieve three-level output voltage with unipolar SPWM.

### B. Operation Principle analysis:

Grid tied photovoltaic system generally operate at unity power factor. Figure 5(b) shows the waveform of the switching pattern for the proposed topology. The operation principle is very similar to the H5 topology shown in Figure 6. Consequently, four operational modes are proposed that produce the output voltage states of $+V_{PV}$, 0, and $-V_{PV}$.

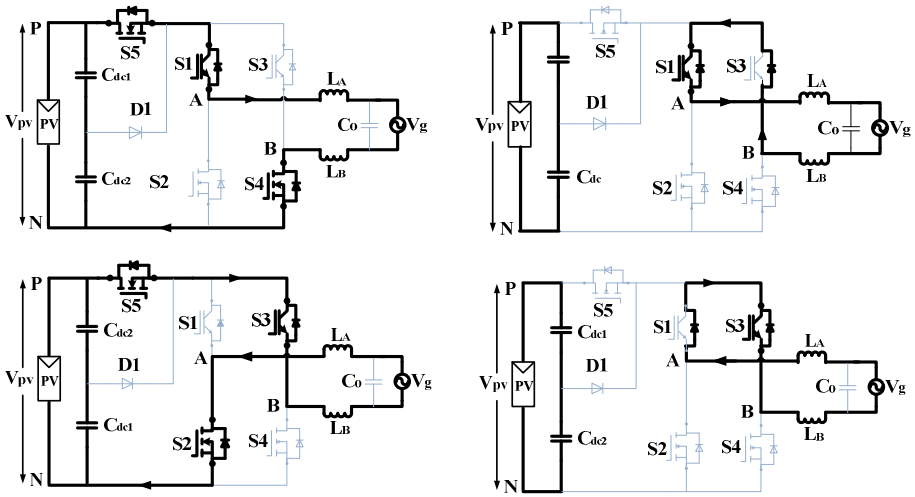



Figure 6. Operation principle of the proposed topology (a) active mode and (b) freewheeling mode in the positive half cycle of grid current (c) active mode and (d) freewheeling mode in the negative half cycle of grid current

1) Mode 1 is the active mode in the positive half cycle of grid current. When S1, S4 and S5 are turned-on, the voltage $V_{AN} = V_{PV}$ and $V_{BN} = 0$, thus $V_{AB} = V_{PV}$ and the CM voltage, $V_{cm} = (V_{AN} + V_{BN})/2 = V_{PV}/2$.

2) Mode 2 is the freewheeling mode in the positive half cycle of grid current as shown in Figure 6(b). The freewheeling current flows through S1 and body diode of S3. In this mode, the freewheeling output voltage is clamped to the half of DC input voltage thorugh diode D1, thus $V_{AN} = V_{BN} = V_{PV}/2$, $V_{AB} = 0$ and the CM voltage, $V_{cm} = (V_{AN} + V_{BN})/2 = V_{PV}/2$.

3) Mode 3 is the active mode in the negative half cycle of grid current. Like as mode 1, when S2, S3 and S5 are turned-on, the voltage $V_{AN} = 0$ and $V_{BN} = V_{PV}$, thus $V_{AB} = -V_{PV}$ and the CM voltage, $V_{cm} = (V_{AN} + V_{BN})/2 = V_{PV}/2$.

4) Mode 4 is the freewheeling mode in the negative half cycle of grid current. When S5 and S2 are turned-off, the freewheeling current flows through S3 and body diode of S1. In this mode, $V_{AN} = V_{BN} = V_{PV}/2$, thus $V_{AB} = 0$ and the CM voltage, $V_{cm} = (V_{AN} + V_{BN})/2 = V_{PV}/2$.

As analysis above, the CM voltage remains constant during the whole operation period of the proposed inverter and equals to $V_{PV}/2$. As a result, the inverter is hardly to generate CM leakage current.

## V. RESULTS

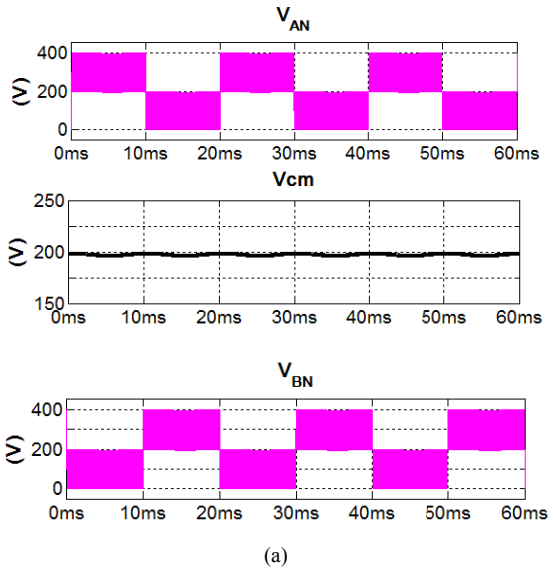
In order to verify the theoretical analysis, the proposed topology has been simulated by MATLAB/Simulink software using the parameters given in Table I. Figure 7 shows the simulated and experimental waveform of $V_{AN}$, $V_{BN}$, and Vcm. From figure 7(a), it can be seen that the CM mode voltage kept constant at the half of DC input voltage 200V. As a result, the generated leakage current is almost zero which is shown in Figure 8(c). Experimental Results of CM mode voltage are shown in Figure 7(b), which shows the same characteristics as simulation. Experimental result of leakage current is shown in Figure 9, which shows that the peak and RMS values are successively limited within 45mA and 8mA, respectively.

Figure 8 shows the simulated waveform of grid current, grid voltage, differential-mode voltage, and leakage current. From Figure 8(a) and Figure 9, it is clear that the output voltage and current of the proposed inverter has low harmonic distortion which can meet the requirements of IEEE STD 1547.1™-2005 [15]. The proposed inverter has three-level output voltage $+V_{PV}$, 0, and $-V_{PV}$ which is shown in Figure 8(c). It designates that the proposed topology is modulated with unipolar SPWM with excellent differential-mode characteristics.

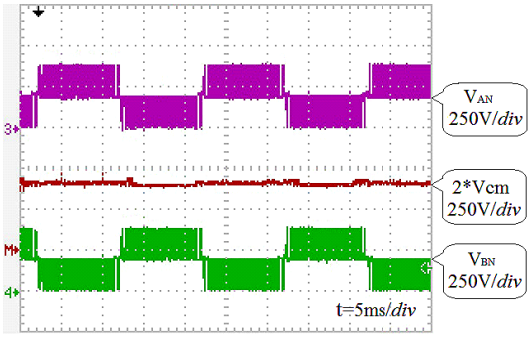
The efficiency comparison among the proposed, HERIC and H6 topologies with the same parameter is illustrated in Figure 10. Note that the presented efficiency diagram covers the total device losses but it does not include the losses for control circuit. It is clear that the efficiency of the proposed topology is higher than the H6 topology and almost same to the HERIC topology.

TABLE I. PARAMETERS USED IN SIMULATION

| Inverter Parameter | Value |
|---|---|
| Input Voltage | 400VDC |
| Grid Voltage / Frequency | 230V / 50Hz |
| Rated Power | 1000 W |
| AC output current | 4.1A |
| Switching Frequency | 20kHz |
| DC bus capacitor | 470µF |
| Filter capacitor | 2.2µF |
| Filter Inductor $L_A$, $L_B$ | 3mH |
| PV parasitic capacitor Cpv1, Cpv2 | 75nF |




(a)




(b)

Figure 7. Simulated waveform of CM voltage and leakage current (a) $V_{AN}$, Vcm, $V_{BN}$ (b) $i_{cm}$

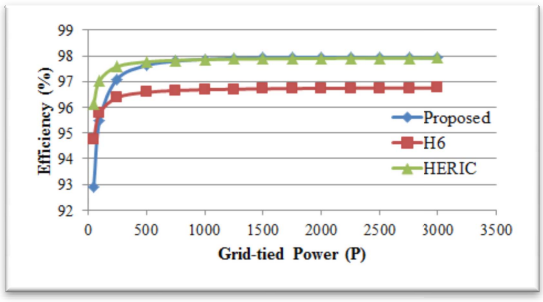
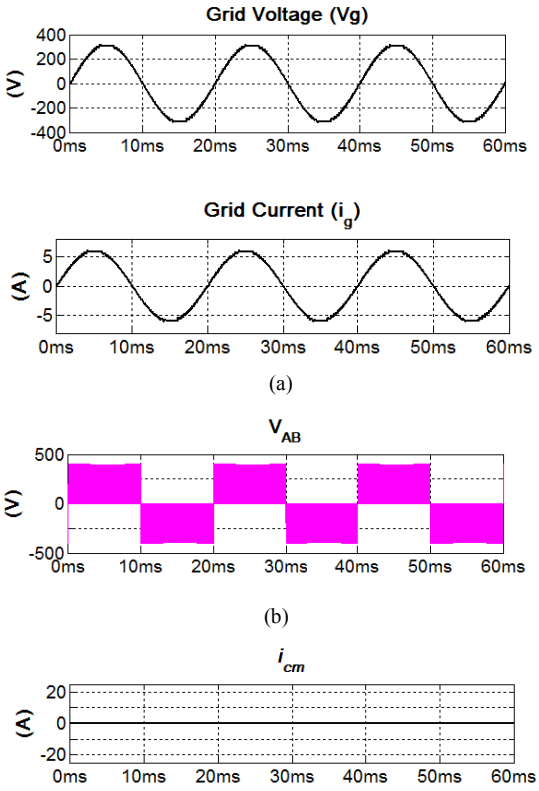



Figure 10. Efficiency comparison of proposed, H6, and HERIC topology




(a)




(b)




(c)

Figure 8. Simulated waveform of output voltage (a) grid voltage and grid current (b) differential output voltage (c) Leakage current

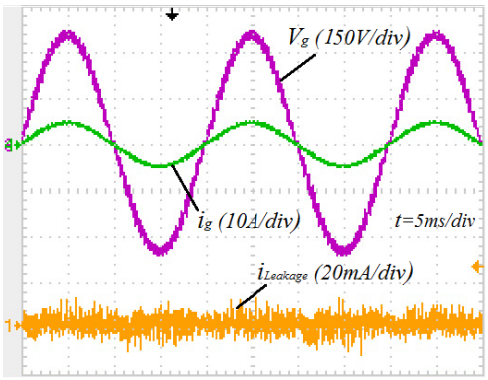



Figure 9. Simulated waveform of output voltage (a) grid voltage and grid current (b) differential output voltage

## VI. CONCLUSION

In this paper, a new single phase transformer-less inverter for grid connected PV system have been presented. Also major transformer-less topologies are reviewed based on leakage current, advantages, and disadvantages. It is found that these topologies have some drawbacks such as leakage current and efficiency. However, The CM voltage of the proposed topology is clamped to half of DC input voltage, thus leakage current is well suppressed. The inverter output voltage has three levels by employing unipolar SPWM with good differential mode characteristics. The waveform of output current shows that the proposed inverter can convert the solar power to a high quality ac power to inject into utility grid.

Furthermore, the proposed inverter achieved higher efficiency. Therefore, it can be concluded that the proposed inverter is very suitable for single phase grid connected PV system.

## REFERENCES


[1] I. Patrao, E. Figueres, F. González-Espín, and G. Garcerá, "Transformerless topologies for grid-connected single-phase photovoltaic inverters," *Renewable and Sustainable Energy Reviews,* vol. 15, pp. 3423-3431, 2011.

[2] Z. Li, S. Kai, F. Lanlan, W. Hongfei, and X. Yan, "A Family of Neutral Point Clamped Full-Bridge Topologies for Transformerless Photovoltaic Grid-Tied Inverters," *IEEE Transactions on Power Electronics,* vol. 28, pp. 730-739, 2013.

[3] R. Gonzalez, J. Lopez, P. Sanchis, and L. Marroyo, "Transformerless Inverter for Single-Phase Photovoltaic Systems," *IEEE Transactions on Power Electronics,* vol. 22, pp. 693-697, 2007.

[4] G. Buticchi, D. Barater, E. Lorenzani, and G. Franceschini, "Digital Control of Actual Grid-Connected Converters for Ground Leakage Current Reduction in PV Transformerless Systems," *IEEE Transactions on Industrial Informatics,* vol. 8, pp. 563-572, 2012.

[5] Y. Bo, L. Wuhua, G. Yunjie, C. Wenfeng, and H. Xiangning, "Improved Transformerless Inverter With Common-Mode Leakage Current Elimination for a Photovoltaic Grid-Connected Power System," *IEEE Transactions on Power Electronics,* vol. 27, pp. 752-762, 2012.

[6] M. Victor, F. Greizer, S. Bremicker, and U. Hübler, "Method of converting a direct current voltage from a source of direct current voltage, more specifically from a photovoltaic source of direct current voltage, into a alternating current voltage," ed: United States Patents, 2008.

[7] T. Kerekes, R. Teodorescu, Rodri, x, P. guez, Va*, et al.*, "A New High-Efficiency Single-Phase Transformerless PV Inverter Topology," *IEEE Transactions on Industrial Electronics,* vol. 58, pp. 184-191, 2011.

[8] D. Barater, G. Buticchi, A. S. Crinto, G. Franceschini, and E. Lorenzani, "A new proposal for ground leakage current reduction in transformerless grid-connected converters for photovoltaic plants," in *35th Annual Conference of IEEE Industrial Electronics (IECON '09)*. 2009, pp. 4531-4536.

[9] M. Kazanbas, C. Noding, H. Can, T. Kleeb, and P. Zacharias, "A new single phase transformerless photovoltaic inverter topology with coupled inductor," in *6th IET International Conference on Power Electronics, Machines and Drives (PEMD 2012),*, 2012, pp. 1-6.

[10] X. Huafeng, X. Shaojun, C. Yang, and H. Ruhai, "An Optimized Transformerless Photovoltaic Grid-Connected Inverter," *IEEE Transactions on Industrial Electronics,* vol. 58, pp. 1887-1895, 2011.

[11] H. Schmidt, S. Christoph, and J. Ketterer, "Current inverter for direct/alternating currents, has direct and alternating connections with an intermediate power store, a bridge circuit, rectifier diodes and a inductive choke," *Germany Patent DE10221592 (A1),* 2003.

[12] S. V. Araujo, P. Zacharias, and R. Mallwitz, "Highly Efficient Single-Phase Transformerless Inverters for Grid-Connected Photovoltaic Systems," *IEEE Transactions on Industrial Electronics,* vol. 57, pp. 3118-3128, 2010.

[13] R. Gonzalez, E. Gubia, J. Lopez, and L. Marroyo, "Transformerless Single-Phase Multilevel-Based Photovoltaic Inverter," *IEEE Transactions on Industrial Electronics,* vol. 55, pp. 2694-2702, 2008.

[14] W. Yu, J.-S. Lai, H. Qian, and C. Hutchens, "High-efficiency MOSFET inverter with H6-type configuration for photovoltaic nonisolated AC-module applications," *IEEE Transactions on Power Electronics,* vol. 26, pp. 1253-1260, 2011.

[15] "IEEE Standard Conformance Test Procedures for Equipment Interconnecting Distributed Resources With Electric Power Systems," *IEEE Std 1547.1-2005,* pp. 0_1-54, 2005.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Grid Frequency Analysis with the Issue of High Wind Power Penetration.

Md. Emdadul Haque[1], Shuva Paul[2], and Md. Rafiqul Islam Sheikh[3]
[1,2,3] Department of Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204 , Bangladesh.
E-mail: emdadieee@gmail.com[1]

*Abstract*— **As wind turbine output is proportional to the cube of wind speed, the wind turbine generator output fluctuates due to random variations of wind speed. Hence, if the power capacity of wind power generator becomes large, wind power generator output can have an influence on the power system frequency. Therefore, this study investigates the influence of the ratio of wind capacity to the power system capacity, on the power system frequency. To do this, a 100[MVA] thermal synchronous generator (SG) is considered and the total capacity of wind power induction generator (IG) increases as 3[MVA], 5[MVA] and 10[MVA] connected to the line. It is seen that the system frequency fluctuation is more severe for 10% capacity of wind power penetration. Again, a hydro generator to the same capacity that of the thermal generator of 100[MVA] is considered and IG capacity increases as before. In this case, it is seen that the system frequency fluctuation is more severe for 10% capacity of wind power. However, SGs total capacity of 100[MVA] is considered combined with 50[MVA] thermal generator and 50[MVA] hydro generator and IG capacity increases as before. In this case, it is also seen that the system frequency fluctuation is more severe for 10% of total wind power capacity. For this reason , as the wind power penetration increases rapidly nowadays for its clean, non-polluting and renewable energy, this study will be helpful for taking preventive measures for the power grid operators to improve the stability and quality of electric power. Considering these point of view, this study plays a vital role for power system applications.**

***Keywords— Power system frequency, AVR, Wind power penetrationl, Cut-inand cut-out speed.***

## I. INTRODUCTION

We are now at a crucial stage of our global energy scenario. Energy has been the important driving force of the continual progress of human civilization. Since the industrial revolution of the two centuries ago, global energy consumption has increased by leaps and bounds to improve our living standards, particularly in the industrialized nations of the world. The growth of wind energy around the world in recent years has been consistently impressive. The increase in wind power penetration level is likely to influence the operation of the existing utilities networks, especially the power system stability[1].The world experienced a great disaster in the year Fukushima nuclear disaster on 11 March, 2011.Considering this hazard the energy consumption has changed its route.

According to the current energy resources shown in Fig.1 about 47.7% energy comes from coal, 26.3% from natural gas, 16.3% from wind and remaining 9.7% comes from oil, biomass and solar of the total energy capacity in the year 2012 and in the year 2013 the wind energy consumption trends to increase considerably and it is 77.8% of the total energy.

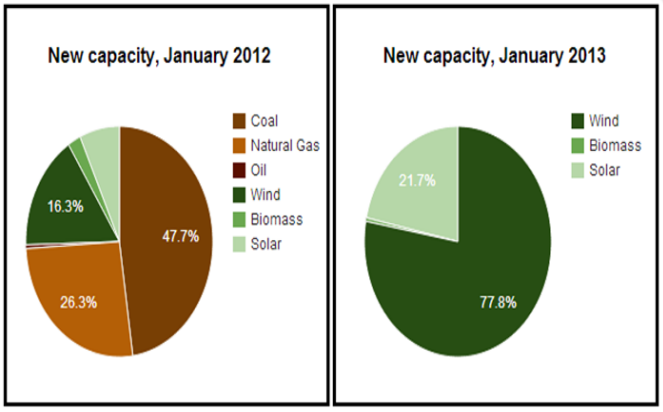



Figure 1. Total energy use.

In the conventional operation of wind power generators, when the wind speed is between the rated speed and the cut out speed, the wind power generator output is controlled at the rated value by a pitch control system [2]. On the other hand, when the wind speed is between the cut in speed and the rated speed, the blade pitch angle is maintained constant (= 0 deg),in general, for the wind turbine to capture the maximum power from the wind turbine. Therefore, the wind power generator output fluctuates due to wind speed variations in the latter condition, because the wind power is proportional to the cube of wind speed. In the previous study [3], it is seen that thermal governor perform better frequency control than hydro governor. Therefore, it is necessary to investigate the influence of the ratio of the wind generator capacity to the power system capacity, on power system frequency. If the power capacity of wind generators becomes large, wind generator output can have an influence on the power system frequency [4]-[7]. In this study, at first thermal generator and hydro generator is considered separately. When the number of wind generator increases as the total capacity of 3[MVA], 5[MVA] and

10[MVA] and keeping constant the total capacity of 100[MVA] thermal generator connected in grid (Case-1, Case-2, Case-3), it shows that the system frequency fluctuation increases with the increasing IG capacity. Again, When the number of wind generator increases as the total capacity of 3[MVA], 5[MVA] and 10[MVA] and keeping constant the total capacity of 100[MVA] hydro generator (Case-4, Case-5, and Case-6), it shows that the system frequency fluctuation also increases with the increasing IG capacity. And now, keeping constant the SGs total capacity of 100[MVA] combination of 50 [MVA] thermal generator and 50[MVA] hydro generator and increases total IG capacity as 3[MVA], 5[MVA] and 10[MVA] (Case-7, Case-8,and Case-9).

It is investigated that in each case the system frequency fluctuation is more severe for 10% of total wind power penetration.

## II. Model System for Simulation Analysis

The model system used in the PSCAD simulation analyses is shown in Fig. 2 One IG actually consists of a number of induction generators (IGs) of capacity [3MVA], [5MVA], and [10MVA] and Two synchronous generators capacity SG1 [50 MVA] Salient pole type with hydro governor and SG2 [50MVA] cylindrical type with thermal governor forming the total capacity of 100 MVA of SGs are used with the network. $Q_{WF}$ and $Q_{Load}$ are capacitor banks. $Q_{WF}$ is used at the terminal of IG to compensate the reactive power demand of wind generator at steady state. The value of the capacitor is chosen so that the p.f. becomes unity, when the wind generator operated in the rated condition [2]. $Q_{load}$ is used at the terminal of load to compensate the voltage drop by the impedance of transmission lines. Core saturations of induction generators and synchronous generators are not considered for simplicity. Parameters of IGs and SGs are shown in table I to II and initial condition is shown in table III. All parameters are expressed in per unit system. The transformers rating used for transmission and distribution are selected wisely so that they are in matched condition when generator to transmission and transmission to distribution respectively.

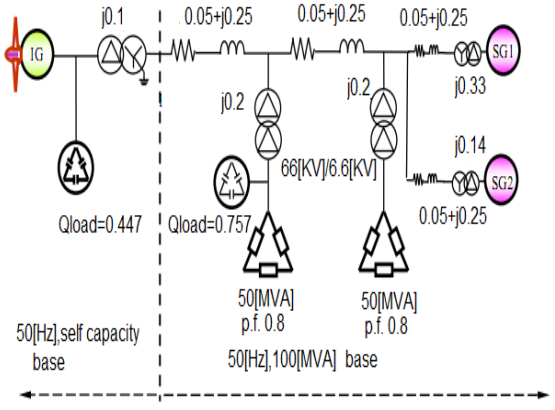



Figure 2. Simulation model.

## III. Synchronous Generator Model

### A. Governor

The governor is a device that automatically adjusts the rotational speed of the turbine and the generator output. When the generator load is constant, the turbine is operated at a constant rotational speed. However, when the load changes, balance between the generator output and the load is not maintained, and the rotational speed changes. When the load is removed, the governor detects the increase of the rotational speed, and then, the valve is closed rapidly so that an abnormal speed increase of the generator is prevented [2].

TABLE I. Parameters of Induction Generator

| **Induction Generator** | | | |
|---|---|---|---|
| [MVA] | 3 | 5 | 10 |
| [MVA] | 3 | | |
| X [pu] | 0.18 | | |
| Xm[pu] | 10 | | |
| R2[pu] | 0.015 | | |
| X2 [pu] | 0.12 | | |
| 2H[sec] | 1.5 | | |

TABLE II. Parameters of Synchronous Generator

| **Synchronous Generator** | | |
|---|---|---|
| | ***Salient Pole Type(HG)*** | ***Cylindrical Type (TG)*** |
| Xd[pu] | 1.2 | 2.11 |
| Xq[pu] | 0.7 | 2.02 |
| Xd[pu] | 0.3 | 0.28 |
| Xd[pu] | 0.22 | 0.215 |
| Xq[pu] | 0.25 | 0.25 |
| Tdo[sec] | 5.0 | 4.2 |
| Tdo[sec] | 0.05 | 0.032 |
| Tqo[sec] | 0.14 | 0.062 |
| H[sec] | 2.5 | 2.32 |

TABLE III. Initial Condition

| | **IG** | **SG1** | **SG2** |
|---|---|---|---|
| P [pu] | 0.03/0.05/0.10 | 1.00 | 1.00 |
| V[pu] | 1.00 | 1.05 | 1.05 |
| Q[pu] | 0.00 | | |
| S(Slip) | -1.733% | | |

*B. Automatic Voltage Regulator (AVR)*

To keep the voltage constant of the synchronous generators, AVR is needed [2]. In the simulation analyses, the AVR is expressed as simplified form. AVR model is shown in Fig.3. Parameters of AVR are shown in table IV. A typical PID controller is adopted in the voltage feedback control loop, considering that there is also a different component in PID controller, a low pass filter is added in this loop [8]. The transfer function of a typical PID controller is shown in equation (1) [9]-[10].

$$G_d(s) = K_G\left(K_P + \frac{K_1}{s} + \frac{K_d s}{1+T_d s}\right) \quad (1)$$

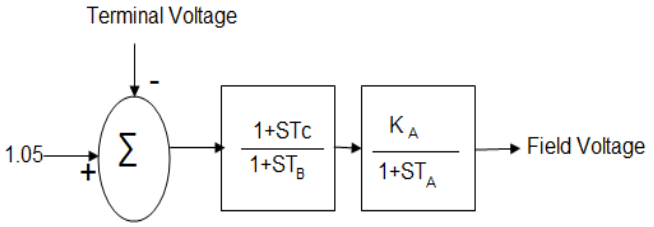



Figure 3. AVR Model

TABLE IV. PARAMETER OF AVR MODEL

| | |
|---|---|
| Gain ,$K_P$ [pu] | 400 |
| Time Constant,$T_A$[sec] | 0.02 |
| Time Constant,$T_B$=$T_C$[sec] | .00 |

*C. Pitch Controller*

The method normally used, in which the angle of the rotor blades can be actively adjusted by the machine control system. This is known as pitch control. Pitch controller is used to maintain the output power of the wind generator at rated level by controlling the blade pitch angle of turbine blade when the wind speed is over the rated speed [11].

## IV. WIND TURBINE MODELING

In this study, the MOD-2 model [12] is considered for the $C_p$-λ characteristics, which is represented by the following equations and shown in Fig. 4.for different values of blade pitch angle, β. Mod-2 system data is shown in table V. The captured power from the wind can be obtained from (2) [13]. Tip speed ratio, λ, and power coefficient, $C_P$, can be expressed as (3) and (4). Since $C_P$ is expressed in feet and mile, Γ is corrected as (5).

$$P_{WTB} = \frac{1}{2}\rho C_P(\lambda)\pi R^3 \quad (2)$$

$$\lambda = \frac{\omega_{wtb} R}{V_w} \quad (3)$$

$$C_P(\lambda) = 0.5(\tau - 0.022\beta^2 - 5.6)e^{-0.17\tau} \quad (4)$$

$$\tau = \frac{R}{\lambda} \cdot \frac{3600}{1609} \quad (5)$$

The torque coefficient and the wind turbine torque are shown as follows

$$C_t(\lambda) = \frac{Cp(\lambda)}{\lambda} \quad (6)$$

$$\tau_M = \frac{1}{2}\rho C_t(\lambda)\pi R^3 V_w^2 \quad (7)$$

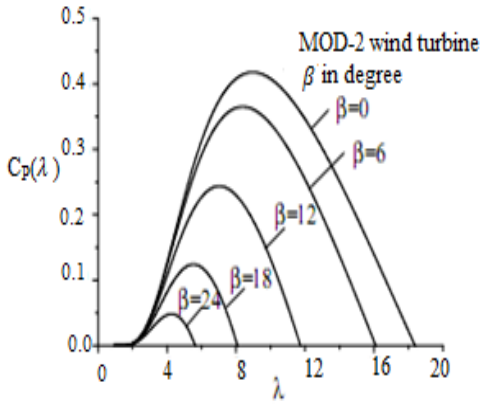



Figure 4. $C_P(\lambda)$ versu $\lambda$ curve for different values of pitch angle.

Where, $P_{wtb}$ is the wind turbine output [W], R is the radius of the blade [m], $\omega_{wtb}$ is the wind turbine angular speed [rad/s], β is the blade pitch angle [deg], $V_w$ is the wind speed [m/s], ρ is the air density[kg/m$^3$], and $\tau_M$ is the wind turbine output torque [N-m].

TABLE V. MOD-2 SYSTEM DATA

| | |
|---|---|
| Power Capacity | 2.5[MW] |
| Rotational speed | 15.55[rpm] |
| Wind turbine radius | 45.73[m] |
| Rated wind speed | 28[mph]*=12.51[m/s] |
| Cut-in wind speed | 13[mph]*=5.81[m/s] |
| Cut-out wind speed | 45[mph]*=20.11[m/s] |

*1mile=1609m, and 1 hour=3600sec

## V. SIMULATION RESULTS

Simulation analyses have been carried out to investigate the performance of the power system frequency with the increased wind power penetration using wind speed data and associated data that is required. In this study, the first 100 [sec] is considered for run up to steady state condition of the wind generator inertia and hence the first 100 [sec] have not considered. For this reason, all the curves shown in the fig. are started from 100 [sec]. The wind speed data applied to the wind generator is shown in Fig. 5.The wind speed data are chosen in such a way that the values vary within the cut-in and cut-out speed. The wind speed data applied on PSCAD (Power

System CAD) from a text file Simulation analyses have been carried out for nine cases shown in Table VI. in order to investigate the influence of the frequency on power system. The simulation analyses have been performed by using PSCAD 4.2.0. Fig.6 shows the wind generator output real power and Fig.7 shows the system frequency for case-1, case-2, and case-3. Fig. 8and 9 shows the IG output real power and frequency response respectively for case-4, case-5 and case-6.It is also evaluated IG output real power and system frequency for cases 7, 8 and 9 in Fig.10 and Fig. 11. Finally, the system frequency for cases 3, 6 and 9 is shown in Fig.12.The evaluation of the results has been presented in Table VII of system frequency. From the table it is seen that system frequency becomes more severe in the case-3, case-6, and case-9. The case-3, case-6, and case-9 are modeled in the simulation block on the bases of wind generator capacity increases 10% of the total penetration. This is the point where different from the other cases.

TABLE VI. SIMULATION PATTERN

| Cases | IG[MVA] | SG1[MVA] with Thermal Governor | SG2[MVA] with Hydraulic Governor |
|---|---|---|---|
| Case-1 | 3 | 100 | × |
| Case-2 | 5 | 100 | × |
| Case-3 | 10 | 100 | × |
| Csae-4 | 3 | × | 100 |
| Case-5 | 5 | × | 100 |
| Case-6 | 10 | × | 100 |
| Case-7 | 3 | 50 | 50 |
| Case-8 | 5 | 50 | 50 |
| Case-9 | 10 | 50 | 50 |

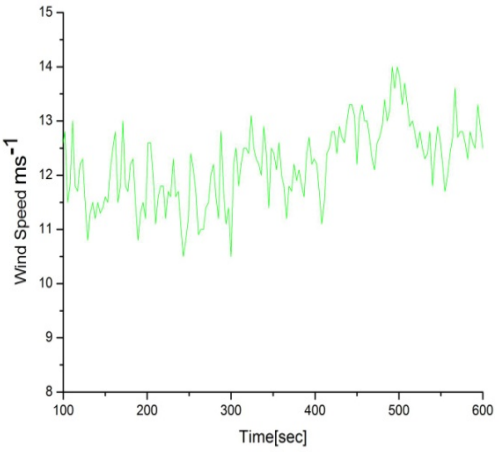



Figure 5. Wind speed data.

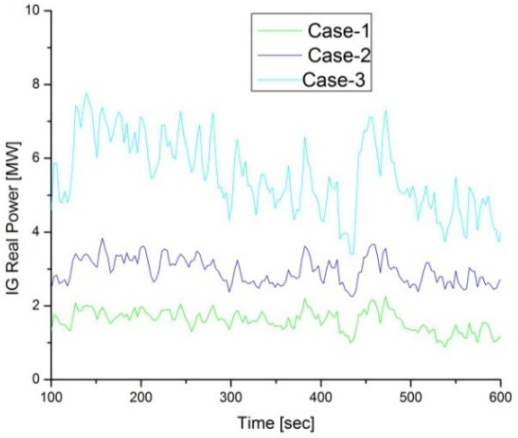



Figure 6. Wind generator output power.

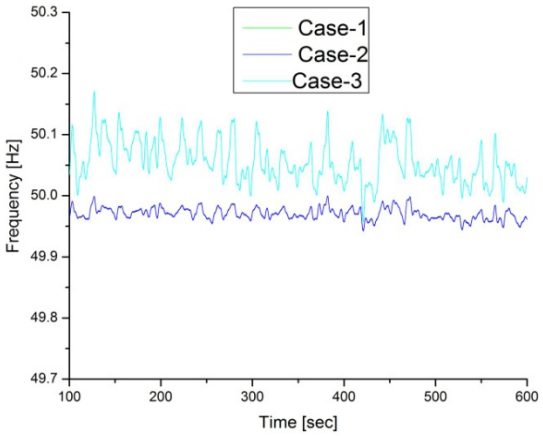



Figure 7. System frequency.

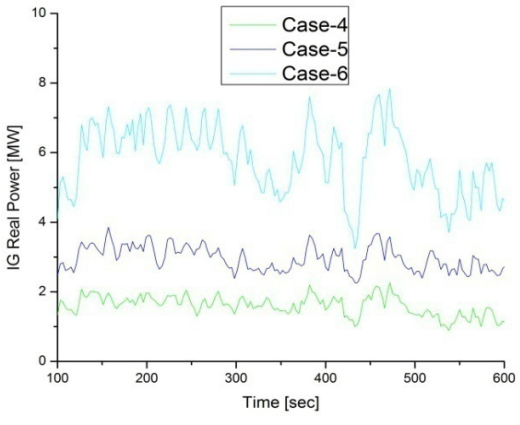



Figure 8. Wind generator output power.

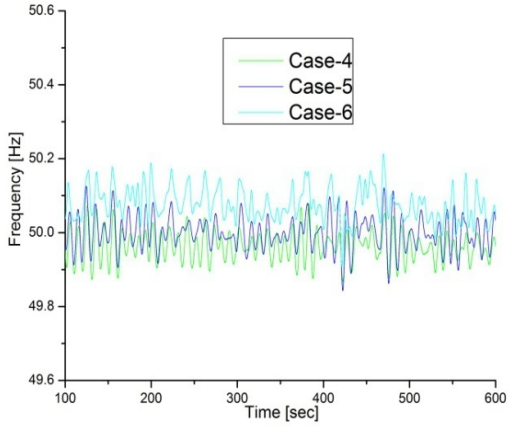



Figure 9. System frequency.

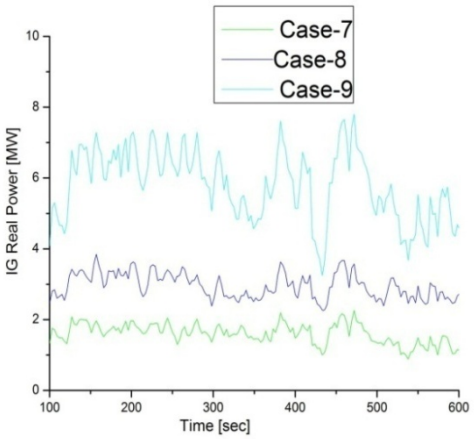



Figure 10. System frequency.

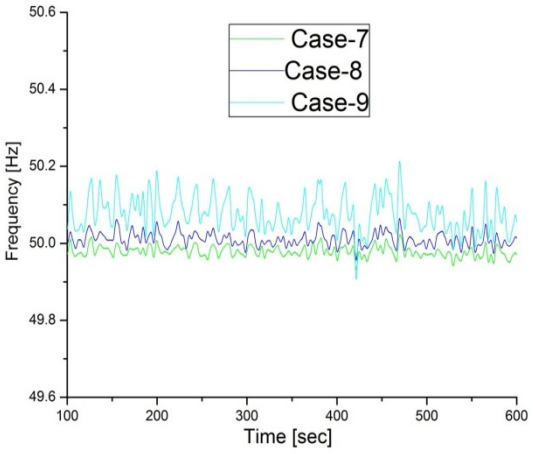



Figure 11. System frequency.

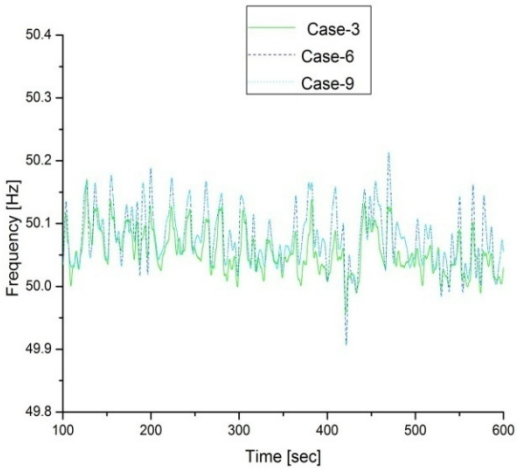



Figure 12. System frequency.

TABLE VII. EVALUATION OF SIMULATION RESULTS

| [Cases] | '0 'and '×' of frequency fluctuations |
|---|---|
| Case-1 | 0 |
| Case-2 | 0 |
| Case-3 | × |
| Case-4 | 0 |
| Case-5 | 0 |
| Case-6 | × |
| Case-7 | 0 |
| Case-8 | 0 |
| Case-9 | × |

'0' means within ±0.2 [Hz] and '×' means beyond ±0.2 [Hz].

## VI. DISCUSSION

A permissible range of power system frequency deviation provided by the most of the developed countries is within ± 0.2[Hz]. Evaluating the frequency response Fig. 8, 11 and 14, it is seen that in case of wind generator capacity of 3[MVA] and 5[MVA] the frequency fluctuation is within the acceptable limit but when the total capacity of 10[MVA] IG, the frequency fluctuation exceeds the permissible range.

## VII. CONCLUSIONS

Since frequency is one of the measures to determine the quality of electric power, this study investigates the influence of frequency fluctuations on power system with high wind power penetration. However, from the results presented, it is confirmed that the high wind power gives rise to the fluctuations of the system frequency as the power capacity of wind becomes large, and system becomes worse when wind power penetration to the grid increases beyond 10% of the total capacity. Therefore some preventive measures should be taken for all the cases 3, 6 and 9 by power grid operators to supply reliable electric power.

.

## REFERENCES


[1] A. Ramkumar, S. Durairraj, ‘‘Coordinated control of interconnected hydro governor synchronous generator with siege,’’Power, Energy and Control (*ICPEC),* Feb.2013. pp. 635-641.

[2] M.R.I. Sheikh, M.M. Haque, and Hossain, ‘‘Performance of governor system on minimizing frequency fluctuations with wind power generation’’, *International Journal of Electrical and Computer Engineering (IJECE),* vol. 2, no. 1, pp. 46–56, Feb. 2012.

[3] M.R.I. Sheikh, R. Takahashi, and J. Tamura, *‘‘*Study on frequency fluctuations in power system with a large penetration of wind power generation*’’, International Energy Journal ,* vol. 1, no. 1, Mar 2011.

[4] Takua Yamazaki, R Takahashi Maruta, Tanji Tamura, Fukushima, Eiichi Ssano, Koji Shinya, and Takatoshi Matstumoto , ‘‘Smoothing control of wind generator output fluctuations by new pitch controller’’, *IEEJ Transactions on Power and Energy,* vol. 129, no. 7, pp. 880-888,Jul. 2009.

[5] C Luo, and B. T Ooi, ‘‘Frequency deviation of thermal power plants due to wind farms’’, IEEE Transaction on Energy Conversion,’’ vol. 21, no. 3, pp. 708-716, Sep. 2006.

[6] C. Carrillo, A.E. Feijoo, J. Cidras, and J. Gonzalez, *‘‘*Power fluctuation in an isolated wind plant’’, , *IEEJ Transactions on Power and Energy,* vol.19, no. 1, pp. 217-221, March 2004.

[7] H. Umeki, R.Takahashi, T. Murata, and J. Tamura, ‘‘Smoothing control of wind generator output fluctuations by new pitch controller’’, *The Int. Conf. of Electrical Machines and Systems(ICEMS),LS3C-1,2006.*

[8] Q-G. Wang,Q. Bi,and Y.Zhang, ‘‘Partial internal model control,’’ *IEEE Transactions on Industrial Electronics*, vol. 48, no. 5, pp. 976–982, Oct.2001.

[9] K. Kim, RC. Schaefer, ‘‘Tuning a PID controller for a digital excitation control system,’’ vol. 41, no. 2, pp. 485–492, March. 2005.

[10] K.Kiyong, P.Rao, J.A.Burnworth, ‘‘self tuning of the PID controller for a digital excitation control system*,’’ IEE Transaction on Industry Applications,* vol. 46, no.4, pp.1518-1524, Jul.2010.

[11] Takua Yamazaki, R Takahashi Maruta, Tanji Tamura, Fukushima, Eiichi Ssano, Koji Shinya, and Takatoshi Matstumoto , ‘‘Smoothing control of wind generator output fluctuations by new pitch controller’’, *IEEJ Transactions on Power and Energy,* vol.129, no. 7, pp. 880-888,Jul. 2009.

[12] Technical Reports, ‘‘Standard models of electrical power system,’’ *IEE of Japan,* vol.754, pp. 40-43.

[13] Z. Bubosny, ‘‘Wind turbine operation in electric power system-advanced modeling,’’ *Spinger Verlag,* New York, 2013.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# A Variable Speed Three Phase Generator Voltage Hook up With a DC Bus by VIENNA Rectifier

Khizir Mahmud[1*], Lei Tao[2], Md. Shamsul Alam[3]
[1,2]Electrical Engineering and Automation School, Northwestern Polytechnical University, Xi'an, P.R. China
[3]Department of Electrical and Electronic Engineering, Chittagong University of Engineering and Technology, Bangladesh
[*] E-mail: khizirbd@gmail.com

*Abstract*—**Different converters like AC-DC, DC-DC or DC-AC is an inevitable part of the modern electronic equipments. Among them the active three-phase converters give a steady DC output voltage and also makes the input current shape sinusoidal. It also improves the power factor by maintaining the input voltage and current in phase. The major concentration of the paper is to develop an interface technique between variable speed three-phase generators and a DC bus. The interface technique is done by VIENNA rectifier topology where it convert a variable voltage and variable frequency to constant output DC voltage. This type of interface can be used in wind generation systems employing AC generators and also in the proposed electrical power systems in automobiles. The paper also derived the designing process of converter to choose semiconductor components for the desired system efficiency. Expected results are verified by the simulation. By the experiments it is found that the VIENNA rectifier is controlled as dual boost rectifier which is suitable for applications that require power factor corrections and simultaneously operate from a wide input voltage range to get a constant DC bus voltage.**

***Keywords—DC, power factor correction, rectifier, three phase generator, variable speed, VIENNA***

## I. INTRODUCTION

In most electronic applications different power converters play a cardinal rule. Generally in low power devices diode bridge rectifier is used. But this kind of rectifier takes non-sinusoidal current from the source which creates stress on the components. Moreover, the voltage and current wave doesn't remain in the same phase and power factor become very bad condition and power loss increases. In high and medium power application it is essential to maintain the input current shape sinusoidal also keep the power factor in tolerance level. Single phase input is enough in most of the medium power conversion process. In the process the unregulated source input is improved higher than the rectified line voltage by suing the single switch non isolated boost topology. For considering the power factor the switching is done in a way that the current from the source is in phase with the source voltage. For this there exist a zero phase angle between voltage and current. For higher power output the component size and the component stress for non sinusoidal current should kept very less. For this reason three phase input and three phase rectifier is used in most cases. This rectifier is based on the principle of single phase active rectifier [1]. It takes sinusoidal current in all its three phase. VIENNA rectifier is also similar of that concept which has three phase, three switch and single quadrant. By that three switch the input current waveform is controlled where the output can get as DC. Some variable speed energy sources like wind generator gives variable voltage and frequency output as speed and frequency has direct relation to each other. So in this case VIENNA rectifier is the best and smart solution for the maximum efficient output to connect variable frequency and variable voltage with fixed frequency and fixed voltage [11]. It can also used in electric vehicle to convert variable frequency and voltage to a fixed frequency and voltage with maximum output [3]. The main goal of the paper is to establish an interface system to get DC voltage with less reactive power consumption from the variable input condition like wind generator or micro hydro generator. This relation is based on the VIENNA rectifier topology where input voltage and current is also sinusoidal in phase so that the power factor remain less and gives the efficient maximum output.

## II. VIENNA RECTIFIER

VIENNA rectifier is a three-switch rectifier which gives the DC output. For controlling the rectifier it only need three switches rather five floating switches of other rectifiers. This switching system gives it more convenience to implement. But in comparing this with single switch controlling is still complex. But the input current distortion of this rectifier is about 8.2% which is significantly less than the single-switching and also than H-bridge and two switch implementation [2]. The main drawback of the VIENNA rectifier is the high boost ratio which needs high output voltage. A Vienna rectifier circuit diagram has shown in fig 1. Basically the VIENNA rectifier works as a two-switch boost rectifier. One of the switches works at line frequency and two switches switched at high frequency. In 60° control block one switch is permanently on. Actually VIENNA rectifier can be explained as two independent boost rectifiers, one for boosting C1 and the other for boosting C2. Thus it can be seen that the minimum boost voltage over C1 and C2 will be the maximum line-to-line voltage of the input.

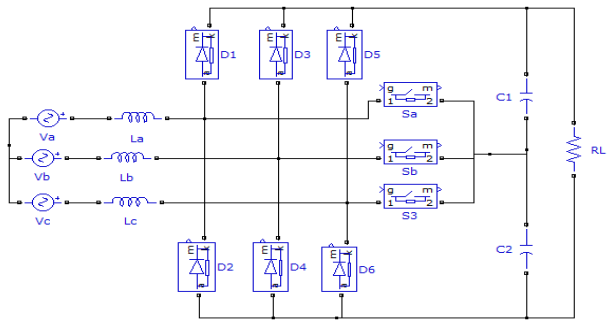



Figure 1. VIENNA Rectifier circuit diagram

The equivalent representation of VIENNA rectifier for a 60° control block in one switch "on" condition is shown in the fig 2. VIENNA rectifier has lower switch and diode currents than all of the other dual-boost rectifiers [5]. For switch losses and diode losses its has the same harmonic distortion like others. But an extra convenience of the VIENNA rectifier is that modules are available having all the power stage bridge leg.

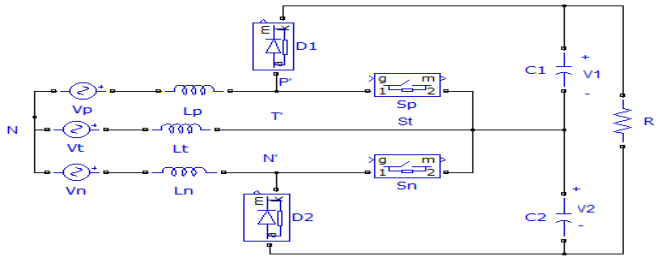



Figure 2. Equivalent circuit of VIENNA rectifier ($60^0$ control block)

Some control algorithm of constant frequency control is shown in the table 1. P, n and t-subscript denotes parts associated with the positive rail, parts associated with the negative rail and parts associated with a transitional period respectively. The small-signal model is derived for a negative duty cycle which is larger than the positive duty cycle, $d_N > d_P$. Then the process is repeated for the positive duty cycle which is larger than the negative duty cycle, $d_P > d_N$. The common parts of two switch cycle off periods ($1-d_N$) and ($1-d_P$) can be associated [6]. The onset of the switching period $T_{SW}$ is represented by the time $t_0$. $V_t$ is denoted as AC source which can be either negative or positive.

TABLE I. VIENNA RECTIFIER CONTROL ALGORITHM

| $\alpha = \omega_L t$ | p | T | N |
|---|---|---|---|
| $-30^0$ to $30^0$ | c | a | b |
| $30^0$ to $90^0$ | a | c | b |
| $90^0$ to $150^0$ | a | b | c |
| $150^0$ to $210^0$ | b | a | c |
| $210^0$ to $270^0$ | b | c | a |
| $270^0$ to $330^0$ | c | b | a |

## III. VIENNA RECTIFIER DESIGN

### A. *Input Inductor Design for Filter*

Input inductor selection is an important part for this rectifier design. In inductor analysis it is assumed that Dp > Dn. It means that the positive side inductance (Lp) is charged for the period $D_P T_{SW}$ and discharged for the period $(1-D_P)T_{SW}$ [7]. The waveform for the positive side current is shown in the following figure 3. The waveform is illustrated by considering the time-domain behavior of the VIENNA rectifier.

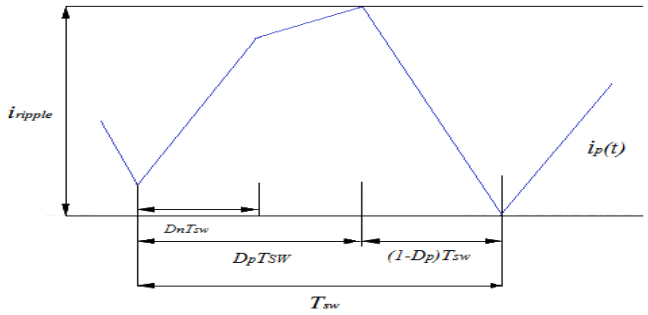



Figure 3. Wave form of the current passing through $L_p$

From the figure 3 it can be seen that during the off cycle ($(1-d_p).T_{SW}$) the current ripple is the current drop. Current waveform can be described mathematically during off cycle by taking the voltage ripple small [6].

$$i_p(t) = i_{p0} + \frac{1}{L}\int_{D_pT_{SW}}^{T_{SW}} V_{differential}\,dt$$

$$= i_{p0} + \frac{1}{L}\int_{D_pT_{SW}}^{T_{SW}} -(V_P(t) - E)dt = i_{p0} - \frac{(V_p(t) - E)(1 - D_p)T_s}{L}$$

Here, the $V_{differential}$ is the voltage over the inductor and $i_{p0}$ is the starting current. From the above equation AC current ripple can be written as.

$$i_{ripple}(t) = -\frac{(V_p(t) - E)(1 - D_p)T_{sw}}{L_p}$$

Positive rail ($V_p$) is shown mathematically below [7].

$$V_p = V_{phase,peak}Cos(k\omega t - 60^0);\ 0^0 \le \omega t < 120^0$$

Here, $V_{phase,\ peak}$ is represented as input source peak to neural voltage and k is an integer. The negative side voltage ($V_n$) is

$$V_n = -V_{phase,peak}Cos(n\omega t + 0^0);\ -60^0 \le \omega t < 60^0$$

Assuming the $60^0$ control block and substituting the value K=1 and n=1 in the previous equation

$$(1 - D_p) = \frac{2V_{phase,peak}Cos(\omega t - 60) - V_{phase,peak}Cos\ (\omega t)}{E}$$

substituting this equation in the previous it can be found

$$= \frac{d}{dt}\left(\frac{2V_{phase,peak}Cos(\omega t - 60) - V_{phase,peak}Cos(\omega t)}{E}\right)\frac{(V_{phase,peak}\ Cos(\omega t - 60) - E)}{L}$$

The above equation has maximum current ripple will coincide in $\omega t=30^0$ . $V_p=0.866V_{phase,\ peak}$ for all K and $V_n=-0.866V_{phase,\ peak}$ for all n. So the maximum current ripple can be found

$$i_{ripple,max} = \frac{0.866V_{phase,peak}}{E}\frac{(0.866V_{phase,peak} - E)T_{sw}}{L}$$

So it can be found that when the transitional voltage (Vt) and the transitional current (It) will be zero then the maximum ripple current will occur. The required inductance can be calculated at 0.866 of the peak phase voltage.

### *B. Output Capacitor Design*

Alike the inductor selection capacitor selection is also a significant task. For the calculation of the output capacitive filter the period is assumed as ωt = [-30° , 0° ]. The output capacitors is assumed to be very small so that the line frequency ripple is dominant [8]. It also assumed as the identical capacitors C1 and C2 whose capacitance equal to C. If the period is ω t = [-30° , 0° ] and Dp < Dn then |In| < |Ip|. Considering the constant output current the capacitors discharge will be

$$V_{C2,discharge} = V_{C1,discharge} = \frac{1}{C}\int_0^{T_L\frac{30}{360}} |I_{OUT}|dt = \frac{30}{360}I_{OUT}T_L$$

Here, C is the output capacitance, $I_{OUT}$ is the output current and $T_L$ is the period of the source. It is considered that the source current waveforms are sinusoidal. Within the period ω t = [-30° , 0° ] capacitor C2 will be charged by the negative side input inductor current and at 0° capacitor C2 will be charged by [7]:

$$V_{C2,charge} = \frac{1}{C}\left|\int_{t_0+D_nT_{SC}}^{t_0+T_{SC}} i_n(t)dt + \int_{t_1+D_nT_{SC}}^{t_2+T_{SC}} i_n(t)dt + \int_{t_2+D_nT_{SC}}^{t_2+T_{SC}} i_n(t)dt + K\right|_{\substack{\omega t = 0^0 \\ \omega t = 30^0}}$$

For the period ωt=[-30$^0$,0$^o$] the negative side current can be analised as

$$i_n(t) = -i_{phase,peak}\cos\left(\frac{2\pi}{T_L}\left(t - \frac{60T_L}{360}\right)\right)$$

Peak to neutral current can be written as:

$$i_{phase,peak} = \sqrt{2}\frac{P_{OUT}}{3V_{phase,rms}}$$

By analyzing the equations the value of C can be found as

$$C = \frac{2\left|I_{OUT}T_L\frac{30}{360} - \left|i_{phase,peak}\frac{T_L}{2\pi}\sum_{x=0}^{y}\left[\frac{\sin\left(\frac{2\pi T_{SW}}{T_L}(x+1) - \frac{120\pi}{360}\right) - }{\sin\left(\frac{2\pi T_{SW}}{T_L}x + \frac{2\pi D_n(x)T_{SW}}{T_L} - \frac{120\pi}{360}\right)}\right]\right|\right|}{V_{ripple,p-p}}$$

## IV. VARIABLE AC TO DC INTERFACING DESIGN WITH VIENNA RECTIFIER

In the experiment a prototype rectifier of 1KW is considered having line to line input voltage VLL = 220V and 50Hz frequency. For generator input the rectifier will be able to supply up to 500W output power at 110V input voltage, linearly from the 220V input voltage and 1KW output power. For non-generator inputs, the rectifier will be able to supply 220V having ±20% variation in input that means it will be able to supply 1kW with input from 176V up to 264V. The peak line to neural voltage with 20% variation is:

$$V_{phase,peak} = \sqrt{2}\left(\frac{V_{LL,max,rms}}{\sqrt{3}}\right)$$

$$= \sqrt{2}\left(\frac{220+220\times0.2}{\sqrt{3}}\right) = 215.55\text{V}$$

So the minimum output voltage can be found $V_{out,\,min} = 3V_{phase,\,peak}$ = 646.65V. For convenience the output voltage is chosen as 700V. The minimum voltage that can be boosted by either bank is :

$$V_{1,min} = \frac{(2V_p + V_n)}{(1 - D_p)} = \frac{(2V_{phase,peak} - 0.5V_{phase,peak})}{(1 - 0)} = 323.325V$$

The allowable ripple voltage ripple for V1 and V2 is equal to (700–2(323.325)) = 53.35V. Input diodes will work in forward mode rather freewheeling if the voltage ripple exceeds 53.35V. The rms phase input current is equal to.

$$i_{phase,rms} = \frac{P_{OUT}}{3\left(\frac{V_{LL}}{\sqrt{3}}\right)}$$

When the voltage is at a minimum, or equal to $V_{LL}$ = 176V then the input current will be maximum. Putting the value of $V_{LL}$ = 176 it can be found:

$$i_{phase,rms} = \frac{1000}{3 \times 101.61} = 3.28A$$

Peak average phase input current is equal to:

$$I_{phase,\,peak} = \sqrt{2}\, i_{phase,\,rms} = \sqrt{2}\times3.28 = 4.64A$$

For rated power 1000W when $V_{OUT}$ = 700V, then the minimum input voltage is

$V_{Phase,\,peak} = 2V_{phase,\,rms} = 2\times101.61 = 143.7$V

From the above equation it follows that the peak to peak current ripple is equal to $i_{ripple,\,max}$ = 0.1×4.64=0.464 A. The switching frequency is arbitrary as 50kHz or 25 KHz. So the period of one switching cycle will be 20us. Now the desired inductance can be get from the equation as L≈3.15mH. So the maximum output current of the rectifier will be: $I_{OUT} = P_{OUT}/V_{OUT}$=1000/700=1.429 and the capacitance can be found as C≥46.842 uF. The mid-point capacitor voltage ripple is proportional to the discharge current minus the charge current [9]. So the capacitor mid-point voltage ripple decreases as the discharge current decreases with the lower input voltage. So the voltage ripple decrease as the input voltage decreases lower than the line to line voltage when the generator connected at the output. In both positive and negative duty cycle the output is symmetrical. So it's better to perform the plant analysis with either Dn or Dp set equal to 1 [10][11]. If the control of the rectifier is done by digitally then the reference voltage should choose half of the operating voltage. The equivalent sense resistance can be choose as:

$$R_s \leq \frac{V_{ref}}{i_{phase,peak} + \frac{i_{ripple,max}}{2}}$$

$$\leq \frac{1.65}{4.64 + \frac{0.464}{2}}$$

$$\leq 0.34\ ohm$$

When analogue controller is used then a standard band gap reference can be used with a reference voltage of 2.5V. If input

current sensors are used, the output voltage should be scaled as (1.65/2) = 0.825V and current as 4.64 + (0.464/2) = 4.872A.

TABLE II. LOW FREQUENCY PARAMETERS OF INTERFACING DESIGNED BY VIENNA RECTIFIER

| Parameters | Output Power | |
|---|---|---|
| | 1KW | 500W |
| $V_{LL}$(rms) | 176V | 110V |
| $V_{p,\ peak}$ | 143.7V | 89.81V |
| $V_{n,\ peak}$ | -71.85V | -44.91V |
| $D_P$ | 0.384 | 0.615 |
| $D_n$ | 1 | 1 |
| L | 3.15mH | 3.15mH |
| C | ≥46.842 uF | ≥22.519 uF |
| $R_L$ | 490ohm | 980ohm |
| $I_P$ | 4.64A | 3.711A |
| $I_n$ | -2.32A | -1.856A |
| $V_1$ | 350V | 350V |
| $V_2$ | 350V | 350V |
| $V_{ref}$ | 2.5V | 2.5V |
| $R_s$ | 0.17ohm | 0.17ohm |
| $V_m$ | 1.92V | 2.456V |
| $V_{OUT}$ | 700V | 700V |

In the experiment three output parallel capacitance is chosen as 22µF. To change the performance the capacitance value can be changed. The capacitors are places in parallel to minimize the ripple current [5]. According to the parameters of the above table bode diagram for the uncompensated open loop transfer function is depicted.

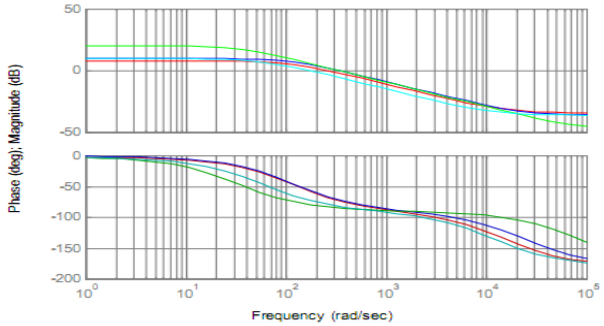



Figure 4. For different load and input voltage uncompensated open-loop transfer diagram

From the above figure it is seen that the gain does not change for different input voltage and power condition except a little drop at 110V. In the frequency of 1 krad/s there will be minimal phase shift for different power levels.

## V. VIENNA RECTIFIER CONTROLLER DESIGN

If the compensated open-loop gain of the system is less than 0dB when the compensated open-loop phase angle is -180° then a system will be stable [7]. Moreover, the compensated open-loop gain roll-off must be -20dB/decade at 0dB to stable the system. Integration control is necessary to ensure the steady-state zero output error [9]. The controller designed in the paper is a lag-lead type compensator which is shown in the following figure. The same compensator is proposed as single pole, single zero type compensator with a DC pole to force the steady-state error to be zero.

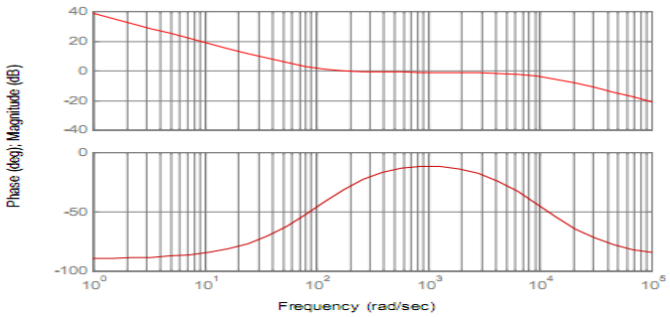



Figure 5. Type II compensator transfer function bode-plot.

This type of compensator provides 0dB/decade slope change at the desired crossover frequency. It also provides the necessary phase-boost to ensure a positive phase margin. For the stability 45º to 60º phase margin is necessary. From the figure it is seen that for a suitable frequency range where compensation can be done is ω = 200 rad/s to ω = 10 krad/s. The gain roll-off for the uncompensated open-loop transfer is -20dB/decade for this frequency range. So 0dB/decade compensation will be stable. The only main disadvantage in the range range ω = [2; 10] krad/s where the phase margin will change if the load change [10]. In the crossover at 1 krad/s the phase shift will be minimal over the full power range. For the controller design the phase margin is chosen to be 60º as this will give optimum transient performance. If the performance of the converter is bad then the phase margin can be decreased, but it should not be less than 45°. If less this level then it will give more overshoot. For this converter the crossover frequency, where the compensated gain plot should be 0dB, is chosen to be ωc = 1 krad/s. The required compensated phase boost for PM=$60^0$ and 5% overshoot at the output is

$$\text{Boost} = -(\angle T_{ol}(s)|_{s\to j\omega c} + 90) + \text{PM} = -(-87.5+90)+60 = 57.5^0$$

The system can be compensated by this type of compensator as the boost requirement is lower than 90°. If the boost requirement goes higher than 90° then the crossover frequency should be decreased. This will result a lower controller bandwidth and hence slower response to step inputs. From the phase plot of the tangent curve and the value of parameter K can be written as:

$$K = \tan\left(45^\circ + \frac{boost}{2}\right) = \tan\left(45^\circ + \frac{57.5^\circ}{2}\right) = 3.431$$

Where, $\omega_c$ =291.5 rad/s and $\omega_p$=3420.8 rad/s and the compensator gain at DC can be written as:

$$|T_c(s)|_{s\to dc} = -|T_{ol}(s)|_{s\to j\omega c} + 20\log_{10}(\omega_c) = 10.5\text{dB} + 49.29\text{dB} = 59.79\text{dB} = 976.42$$

The gain become

$$A_G = |T_c(s)|_{s\to j\omega c}\frac{\omega_p}{\omega_c} = 976.42 \times \frac{3430.8}{291.5} = 11492.00$$

The compensator transfer function is

$$T_c(s) = 11492\frac{(S+291.5)}{S(S+3430.8)}.$$

From the following diagram it can be shown that the phase boost is 57.5° and the gain 10.5dB at ωc. The bode diagram for the compensated open-loop gain and phase response is shown in the following diagram. From the MATLAB analysis of the following figure it can be analyzed that seen that the phase margin is 59.9° and the crossover frequency is ωc = 979.7 rad/s. It is also analyzed that the roll-off of the gain plot is 20dB/decade at 0dB and that the gain margin is positive.

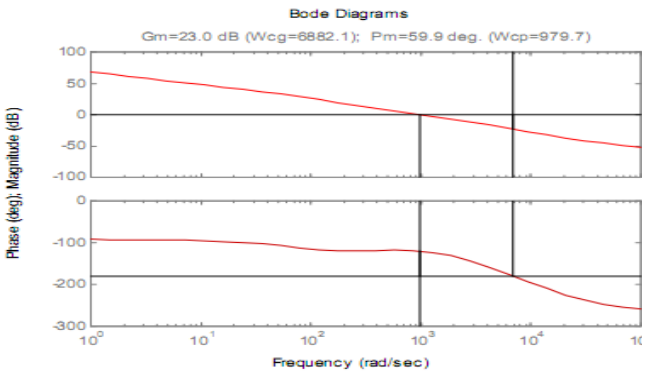



Figure 6. VIENNA Rectifier open-loop compensated system gain and phase plot.

Compensator of the interfacing can be implemented by the following circuit having operational amplifier.

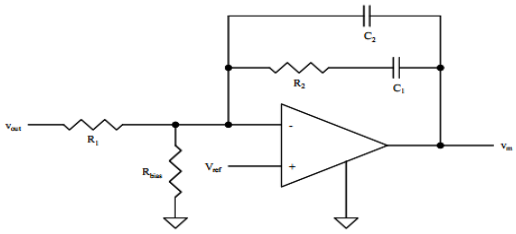



Figure 7. Interfacing compensator circuit diagram by operational amplifier

If the compensator resistance $R_1$ value is considered as the 10 K ohm then the value of $C_1$, $C_2$ and $R_2$ can be found as following.

$$C_2 = \frac{|T_{OL}(s)|_{s\to j\omega c}}{KR_{1\omega c}} = \frac{10^{\frac{-10.5}{20}}}{3.431\times10000\times1000} = 8.70nF \approx 8.2nF$$

And

$$C_1 = C_2(K^2 - 1) = 8.7\times10^{-9}(3.431^2 - 1) = 93.7nF \approx 100nF$$

$$R_2 = \frac{K}{C_{1\omega c}} = \frac{3.431}{(93.7\times10^{-9}\times1000)} = 36.606k\Omega \approx 36.5k\Omega$$

## VI. Experimental Result and Interfacing

The interface circuit diagram between the variable AC and the DC Bus has been depicted in the following figure 8.

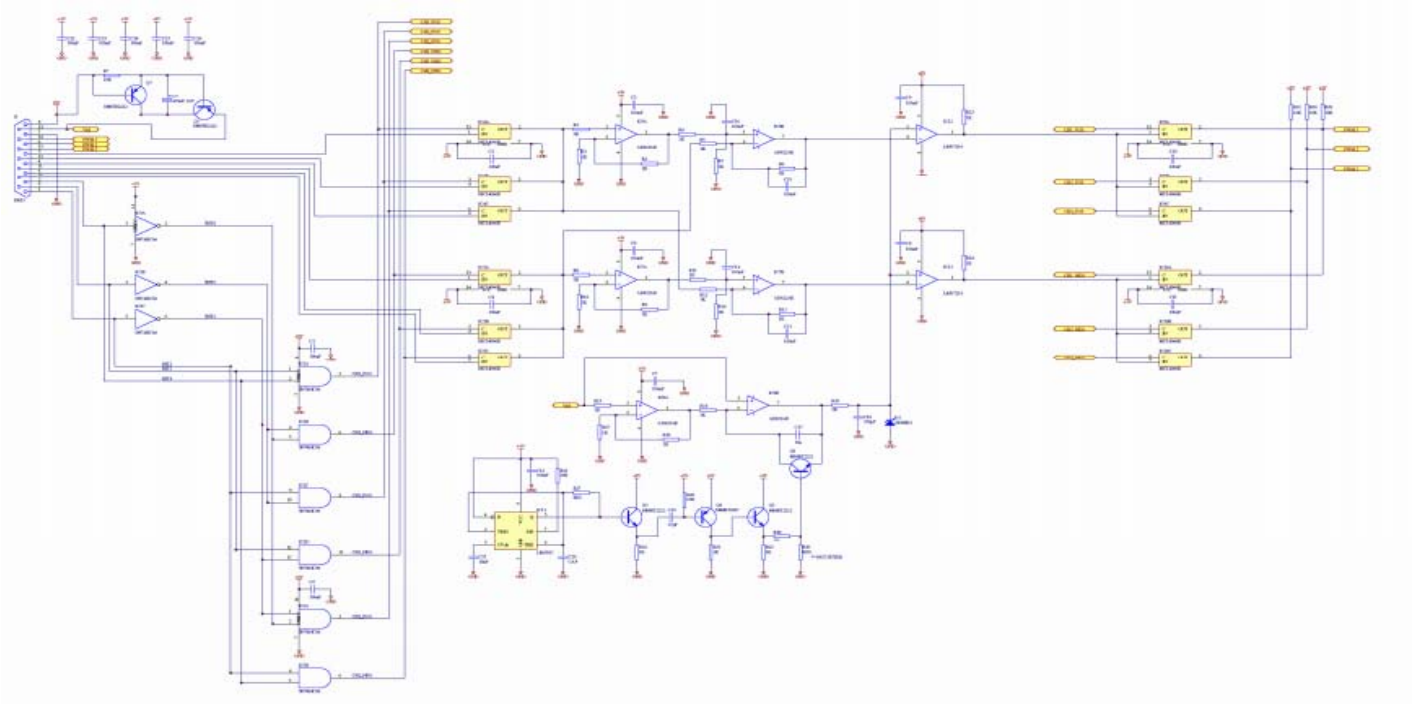


Figure 8. Interfacing diagram

The experimental output for the interfacing between variable input to fixed output is depicted below. Different input has been chosen to observe the output value and the performance and effectiveness of the rectifier. In the paper the visual waveform for the output power of 200W and line to line input voltage of VLL = 176V is shown.

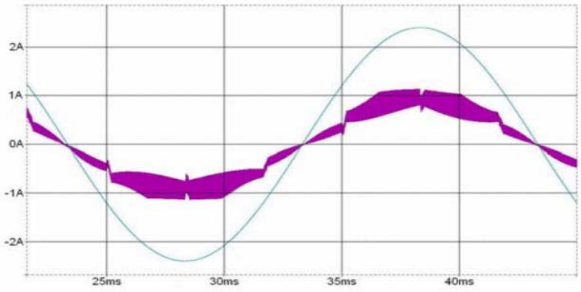



Figure 9. Input voltage (cyan) and input current (violet) wave form having the same phase. With a load of 200W and 176V input.

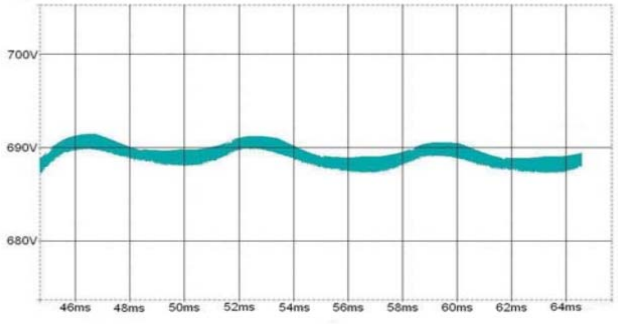



Figure 10. Output voltage waveform with a load of 200W and 176V input.

## VII. CONCLUSION

The main concentration in the paper is to develop an interface between variable speed three-phase generators and a DC-bus. It has been explored by the experiment that the VIENNA rectifier is more suitable to convert a wide input voltage source to a DC output than a 6-switch topology. VIENNA rectifier is also suitable for boosting very low input voltages and still maintain a constant DC-bus voltage. While implement the interface snubbers should add for improving the overall efficiency and noise performance. For better efficiency and noise performance soft-switching techniques can be introduced.

## REFERENCES

[1] Johann W. Kolar, Uwe Drofenik, Franz C. Zach, "VIENNA Rectifier II—A Novel Single-Stage High-Frequency Isolated Three-Phase PWM Rectifier System", IEEE transactions on industrial electronics, vol. 46, no.4, August 1999

[2] Johann W. Kolm, Hans Ertl, "Status of the Techniques of Three-phase Rectifier Systems with Low Effects on the Mains", 21st INTELEC, June 6-9, 1999, Copenhagen, Denmark

[3] Uwe Drofenik, Johann W. Kolar, "Comparison of Not Synchronized Sawtooth Carrier and Synchronized Triangular Carrier Phase Current Control for the VIENNA Rectifier I", IEEE Transactions, 1999

[4] Chongming Qiao, Keyue Ma Smedley, "A General Three-Phase PFC Controller for Rectifiers With a Series-Connected Dual-Boost Topology", IEEE transactions on industry applications, vol.38, no.1, 2002

[5] Chongming Qiao, Keyue M. Smedley, "A General Three-Phase PFC Controller for Rectifiers With a Parallel-Connected Dual Boost Topology", IEEE transactions on power electronics, vol.17, no.6, November 2002.

[6] Chongming Qiao, Keyue M. Smedley, "Three-Phase Unity-Power-Factor Star-Connected Switch (VIENNA) Rectifier With Unified Constant-Frequency Integration Control", IEEE transactions on power electronics, vol.18, no.4, July 2003.

[7] Johann Miniböck, Johann W. Kolar, "Novel Concept for Mains Voltage Proportional Input Current Shaping of a VIENNA Rectifier Eliminating Controller Multipliers", IEEE transactions on industrial electronics, vol.52, no.1, February 2005.

[8] Ralph Teichmann, Mariusz Malinowski, "Evaluation of Three-Level Rectifiers for Low-Voltage Utility Applications", IEEE transactions on industrial electronics, vol.52, no.2, April 2005.

[9] Nesrine Bel Haj Youssef, Kamal Al-Haddad, Hadi Y. Kanaan, "Large-Signal Modeling and Steady-State Analysis of a 1.5-kW Three-Phase/Switch/Level(Vienna) Rectifier With Experimental Validation", IEEE transactions on industrial electronics, vol.55, no.3, March 2008.

[10] Michael Hartmann, Simon D. Round, Hans Ertl, Johann W. Kolar, "Digital Current Controller for a 1MHz, 10kW Three-Phase VIENNA Rectifier", IEEE transactions on power electronics, vol.24, no.11, November 2009.

[11] HaoChen, DionysiosC.Aliprantis, "Analysis of Squirrel-Cage Induction Generator With Vienna Rectifier for Wind Energy Conversion System", IEEE Transactions On Energy Conversion, Vol. 26, No.3, September, 2011

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Design and Performance Analysis of Ultra Wideband Double Inverted-FL Micro strip Antenna for Wi-Fi, WLAN, WiMAX and UMTS Applications

Farzana Khanam[1], Sathi Rani Mitra[2], Md. Asadur Rahman[3] and Md. Selim Hossain[4]
Department of Electrical Department of Electrical and Electronic Engineering,
Khulna University of Engineering and Technology,
Khulna-9203, Bangladesh.
Email: farzana0703047@gmail.com[1], sathi_mittro@yahoo.com[2], eeeasadur@yahoo.com[3] and selimkueteee@yahoo.com[4].

*Abstract*—**A survey of microstrip antenna is presented, with emphasis on theoretical and simulation design techniques. Among several simulation techniques, Numerical Electromagnetic Code-2 (4NEC2) software is used by means of analyzing performance results. The wide band low profile modified Double Inverted- FL (DIFL) antenna covers 2.4 GHz for Wi-Fi, 5.2 GHz serving for WLAN, 2.5 GHz for WiMAX and also 2.1 GHz for UMTS applications. The antenna provides a total bandwidth of 850 MHz (2.0 GHz~2.85 GHz) which fully covers the 2.1/2.4/2.5 GHz bands and also 450 MHz (4.9 GHz~5.35 GHz) which covers 5.2 GHz band. The compact antenna is suitable in mobile phones and laptop computer for its small size, 31×21 mm$^2$. The antenna maintains a return loss more than -10 dB throughout the bands. The measured results also demonstrate that the proposed antenna has good omnidirectional radiation patterns with appreciable gain across the operating bands and is capable to implant within different portable devices for WiMAX/WLAN applications.**

***Keywords—Double Inverted-FL antenna (DIFL antenna), WiMAX, WLAN, UMTS.***

## I. INTRODUCTION

An antenna is a device designed to transmit or receive electromagnetic waves [1]. There are different sizes of antennas for particular practical purposes. At present some practical antenna designing requirements like multiband operation and miniaturized size have attracted high attention. Several researchers have devoted large efforts to improve performance of antennas that satisfy the above demands [9], [10]. In this communication revolutionary decade, the demands of wireless local area networks (WLANs), Universal Mobile Telecommunications System (UMTS) and worldwide interoperability for microwave access (WiMAX) are increasing numerously worldwide for commercial communication. WLAN provides high speed connectivity and easy access to networks without wiring. Else WiMAX can provide a long operating range with a high data rate for mobile broadband wireless access [4]. The IEEE 802.11 Wi-Fi and WLAN standard allocates the license-free spectrum of 2.4 GHz (2.40-2.48 GHz), 5.2 GHz (5.15-5.35 GHz) and 5.8 GHz (5.725-5.825 GHz).WiMAX, based on the IEEE 802.16 standard, has been evaluated by companies for last mile connectivity, which can reach a theoretical up to 30 mile radius coverage [6]. Now a days, state of the art antenna for mobile phone has covered frequency bands including UMTS2100 (1920~2170 MHz), UMTS2600 (2500~2690MHz), and Mobile-WiMAX [5].

A microstrip antenna is a low cost tiny size antenna used to process ultra-high frequency signals. The microstrip antenna like inverted-F antenna is widely used as a built-in antenna for mobile phones [2], [3]. Else, the inverted-L element operates on a frequency band of 824~ 960MHz [5].

To congregate most of the mentioned requirements, the proposed Double Inverted-FL antenna is one of the high-quality candidates within the micro-strip printed antennas because of its compact size and good input impedance than other printed antennas.

## II. ANTENNA DESIGN

In this design, the proposed ultra wide band Double Inverted –FL antenna provides the possibility of increasing antenna bandwidth. With simplifying its structure, prolongation of the input impedance near about 50 Ω, apposite gain all the way through the application bands have been examined. Different parameter studies had been conducted to ascertain the effect of different loading on the antenna performance to find out the optimal design where, optimum segmentation of each geometrical parameter is used.

Fig. 1 represents the basic geometry of the IFA [4]. Here one leg of IFA is directly connected to the feed and another leg is spaced from the ground plane. IFAs are commonly used in mobile phones due to their diminutive size (quarter-wavelength). An example of several IFAs in a mobile phone can be represented as Fig. 2. These antennas are visible once the back cover is removed [12].

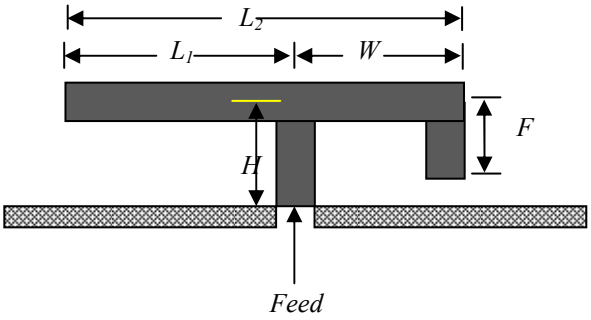



Figure 1: Structure of Inverted-F antenna (IFA).

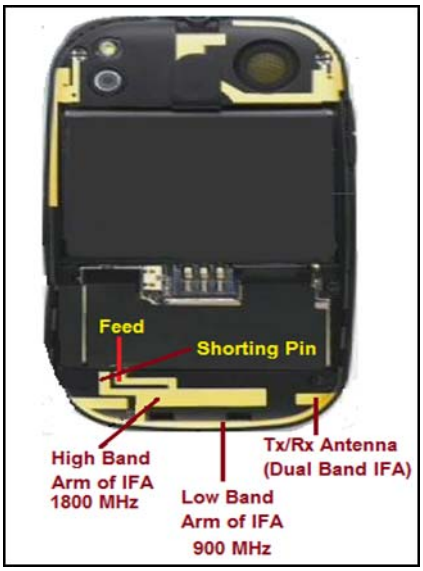



Figure 2: IF Antennas viewable by removing back cover of a cellular phone.

For the simulation, we consider portable circuit board (PCB) with permittivity of $\varepsilon_r$ = 2.2 and substrate thickness of 1.58 mm. The antenna is assumed to feed by 50 Ω coaxial connector, with its central conductor connected to the feeding point. Its outer conductor is connected to the ground plane just across the feeding point by using method of moments (MoM's) in Numerical Electromagnetic Code (NEC) [7]. In the analysis, the dimensions of the ground plane are considered as 60 mm × 60 mm.

For IFA of figure 1, the resonant frequency is proportional to the effective length of current distribution. There are two cases in which it is easy to formulate an expression of the resonant frequency with respect to the size of the IFA. The first case is when the width of the short-circuit plate W is equal to the length of the planar element, say $L_1$. From figure 1, the effective length of the antenna is +H where H is the height of the short-circuit plate. The resonance condition then is expressed by [8],

$$L_2 + H = \frac{\lambda_0}{4} \qquad (1)$$

Where, $\lambda_0$ is the wavelength. Resonant frequency associated with $W=L_1$ calculated from (1)

$$f_1 = \frac{c}{4(L_2 + H)} \qquad (2)$$

Where, c is the speed of light. The other case is for $W$=0. The effective length of the current is then $L_1+L_2+H$. For this case, the resonance condition is expressed by

$$L_1 + L_2 + H = \frac{\lambda_0}{4} \qquad (3)$$

The other resonant frequency that is part of the linear combination is associated with the case $0<W<L_1$ and is expressed as

$$f_2 = \frac{C}{4(L_1 + L_2 + H - W)} \qquad (4)$$

For the case, when $0<W/L_1<1$, the resonant frequency, $f_r$ is a linear combination of the resonant frequencies associated with the limiting cases. The resonant frequency $f_r$ is found using the experiment for $f_1$ and $f_2$ above in the following [9].

$$f_r = r.f_1 + (1 - r)f_2 \qquad (5)$$

Where, $r= W/L_1$. With the help of resonant frequency theory of IFA and impedance matching concept, we considered the dimension of the proposed antenna. The basic structure of proposed antenna is presented in Fig. 3. The low profile antenna is connected to the feed by the arm K and the arm K is connected an L shaped region by another arm C. The arm K supports the whole antenna. The arm is spaced from the ground plane. Fig. 4 represents the 3-D view of the proposed antenna.

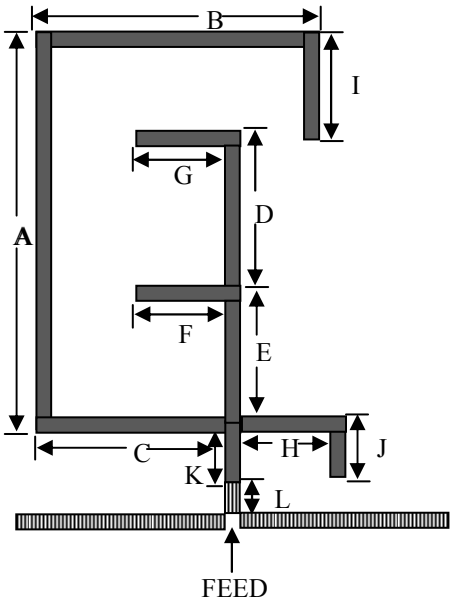



Figure 3: Structure of the DIFL microstrip antenna in 2-D view (proposed).

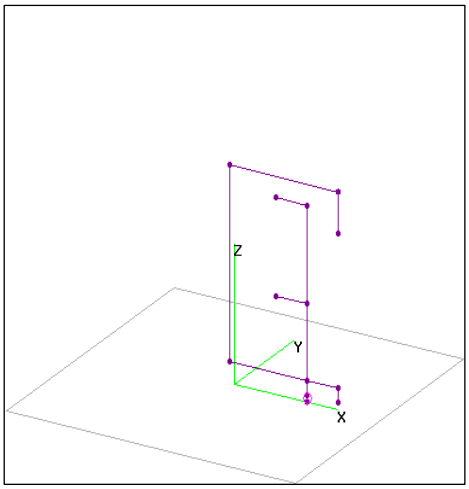



Figure 4: 3-D structure of the proposed modified Double Inverted-FL strip antenna.

The antenna arms are elected in such a way that the best results can be obtained.

The arm A has a great effect on the return loss and the bandwidth. After analyzing Fig. 5, the length of the antenna, A=28 mm is chosen with high return loss. In Fig. 6, B=23 mm cause high return loss than B=21 mm. But among them, B=21 mm is chosen because, this length compatibles with other parameters.

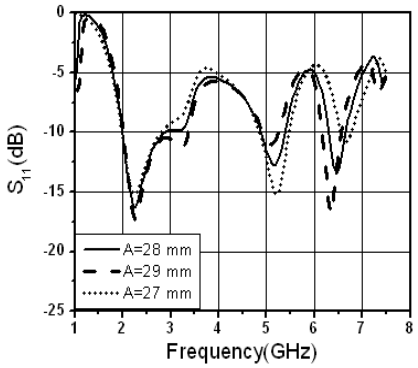



Figure 5: Variation of return loss at different values of A.

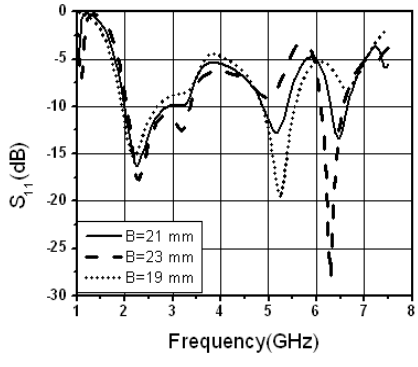



Figure 6: Variation of return loss different values of B.

From Fig. 7, C=15 mm is chosen. Fig. 8 represents variation of return loss at different values of D. Here D=14 mm is chosen for better performance. According to Fig. 9, variation in E causes little difference. E= 11 mm is chosen. From Fig. 10 to Fig. 13 shows that for F, G, H, I=6 mm, the antenna has the bandwidth for expected operation with a high return loss. In Fig. 11, although G=8 mm shows better performance, but G=6 mm is selected by analyzing all conditions. The variation of length of H and I affects same for the return loss and bandwidth from 2.4 GHz to 5.2 GHz as shown in Fig. 12 and Fig. 13. In Fig. 14, J= 1 mm covers better return loss, but J=2 mm covers better bandwidth & adjusts considerably. In Fig. 3, we can see that the length of K is too small to be varied to achieve return loss.

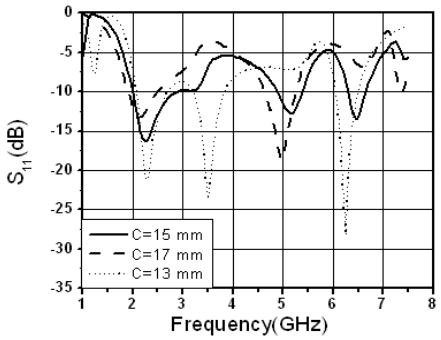



Figure 7: Variation of return loss at different values of C.

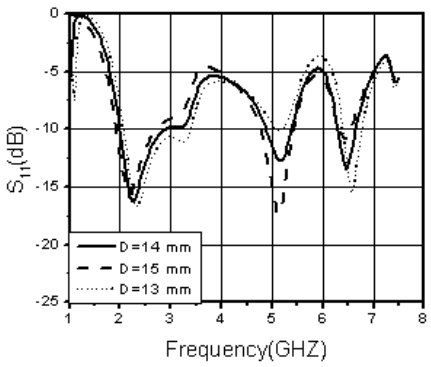



Figure 8: Variation of return loss at different values of D.

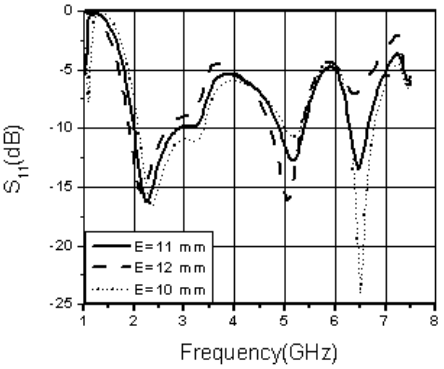



Figure 9: Variation of return loss at different values of E.

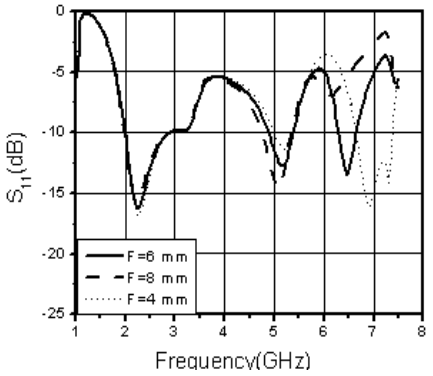



Figure 10: Variation of return loss at different values of F.

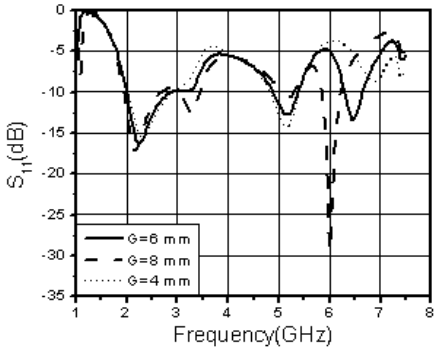



Figure 11: Variation of return loss at different values of G.

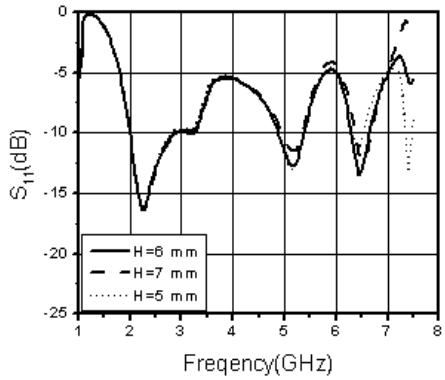



Figure 12: Variation of return loss at different values of H.

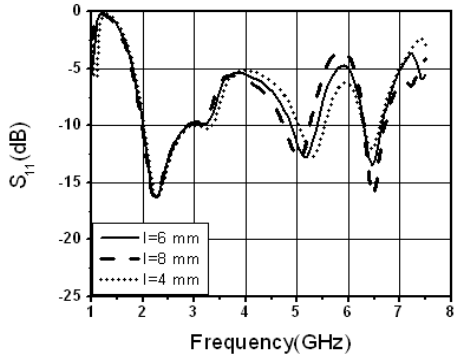



Figure 13: Variation of return loss at different values of I.

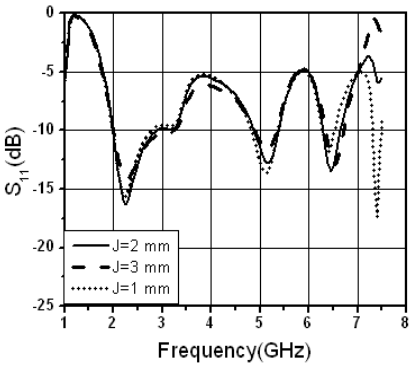



Figure 14: Variation of return loss at different values of J.

From overall analysis, we see that modified Double Inverted FL antenna provides best performance for the desired applications.

The optimized dimensions of the proposed DIFL strip antenna are listed in Table I.

Design Parameters of the Proposed Double Inverted-FL Antenna

TABLE I. DESIGN PARAMETERS OF THE PROPOSED DOUBLE INVERTED-FL ANTENNA

| **Antenna Name** | **Antenna Parameters** | **Values (mm)** | **Dimension ($mm^2$)** |
|---|---|---|---|
| Proposed DIFL Antenna | A | 28 | 31×21 |
| | B | 21 | |
| | C | 15 | |
| | D | 14 | |
| | E | 11 | |
| | F,G,H,I | 6 | |
| | J,K | 2 | |
| | L | 1 | |
| | d | 2 | |
| | A+K+L | 31 | |
| | C+H | 21 | |

## III. NUMERICAL SIMULATION RESULTS

The proposed low profile antenna is designed and numerically analyzed by using Numerical Electromagnetic code (NEC) [7] and their important characteristics are determined here. Return loss is the most important characteristics for the antenna shown in Fig. 15. For UMTS operation, the antenna has a high return loss -15.93 dB at 2.1 GHz. For WLAN and Wi-Fi operation, the antenna has a return loss -16.294 dB at 2.4, 2.5 GHz band and -12.793 dB at 5.2 GHz band. For WiMAX operation it has return loss -12.89 dB at 2.5 GHz band. In Fig. 16, the gain variation is shown. Here

3.53 dBi at 2.1 GHz band, 3.54 dBi at 2.5 GHz band, 6.2 dBi at 5.2 GHz band. The VSWR of the antenna is in between 1.36 to 1.90, which we can monitor from the next Fig. 17. In Fig. 18 we illustrate in 2.1 and 2.4 GHz band, the impedance is 40.498 Ω and 49.3801 Ω. But we get perfect impedance in 5.2 GHz band, which is 49.9353 Ω nearest 50 Ω for WiMAX operation.

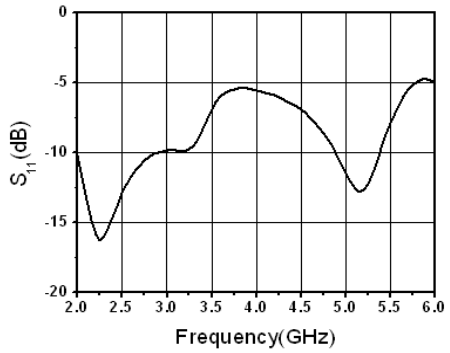



Figure 15: Variation of return loss with frequency.

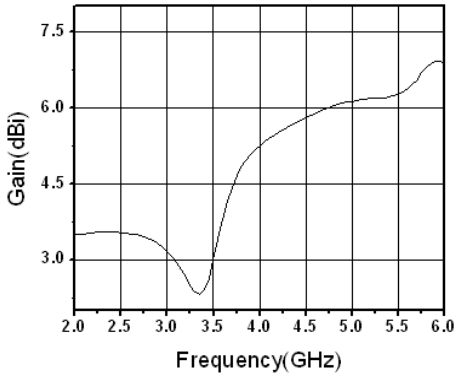



Figure 16: Variation of gain with frequency.

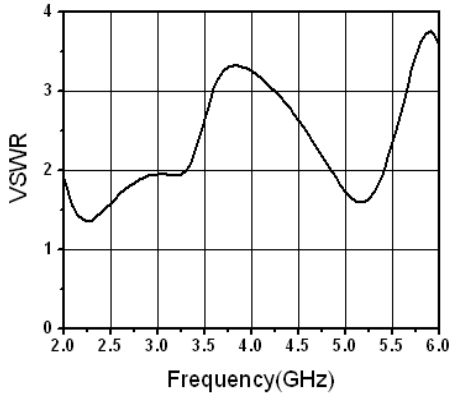



Figure 18: Variation of input impedance with frequency.

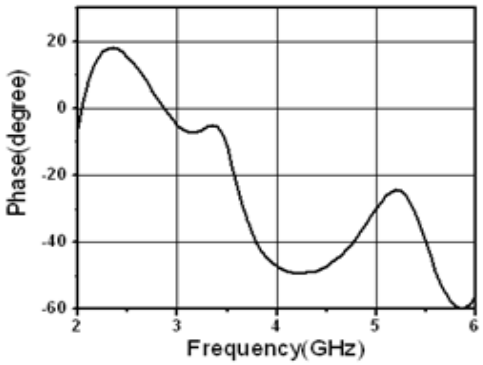



Figure 19: Variation of phase with frequency.

## IV. RADIATION PATTERN

Fig. 20 and Fig. 21 show the normalized radiation patterns of DIFL antenna at 2.1/2.4/2.5 and 5.2 GHz bands respectively. Normalized radiation patterns for four resonant frequencies are shown as: total gain in vertical (YZ/XZ plane) and horizontal plane (XY plane). The radiation pattern is omnidirectional in horizontal plane. In vertical plane it covers half cycle.

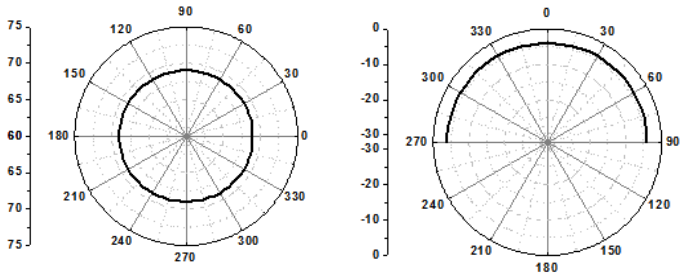


Figure 20: Radiation pattern (normalized) (a) Total gain in horizontal (XY) plane and (b) total gain in vertical (YZ/ XZ) plane of DIFL antenna at 2.1/2.4/2.5 GHz bands.

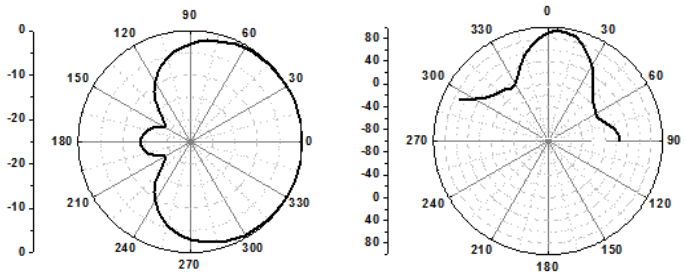


Figure 21: Radiation pattern (normalized) (a) Total gain in horizontal (XY) plane and (b) total gain in vertical (YZ/ XZ) plane of DIFL antenna at 5.2 GHz band.

TABLE II. COMAPARISON BETWEEN THE PROPOSED ANTENNA AND DIFFERENT REFFERENCE ANTENNAS

| Antenna Name | Peak Gain | Bandwidth |
|---|---|---|
| Proposed Ultra wideband Double Inverted-FL Antenna | 3.53 dBi at 2.1 GHz band<br>3.54 dBi at 2.5 GHz band<br>6.2 dBi at 5.2 GHz band | 850 MHz (2.0GHz~2.85 GHz) covers the 2.1/2.4/2.5 GHz bands and 450 MHz (4.9 GHz~5.35 GHz) which fully covers 5.2 GHz band. |
| Dual Band metamaterial –Inspired small monopole antenna [10] | 0.71 dBi at 2.4 GHz band<br>1.53 dBi at 5.2 GHz band. | 90 MHz (2420MHz~2510MHz)<br><br>320MHz (4520MHz~7720MHz) |
| A CPW –fed planar monopole antenna [11] | 2.15 dBi at 2.4 GHz band<br>2.47 dBi at 3.5 GHz band &<br>4.31 dBi at 5.5 GHz band | 300 MHz (2400~2700 MHz),<br>500 MHz (3300MHz~3380 MHz),<br>700 MHz (5150MHz~5850 MHz) |

## V. CONCLUSIONS

In this work, observing different antenna designs and their numerical analysis, the ultra wideband Double Inverted-FL low profile microstrip antenna has been proposed and analyzed all numerical simulations using NEC software and measured by standard techniques. We have analyzed antenna geometry by varying the four major geometry parameters (length, height, tap distance and spacing). The antenna geometry has been chosen and proposed the antenna operations depending on these four parameters analysis. We have focused to propose an antenna with larger bandwidth, good return loss characteristics, high gain, negligible phase shift, good radiation characteristics & unit VSWR. Our designed antenna occupies a small area of 31 $\times$21 $mm^2$ and has a larger bandwidth of 850 MHz (2.0 GHz~2.85 GHz) covers the 2.1/2.4/2.5 GHz bands and also 450 MHz (4.9 GHz~5.35 GHz) which covers 5.2 GHz band. For compact size, it is promising to be embedded within different mobile devices & laptop computer employing Wi-Fi, mobile WiMAX, WLAN & as well as UMTS applications. As compared with other proposed antennas by different researchers, our proposed antenna is smaller in size & provides larger bandwidth with a high gain.

## REFERENCES


[1] K.D. Prasad, Deepak Handa, *Antenna and Wave Propagation,* Satya Prakashan, 1996.

[2] R. King, C. W. Harrison , Jr, and D. H. Denton , Jr, "Transmission-Line missile antennas," *IRE Trans. Antennas Propag.*, pp. 88–90, Jan. 1960.

[3] R. J. F. Guertler, "Isotropic transmission-line antenna and its toroid pattern modification," *IEEE Trans. Antennas Propag.*, vol. TAP-25, no.3, pp. 386–392, May 1977.

[4] Farzana Khanam, Sathi Rani Mitra, Md. Selim Hossain and Debabrata Kumar Karmokar, "Multiband Low Profile Modified Inverted-FL Strip Antenna for 5.2/5.8 GHz WLAN and 5.5 GHz WiMAX Applications in Laptop Computer," *15th ICCIT 2012, University of Chittagong, 22-24 December 2012.*

[5] Nobuyasu Takemura, "Inverted-FL Antenna with Self-complementary Structure",*IEEE Transactions on Antenna And Propagation,*Vol.57,No.10,pp.3029-3034,October 2009.

[6] Sarawuth Chaimool and Prayoot Akkaraekthalin, "CPW-Fed Antennas for Wi-Fi and WiMAX", [Online]. Available:www.intechopen.com/download/pdf/pdfsid/26580.

[7] G.J. Burke and A. J. Poggio, "Numerical Electromagnetic Code-2," *Ver. 5.7.5, ArieVoors*, 1981.

[8] Minh-Chau T. Huynh, "A Numerical and Experimental Investigation of Planar Inverted-F Antennas for Wireless Communication Applications," *M.Sc. Thesis, Virginia Polytechnic Institute and State University*, October 2000.

[9] K. Hirisawa and M. Haneishi, "Analysis, Design, and Measurement of small and Low-Profile Antennas," *ArtechHouse, Boston*, 1992.

[10] J.Zhu and G.V Eleftheriades "Dual band material inspired small monopole antenna for Wi-Fi applications", Electronics Letters VOl.45 No 22 2009.

[11] Woo Chan Kim and Woon Geun Yang, "Design and Implementation of CPW-Fed Planar Monopole Antenna with Dual Band Rejection Characteristics", *University of Incheon, Republic of Korea*, 2010.

[12] Antenna theory tutorial webpage.[online] Available: http://www.antennatheory.com/antennas/aperture/ifa.php.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Comparative Study of Optical Cross Connect Architectures Regarding Wavelength Interchanging Capability

Md. Ishtiaque Aziz Zahed* and S. P. Majumder
Department of Electrical & Electronic Engineering, Bangladesh University of Engineering & Technology, Bangladesh.
* E-mail: ishtiaque2307@gmail.com

*Abstract* — **Wavelength conversion capability is desired at the network nodes in order to utilize the large bandwidth available with optical fiber and to use network resources efficiently. To provide wavelength conversion in an optical cross-connect node; the converters may be shared on share-per node or share-per link basis by the use of space switching matrix (SSM). In this paper the consequence of homodyne crosstalk in wavelength division multiplexing (WDM) optical network, where an optical signal is passed through optical cross-connect nodes (OXC's) has been examined. The study has been carried out for four different SSM (space switching matrix) OXC architectures namely WSXC, WIXC, Share-per-node (SSM-1) and Share-per-link (SSM-2). In the performance analysis the loss characteristics are measured through experimental measurements in the form of Bit Error Rate (BER). The BER is determined for three different cases regarding delay differences, no. of channels and in terms of relative intensity noise.**

***Keywords— Homodyne Crosstalk, Optical Cross Connect (OXC), Space Switching Matrix (SSM),Wavelength Interchanging Cross-Connect (WIXC), Wavelength Selective Cross-Connect (WSXC).***

## I. INTRODUCTION

Wavelength Division Multiplexing (WDM) optical network has turned out to be a field of widespread interest due to not only of its large bandwidth but also due to the increased capacity and flexibility. The high cost of wavelength converters has motivated the efficient design of optical networks that require only limited wavelength conversion capability. The optical cross-connect (OXC) at each node carries out wavelength sensitive switching in optical form without restoring to electro optical conversion. A number of OXC architectures can be designed who have their own features, strengths and limitations [1].

While cross-connecting wavelengths from input to output fibers OXC introduces inband and intraband crosstalk. The inband crosstalk which is also known as homodyne crosstalk has the same wavelength as the signal and degrades the transmission performance seriously [2]. When an optical signal passes through an OXC, many crosstalk contributions are combined with the signal. The optical propagation delay differences and polarization states of the crosstalk contributions are also in random and drift with respect to one another due to thermal and mechanical fluctuations in minutes on timescale [3]. Therefore, the transmission performance also varies from time to time.

In this paper, at the presence Homodyne crosstalk the performance of wavelength selective cross-connect (WSXC), wavelength interchanging cross-connect (WIXC), Share-per-node (SSM-1) and Share-per-link (SSM-2) is investigated in the form of Bit Error Rate (BER). Afterwards, with a specific Bit Error Rate (BER), the amount of power penalty as a function of the component parameters of the OXC is evaluated.

## II. OPTICAL CROSS CONNECT ARCHITECTURES

The OXC guides the light from one link to another and enables reconfigurable optical networks. OXCs need to be all optical and internally they can use either an electrical or an optical switch [1], [4].

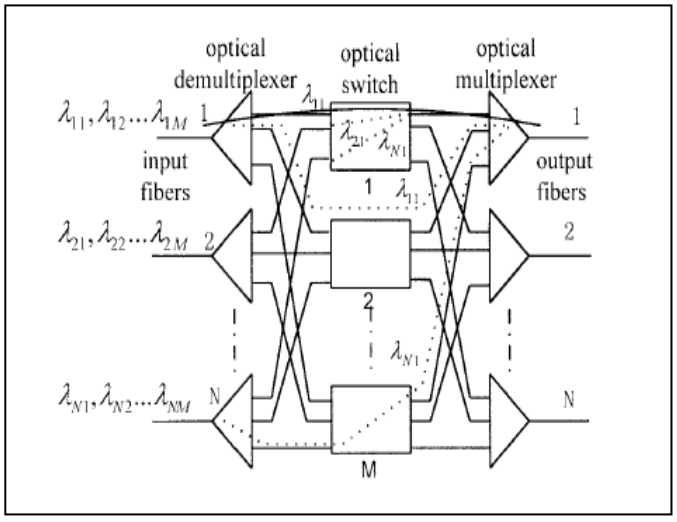



Figure 1. Typical Structure of an OXC

A typical OXC of the Fig. 1 uses N number of (M × 1) multiplexers and (1 × M) demultiplexers and M number of (N × N) optical switches. The demultiplexers of OXC separate M

wavelengths. The optical switch takes N number of same wavelength signal coming from all the N input fibers and connect each wavelength to any of the N output fibers according to the destination address [5].The multiplexer of OXC again combines M number of wavelengths and send them to a single fiber. The output demux separates the M wavelengths and send them to the individual user terminal.

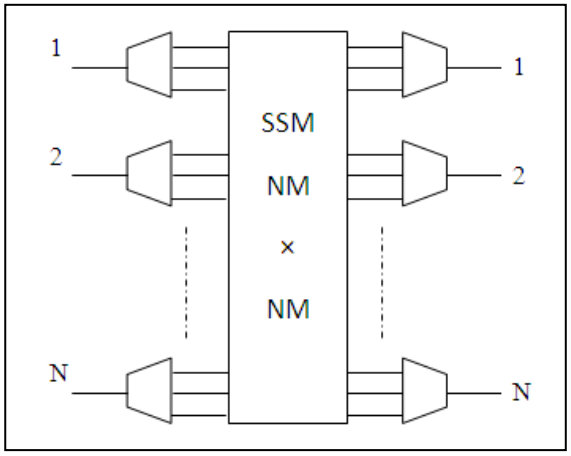



Figure 2. WSXC OXC Architecture

The OXC is a switch, which is generally controlled by the management layer while setting up or terminating light paths. It therefore requires less frequent and slower reconfiguration than if it was to be controlled by call traffic signaling. The lightpath, representing the optical layer connection between the source-destination node pairs, can be set up through the intermediate OXCs in either a wavelength continuous (WC or VWP, virtual wavelength path) or non wavelength-continuous (NWC or WP, wavelength path) fashion [4]. In the WC case, the same wavelength is used over the entire lightpath whereas, in the NWC case, different wavelengths may be used in different optical links along the given path.

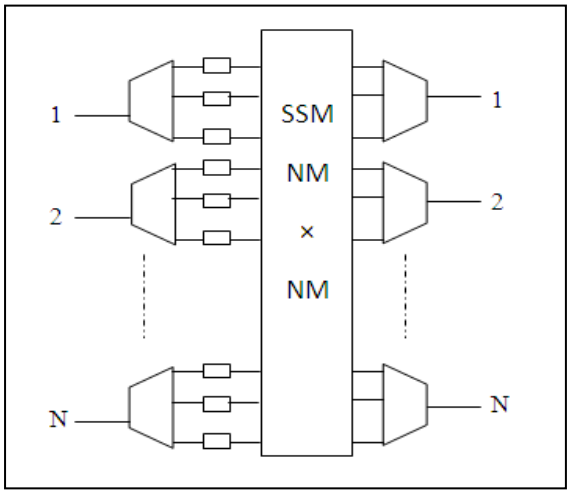



Figure 3. WIXC OXC Architecture

Setting up the lightpath would not only involve selecting the route to be followed but also the wavelengths to be used along the selected route. Wavelength conversion at the intermediate node is necessary if NWC (WP) lightpaths are to be supported. This, however, would require the OXCs to do wavelength conversion in addition to their switching functions [6].

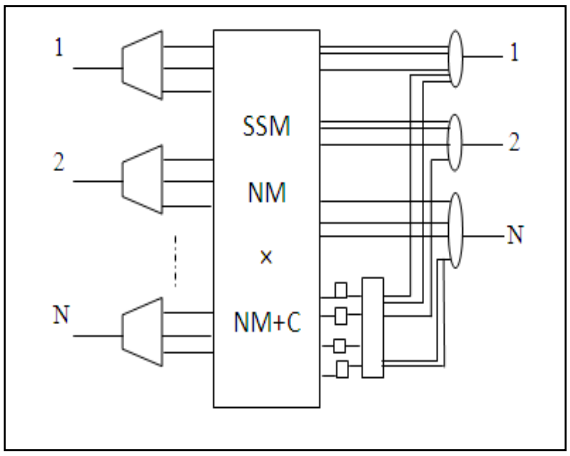



Figure 4. Share per Node (SSM-1) OXC Architecture

The OXCs may, in turn, be classified based on their wavelength conversion capability. An OXC without any wavelength conversion capability is called a wavelength selective cross-connect (WSXC) whereas an OXC with full conversion capability (i.e. capable of changing any wavelength on any incoming link to any wavelength on any outgoing link) is referred to as a wavelength interchanging cross-connect (WIXC) [4]. Examples of these have been shown in Fig. 2 and Fig. 3

SSM refers to the space switching matrix used to switch the optical signals without doing any wavelength conversion. The wavelength converters required have been shown separately.

Given the cost and complexity of wavelength conversion, an OXC with limited wavelength conversion capability may also be used as these have been observed to perform almost as well as WIXCs with full conversion capability in typical network environments. These are referred to as limited wavelength interchanging cross connect (L-WIXC). The limitation here is in the number of converters available for changing wavelengths between the inputs and outputs [4], [6].

In an L-WIXC, a limited number of wavelength converters are shared instead of being dedicated as in a WIXC. The wavelength converters are therefore more efficiently utilized. The number of wavelength converters required can be further reduced with some optimization techniques [4]. Two architectures for L-WIXC are the share per node and share per link architectures as illustrated in Fig.4 and Fig. 5.

In the share per node architecture of Fig. 4, any converter from the common pool of C converters may be accessed by any of the incoming lightpath requests by appropriately configuring the NM × (NM+C) SSM. Only the lightpaths requiring conversion are actually directed to the converters. A second stage C × NC SSM is used to switch the converted light paths to the desired output link [4].

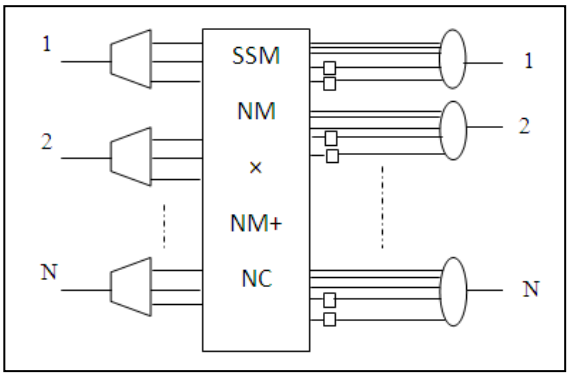



Figure 5. Share per Link (SSM-2) OXC Architecture

The share per link architecture of Fig. 5 provides a set of converters, dedicated for each output link. These may be used only by the lightpaths intended for that link. In terms of sharing efficiency of wavelength converters, share per node is the best followed by share per link and WIXC is the worst. In terms of complexity of switching, WIXC is the least complex, share per link comes next and share per node the most complex [4].

Practical implementation of the OXCs often employs multi stage structures to achieve the required size with less complexity.

## III. HOMODYNE CROSSTALK ANALYSIS

When the crosstalk signal is of same wavelength of the desired signal or sufficiently close to it that the difference in wavelength is within the receiver's electrical bandwidth, is called homodyne crosstalk. Homodyne crosstalk can be divided into coherent crosstalk and incoherent crosstalk. When the phase of the crosstalk signal is correlated with that of the main signal, it is called coherent crosstalk. When the phase of the crosstalk signal is not correlated with that of the main signal, it is called incoherent crosstalk [6], [7].

In the structure of OXC in Fig. 1, each of the input fibers to an optical demultiplexer contains M different wavelengths. Each of these passes through an optical switch before they are combined with the outputs from the other M-1optical switches. Assuming the OXC is fully loaded the OXC will be interfered by M+N-2 homodyne crosstalk contributions, N-1 of which are leaked by the optical switch and the other M-1are leaked from demultiplexer / multiplexer pair [5] .

If we consider the signal with wavelength 1 in input fiber 1, noted as $\lambda_{11}$ or the main signal. $\lambda_{11}$ will be interfered by N-1 crosstalk contributions leaked from the N-1 signals with wavelength 1 in the other *N-1* input fibers. Similarly, when each signal with wavelength 1 is demultiplexed to one path, there will be a fraction of it in each of the other M-1 outputs of the corresponding demultiplexer, because of the non-ideal crosstalk specification of optical demultiplexers. These M-1 crosstalk contributions can be leaked from any signal with wavelength 1 in all the N input fibers. The number of contributions leaked from each signal is random, from 0 to *M-1*, depending on the cross connecting state of the OXC [5]-[7]. Defining $X_1$ as the number of contributions leaked from $\lambda_{11}$ in a given state of the OXC,

$$X_1 \in [0, M-1]$$

Defining, Xj( j=[2,N] ) as the number of contributions leaked from λj1 in the same state of the OXC, taking into account the N-1 contributions leaked by the optical switch 1, we have

$$Xj \in [1, M] \quad \text{and}$$

$$X_1 + \sum_{j=2}^{N} X_j = M + N - 2$$

The field of the main signal and all the M+N-2 crosstalk contributions can be expressed as

$$\vec{E}(t) = Eb_s(t)\cos[\omega_s t + \Phi_s(t)]\,\overrightarrow{P_s} + \sum_{i=1}^{X_1} \sqrt{\varepsilon} E\, b_s(t-\tau_i)\cos[\omega_s(t-\tau_i) + \Phi_s(t-\tau_i)]\,\overrightarrow{P_i} + \sum_{j=2}^{N}\sum_{k=1}^{X_j} \sqrt{\varepsilon} E b_j\,(t-\tau_{jk})\cos[\omega_j(t-\tau_{jk}) + \Phi_j(t-\tau_{jk})]\,\overrightarrow{P_{jk}}$$

Where E is the signal field amplitude which is assumed to be unchanged as the leaked power is rather low; bs(t) and bj(t) (j=[2,N]) are the binary data sequences with values of 0 or 1 in a bit period T of $\lambda_{11}$and $\lambda_{j1}$, respectively, $\omega_s(t), \Phi_s(t),$ and $\omega_j(t), \Phi_j(t)$ are the center frequencies and phase noises of the lasers, respectively, $\overrightarrow{P_s}$ is the unit magnitude polarization vector of the signal; $\tau_i, \tau_{jk}$ and $\overrightarrow{P_i}$, $\overrightarrow{P_{jk}}$ are the propagation delay differences and unit magnitude polarization vectors of the contributions, respectively; $\varepsilon$ is the optical power ratio of each crosstalk contribution to the signal and for simplicity we assume all the crosstalk contributions have the same power, $\overrightarrow{P_s}, \overrightarrow{P_i}$ and $\overrightarrow{P_{jk}}$ are treated as time invariant here as they change rather slowly compared to the bit period. Now depending on relation between $\tau_i, \tau_{jk}, \tau_{coherent}$ and T three cases may be considered for which the laser relative intensity noise (RIN) will get different values [5].

**Case 1:** If $\tau(\tau_i and \tau_{jk}) > \tau_{coherent}$ : As $\Phi_s(t)$ is uncorrelated with $\Phi_s(t-\tau_i)$ *and* $\Phi_j(t-\tau_{jk})$ are also uncorrelated with each other for different k. In that case the noise power can be expressed as [5]

$$\sigma^2_{RIN,1} = \varepsilon \sum_{l=1}^{M+N-2} cos^2\theta_l \dots\dots\dots (1)$$

$$\cos\theta_l = \vec{P_s}.\vec{P_l}$$

where, $\theta_l$ is the polarization angle difference between the *l*th crosstalk contribution and the signal.

If $\tau(\tau_i and \tau_{jk}) < \tau_{coherent}$ : Depending on the relation between $\tau$ and T two cases may arise.

**Case 2(a):** If $\tau(\tau_i and \tau_{jk}) \ll T$ : As $b_s(t-\tau_i)$ equal to $b_s(t)$ approximately in this case, so coherent crosstalk do not cause noise but causes fluctuation. So, noise power will be [5]

$$\sigma^2_{RIN,2a} = \varepsilon \sum_{j=2}^{N} \left(\sum_{k=1}^{X_j} cos\phi_{jk}\, cos\theta_{jk}\right)^2 \dots\dots\dots (2)$$

**Case 2(b):** If $\tau(\tau_i and \tau_{jk}) > T$ : As $b_s(t-\tau_i)$ becomes completely incorrelated with $b_s(t)$ due to unsynchronus nature of $b_s(t)$ the noise power will be [5]

$$\sigma^2_{RIN,2b} = \frac{1}{3}\varepsilon \sum_{i=1}^{X_1} \left(cos\phi_i cos\theta_{jk}\right)^2 + \varepsilon \sum_{j=2}^{N} \left(\sum_{k=1}^{X_j} cos\phi_{jk}\, cos\theta_{jk}\right)^2 \dots\dots (3)$$

In this paper the worst case scenario with fully loaded OXC has been considered to get the maximum power penalty when the polarization angle is the minimum. For the above three cases at the worst condition the crosstalk can be expressed as

$$\sigma^2_{RIN,1} = \varepsilon(M+N-2) \dots (4)$$

$$\sigma^2_{RIN,2a} = \varepsilon M(N-1) \dots (5)$$

$$\sigma^2_{RIN,2b} = \frac{1}{3}\varepsilon M + \varepsilon M(N-1) \dots (6)$$

## IV. Results and Discussions

In this section, the BER performance due to the presence of Homodyne crosstalk (both coherent and incoherent) in OXCs of a WDM system is investigated. The simulation results and brief discussion on these results are also presented.

Homodyne crosstalk induced RIN due to these OXCs is given by equations 1-3 for both coherent and incoherent case. Case 1 represents the incoherent homodyne crosstalk while there are 2 cases for coherent homodyne crosstalk. Case 2a occurs when optical propagation delay differences are much less than the time duration of one bit ($\tau$<<T) which means $b_s(t-\tau_i)$==$b_s(t)$. Again case 2b represent the case when $\tau$>T and $b_s(t-\tau_i)$ become uncorrelated completely with $b_s(t)$ as $b_s(t)$ is a random sequence and they are not synchronized. The effect of the RIN caused by these cases on BER is described in section 3. To observe the BER performance we assumed the worst case scenario and simulated equation 4 - 6 incorporating the Homodyne crosstalk induced RIN into these equations. To evaluate the expressions, we assumed, T=300K, k=1.38x$10^{-23}$, $B_e$=$10^9$ Hz and $R_L$=50 Hz.

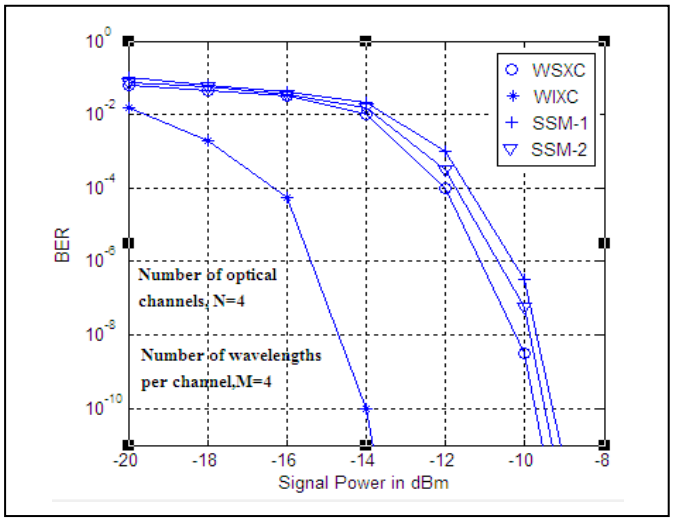



Figure 6. BER for Different OXC Architectures for case 1

Fig. 6 shows the BER performance of all the architectures discussed above for incoherent homodyne crosstalk (case1). Here, number of optical channels, N=4 and number of wavelength per channel, M=4. WIXC shows better response comparing with other architectures. The SSM architectures having limited conversion capability shows almost similar performance to the WSXC OXC.

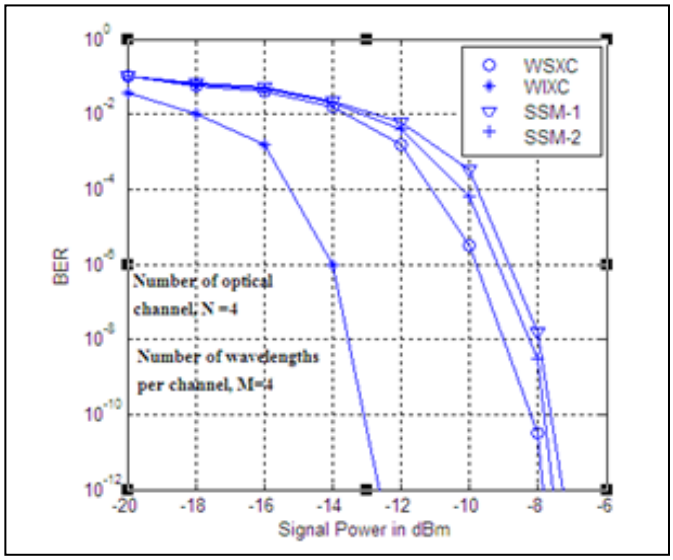



Figure 7. BER for Different OXC Architectures for case 2(a)

On the other hand, Fig. 7 shows the BER performance of the discussed architectures for coherent homodyne crosstalk while coherent time is smaller (case 2 a). Other parameters of simulation have been kept similar. Again, WIXC shows better response comparing with other architectures and SSM-1 shows the worst performance. The BER curves have been found to be shifted by almost 1dBm towards right. It reflects, to attain similar BER more power is needed than previous case.

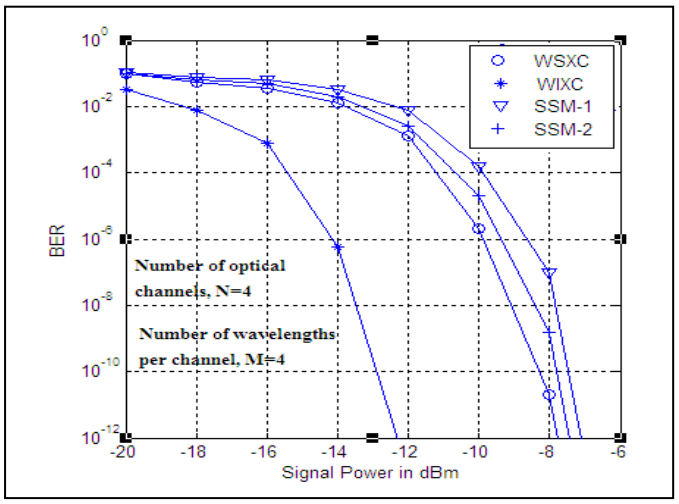



Figure 8. BER for Different OXC Architectures for case 2(b)

Fig. 8 shows similar BER performance of the discussed architectures for coherent homodyne crosstalk while coherent time is greater (case 2 b). Other parameters of simulation have been kept similar. The comparison among the architectures shows similar result but the BER curves have been found to be shifted by fraction of 1 dBm towards right.

From the three curves it evident that, WIXC has the best wavelength interchanging capability, but by using OXCs of limited wavelength changing capability (SSM-1, SSM-2) we can get similar result as that of the Wavelength Selective Cross Connect (WSXC). The curves also reflect, to achieve an acceptable BER we need more signal power to overcome coherent crosstalk rather than incoherent crosstalk.

## V. CONCLUSION

The effect of homodyne crosstalk on BER performance is investigated for both coherent and incoherent homodyne crosstalk. This study enables us to do compromise between different system parameter to achieve a particular network performance.

There is a scope for analyzing performance of the described OXCs by considering scenario other than worst case. This means if we consider that not all optical switches and mux/demux pair contributes homodyne crosstalk than performance of the OXCs will show difference pattern especially WSXC, SSM-1, SSM-2 may change their relative positions. By assuming different probability distribution of the cross connect state and imposing a relaxation factor or quintiles we will get relaxation in terms of power penalty. So, future work can also be done by statistical analysis to relax the extreme power penalty requirements and to find out a better OXC from much available architecture to meet the network performance requirements. The number of fibers can be increased without significant penalty if the performance of the switch is improved. The number of wavelengths can be increased but requires higher suppression of other channels or regeneration. Realistic systems require a larger number of wavelengths compared to the number of fibers.

## REFERENCES


[1] E. Iannone and R. Sabella, "Optical Path Technologies: A comparison among different cross-connect architectures," *J. Lightwave Technol.*, vol. 14, no. 10, pp. 2184–2194, Oct. 1996.

[2] J. Zhou, "Crosstalk in multiwavelength optical cross connect networks," *J. Lightwave Technol.*, vol. 14, no. 6, pp. 1423–1435, Jun. 1996.

[3] C. S. Li and F. Tong, "Crosstalk and Interference penalty in all optical networks using static wavelength routers," *J. Lightwave Technol.*, vol. 14, no. 6, pp. 1120–1126, Jun. 1996.

[4] T. Y. Chai, T. H. Cheng, G. Shen, S. K. Bose and C. Lu, "Design and Performance of optical crossconnect architectures with converter sharing," *Optical Networks Mag.*, pp. 73–84, Jul/Aug. 2002.

[5] Y. Shen, K. Lu and W. Gu, "Coherent and Incoherent Crosstalk in WDM optical networks," *J. Lightwave Technol.*, vol. 17, no. 5, pp. 759–764, May. 1999.

[6] T. Y. Chai, T. H. Cheng, S. K. Bose, C. Lu and G. Shen, "Crosstalk Analysis for Limited Wavelength Interchanging Cross Connects," *IEEE Photon. Technol Lett..*, vol. 14, no. 5, pp. 696–698, May. 2002.

[7] E. Iannone , R. Sabella, M. Avattaneo and G. D. Paolis, "Modeling of In-Band Crosstalk in WDM Optical Networks," *J. Lightwave Technol.*, vol. 17, no. 7, pp. 1135–1141, Jul. 1999.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# DSP Aided Chromatic Dispersion Reckoning in Single Carrier High Speed Coherent Optical Communications

Amir Hamja[1], Md. Siam Uddin[2,*], Jakia Sultana[3], Md. Monjurul Islam[4], and Shahid Iqbal[5]
[1, 2, 3, 5] Department of EEE, Mymensingh Engineering College, Bangladesh
[4] Department of CSE, Mymensingh Engineering College, Bangladesh
[*] E-mail: siam@eee.mec.ac.bd

*Abstract*-**This paper demonstrates digital signal processing (DSP) based chromatic dispersion (CD) reckoning technique for single carrier high speed coherent optical communications. Coherent detection permits the optical field parameters (amplitude, phase and polarization) to be available in the electrical domain enabling new opportunity for multi-level signaling (M-ary PSK and M-ary QAM modulation), as well as the possibility of exploring polarization multiplexing. Again, it enables quasi-exact compensation of linear transmission impairments by a linear filter (equalizer), which can operate adaptively to overcome time-varying impairments. That is why Coherent detection employing QAM modulation formats has become one of the most promising technologies for next generation high speed transmission systems due to the high power and spectral efficiencies. With the development of DSP, coherent optical receivers allow the significant estimation of CD. In recent days, the realizations of these DSP algorithms for mitigating the chromatic dispersion in the coherent transmission systems are the most attractive investigations. In this paper, CD estimation technique has been improved using DSP algorithm and optimal OSNR, BER have been achieved.**

***Keywords—Chromatic dispersion (CD), coherent detection, digital signal processing (DSP).***

## I. INTRODUCTION

In order to cope with the increasing global information exchange it's becoming crucial to transmit information over longer distance. A solution to this issue is optical fibers that are already used for most of the voice and data traffic all over the world. Optical fibers are exceptionally advantageous for long-haul communication systems.

The reason behind it, when light propagates through an optical fiber it suffers from less attenuation than the case of an electrical cable. Moreover it's possible to transmit several channels, at different wavelength, on the same medium using wave division multiplexing (WDM). This enables to reach a capacity system of several Tbps. Recently, digital coherent optical communication has become the main technology for optical transport networks [1]. Moreover, digital signal processing is under consideration as a promising technique for optical signal modulation, fiber transmission, signal detection and dispersion compensation. There are different reasons why the utilization of coherent detection associated digital signal processing can be very advantageous.

Firstly, coherent detection is a promising technology to increase optical receiver sensitivity, permitting a greater span loss to be tolerated. Secondly, coherent detection enables supporting of more spectrally efficient modulation formats such as quadrature phase shift keying (QPSK) and quadrature amplitude modulation (QAM). And finally, instead of implementing costly physical impairments compensation links, coherent detection allows digital signal processing for compensation of transmission impairments such as CD, polarization mode dispersion (PMD), signal carrier offset, spectrum narrowing, etc. Furthermore, next generation optical transmission systems require adaptive fitting for time varying transmission impairments such as channel spectrum narrowing and random phase noise. Digital signal processing is a powerful solution for future adaptive optical transmission links.

Recently, 100-Gb/s technologies using polarization-division-multiplexed quadrature-phase-shift-keying (PDM-QPSK) and digital coherent detection have been commercialized, and the focus of the optical communication industry is moving beyond 100-Gb/s. In a digital coherent optical communication system, no optical dispersion compensation is required and the large amount of chromatic CD accumulated along the link can be compensated by DSP in a coherent optical receiver [2]. As there occurs large amount of accumulated CD in the system, a relative small error in CD estimation and compensation can cause failure of subsequent clock recovery, dynamic butterfly equalizer and carrier recovery, and thus entire digital detection process of the receiver. Moreover, CD is mainly responsible for pulse broadening which gives birth to bit error rate (BER). Therefore, accuracy in CD estimation and compensation are requisite for a coherent optical receiver [3].

## II. ASSIMILATING DIFFERENT CD ESTIMATION METHOD

Many methods for CD estimation have been proposed and demonstrated [4]. A commonly adopted solution is the using Dispersion-Compensating Fiber (DCF) however DCF introduces additional loss, therefore requiring additional optical amplifiers increasing additional noise and cost of the system. An alternative approach is to compensate entirely CD

in the electric domain. At the very beginning, equalization with training sequence was used but this method is found to be good only for low dispersion system, as any system has a relative high CD in practice, or system with DCF and residual CD. The mean advantages of utilizing coherent detection with DSP algorithms can be summarized as follows. Firstly, coherent detection with DSP algorithms is able to detect advanced modulation formats with improved spectral efficiency, such as n phase-shift keying (PSK). Secondly, coherent detection with DSP algorithms provides potential for superior receiver sensitivity. Furthermore, coherent detection with DSP algorithms enables electrical offline compensation for impairments arising due to fiber transmission [5]. Coherent detection with DSP algorithms can take advantage from continuously increasing electrical processing speed. Moreover optical coherent detection can benefit from intensive researches concerning digital signal processing algorithms. Consequently, it is efficient and simple to implement optical signal processing in digital domain.

## III. PROPOSED MODEL

The aim of our research is to develop and analyze digital signal processing techniques for enhancing performance in coherent optical communication systems. This scheme allows the use of PDM without the need for adaptive optics, and also enables full compensation of arbitrary amounts of previously limiting effects such as PMD and CD. This DSP method allows us to reduce the Polarization mode dispersion, Chromatic dispersion. By implementing this DSP algorithm it is possible to achieve high data rate communication applications.

## IV. TECHNICAL DESCRIPTION

The common configuration of optical coherent receiver associated with DSP algorithms is shown in Fig. 2 Coherent detection with DSP algorithms can take advantage from continuously increasing electrical processing speed. Consequently, it is efficient and simple to implement optical signal processing in digital domain. Four main functions are performed in digital domain: 1) Dispersion compensation, 2) Clock recovery, 3) Polarization de-multiplexing, and 4) Carrier phase estimation.

### *A. Dispersion Compensation*

The frequency response for an all-pass filter to compensate fiber CD can be expressed as in the absence of fiber nonlinearity:

$$G(z,\omega)=\exp[-jD\frac{\lambda^2}{2\pi c}\frac{\omega^2}{2}z+jS\left(\frac{\lambda^2}{2\pi c}\right)^2\frac{\omega^3}{6}z] \qquad (1)$$

Where D is the dispersion coefficient, S is the dispersion slop, $\omega$ is the angular frequency, $\lambda$ is the light wavelength, c is the light velocity, and z is the fiber length. In order to compensate for the dispersion, the output field is multiplied to the inverse of the channel transfer function (FIR filter) [6]. After CD compensation at frequency domain, IFFT inverts the sequence back to the time domain.

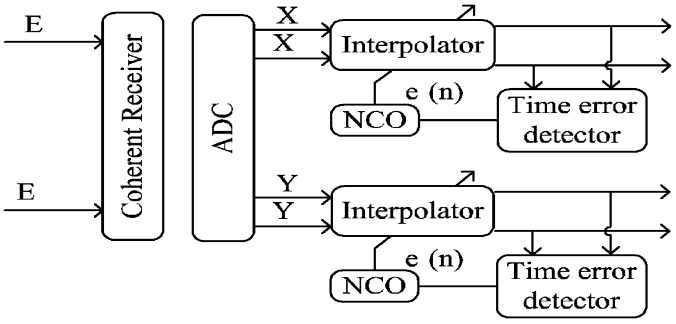



Figure 1. Structure of Gardner clock recovery DSP algorithm

### *B. Clock Recovery*

In general, any sampling clock errors significantly reduce system BER performance. Therefore clock recovery DSP algorithm is demanded to determine the suitable sampling clock. The clock recovery DSP algorithm implemented in this paper is known as Gardner algorithm [7], which is widely used in the field of wireless communication systems.

$$e_x(n)=X_I(n)[X_I(n+1)-X_I(n-1)]+ X_Q(n)[X_Q(n+1)-X_Q(n-1)] \qquad (2)$$

$$e_y(n)=Y_I(n)[Y_I(n+1)-Y_I(n-1)]+ Y_Q(n)[Y_Q(n+1)-Y_Q(n-1)] \qquad (3)$$

### *C. Polarization Demultiplexing*

In order to emulate the cross-talk between the signals carried on two polarizations, Jones matrix is employed, which is given as:

$$\begin{pmatrix}\sqrt{\alpha e^{i\delta}} & -\sqrt{1-\alpha}\\ \sqrt{1-\alpha} & \sqrt{\alpha e^{-i\delta}}\end{pmatrix} \qquad (4)$$

Where $\alpha$ and $\delta$ denote the power splitting ratio and phase difference between two polarizations. Therefore the polarization multiplexed signal at the receiver side after fiber propagation can be presented as [8].

$$\begin{bmatrix}E_x\\ E_y\end{bmatrix}=\begin{pmatrix}\sqrt{\alpha e^{i\delta}} & -\sqrt{1-\alpha}\\ \sqrt{1-\alpha} & \sqrt{\alpha e^{-i\delta}}\end{pmatrix}\begin{bmatrix}E_{in,x}\\ E_{in,y}\end{bmatrix} \qquad (5)$$

So if the inverse of Jones matrix is found, polarization de-multiplexing can be performed.

$$\begin{bmatrix}E_X\\ E_Y\end{bmatrix}=\begin{pmatrix}P_{xx} & P_{xy}\\ P_{yx} & P_{yy}\end{pmatrix}\begin{bmatrix}E_x\\ E_y\end{bmatrix} \qquad (6)$$

The matrix elements are updated symbol by symbol according to

$$P_{xx}(n+1) = P_{xx}(n) + \mu(1 - |E_x(n)|^2) E_x(n) E_x^*(n) \quad (7)$$

$$P_{xy}(n+1) = P_{xy}(n) + \mu(1 - |E_x(n)|^2) E_x(n) E_y^*(n) \quad (8)$$

μ is the step-size parameter and n is the number of symbols. The P matrix is basically an adaptive FIR filter and we use CMA for blind estimation [9]. The initial values for $P_{xx}(0)$ and $P_{yy}(0)$ are: $P_{xx}(0)$ = [00…010..00];$P_{yy}(0)$ = [00…010..00];again $P_{xy}(0)=P_{yx}(0)$= [00…000..00]. In this simulation a 3-tap FIR filter, however the order can be changed is chosen.

### *D. Carrier phase estimation*

Phase locking in the hardware domain can be replaced by phase estimation in digital domain by DSP [10]. The received QPSK signal can be presented to estimate the phase of the signal in digital domain by

$$E(t) = A \exp\{j[\theta_S(t) + \theta_C(t)]\} \quad (9)$$

## V. SYSTEM DESIGN

The configuration of optical coherent receiver associated with DSP algorithms is shown in Fig. 2. In this block diagram, several DSP algorithms are under consideration. CD compensation block is used to compensate for chromatic dispersion. Clock recovery block is implemented to correct digital sampling error which is made by analog to digital converters (ADC). Polarization de-multiplexing is realized by using polarization de-multiplex algorithm. Phase and frequency offset recovery block is employed to correct phase and frequency difference between received signal and Local Oscillator.

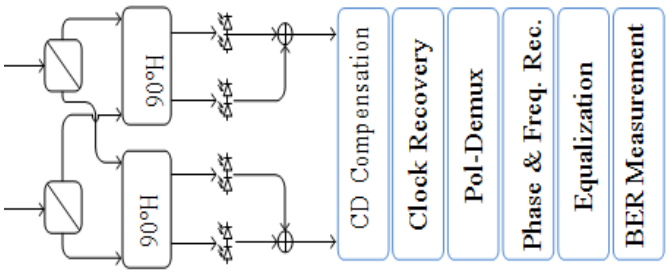



Figure 2. Configuration of coherent detection with DSP.

## VI. SIMULATION DESCRIPTION

In this subsection, the combination of coherent detection with DSP algorithm is proposed to compensate for CD in dual polarization (DP) QPSK systems. The configuration diagram of optical DP-QPSK system with coherent digital receiver is shown in Fig. 2.

The 100 Gbps DP-QPSK system can be divided into five main parts: DP-QPSK Transmitter, Transmission Link, Coherent Receiver, Digital Signal Processing, and Detection & Decoding (which is followed by direct-error-counting). The signal is generated by an optical DP-QPSK Transmitter, and is then propagated through the fiber loop where dispersion and polarization effects occur. It then passes through the Coherent Receiver and into the DSP for distortion compensation. The fiber dispersion is compensated using a simple transversal digital filter, and the adaptive polarization de-multiplexing is realized by applying the constant-modulus algorithm (CMA). A modified Viterbi-and-Viterbi phase estimation algorithm (working jointly on both polarizations) is then used to compensate for phase and frequency mismatch between the transmitter and local oscillator (LO) [11]. After the digital signal processing is complete, the signal is sent to the detector and decoder, and then to the BER Test Set for direct-error-counting.

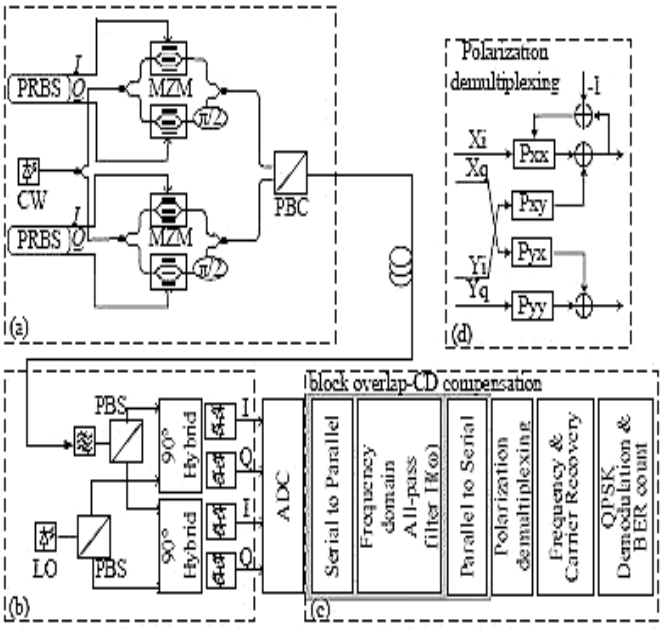



Figure 3. Coherent DP-QPSK system configuration: (a) DP-QPSK transmitter, (b) Coherent receiver optical front end, (c) Off-line DSP algorithms blocks, (d) Polarization de-multiplexing buttery structure

## VII. RESULTS AND DISCUSSIONS

Below some images of the optical spectrum of the 100 Gbps DP-QPSK signal after the transmitter, as well as the RF spectrum obtained after the Coherent DP-QPSK Receiver. Fig. 4 shows the analyzed optical spectrum after optical DP-QPSK transmitter of frequency 1550 nm and Fig. 5 is the analyzed RF signal spectrum of transmitter. Fig. 6 is the electrical constellation visualizer of polarization-X before the DSP algorithm. Fig. 7 unveils the receiver constellation diagram after DSP compensation, it gives effective spectrum efficiency in the view of high OSNR.

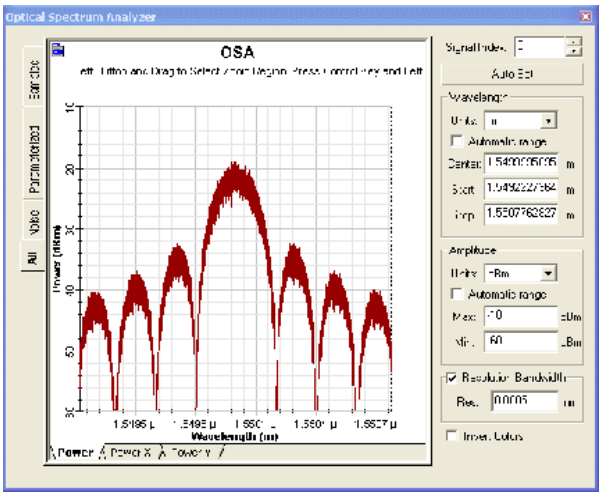



Figure 4. Optical Spectrum Analyzer after transmitter

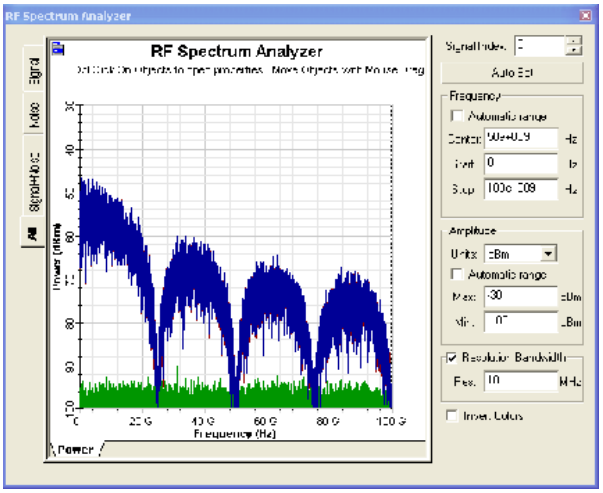



Figure 5. RF Spectrum Analyzer after transmitter

The electrical constellation diagrams (for polarization X) before and after the DSP are as follows:

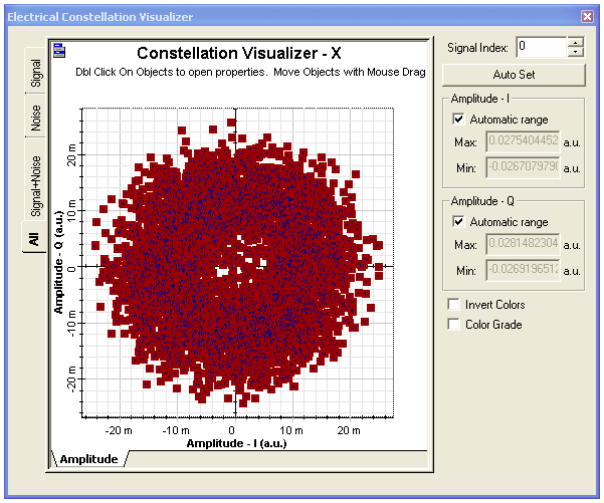



Figure 6. Electrical constellation visualize-X before DSP

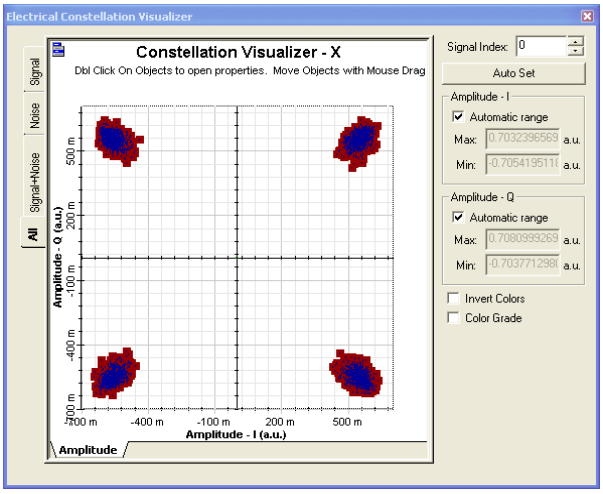



Figure 7. Electrical constellation visualize-X after DSP

The algorithms used for digital signal processing are implemented through a Matlab component. By setting the Matlab component to debug mode, the generated electrical constellation diagrams after each step (CD compensation, Polarization De-multiplexing, and Carrier Phase Estimation) are shown here:

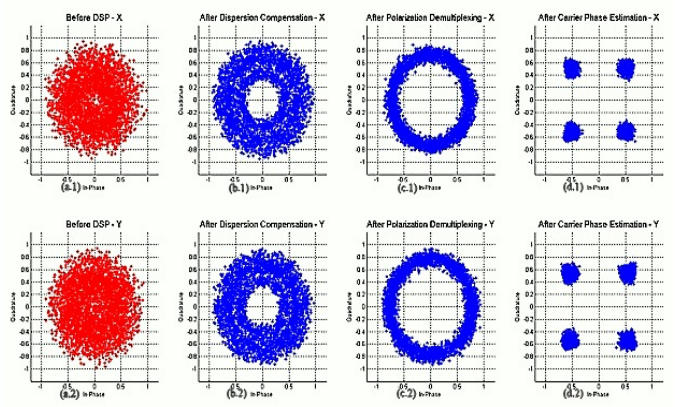



Figure 8. Matlab implementation of DSP algorithm. (a.1, a.2) Before DSP of X & Y. (b.1, b.2) After dispersion compensation. (c.1, c.2) After polarization demultiplexing. (d.1, d.2) After carrier phase estimation

## VIII. CONCLUSION

Coherent detection at 100 Gbps will become feasible in the future using DSP for linear impairments compensation. DSP facilitates polarization de-multiplexing, compensation of linear transmission impairments, as CD and PMD, and also gives higher OSNR tolerance. Improved OSNR tolerance leads to increasing the maximal propagation distance with less optical amplifiers, less noise and less costs of the system. On the other hand a more complex receiver is required as polarization tracking has to be performed but this is well paid by the improvement on system performance. The performance of DSP algorithm in nonlinear compensation deteriorates when inter channel nonlinearities become predominant.

Considering the future work, there are a lot of issues that can be addressed. First is to implement more complex and advanced adaptive algorithms for the PMD and residual CD

compensation (especially step size updating algorithms). Investigation over nonlinear effects compensation would also be interesting. Advanced techniques based on Volterra compensator are interesting because can avoid this problem.

## REFERENCES


[1] Chongjin Xie, "Chromatic dispersion estimation for single-carrier coherent optical communication," *IEEE Photonics Technology Letters*, vol. 25, no. 10, pp. 992-995, May. 15, 2013.

[2] F. N. Hauske, M. Kuschnerov, B. Spinnler, and B. Lankl, "Optical performance monitoring in digital coherent receivers," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3623–3631, Aug. 15, 2009.

[3] R. A. Soriano, F. N. Hauske, N. G. Gonzalez, Z. Zhang, Y. Ye,and I. T. Monroy, "Chromatic dispersion estimation in digital coherent receivers," *J. Lightw. Technol.*, vol. 29, no. 11, pp. 1627–1637, Jun. 1, 2011.

[4] E. Ibragimov, G. Zarris, S. Khatana, and L. Dardis, "Blind chromatic dispersion estimation using a spectrum of a modulus squared of the transmitted signal," *In Proc. of Eur. Conf. Exhibit. Opt. Commun*, Amsterdam, The Netherlands, pp. 1–3, paper Th.2.A.3, Sep. 2012.

[5] X. Zhou and J. Yu., "Advanced coherent modulation formats and algorithms: higher-order multi-level coding for high-capacity system based on 100gbps channel," *In Proc. of 2010 OFC/NFOEC Conference on Optical Fiber Communication(OFC), Collocated National Fiber Optic Engineers Conference*, pp. 1-3, Mar. 2010

[6] Seb J. Savory, "Digital filters for coherent optical receivers," *Optics Express*, vol. 16, no. 2, pp. 804-817, Jan. 21, 2008.

[7] F. Gardner, "A BPSK-QPSK timing-error detector for sampled receivers," *IEEE Transactions on Communications,* vol. 34, no. 5, pp. 423-429, May. 1986.

[8] K. Kikuchi, "Polarization de-multiplexing algorithm in the digital coherent receiver," *IEEE/LEOS Summer Topical Meetings*, pp.101-102, 21-23 Jul. 2008.

[9] L. Liu, Z. Tao, W. Yan, S.Oda, T. Hoshida, and J. C. Rasmussen, "Initial tap setup of constant modulus algorithm for polarization de-multiplexing in optical coherent receivers," *In Proc.of IEEE Optical Fiber Communication*, pp.1-3, 22-26 Mar. 2009.

[10] Guifang Li, "Recent advances in coherent optical communication," *Adv. Opt. Photon*, vol. 1, no. 2, pp. 279-307, Apr. 15, 2009.

[11] J. H. Lee and M. H. Sunwoo. "High-speed and low complexity carrier recovery for DP-QPSK transmission," *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 438-441, May. 2011.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Design of a Photonic Crystal Fiber for Dispersion Compensation over Telecommunication Bands

Redwan Ahmad[1,*] , A. H. Siddique[1] , Md. Sharafat Ali[1] , Aminul Islam[1], K.M.Nasim[1] , M. Samiul Habib[1]
[1]Dept of Electrical and Electronic Engineering , Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh.
[*] E-mail: dipto081063@gmail.com

*Abstract*— **This paper presents a microstructure optical fiber based on an hexagonal structure for dispersion compensation in a wideband transmission system. According to simulation, negative dispersion coefficient of - 562 ps/(nm.km) and relative dispersion slope (RDS) close to that of single mode fiber (SMF) of about 0.0036 nm$^{-1}$ is obtained at 1550 nm wavelength. Besides the proposed hexagonal microstructure optical fiber (H-MOF) offers high birefringence of $3.06\times10^{-2}$ . Due to having better optical properties, this proposed fiber can be effectively used in broadband dispersion compensation and sensing applications.**

***Keywords— Photonic crystal fiber, Confinement loss, Birefringence, Dispersion compensating fiber.***

## I. INTRODUCTION

Photonic crystal fiber have drawn increased attention due to many attractive properties such as wideband dispersion flattened characteristics, high or low birefringence [1] . In Wavelength division multiplexing (WDM) and high bit rate transmission systems dispersion is one of the major problem because it broadening the optical pulse and limit the bandwidth of the system. Dispersion compensating fiber (DCF) with high negative dispersion is used to nullify the accumulated positive dispersion of SMFs [2]. Photonic crystal fiber (PCF) or holey fiber offers flexibility in tuning dispersion, because of by varying the size of air-holes and their position, the dispersive properties can be controlled [3]-[4] which is crucial for dispersion compensating fiber design. The negative dispersion coefficient of conventional fiber is about - 100 to - 300 ps/(nm.km) at 1550 nm [5]. To minimize the losses and reduce the cost, the dispersion compensating fiber (DCF) should be as short as possible with high negative dispersion [6]-[9].

Several attempts have been made by different groups to achieve high negative dispersion as well as a suitable bandwidth for dispersion compensation. For example an Octagonal MOF structure in [10] exhibits negative dispersion coefficient of - 239.5 ps/(nm.km) with high birefringence of $1.67\times10^{-2}$, which is less than our proposed MOF. Another design proposed by [11] which offers negative dispersion coefficient of - 300 ps/(nm.km) at 1550 nm but birefringence is not accounted here. The MOF designed by Matsui *et al.* simultaneously covers all three communication band but due to its low dispersion peak it requires a long fiber to compensate the dispersion [12].

In this paper, we propose an hexagonal MOF structure that is suitable in compensating the dispersion of SMF over a wide range of wavelengths. The main advantage of our proposed MOF is the design flexibility with high negative dispersion of - 562 ps/(nm.km) and high birefringence of the order $3.06\times10^{-2}$ which is very crucial in high bit rate transmission network and sensing applications. . Another advantage of proposed MOF is , in our design to reduce the complexity of fabrication process we have used only circular air holes with less optimized parameters.

## II. DESIGN METHODOLOGY

Fig. 1 shows the air holes distribution of the proposed H-MOF which contains five air hole rings. Where $\Lambda$ is the pitch of the lattice, $d_3$ is the air hole diameter of the 3$^{rd}$ ring and $d$ is the air hole diameter of rest of the ring. The hot material in our proposed structure is silica and air holes are arranged in hexagonal structure symmetry. To increase the birefringence four air holes along the y axis in first ring make small. In order to achieve large negative dispersion, air holes near the fiber core are chosen higher [13]. The air holes diameter in the third ring are relatively smaller for controlling RDS and rest of the ring's air hole diameter keeps same to lessen fabrication complexity [14].

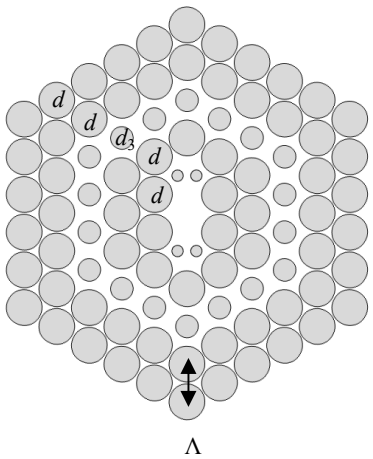



Figure 1. Transverse cross section of proposed O-MOF

## III. NUMERICAL METHOD

Finite element method (FEM) with perfectly matched layers (PML) boundary condition is used to calculate the chromatic dispersion, effective area and confinement loss of the proposed MOFs. Once the modal effective index $n_{eff}$ is obtained by solving an eigen value problem using FEM, the Chromatic dispersion $D(\lambda)$ , confinement loss $L_c$ and birefringence $B$ can be calculated by the following equations [14].

$$D(\lambda) = -\lambda/c \, (d^2 \mathrm{Re}[n_{\mathrm{eff}}]/ d\lambda^2) \quad (1)$$

$$L_c = 8.686 \times k_0 \, \mathrm{Im}[n_{\mathrm{eff}}] \times 10^3 \text{ dB/ km} \quad (2)$$

$$B = \left| n_x - n_y \right| \quad (3)$$

Where Re[$n_{eff}$] and Imag[$n_{eff}$] is the real part and imaginary of effective refractive index $n_{eff}$ respectively , $\lambda$ is the wavelength in vacuum, c is the velocity of light in vacuum and $k_0$ is the free space number.

Because of the positive dispersion and dispersion slope of the SMF, the requirements of a DCF for WDM operation are large negative dispersion and a dispersion slope over a wide range of wavelengths.

$$D_{\mathrm{SMF}} \, . \, L_{\mathrm{SMF}} + D_{\mathrm{DCF}} \, . L_{\mathrm{DCF}} = D_{\mathrm{T}} \quad (4)$$

Where $D_{SMF}$ and $D_{DCF}$ are the dispersion coefficient of the mode fiber and dispersion compensating fiber respectively. If the total compensation of dispersion is required , length of DCFs ($L_{DCF}$) is so chosen that total residual dispersion becomes zero. Besides the dispersion , it is also necessary to compensate for the dispersion slope [10].

$$S_{\mathrm{slope}} = S_{\mathrm{SMF}} \, . \, L_{\mathrm{SMF}} + S_{\mathrm{DCF}} \, . L_{\mathrm{DCF}} \quad (5)$$

Where $S_{SMF}$ and $S_{DCF}$ are the dispersion coefficient of the single mode fiber and dispersion compensating fiber respectively. The relative dispersion slope of SMF is 0.0036 $nm^{-1}$. The proposed design can be efficiently used when RDS of the proposed DCF is exactly equal or very close to that of SMF.

## IV. SIMULATIONN RESULTS AND DISCUSSION

Fig. 2(a) shows the dispersion characteristics of both x and y polarization for optimum design parameters with $d/\Lambda$ =0.95, $d_3/\Lambda$ = 0.59 and pitch $\Lambda$=0.90 µm. Global diameter of the four air holes in 1st ring along the y axis is kept as 0.3. From curve it is seen that, the proposed DCF exhibits negative dispersion coefficient about - 562 ps/(nm.km) along the y polarization at 1550 nm. Due to having high negative dispersion coefficient our proposed DCF could be suitable candidate for dispersion compensating in high bit rate transmission network.

In PCF ± 1% variation in global diameters may be occurred during fabrication [16]. Due to this reason, we have analyzed the effect on dispersion and birefringence by varying different parameters ± 2% to ± 5%, which is discussed in the following section.

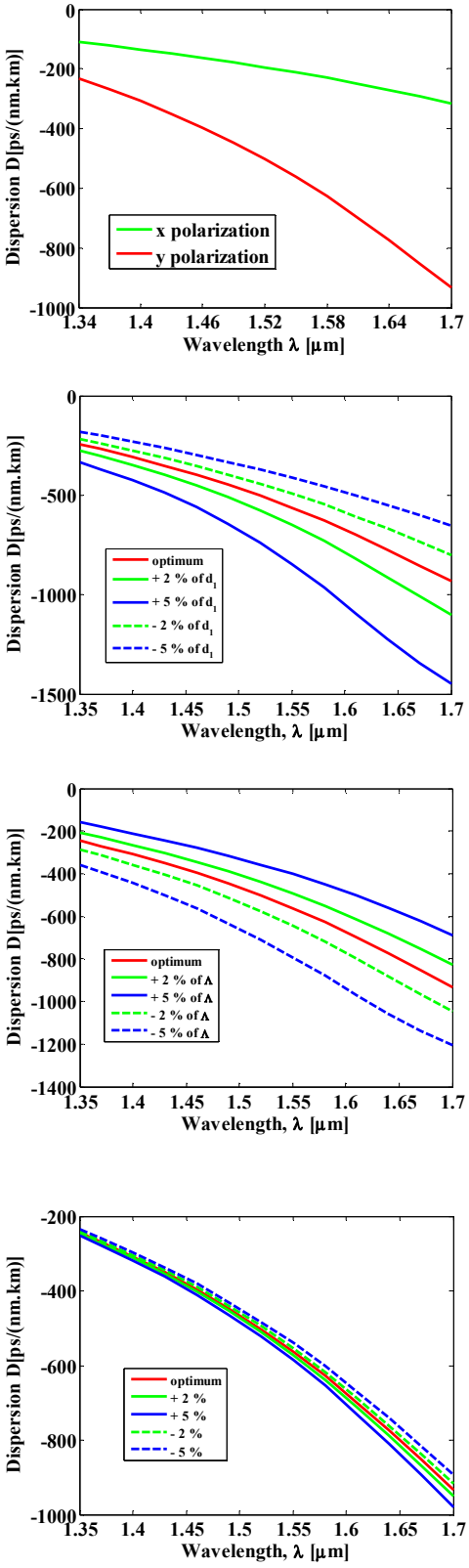



Figure 2. (a) Wavelength dependence dispersion curve for both x and y polarization (b) Effect on dispersion by changing $d_1/\Lambda$ (c) Effect on dispersion by changing $\Lambda$ (d) Effect on dispersion by changing four air holes along the y axis.

Fig. 2(b) reveals the effect by varying global diameter of first ring ± 2% to ± 5%, while other parameters are kept constant. . Solid line depicts due to increment in parameters whereas dashed line for decrement. When $d_1/\Lambda$ is varied as ± 2% to ± 5%, their corresponding dispersion value becomes - 650 , - 845 , - 493 and - 412 ps/(nm.km) respectively. From Fig. 2(c) it is seen that by varying pitch has a remarkable effect in dispersion value, but there is a insignificant effect in dispersion by varying global diameter of four air holes in 1st ring along the y axis (Fig. 2(d)).

Birefringence characteristics of the proposed DCF are also shown in fig. 3(a). From curve it is seen that this proposed DCF shows birefringence about $3.06\times10^{-2}$ at 1550 nm. High birefringence makes our proposed MOF very efficient in sensing applications. Fig. 3(b) shows the effect in birefringence by varying global diameter of first ring . As $d_1/\Lambda$ is varied as ± 2% to ± 5% from optimum value , birefringence at 1550 nm becomes 0.0329, 0.0365, 0.0284 and 0.02533 respectively. Due to asymmetric core design, the proposed design exhibits high birefringence, which is essential in polarization maintaining applications. However current conventional PM fibers shows a modal birefringence about $5\times10^{-4}$ [13]. Moreover our proposed PCF exhibits birefringence about $3.06\times10^{-2}$, which could be a suitable candidate in sensing applications. From Fig. 3(d) it is seen that there is a significant effect in birefringence by varying global diameter of four air holes in 1st ring along the y axis. The effect of pitch on birefringence also studied, which is shown in fig. 3(c).

While designing Residual dispersion compensating fiber (RDCF) , splice loss is one of the challenges, due to dissimilar mode field diameter results spilce loss. Reduction of splice is required to ensure effective performance of the optical fiber system. Fig. 4(a) shows the effective area of the proposed MOF along the y polarization . At 1550 nm the effective area of the proposed MOF is 2.06 $\mu m^2$ . The Splice free interconnection techniques between PCFs of almost any structure and the SMFs provided the PCF is drawn from individual stackable units reported by [15]. We believe that, our proposed design can easily be fabricated without any fabrication complexity. While designing RDCF, confinement loss is one of the issue. In our simulation we have used only five rings, confinement loss can be decreased by increasing the air hole ring without affecting the dispersion shape [19].

From fig. 4(b), it is seen that our proposed DCFs RDS value is 0.0036 $nm^{-1}$ . Which is exactly equal to that of SMFs RDS value. For better compensation, it is necessary to obtain high negative dispersion as well as matched RDS to that of SMFs.

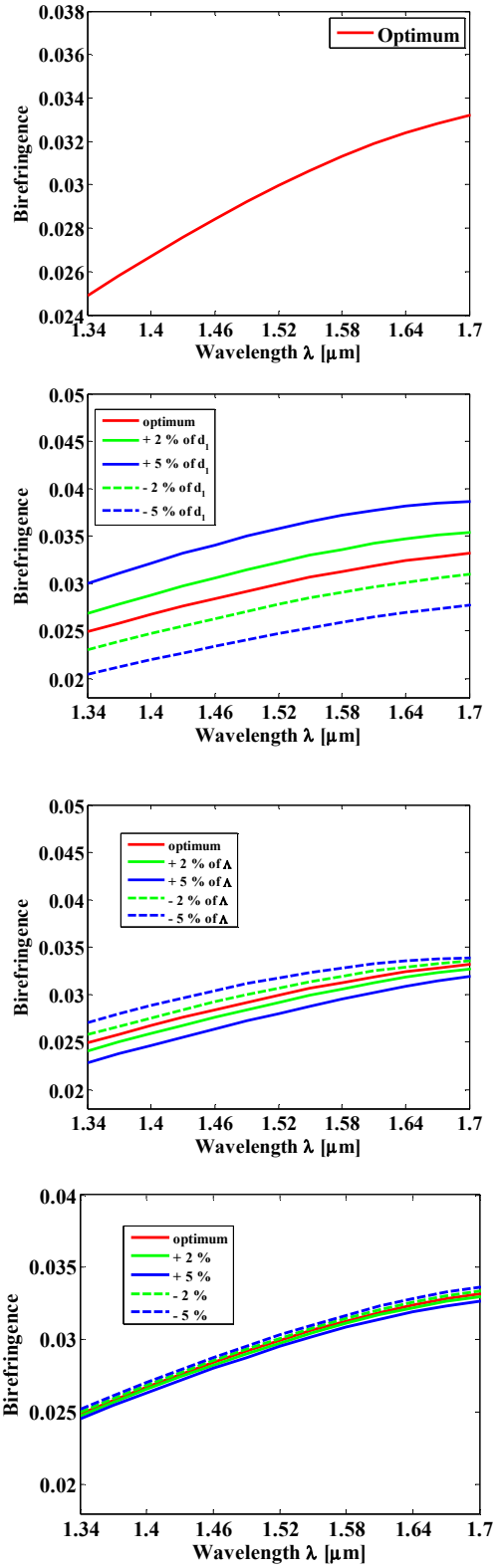



Figure 3. Birefringence as function of wavelength (a) optimum (b) by varying global diameter of 1st ring's airhole (c) by varying pitch (d) by varying four holes along the y axis.

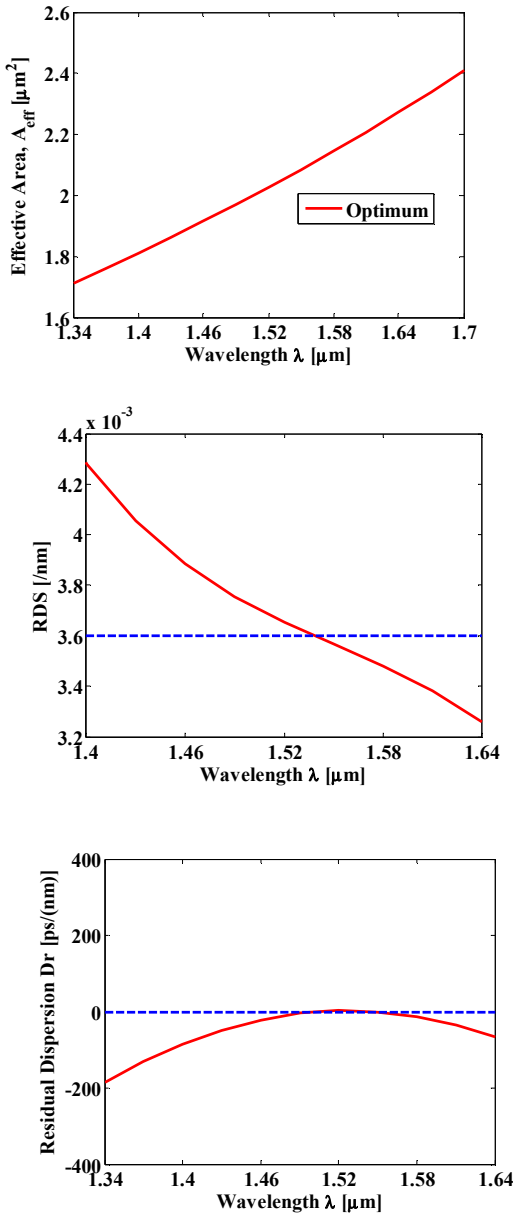



Figure 4. (a) wavelength dependence effective area (b) Relative dispersion slope of propsed MOF for optimum design parameters (c) Residual dispersion curve after compensating for 40 km SMFs.

After the dispersion compensation by 1.23 km long proposed O-MOF for the dispersion accumulated in 40 km long SMF, the residual dispersion curve is shown fig. 4(c). From simulation result it is clear that, our proposed design could be a suitable candidate for high bit rate transmission systems covering S, C and L bands [13].

Comparision between properties of the proposed MOF and other MOF at 1550 nm is shown in Table I.

TABLE I. COMPARISON OF MODAL PROPERTIES BETWEEN PROPOSED MOF AND OTHER DESIGNS

| MOFs | Comparison of modal properties | | | |
|---|---|---|---|---|
| | $D(\lambda)$ Ps/(nm.km) | $B=\lvert n_x-n_y\rvert$ | $A_{eff}$ ($\mu m^2$) | NDP ($N_1$, $N_A$,$N_d$) |
| Ref. [11] | -300 | ----- | 1.55 | 5,1,1 |
| Ref. [13] | -588 | $1.81\times10^{-2}$ | 3.41 | 5,1,2 |
| Ref. [18] | -474.5 | ----- | 1.60 | ---- |
| Proposed MOFs | -562 | $3.06\times10^{-2}$ | 2.08 | 5,1,1 |

## V. CONCLUSION

In summary, we have reported a relatively simple highly birefringent broadband dispersion compensating MOF. It has been shown through simulation results, by using proposed broadband dispersion compensating MOF, negative dispersion coefficient of about - 396 to - 735 ps/(nm.km) is successfully obtained over the S, C and L bands and an RDS close to that of a conventional SMF simultaneously. Another excellent feature of our designed fiber is that it offers high birefringence as $3.06\times10^{-2}$ along with the property of dispersion compensation. Due to having excellent guiding properties, our proposed MOF could be a suitable candidate for sensing and broadband dispersion compensation in high bit rate transmission network

## REFERENCES


[1] S. M. A. Razzak, Y. Namihira and F. Begum, " Ultra flattened Dispersion Photonic Crystal Fiber ," *Electron. Lett.*, vol. 43 , pp. 615-617, 2007.

[2] M. Koshiba and K. Saitoh, " Structural dependence of effective area and mode field diameter for holey fibers", *Opt. Express* 11 (2003) 1746-1756.

[3] S. M. A. Razzak and Y. Namihira , "Proposal for highly nonlinear dispersion- flattened octagonal photonic crystal fibers," *IEEE photon. Technol. Lett.*, 20 (2008) 249-251.

[4] K. Saitoh , M. Koshiba , T. Hasegawa and E. Sasaoka, "Chromatic dispersion control in photonic crystal fibers: application to ultra-flattened dispersion,"*Opt. Exp.*, 11 (2003) 843-852.

[5] J. Laegsgaard, S. E. BarkouLibori, K. Hougaard, J. Riished, T. T. Lasen, and T. Sorensen, "Dispersion Properties of photonic crystal fibers- issues and opportunities," *Mater. Res. Soc. Symp. Proc.*, (2004) 797.

[6] L. Gruner- Nielsen, S. N. Knudsen , B. Edvold , T. Veng, D. Magnussen, and C. C. Larsem, " Dispersion compensating fibers," *Opt. Fiber Technol.*, (2000) 164.

[7] S .G. Li, X. D. Liu, L. T. Hou, "Numerical study on dispersion compensating property in photonic crystal fibers," *Acta Phys. Sin.*, (2004) 1880.

[8] Z. W. Tan , T. G. Ning , Y. Liu, Z. Tong and S. S. Jian , "Suppression of the interactions between fiber gratings used as dispersion compensators in dense wave length division multiplexing systems," *Chin. Phys.*, (2006) 1819-1823.

[9] Y. Ni, L. Zhang, L. An, J. Peng and C. C. Fan , "Dual- core photonic crystal fiber for dispersion compensation," *IEEE Photon. Technol. Lett.*, (2004) 1516.

[10] S. F. Kaijage , Y. Namihira, N. H. Hai , F. Begum , S. M. A. Razzak , T. Kinjo, K. Miyagi and N. Zou, ''Broadband dispersion compensating octagonal photonic crystal fiber for optical communication apllications," *Japan. J. of Appl. Phy.*, vol.48, 052401-052408, 2009.

[11] M. Selim Habib , M. Samiul Habib , S. M. A. Razzak , Y. Namihira, M. A. Hossain and M. A. G. khan , "Broadband dispersion compensation of conventional single mode fibers using microstructure optical fibers," *Optik*, (doi.org/10.1016/j.ijleo.2012.12.014), Feb 2013

[12] T. Matsui, K. Nakajima and I. Sankawa, "Dispersion Compensation Over All the Telecommunication Bands With Double-Cladding Photonic-Crystal Fiber," J. Lightw. Technol. 25, 757-762 (2007).

[13] M. Selim Habib , M. Samiul Habib , S. M. A. Razzak and M. A. Hossain, ''Proposal for highly birefringent broadband dispersion compensating octagonal photonic crystal fiber," *Opt. Fiber Technol.*, (dx.doi.org/10.1016/j.yofte.2013.05.014), May 2013.

[14] M. Selim habib, M. Samiul Habib, S. M. A. Razzak, M. I. Hasan, R. R. Mahmud and Y. Namihira, ''Microstructure holey fibers as wideband dispersion compensating media for high speed transmission system," *Optik* ( 2013 ).

[15] S. G. leon- Saval , T. A. Birks , N. Y. Joy, A. K. Geotge, W .J. Wadsworth , G. Jakarantzas, P. St. J. Rusell , "Splice free interfacing of photonic crystal fibers," *Opt. Lett.* 30 (2005) 1629-1631.

[16] W. H. Reeves, J. C. Knght and P. S. J. Russell, ''Demonstration of ultra-flattened dispersion in photonic crystal fibers," *Opt. exp*., 10(2002) 609-613.

[17] T. Fujisawa, K. Saitoh, K. Wada and M. Koshiba, "Chromatic dispersion profile optimization of dual core photonic crystal fibers for broadband dispersion compensation," *Opt. Exp*., 14 (2006) 893-900.

[18] L .P. Shen , W. P. Huang , G. X. Chen and S. S. Jian , "Design and optimization of photonic crystal fibers for broadband dispersion compensation," *IEEE Photon. Technol. Lett.*, 15(2003) 540-542.

[19] G. Renversez, B. Kuhlmey, and R. McPhedran, "Dispersion management with microstructured optical fibers: Ultraflattened chromatic dispersion with low losses," Opt. Lett., 28 (2003) 989-991.

# Estimation of Second-Order PMD from Adaptive FDE in Coherent Optical Receivers

Md. Saifuddin Faruk
Department of Electrical and Electronic Engineering, Dhaka University of Engineering and Technology, Bangladesh
E-mail: faruk@duet.ac.bd

*Abstract*—**Recently digital signal processing (DSP) based channel parameter estimation in coherent optical receivers attacks significant attention. Second-order polarization mode dispersion (PMD) is an important parameter to estimate for high-speed optical communication systems. In this paper, we propose a new second-order PMD estimation technique from the adaptive frequency-domain equalizer (FDE). The proposed method is verified with 100-Gbits/s dual-polarization QPSK experiments.**

***Keywords—Optical fiber communication, digital coherent receivers, channel parameter estimation, digital signal processing.***

## I. INTRODUCTION

Channel parameters in an optical communication system such chromatic dispersion (CD), polarization-mode dispersion (PMD) and polarization-dependent loss (PDL) are needed to be estimated because such parameters monitoring serves information of channel quality and also helps network management functionalities such as impairment-aware-routing [1]. For long-haul high-speed communication systems, estimation of PMD up to second-order is essential.

Recent advances in coherent receivers allow for fiber transmission impairments such as CD, PMD and PDL to be mitigated and estimated by using digital signal processing (DSP) techniques. Several efforts for multi-impairment estimation from the adaptive finite-impulse-response (FIR) filter based equalizer have already been demonstrated [2-4]. As to the second-order PMD, it is estimated either by using training sequences [5] or blindly using adaptive FIR filter [6]. However, training sequences reduces transmission efficiency. On the other hand, the key requirement of FIR based estimation is that the filter delay-tap length should be long enough to equalize all the linear impairments. However, with increasing number of delay taps, the computational complexity of the time-domain FIR filters increases sharply [7]-[8].

To overcome such difficulties, in this paper, second-order PMD estimation from adaptive frequency-domain equalizer (FDE) is proposed. The adaptive FDE can reduce the computational cost significantly relying on block-by-block processing and fast implementation of discrete Fourier transform (DFT) [9]. In fact, the complexity in terms of required number of complex multiplications for adaptive TDE is in the order of $N^2$ while that for FDE is only *log* ($N$) where $N$ is the number of taps. Thus, adaptive FDE is a good choice for multi-impairment estimation in digital coherent receivers. Effectiveness of the proposed estimation scheme is validated with 100-Gbits/s polarization-division multiplexed QPSK transmission experiments and estimation results are compared with those from time-domain adaptive FIR case.

## II. ADAPTIVE FDE

Employing the overlap-save method, adaptive FDE can be constructed by calculating the gradient vector in the time domain. To allow an adaptive FDE that can work on two-fold oversample-sampled input sequence, as shown in Fig.1, it is configured by even- and odd- sub-equalizer concept where each sub-equalizer operates on symbol-spaced input sequence [8].

Consider that the incoming optical signal after homodyne coherent detection is sampled with ADCs at the rate of twice symbol rate. Then, the two-fold oversampled sequence $u_{x,y}(n)$ are sub-divided into two even and odd sequences as

$$\mathbf{u}^{e}_{x,y}(m) = [u_{x,y}(2m), u_{x,y}(2m-2), \cdots, u_{x,y}(2m-N)]^{\mathrm{T}}, \quad (1)$$

$$\mathbf{u}^{o}_{x,y}(m) = [u_{x,y}(2m+1), u_{x,y}(2m-1), \cdots, u_{x,y}(2m-N+1)]^{\mathrm{T}}. \quad (2)$$

where, $m$ is the symbol-spaced sampling time index related to $n$ as $n$=2$m$+1 ($m$=0,1,2$\cdots$). Consider that each filter length is $N$ and tap-weight vectors are given as

$$\mathbf{h}_{pq}(m) = [h_{pq,0}(m), h_{pq,1}(m), \cdots h_{pq,N-1}(m)], \quad (3)$$

where $p$ and $q$ are either $x$ or $y$. Correspondingly, the tap coefficient vectors are partitioned into even and odd sub-equalizer coefficients of length $L$=$N/2$ and can be written as

$$\mathbf{h}^{e}_{pq}(m) = [h_{pq,0}(m), h_{pq,2}(m), \cdots h_{pq,2L-2}(n)], \quad (4)$$

$$\mathbf{h}^{o}_{pq}(m) = [h_{pq,1}(m), h_{pq,3}(m), \cdots h_{pq,2L-1}(m)]. \quad (5)$$

Considering 50% overlapping factor, frequency domain input vectors, $\mathbf{U}^{e,o}_{x,y}(k)$ of length $2L$ ($L$ samples from current block and $L$ samples from previous block) for the sub-equalizers can be written as

$$\mathbf{U}^{e,o}_{x,y}(k) = \mathrm{FFT}[u^{e,o}_{x,y}(kL-L) \cdots u^{e,o}_{x,y}(kL+L-1)], \quad (6)$$

where $k$ is block index and relates to $m$ as $m = kL + i$ ($i = 0, 1, \cdots, L-1$). Then, the $L$ tap-weights of the sub-equalizers are padded with $L$ zeros and $2L$ point *FFT* is

performed. Let $\mathbf{H}_{pq}^{e,o}(k)$ is the frequency-domain coefficients vector of zero padded tap-weight vector **h**(k) which is calculated as

$$\mathbf{H}_{pq}^{e,o}(k) = \mathrm{FFT}[\mathbf{h}_{pq}^{e,o}(k); \mathbf{O}_M]. \tag{7}$$

Then, by using overlap-save method the time-domain output vector of length $L$ can be found as

$$\mathbf{v}_{x,y}(k) = \text{last } L \text{ elements of IFFT}\{\mathbf{V}_{x,y}(k)\}, \tag{8}$$

where,

$$\begin{aligned}\mathbf{V}_x(k) &= \mathbf{H}_{xx}^e(k)\otimes\mathbf{U}_x^e(k) + \mathbf{H}_{xx}^o(k)\otimes\mathbf{U}_x^o(k) \\ &+ \mathbf{H}_{xy}^e(k)\otimes\mathbf{U}_y^e(k) + \mathbf{H}_{xy}^o(k)\otimes\mathbf{U}_y^o(k),\end{aligned} \tag{9}$$

$$\begin{aligned}\mathbf{V}_y(k) &= \mathbf{H}_{yx}^e(k)\otimes\mathbf{U}_x^e(k) + \mathbf{H}_{yx}^o(k)\otimes\mathbf{U}_x^o(k) \\ &+ \mathbf{H}_{yy}^e(k)\otimes\mathbf{U}_y^e(k) + \mathbf{H}_{yy}^o(k)\otimes\mathbf{U}_y^o(k).\end{aligned} \tag{10}$$

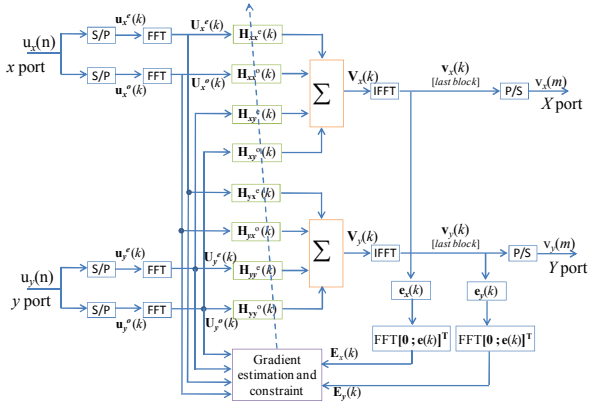



Fig.1: Schematic of adaptive FDE enabling two-fold oversampled input sequences for dual-polarization transmission systems

The error signal for CMA is then calculated in time-domain as

$$\mathbf{e}_{x,y}(k) = [1 - \mathbf{v}_{x,y}(k)\otimes\mathrm{conj}\{\mathbf{v}_{x,y}(k)\}]\otimes\mathbf{v}_{x,y}(k). \tag{11}$$

Then augmenting $\mathbf{e}_{x,y}(k)$ with $L$ zeros, we convert it to the frequency-domain vector with the column length of $2L$ as

$$\mathbf{E}_{x,y}(k) = \mathrm{FFT}[\mathbf{O}_L; \mathbf{e}_{x,y}(k)]^{\mathrm{T}}. \tag{12}$$

Applying the overlap-save method, we calculate the gradient vector $\nabla_{pq}^{e,o}(k)$ as

$$\nabla_{pq}^{e,o}(k) = \text{first } L \text{ terms of IFFT}[\mathbf{E}_p \otimes \mathrm{conj}\{\mathbf{U}_q^{e,o}(k)\}]^{\mathrm{T}}. \tag{13}$$

Finally, tap weights are updated in the frequency domain by using the gradient decent algorithm as

$$\mathbf{H}_{pq}^{e,o}(k+1) = \mathbf{H}_{pq}^{e,o}(k) + \mu\,\mathrm{FFT}[\nabla_{pq}^{e,o}(k); \mathbf{O}_L]^{\mathrm{T}}, \tag{14}$$

where $\mu$ is the step-size parameter.

## III. Second-Order PMD Estimation Algorithm

After the convergence of the filters, the impulse responses of eight sub-equalizers can be found as

$$\mathbf{h}_{pq}^{e,o}(k) = \mathrm{IFFT}\{\mathbf{H}_{pq}^{e,o}(k)\}. \tag{15}$$

Then we can construct the monitoring matrix as

$$\mathbf{M}(\omega) = \left\{ DFT \begin{bmatrix} \mathbf{h}_{xx}(k) & \mathbf{h}_{xy}(k) \\ \mathbf{h}_{yx}(k) & \mathbf{h}_{yy}(k) \end{bmatrix} \right\}^{-1}. \tag{16}$$

where the elements of the matrix are calculated as

$$\begin{aligned}\mathbf{h}_{pq}(k) = [&h_{pq,0}^e(k), h_{pq,0}^o(k), h_{pq,1}^e(k), h_{pq,1}^o(k)\cdots \\ &h_{pq,L-1}^e(k), h_{pq,L-1}^o(k)]^T\end{aligned} \tag{17}$$

The monitoring matrix is essentially the inverse of the channel transfer function. Therefore, the eigen values of the matrix $\mathbf{M}(\omega+\Delta\omega)\mathbf{M}^{-1}(\omega)$ , $\rho_{1,2}$, are associated with group delays of the two principle state of polarization (PSP) [2]. The first-order PMD or differential group delay (DGD), $\Delta\tau$ can be estimated as

$$\Delta\tau(\omega+\Delta\omega/2) = \left|\frac{\arg(\rho_1/\rho_2)}{\Delta\omega}\right|. \tag{18}$$

The necessary condition to avoid ambiguities arising from the multi-valued argument function is $\Delta\tau\Delta\omega < \pi$.

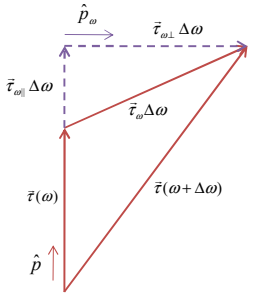



Fig.2: Definition of second-order PMD vector. Second-order PMD vector is shown with its perpendicular and parallel components for two PMD vectors at angular frequencies $\omega$ and $\omega+\Delta\omega$. This is given as the derivative of the first-order PMD vector with respect to the angular frequency $\omega$.

As shown in Fig. 2, the second-order PMD vector is described by the derivative of the fiber PMD vector $\vec{\tau} = \Delta\tau\hat{p}$ and can be written as

$$\vec{\tau}_\omega = \hat{p}\Delta\tau_\omega + \Delta\tau\hat{p}_\omega, \tag{19}$$

where $\hat{p}$ is the unit vector pointing the direction of the slow PSP. The magnitude of the first component $|\Delta\tau_\omega|$ is the change of DGD with the angular frequency, causing polarization-dependent chromatic dispersion (PCD). Since frequency-dependent DGD can be estimated from Eq. (18), PCD can be calculated as

$$PCD = |\Delta\tau_\omega| = \left|\frac{\Delta\tau(\omega_1) - \Delta\tau(\omega_2)}{\omega_2 - \omega_1}\right|. \quad (20)$$

The second-term $\Delta\tau\hat{p}_\omega$ relates to the PSP depolarization, that is, a rotation of the PSP with the angular frequency. The eigen vectors of the matrix $\mathbf{M}(\omega+\Delta\omega)\mathbf{M}^{-1}(\omega)$ locate the PSP as a function of $\omega$. Consider the eigen vectors are $|t_{1,2}\rangle$ and $\boldsymbol{S}$ is the Stokes vector corresponding to the slow PSP $|t_1\rangle$ ; then the unit vector $\hat{p}$ can be found as $\hat{p} = S/|S|$. The angular rate of PSP rotation $|\hat{p}_\omega|$ from $\omega_1$ to $\omega_2$ can be found as

$$|\hat{p}_\omega| = \frac{\cos^{-1}\{\hat{p}(\omega_2).\hat{p}(\omega_1)\}}{\omega_2 - \omega_1}. \quad (21)$$

Therefore, the degree of depolarization (DEP) is calculated as

$$DEP = |\hat{p}_\omega|\Delta\tau. \quad (22)$$

Finally, the magnitude of total second-order PMD can be estimated as

$$|\vec{\tau}_\omega| = \sqrt{PCD^2 + DEP^2}. \quad (23)$$

## IV. Experimental Results and Discussions

### A. Experimental Setup

To verify the proposed algorithm, we conduct second-order PMD monitoring experiments. We use a commercial all-order PMD emulator (PMDE) between the 100-Gbit/s dual-polarization transmitters and a coherent optical receiver. The transmitter as well as local oscillator lasers are distributed-feedback laser diodes (DFB-LD), each having a 3-dB linewidth of 150 kHz and a center wavelength of 1552 nm. A 50-Gbit/s NRZ QPSK signal is generated using a $LiNbO_3$ optical IQ modulator (IQM) from two streams of pre-coded data from a pulse-pattern generator (PPG) with $2^9$-1 pseudo-random binary sequences (PRBS). A 100-Gbit/s dual-polarization signal is then produced in the split-delay-combine manner. Then the signal passed through the PMDE. The PMDE uses three programmable DGD sections separated by polarization controller to generate all-order PMD with tunable statistics. The PMDE is set to generate a Maxwellian-distributed DGD with the mean value of 35 ps and a corresponding second-order PMD with a refresh rate of 10 ms. The signal is then pre-amplified by an erbium-doped fiber amplifier (EDFA) and incident on a coherent optical receiver employing phase and polarization diversities. The received power is fixed so that BER is around $3\times10^{-4}$. The outputs are sampled and digitized at 50 GSample/s with analog-to-digital converters (ADCs), and stored for offline DSP.

In the DSP circuit, sampling-phase adjustment, polarization demultiplexing, and signal equalization are done simultaneously either by the adaptive FDE or by the conventional adaptive FIR filters, where CMA adapted filter-tap weights. In both cases, the singularity problem [10] inherent in the CMA is handled by introducing the training mode prior to the blind CMA mode [11]. The delay-tap length for the FIR is $N$=32 and the block length of each sub-equalizer for the FDE is $N/2$=16. The step-size parameter for both cases is $2^{-10}$. After the convergence, filter tap coefficients are used for second-order PMD estimation with the algorithm described in the previous section.

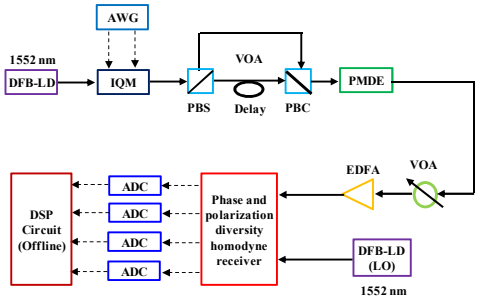



Fig.3: Schematics of experimental setup to verify the proposed second-order PMD estimation scheme

### B. Results and Discussions

If $B$ is the symbol rate, the spectrum of the transfer function of a $T/2$-spaced filter covers the range from $-B$ to $+B$. However, the parameter estimation is concentrated to several center taps only where low-pass electrical filters in the transmitter and the receiver have the linear phase response.

To verify the proposed second-order PMD estimation scheme, we estimate the statistical behavior of second-order PMD and compare it with the theoretical density as described in [12]-[13].

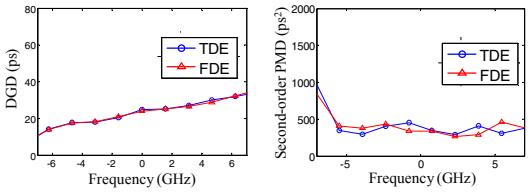



Fig. 4: Measured DGD and second-order PMD spectrum from an arbitrary sample

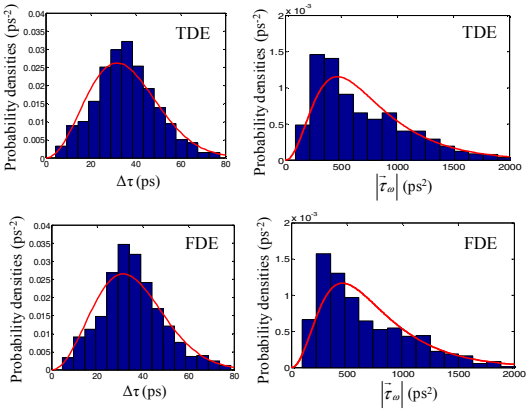



Fig.5: Probability densities of the first- and second-order PMD. Bars show those estimated from monitored values and solid curves are theoretical ones.

Figure 4 shows the estimated spectrum of DGD and second-order PMD. For both the cases, the estimated value from adaptive FIR and FDE are almost same. From such spectrum, we count the DGD and second-order PMD at centre frequency from 700 samples and calculate their probability densities as shown in Fig. 5. The measured probability densities for all the cases match well with the theory. The mean DGD estimated from the FIR is 35.64 ps while that from FDE is 35.36 ps; both values are very close to the set value of 35 ps. As of the second-order PMD, the mean value estimated from FIR is 627 $ps^2$ while that from FDE is 616 $ps^2$ which are about 20% less than the predicted mean value. The dissimilarity between the estimated second-order PMD and the theoretical prediction may stem from the inadequate number of statistical samples and using only three DGD sections in our PMDE which is insufficient to generate the second-order PMD similar to that of the real fiber.

## V. CONCLUSIONS

A novel approach of second-order PMD estimation from adaptive FDE is proposed and experimentally verified. The estimation performance is found similar to that from adaptive FIR filter; however, proposed concept has less computational complexity.

## REFERENCES


[1] D.C. Kilper, R. Bach, D. J. Blumenthal, D. Einstein, T. Landolsi, L. Ostar, M. Preiss and A. E. Willner, "Optical performance monitoring," *J. Lightwave Technol.*, vol.22, no.1, pp.294-304, Jan. 2004.

[2] Md. S. Faruk, Y. Mori, C. Zhang, K. Igarashi, and K. Kikuchi, "Multi-impairment monitoring from adaptive finite-impulse-response filters in a digital coherent receiver," *Opt. Express*, vol.18, no.26, pp. 26929-26936, Dec. 2010.

[3] F. N. Hauske, M. Kuschnerov, B. Spinnler, and B. Lankl, "Optical performance monitoring in digital coherent receivers," *J. Lightwave Technol.*, vol.27, no.16, pp.3623-2631, Aug. 2009.

[4] J. C. Geyer, F. N. Hauske, C. R. S. Fludger, T. Duthel, C. Schulien, M. Kuschnerov, K. Piyawanno, D. van den Borne, E.-D. Schmidt, B. Spinnler, B. Lankl, and B. Schmauss, "Channel parameter estimation for polarization diverse coherent receives," *IEEE Photon. Technol. Lett.*, vol.20, no.10, pp.776-778, May 2008.

[5] C. Do, A.V. Tran, T. Anderson, and E. Skafidas, "Data-aided second-order polarization-mode dispersion estimation for QPSK and 16-QAM coherent optical systems," in *Proc. Opto-Electron. and Commun. Conf.,* Jul. 2013, paper WR3-3.

[6] Md. S. Faruk, Y. Mori, C. Zhang, K. Igarashi, and K. Kikuchi, "Second-order PMD monitoring from adaptive FIR-filter tap coefficients in a digital coherent receiver," in *Proc. OFC,* Mar. 2011, paper OWN3.

[7] B. Spinnler, "Equalizer design and complexity for digital coherent receivers," *IEEE J. Sel. Topics Quantum Electron.*, vol.16, no.5, pp.1180-1192, Sept. 2010.

[8] Md. S. Faruk and K. Kikuchi, "Adaptive frequency-domain equalization in digital coherent receivers," *Opt. Express*, vol.19, no.13, pp.12789-12798, Jun. 2011.

[9] S. Haykin, *Adaptive Filter Theory,* Prentice Hall, 2001.

[10] K. Kikuchi, "Performance analyses of polarization demultiplexing based on constant-modulus algorithm in digital coherent optical receivers," *Opt. Express*, vol.**19**, no.10, pp. 9868-9880, May 2011.

[11] Md. S. Faruk, Y. Mori, C. Zhang, and K. Kikuchi, "Proper polarization demultiplexing in coherent optical receiver using constant modulus algorithm with training mode," in *Proc. Opto-Electron. and Commun. Conf.,* Jul. 2010, pp.768-769.

[12] G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Probability densities of second-order polarization mode dispersion including polarization dependent chromatic fiber dispersion," *IEEE photon. Technol. Lett*., vol. 12, no. 3, pp. 293-295, Mar. 2000.

[13] G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Statistics of second-order PMD depolarization," *J. Lightwave Technol.,* vol. 19, no. 12, pp. 1882-1886, Dec. 2001.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Session-Based Improved Test Case Generation Using Genetic Algorithm

Arif Mahmud[1,3], Fardina Fathmiul Alam[1,2,4], and Muhammad Masroor Ali[2,5]
[1]Military Institute of Science and Technology, Bangladesh
[2]Bangladesh University of Engineering and Technology, Bangladesh
* E-mail: [3]arif_0964@yahoo.com , [4]dreamy.eye30@gmail.com , [5]mmasroorali@cse.buet.ac.bd

*Abstract*— **Generation of test data is a central theme in testing field. One of the major testing methodologies is user-session-based testing. The main goal of user-session-based testing is to capture and replay real user sessions. It is an acceptable method to test web applications. However, as a black-box testing, test case generation totally based on user session data may not be qualified for ensuring the reliability of web applications. We discuss here an approach where user session data with Request Dependence Graph (RDG) of web application have been combined, to generate test cases automatically. Based on an existing Genetic Algorithm (GA) heuristics, we propose an approach which is more effective. A model of Request Dependency Graph (RDG) is first constructed according to the structural analysis of the application under test; and then transition relations between pages and requests are extracted based on the request dependence graph (RDG). Our experimental result has indicated that our approach is better than the aforementioned existing user-session-based testing algorithm.**

***Keywords—Genetic Algorithm (GA), User Session based Request Dependency Graph (US-RDG), Request Dependency Graph (RDG)***

## I. INTRODUCTION

Web service [20] is a software system that is designed to support interoperable machine to machine interaction over a network. Web services which are Web APIs can be accessed over a network, such as the Internet, and executed on a remote system hosting the requested services.

Web services hold the promise of integrating software applications from heterogeneous networks and exchanging information in a simple, standardized manner. Now-a-day, majority of companies are moving to services-oriented architectures (SOAs) i.e. a web page or a website and manipulating web services within and across their IT infrastructure.

Different integrations and innovations have been made possible by web services. Web services have also determined but how these services affect the quality and performance of the mission-critical applications with which they interface.

As such, it is crucial to thoroughly test Web services before they are deployed in order to ensure service level compliance in production. XML is used to encode all communications in a web service. Web Services, as self-contained, modular, distributed, dynamic applications, can be described, published, located, or invoked over the network to create products, processes, and supply chains. These applications can be local, distributed, or Web-based.

Web services are built on top of open standards such as TCP/IP, HTTP, Java, HTML, and XML. Various web service development options like PHP web development, ASP, .NET, web applications, computer languages, programming and web application platforms all together makes online presence of a particular website over the net through server networking, hosting and other web services.

With the increasing sophistication of software applications and the expanding role of database, testers require high volumes of high quality, realistic test data that can faithfully represent existing, and stress-test new, platforms. So, realistic data must be correctly formed, according to structural rules and must also have semantics that represent real-world entities. If an automated tool generates many unusual and endemic inputs that a user is unlikely to use, then the tester will soon abandon the tool or disregard its findings.

Session-based test case generation [1], [2] is a mechanism of replaying the user session to ensure the relational transition for a web application. To remove the redundancy problem in user – session Genetic Algorithm (GA) heuristic has been applied in [1], where two point crossover and mutation property of GA have been used to increase the transition coverage number in a web application for a user – session. Genetic Algorithm (GA) heuristic actually consists of two basic properties, crossover and mutation. There are many ways how to perform crossover and mutation. Worth mentioning ones are, single point crossover, two point crossover, selective crossover, uniform crossover for crossover property as well as order change mutation, changing operator mutation, adding mutation for mutation property. Our contribution in this paper lies in generating more effective session based test case with a new approach over [1] of GA heuristic where we have used selective crossover and adding mutation property. We have successfully increased the transition coverage without participating in a lengthy and time consuming calculation. This improvement has been evidenced by experimental results.

The organization of rest of the paper is as follows. In section II, we describe the most relevant works along with their limitations. Then in section III, we have described the setup of our proposed approach using some steps, which is largely

adapted from [1]. After that, we have shown experimental result of our proposed approach in section IV based on an open source online Book Store which is available at gotocode.com. Finally in section V, we give a conclusion based on our approach.

## II. Related Works

User-session-based testing [1], [2] has been researched recently to make use of user session data to generate test cases from a functional standpoint. Session information transparently collects user interactions, and is stored in a server log.

A user session is composed of a sequence of user requests. A request is generally in the form of user IP, timestamp, request pattern (GET/POST), URL, parameter-value pairs, data transmission protocol, etc. User-session based testing automatically generates test cases based on user-session. By capturing and replaying user sessions, this technique reruns real user induced events to meet the functional requirements of web application. When compared with traditional white box testing, user-session-based testing reduces much human effort of manual test case generation. Additionally, test cases generated from user sessions are more representative of the real field usage application.

That's why, several works have been done on web testing area and some of them emphasizes on user-session based testing and test case generation, as thus are likely to detect faults effectively. Works on automated test case generation using Genetic Algorithm (GA) heuristic have remarkable importance on web testing era.

However, some other effective works have been done though they have some limitations. Some approaches in [1], [2] have been proposed to generate test data relying on the structure analysis of web application in terms of traditional white-box testing.

An approach has proposed in [19] to create a model based on Unified Modeling Language (UML), in which path expression was created to generate test cases for web application. Web applications have been partitioned into logical components (LCs) in [2] at different levels of abstraction.

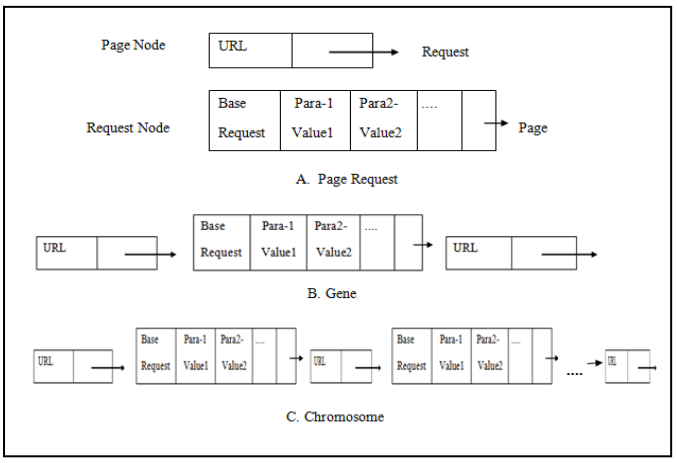



Figure 1. Genetic Encoding and Chromosome (reproduced from [1])

To partition into different abstraction, it used Pages-Flow-Diagram (PFD) and then used an automaton to model those LCs, with which test cases could be generated automatically.

But generating test case manually, changeability of web structure, randomness of web application result, deploying environment and some volatile issue have put some limitations on these works.

In this paper, a user-session is considered as a chromosome as like Fig. 1. Whole web application has been drawn as a User-Session Request Dependency Graph (US-RDG). Then they approached for Genetic Algorithm (GA) heuristics on each chromosome.

For generating optimal test cases using Genetic Algorithm (GA) heuristic on the basis of Gray-Box testing an approach in [1] has been made.

Though it covers most of the previously mentioned problems, still it has some limitations. Specifically, performances of Genetic Algorithm (GA) heuristics depend on them very much. The type and implementation of operators depends on the encoding and also on the problem. In particular the problems are as follows,

1. In the Crossover operation they used two-point crossover method and considered all common nodes between two user-sessions for crossover as shown in Fig. 2. There are two common nodes and after crossover four offsprings (user-session) will be generated and fitness of each user-session will be calculated to decide whether it will be accommodated in the test-suite or not. But what will happen if there exists more common nodes? This is very usual for a user-session and for 10 common nodes we will get more than 20 new user-sessions. For each of these, we have to calculate the fitness for the decision whether it will be discarded or collected, which is definitely cumbersome and time consuming.

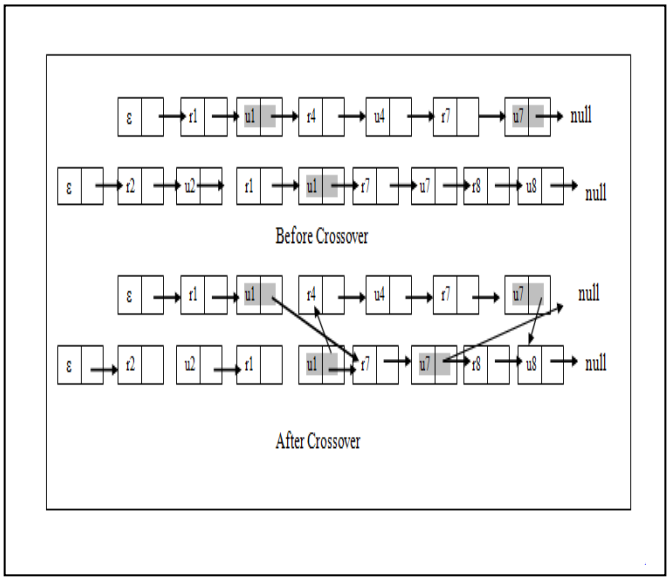



Figure 2. Crossover Process (reproduced from [1])

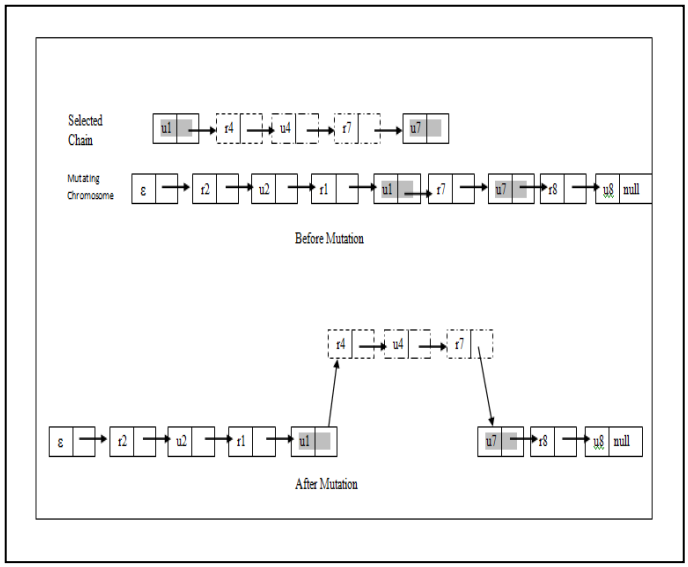



Figure 3. Mutation Process (reproduced from [1])

2. In mutation, they have selected a random chain from user-sessions to change the request node in order to increase the fitness of a user-session as shown in Fig. 3. But if a user-session has more than one identical starting and ending nodes similar to the selected chain, then according to the process, the chain will be inserted to all the identical nodes. As a result it will create redundant tests for replaying a user-session. That is why; fitness of that user-session will be demurred.

We have tried to overcome these problems and generated more effective user – session for replaying to test web services. Detailed description of our methodology is produced in the following section.

## III. PRILIMINARIES

The setup described below is adapted from [1] with necessary modifications to incorporate our proposed improvements.

### *A. Constructing User Session based Request Dependency Graph (US-RDG)*

It's easy to depict a web application dependency relationship by a graph. There exists a request dependence relationship between pages, and this relationship can be attained from the developers, that is, a specification document or source code analysis of the application. A web page A is request dependent on a web page B if an execution of B will trigger a request for A. To depict the request dependence relationship clearly, a Request Dependence Graph (RDG) has been constructed in [1]. In this graph, a node represents a web page, a directed edge represents a request dependence relationship between two pages, and the request itself is presented to identify each directed edge. Fig. 4 illustrates Partial Request Dependence Graph. Though it may seem a little awkward to reproduce these 4 figures from another paper, but we think that in order to create a solid impression on the present scenario inclusion of these figures is important.

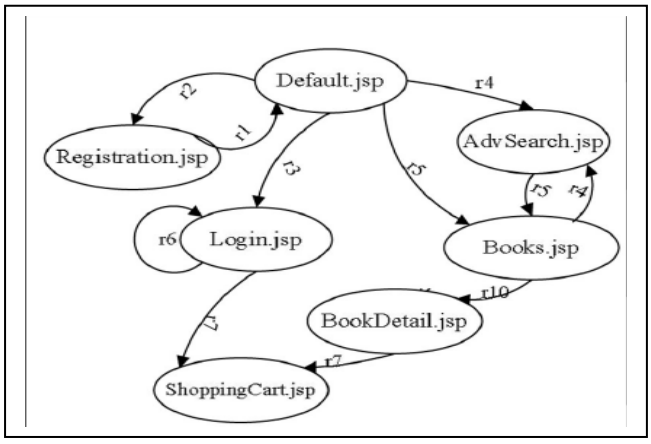



Figure 4. Partial Request Dependence Graph (reproduced from [1])

### *B. Selecting Effective Population Generating and Genetic Coding*

In our improved methodology over [1], to encode chromosome, pre-analysis is carried out over the request sequence of each user session. Before constructing user session based request dependency graph (US-RDG), we will make sure that there is no existence of duplicate user-session (exactly identical user-session) in the total collection of user-sessions. We will consider particularly for CDTR and CLTR of each chromosome, where CDTR and CLTR separately denotes the number of data and link dependence transition relations covered in the chromosome. We will rearrange the request nodes of a user-session to check whether it is same as other in the total collection or not. After that we will get a collection of user-session without duplication. We will consider the first page node of a chromosome with Uniform Resource Location (URL) □ since the first request of a user session has no previous request to fix a page which triggers it.

### *C. Fitness Function and Selection Strategy*

In [1], CDTR has been given more importance over CLTR using a coefficient. But we want to put same emphasis on both, because, in present days, a substantial number of web applications are used through mobile phones or other smart devices. In these cases, link transitions are important as the data over a User Interface (UI) to create an effect on another UI.

We evaluate fitness of each chromosome based on the transition relation coverage analysis. For a chromosome, we calculate how many transition relations were covered in it. In addition, we introduce the dependence property of each covered transition relation into fitness calculation, because there is difference between data dependence transition relation and link dependence transition relation in fault detection.

Thus we enhanced the fitness function by using some factors.

$$Fitness \quad value = (CDTR + CLTR)/(DTR + LTR) \quad (1)$$

CDTR and CLTR separately denote the number of data transition relations and the number of link transition relations, covered by a chromosome. DTR and LTR separately denote the number of data transition relations and the number of link transitions relations existing in the application which are attained from the structural analysis in request dependence graph construction. From the fitness function, we can see that the highest fitness value is achieved as 1 when a chromosome covers all the data dependence transition relations and link dependence transition relations existing in the application.

### *D. Crossover*

In our proposed method, in contrast with two point crossover in [1], we will use selective crossover method for the crossover property. At first, we will measure a primary fitness value for each chromosome in our collection. Then we will take the highest fitness value possessing two chromosomes as Parent Chromosome 1 and Parent Chromosome 2 for crossover operation and we will ensure single point crossover for these two chromosomes as shown in Fig. 5.

Then two offsprings will be generated and we will calculate the fitness value for both. If the new fitness value is greater than the fitness value of any other chromosome in the collection, then the new one will replace it. We will always look for the highest fitness value. Once a crossover operation is finished, the collection will be resorted as per the fitness value of each chromosome.

In [1], they considered two parent chromosomes randomly from the collection of user sessions and two point crossover. After a successful crossover operation, it cannot be deterministically determined whether the fitness value of the newly created offspring will increase or not as a result of random selection of parent chromosome and producing useless offsprings. We considered the highest fitness value possessing two chromosomes as parent chromosome. After a single point crossover there is a probability for increasing the fitness value for the newly created offspring and so far iteratively fitness value is increasing.

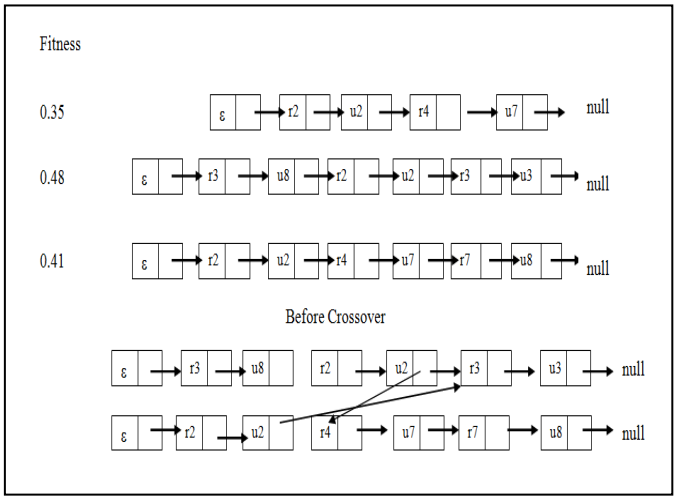



Figure 5. Proposed Crossover Process

As in our methodology, there is a very little chance for producing useless offspring, so the probability of increasing the fitness value in each iteration for a newly created offspring is high and within a few iteration the fitness value will reach to a considerable level. Therefore, the more the fitness value, the more improvement in test-suite. So it is better to select the highest fitness valued chromosomes as parent chromosome than selecting two parent chromosomes randomly.

Here, we reduce the probability of generating useless offsprings and decrease the calculation of fitness in every step of crossover between two parent chromosomes.

### *E. Mutation*

In mutation, as elaborated in section II, [1] suffers from the problem of inserting redundant chains. In order to circumvent this problem, we will randomly select a chain from a chromosome whose presence is comparatively less among the chromosomes and whose starting node has a match with the last node of the chromosome which we have got after the crossover operation. Thus it will accelerate the probability of creating a new offspring which covers more transitions.

Fig. 6 shows the mutation process. In the mutation process, a mutation probability is first predefined, and for each chromosome a mutation score is randomly generated to compare with the mutation probability to decide whether or not a mutation will be processed. After that, the common transition relations covered between the mutating chromosome and the selected chain are counted, and the common ratio in the selected chain is calculated. If the common ratio is smaller than a predefined common threshold, then mutation process will be carried out.

Otherwise, another chain will be selected to be judged as the foregoing one, until the mutation process is successfully finished or the iteration count reaches a limiting value. Here we have decreased the redundancy probability of a chain within a chromosome and unnecessary calculations

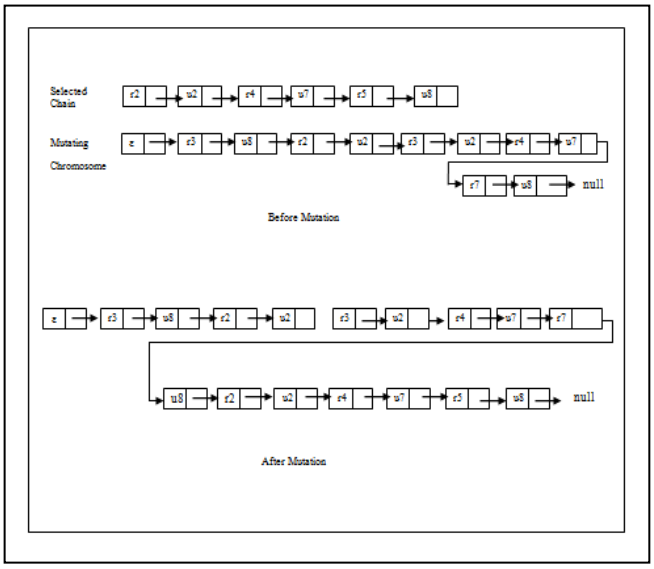



Figure 6. Proposed Mutation Process

### F. *Acceptance and Replacement*

For the fact that the aforementioned two processes of crossover and mutation are fraught with uncertainty due to GA heuristics, it is not sure that the generated offsprings are superior to their parents. We will recalculate the fitness values of the new ones with the fitness function in step B; and then their fitness values will be compared with that of the parents, after which two chromosomes with relatively high fitness value will be selected and placed in a new population. This process assures that in any case, the optimal chromosome is inherited. The new population is used for the next iteration of the heuristic.

### G. *Stop*

If the average fitness value reaches a steady state in recent several populations or a predefined maximum generation has been achieved, the mutation should be stopped, and the best solution should be returned in the current population.

### H. *Loop*

Go to step C (Fitness Function and Selection Strategy).

## IV. EXPERIMENTAL RESULT

In order to substantiate our theoretical claim in an actual scenario, we conducted the empirical study on the same web application (an open source online book store available at gotocode.com) as [1] had used to check the performance of user session based request dependency graph (US-RDG) and effectiveness of the generated test suite. The online book store allows users to register, sign in/out, search books, order books, edit shopping cart and edit user profile. But we have collected more number of user-sessions than [1] and applied our methodology on them. Details of our experiment have been illustrated in Table 1.

The graphs in Fig. 7 and Fig. 8, present the results of our experiment. Our GA heuristic parameters were: (1) fitness percentage threshold: 0.92; (2) mutation probability: 0.25. Compare to [1], our results show improvement both in fitness value generation speed and final fitness value. After seventeen iterations, we have achieved 70% coverage of the application, where in [1], 70% coverage has been achieved after twenty three iterations. When we run more than seventeen iterations then we achieved 79% coverage successfully. As a consequence, we have been able to cover more transitions within a shorter time period and less number of iterations.

TABLE I. CHARACTERISTICS OF THE COLLECTED USER SESSIONS

| Characteristic | Values |
|---|---|
| Total number of user sessions | 97 |
| Total number of requests accessed | 1682 |
| Largest user session in the number of requests | 56 |
| Average user session in the number of requests | 14.4 |
| Number of unique requests covered (coverage percentage) | 15 (100%) |
| Number %of data dependence transition relations covered (coverage percentage) | 18 (100%) |
| Number of %link dependence transition relations covered (coverage percentage) | 29 (78.64%) |
| Number of transition relations covered (coverage percentages) | 47 (87.24) |

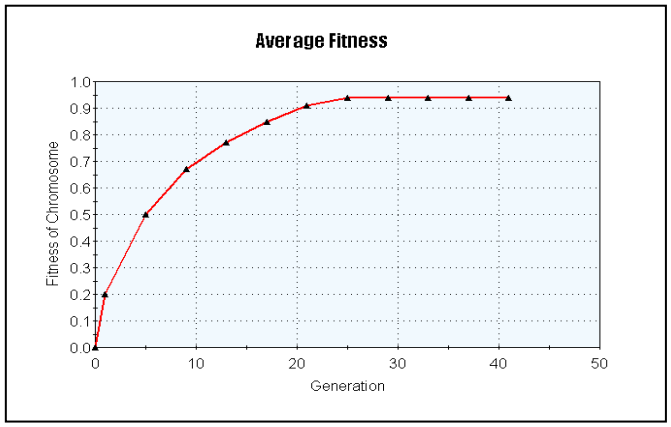



Figure 7. Generation Table for Mutation Process (Average Fitness)

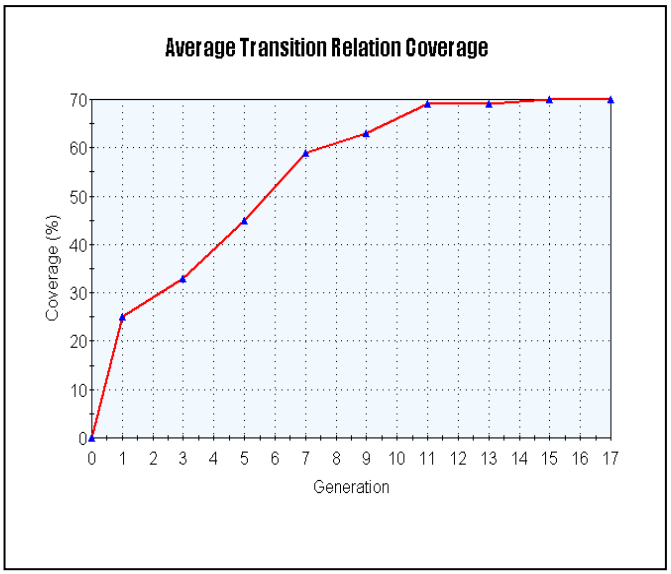



Figure 8. (A) Generation Table for Average Transition Relation Coverage

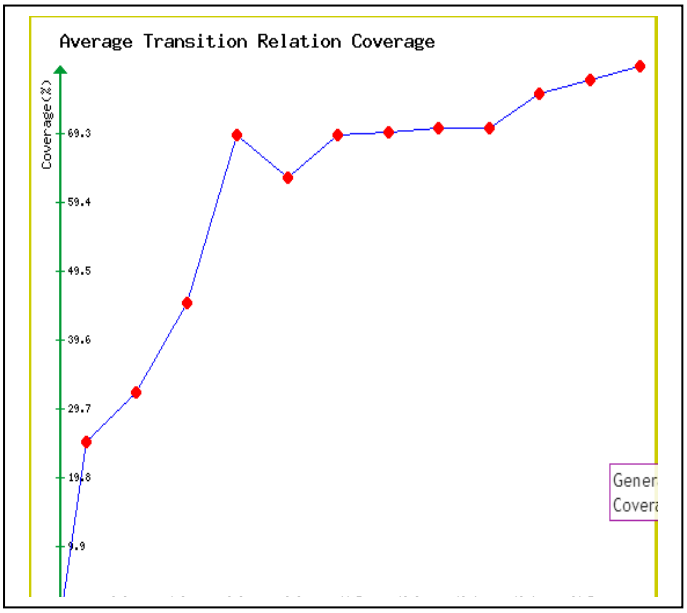



Figure 9. (B) Generation Table for Average Transition Relation Coverage

## V. CONCLUSION AND FUTURE WORKS

In our approach we have emphasized on the both data and link transitions. Selective crossover method has been applied to reduce the number of generating worthless or unnecessary offsprings and thus avoid the cumbersome calculation of fitness value. Though it is uncertain and risky for a small number of user-session but it carries remarkable output for the web applications those are used very effectively and create a large number of user-sessions. In mutation we have tried to apply a different process, adding chain in the end of a chromosome which has a large possibility of increasing the number of transition coverage and as well as the fitness value of the chromosome. Superiority of our proposed method has been evidenced by experimental results. In the future, we plan to apply Uniform Crossover method where nodes will be changed randomly between two parent chromosomes according to their common node to find some better solution over this methodology.

## REFERENCES


[1] X. Peng and L. Lu, "User-Session-Based Automatic Test Case Generation Using GA," in *Proc. International Journal of the Physical Sciences* ,vol. 6 ,no. 13, pp. 3232-3245, 4 July.2011.

[2] S. B. Chen, H. K. Miao, and Z. S. Qian, "Automatic Generating Test Cases for Testing Web Application," in *Proc Int. Conf. Comput. Intell. Secur. Workshops Harbin*, 2007, pp. 881-885.

[3] M. H. Chen, C. Song, X. M. Luo and X. Y. Zheng (2008), WebMTA homepage on CS.ALBANY. [Online]. Available: http:// www.cs .albany. edu/~mhc/WebMTA/docs/tool.pdf, 2008.

[4] Z. Y. Dai and M. H. Chen, "Automatic Test Case Generation for Multi-Tier Web Applications,"in *Proc. 9th IEEE Int. Workshop Web Site Evol., 2007,* pp. 39-43.

[5] D. G. A. Lucca, AR. Fasolino, and P. Tramontana P, "A Technique for Reducing User Session Data Sets in Web Application Testing," in *Proc. 8th IEEE Int. Symp*. 2006, pp. 7-13.

[6] C. Doungsa-ard, K. Dahal, Hossain , and T. Suwannasart, "Test Data Generation from UML State Machine Diagrams Using Gas," in *Proc. 2nd Int. Conf. Softw. Eng. Advances Cap Esterel*, 2007, pp. 47-52.

[7] S. Elbaum, S. Karre, and G. Rothermel, "Improving Web Application Testing with User Session Data," in *Proc. 25th Int. Software Eng. Portland*, 2003, pp. 49-59.

[8] S. Elbaum, S. Karre, E. Gibson, and L. Pollock,"An Empirical Comparison of Test Suite Reduction Techniques for User-Session-Based Testing of Web Applications," in *Proc. 21st IEEE Int. Conf.Softw. Mainten.*, 2005, pp.587-596.

[9] S. Elbaum, G. Rothermel, and S. Karre, "Leveraging User-Session Data to Support Web Application Testing," *IEEE Tran. Softw. Eng*., vol. 31, no. 3, pp.187-202, 2005.

[10] A. S. Ghiduk, M. J. Harrold, and M. R. Girgis, "Using Genetic Algorithms to Aid Test-Data Generation for Data-Flow Coverage," in *Proc 14th Asia-Pacific Software Eng. Conf.*, 2007, pp. 41-48.

[11] J. Hollan., *Adaptation in Natural and Artificial Systems*. CA: Uni. Michigan Press, Ann Arbor, USA, 66-72, 1975.

[12] A. S. Kalaji, R. M. Hierons, and S. Swift, "Generating Feasible Transition Paths for Testing from an Extended Finite State Machine," in *Proc Inte. Conf. Softw. Test. Verification Validation,* 2009, pp. 230-239.

[13] S. Khor and P. Grogono, "Using a Genetic Algorithm and Formal Concept Analysis to Generate Branch Coverage Test Data Automatically," in *Proc. 19th Int. Conf. Automated Softw*. 2004, pp. 346-349.

[14] X. M. Luo, F. Ping, and M. H. Chen, "Clustering and Tailoring User Session Data for Testing Web Applications," in *Proc 2nd Int. Conf. Softw. Test. Verification Validation,* 2009, pp. 336-345.

[15] H. K. Miao, Z. S Qian, and B. Song, "Towards Automatically Generating Test Paths for Web Application Testing," in *Proc 2nd IFIP/IEEEInte. Symp. Theor. Aspects Softw. Eng Nanjing, 200,* pp. 211-218.

[16] S. Sampath, and S. Sprenkle, "Applying Concept Analysis to User-Session-Based Testing of Web Applications." *IEEE Tran. Softw. Eng,* pp.643-658, 2007.

[17] S. A. Yevtushenko, *System* of Data Analysis "Concept Explorer" (in Russian)," in *Proc. 7th national conf. Artif. Intell.* , 2000, pp.127-134.

[18] P. Ammann and J. Offutt, *Introduction to Software Testing,* CA: China Machine Press, Beijing, China, pp. 246-267, 2008.

[19] F. Ricca and P. Tonella, "Analysis and Testing of Web Applications,"in *Proc. 23rd Int. Conf. Softw. Eng.,* 2001, pp. 25-34.

[20] The wikipedia website. [Online]. Available: http:// en.wikipedia.org /wiki/Web_service.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# An Approach to Develop an Effective Job Rotation Schedule by Using Genetic Algorithm

Pritom Kumar Mondal, A.M.M. Nazmul Ahsan*, and Kismot Abdul Quayum
Department of Industrial Engineering and Management, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh
* E-mail: ahsan.ipe@gmail.com

*Abstract*—**Job rotation is a significant approach of job design used in manufacturing, assembly or any service providing system, which requires the workers to move between different tasks, at fixed or irregular periods of time. Now a days it is increasingly employed in organizations because it benefits not only the workers by preventing musculoskeletal disorder, eliminating boredom and increasing job satisfaction but also the management by increasing productivity, creating morale and loyalty in workers and decreasing turnover rate. In this paper a job rotation schedule of a construction firm has been developed by using genetic algorithm to eliminate repetitive tasks for each of the workers. This genetic algorithm has been implemented as a multi-factor algorithm since it takes into account ergonomic, competence, environmental, and physical skill factors to develop an optimum job rotation schedule. Genetic operators- crossover and mutation are used to create new solutions in each generation maintaining some conditions. Finally a case study has been conducted to validate the proposed methodology.**

***Keywords—Genetic algorithm, Job rotation schedule, Ergonomic factor, Musculoskeletal (MSK) disorder.***

## I. INTRODUCTION

Flexibility at workplace is today's most common word all over the world. It's very crucial to provide a suitable working environment for workers and to achieve certain human resource objectives. Working with repetitive movement of same body part for a long time is dangerous to health and causes musculoskeletal (MSK) disorders that result into job boredom, job absenteeism and job turnover. To protect workers from MSK disorders and job boredom and to motivate workers towards workplace who are assigned to repetitive tasks, job designing through job rotation is a must. As a result, it ensures the highest productivity possible without sacrificing quality, service, or responsiveness. Job rotation scheduling problem is generally structured as a multi-period assignment model, which has a great deal of feasible rotations impossible to be enumerated and evaluated for optimality by mathematical programming approaches particularly for large-size problems [6]. Thus, a genetic algorithm model has been approached to schedule a job rotation.

Job rotation can be defined as working at different tasks or in different positions for set periods of time [5] in a planned way using lateral transfers aiming to allow employees to gain a range of knowledge, skills and competencies and is also seen as an on-the-job training technique [5]. Workers must do something completely different during rotation. Different tasks must engage different muscle groups in order to allow recovery for those already strained. The design of efficient rotation schedules is a complex problem due to the many factors to be considered for the assignment of workers to jobs in each rotation. For example, four factors have been considered in this research to develop a job rotation schedule. In previous works, job rotation scheduling problem was solved using integer programming [2, 7], heuristics [8], metaheuristics such as genetic algorithms [2] or by combining integer programming and metaheuristic models [3]. Tharmmaphornphilas proposed an integer programming model [7] and a method based on heuristics [8] for developing robust job rotation schedules to reduce the likelihood of low back injury due to lifting. Carnahan et al. [2] developed a genetic algorithm for the design of rotation schedules to prevent back injuries. John S. Dean [4] designed a job rotation schedule by implementing genetic algorithm staff-scheduling solutions for scheduling nurses at a hospital. A traditional bit-string chromosome structure and a two-dimensional array chromosome structure have been used as solution to represent each schedule. Boyd and Savory [1] applied a genetic algorithm for scheduling laboratory personnel and the genetic algorithm approach appears to be useful for scheduling in highly technical work environments that employ multi-skilled workers. In this study, the schedule is developed by a genetic algorithm which considers four types of factors: ergonomic, competence, environmental and physical skill factors. Ergonomic factor shows how frequently a body movement is required to accomplish a job and shows how much difficulties a worker is facing with this body movement. Work-related competences are the knowledge and skills that can be useful to accomplish the job. Environmental factor defines whether the environment has a suitable work condition or not. If the condition is not suitable then it determines how much bad it is. Physical skill factor refers to physical ability of the worker either he possess enough ability or not. It also determines how much difficulties a worker is facing of doing a specific job. No previous works has included all of these four factors. Here an important concept has come in light that the ability of worker cannot always be judged by the fact of having ability or no ability, the worker can have more or less ability than the required amount. This study has showed that a worker can be assigned to a job even though he possess less ability, initially it will cause more cost but with time the cost can be minimized.

## II. PROBLEM STATEMENT

Job rotation schedule is generally applied in all those fields where each worker possesses the ability to work in all the jobs of that particular field with or without small training. If there are a several number of workers & jobs and total working hour is divided in more than one shift which is called rotation, then it is possible to develop several number of job rotation schedules among which only one always gives the best result. All the workers have the training level required to perform the activities of the jobs selected. The working day is of 8 hours, with a break of 1 hour for lunch at midday. The first two rotations are 2 hours long and last two rotations are one and a half hours long, making the start of third rotation coincide with the end of lunch break. The first two rotations have an allowance time of 12 minutes and last two have an allowance of 9 minutes. If a body movement is done while performing a job in the earlier rotation, then it definitely affects that body movement in the following rotations. That means if a body movement is required in first rotation and again if it is required in second rotation, then that ergonomic factor or body movement in second rotation is surely to be influenced by the frequency of body movement required by the job in the first rotation. The scale of influence factor is shown in Table I. If there is no influence then the score becomes 1 and if totally influenced then the score becomes 0. The assignments of workers to different jobs should minimize the repetitive movements of the same body parts and should also consider the physical ability and competences of workers to do the jobs.

TABLE I. DIFFERENT INFLUENCE LEVELS AND CORRESPONDING SCORES

| Influence Level | Relative Score |
|---|---|
| No Influence | 1 |
| Very Small Influence | 0.8 |
| Relatively Small Influence | 0.6 |
| Relatively High Influence | 0.4 |
| Very High Influence | 0.2 |
| Totally Influenced | 0 |

All possible schedules are not to be granted, only those schedules are granted which satisfy three specific conditions which are– 'same worker can't be assigned on same job in two different rotations', 'same worker can't be assigned on two different jobs in a single rotation', and 'two distinct workers can't be assigned on same job in the same rotation'. There are different jobs in different fields. The jobs of construction firm which can be done by almost all the workers can be listed as brick, sand, and cement loading as well as unloading, washing and breaking bricks, bending and cutting rods, making mixtures, earth excavation, sand compaction, curing etc. The basic four types of factors which are ergonomic, competence, environmental and physical skill factors are same for all cases but the sub factors under them vary from field to field. Arm abduction, extension and flexion, neck extension, flexion and turning, hand turning, finger pinching etc. are the construction work related ergonomic factors. Company strategy and policy, self-control, motivation, learning skill etc. are relevant competence factors and high temperature and pressure, sunny and dusty weather, noisy working area etc. are relevant environmental factors. In this particular field, climbing (ladder or stair), exerting force during moving and working, working in height, taking bulk or piece load on head etc. can be listed as relevant physical ability.

## III. FITNESS ASSESSMENT MODEL FORMULATION

The job rotation model is formulated as a maximization problem and the algorithm determines the optimum job rotation schedule on the basis of the overall fitness value of the schedules. The schedule which has the highest fitness value is the most optimum one. To compute the overall fitness of a schedule it is needed to compute the fitness value corresponding to each of ergonomic, competence, Physical skill and environmental factors. To compute these fitness values, it is needed to judge the factors against jobs and workers. Different kinds of jobs involve different body movements. A particular job does not necessarily require all kind of body movements, besides it may require some movement quite frequently and some movements infrequently. A worker may face some difficulties in some specific body movements and there are also some movements which can be accomplished by workers without any sort of problem. For each ergonomic factor, each job is scaled from 0 to 3 depending upon the required frequency of the body movement and each worker is scaled from 0 to 5 depending upon the difficulty he is facing to accomplish that particular body movement. So the highest possible scale value for an ergonomic factor is 15 which means the job does not require that particular body movement and the worker is facing no difficulty. The value is 0 for opposite condition and it is the minimum value. For competence, environmental and physical skill factors, score for each factor is determined by subtracting the score of a selected job for that factor from the corresponding score of a selected worker. The scoring scale of job and worker are 0 to 1, and 0 to 4 respectively, and 0 to 2 for both competence, environmental, and physical skill factors. For competence, environmental and physical skill factors, the maximum scores are 1, 4, and 2 and the minimum scores are -1, -4 and -2 respectively. Table II and III shows the scaling system of jobs and workers for ergonomic factors respectively. Different scores for workers and jobs for competence factors are shown in Table IV and V respectively. Table VI and VII shows the scaling system of workers and jobs for environmental factors respectively. Scores for workers and jobs for physical factors are shown in Table VIII and IX respectively.

TABLE II. SCORING SCALE OF ERGONOMIC FACTORS FOR JOBS

| Frequency of movement/min | Description | Score |
|---|---|---|
| 10 | Very high frequency | 0 |
| 8-9 | High frequency | 1 |
| 5-7 | Medium frequency | 2 |
| 3-4 | Less frequency | 3 |
| 1-2 | Infrequent | 4 |
| 0 | Never | 5 |

TABLE III. SCORING SCALE OF ERGONOMIC FACTORS FOR WORKERS

| Limitation level | Score |
|---|---|
| High limitation | 0 |
| Medium limitation | 1 |
| Low limitation | 2 |
| Without limitation | 3 |

TABLE IV. SCORING SCALE OF COMPETENCE FACTORS FOR WORKERS

| Competence level | Score |
|---|---|
| Perfect | 1 |
| Very good | 0.9 |
| Pretty good | 0.8 |
| Relatively good | 0.7 |
| Rather good | 0.6 |
| Regular | 0.5 |
| Rather bad | 0.4 |
| Relatively bad | 0.3 |
| Pretty bad | 0.2 |
| Very bad | 0.1 |
| Worst | 0 |

TABLE V. SCORING SCALE OF COMPETENCE FACTORS FOR JOBS

| Required competence level | Score |
|---|---|
| Very High | 1 |
| Relatively high | 0.75 |
| Medium | 0.5 |
| Low | 0.25 |
| Negligible | 0 |

TABLE VI. SCORING SCALE OF ENVIRONMENTAL FACTORS FOR WORKERS

| Adaptation level | Score |
|---|---|
| Normal condition | 4 |
| Easy to adapt | 3 |
| Adaptable | 2 |
| Difficult to adapt | 1 |
| Impossible to adapt | 0 |

TABLE VII. SCORING SCALE OF ENVIRONMENTAL FACTORS FOR JOBS

| Required adaptation level | Score |
|---|---|
| Very High | 4 |
| Relatively high | 3 |
| Medium | 2 |
| Low | 1 |
| Negligible | 0 |

TABLE VIII. SCORING SCALE OF PHYSICAL SKILL FACTORS FOR WORKERS

| Difficulty level | Score |
|---|---|
| Very easy | 2 |
| Easy | 1.6 |
| Normal | 1.2 |
| Relatively difficult | 0.8 |
| Very difficult | 0.4 |
| Not possible | 0 |

TABLE IX. SCORING SCALE OF PHYSICAL SKILL FACTORS FOR JOBS

| Probable hardship level | Score |
|---|---|
| Very High | 2 |
| Relatively high | 1.5 |
| Medium | 1 |
| Low | 0.5 |
| Negligible | 0 |

### *A. The solution fitness model*

The first step to determine the solution fitness model is to calculate the fitness value for each factor. Equation 1 shows the way of calculating ergonomic fitness and the process is accomplished by multiplying the influence factor, score and actual working time together for all factors, worker, job and rotation. The assessments of competence, environmental, and physical fitness are accomplished by using Equation 2, 4, 6 respectively and in each case the comparative scores of all factors of each kind for all jobs and all workers are summed to get the corresponding total fitness value. The comparative score of a factor is determined by using Equation 3, 5, or 7 depending upon whether the factor is competence, environmental, or physical.

The Ergonomic Fitness model:

$$F_e = \sum_{c=1}^{n_c} \sum_{l=1}^{n_l} \sum_{j=1}^{n_j} \sum_{p=1}^{n_p} INF_p^c * LS_l^c * JS_j^c * (T_p - A_p) \quad (1)$$

The competence Fitness model:

$$F_c = \sum_{c=1}^{n_c} \sum_{l=1}^{n_l} \sum_{j=1}^{n_j} \sum_{p=1}^{n_p} CS_{ljp}^c \quad (2)$$

With:

$$CS_{ljp}^c = LS_{lp}^c - JS_{jp}^c \quad (3)$$

The Environmental Fitness model:

$$F_{ef} = \sum_{c=1}^{n_c} \sum_{l=1}^{n_l} \sum_{j=1}^{n_j} \sum_{p=1}^{n_p} ES_{ljp}^c \quad (4)$$

With:

$$ES_{ljp}^c = LS_{lp}^c - JS_{jp}^c \quad (5)$$

The physical skill fitness model:

$$F_{ps} = \sum_{c=1}^{n_c} \sum_{l=1}^{n_l} \sum_{j=1}^{n_j} \sum_{p=1}^{n_p} PS_{ljp}^{c} \quad (6)$$

With:

$$PS_{ljp}^{c} = LS_{lp}^{c} - JS_{jp}^{c} \quad (7)$$

Where, $F_e$, $F_c$, $F_{ef}$, $F_{ps}$ are the ergonomic, competence, environmental, and physical fitness respectively; $n_e, n_c, n_{ef}, n_{ps}$ are the number of ergonomic, competence, environmental and physical skill factors respectively; $n_l$, $n_j$, $n_p$ are the number of worker, job and rotation respectively; $INF_p^c$ is the influence factor of factor $c$ for $p$ rotation; $CS_{ljp}^{c}$, $ES_{ljp}^{c}$, $PS_{ljp}^{c}$ are the score of competence, environmental and physical skill factors $c$ respectively for $j$ job assigned to worker $l$ in rotation $p$; $t_p$ is the duration of rotation $p$; $a_p$ is the allowance time of rotation $p$; $LS_l^c$ is the score of ergonomic factors $c$ for worker $l$; $JS_j^c$ is the score of ergonomic factors c for job $j$; $LS_{lp}^{c}$ is the score of competence, or environmental, or physical factor $c$ for worker $l$ in rotation $p$; $JS_{jp}^{c}$ is the score of competence, or environmental, or physical factor $c$ for job $j$ in rotation $p$.

The overall fitness is calculated by equation 8 where weighting coefficients are used to prioritize the factors and homogenization coefficients are used to homogenize different ranking scales of different factors. Equation 9, 10, 11 & 12 corresponds to the way of determining homogenization coefficients for ergonomic, competence, environmental, and physical skill factors.

The overall solution fitness model:

$$F = (C_{ew} * C_{eh} * F_e) + (C_{cw} * C_{ch} * F_c) + (C_{efw} * C_{efh} * F_{ef}) + (C_{psw} * C_{psh} * F_{ps}) \quad (8)$$

With:

$$C_{eh} = \frac{n_c + n_{ef} + n_{ps}}{n_e + n_c + n_{ef} + n_{ps}} * \frac{S_{max\,c} + S_{max\,ef} + S_{max\,ps}}{S_{max\,e} + S_{max\,c} + S_{max\,ef} + S_{max\,ps}} \quad (9)$$

$$C_{ch} = \frac{n_e + n_{ef} + n_{ps}}{n_e + n_c + n_{ef} + n_{ps}} * \frac{S_{max\,e} + S_{max\,ef} + S_{max\,ps}}{S_{max\,e} + S_{max\,c} + S_{max\,ef} + S_{max\,ps}} \quad (10)$$

$$C_{efh} = \frac{n_e + n_c + n_{ps}}{n_e + n_c + n_{ef} + n_{ps}} * \frac{S_{max\,e} + S_{max\,c} + S_{max\,ps}}{S_{max\,e} + S_{max\,c} + S_{max\,ef} + S_{max\,ps}} \quad (11)$$

$$C_{psh} = \frac{n_e + n_c + n_{ef}}{n_e + n_c + n_{ef} + n_{ps}} * \frac{S_{max\,e} + S_{max\,c} + S_{max\,ef}}{S_{max\,e} + S_{max\,c} + S_{max\,ef} + S_{max\,ps}} \quad (12)$$

Where, $C_{ew}$, $C_{cw}$, $C_{efw}$, $C_{psw}$ are the weighting coefficients of ergonomic, competence, environmental, and physical factors respectively; $C_{eh}$, $C_{ch}$, $C_{efh}$, $C_{psh}$ are the homogenization coefficients of ergonomic, competence, environmental, and physical factors respectively; $S_{max\,e}$, $S_{max\,c}$, $S_{max\,ef}$, $S_{max\,ps}$ are the maximum scores of an assignment for ergonomic, competence, environmental, and physical factors respectively. This model has been solved by using a genetic algorithm which is discussed in the following section.

## IV. The Genetic Algorithm

At first, the number of jobs, workers, and rotations are identified as the inputs of the algorithm. Then the scores of different factors like ergonomic, competence, environmental and physical skill factors, for different jobs, workers and rotations are also determined depending upon different scales of different factors as stated in section 3. After taking all necessary inputs the algorithm is performed through the following steps.

Step 1: (Initialization):
Initial population of 20 solutions is developed randomly.

Step 2: (Fitness Assessment):
Fitness value of each solution is calculated by using our developed model.

Step 3: (Selection):
14 elite solutions which have the highest fitness values are selected for next generation.

Step 4: (Reproduction):
6 descendant solutions are created for the next generation from 6 parent solutions of the current generation.

Step 5:
New population with 14 elite solutions from previous generation and 6 descendant solutions produced at the present generation is built and the fitness value of each solution is determined.

Step 6:
Steps 3, 4, and 5 are repeated up to maximum number of generation.

Step 7:
The solution which has the highest fitness value is identified as the optimum solution.

## V. Case Study

In this paper, a case study for 7 workers and 7 jobs has been shown though this algorithm is able to develop a job rotation schedule for a huge number of workers and rotation. 4 sub-factors are selected from each type of factor. For ergonomic, competence, environmental and physical skill factors fitness value is calculated by giving different scores to different jobs and workers for different sub factors on the basis of scales stated in section 4. Influence levels are given according to the predetermined scale and weighting coefficients are given according to the importance of each type of factor.

### A. Results and Discussion

The algorithm has been run for 700 generations and during this execution, the optimum solution is achieved. The optimum job rotation schedule has been found at 349$^{th}$ generation and the fitness value of the schedule has been observed up to 700 generations to ensure that the solution has conversed. The optimum job rotation schedule shown in table 1 has a fitness value of 69.51212 which is the highest one and this schedule satisfies the three basic conditions which are 'same worker is not to be assigned on same job on two different rotations', 'same worker is not to be assigned on two different jobs on a single rotation', and 'two distinct workers are not to be assigned on same job on the same rotation'. From Table 1 it is clear that all workers are assigned to different jobs in different rotations like worker 4 is assigned to job 1 in rotation 1, then in rotation 2 he is assigned to job 4, then in rotation 3 he is assigned to job 2 and at rotation 4 he is assigned to job 3. So there is no repetition of jobs for each of the workers.

TABLE X. THE OPTIMIMUM JOB ROTATION SCHEDULE

| | Job 1 | Job 2 | Job 3 | Job 4 | Job 5 | Job 6 | Job 7 |
|---|---|---|---|---|---|---|---|
| Rotation 1 | Worker 4 | Worker 5 | Worker 6 | Worker 7 | Worker 3 | Worker 2 | Worker 1 |
| Rotation 2 | Worker 7 | Worker 3 | Worker 2 | Worker 1 | Worker 4 | Worker 5 | Worker 6 |
| Rotation 3 | Worker 1 | Worker 4 | Worker 5 | Worker 6 | Worker 7 | Worker 3 | Worker 2 |
| Rotation 4 | Worker 2 | Worker 1 | Worker 4 | Worker 5 | Worker 6 | Worker 7 | Worker 3 |

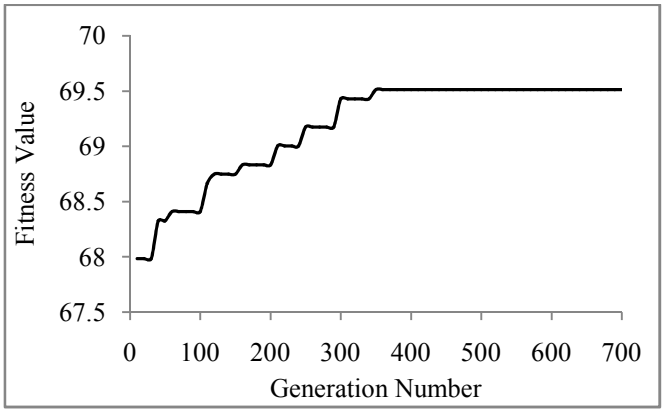



Figure 1. Convergence history of overall fitness value

From Figure 1 it is clearly visible that the best fitness value has increased from generation to generation. After generating initial population randomly, the elite solutions are passed to next generation because they have the highest fitness value. In each generation 6 solutions have been reproduced by combined crossover and mutation operation. Sometimes one or more than one of these newly generated solutions have become the elite solutions because of having higher fitness value than some of the elite solutions of the previous generation. Owing to this property the maximum fitness value has been increased at each generation. But after 349th generation we can see that the fitness curve becomes straight line and it remains straight up to the 700th generation which indicates that the algorithm has conversed enough to give the global optimal solution. The algorithm thus provides the maximum fitness value and corresponding optimal job rotation schedule.

## VI. CONCLUSION

This genetic algorithm has been proved to be an efficient tool of designing job rotation schedules which not only minimize repetitive body movements by diversifying the task of a worker throughout the day but also maximize the worker performance and productivity of the organization. The algorithm requires a very short computation time and considers the movement of the body parts as well as the psychological behavior and competences of the workers. The algorithm eliminates worker exposure to MSK disorder risk factors such as bending, weight-lifting etc.

## REFERENCES


[1] J. C. Boyd and J. Savory, "Genetic algorithm for scheduling of laboratory personnel," *Clinical Chemistry*, vol. 47, no. 1, pp. 118–123, 2001.

[2] B. J. Carnahan, M. S. Redfern, and B. Norman, "Designing safe job rotation schedules using optimization and heuristic search," *Ergonomics*, vol. 43, pp. 543–560, 2000.

[3] A. M. Costa and C. Miralles, "Job rotation in assembly lines employing disabled workers," *International Journal of Production Economics*, vol. 120, no. 2, pp. 625–632, 2009.

[4] J. S. Dean, "Staff scheduling by a genetic algorithm with a two-dimensional chromosome," in E. K. Burke, M. Gendreau, (Eds.), *Proceedings of the 7th International Conference on the Practice and Theory of Automated Timetabling– PATAT 2008*, Montreal, Canada, August 18-22, 2008, pp. 1-5.

[5] T. Eriksson and J. Ortega, "The adoption of job rotation: testing the theories," *Industrial and Labor Relations Review*, vol. 59, no. 4, pp. 653–666, 2006.

[6] S. U. Sekiner and M. Kurt, "A simulated annealing approach to the solution of job rotation scheduling problems," *Appl. Math Comput*, vol. 188, pp. 31–45, 2007.

[7] W. Tharmmaphornphilas and B. A. Norman, "A quantitative method for determining proper job rotation intervals," *Ann Oper Res*, vol. 128, pp. 251–266, 2004.

[8] W. Tharmmaphornphilas and B. A. Norman, "A methodology to create robust job rotation schedules," *Ann Oper Res*, vol. 155, pp. 339–360, 2007.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Autonomous Robot Path Planning in Dynamic Environment Using a New Optimization Technique Inspired by Bacterial Foraging Technique

Md. Arafat Hossain[1,*], Israt Ferdous[1]
[1]Rajshahi University of Engineering & Technology, Bangladesh.
[*] E-mail: arafat.ruet03@gmail.com

*Abstract*—**Path planning is one of the basic and interesting functions for a mobile robot. This paper explores the application of Bacterial Foraging Optimization to the problem of mobile robot navigation to determine shortest feasible path to move from any current position to target position in unknown environment with moving obstacles. It develops a new algorithm based on Bacterial Foraging Optimization (BFO) technique. This algorithm finds a path towards the target and avoiding the obstacles using particles which are randomly distributed on a circle around a robot. The criterion on which it selects the best particle is the distance to target and the Gaussian cost function of the particle. Then, a high level decision strategy is used for the selection and thus proceeds for the result.**

**It works on local environment by using a simple robot sensor. So, it is free from having generated additional map which adds cost. Furthermore, it can be implemented without requirement to tuning algorithm and complex calculation. To simulate the algorithm, the program is written in C language and the environment is created by OpenGL.**

**To test the efficiency of proposed technique, results are compared with Basic Bacterial Foraging Optimization (BFO) and another well-known algorithm called Particle Swarm Optimization (PSO). From the experimental result it can be told that the proposed method gives better path or optimal path.**

***Keywords—Mobile robot, Path planning, Bacterial Foraging Optimization (BFO), Optimization.***

## I. INTRODUCTION

An autonomous mobile robot is a programmable and multi-tasks mechanical device capable of moving freely in the environment including obstacles; executing various functions and acquiring the environment information through the sensors. Usually, they are used in higher, deeper and riskier environment where human cannot imagine treading. For this, the last few decades have witnessed ambitious research efforts in this area of mobile robotics.

In order to achieve tasks, they are intelligent in deciding their own actions. Especially, the topic of navigation is one of the focused points in the correlation of mobile robots. As their use is increasing day by day and navigation is an essential and obvious task for initial interaction, the studies in this area deals now with autonomous indoor and outdoor navigation.

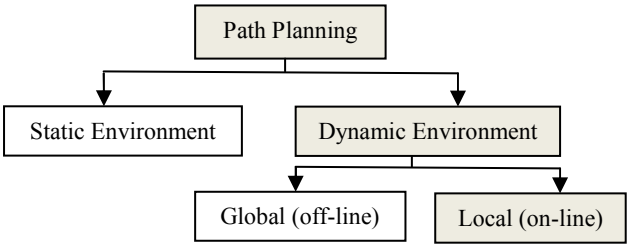
Navigation consists of two essential components known as

**Localization** which refers to the ability of determining accurate position at any moment relative to the search space according to the environment perceptions gathered by sensors.

**Path Planning** is considered as the computation of an optimal path according to some criteria such as distance, time, cost, energy etc. Distance and time being the most commonly adopted criterion. So, for designing a fast and efficient procedure for navigation, path planning is an essential aspect. It can be divided into classes according to the atomicity and availability of knowledge of the information about the environment which is shown below:

In the static environment, all the obstacles would be static, while in dynamic environment, obstacles can be both static and dynamic and may move at arbitrary directions with varying speeds.

Again, in global planning, the map of the environment is totally known, where, in local planning, the information of the environment is not known in advance. So, the robot needs the capability to build a map of the environment.




As the real application contains arbitrary obstacles (moving and fixed) and most of the cases the environment remains unknown, so establishing an efficient path planning algorithm on local dynamic environment would be meaningful and significant; which is the main concern of this paper.

A great number of different techniques in this regard have been and are being developed which can be classified as

1) Analytical Method
2) Enumerative Method
3) Evolutionary Algorithm

Each of these has its own strengths and weaknesses. But the main weakness arise from the fact that the analytical methods are too complex to be used in tangible and enumerative search methods are overwhelmed by the size of the search space. On the other hand, many evolutionary algorithms have been shown to be ineffective in path planning when the search space is large. To overcome these drawbacks, meta-heuristic approaches have been attracting considerable research interest in recent years like Ant colony algorithm, Particle Swarm Optimization, Bacterial Foraging Optimization. Among of these, Bacterial Foraging Optimization is relatively new and has much scope. It meets all the major concerns e.g. it can find the destination in a relatively short time which provides **efficiency**. By using this, a feasible path can avoid all known obstacles in the area which ensures **safety** of planning. Finally, it **accurately** and always finds and follows the determined path.

It also reduces the difficulties related to path planning problem like it is easy computed avoiding most of the computational complexity, free from getting stuck in the local minima and no need to tuning algorithm which offers availability.

Therefore, this paper presents a new and modified algorithm on the basis of Bacterial Foraging Optimization. Using the facial appearance of BFO and the proposed method provide better solution in terms of time and path's feasibility.

The remainder of the paper is organized as follows. In Section 2, recent works on BFO is discussed, Section 3 explains the problem formally, full description and implementation is presented in Section 4, Section 5 shows the experimental setup and result and the paper is concluded in Section 6.

## II. Related Work

Researchers could apply the BFO algorithm successfully in various fields. For example, in the context of biologically inspired optimization methods, several models of bacterial chemo taxis algorithm based on pioneered work of *Bremermann* [2] have been proposed in literature for many applications including robotics [3]. In this paper, the foraging theory is applied to bacteria, adopting the bacteria colonies nomenclature. The fact that bacteria are one of the simplest living beings existing in Earth and they use the forage theory to benefit the group motivated this study. Bacteria own a control system that allows the foraging control and avoidance of noxious substances. In this context, the cooperation activity in a bacteria colony may be used in an optimization procedure, based in the forage strategy, as proposed by *Passino* [4].

So, we focus on the problem of path planning in a practical surroundings which can be categorized as an optimization problem to apply the BFO technique and modify it as necessary to solve the problem.

## III. Problem Representation

To search for an optimal robot path successfully, the model of an environment should be first constructed.

Environment

Fig.1 shows the environment. It is a two-dimensional rectangular workspace where all the objects including robot, obstacles, source and destination point are located

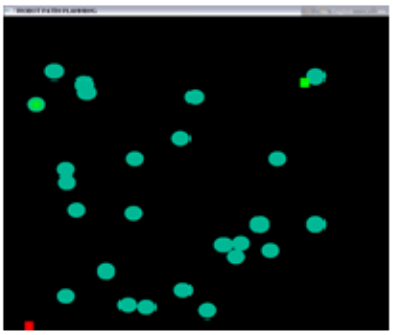


Figure 1. Model of Environment

**Mobile Robot** is defined as a white square object which is represented by *C(Cx, Cy)* in the 2D space.

**Obstacles** are unpredictable objects that the robot may encounter during the execution of the task. In general, obstacles can be of any shape and size with a representation of each point by *O (Ox, Oy)*.

To avoid a safe region around obstacles, the obstacle is wrapped by a circle. The radius, R of the circle is chosen in accordance with the size of the obstacle.

**Goal** is defined as a green square object which is represented by *G (Gx, Gy)* in the 2D space.

### A. Assumption

- The obstacles are detected by the sensor which is mounted on the robot. So, the obstacles within the sensor range, $\beta$ can be detected with reference to the coordinate system.
- The obstacles can be overlapping.
- The path planning program run until the goal has been returned.
- Both the goal and the obstacles should be dynamic.

### B. Input

- Initial location of robot (Red square in bottom-left).
- Location of Destination (Green square in Fig. 1) and
- Position of the obstacles detected by sensor.(Cyan circles)

### C. Outcome of each step

Next step of robot towards the feasible path by avoiding obstacle/s.

### D. Final Output

A shortest pathway which is not crashed with the given obstacle/s.

## IV. METHODOLOGY

The model and concept of the proposed algorithm is as follows:

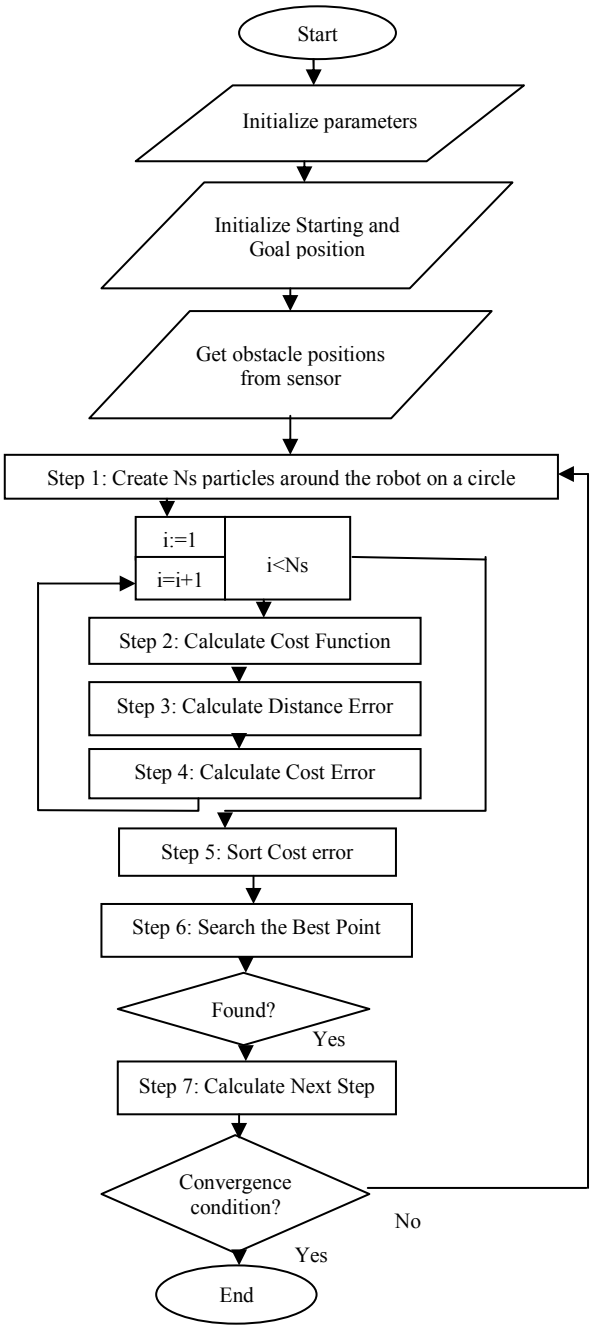



Figure 2. Flowchart of Proposed Technique

The algorithm is based on particles randomly distributed around a robot. At first, all the initial inputs are taken into consider and the parameters are initialized:

TABLE I. PARAMETERS

| |
|---|
| Height of repellant, $H_{obstacles}$=1 |
| Width of repellant, $W_{obstacles}$=4 |
| Height of attractant, $H_{goal}$=1 |
| Width of attractant $W_{goal}$=2300 |
| Radius of circle, R=0.5 |
| Particle Generated around robot, Ns=100 |

TABLE II. INPUTS

| |
|---|
| Initial Position, ($C_x$, $C_y$)=(5.0,5.0) |
| Goal Position, ($G_x$, $G_y$)=(150.0, 160.0) |

### A. Steps

*a) Step 1*: (S=1, 2,……., n) virtual particles are generated and distributed randomly on a circle with radius of R around the robot's current position, (initially, C (Cx,Cy).

Each particle position(S) in time t could be defined as $P_s(t)$ and the next position is calculated as

$$P_s(t+dt) = P_s(t) + R(\Delta(t) / ||\Delta(t)||). \quad (1)$$

where, $\Delta(t)$ is a unit length random vector which is used to define the direction of particle in each time and $||\Delta(t)||$ is the magnitude of the vector.

*b) Step 2:* For choosing the best particle two different strategies are combined described in this and the next two steps.

When the robot arrives at the moving obstacles, its sensor detect obstacles (i=1, 2,……, n) and a repellant Gaussian cost function is virtually assigned to each obstacle.

So, formula of the function in all through the mission is defined as-

$$J_{obstacle} = H_{obstacle} * \exp(—W_{obstacle}(||\theta_i(t)—P_o(t)||^2). \quad \text{if } ||P_o(t)—C(t)||^2 <= \beta$$

$$J_{obstacle} = 0. \quad \text{otherwise} \quad (2)$$

where, $H_{obstacle}$ and $W_{obstacle}$ are constant values defining height and width of the repellant, $P_o(t)$ denotes the obstacle position detected by the sensor and $\beta$ be the sensor range.

$$J_{goal} = — H_{goal} * \exp(—W_{goal}(||\theta_i(t)—P_G(t)||^2). \quad (3)$$

where, $H_{goal}$ and $W_{goal}$ are height and width of the attractant.

So, the total cost function can be calculated as,

$$J = J_{obstacle} + J_{goal}. \quad (4)$$

*c) Step 3:* A decision making strategy designed for finding the best particle among the others is based on the distance error to target which can be showed as,

$$e_s^d(t) = d_s(t + dt) - d_s(t). \quad (5)$$

where, $d_s(t)$ is the distance from particle, S to goal at time,t which can be computed by $(d_s(t) = \|P_s(t)—P_G(t)\|^2)$, and again $d_s(t + dt)$ is the distance from the same particle to goal at time $(t + dt)$ which can be calculated by the formula, $(d_s(t + dt) = \|P_s(t + dt)—P_G(t)\|^2)$.

*d) Step 4*: In a similar manner, cost function error should also be determined by using this one,

$$e_s^J(t) = J(P_s(t +dt) - J(P_s(t)). \quad (6)$$

*e) Step 5:* After calculating all particles at time, t, they are sorted in ascending order of cost error in a vector $S_{sort}$, such that particle with the lowest distance error is at top which is considered as the fit ones.

*f) Step 6:* Since, when an obstacle is not in the robot's range, $J_{obstacle}$ is zero and as a result $-2H_{goal} < e_s^J(t) < 0$, in this situation, robot freely move towards target by choosing the first element of vector, Ssort, as all of them have $e_s^J(t) < 0$.

Again, when sensor detect an obstacle, $J_{obstacle}$ gets a value and consequently, $-2H_{goal} < e_s^J(t) < 2H_{obstacle}$. In this case, a search on $S_{sort}$ from top to bottom is performed and first particle with $e_s^J(t) < 0$ is selected as the best one.

To get more accuracy, some modifications would have to be done i.e. distances from robot to detected obstacles which can be evaluated as, $d_{so}(t) = \|C(t)—P_o(t)\|$, is also sorted in $O_{sort}$.

Then, if both the first and second element of $O_{sort}$ is on the left side, then, sort the cost errors $e_s^J(t)$ in descended order and pick the first particle with negative error as there is no obstacle.

Again, if both the first and second element of $O_{sort}$ is on the right side, then, sort the cost errors $e_s^J(t)$ in ascended order and pick the first particle with negative error.

But, if the first element is on the right side and the second on the left at a safe distance or vice versa, then, sort the cost error from the middle in ascended order and as usually select the first one with negative error.

Otherwise, calculate the best one in normal procedure.

Again to get optimal path, a new parameter, $\eta$ can be introduced which is a co-efficient of $H_{obstacle}$ in decision equation $e_s^J(t) < \eta$ and let the particle not to be omitted like before which can provide optimal path.

*g) Step 7:* Lastly, add the particle to the desired path and let the new position be robot's current position.

Thus, continue the process repeatedly until it satisfies the convergence condition.

## V. SIMULATION RESULT

In this section, the result is included and the efficiency of the proposed algorithm is analyzed. The results are considered in a simulation environment.

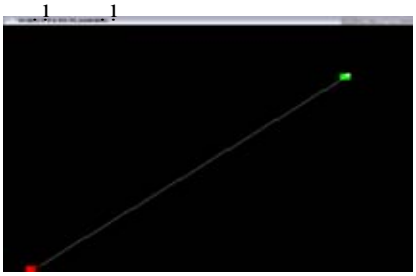


Figure 3. Workspace showing starting and final point with shortest path.

### A. *Fixed obstacle and Target:*

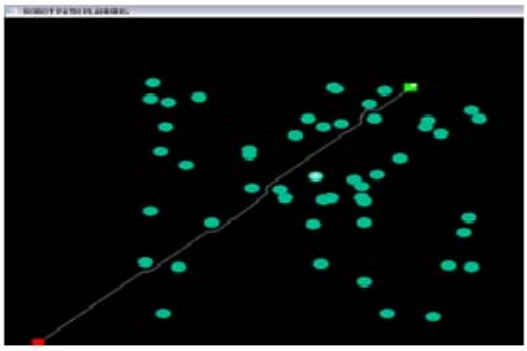


Figure 4. Path generated by the algorithm for fixed obstacles.

### B. *Randomly Moving obstacles and fixed target:*

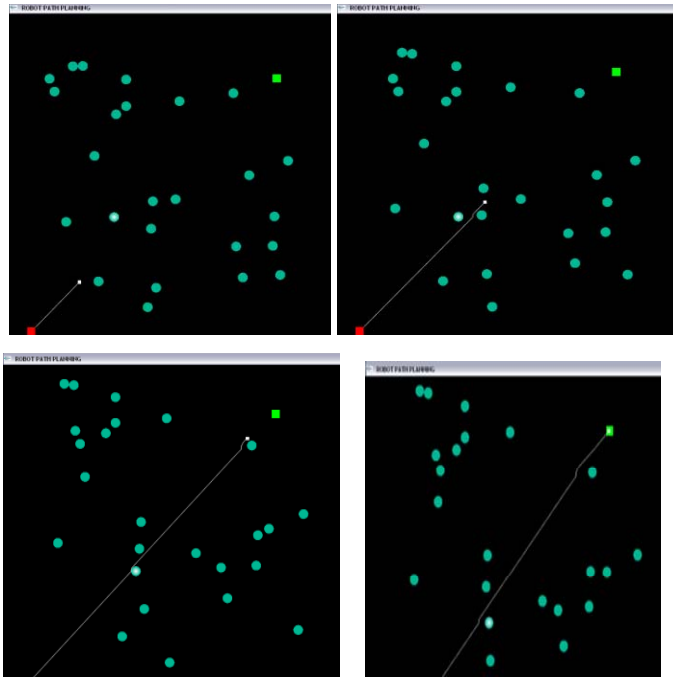


Figure 5. Path traveled by the robot at 24th, 58th , 103rd and final iteration

Each of the tests on each benchmark problem was allocated 10 standard deviation values of path lengths and the time of runs and the results collected are the best, worst, average values.

TABLE III. SIMULATION RESULTS IN TERMS OF PATH LENGTHS AND COLLISION

| Benchmark | Generated path length | | | | Collision |
|---|---|---|---|---|---|
| | *Best* | *Worst* | *Avg.* | *Std. Dev* | |
| 0 obstacles | 212.3 | | | | No |
| 50 obstacles | 224.6 | 267.4 | 245.84 | 19.81 | No |
| 100 obstacles | 261.1 | 444.6 | 310.46 | 75.85 | No |
| 200 obstacles | 264.6 | 394.2 | 329.26 | 46.13 | No |
| 300 obstacles | 365.3 | 577.0 | 480.64 | 103.58 | No |

TABLE IV. SIMULATION RESULTS IN TERMS OF TIME

| Benchmark | Elapsed Time(s) | | | |
|---|---|---|---|---|
| | *Best* | *Worst* | *Avg.* | *Std. dev* |
| 0 obstacles | 7 | | | |
| 50 obstacles | 15 | 18 | 16.2 | 1.30 |
| 100 obstacles | 27 | 44 | 34.8 | 12.65 |
| 200 obstacles | 53 | 77 | 64.8 | 8.55 |
| 300 obstacles | 103 | 160 | 138.2 | 23.14 |

As it is seen from Table III for environments with low and medium density (in terms of obstacles), the proposed algorithm generates solutions with no collision and reasonable average path lengths in relatively short time as well as high density environments where collision occur very rare.

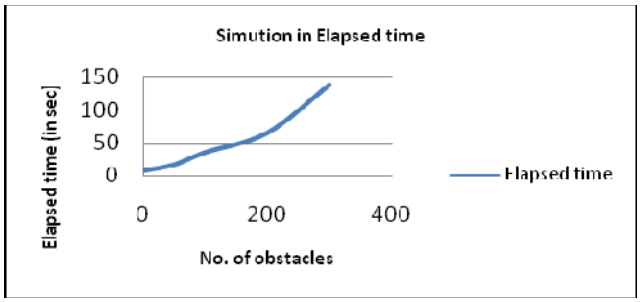



Figure 6. Graph of Elapsed time in various density environment

Sometimes rare fluctuation happens, as the environment contains movable obstacles of various motions. But high density does not frequently occur in practical situations.

Similarly, from the time curve, we see that, the execution time of reaching to goal is almost linear in comparison to the density of environment. So, the time complexity of the algorithm is NP complete. Therefore, we can say that, the algorithm can solve the planning problem in real time which let us to state that the overall performance of the new algorithm is reliable and satisfactory.

Now, the comparison of the technique with other algorithm, Basic BFO and PSO, is shown. Here, is can be seen that, the length is almost same in all of them, but the proposed technique provides a much smoother and less jagging path than the others.

The comparison of calculated distance and time by the technique in different environment among the existing algorithms is shown in Table V and the corresponding curve is shown in figure .7.

TABLE V. SIMULATION RESULTS OF THREE ALGORITHMS IN TERMS OF DISTANCE AND TIME

| Obstacles | Proposed BFO | | Basic BFO | | PSO | |
|---|---|---|---|---|---|---|
| | *Dist.* | *Time (in sec.)* | *Dist.* | *Time (in sec.)* | *Dist.* | *Time (in sec.)* |
| 50 | 219 | 26.4 | 225 | 29 | 215 | 28 |
| 100 | 248 | 66 | 261 | 76 | 241 | 78 |
| 200 | 232 | 118 | 491 | 230 | 276 | 122 |
| 300 | 311 | 222 | 677 | 477 | 328 | 265 |

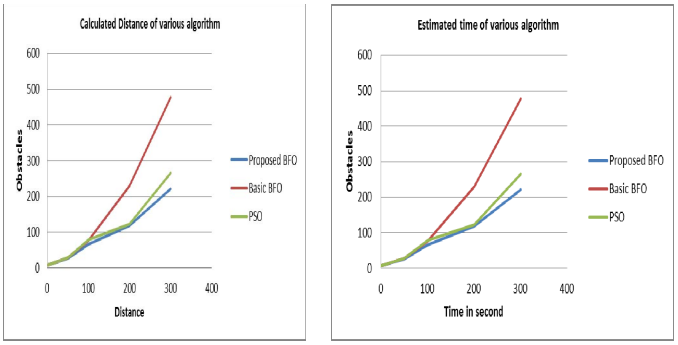



Figure 7. Graprh of estimation of three different algorithm in various density environment

On the basis of the table and curve, it is clear that, the proposed BFO performs better and more effective than the basic one and also than the PSO, especially in high density environment with multiple obstacles and provides the shortest path.

## CONCLUSION

BFO algorithm is extremely resourceful in finding the solution in path navigation. Since, the proposed algorithm uses a non-linear fitness function for making decisions; therefore, it can be used in real complex environment in a feasible time.

In future work, many Nature Inspired Algorithm (NIA) and comparison can be applied to the results which come from the BFO algorithm. Some NIA with our proposed algorithm will be hybridized to improve the performance and also can be added the feature of motion planning to give it a more genuine look.

## REFERENCES


[1] Alon, U.; Surette, M.G.; Barkal; Leibler, S. "*Robustness in bacterial chemotaxis*", Nature 397, 14 January 1999, pp. 168-171, 1999.

[2] Bremermann, H. J. "*Chemotaxis and optimization*", Journal of Franklin Institute 297, pp. 397-404, 2004.

[3] Dhariwal, A.; Sukhatme, G.S.; Requicha, A.A.G. "*Bacterium-inspired robots for environmental monitoring*", Proceedings of the IEEE

International Conference on Robotics & Automation, New Orleans, LA, pp. 1496-1443, 2004.

[4] Passino, K.M. "*Bio-mimicry of bacterial foraging for distributed optimization and control*", IEEE Control Systems 22(3), pp. 52-67, 2002.

[5] O. Ibidapo-Obe and O. S. Asaolu, "*Optimization problems in applied sciences: from classical through stochastic to intelligent meta-heuristic approaches*," in Handbook of Industrial and Systems Engineering, A. B. Badiru, Ed., chapter 22, pp. 1–18, CRC Press, Taylor and Francis Group, New York, NY, USA, 2006.

[6] V. O. S. Olunloyo and M. K. O. Ayomoh, "*Autonomous mobile robot navigation using hybrid virtual force field concept*," European Journal of Scientific Research, vol. 31, no. 2, pp. 204–228, 2009.

[7] G.Kokila, Mr.M.Karnan, Mr.R.Sivakumar, "*Immigrants and Memory Schemes For Dynamic Shortest Path Routing Problems In Mobile Adhoc Networks Using PSO, BFO*", International Journal of Computer Science and Management Research, Vol 2, Issue 5, 2013.

[8] Anupama sharma, Miss Sampada Satav "*Path Navigation Using Computational Intelligence*", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 7, 2012.

[9] Vanitha Aenugu, Peng-Yung Woo, "*Mobile Robot Path Planning with Randomly Moving Obstacles and Goal*", I.J. Intelligent Systems and Applications, 2012

[10] F. Lingelbach, "*Path planning using probabilistic cell decomposition*", in Proceedings of the *IEEE International Conference on Robotics and Automation*, pp.467–472, New Orleans, La, USA, 2004.

[11] O. Khatib, "*Real-time obstacle avoidance for manipulators and mobile robots*", in Proceedings of the IEEE International Conference on Robotics and Automation, pp.500–505, St. Louis, Mo, USA, March 1990.

[12] M. Bikdash, S. Karagol, and M. Charifa, "*Mesh analysis with applications in reduced-order modeling and collision avoidance*", in Proceedings of the COMSOL Users Conference, pp.1–7, Boston, Mass, USA, 2006.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Velocity Tentative Particle Swarm Optimization to Solve TSP

M. A. H. Akhand[1,*], Shahina Akter[1], and M. A. Rashid[2]
[1]Khulna University of Engineering & Technology (KUET), Bangladesh
[2]University Malaysia Perlis (UniMAP), Malaysia
[*] E-mail: akhand@cse.kuet.ac.bd

*Abstract*—**Particle Swarm Optimization (PSO) is a population based optimization technique on metaphor of social behavior of flocks of birds or schools of fishes and has found popularity in solving difficult optimization problems. A number of PSO based methods have been investigated for solving Traveling Salesman Problem (TSP), the popular combinatorial problem. The velocity for solving TSP is the Swap Sequence (SS) with several Swap Operators (SOs) and all SOs of a SS are applied maintaining order on a particle and gives a new tour i.e., a particle having new solution in the PSO. This study investigates a new PSO based method, called Velocity Tentative PSO (VTPSO), for solving TSP where calculated velocity SS is considered as tentative velocity. VTPSO evaluates a number of tentative tours applying SOs one after another sequentially and the final tour is considered as the best tentative tour. VTPSO is shown to produce optimal solution within a minimal time when compared with traditional PSO based methods in solving benchmark TSPs. The reason behind the less time requirement is revealed from the experimental analysis is that VTPSO converge faster due to intermediate tour evaluation.**

***Keywords—Particle Swarm Optimization, Traveling Salesman Problem and Swap Sequence.***

## I. INTRODUCTION

Particle Swarm Optimization (PSO) is a population based optimization technique on metaphor of social behavior of flocks of birds or schools of fishes [1]. PSO is a simple model of social learning whose emergent behavior has found popularity in solving difficult optimization problems. It firstly generates a random initial population, the population contains numbers of particles, each particle represents a potential solution of system, and each particle is represented by three indexes: position, velocity and fitness. PSO has been proven to succeed in continuous problems (e.g., function optimization) as it was proposed for such problems and much work has been done effectively in this area [2-3]. In function optimization domain, the position of the $i$-th particle is represented as $X_i = \{x_{i1}, x_{i2}, x_{i3} \dots, x_{iD}\}$ in the $D$-dimensional search space. At every step each particle changes position based on its velocity represented as $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$. The velocity of a particle depends on its previous best position, $P_i = (p_{i1}, p_{i2} \dots p_{iD})$ and the best one among all the particles in the population $G = (g_1, g_2 \dots g_D)$.

Recently, a number of PSO based methods [4-10] have been investigated for solving Traveling Salesman Problem (TSP), the most popular combinatorial problem. In TSP a salesman is required to complete a tour with the minimum distance visiting all the assigned cities exactly for once. Each PSO particle represents a complete tour to solve TSP and velocity is something to change a tour to a new tour. A number of PSO based methods use Swap Sequence (SS) for velocity operation [4-5]. A Swap Sequence (SS) is a collection of several Swap Operators (SOs) and a SO indicates two positions in the tour that might be swap. All SOs of a SS are applied maintaining order on a particle and gives a new tour i.e., a particle having new solution in the PSO.

The pioneer PSO based method for TSP, basic SS based PSO (SSPSO) [4], transforms the operations of PSO of function optimization to TSP. The SSPSO calculates velocity SS for each particle considering its previous best position and the best one among all the particles in the population. It does not conceive any additional operation in PSO. Conceiving the idea of SSPSO other algorithms to solve TSP are Self-Tentative PSO and Enhance Self Tentative PSO. Self-Tentative PSO (STPSO) introduces tentative behavior in SSPSO that tries to improve each particle placing a node in a different position. Enhance Self-Tentative PSO (ESTPSO) tries to improve individual particle with block of nodes adjustment in addition to individual node adjustment of STPSO [5]. The above discussed methods apply all the SOs of the velocity SS on a particle maintaining order for new solution (i.e., tour).

In the existing methods, the new tour is considered after applying all the SOs of a SS and no intermediate measure is considered. It is notable that every SO implementation gives a new tour, and therefore, there might be a chance to get a better tour (heaving better tour cost) with some of SOs instead of all the SOs. This study presents a new PSO based algorithm for solving TSP in which calculated velocity SS is considered as tentative velocity and new solution is best tour from tentative solutions applying the SOs of velocity SS.

This paper belongs four chapters. An introduction to PSO and few PSO based methods to solve TSP has been given in this section. The rest of the paper is organized as follows. Section II explains the proposed Velocity Tentative Particle Swarm Optimization (VTPSO) to solve TSP in detail. Section III represents the experimental result of VTPSO and compares performance of the proposed VTPSO with exiting PSO methods to solve benchmark TSPs. Finally, Section IV concludes this paper with some remarks and outlines future directions of research opened by this work.

## II. VELOCITY TENTATIVE PARTICLE SWARM OPTIMIZATION TO SOLVE TSP

Figure 1 shows the steps of the proposed Velocity Tentative Particle Swarm Optimization (VTPSO) to solve TSP considering the partial search technique. Like other population based algorithm, VTPSO initializes the population with random solutions and try to improve solutions at every generation step. In initialization (Step 1 of Fig. 1) VTPSO defines number of particles in the population, termination criterion, tour cost. It also assigns a random solution (here tour) and a random velocity (here Swap Sequence) to each of the particle. At this initial stage, previous best solution of each particle ($P_i$) is considered as the current random tour of it and Global best solution ($G$) is the best tour among them.

At each iteration step, VTPSO calculates velocity Swap Sequence (SS) (Step 2.a) using Eq. 1 that is same as traditional methods. For the velocity calculation of a particle, VTPSO considers (i) Last applied velocity ($v^{(t-1)}$), (ii) Previous best ($P_i$) solution of the particle and (iii) Global best ($G$) solution of the swarm. In the equation $\alpha$, $\beta$ are random number between 0 and 1.

$$V_i^{(t)} = \omega V_i^{(t-1)} \otimes \alpha \left(P_i^{(t-1)} - X_i^{(t-1)}\right) \otimes \beta \left(G^{(t-1)} - X_i^{(t-1)}\right) \quad (1)$$

The calculated velocity using Eq. 1 does not apply on a position of the particle to get the new position like a traditional method. But VTPSO considers the calculated velocity as tentative velocity as its name regards. With the tentative velocity SS a number of tentative tours is evaluated and particle moves to the best one among those. The velocity SS that gave the best tour considers previous velocity ($V_i^{(t-1)}$) of the next iteration of Eq. 1. Finding the best tour is considered as the partial search and is the main attraction of VTPSO; therefore, Step 2.b in Fig. 1 for the operation is marked in bold-faced.

Partial search seeks better result with portions of calculated tentative velocity Swap Sequence (SS). A SS is a collection of several Swap Operators (SOs) and consider as velocity that may apply on a solution (i.e., a tour) to get a new tour. In the tradition methods, all the SOs of a velocity SS applied on a tour and generate a new tour. But implementation of every successive SO transforms the previous tour to a new tour and reaches the final tour for the SS. While a traditional method ignores the intermediate tours, the partial search technique explores the option of getting better tour considering those. Therefore, the partial search enhances the capability of getting better tour from a SS applying SOs one by one.

**Step 1:** Initialization

**Step 2:** For each particle $X_i$ in the swarm

a. Calculate tentative velocity $V_i^{(t)}$

**b. Update Solution $X_i^{(t)}$ Partial Search manner**

c. Self-Tentative Operation and Update $P_i$ if the new solution $X_i^{(t)}$ is superior to $P_i$

**i. Apply Self-Tentative Operation on $X_i^{(t)}$**

ii. Update $P_i$ with new solution $X_i^{(t)}$

d. Update $G$ if the new solution $X_i^{(t)}$ is superior to $G$

**Step 3:** Loop to **Step 2** until a termination criterion is met.

**Step 4:** Take the global best solution $G$ as an outcome.

Figure 1. Velocity Tentative PSO (VTPSO) Algorithm for TSP.

Suppose $V_i^{(t)} = SO_1, SO_2, SO_3, \ldots SO_n$ then

$$X_i^{1(t)} = X_i^{(t-1)} + SO_1$$
$$X_i^{2(t)} = X_i^{1(t)} + SO_2 = X_i^{(t-1)} + SO_1 + SO_2$$

……

$$X_i^{n(t)} = X_i^{n-1(t)} + SO_n$$

In the above cases $X_i^{1(t)}, X_i^{2(t)}, \ldots\ldots, X_i^{(t)}$ are the tentative tours and the final tour $X_i^{(t)}$ is a tentative tour having the minimum tour cost.

$$X_i^{(t)} = X_i^{j(t)} \quad (2)$$

where $X_i^{j(t)}$ belongs minimum tour cost among $X_i^{1(t)}, X_i^{2(t)}, \ldots X_i^{j(t)} \ldots X_i^{n(t)}$

Finally the velocity considered is

$$V_i^{(t)} = SO_1, SO_2, SO_3, \ldots SO_j \quad 1<j<n$$

The final velocity may also get from new and previous position of the particle

$$V_i^{(t)} = X_i^{(t)} - X_i^{(t-1)}. \quad (3)$$

The above scenario may describe clearly with an example. Consider a tour $X_i^{(t-1)}$ of 10 cities where city position is the number of the city.

$$X_i^{(t-1)} = 1-2-3-4-5-6-7-8-9-10$$

If velocity SS is $SS = SO(1,4),\ SO(2,5), SO(2,4)$ the tentative tours are:

$$X_i^{1(t)} = X_i^{(t-1)} + SO(1,4) = \mathbf{4}-2-3-\mathbf{1}-5-6-7-8-9-10$$
$$X_i^{2(t)} = X_i^{1(t)} + SO(2,5) = 4-\mathbf{5}-3-1-\mathbf{2}-6-7-8-9-10$$
$$X_i^{3(t)} = X_i^{2(t)} + SO(2,4) = 4-\mathbf{1}-3-\mathbf{5}-2-6-7-8-9-10$$

The implementation of the SS gives two intermediate tours (i.e., $X_i^{1(t)}$ and $X_i^{2(t)}$) and finally reached at $X_i^{3(t)}$. Traditional methods only consider the last one as the final tour (i.e., $X_i^{(t)} = X_i^{3(t)}$) without evaluating the intermediate tours. On the other hand, the PS evaluates the intermediate tours too since all are the complete tours and an intermediate one (here $X_i^{1(t)}$ or $X_i^{2(t)}$) might be better than the final one (i.e., $X_i^{3(t)}$). In PS technique all three tours (i.e., $X_i^{1(t)}$, $X_i^{2(t)}$ and $X_i^{3(t)}$) are considered as tentative tours and velocity SS is considered as the tentative velocity. The final tour in PS technique is the best one among the three. If tour cost of $X_i^{2(t)}$ is better (i.e., lower) than $X_i^{1(t)}$ and $X_i^{3(t)}$ then $X_i^{2(t)}$ is considered as the final tour (i.e., $X_i^{(t)} = X_i^{2(t)}$) and final

velocity is $V_i^{(t)} = SO(1,4),\ SO(2,5)$ that contains first two SOs of the calculated velocity SS.

Partial search might not increase computational cost. It seems that evaluating the intermediate tours might increase computational cost of PS technique because it evaluates all the tentative tours while traditional methods only evaluate last one. But the technique of tour cost calculation minimizes the gap. A tour cost is the sum of individual link costs. To evaluate a tour, if all the links' cost are accumulated PS technique will take much time, in general (*n-1*) times of a traditional method if velocity SS contains *n* SOs. But it does not require considering all the links to get cost of a new tour from a tour which cost is known because a SO only changes four links because it index two cities in the tour which may interchange positions. So implementation of a SO need to discard costs of fours links and add costs four new links that associate with the indexed cities before and after interchange, respectively. As an example, to apply *SO(1,4)* on $X_i^{(t-1)} = 1-2-3-4-5-6-7-8-9-10$ to get $X_i^{1(t)} = X_i^{(t-1)} + SO(1,4) = 4-2-3-1-5-6-7-8-9-10$ it needs to discard cost of 1-2, 10-1, 3-4 and 4-5; and need to add cost of 4-2, 10-4, 3-1 and 1-5.

In VTPSO the fitness of every new position of a particle is checked with $P_i$. If $X_i$ is found better than $P_i$ it applies Self-Tentative operation on ($X_i$) owing to improve it furthermore and then updates $P_i$ (Step 2.c). The Self-Tentative (ST) operation of VTPSO (Step 2.c(i)) consists with Single Node Adjustment of STPSO/ESTPSO [5] and a simplified version of ESTPSO Block Node Adjustment. The block size *k* is determined as a random number between 2 and $K_{max}$ and the value of $K_{max}$ is defined as *N/2*, i.e., half of total cities of the problem. Moreover, ST operation is applied on a particle (i.e., a solution $X_i^{(t)}$) when it is found superior to $P_i$ instead of all the particles as of ESTPSO. The ST operation on the selected particles might be helpful to improve overall performance of VTPSO with a minimal time complexity. After ST operation $P_i$ is updated (Step 2.c(ii) of Fig. 1) with new solution $X_i$. The fitness of every new position ($X_i$) of a particle is checked with *G* and updates *G* if $X_i$ is shown better than *G* (Step 2.d).

VTPSO checks termination criterion at the end of each iteration (in Step 3 of the Fig. 1) and terminates if criterion is met. Usually a sufficiently good fitness of *G* or a maximum number of iterations is considered as the termination criteria. If a termination criterion does not meet, VTPSO continues updating positions of the particles again as indicates the loop to Step 2 from Step 4 in Fig. 1.

## III. EXPERIMENTAL STUDIES

This chapter experimentally investigates the proficiency of proposed Velocity Tentative PSO (VTPSO) algorithm to solve TSP. A set of benchmark problems were chosen as a test bed and the performance of VTPSO compared with two popular PSO based methods that are Swap Sequence based PSO (SSPSO) and Enhanced Self-Tentative PSO (ESTPSO). For fair comparison the experimental methodology were chosen carefully. Finally, an experimental analysis has been given for better understanding of the way of performance improvement in proposed method for solving TSP.

### A. General Experimental Methodology

In this study a suite of 20 benchmark problems are considered from TSPLIB [11] where number of cities varied from 14 to 150 and give diverse test bed. A numeric value in the problem name presents the number of cities in that tour. For example, burma14 and ch150 have 14 and 150 cities, respectively. A city is represented as a coordinate in a problem. Therefore the cost is found after calculating distance using the coordinates.

To investigate the proficiency of the proposed VTPSO, the study also implemented SSPSO and ESTPSO, and compared the experimental results among the three methods. The algorithms are implemented on Visual C++ of Visual Studio 2010. The experiments have been done on a Computer (Dell RM710, Intel (R) Xeon (R) CPU E5620 2.40 GHz, 16 GB RAM) with Windows Server 2008 Datacenter OS. For the fair comparison, the population size was 100; the number of generation was set at 500 as termination criteria for the algorithms. The selected parameters are not optimal values, but selected for simplicity as well as for fairness in observation.

### B. Experimental Results and Performance Comparison

This section presents experimental results of SSPSO, ESTPSO and proposed VTPSO; and makes a comparison among the results. SSPSO is the pioneer PSO based method to solve TSP and ESTPSO is the improved version of SSPSO and so far achieved good result. Table 1 compares tour cost achieved and required time to solve by the methods from 20 individual runs. The bottom of the table shows the summary of the presented results. Since experiments are conducted in a single machine with same experimental settings for all the methods, comparison of required time is a good choice to identify proficiency of a method.

The results presented in the Table 1 clearly indicate the effectiveness of the proposed VTPSO to solve TSP. The proposed method is shown the best method on the basis average tour cost of 20 problems. The average tour cost achieved by VTPSO is 4640.52; on the other hand SSPSO and ESTPSO achieved average tour costs of 25335.75 and 4737.96, respectively. Pairwise Win/Draw/Lose summary in the bottom of the table identified that VTPSO is better than SSPSO for all 20 problem individually. VTPSO is found better than ESTPSO for 11 cases out of 20 problems and preformed worse or equal for rest nine problems which are small sized problems such as burma14 (problem with 14 cities) and ulysses16 (problem with 16 cities).

Between two conventional methods, ESTPSO outperformed SSPSO for all the cases as it is observed from Table 1. As an example, SSPSO achieved tour cost of 1989.80 for eil76 problem, but ESTPSO achieved much better than SSPSO for the problem that is 583.93. Self-Tentative (ST) operation in ESTPSO is the element to get good result since such ST employment in ESTPSO is the main difference from SSPSO. But ESTPSO took on average 63.69 seconds to solve the same eil76 problem which is much larger than SSPSO of 350.66 seconds. On average, ESTPSO took near about double

Table 1: Comparison among SSPSO, ESTPSO and VTPSO on basis of average tour cost achieved and required time to solve benchmark TSPs. The results are the average of 20 independent runs.

| Sl. | Problem | Average Tour Cost (Standard Deviation) | | | Average Required Time in second | | |
|---|---|---|---|---|---|---|---|
| | | SSPSO | ESTPSO | VTPSO | SSPSO | ESTPSO | VTPSO |
| 1 | burma14 | *34.61* (1.16) | **30.87** (0.00) | **30.87** (0.00) | *47.68* | 30.41 | **24.39** |
| 2 | ulysses16 | *83.94* (3.69) | **73.99** (0.00) | 74.00 (0.01) | *56.78* | 44.41 | **28.12** |
| 3 | gr17 | *3215.50* (260.48) | **2332.60** (0.00) | 2342.80 (20.90) | *60.82* | 39.23 | **29.56** |
| 4 | ulysses22 | *110.19* (4.59) | 75.36 (0.09) | **75.35** (0.08) | *82.01* | 57.93 | **38.47** |
| 5 | gr24 | *2295.50* (105.69) | **1249.80** (0.00) | **1249.80** (0.00) | *89.77* | 51.87 | **42.77** |
| 6 | fri26 | *1229.10* (73.02) | **635.58** (0.00) | 637.41 (4.94) | *98.74* | 67.38 | **45.85** |
| 7 | bays29 | *17777.00* (655.35) | **9074.20** (0.00) | 9092.90 (52.32) | *109.57* | 82.99 | **56.29** |
| 8 | eil51 | *1218.60* (61.57) | **408.26** (4.86) | 419.30 (5.76) | 214.69 | *273.49* | **117.35** |
| 9 | berlin52 | *22046.00* (769.74) | **7750.70** (173.37) | 7864.40 (192.15) | 220.07 | *295.32* | **125.21** |
| 10 | st70 | *2831.80* (77.18) | **709.15** (17.70) | 721.29 (19.43) | 314.78 | *467.49* | **188.13** |
| 11 | eil76 | *1989.80* (54.33) | 583.93 (11.87) | **574.67** (6.90) | 350.66 | *563.69* | **230.90** |
| 12 | gr96 | *2693.10* (60.72) | 561.68 (22.45) | **547.87** (16.43) | 473.45 | *777.36* | **316.78** |
| 13 | rat99 | *6615.80* (186.85) | 1402.30 (35.13) | **1337.80** (31.52) | 493.56 | *740.07* | **325.03** |
| 14 | rd100 | *44747.00* (933.99) | 8825.50 (265.82) | **8470.50** (162.87) | 423.63 | *673.56* | **318.30** |
| 15 | eil101 | *2800.80* (51.55) | 694.07 (13.45) | **678.21** (11.08) | 503.45 | *895.13* | **372.05** |
| 16 | lin105 | *96405.00* (2576.40) | 16249.00 (811.70) | **15805.00** (431.74) | 540.92 | *919.78* | **367.17** |
| 17 | ch130 | *38976.00* (570.14) | 6640.60 (135.48) | **6455.90** (132.63) | 716.89 | *1361.80* | **561.55** |
| 18 | gr137 | *4948.20* (128.08) | 807.59 (21.75) | **775.87** (23.03) | 765.93 | *1437.20* | **557.98** |
| 19 | ch150 | *45170.00* (1101.10) | 7321.00 (191.25) | **7045.50** (183.45) | 874.67 | *1610.00* | **728.94** |
| 20 | kroA150 | *211527.00* (5224.30) | 29333.00 (752.22) | **28611.00** (453.60) | 872.64 | *1485.20* | **726.30** |
| | **Average** | **25335.75** | **4737.96** | **4640.52** | **365.54** | **593.72** | **260.06** |
| | **Best/Worst** | **0/20** | **9/0** | **11/0** | **0/7** | **0/13** | **20/0** |

| **Pairwise Win/Draw/Lose Summary** | | | | | | |
|---|---|---|---|---|---|---|
| **Method** | **on Tour Cost** | | | **on Required Time** | | |
| | **SSPSO** | **ESTPSO** | **VTPSO** | **SSPSO** | **STPSO** | **VTPSO** |
| SSPSO | - | **20/0/0** | **20/0/0** | - | **0/0/20** | **20/0/0** |
| ESTPSO | | - | **11/2/7** | | - | **20/0/0** |

time than SSPSO for most of the problems although SSPSO took larger time for few small sized problems such as burma14. That indicates ST operation is costly solution to improve performance.

The interesting observation from results of Table 1 is that VTPSO is much more time efficient with respect to SSPSO and ESTPSO but provide suitable solution with minimal tour costs. It is already mentioned that population size was 100 and the number of generation was set at 500 as termination criteria for each of SSPSO, ESTPSO and VTPSO for results presented in the Table 1. Since the algorithms were tested on the same machine with defined fair setting (unbiased to any one of those), the time to get solution is also a good measure to identify proficiency of a method that differs due to algorithmic matter. Moreover, finding better result with lesser time is more interesting that provides VTPSO. Partial search and moderate Self-Tentative operation are the additional operation in VTPSO with respect SSPSO. Both operations help to find better solution early and size of velocity SS reduces with generation that might be reason to take less time by VTPSO than SSPSO for same number of generations with fixed population size. On the other hand, Self-Tentative operation on each particle at each iteration make ESTPSO computationally much expensive than SSPSO in general. Self-Tentative operation on selected particles and velocity partial search makes VTPSO faster convergence and therefore return good result with less time. As example, VTPSO took 318.30 seconds for rd100 problem. For the same problem, SSPSO took 423.63 seconds and ESTPSO took more than double time of VTPSO i.e., 673.56 seconds. But VTPSO achieved best result for the problem with tour cost of 8470.50. In general, VTPSO took 80% time of SSPSO and less than half of ESTPSO. Finally, on the basis of tour cost and required time VTPSO is the best method.

### *C. Experimental Analysis*

This section first investigates why VTPSO require less time than SSPSO and ESTPSO to get the solution on the basis of size of velocity Swap Sequence over generation. After that it investigates the effect of different parameters on the performance of SSPSO, ESTPSO and VTPSO. The size of swarm (i.e., number particles in the population) and the number of iterations were varied to observe their effect on the methods. The experiments are performed on a selected problem that is eil101.

#### *1) Velocity Swap Sequence and Time over Generation*

Figure 2 presents size of average velocity Swap Sequence (SS), i.e., average number of Swap Operators (SOs) in it, and time elapsed from the beginning for sample runs of three different problems. Total 100 particles are trained for 500 generations for each of SSPSO, ESTPSO and VTPSO; SS in the figure is the average value of the particles at different generations. A velocity SS holds several SOs; operation of a SS is the collective operations of individual SOs of it. Therefore, a large velocity SS (having many SOs) requires more time to calculate as well as to implement for getting a new tour than a small one. A particle's position or solution ($X_i$) closer to particle best ($P_i$) and/or Global best ($G$) generates smaller velocity SS.

The velocity SS in the SSPSO is the accumulation of previous velocity ($v^{t-1}$) and a portion of calculated SS considering present solution ($X_i$), particle best ($P_i$) and Global best ($G$) [4]. Therefore, size of SS increases for any problem and reaches saturation level as it is seen in the Figure 2. ESTPSO tries to improve every individual particle through Self-Tentative operation and considers a portion of previous velocity in the present velocity calculation [5]. Both the things might be the reason to give smaller SS than SSPSO as it is seen the Fig. 2. But due to Self-Tentative operation on all the particles at each generation ESTPSO demands much time than SSPSO as it is seen in figure although size of velocity SS of ESTPSO is smaller than SSPSO.

VTPSO induce partial search in velocity SS implementation and apply Self-Tentative operation on a particle when it is found better than particle best ($P_i$). Therefore, VTPSO seems slower than ESTPSO at early stage of generation when Self-Tentative operation is perform on most of the particles and later on (after 100 generations) it took less time than ESTPSO. On the other hand, the size of velocity SS for VTPSO decreases over generation, at the initial stage it is equal to SSPSO and after 100-200 generations it is less than ESPSO, in general. Finally, Partial Search and selected Self-Tentative operations make VTPSO faster convergence as well as time efficient.

*2) Effect of Population Size and Total Generation on Tour Cost and Required Time*

This section investigates performance of SSPSO, ESTPSO and VTPSO varying population size (i.e., number of particles in the swarm) and number of generation. The result presented in Tables 1 are for the fixed number of population size (=100) and generation (=500) for all the problems. It is interesting to observe how the algorithms perform on the variation of both the parameters. The experiments performed on the same machine explained before.

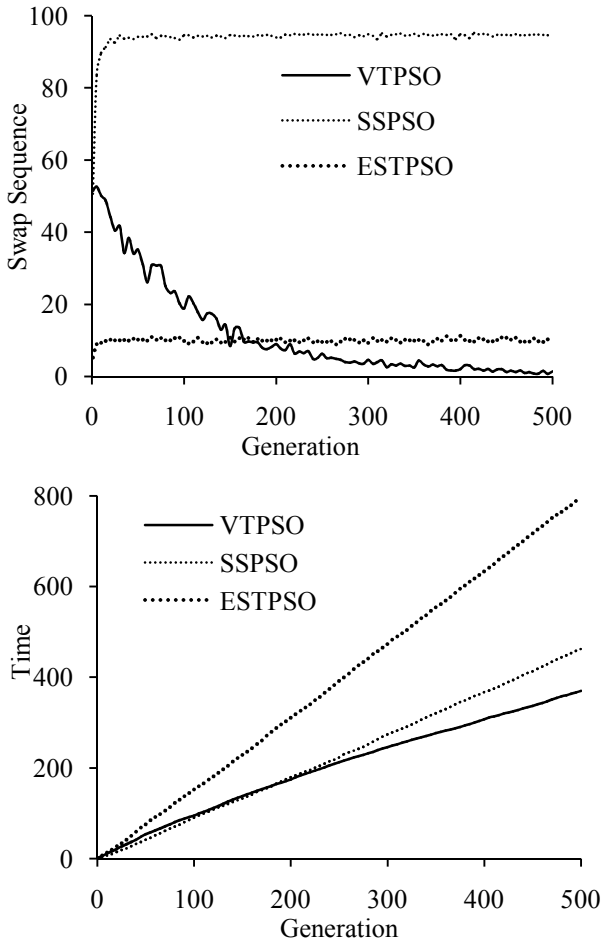



Figure 2. Velocity Swap Sequence size and time over generation for eil101 problem.

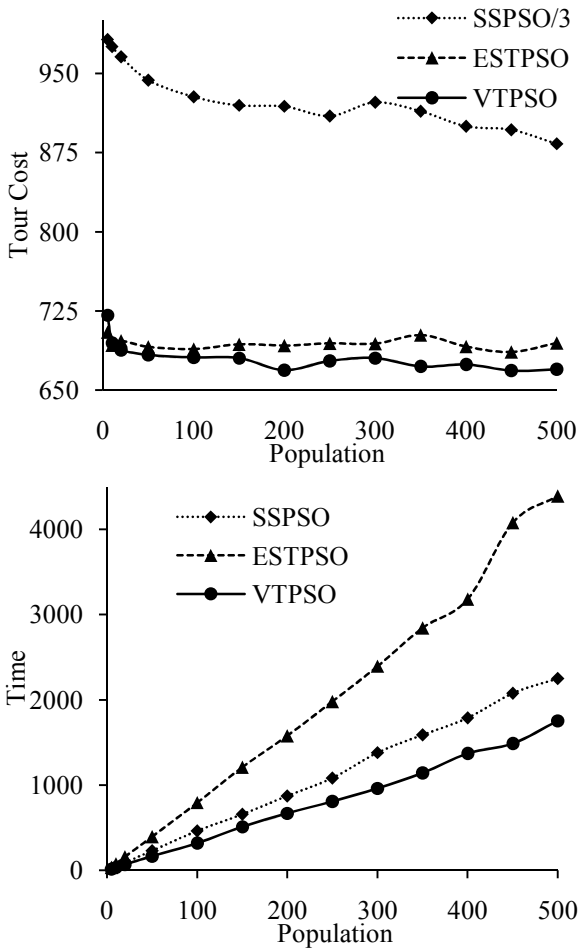



Figure 3. Variation effect of population size on tour cost and require time for eil101 problem.

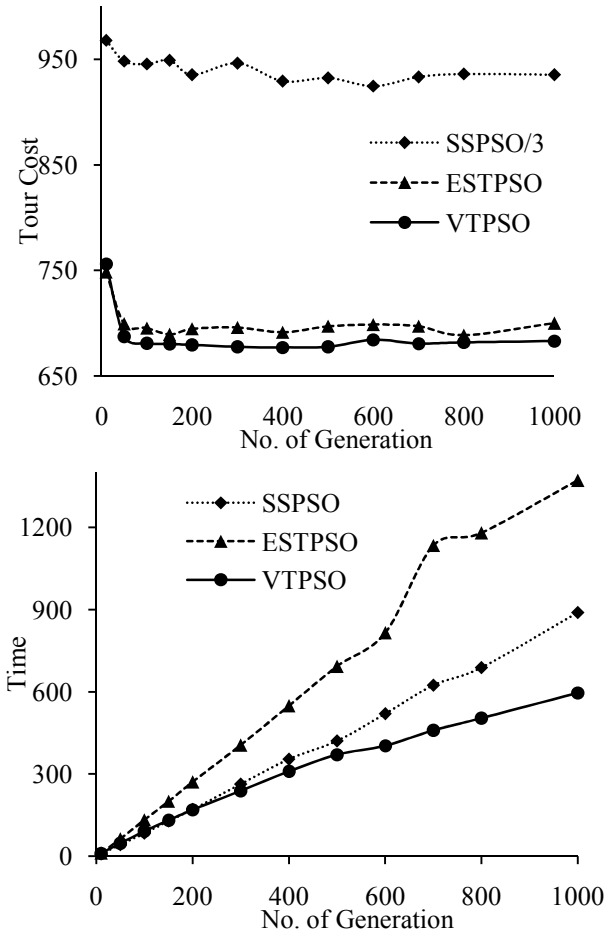



Figure 4. Variation effect of fixed no of generation on tour cost and require time for eil101 problem.

Figure 3 shows the achieved tour cost and required time for different population sizes from five to 500. The presented results are the average for ten independent runs. Since SSPSO perform is much worse result (having larger tour costs) than ESTPSO and VTPSO, its values are presented after scaling down for better visibility. It is seen from Fig. 3 that for very small population (e.g., 5) all three methods found to show worst tour cost and improve up to a certain population size. With tour cost improvement with population, VTPSO is shown better than ESTPSO for any population size. Moreover, on the basis of required time, VTPSO is the best and ESTPSO is the worst on the basis of required time.

Figure 4 shows the achieved tour cost and required time for different fixed number of generations from 10 to 1000. The presented results are the average for ten independent runs. As like Fig. 3, the tour costs of SSPSO are presented in Fig. 4 after scaling down for better visibility. It is seen from the figure that all three methods are shown the worst tour costs for generation 10 and improve rapidly up to a certain value (e.g., 250) and after that improvement is not significant. However, on the basis of achieved tour cost, SSPSO is the worst and VTPSO is the best for any value of generation. VTPSO is also the best on the basis of time requirement taking lowest time. Finally, the Figs. 3 and 4 ascertain VTPSO is a good method for solving TSP.

## IV. CONCLUSIONS

Traveling Salesman Problem (TSP) is the most popular combinatorial optimization problem and interest grows in recent years to solve it new ways. This study investigates a new Swap Operator (SO) based Particle Swarm Optimization (PSO) method for solving TSP. In the existing method, a new tour is considered after applying a complete SS with all its SOs of the velocity calculated. Since every SO implementation on a solution gives a new solution, this study introduced a Partial Search and checked all the solutions for the calculated velocity SS to get best solution with velocity SS. The proposed Velocity Tentative PSO (VTPSO) is shown to produce optimal solution within a minimal time than conventional methods in solving benchmark TSPs. The reason behind the less time requirement is revealed from the experimental analysis is that VTPSO converge faster due to intermediate tour evaluation.

## REFERENCES

[1] R. Eberhart and J. Kennedy, "A New Optimizer Using Particles Swarm Theory," in *Proc. Sixth International Symposium on Micro Machine and Human Science,* Nagoya, Japan, October 1995, pp. 39-43.

[2] X. Yan, C. Zhang, W. Luo, W Li, W, Chen and H. Liu,"Solve Traveling Salesman Problem Using Particle Swarm Optimization Algorithm," *International Journal of Computer Science Issues*, vol. 9, no. 6(2), pp. 264-271, 2012.

[3] J. J. Liang, A. K Qin, P. N Suganthan and S Baskar, "Comprehensive learning particle swarm optimizer for global optimization of multimodal functions". Journal of evolutionary computation, vol. 10, issue. 3, pp. 281 – 295,2006.

[4] K. P. Wang, L. Huang, C. G. Zhou and W. Pang. "Particle swarm optimization for traveling salesman problem," in *Proc. International Conference on Machine Learning and Cybernetics*, November 2003, pp. 1583–1585.

[5] J. Zhang and W. Si, "Improved Enhanced Self-Tentative PSO Algorithm for TSP," in *Proc. Sixth IEEE International Conference on Natural Computation,* 2010, pp. 2638-2641.

[6] X. H. Shi, Y. C. Liang, H. P. Lee, C. Lu and Q. X. Wang, "Particle swarm optimization-based algorithms for TSP and generalized TSP," *Information Processing Letters*, vol. 103, pp. 169–176, 2007.

[7] H Fan, "Discrete Particle Swarm Optimization for TSP based on Neighborhood," *Journal of Computational Information Systems (JCIS)*, vol. 6, pp. 3407-3414, 2010.

[8] W. Zhong, J. Zhang and W. Chen, "A Novel Discrete Particle Swarm Optimization to solve Traveling Salesman problem," *IEEE Congress on Evolutionary Computation*, 2007, pp. 3286-3287.

[9] M. F. Tasgetiren, P. N. Suganthan and Q. Pan, "A Discrete Particle Swarm Optimization Algorithm for the Generalized Traveling Salesman Problem," *in* Proc. 9th annual conference on Genetic And Evolutionary Computation, 2007, pp. 158–167.

[10] E. F. G. Goldbarg, M. C. Goldbarg and G. R.de Souza, "Particle Swarm Optimization Algorithm for Traveling Salesman Problem," Traveling Salesman Problem, Federico Greco(Ed.), InTech, 2008.

[11] TSPLIB - a library of sample instances for the TSP.
Available: http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Localization of FACTS Devices for Optimal Power Flow Using Genetic Algorithm

A.K.M. Rezwanur Rahman, Md. Shahabul Alam, Md. Zakir Hossain * and Md. Shahjahan
Dept. of Electrical and Electronic Engineering, Khulna University of Engineering & Technology, Bangladesh
* E-mail: zakireee62@yahoo.com

*Abstract*—**This paper presents about the effective localization of the FACTS (Flexible AC Transmission Systems) devices in power system by a global search GA (Genetic Algorithm) technique. Ultimate goal is to improve the stability of power system as well as to reduce the generation cost, transmission losses by increasing loadability and improving voltage profile with introduce FACTS device at the most effective region. This method is employed by considering the cost of FACTS and their optimal utilization in the system. Optimal location of FACTS , their types and rated values are optimized simultaneously. Three FACTS devices such as TCSC, TCPAR and SVC are simulated in this study. Simulation is carried out on IEEE30 bus and IEEE 118 bus power system with different increased load-ability. We search the efficiency of this method on the basis of power generation cost, FACTS investment cost and transmission loss reduction. The employed algorithm is emerged as an effective and practical method for the choice and allocation of FACTS in large power systems.**

***Keywords—FACTS devices, Genetic Algorithm (GA), Optimal Power Flow (OPF), Improved loadability, Load Flow.***

## I. INTRODUCTION

In recent years, deregulation of electricity has emerged for its huge demand. Due to the deregulation of the electricity market, study regarding this matter has become imperative. Various initiatives are taken to overcome that, but utilization of FACTS device attracts everyone's attention. FACTS devices are being played vital role for better utilization of the existing power system with the increased demaned [1], [2].

On the other hand, transmision and distribution orientation has become more severe due to the lack of proper arrangement. The major power loss occurs for system loss which is increasing day by day around the world and has emerged as a challange for the developing countries to run with limited resources. To minimise this transmission power losses and ensure optimal power flow, FACTS is introduced in power system. Different parameters and variables of the transmission line such as line impedance, terminal voltages and voltage angle can be controlled by FACTS devices in a fast and effective way [3].

Various types of FACTS devices such as Thyristor Controlled Series Compensations (TCSC), Thyristor controlled phase angle Regulators (TCPR), Unified Power Flow Controllers (UPFC) and Static Var Compensator (SVC) etc, are used to control the power flow in the network. These increase the flow in heavily loaded lines, there by resulting in increase load ability, lowering system losses, improved stability of network and reduced cost of production [4]. Although FACTS device has an great impact on power system for optimal power flow, but it requires optimal alocation for proper stabiliztion and localization. For that reason, many researches were made on the optimal location of FACTS devices with many different ways. We use GA technique to search the optimal localization of FACTS devices.

Genetic algorithms (GA) is a parallel and global search technique [5], [6], which generate solutions to optimization problems using natural evolution, such as inheritance, mutation, selection, and crossover. It is more likely to use for converging toward the global solution because it evaluates many points in the parameter space simultaneously. GA differs from other optimization and search procedures in four ways [7], such as (i) it can easily handle the integer or discrete variables because of coding of the parameter set, not the parameters themselves, (ii) it searches within a population of points, not a single point which may provide a globally optimal solution, (iii) it can deal with the non-smooth, non-continuous and no differentiable functions which are actually exist in a practical optimization problem, because it utilizes only objective function information, not derivatives or other auxiliary knowledge, (4) it uses probabilistic transition rules, not deterministic rules. Although GA seems to be a good method to solve optimization problem, but sometimes the solution obtained from GAs is only a near global optimum solution.

In this paper general GA is applied to improve stability, voltage profile of power system as well as to reduce transmission and generation losses using three types of FACTS devices. In [1], locatioin of FACTS devices are found on the basis of two parameters - overload and over voltage while in [11], only instalation and generation costs are considered. Our optimization includes all of above parameters. Moreover, reduction of generation cost**,** transmission losses, improved loadability and system stability were considered in the objective function of the GA for better improvement of the power system which are not considered altogether by previous work. IEEE 30 and IEEE 118 bus of power system are used for simulations.

The rest of the papers are organized as follows. Section II describes about optimal power flow where subsections A and B analyzes about optimization and analysis respectively. FACTS devices are described in section III where generalities & choice and modeling are explored in subsections A & B accordingly.

We described about GA in section IV. Results are optimized in section V. Finally we conclude the work in section VI.

## II. Optimal Power Flow

### A. Power Flow Optimization

Optimal power flow (OPF) is a nonlinear programming problem which is also called load flow analysis, is very

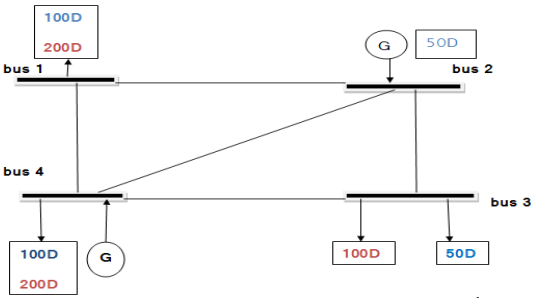



Figure 1. Configuration of 4-bus power system.

important for analysis of an electrical system. By analysis this the characteristics of the power flow direction in different line and the voltage at different buses can be easily obtain [8]. Fig. 1 shows an example of configuration of 4-bus power system where two bus has alternator to generate power and all bus share this power with different demand.Where red and blue indicate peak and off peak demand respectively. Due to increase in sudden load some buses are overloaded and some are with lag of load which causes difference in bus voltages and power flow in transmission line. So we need to calculate overall generation of power, cost, transmission loss, voltage stability in that time in an economic way. For that reasson power flow analysis is necessary.

### B. Optimal Power Flow Analysis

The line resistance is small compared to the reactance and transverse capacitance is close to zero for intereconected power system network that obyes the Kirchoffs law. we consider only line reactances for interconection between i and j bus which is shown in Fig. 2. $P_{ij}$ is the real power flow and $Q_{ij}$ is the reactive power flow between two buses by a line is related by the following equations:

$$P_{ij} = \frac{V_i V_j}{X_{ij}} \sin\theta_{ij}, \qquad (1)$$

$$Q_{ij} = \frac{1}{X_{ij}}\left(V_i^2 - V_i V_j \cos\theta_{ij}\right), \qquad (2)$$

Where, $V_i$ and $V_j$ voltages at buses i and j respectively, $X_{ij}$ reactance of the line, $\theta_{ij}$ angle between $V_i$ and $V_j$. Here $X_{ij}$ & $\theta_{ij}$ controls real and reactive power. We analyze these by Matlab power simulation package Matpower 4.1 [9].

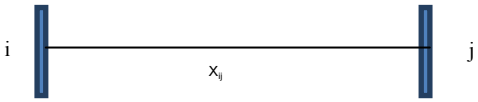



Figure 2. Bus i and bus j is conected by line with reactance Xij.

## III. Facts Devices

### A. Introduce with FACTS Devices

The FACTS is power electronics based device which is used for AC power transmission and distribution. Due to rapid response, ability for frequent variations and smooth adjustability with output the applications of FACTS devices are increased day by day [10].

Different types of FACTS devices are used for power transmission and distribution such as SVC (Static Var Compensators), Fixed Series Capacitors (SC), Thyristor-Controlled Series Compensator (TCSC), Thyristor Controlled Phase Angle Regulator (TCPAR), Unified Power Flow Controller (UPFC), STATCOM etc. Power quality, availability, system stability, transmission capability can be improved using FACTS as well as it minimizes transmission and distribution losses. Here just TCSC,TCPAR and SVC are used for the optimal power flow analysis.

### B. Choice and Modeling

Three different types of FACTS devices have been chosen for the controlling of power flow. These are TCSC, TCPAR, and SVC. TCSC is used to modify the reactance of the transmission line $X_{ij}$. For controlling the phase angle $\theta_{ij}$ the TCPAR is used. The SVC is used to absorb or inject reactive power which is connected in shunt with the line. Both the TCSC and TCPAR are connected in series with the line.

Each FACTS device is represented with fixed discrete values for mathematical analysis, where it has two possible characteristics, capacitive or inductive accordingly in order to decrease or increase the line reactance, phase angle, reactive power in line using TCSC, TCPAR and SVC respectively. Maximum and minimum value of each FACTS device is fixed and type of each device is also specified. TABLE I represents the specification of FACTS devices. Fig. 3 shows the connection model of FACTS devices.

TABLE I. Specification of FACTS Devices

| Name/Specification | *TCSC* | *TCPAR* | *SVC* |
|---|---|---|---|
| **Device type** | 1 | 2 | 3 |
| **Minimum value** | -0.8 $X_L$ (Capacitive) | - 5 deg. | -100 MVar |
| **Maximum value** | 0.2 XL (Inductive) | +5 deg. | +100 MVar |

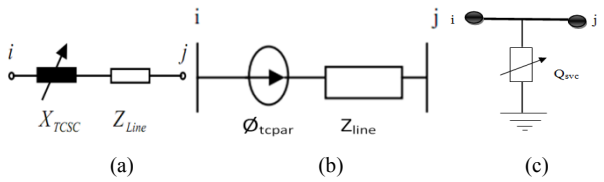



Figure 3. Connection models of FACTS devices. (a) TCSC (b) TCPAR (c) SVC

The real value of the FACTS device [1] $V_{realF}$ is calculated with there loccation according to the model of the FACTS by

$$V_{realF} = V_{\min F} + (V_{\max F} - V_{\min F})V_F \qquad (3)$$

Where, the normalized value is $V_F$, the maximum and minimum setting value are $V_{maxF}$ and $V_{minF}$ respectively.

## IV. GENETIC ALGORITHM

### A. GA overview

GA was proposed on the basis of the evolutionary ideas of natural selection and genetics [5], [6]**.** It represents an intelligent exploitation of a random search used to solve optimization problems which is a rapidly growing area of artificial intelligence (AI). Because of independent of the choice of the initial configurations GA finds high quality solutions**.** Moreover, they are computationally simple and easy to implement. Genetic diversity or variation is configured using different operators such as reproduction (selection), crossover (recombination) and mutation.

### B. Model of GA

The main objective of the optimization is to find the best location for a given number of FACTS devices in the power system based on defined criterion. Three parameters are utilized for encoding individual these are the location, type of device and rated value [1], [12]. Each individual is represented by *nF* number of three strings, where *nF* is the number of FACTS devices installed in the power system. TABLE II shows the individual format. Individual is made in three stages, first a set of branches are randomly selected and is put in the first string. In the second string type is also randomly selected. In the third string device setting value is randomly selected. This approach is repeated for obtaining desired population. Then the entire population is computed with respect to objective function which is the measure of obtaining best location for the FACTS device.

TABLE II. INDIVIDUAL FORMAT FOR GA

| *Location* | *FACTS Type* | *Normalized value* |
|---|---|---|
| 2 | 1 | 0.3 |
| 5 | 3 | 0.6 |
| 11 | 1 | 0.2 |
| 22 | 2 | 0.9 |
| 26 | 1 | 0.5 |
| 18 | 3 | 0.1 |

New individual is generated based on the results obtained from the old generation. For this 1st GA operator selection is used. In this case Proportional Roulette Wheel Selection technique is used. In proportional roulette wheel, individuals are selected with a probability that is directly proportional to their fitness values i.e. an individual selection corresponds to a portion of a roulette wheel. Let $F_i$ be the fitness value and $P_i$ be the selection probability, then

$$P_i = \frac{F_i}{\sum_{i=1}^{N} F_i} \qquad (4)$$

Based on the selection probability $P_i$ individual is randomly selected by roulette wheel. After that 2nd GA uniform crossover is applied [5]. Uniform crossover with some probability knows as the mixing ratio. The crossover operator allows the parent chromosomes to be mixed at the gene level. Consider the two parents selected for crossover. If the mixing ratio is about 0.5, then half of the genes in the offspring will come from parent 1 and rest from parent 2. Then boundary mutation is applied. Fig.4. shows the uniform crossover and boundary mutation technique clearly. Crossover and mutation is done after selection of parents. Blue and black colors represent parent 1 and parent 2 respectively. Then we obtain offspring according to crossover of parents and mutation. The red color shows the mutation results where other for crossover.

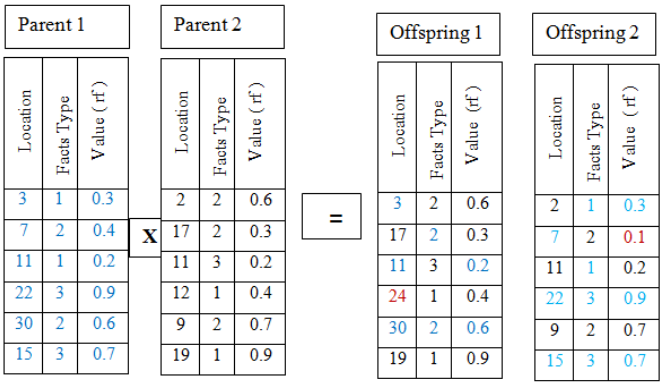



Figure 4. Crossover and Mutation approach.

The entire methodology of GA is explained in Fig. 5. Firstly individuals are selected randomly. Then fitness value is calculated for each individual based on the fitness function. Best individual is found according to selection, crossover and mutation when final criterion is reached.

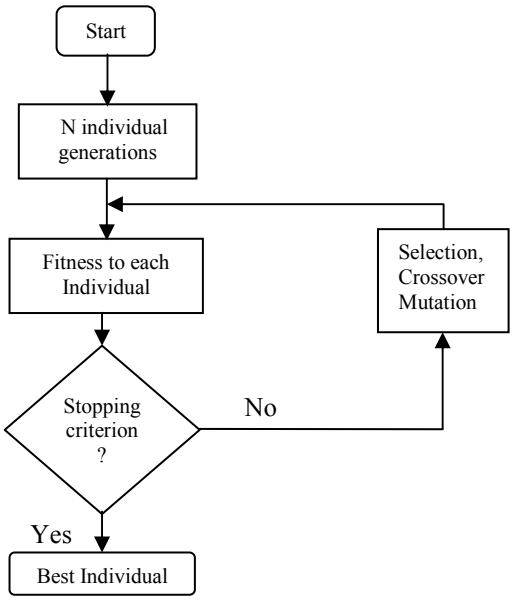



Figure 5. Flow chart of the optimization strategy.

## V. OPTIMIZATION AND RESULT

### A. Objective function

The aim of the optimization is to find the best locations for the given number of FACTS devices within the defined

constrains for the best utilization of the existing system. We want to minimize the power generation costs and reduce the transmission and distribution losses. So the objective function is based on the minimization of cost which can be expressed as

$$C_{Total} = C_1(f) + C_2(PG) + C_3(PL), \quad (5)$$

Where, $C_{\text{Total}}$, $C_1$ (f), $C_2$ (P*G*), $C_3$ (PL) are the total cost of objective function, average installation costs of FACTS devices at each observation per hour, total generation costs and cost of power transmission losses respectively.

The cost functions for SVC, TCSC and TCPAR are developed on the basis of the Siemens AG Database [8], [15]

The cost function for SVC and TCSC are:

$$C_{svc} = 0.0003S^2 - 0.3051S + 127.38(US\$/K\text{ var}) \quad (6)$$

$$C_{t\,csc} = 0.001S^2 - 0.713S + 153.75(US\$/K\text{ var}) \quad (7)$$

Where S is the operating range of the FACTS controllers in kVar. Depending on the installment cost, the cost function of TCPAR can be expresed as

$$C_{tcpar} = 140.5(US\$/K\text{ var}) \quad (8)$$

The cost function for SVC, TCSC and UPFC from Siemens AG Database is shown in Fig. 6.

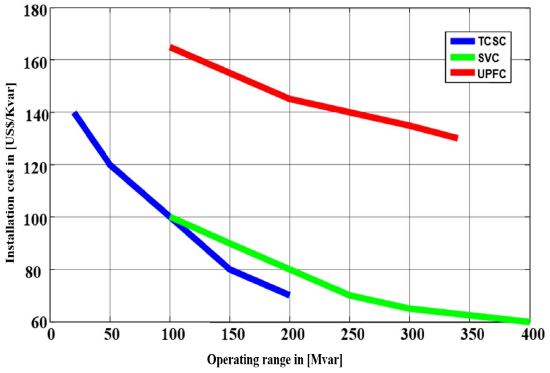



Figure 6. Installation cost curve

Now if C(f) is the summetion of the used FACTS device instalation cost. The generation cost is calculated in per unit that is *US$/Hour* and the installation costs of FACTS devices are in *US$*. For that reason life time of the FACTS is considered. In this paper, three years is applied to evaluate the cost function [10], [11]. We calculate the average values of the installation costs using the following equation, where 8760 is the total hour in a year.

$$C_1(f) = \frac{C(f)}{8760 \times 3}(US\$/Hour) \quad (9)$$

The generation cost function is represented by a quadratic polynomial as follows:

$$C_2(PG) = \alpha_0 + \alpha_1 PG + \alpha_2 (PG)^2 \quad (10)$$

Where PG is the output of the generator (MW), and α0, α1 and α2 are cost coefficients.

The cost function for power loss is represented as

$$C_3(PL) = \sum_{i=1}^{N} PL * Eloss * dt \quad (11)$$

Where, N, P*L,* Eloss and dt are denotes the number of used FACTS devices, transmissin losses, cost of the losses in per hour and FACTS devices utilization time respectively.

Now the fitness function for the genetic algorithm is found as

$$\text{Fitness} = 1/C_{\text{Total}} \quad (12)$$

*B. Results*

According to variation of fitness function individuals are generated using GA to optimize the power flow. The simulation is carried out by free Matlab power simulation package Matpower 4.1 [9]. Based on the GA, best fittest individual is found for the optimal power flow in IEEE 30 and IEEE 118 bus power system with increased amount of demand. Reduction of the power loss and improvement of the voltage profile during transmission are introduced here those are shown in Figs. 9 and 10 accordingly.

Figs.7 and 8 show the fitness value of the fitness function with respect to generation for IEEE 30 and IEEE 118 bus power system respectively.

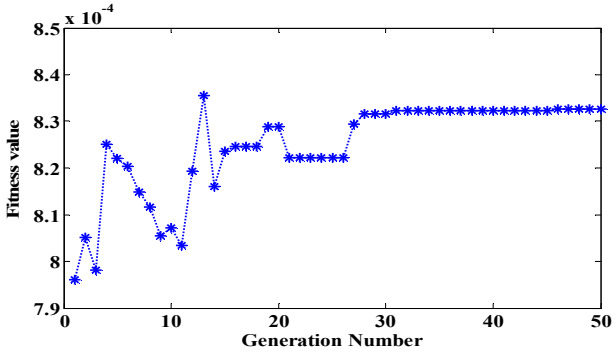



Figure 7. Fitness function curve with generation for IEEE 30 bus system

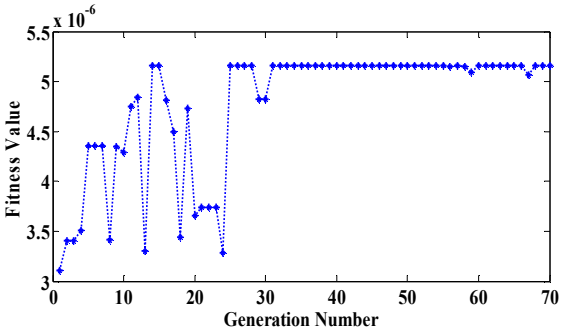



Figure 8. Fitness function curve with generation for IEEE 118 bus system.

Fig. 9 shows both for the IEEE 30 bus and IEEE 118 bus that before using FACTS device power loss through line is more which is showed with red mark. After using FACTS device power loss through line reduced and it is showed by blue mark. Although in some case little increasement of power loss after using FACTS device but it is negligible. So overall

performance is much better after using FACTS device. It has found that obtained location is showing satisfactory output.

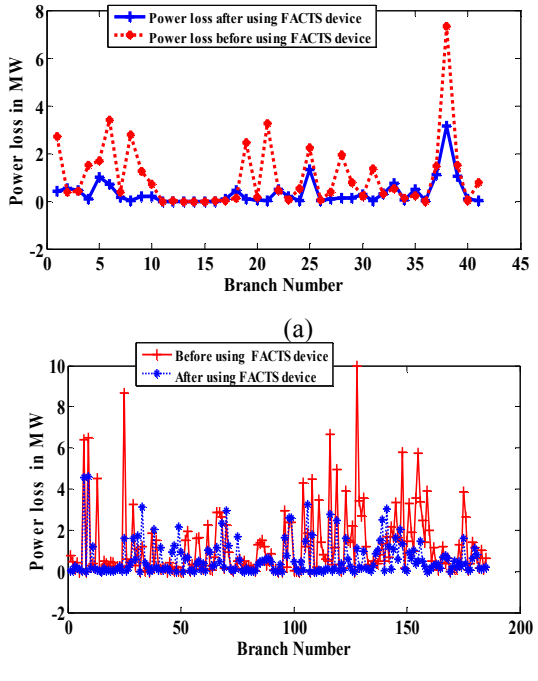



Figure 9. Comparison of power loss before and after using FACTS device in (a) IEEE 30 bus (b) IEE118 bus.

Fig. 10 shows both for IEEE 30 and IEEE 118 bus it is found that voltage magnitude (VM) profile is better while using FACTS device than without FACTS device. As voltage magnitue sholud stay in the limit 1.05 and 0.95 per unit which is marked by color green and black, without FACTS device VM is marked by red color which is very poor. On the other hand after FACTS device utilization voltage profile is marked by blue and it is more stable.

Optimal locations in the power system are detected by using genetic algorithm for the FACTS device. At the same time specifed FACTS device with specifed value which is highly effective for optimal power flow is too determined. After applying these obtainted outcome in the power system, final optimal power flow is observed in the IEEE 30 and IEEE118 bus system. In the IEEE 30 and IEEE118 bus system it was observed that transmission loss reduced in the system after iinstalling FACTS device during certain increase in the load in the system. It was also too observed that voltage profile at each bus of the system improved after installing FACTS device.

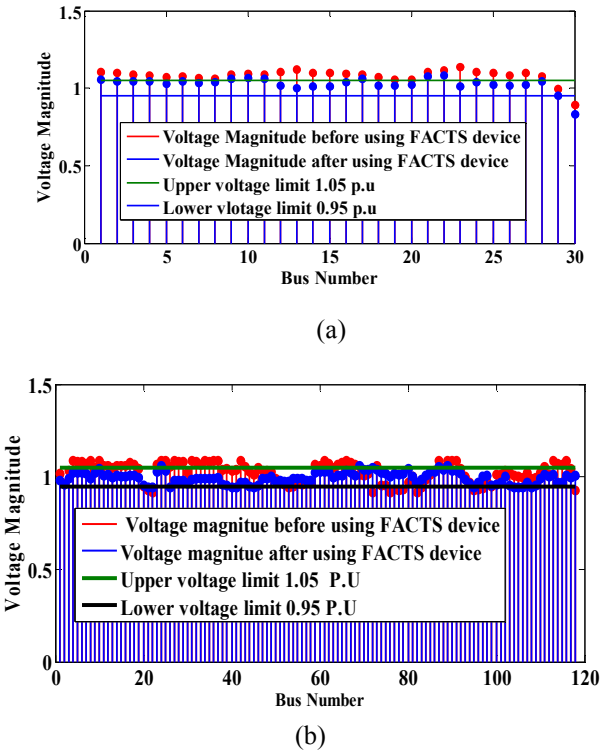



Figure 10. Voltage Magnitude profile comparision before and after using FACTS device in (a) IEEE 30 bus (b) IEEE118 bus

## VI. CONCLUSION

A genetic algorithm has been presented with larger parameters than previous methods to optimally locate FACTS devices in the power system. Here only three types of FACTS devices are used and simulation is carried out on IEEE 30 bus and IEEE 118 bus. After simulation FACTS devices are used in the obtained location and the power flow of the system is observed. It is found that power transmission loss has reduced for using FACTS devices in case of IEEE 30 bus for 90% cases and in case of IEEE 118 it is about 75%. So transmission losses are reduced. In case of bus voltage profile, about 95% bus voltage remain within the limit due to use of FACTS device in IEEE 30 bus and in case of IEEE 118 bus it is about 80%. So it can be said that overall FACTS device has a great impact in power system for optimal power flow and in that case Genetic algorithm exhibits a great impact for selecting the perfect location.

## REFERENCES


[1] S. Gerbex, R. Cherkaoui and A. J. Germond, "Optimal location of multi-type FACTS devices in a power system by means of genetic algorithms," *IEEE Trans. Power Systems*, vol. 16, pp. 537-544, August. 2001.

[2] K. Ghoshand and V. C.Ramesh, "An option model for electric power markets," *Electrical Power and Energy Systems*, vol.19, no.2, 1997.

[3] F. D. Galiana, K. Almeida, M. Toussaint, J. Griffin and D. Atanackovic, "Assessment and control of the impact of FACTS devices on power system performance" *IEEE Trans. Power Systems*, vol. 11, no. 4, Nov. 1996.

[4] D. J. Gotham and G. T. Heydt, "Power flow control and power flow Studies for systems with FACTS devices," *IEEE Trans. Power Systems*, vol.13, no.1, Feb.1998.

[5] R. C. Chakrabarty. Fundamental of Genetic Algorithms AI course, Lecture39-[Online],Availale: http//www.myreaders.info/html/artificial_intelligence.html.

[6] X. P. Wang and L. P. Cao, "Genetic Algorithms – Theory, Application and Software Realization", Xi'an Jiaotong University, Xi'an, China, 1998.

[7] T. S. Chung and Y. Z. Li, "A hybrid GA approach for OPF with consideration of FACTS devices," *IEEE power engineering Review*, pp. 47-57, February, 2001.

[8] I. O. Elgrd, *Electric Energy System Theory- An Introduction*, McGraw Hill Inc., New York, 1971.

[9] R. D. Zimmerman, C. E Murillo-Sanchez and R. J. Thomas, "MATPOWER: Steady State Operations, Planning & Analysis. Tools for Power Systems Research & Education," *IEEE Transactions on Power Systems,* vol. 26, no.1, pp.12-19, Feb. 2011.

[10] N. G. Hingorani and L. Gyugyi,. "Understanding, FACTS - Concepts and Technology of Flexible AC Transmission Systems". *IEEE*, 1999

[11] L. Cai, I. Erlich, G. Stamtsis and Y. Luo, "Optimal Choice and Allocation of FACTS Devices in Deregulated Electricity Market using Genetic Algorithms", *Conf. on Bulk Power System*, Daynamics & Control-VI, August 22-27, 2004.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Speckle Noise Modeling in the Contourlet Transform Domain

Shahriar Mahmud Kabir[1] and Mohammed Imamul Hassan Bhuiyan[1]
[1]Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh
E-mail: shahriar_buet_msc@yahoo.com, imamhas@gmail.com

***Abstract*— Speckle noise is an inherent phenomenon in medical ultrasound (US) images. Since it degrades an ultrasound image quality and reduces its diagnostic value, reduction of speckle noise is a very important pre-processing step in ultrasound image processing. For this purpose, the knowledge of the statistics of speckle noise is necessary; especially in the multi-resolution transform domain due to their sparse and efficient representation of images. In this paper a Bessel K-Form (BKF) probability density function (*pdf*) is proposed as a highly suitable prior for modeling the log-transformed speckle noise in the well-known contourlet transform domain. A maximum likelihood based method is presented for estimating the parameters of the BKF *pdf*. The appropriateness of the BKF *pdf* in modeling the speckle is studied for different noise levels in the contourlet transform domain, in addition the suitability of BKF model is investigated for the case of real US images. It is shown that, in general the BKF can model the statistics of the contourlet transform coefficients corresponding to log-transformed speckle better than the traditional Gaussian and normal inverse Gaussian *pdf*s.**

***Keywords*—Bessel K-form *pdf*, Speckle Noise, Maximum Likelihood (ML), Contourlet transform.**

## I. INTRODUCTION

Medical ultrasound (US) has become an ubiquitous imaging modality for diagnostic purposes due to its non-invasiveness, use of safe non-ionizing sound waves, low cost and portability. It is specially used for visualizing subcutaneous body structures that include tendons, muscles, joints, blood vessels, testes, breast, liver, kidneys, thyroid, parathyroid glands and the neonatal brain, among others [1]. However, the US images have a granular appearance due to the inherently generated speckle noise. It obscures diagnostically important details, reduces object detectibility and makes image processing tasks such as compression and segmentation quite difficult. Hence, speckle removal is an important pre-processing step in the analysis, processing and interpretation of US images. Among the various numerous approaches, the homomorphic ones are most popular [2]-[10]. Basically, the multiplicative speckle noise is converted to an additve one by applying log-transformation. Thus, one can employ additive noise-suppression techniques to reduce the noise. The corresponding output is subjected to an exponential transformation to obtain the despeckled image. As for the additive part, the most promising results are obtained in the multi-resolution time-frequency transform domains and using the appropriate statistics of the log-transformed signal and noise. The most widely used model is the Gaussian probability density function (*pdf*) due to its mathematically tractability and ability to capture the noise statistics when the noise level is low [2]-[6]. However, unlike Gaussian, the statistics of these coefficients is actually heavy-tailed and can be described more accurately by a double-exponential *pdf*, commonly known as Fisher-Tippet *pdf* [7], [8]. The disadvantage of using this *pdf* is its mathematical intractability and associated difficulty in parameter estimation that complicates the development of an effective denoising processor. In fact, in [7], the authors estimate the noise outliers responsible for the heavy-tailedness and subtracts it from the ultrasound image to make the noise Gaussian. In [9], [10] the wavelet coefficients corresponding to log-transformed speckle are modeled with a bimodal Rayleigh and Maxwell *pdf*s respectively. However, it is unrealistic since the noise is unimodal. A normal inverse Gaussian (NIG) *pdf* is used in [11]. A generalized Nakagami *pdf* adopted in [12] to model the speckle wavelet coefficients.

The BKF *pdf* introduced by Srivastava [13] has attracted the attention of researchers for its ability to effectively model the heavy-tailed statistics of image data [14]. In this paper, the BKF *pdf* is proposed as a highly suitable prior for modeling the statistics of the log-transformed speckle in the multi-resolution contourlet transform domain. The traditionally used discrete wavelet transform (DWT) can give a good time-frequency representation of the non-stationary signal, but it has limited directional informations, only along horizontal, vertical, and diagonal directions. On the other hand the contourlet transform has the ability to describe the directionalities of image signals significantly better than the wavelet transform. In comparison to other efficient directional transforms such as the 2-D dual-tree complex wavelet transform (DT-CWT) [15], it gives more directional information, which is not fixed and rather increases along with the increase of the pyramidal decomposition levels. Also it provides a better description of arbitrary shapes and contours. In other words, it is a better descriptor of directionality and anisotropy. A maximum likelihood (ML)-based estimation technique is presented here for obtaining the parameters of the BKF *pdf*. The suitability of the *pdf* in modeling the contourlet coefficients is studied for different noise levels and compared with those of Gaussian and NIG *pdf*s using simulated noise and speckle extracted from US images.

## II. THE BKF PDF

The BKF *pdf* is expressed as [13]

$$f(x;p,c)=\frac{1}{\sqrt{\pi}\,\Gamma(p)}\left(\frac{c}{2}\right)^{-\frac{p}{2}-\frac{1}{4}}\left|\frac{x}{2}\right|^{p-\frac{1}{2}}K_{p-\frac{1}{2}}\left(\sqrt{\frac{2}{c}}|x|\right) \quad (1)$$

where $K_{p-\frac{1}{2}}$ denotes the modified Bessel function of the second kind of order $p-\frac{1}{2}$, given by [17]

$$K_{p-\frac{1}{2}}(z)=\frac{\pi^{1/2}\left(\frac{1}{2}z\right)^{p-\frac{1}{2}}}{\Gamma\left\{\left(p-\frac{1}{2}\right)+\frac{1}{2}\right\}}\int_1^{\infty}e^{-zt}\,(t^2-1)^{\left(p-\frac{1}{2}\right)-\frac{1}{2}}\,dt \quad ;\left(\Re\left(p-\frac{1}{2}\right)>-\frac{1}{2},\ |\arg z|<\frac{\pi}{2}\right) \tag{2}$$

$p$ and $c$ being scale and shape parameters, respectively, and Γ represents the gamma function. The BKF *pdf* is unimodal, symmetric around the mode, the mode necessarily not being zero. Its peakedness increase when $p$ increases. When $p = 1$, it tends towards the double exponential *pdf*, while for $p > 1$, it turns into nearly Gaussian and $p < 1$ causes a sharper peak and heavier tails. Clearly, the BKF *pdf* can be considered as the $p$-th convolution power of the double exponential [14]. Fig. 1 shows plots of a BKF *pdf* for various values of $p$ and $c$.

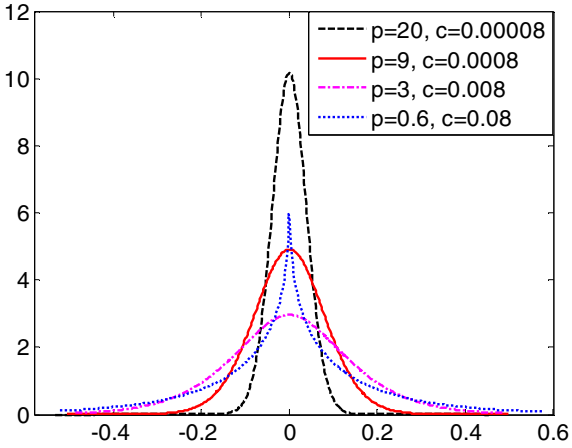



Figure 1. Plots of a BKF *pdf* for different values of $p$ and c.

The cumulants of a BKF *pdf* are given by

$$K_{2i}=p\left(\frac{c}{2}\right)^{i}\frac{(2i)!}{i},\quad i\geq 1 \tag{3}$$

the odd cumulants are zero and the even ones nonzero. Particularly, the first four cumulants are given by

$$K_1=0\ ,\ K_3=0 \tag{4}$$

$$K_2=m_2\ ,\ K_4=m_4-3{m_2}^2 \tag{5}$$

here, $m_2$ and $m_4$ are the 2nd and 4th order moments of the *pdf*. From (4) and (5), the variance and kurtosis of a BKF random variable $X$ are determined as

$$Var(X)=K_2=pc,\ \ Kurt(X)=\frac{K_4}{K_2^2}+3=\frac{3}{p}+3 \tag{6}$$

using (5) and (6), the parameters $p$ and $c$ are estimated as

$$\hat{p}=\frac{3}{Kurt(x)-3}\ ,\ \hat{c}=\frac{Var(x)}{\hat{p}} \tag{7}$$

From (7) it is seen that for a Gaussian *pdf*, since the value of kurtosis is 3, p→∞ approaches infinity whereas c→0. The moment-based estimators of the parameters $p$ and $c$, given by (7) are biased especially for a small-sized data set. This is important considering the small size of the wavelet sub bands. In this paper, a *Maximum Likelihood Estimation* (MLE)-based method is presented for estimating the parameters of a BKF *pdf*. Let the observations corresponding to a BKF random variable $X$ represented by $x_i$ where *i=1,2,3,……n.* The *Log Likelihood function* for $X$ is given by

$$log_e(L)=n\,log_e\left(\frac{1}{\sqrt{\pi}\,\Gamma(p)}\left(\frac{c}{2}\right)^{-\frac{p}{2}-\frac{1}{4}}\right)+\left(p-\frac{1}{2}\right)\sum_{i=1}^{n}log_e\left|\frac{x_i}{2}\right| +\sum_{i=1}^{n}log_e\left\{K_{p-\frac{1}{2}}\left(\sqrt{\frac{2}{c}}|x_i|\right)\right\} \tag{8}$$

Applying derivatives on (8) with respect to $c$ provides

$$\frac{\partial}{\partial c}\{log_e(L)\}=\frac{n}{c}\left(-\frac{p}{2}-\frac{1}{4}\right)+\sum_{i=1}^{n}\left(\frac{1}{K_{p-\frac{1}{2}}\left(\sqrt{\frac{2}{c}}|x_i|\right)}\cdot\frac{-\left(p-\frac{1}{2}\right)K_{p-\frac{1}{2}}\left(\sqrt{\frac{2}{c}}|x_i|\right)-\left(\sqrt{\frac{2}{c}}|x_i|\right)K_{p-\frac{3}{2}}\left(\sqrt{\frac{2}{c}}|x_i|\right)}{\left(\sqrt{\frac{2}{c}}|x_i|\right)}\cdot\left\{-\frac{|x_i|}{\sqrt{2}\,c^{3/2}}\right\}\right) \tag{9}$$

Taking derivatives on both sides of (8) with respect to $p$ yields

$$\frac{\partial}{\partial p}\{log_e(L)\}=-n\psi(p)-\frac{3n}{4}log_e\left(\frac{c}{2}\right)+\frac{1}{2}\sum_{i=1}^{n}log_e\left|\frac{x_i}{2}\right| +\sum_{i=1}^{n}\left\{\frac{1}{K_{p-\frac{1}{2}}\left(\sqrt{\frac{2}{c}}|x_i|\right)}\cdot\left(\frac{n!\left\{\frac{1}{2}\left(\sqrt{\frac{2}{c}}|x_i|\right)\right\}^{-n}}{2}\sum_{k=0}^{n-1}\frac{\left\{\frac{1}{2}\left(\sqrt{\frac{2}{c}}|x_i|\right)\right\}^{k}K_k\left(\sqrt{\frac{2}{c}}|x_i|\right)}{(n-k)k!}\right)_{p-\frac{1}{2}=n}\cdot\left(\frac{1}{2}\right)\right\} \tag{10}$$

where $\psi$ denotes the digamma function, given by

$$\psi(z)=\frac{\partial}{\partial z}\{log_e(\Gamma(z))\} \tag{11}$$

The solutions to (9) and (10) are found numerically using the Aitken's $\Delta^2$ process of acceleration method [18] which accelerates the convergence of the first-order iterative method. For this purpose, define:

$$F_1(x;\hat{p},\hat{c})=0 \tag{12}$$

$$F_2(x;\hat{p},\hat{c})=0 \tag{13}$$

where, $F_1$ and $F_2$ are the left hand side of (9), (10). The value of $\hat{p}$ and $\hat{c}$ at a given iteration are obtained as [18]

$$\left(\hat{c}_{k+2}=\hat{c}_{k+1}-\frac{(\hat{c}_{k+1}-\hat{c}_k)^2}{\hat{c}_{k+1}-2\hat{c}_k+\hat{c}_{k-1}}\right) \tag{14}$$

$$\left(\hat{p}_{k+2}=\hat{p}_{k+1}-\frac{(\hat{p}_{k+1}-\hat{p}_k)^2}{\hat{p}_{k+1}-2\hat{p}_k+\hat{p}_{k-1}}\right) \tag{15}$$

The values $\hat{p}$ and $\hat{c}$ are estimated at the $k$-th iteration of (14) and (15). The initial values, $\hat{p}_{k-1}$ and $\hat{c}_{k-1}$ are estimated from the moment-based estimator of (7). In solving (14) by subsequent iterations, $\hat{c}_k=F_1(x;\hat{p}_{k-1},\hat{c}_{k-1})$ and $\hat{c}_{k+1}=F_1(x;\hat{p}_{k-1},\hat{c}_k)$. After, In solving (15) by subsequent iterations $\hat{p}_k=F_1(x;\hat{p}_{k-1},\hat{c})$ and $\hat{p}_{k+1}=F_1(x;\hat{p}_k,\hat{c})$, where $\hat{c}$ is found from solving (14). This iterative process is continued until the following condition is satisfied:

$$|(\hat{p}_{k+2}-\hat{p}_{k+1})+(\hat{c}_{k+2}-\hat{c}_{k+1})|\leq 1\text{x}10^{-8} \tag{16}$$

A summary of the parameter estimation method is given below:

1) Find the initial values $\hat{c}_{k-1}$ , $\hat{c}_k$ and $\hat{p}_{k-1}$ , $\hat{p}_k$ .
2) Estimate $\hat{c}_{k+2}$ using (14).
3) Estimate $\hat{p}_{k+2}$ employing (15).
4) Check whether (16) is satisfied. If so, stop the iteration. Otherwise, again start the parameter estimation method from Step 1 where use the values $\hat{c}_{k-1}=\hat{c}_{k+2}$ found in Step 2 and $\hat{p}_{k-1}=\hat{p}_{k+2}$ found in step 3 as the initial values.

## III. Speckle Noise Modeling in Contourlet Domain

The contourlet transform is implemented by using a filter bank that decouples the multiscale and the directional decompositions proposed in [16]. A conceptual set up of a contourlet transform is shown in Fig. 2. The decoupling operation includes a multiscale decomposition by a Laplacian pyramid and a subsequent directional decomposition employing a directional filter bank. Basically, the contourlet transform is constructed by grouping of nearby wavelet coefficients, since they are locally correlated due to the smoothness of the contours. Therefore, a sparse expansion is obtained for natural images by first applying a multiscale transform, followed by a local directional transform to gather the nearby basis functions at the same scale into linear structures. Thus it constitutes a wavelet-like transform for *edge* detection and then a local directional transform for *contour segment* detection. The overall result is an image expansion using basic elements that are like contour segments, and hence the name *contourlets*.

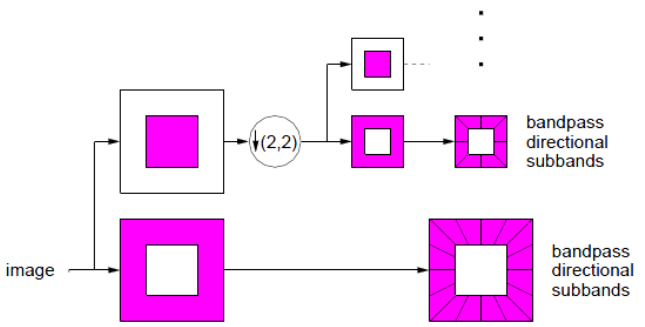



Figure 2. The contourlet filter bank: first, a multiscale decomposition into octave bands by the Laplacian pyramid is computed, and then a directional filter bank is applied to each bandpass channel [16].

Generally, the speckle noise is described as a multiplicative phenomenon. The reflectivity image $\mathcal{F}$ is assumed to be corrupted by the speckle noise $\eta$ and an additive noise (such as thermal noise) $\eta_a$ as [5]

$$\mathcal{F}_n = \mathcal{F}\,\eta + \eta_a \qquad (17)$$

Since the effect of additive noise is insignificant as compared to the multiplicative speckle**,** (17) can be written as [5]

$$\mathcal{F}_n = \mathcal{F}\,\eta \qquad (18)$$

After applying log-transformation on (18), we obtain

$$f = g + \Upsilon \qquad (19)$$

where $f=log(\mathcal{F}_n)$, $g=log(\mathcal{R})$ *and* $\Upsilon=log(\eta)$. As the log-transformed image is subjected to wavelet transform, one gets

$$y = \varepsilon + x \qquad (20)$$

where $y, \varepsilon$ and $x$, respectively, represent the coefficients corresponding to *d, S* and Υ. For the purpose of modeling, the BKF parameters, $p$ and $c$, are estimated using the proposed MLE-based method from the contourlet transform coefficients of the log-transformed noise. The speckle noise can be simulated by low-pass filtering a complex Gaussian random field, and then taking the magnitude of the filtered output. The filtering is carried out using a 3x3 window, since such a short-term correlation is sufficient to account for real speckle noise [3]. The log-transformed noise is decomposed in the contourlet transform domain using the contourlet toolbox [19] with many different orientations. The modeling performance of the BKF *pdf* is compared with that of the Gaussian and normal inverse Gaussian (NIG) *pdf*s using the well-known Kolmogorov-Smirnov (KS) statistics and variance stabilized *pp-plot*. The *pdf* of a zero-mean Gaussian distributed random variable, x, is given by

$$P_x(x) = \frac{1}{\sigma_x\sqrt{2\pi}}\, exp\left(-\frac{x^2}{2\sigma_x^2}\right)\ ;\ -\infty < x < \infty \qquad (21)$$

where $\sigma_x$ is the standard deviation of signal $x$, which determines the spread of the density function. The value of $\sigma_x$ is estimated as

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i)^2} \qquad (22)$$

TABLE I
Values of the Kolmogorov-Smirnov (*KS*) statistics ($d_{KS}$) calculated in the contourlet transform domain. The subscripts represent the corresponding decomposition level

| Contourlet Sub bands P=Pyramidal D=Directional | | *Values of the Kolmogorov-Smirnov (KS) Statistics ($d_{KS}$) for Noise Standard Deviation 0.3* | | |
|---|---|---|---|---|
| | | BKF | Gaussian | NIG |
| $P_3$ | $D_1$ | **0.0413** | 0.0417 | 0.0441 |
| | $D_2$ | **0.0528** | 0.0593 | 0.0659 |
| | $D_3$ | 0.0551 | 0.0517 | **0.0474** |
| | $D_4$ | **0.0401** | 0.0559 | 0.0404 |
| $P_4$ | $D_1$ | **0.0735** | 0.0866 | 0.0753 |
| | $D_2$ | **0.0408** | 0.0467 | 0.0425 |
| | $D_3$ | **0.0281** | 0.0369 | 0.0320 |
| | $D_4$ | **0.0185** | 0.0196 | 0.0199 |
| | $D_5$ | **0.0314** | 0.0334 | 0.0321 |
| | $D_6$ | **0.0243** | 0.0270 | 0.0309 |
| | $D_7$ | **0.0420** | 0.0454 | 0.0461 |
| | $D_8$ | **0.0221** | 0.0311 | 0.0271 |
| $P_5$ | $D_1$ | **0.0148** | 0.0231 | 0.0186 |
| | $D_2$ | **0.0344** | 0.0353 | 0.0348 |
| | $D_3$ | **0.0332** | 0.0340 | 0.0337 |
| | $D_4$ | **0.0151** | 0.0177 | 0.0175 |
| | $D_5$ | **0.0142** | 0.0158 | 0.0159 |
| | $D_6$ | **0.0190** | 0.0241 | 0.0200 |
| | $D_7$ | **0.0275** | 0.0289 | 0.0344 |
| | $D_8$ | **0.0234** | 0.0267 | 0.0402 |
| | $D_9$ | **0.0152** | 0.0153 | 0.0190 |
| | $D_{10}$ | **0.0182** | 0.0187 | 0.0190 |
| | $D_{11}$ | **0.0157** | 0.0166 | 0.0220 |
| | $D_{12}$ | **0.0186** | 0.0264 | 0.0197 |
| | $D_{13}$ | **0.0277** | 0.0312 | 0.0297 |
| | $D_{14}$ | **0.0201** | 0.0209 | 0.0227 |
| | $D_{15}$ | 0.0217 | 0.0281 | **0.0214** |
| | $D_{16}$ | **0.0310** | 0.0314 | 0.0326 |
| $P_6$ | $D_1$ | **0.0129** | 0.0148 | 0.0474 |
| | $D_2$ | **0.0144** | 0.0175 | 0.0411 |
| | $D_3$ | **0.0117** | 0.0133 | 0.0518 |
| | $D_4$ | **0.0179** | 0.0197 | 0.0451 |
| | $D_5$ | **0.0214** | 0.0223 | 0.0413 |
| | $D_6$ | **0.0130** | 0.0137 | 0.0350 |
| | $D_7$ | **0.0257** | 0.0320 | 0.0473 |
| | $D_8$ | **0.0101** | 0.0113 | 0.0152 |
| | $D_9$ | **0.0246** | 0.0249 | 0.0470 |
| | $D_{10}$ | **0.0175** | 0.0182 | 0.0382 |
| | $D_{11}$ | **0.0116** | 0.0118 | 0.0435 |
| | $D_{12}$ | **0.0229** | 0.0246 | 0.0454 |
| | $D_{13}$ | **0.0267** | 0.0290 | 0.0403 |
| | $D_{14}$ | **0.0122** | 0.0138 | 0.0375 |
| | $D_{15}$ | **0.0178** | 0.0195 | 0.0443 |
| | $D_{16}$ | **0.0126** | 0.0172 | 0.0344 |
| | $D_{17}$ | **0.0120** | 0.0122 | 0.0331 |
| | $D_{18}$ | **0.0286** | 0.0309 | 0.0476 |
| | $D_{19}$ | **0.0275** | 0.0283 | 0.0348 |
| | $D_{20}$ | **0.0131** | 0.0139 | 0.0267 |
| | $D_{21}$ | **0.0107** | 0.0116 | 0.0353 |
| | $D_{22}$ | **0.0125** | 0.0127 | 0.0433 |
| | $D_{23}$ | **0.0119** | 0.0136 | 0.0351 |
| | $D_{24}$ | **0.0110** | 0.0121 | 0.0463 |
| | $D_{25}$ | **0.0115** | 0.0117 | 0.0168 |
| | $D_{26}$ | **0.0132** | 0.0136 | 0.0262 |
| | $D_{27}$ | **0.0121** | 0.0134 | 0.0254 |
| | $D_{28}$ | **0.0102** | 0.0108 | 0.0367 |
| | $D_{29}$ | **0.0108** | 0.0126 | 0.0445 |
| | $D_{30}$ | **0.0159** | 0.0184 | 0.0434 |
| | $D_{31}$ | **0.0188** | 0.0200 | 0.0379 |
| | $D_{32}$ | **0.0145** | 0.0154 | 0.0272 |

The NIG *pdf* is expressed as [20]

$$f_x(x;\theta) = \frac{\alpha\delta}{\pi}\frac{exp\{p(x)\}}{q(x)}K_1[\alpha q(x)] \quad (23)$$

where $p(x) = \delta\sqrt{(\alpha^2-\beta^2)} + \beta(x-\mu)$ and $q(x) = \sqrt{(x-\mu)^2+\delta^2}$.

The parameters of the NIG *pdf* are obtained as [17]

$$\hat{\delta} = \sqrt{\hat{\kappa}^{(2)}\xi(1-\rho^2)} \quad ; \ (\delta > 0) \quad (24)$$

$$\hat{\alpha} = \frac{\xi}{\hat{\delta}(1-\rho^2)} \quad (25)$$

$$\hat{\beta} = \hat{\alpha}\rho \quad ; \ (0 \le |\beta| < \alpha) \quad (26)$$

$$\hat{\mu} = \hat{\kappa}^{(1)} - \rho\sqrt{\hat{\kappa}^{(2)}\xi} \quad ; \ (-\infty < \mu < \infty) \quad (27)$$

where $\hat{\kappa}^{(1)}, \hat{\kappa}^{(2)}, \hat{\kappa}^{(3)}, \hat{\kappa}^{(4)}$ are the first four cumulants from sample data, the skewness $\hat{\gamma}_3 = \hat{\kappa}^{(3)}/\left[\hat{\kappa}^{(2)}\right]^{3/2}$, normalized kurtosis $\hat{\gamma}_4 = \hat{\kappa}^{(4)}/\left[\hat{\kappa}^{(2)}\right]^2$ and auxiliary variables $\xi = 3\left(\hat{\gamma}_4 - \frac{4}{3}\hat{\gamma}_3^2\right)^{-1}$, $\rho = \frac{\hat{\gamma}_3}{3}\sqrt{\xi}$.

## IV. Results

In this section, the results of our experiments to determine the efficacy of various *pdf*s in modeling the speckle noise in contourlet transform domain are described. The well-known Kolmogorov-Smirnov (KS) statistics and *p-p* plots are used as the matrices for comparing the performance of different *pdf*s in modeling the speckle noise. The KS statistics is defined as

$$d_{KS} = \max_{x\in R} |F_e(x) - F_a(x)| \quad (28)$$

Here, $d_{ks}$, $F_a(x)$ and $F_e(x)$ denote the KS statistics, cumulative density function (*cdf*) of the modeling *pdf* and the empirical *cdf*, respectively [20], [21].

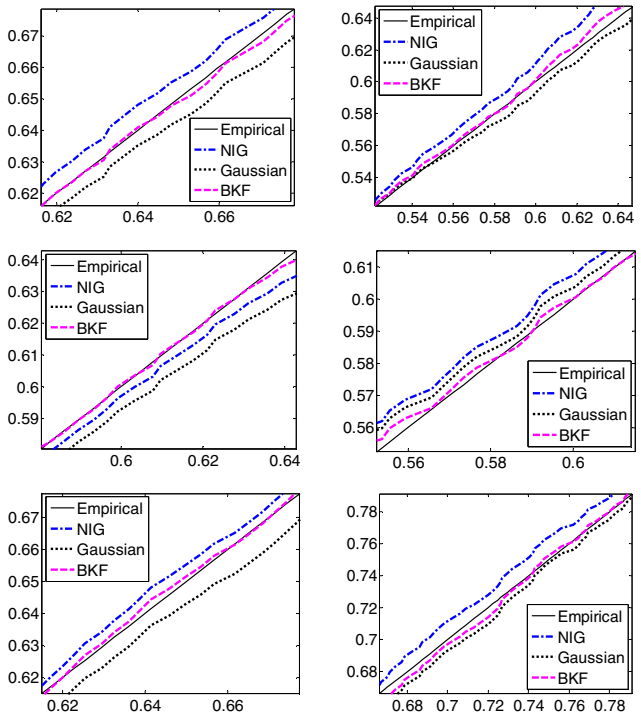



Figure 3. *PP-plots* for the contourlet transform sub bands of some different orientations. The figures in the first, second and third rows show the contourlet transform at pyramidal sub bands-4,5,6 where as the first and second columns show the directional sub bands-4,8 at noise standard deviation of 0.3 and 0.5 respectively.

The *pp-plot* is obtained by plotting $F_a(x)^t$ against $F_e(x)^t$ where a linear plot means excellent fitting [5]:

$$F_a(x)^t = \frac{2}{\pi}\, arcsin\left(\sqrt{F_a(x)}\right) \quad (29)$$

$$F_e(x)^t = \frac{2}{\pi}\, arcsin\left(\sqrt{F_e(x)}\right) \quad (30)$$

where *t* denotes the transpose operation.

TABLE II
Values of the Kolmogorov-Smirnov (*KS*) statistics ($d_{KS}$) calculated in the contourlet transform domain. The subscripts represent the corresponding decomposition level

| Contourlet Sub bands P=Pyramidal D=Directional | | *Values of the Kolmogorov-Smirnov (KS) Statistics ($d_{KS}$) for Noise Standard Deviation 0.5* | | |
|---|---|---|---|---|
| | | BKF | Gaussian | NIG |
| $P_3$ | $D_1$ | **0.0413** | 0.0421 | 0.0443 |
| | $D_2$ | **0.0422** | 0.0426 | 0.0425 |
| | $D_3$ | **0.0317** | 0.0411 | 0.0330 |
| | $D_4$ | **0.0310** | 0.0339 | 0.0322 |
| $P_4$ | $D_1$ | **0.0887** | 0.0896 | 0.0892 |
| | $D_2$ | **0.0493** | 0.0500 | 0.0496 |
| | $D_3$ | **0.0406** | 0.0417 | 0.0420 |
| | $D_4$ | 0.0328 | 0.0330 | **0.0324** |
| | $D_5$ | **0.0313** | 0.0315 | 0.0319 |
| | $D_6$ | **0.0227** | 0.0278 | 0.0236 |
| | $D_7$ | **0.0296** | 0.0311 | 0.0319 |
| | $D_8$ | **0.0337** | 0.0344 | 0.0345 |
| $P_5$ | $D_1$ | **0.0225** | 0.0240 | 0.0232 |
| | $D_2$ | **0.0340** | 0.0360 | 0.0371 |
| | $D_3$ | **0.0129** | 0.0132 | 0.0170 |
| | $D_4$ | **0.0236** | 0.0247 | 0.0252 |
| | $D_5$ | **0.0407** | 0.0417 | 0.0418 |
| | $D_6$ | **0.0132** | 0.0141 | 0.0144 |
| | $D_7$ | **0.0141** | 0.0244 | 0.0154 |
| | $D_8$ | **0.0212** | 0.0290 | 0.0224 |
| | $D_9$ | **0.0090** | 0.0094 | 0.0107 |
| | $D_{10}$ | **0.0173** | 0.0208 | 0.0188 |
| | $D_{11}$ | **0.0122** | 0.0156 | 0.0128 |
| | $D_{12}$ | **0.0170** | 0.0233 | 0.0200 |
| | $D_{13}$ | **0.0145** | 0.0177 | 0.0147 |
| | $D_{14}$ | **0.0318** | 0.0325 | 0.0338 |
| | $D_{15}$ | **0.0308** | 0.0318 | 0.0311 |
| | $D_{16}$ | **0.0305** | 0.0305 | 0.0307 |
| $P_6$ | $D_1$ | **0.0300** | 0.0305 | 0.0345 |
| | $D_2$ | **0.0322** | 0.0362 | 0.0382 |
| | $D_3$ | **0.0124** | 0.0131 | 0.0161 |
| | $D_4$ | **0.0190** | 0.0216 | 0.0228 |
| | $D_5$ | **0.0188** | 0.0207 | 0.0288 |
| | $D_6$ | **0.0201** | 0.0208 | 0.0218 |
| | $D_7$ | 0.0187 | 0.0213 | **0.0119** |
| | $D_8$ | **0.0179** | 0.0198 | 0.0222 |
| | $D_9$ | **0.0181** | 0.0196 | 0.0319 |
| | $D_{10}$ | **0.0168** | 0.0270 | 0.0177 |
| | $D_{11}$ | **0.0119** | 0.0168 | 0.0125 |
| | $D_{12}$ | **0.0166** | 0.0214 | 0.0180 |
| | $D_{13}$ | **0.0264** | 0.0347 | 0.0289 |
| | $D_{14}$ | **0.0137** | 0.0285 | 0.0143 |
| | $D_{15}$ | **0.0140** | 0.0227 | 0.0154 |
| | $D_{16}$ | **0.0193** | 0.0262 | 0.0210 |
| | $D_{17}$ | **0.0249** | 0.0314 | 0.0266 |
| | $D_{18}$ | **0.0214** | 0.0251 | 0.0221 |
| | $D_{19}$ | **0.0116** | 0.0151 | 0.0120 |
| | $D_{20}$ | **0.0244** | 0.0294 | 0.0267 |
| | $D_{21}$ | **0.0201** | 0.0210 | 0.0206 |
| | $D_{22}$ | **0.0184** | 0.0232 | 0.0205 |
| | $D_{23}$ | **0.0178** | 0.0217 | 0.0207 |
| | $D_{24}$ | **0.0220** | 0.0337 | 0.0222 |
| | $D_{25}$ | **0.0249** | 0.0382 | 0.0277 |
| | $D_{26}$ | **0.0159** | 0.0253 | 0.0197 |
| | $D_{27}$ | **0.0185** | 0.0293 | 0.0208 |
| | $D_{28}$ | **0.0155** | 0.0199 | 0.0171 |
| | $D_{29}$ | **0.0183** | 0.0228 | 0.0215 |
| | $D_{30}$ | **0.0199** | 0.0224 | 0.0218 |
| | $D_{31}$ | **0.0161** | 0.0312 | 0.0198 |
| | $D_{32}$ | **0.0183** | 0.0275 | 0.0208 |

The values of the KS statistics for various orientations and noise levels are provided in Tables I and II. It is seen that the BKF *pdf*, in general, gives lower values as compared to those

of the other *pdf*s, indicating a closer match with the empirical *pdf*. From the *p-p plots* shown in Fig. 3, it is seen that the BKF *pdf* provides a better match with the underlying empirical one.

TABLE III
Values of the Kolmogorov-Smirnov (*KS*) statistics ($d_{KS}$) calculated in the contourlet transform domain. The subscripts represent the corresponding decomposition level

| Contourlet Sub bands P=Pyramidal D=Directional | | *Values of the Kolmogorov-Smirnov (KS) Statistics ($d_{KS}$)* | | |
|---|---|---|---|---|
| | | **BKF** | **Gaussian** | **NIG** |
| $P_3$ | $D_1$ | 0.0831 | 0.0821 | **0.0810** |
| | $D_2$ | **0.0582** | 0.0687 | 0.0584 |
| | $D_3$ | **0.0917** | 0.0919 | 0.0940 |
| | $D_4$ | **0.0910** | 0.1032 | 0.0989 |
| $P_4$ | $D_1$ | **0.0707** | 0.0863 | 0.0770 |
| | $D_2$ | **0.0393** | 0.0569 | 0.0393 |
| | $D_3$ | **0.0506** | 0.0855 | 0.0550 |
| | $D_4$ | 0.0621 | 0.0731 | **0.0440** |
| | $D_5$ | **0.0403** | 0.0659 | 0.0524 |
| | $D_6$ | **0.0406** | 0.0723 | 0.0465 |
| | $D_7$ | **0.0256** | 0.0477 | 0.0313 |
| | $D_8$ | **0.0587** | 0.0844 | 0.0666 |
| $P_5$ | $D_1$ | **0.0505** | 0.1016 | 0.0550 |
| | $D_2$ | **0.0210** | 0.0996 | 0.0237 |
| | $D_3$ | **0.0379** | 0.1287 | 0.0500 |
| | $D_4$ | **0.0296** | 0.1060 | 0.0516 |
| | $D_5$ | **0.0167** | 0.1029 | 0.0180 |
| | $D_6$ | **0.0200** | 0.1086 | 0.0300 |
| | $D_7$ | **0.0141** | 0.0862 | 0.0154 |
| | $D_8$ | **0.0162** | 0.0983 | 0.0164 |
| | $D_9$ | **0.0250** | 0.1409 | 0.0259 |
| | $D_{10}$ | 0.0243 | 0.0688 | **0.0238** |
| | $D_{11}$ | **0.0422** | 0.0853 | 0.0573 |
| | $D_{12}$ | **0.0470** | 0.1216 | 0.0761 |
| | $D_{13}$ | **0.0345** | 0.1094 | 0.0632 |
| | $D_{14}$ | **0.0438** | 0.0778 | 0.0502 |
| | $D_{15}$ | **0.0618** | 0.0981 | 0.0730 |
| | $D_{16}$ | **0.0325** | 0.0489 | 0.0328 |
| $P_6$ | $D_1$ | **0.0390** | 0.1010 | 0.1704 |
| | $D_2$ | **0.0522** | 0.1172 | 0.1311 |
| | $D_3$ | **0.0384** | 0.1122 | 0.1249 |
| | $D_4$ | **0.0490** | 0.1231 | 0.0825 |
| | $D_5$ | **0.0348** | 0.1101 | 0.1147 |
| | $D_6$ | **0.0251** | 0.0897 | 0.1280 |
| | $D_7$ | **0.0832** | 0.1264 | 0.0851 |
| | $D_8$ | 0.0278 | 0.1337 | **0.0180** |
| | $D_9$ | **0.0511** | 0.1196 | 0.1319 |
| | $D_{10}$ | **0.0388** | 0.1270 | 0.0777 |
| | $D_{11}$ | **0.0329** | 0.1068 | 0.1125 |
| | $D_{12}$ | **0.0386** | 0.1114 | 0.1020 |
| | $D_{13}$ | **0.0264** | 0.1047 | 0.1589 |
| | $D_{14}$ | **0.0477** | 0.1285 | 0.1143 |
| | $D_{15}$ | **0.0610** | 0.1227 | 0.1454 |
| | $D_{16}$ | **0.0550** | 0.1262 | 0.1910 |
| | $D_{17}$ | **0.0559** | 0.1414 | 0.0566 |
| | $D_{18}$ | **0.4714** | 0.1251 | 0.0721 |
| | $D_{19}$ | **0.0176** | 0.0951 | 0.0320 |
| | $D_{20}$ | **0.0244** | 0.0694 | 0.0267 |
| | $D_{21}$ | **0.0401** | 0.0810 | 0.0406 |
| | $D_{22}$ | **0.0374** | 0.0932 | 0.0505 |
| | $D_{23}$ | **0.0378** | 0.1117 | 0.0607 |
| | $D_{24}$ | **0.0320** | 0.1337 | 0.0722 |
| | $D_{25}$ | **0.0499** | 0.1382 | 0.0877 |
| | $D_{26}$ | **0.0359** | 0.1053 | 0.0597 |
| | $D_{27}$ | **0.0375** | 0.0793 | 0.0408 |
| | $D_{28}$ | **0.0455** | 0.0699 | 0.0471 |
| | $D_{29}$ | **0.0283** | 0.0728 | 0.0295 |
| | $D_{30}$ | **0.0199** | 0.0624 | 0.0218 |
| | $D_{31}$ | 0.0241 | 0.0612 | **0.0198** |
| | $D_{32}$ | **0.0473** | 0.0975 | 0.0608 |

Real US images are also used in our investigation. For this purpose, a number of ultrasound images obtained from the authors of [3] is used in our study. Fig. 4 shows a real US image and it's filtered version obtained by the Homomorphic Weiner method [2]. Since the true reflectivity is unknown, the filtered image is considered approximately noise-free; the corresponding speckle is extracted by dividing the noisy US image with this filtered image. The log of this speckle noise is modeled using the various *pdf*s in contourlet transform domain. Table III shows the values of the KS statistics where lower values of KS statistics indicates a better match with the underlying statistics of noise in contourlet transform domain. The efficacy of BKF *pdf* in modeling the noise is further illustrated in Fig. 5 which shows the *pp-plots* for various *pdf*s of different orientations. It is seen that the plot for the BKF *pdf* is in general, close to linear one, thus indicating a better match with the empirical *pdf*.

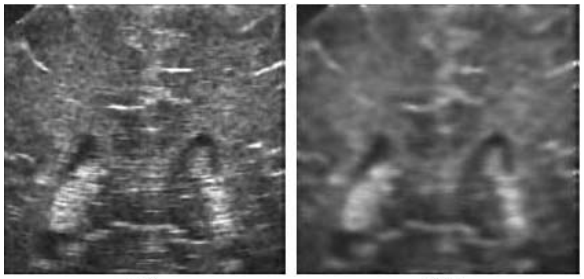



Figure 4. US Image (a) healthy Neonatal Brain (Coronal View) & corresponding (b) denoised image obtained by Homomorphic Wiener filter using a 5x5 window.

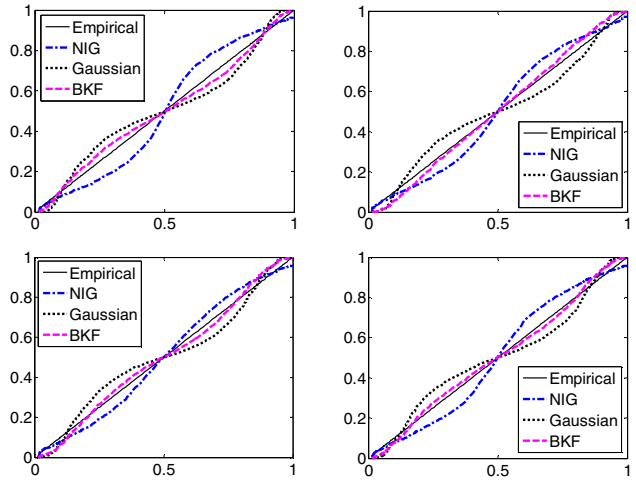



Figure 5. *PP-plots* for the contourlet transform sub bands of some different orientations. The figures in the first and second rows show the contourlet transform at pyramidal sub bands-5,6 whereas the first and second columns show the directional sub bands-12, 24 respectively.

## V. Conclusion

In this paper, the appropriateness of the Bessel K-Form (BKF) *pdf* as a highly suitable model for describing the statistics of log-transformed speckle noise in contourlet transform domain has been demonstrated. A Maximum Likelihood (ML)-based Estimator (MLE) has been developed for this purpose. The MLE equations have been solved using the Aitken's $\Delta^2$ process of acceleration method. For the case of simulated noise, it has been shown that the BKF *pdf* is highly suitable for modeling the log-transformed speckle in contourlet transform domain, better than the NIG and the Gaussian *pdf*s. The suitability of the BKF *pdf* has also been illustrated for the case of real US images. The findings of this study may help researchers in developing effective statistical methods for reducing speckle noise from medical US images. There are some limitations regarding the parameter estimation process

since it does not have a closed-form expression, necessary to have reduced complexity.

## REFERENCES


[1] Kayvan Najarian And Robert Splinton, " Biomedical Signal and Image Processing" *CRC Press., Taylor and Francis,* 2006.

[2] A. K., Jain *Fundamentals of Digital Image Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1989.

[3] A. Pizurica, W. Philips, I. Lemahieu. And M. Acheroy, 'A Versatile Wavelet Domain Noise Filtration Technique for Medical Imaging'', *IEEE Trans. Med. Imag.*, 22, pp. 323–331, 2003.

[4] S. Gupta, R.C. Chauhan and S.C. Saxena, ''Wavelet-based statistical approach for speckle reduction in medical ultrasound images,'' *Med. Biol. Eng. Comput.*, 42, 189-192, 2004.

[5] A. Achim, P. Tsakalides, and A. Bezarianos, ''Novel Bayesian multiscale method for speckle removal in medical ultrasound images,'' *IEEE Trans. on Medical Imaging*, vol. 20, pp. 772–783, 2001.

[6] M.I.H. Bhuiyan, M.O. Ahmad and M.N.S. Swamy, ''Spatially Adaptive Thresholding in Wavelet Domain For Despeckling Of Ultrasound Images,'' *IET Image Processing*, vol. 3, no. 3, pp. 147–162, 2009.

[7] Oleg V. Michailovich and Allen Tannenbaum, ''Despeckling of Medical Ultrasound Images,'' *IEEE Transactions On Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 53, no.1, 2006.

[8] Peter C. Tay, C. D. Garson, S. T. Acton and J. A. Hossack, ''Ultrasound despeckling for contrast enhancement,'' *IEEE Trans. on Image Processing*, vol. 19, no. 7, pp. 1847-1860, 2010.

[9] S. Gupta, R.C. Chauhan and S.C. Saxena, ''Locally adaptive wavelet domain Bayesian processor for denoising medical ultrasound images using Speckle modeling based on Rayleigh distribution,'' IEE Proc.-Vis. Image, and Signal Process., vol. 152, no. 1, 2005.

[10] M.I.H. Bhuiyan, M.O. Ahmad and M.N.S. Swamy, " New Spatially Adaptive Wavelet-based Method for despeckling of Ultrasound Images,"Proceedings of IEEE Symposium on Circuits and Systems, pp. 2347-2351, 2007.

[11] Stian Solbø and Torbjørn Eltoft, ''Homomorphic Wavelet-Based Statistical Despeckling of SAR Images,'' *IEEE Trans. on Geoscience And Remote Sensing*, vol. 42, no. 4, 2004.

[12] S. Gupta, L. Kaur, R.C. Chauhan and S.C. Saxena, "A versatile method for visual enhancement of medical ultrasound images," *Digital Signal Processing*, vol. 17, no.3, pp. 542-560, 2007.

[13] A. Srivastava, A.B. Lee, E.P. Simoncelli and S.-C. Zhu, "On Advances in Statistical Modeling of Natural Images," *Journal of Mathematical Imaging and Vision*, vol. 18, pp.17–33, 2003.

[14] J. M. Fadili and L. Boubchir, "Ananytical form for a Bayesian wavelet estimator images using the Bessel K-Form densities," *IEEE Trans. On Image Processing*, vol. 14, no. 2, 2005.

[15] Ufuk Bal, "Dual tree complex wavelet transform based denoising of optical microscopy images,"Biomedical Optics Express, vol. 3, no. 12, pp. 3231-3239, 2012.

[16] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," IEEE Transactions on Image Processing, vol. 14, no. 12, pp. 2091–2096, Dec 2005.

[17] Milton Abramowitz and Irene A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 1970.

[18] S. Balachandra Rao, C. K. Shantha, "Numerical Methods with Programs in Basic, Fortran and Pascal"

[19] M. N. Do, "Contourlet toolbox." [Online]. Available: http://www.ifp.uiuc.edu/minhdo/software/contourlettoolbox.tar

[20] A. Hanssen and T. A. Øigard, "The normal inverse Gaussian distribution: A versatile model for heavy tailed stochastic processes", *Proceedings of IEEE ICASSP,* vol. 6, pp. 3985-3988, 2001.

[21] W. H. Press, S. A. Teukolsky, W. T. Vellerling and B. P. Flannery, *Numerical recipes in C: The Art of Scientific Computing*, UK: Cambridge University Press, 1999.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Detection of Cognitive State for Brain-Computer Interfaces

Md. Abu Baker Siddique Akhanda* , Shaon Md. Foorkanul Islam and Md. Mostafizur Rahman
Department of Electronics & Communication Engineering
Khulna University of Engineering & Technology, Khulna-9203, Bangladesh
*E-mail: Shumel0809@ymail.com

***Abstract*—Brain–Computer Interface (BCI) requires generating control signals for external device by analyzing and processing the internal brain signal. Cognitive or Mental State detection has its advantages in asynchronous BCI where the subjects are involved to the interface in response to some external stimulation. But the main problem is feature extraction and classification for different Cognitive State. In this research work, four Cognitive States Resting State (RS), Thought (TH), Memory (MR) and Emotion (EM) have been detected by collecting, processing and classifying Electroencephalogram (EEG) signals from six subjects. EEG signals were analyzed to find out the features such as spectral Power, frequency band combination ratios and linear combination of power of EEG frequency bands. A three layer BP neural network was structured to use as classifier for pattern recognition. Results indicate that different Cognitive States were perfectly identified with higher classification performance and classification performance remains approximately invariant to the number of NN hidden layer units.**

***Keywords— Cognitive State, Brain-Computer Interface (BCI), EEG, Neural Network classifier.***

## I. INTRODUCTION

BRAIN-computer interface (BCI) systems are the recent field of the modern technology that creates an interface or communication path between human brain and external devices. It is the process where internal brain signals are processed to generate control signal for the external devices. BCI systems have its potential applications for the patients suffering spinal cord injury, paralysis and brainstem stroke etc. Because this is a "hands off" control procedure people who have no or less control over their muscle can use BCI system to convert their thoughts in to action [1] & [2]. BCI system also has similar application for creating virtual environment, keyboard etc. [3]. So the advantages of BCI systems for injured or paralyzed patients encourage more development of BCI systems.

Cognitive states or mental states are the inside human mental state during the interaction with the system through men machine interface [4]. Real time BCI application requires non-invasive asynchronous detection of cognitive states where a subject has to undergo a specific sequence of mental state in response to some external stimulation [5] & [6]. One and most popular method of asynchronous BCI system is collecting and processing EEG signals non-invasively. But the key problems are, it's very difficult to create cognitive state environment, collected signals can be affected by the motion artifact and process to find out the features from EEG and classification to detect different cognitive states.

EEG signals are divided in to five frequency bands named Delta 0-3.5Hz, Theta 4-7Hz, Alpha 8-13Hz, Beta 14-34Hz and Gamma > 35Hz [7]. Spectral power of different frequency band of EEG was used by different researchers [8] & [9]. Spectral power and linear combination of this frequency bands to reduce the feature number were used for BCI system [10]. Using gamma band to increase the accuracy rate of mental task detection were also introduced in [11]. Using Approximate Entropy (ApEn) to classify different mental task was also introduced in [12].

In this study, four cognitive states Resting State (RS), Thought (TH), Memory (MR) and Emotion (EM) have been detected by collecting, processing and classifying Electroencephalogram (EEG) signals. EEG signals were collected by using BIOPAC system, processed to find out features by spectral power, frequency band combination ratios and linear combination of five frequency bands alpha, beta, theta, delta and gamma. Using gamma band to increase the classification performances of mental task are described in [11] as gamma band also shows significant variations with different mental state. So we used gamma band with other frequency bands of EEG to extract features for classification. A three layer back propagation neural network was used for training and testing the collected features for the detection of different cognitive states.

## II. METHODOLOGY

### *A. Subjects*

EEG signals of 6 healthy subjects for different cognitive states were collected by using BIOPAC software at Biomedical Signal Processing Laboratory, Department of BME of EEE Faculty in Khulna University of engineering and technology (KUET), Bangladesh. All the subjects are

undergraduate student and are of good mental and physical condition.

*B. Data Acquisition*

EEG signals were collected from different subjects where each subject has to undergo a specific sequence of data collection procedure. As it is said earlier, EEG signals were collected for four different cognitive states as Resting State (RS) which includes eye close (RSEC) and eye open (RSEO) state, Thought (TH), Memory (MR) and Emotion (EM) state. All the subjects were asked to sit in an armchair and stared at a computer monitor placed approximately 60 centimeters away from the subject at eye level. They were asked to keep their arms relaxed and comfortable and to avoid eye movements during the recordings. For RS state, EEG signals were collected for sitting resting state but for two conditions: (a) eyes close and (b) eyes open. Then for TH, EEG signals were collected when the subjects were challenged by different mental mathematical problem i.e., $x^2 + 5x + 6$, here what is the value of x. EEG signals for MR state were collected when the subjects asked to memorize some positive random number sequence displaying on the CRT screen. Twenty seconds after the positive set disappears from the screen, one character will be displayed on and subject was asked to judge whether it was in the positive set or not. This process of memory data collection was described in [4]. EM is defined as the state when the emotional aspect is dominating in the mental activities. Happiness, laughing, confusion and embarrassment are supposed to be major factors evoking the state of EM. EEG signals for EM state were collected by providing subject with different audio tracks and video slide.

All the physiological variables are collected by using BIOPAC MP36 unit. Required equipment's are BIOPAC electrode lead set (SS2L), BIOPAC disposable vinyl electrodes (EL503), BIOPAC data acquisition unit (MP36) with cable and power. Three electrodes were placed on the subject's scalp according to the [13]. Red electrode was placed on the occipital lobe, the white electrode was placed behind ear and the ground black lead was placed on ear lobe. Practical view of EEG collection of a subject is shown in figure 1.

Figure 1: Pictorial view of EEG measurement in BME Lab KUET

*C. Feature extraction*

Feature from collected EEG signal were generated by two methods:

A. Calculating the linear combination of frequency band power for Alpha, theta and delta band to reduce the feature numbers described in [10].
B. Using gamma band with alpha, theta and delta to generate features to increase the classification performance and also to increase the timing performance (proposed method).

❖ Beta band was taken for both method but has less variation than other frequency band.

For effective analysis of collected EEG signals, in method B feature extracted from the EEG signal in four steps:

i) Filtering to an unfiltered EEG lead signal to generate the following five standard EEG bands: Alpha, Beta, Theta, Delta, and Gamma. Filtering is performed using IIR low pass + high pass combination filters;
ii) Calculation of band power in predefined frequency bands in 1 second intervals;
iii) Calculation of frequency band power combination ratio for seven different combinations.
iv) Calculation of linear combination of frequency band powers.

So, all the collected EEG signals at first filtered to generate five EEG frequency bands Alpha, Beta, Theta, Delta, and Gamma. Spectral power of this frequency bands were calculated by using Power spectral density.

**Power spectral density** (PSD), which describes how the power of a signal or time series is distributed over the different frequencies. Here, power can be the actual physical power, or more often, for convenience with abstract signals, can be defined as the squared value of the signal. The total power $P$ of a signal $x(t)$ is the following time average:

$$P = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)^2 dt \quad (1)$$

If $\check{x}(w)$ is the Fourier transform of signal $x(t)$ then

$$\check{x}(w) = \frac{1}{\sqrt{T}} \int_{0}^{T} x(t) e^{-i\omega t} dt \quad (2)$$

Now PSD can be defined as,

$$S(w) = \lim_{T\to\infty} E[|\check{x}(w)|^2] \quad (3)$$

Where $E$ is the expected value of $\check{x}(w)$

So, five set of spectral power value for each cognitive state were generated by using this process.

In the next step, frequency band power combination was calculated by the method described in [10]. According to the [10] alpha, theta and delta band generates large variations with different cognitive states in comparison with the beta and gamma band. Now frequency band power combination can be calculated by using $(C-D)/(C+D)$ ratio, where C & D are two different frequency bands. So by putting alpha, beta and delta band in C group and beta, gamma in D group six different combination were calculated.

TABLE I. FREQUENCY BAND POWER COMBINATION

| Frequency band combination | Combination ratio |
|---|---|
| Alpha & Beta | $(\alpha-\beta)/(\alpha+\beta)$ |
| Alpha & Gamma | $(\alpha-\gamma)/(\alpha+\gamma)$ |
| Delta & Beta | $(\delta-\beta)/(\delta+\beta)$ |
| Delta & Gamma | $(\delta-\gamma)/(\delta+\gamma)$ |
| Theta & Beta | $(\theta-\beta)/(\theta+\beta)$ |
| Theta & Gamma | $(\theta-\gamma)/(\theta+\gamma)$ |

Again according to the [11] gamma band can also be used for increasing the classification ratio. So another power combination ratio $(\gamma-\beta)/(\gamma+\beta)$ was used with above six combinations.

So, five set of power combination data were generated for each cognitive state by using this process.

In the next step linear combination of frequency bands were calculated by using following formula,

Linear Power combination =
$$((\alpha+\delta+\theta)-(\beta+\gamma))/(\alpha+\delta+\theta+\beta+\gamma) \quad (4)$$

Now if we include gamma band then equation 4 will be like,

Linear Power combination =
$$((\alpha+\delta+\theta+\gamma)-(\beta))/(\alpha+\delta+\theta+\beta+\gamma) \quad (5)$$

So by using this process another five set of data were generated for different cognitive state EEG signals. At last PCA was used for better comparison of collected data sets. At the end total five set of data where input parameters were frequency band power, Frequency band combination ratio and Linear Power combination of five frequency band were generated for each cognitive state for a single subject and for each subject total data was collected twice. So for six subjects total 30× 2 = 60 data sets for five cognitive states were generated for classification and each cognitive state contains 12 set of data.

### D. *Classification precedure*

A three layer Back Propagation Neural Network (NN) was structured for the classification of EEG signals using the features collected by analyzing different cognitive state EEG signals. Input layer consist spectral power, frequency band combination ratio and linear combination of five frequency band of EEG signal. Output layer consist five outputs for five cognitive states RSEC, RSEO, TH, MR and EM. A sigmoid transfer function is applied between each layer. 70% of total data were used as training data and others were used as testing data to find out the classification performance of the system.

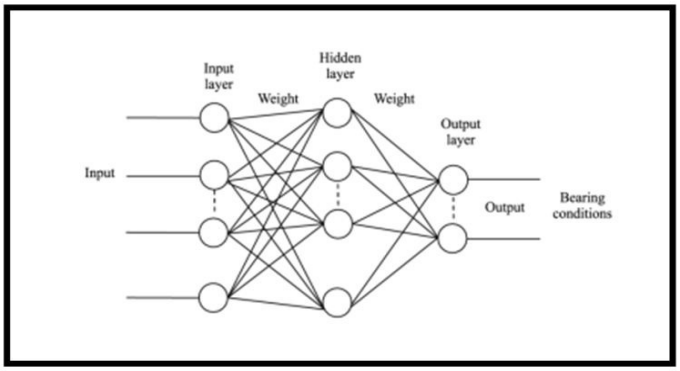



Figure 2: A basic three layer Back Propagation Neural Network structure [14]

## III. RESULTS

### A. *Training the NN*

As it is said earlier 70% of total data used as training data. Each data set contains 1× 52 input values. Iteration number to Mean Square Error (MSE) graph for training the data set is given in fig 3.

For 1000 iteration MSE is less than 4.7402e-04.

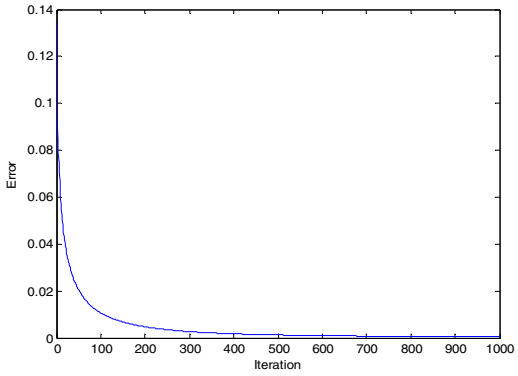



Figure 3: Iteration vs. error curve for training data set.

### B. *Testing the NN*

After training the NN with different cognitive state data sets, it is required to test the NN to find out the classification performance. For this purpose each state data were marked by

a specific sequence. Testing data for different cognitive state with mental state marking are given in TABLE II.

TABLE II. TESTING DATA WITH COGNITIVE STATE MARKING

| No. | Data | RSEC | RSEO | TH | MR | EM |
|---|---|---|---|---|---|---|
| 1 | Data's of RSEC | 1 | 0 | 0 | 0 | 0 |
| 2 | Data's of RSEO | 0 | 1 | 0 | 0 | 0 |
| 3 | Data's of TH | 0 | 0 | 1 | 0 | 0 |
| 4 | Data's of MR | 0 | 0 | 0 | 1 | 0 |
| 5 | Data's of EM | 0 | 0 | 0 | 0 | 1 |

*RSEC: Resting state Eye close; RSEO: Resting state Eye open; TH: Thought; MR: Memory; EM: Emotion*

For example, output given by the TH state by NN is,

0.0084 0.1224 0.8673 0.0048 0.0264

NN classification results for method A are given in TABLE III. This method described in [10], where features collected for different mental task by using spectral power and linear combination of power for EEG frequency band alpha, theta and delta. 1×44 values used as NN input.

TABLE III. NN CLASSIFICATION RESULT USING METHOD A FOR 6 SUBJECTS

| Subject states | RSEC | RSEO | TH | MR | EM |
|---|---|---|---|---|---|
| Subject 1 | 80% | 65% | 70% | 82% | 67% |
| Subject 2 | 83% | 63% | 65% | 85% | 53% |
| Subject 3 | 85% | 60% | 78% | 76% | 61% |
| Subject 4 | 77% | 62% | 55% | 80% | 65% |
| Subject 5 | 80% | 70% | 64% | 73% | 67% |
| Subject 6 | 84% | 63% | 80% | 82% | 70% |

NN classification results of method B for 6 subjects are given in TABLE IV, where single target state was used for testing the NN. Features were collected from EEG signal by using spectral power, frequency band combination ratio and linear combination of EEG frequency band alpha, theta, delta and gamma. Here additional gamma band was used as well as with other frequency band. Though in this method feature number increased but classification performance for different cognitive state also increased.

TABLE IV. NN CLASSIFICATION RESULT OF SINGLE TARGET STATE FOR 6 SUBJECTS USING METHOD B

| Cognitive state | Subject 1 | Subject 2 | Subject 3 | Subject 4 | Subject 5 | Subject 6 |
|---|---|---|---|---|---|---|
| RSEC | 90% | 95% | 96% | 95% | 95% | 94% |
| RSEO | 72% | 65% | 67% | 68% | 82% | 71% |
| TH | 89% | 82% | 85% | 82% | 85% | 85% |
| MR | 96% | 97% | 97% | 97% | 96% | 97% |
| EM | 79% | 81% | 77% | 79% | 75% | 78% |

TABLE V represents NN classification results for subject 1 when two target states were used for testing the NN. This table also represents the variations in classification results with different hidden layer unit number. From the average value it is evident that classification performance remains approximately invariant to the number of NN hidden layer units. For subject 1 classification performance is maximum for RSEC MR State.

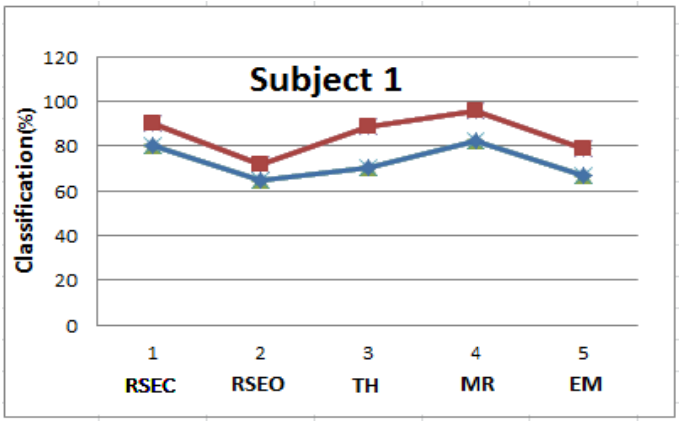
TABLE V: NN CLASSIFICATION RESULT FOR DOUBLE TARGET STATE OF SUBJECT 1

| Hidden layer unit / Cognitive state | | 20 | 40 | 60 | 80 | 100 | Average |
|---|---|---|---|---|---|---|---|
| RSEC | RSEO | 70% | 67% | 80% | 76% | 78% | 74% |
| RSEC | TH | 84% | 90% | 86% | 77% | 82% | 84% |
| RSEC | MR | 95% | 96% | 97% | 96% | 97% | 96% |
| RSEC | EM | 73% | 85% | 68% | 71% | 67% | 73% |
| TH | RSEO | 70% | 60% | 69% | 68% | 70% | 67% |
| TH | MR | 92% | 95% | 98% | 97% | 96% | 97% |
| TH | EM | 75% | 67% | 71% | 73% | 69% | 72% |
| MR | RSEO | 63% | 65% | 62% | 71% | 75% | 67% |
| MR | EM | 62% | 70% | 76% | 74% | 70% | 70% |
| EM | RSEO | 56% | 68% | 72% | 67% | 71% | 67% |

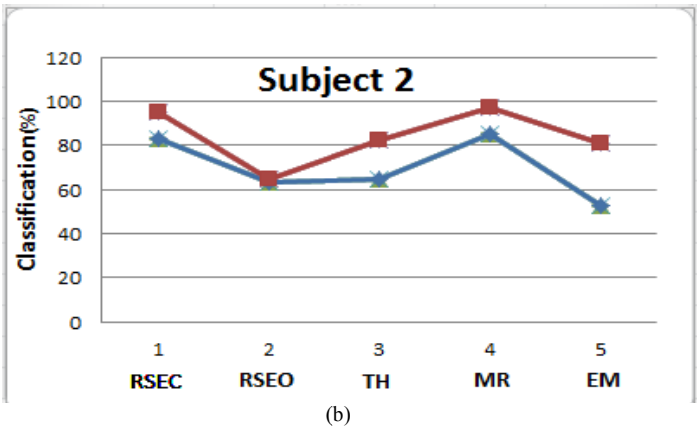
Comparisons of classification performance are given in figure 4. From the figure it is evident that classification performance is better in method B for different mental state than method A. Only in case of subject 6, EM and MR classification results are less in method A. But for other subjects overall performances are good. For subject 1, 2 and 3 MR state is dominant, gives maximum detection performance where in case of subject 4, 5 and 6 RSEC state gives maximum detection performance. All this processing were done in MATLAB R2012a running in a computer with 2-GB RAM, Intel core to duo processor. Where Elapsed time is 1.137012 seconds for method B and method A it is 2.16 seconds because method A requires more iterations than method B to reach the required MSE value. So In terms of processing time method B gives better performance than method A.

## IV. CONCLUSION

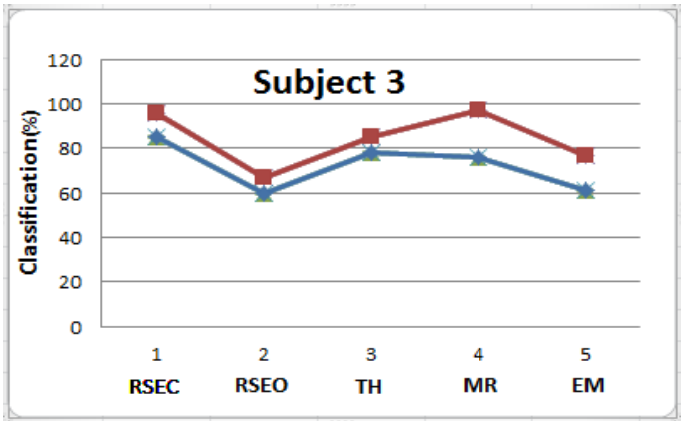
In this paper a cognitive state detection system is structured for BCI applications. Cognitive states are efficiently detected where EEG signals from 6 different subjects were collected for Resting State with eye close and eye open condition, Thought, Memory and Emotion states. Collected EEG signals were filtered to generate five EEG frequency band alpha, beta, theta, delta and gamma. Spectral power, power combination ratio and linear power combination of 4 frequency band alpha, delta, theta and gamma were used to generate the EEG features for cognitive state detection. Classification performance of this method was compared with existing method which uses spectral power and linear combination of 3 frequency band alpha, theta and delta for mental task detection. This study shows that using gamma as additional band though increase the feature numbers but also increase the classification and timing performance.



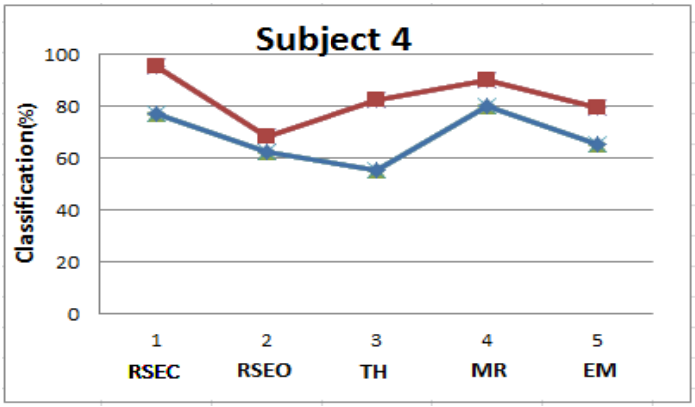

(a)



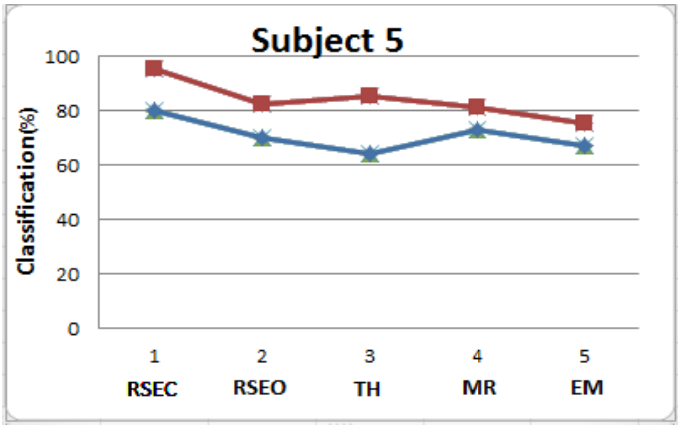

(b)



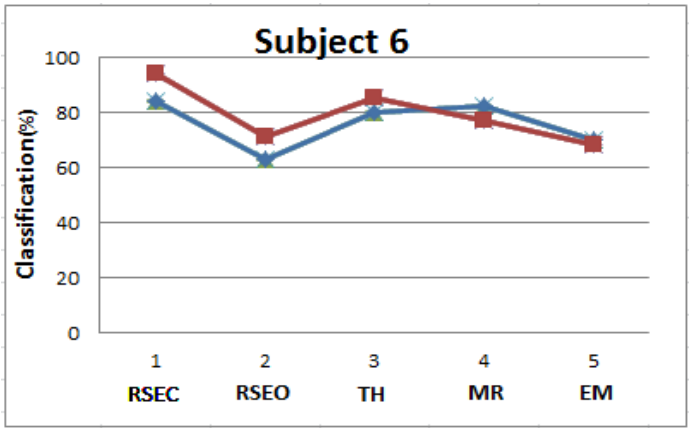

(c)




(d)




(e)




(f)

Figure 4: Comparison of results for (using Average value) method 1 and 2 in terms of classification performance. —— Line indicates method A and ——Line indicates method B. (a) subject 1; (b) subject 2; (c) subject 3; (d) subject 4; (e) subject 5; (f) subject 6.

For future work we want use other features like approximate entropy and want to collect data for more subjects to increase the systems classification performance and decrease the feature numbers required to detect each Cognitive States.

## REFERENCES

[1] C. T. Moritz, S. I. Perlmutter, and E. E. Fetz, "Direct control of paralysed muscles by cortical neurons,"Nature, vol. 456, no. 7222, pp. 639–642, Dec. 2008

[2] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. K¨ ubler, J. Perelmouter, E. Taub, and H. Flor, "A

spelling device for the paralysed," Nature, vol. 398, no. 6725, pp. 297–298, Mar. 1999

[3] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, "Brain–computer communication: Motivation, aim, and impact of exploring a virtual apartment," vol. 15, no. 4, pp. 473–482, Dec. 2007.

[4] M. Takahashi, O. Kubo, M. Kitamura, and H. Yoshikawa, "Neural Network for Human Cognitive State Estimation", Proceedings of the *IEEE Conference on Intelligent Robots and Systems, Advanced Robotic Systems and Real World, IROS*, vol. 3, pp. 2176 – 2183, Sept 1994.

[5] Koel Das, Daniel S. Rizzuto, and Zoran Nenadic, Member, IEEE, "Mental State Estimation for Brain–Computer Interfaces*",IEEE Transactions on Biomedical engineering*, vol. 56, no. 8, august 2009

[6] J. F. Borisoff, S. G. Mason, and G. E. Birch, "Brain interface design for asynchronous control," inToward Brain–Computer Interfacing,G.Dornhege,J.D.R.Mill´ an, T. Hinterberger, D. J. McFarland, and K.-R. M¨ uller, Eds. Cambridge, MA: MIT Press, 2007, ch. 7, pp. 111–121.

[7] E. Niedermeyer, F. Lopes da Silva, "Electroencephalography: basic principles, clinical applications and related fields", Urban and schwarzenberg, 1987.

[8] T. T. K Munia, A. Islam, M. M. Islam, S. S. Mostafa and M. Ahmad "Mental States Estimation with the Variation of Physiological Signals" *,IEEE/OSA/IAPR International Conference on Informatics, Electronics & Vision* (ICIEV) 2012.

[9] S. Solhjoo, M.H. Moradi, "Mental task recognition: a comparison between some of classification methods", BIOSIGNAL 2004.

[10] F. Abdollahi, A. M. Nasrabadi,"Combination of Frequency Bands in EEG for Feature Reduction in Mental Task Classification",Proceedings of the 28th *IEEE EMBS Annual International Conference New York City*, USA, Aug 30-Sept 3, 2006.

[11] Ramaswamy Palaniappan, Member, IEEE, "Utilizing Gamma Band to Improve Mental TaskBased Brain-Computer Interface Design*",IEEE Transactions on Neural systems and rehabilitation engineering*, vol. 14, no. 3, september 2006.

[12] Lei Wang, Guizhi Xu, Jiang Wang, Shuo Yang and Weili Yan, "Feature Extraction of Mental Task in BCI Based on the Method of Approximate Entropy",Proceedings of the *29th Annual International Conference of the IEEE EMBSCité Internationale*, Lyon, FranceAugust 23-26, 2007.

[13] Reference manuals of AcqKnowledge Software Guide and MP System Hardware Guide, BIOPAC Systems Inc., 42 Aero Camino, Goleta, CA –0067.

[14] www.intechopen.com/source/html/38727/media/image2.jpg

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Statistical Parameters in the Dual Tree Complex Wavelet Transform Domain for the Detection of Epilepsy and Seizure

Anindya Bijoy Das[†],[*], Mohammed Imamul Hassan Bhuiyan[†] and S. M. Shafiul Alam[‡]
[†]Department of Electrical and Electronic Engineering(Bangladesh University of Engineering and Technology, Bangladesh)
[‡]Department of Electrical and Computer Engineering (Kansas State University, United States of America)
Email: *anindya149@yahoo.com

***Abstract*—In this paper, a comprehensive statistical analysis of electroencephalogram (EEG) signals is carried out in the dual tree complex wavelet transform domain using a publicly available EEG database. It is shown that variance and kurtosis can be effective in distinguishing EEG signals at sub-band levels. It is further shown that the parameters of a normal inverse Gaussian probability density function can equally discriminate the EEG signals at sub-band levels. Thus, these statistical quantities may be used to characterize EEG signals and help the researchers in developing improved classifiers for the detection of epilepsy and seizure and building a better understanding of the diverse process of EEG signals.**

***Index Terms*—Electroencephalogram(EEG), Seizure, Dual Tree Complex Wavelet Transform(DT-CWT), Normal Inverse Gaussian(NIG).**

## I. Introduction

Epilepsy is a disorder of the brain characterized by a predisposition to generate epileptic seizures; a transient occurrence of signs and/or symptoms, that arises due to abnormal and excessive or synchronous neuronal activity in the brain [1].It has considerable economic implications in terms of healthcare needs, premature mortality, and losses in productivity and sometimes, may result in death of matured persons. Around 1% of the world's population is affected by epilepsy, and nearly 25% of them cannot be treated effectively by available therapies due to resistance to drugs [2].Interestingly, most of the drug-resistant patients respond quite well to focal surgery where the epileptogenic zone is resected. However, the resection process is rather difficult due to the overlay of the epileptogenic zone with other eloquent areas responsible for language, primary motor and vision, and in some other cases, because of the presence of multifocal epilepsy. Closed-loop neuro-stimulators, for example, the cranially implanted responsive neurostimulator (RNS) [3], can be a possible effective solution in such cases. Here, the seizure detection is carried out by analyzing the EEG signals that reveal the seizures as rhythmic discharges. As soon as the seizure is detected, a closed-loop device implanted in the body [3], triggers appropriate stimulation to the epileptogenic zone to suppress the neuronal discharge and thus, abort the seizure. Obviously, this requires an effective and fast detection of seizure onset. In addition, diagnosing epilepsy requires the observation of a voluminous EEG data by an expert neurologist which may last for days to weeks. This is time consuming and may lead to error. So, automatic detection of seizure can be helpful by generating alarms, thus relieving the neurologists of the burden of time consuming observations. Also, it helps to reduce the effect of misinterpretation, since manual detection of seizure is highly subjective [4]. Thus, automatic detection of seizure is a very important problem in health-care area.

Various algorithms have been proposed in the literature for automatic detection of seizures [5]- [11]. The indispensable part of a detection algorithm is the extraction of the appropriate features to discriminate the EEG signals. The detection process is carried out by extracting the features from an EEG signal and classify into the appropriate categories such as seizure (ictal) and non-seizure (non-ictal). Time-frequency-based features have been shown to be highly promising in the detection of seizures [4], [8]. One of the main reasons for this is that the diverse process of brain dynamics and associated neuronal activities are more precisely represented in frequency sub-bands as compared to the original EEG. Certain changes which may include the seizure are found more evident while analyzing the bands separately than the original EEG.

The objective of this paper is to analyze the EEG signals in the dual tree complex wavelet transform (DT-CWT) domain and develop features for discriminating EEG signals into various classes, for example, ictal and non-ictal. This transform offers a better and sharper time-frequency representation of a signal as compared to that of discrete wavelet transform (DWT). DT-CWT has widely been used in the analysis of processing of image and video signals [12]- [16]. However, to the best of our knowledge, limited research has been reported in the literature by using DT-CWT for processing of biological signals such as EEG, EMG or ECG [17]. Analysis of an EEG in the related sub-bands signal of DT-CWT domain could provide with effective features that may result in understanding the various neural activities related to the diagnosis of epilepsy and possible detection of seizure, and help researchers in developing improved algorithms for seizure detection.

## II. Description of the EEG Database

In this paper, features in DT-CWT domain are picked out and their statistical significance in discriminating EEG signals

is investigated using a publicly available database [18], [19]. For the convenience of the readers a brief description of the database is provided in this section. The reasons behind using this database are: (i) availability in public domain, (ii) a considerable volume [9], which is necessary to provide the statistical significance of the discriminating features, and (iii) widespread use in the literature of EEG (See references [4], [6], [8] and the references therein). The EEG database consists 500 single-channel EEG segments of 23.6-seconds duration each. There are five sets of grouped data, namely A, B, C, D and E each containing 100 EEG segments. Sets A and B consist of surface EEG segments collected from five healthy volunteers in awake and relaxed state, with their eyes open and closed, respectively. The standard 10-20 electrode placement scheme is used to collect the data. An archive of EEG signals collected during presurgical diagnosis is used to obtain the Sets C, D and E. These sets correspond to five patients, who gained complete control of seizures after resection. These resected sites are thus diagnosed as epileptogenic zone. EEG data in Sets C and D are obtained from the electrodes placed in the epileptogenic zone and hippocampal formation of the opposite hemisphere and correspond to seizure-free epochs. Data in Set E are collected from these electrodes as well as those implanted in temporal and basal regions of the neocortex, and correspond to seizure attacks. However the intra-cranial electrodes containing pathological activity are not considered in sets C,D and E. All of the signals are recorded in digital format at a sampling rate of 173.61Hz.Thus, the sample length of each segment is $173.61 \times 23.6 \approx 4097$, and the corresponding bandwidth is 86.8 Hz.

## III. Dual Tree Complex Wavelet Transform (DT-CWT)

In this section, a brief description of the dual tree complex wavelet transformation (DT-CWT) is provided. The DT-CWT is a relatively recent enhancement to the DWT with important additional properties. It introduces limited redundancy and provides approximate shift invariance and directionally selective filters (properties lacking in the traditional wavelet transform) while preserving the usual properties of perfect reconstruction and computational efficiency.

The 1-D DT-CWT employs two DWTs (Fig. 1); the first DWT gives the real part of the transformed coefficients, while the second DWT gives the imaginary part. Fig. 2 shows the plots of sample EEG segments for six seconds from the datasets C (top left) and E (top right) in the first row and the plots of the corresponding first level DT-CWT real and imaginary coefficients, respectively in the second and the third row.

## IV. Analysis of EEG in DT-CWT domain

The frequency range of an EEG signal spans over 0 to 60 Hz. The frequencies greater than 60 Hz may be considered as noise [9], [20]. On the other hand, the highest frequency component of an EEG segment of the database is 86.8 Hz. The frequencies beyond 60 Hz are removed by using a 6th

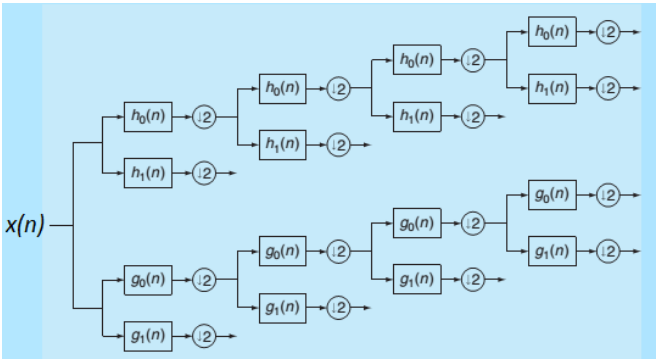



Fig. 1. 1-D Dual Tree Complex Wavelet Transformation [13]

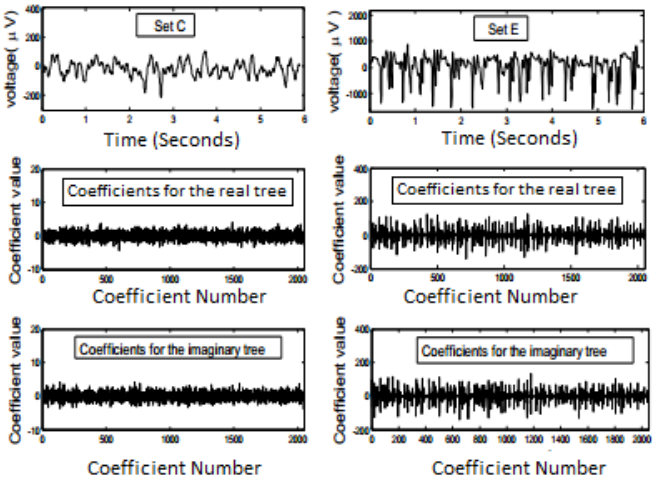



Fig. 2. Sample EEG Signals and the corresponding DTCWT coefficients

order Butterworth filter. Fig. 3 shows the power spectrum density of two EEG segments and their corresponding filtered versions. The first row shows the power spectrum density of the actual EEG segment from sets C and E respectively, and the second row shows the power spectrum density (PSD) of the corresponding filtered signals. Note that the difference between the two spectrums is negligibly small for each of the datasets.

For the purpose of analysis, the band-limited signals are subjected to a 4 level DT-CWT decomposition. After the first level of decomposition, the EEG signal, X (0-60 Hz), will be decomposed into its higher resolution components y1 (30-60 Hz) and lower resolution components, z1 (0-30 Hz). In the second level of decomposition, the z1 component is then decomposed into higher resolution components, y2 (15-30 Hz) and lower resolution components, z2 (0-15 Hz). So, the components obtained after four levels of decomposition, are z4 (0-4 Hz), y4 (4-8 Hz), y3 (8-15 Hz), y2 (15-30 Hz), and y1 (30-60 Hz). Reconstructions of these five components using the inverse DTCWT approximately correspond to the five physiological EEG sub-bands delta, theta, alpha, beta, and gamma respectively [9]. Although, the sub-bands might overlap, it is insignificant considering their physiologically

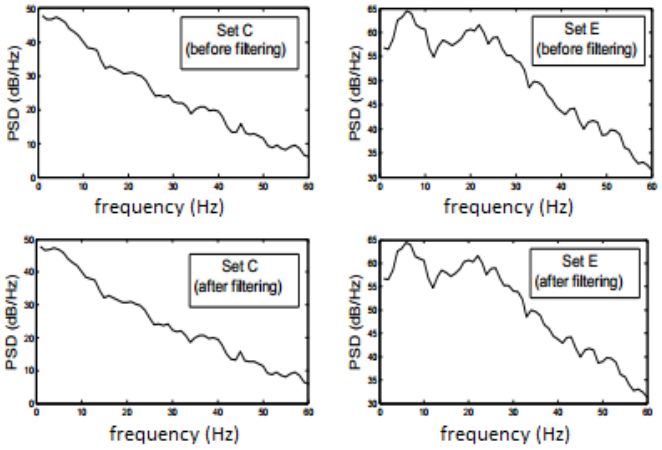



Fig. 3. Power spectrum density of two sample EEG Signals and their filtered versions

approximate nature of the sub-bands. Since, each DT-CWT coefficient has two parts, real and imaginary, the 4-level decomposition yields ten sub-bands in total (five for real and five for imaginary). The DT-CWT coefficients for any sub-band are characterized with higher order statistical moments such as variance, skewness and kurtosis. These moments have been shown to be effective features for describing EEG signals [4], [8]. The dispersion, asymmetry and peakedness of a dataset are well described by these moments. In order to understand the underlying diverse process of an EEG signal, it might be more useful to analyze it in the frequency domain (as in the DT-CWT sub-bands) as compared to the band-limited ones in terms of these higher order moments. In this paper the EEG signals are analyzed in the DTCWT domain using the higher order moments. The expected outcome of this analysis is the characterization of the sub-bands of EEG signals in form of the moments and thus shows their ability to describe the diverse processes underlying the EEG signals and discriminate those signals.

For a given dataset X with length N and mean $\mu$, the corresponding variance($\sigma^2$), skewness($\beta_1$) and kurtosis($\beta_2$) are expressed as:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{1}$$

$$\beta_1 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^3 \tag{2}$$

$$\beta_2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^4 \tag{3}$$

The analysis is carried out with the values of the higher- order moments calculated using the equations (1), (2) and (3), for each sub-band and their respective averages. Significance of the statistical analysis may be established from the following. First, there are 10 sub-bands for a 4 level decomposition, thus giving 10×100=1000 sub-bands for a given data-segment. As we are using five sets, there are total 1000×5 =5000 sub-bands in total. Secondly, the minimum and maximum length of a sub-band is 256 and 2048 respectively. Thus, the statistical analysis of the EEG data in DT-CWT domain is comprehensive enough for the corresponding results being statistically significant.

TABLE I
VALUES OF THE VARIANCE FOR REAL AND IMAGINARY TREES FOR DIFFERENT SUB-BANDS AND BAND-LIMITED SIGNALS

| Signals | Set A | Set B | Set C | Set D | Set E |
|---|---|---|---|---|---|
| Band-limited | 1721.9 | 4057.6 | 2946.1 | 7610.2 | 115580 |
| (y1,1) | 10.901 | 21.848 | 5.749 | 6.865 | 345.62 |
| (y1,2) | 10.902 | 22.062 | 5.709 | 7.088 | 344.30 |
| (y2,1) | 218.06 | 368.239 | 90.345 | 127.87 | 10665 |
| (y2,2) | 222.10 | 360.037 | 88.637 | 128.53 | 10629 |
| (y3,1) | 1490.2 | 6091.6 | 728.15 | 1554.1 | 121740 |
| (y3,2) | 1465.3 | 6303.3 | 767.595 | 1468.6 | 121890 |
| (y4,1) | 2753.6 | 11078 | 3443.1 | 7668.9 | 294030 |
| (y4,2) | 2788.4 | 10649 | 3368.7 | 8986.3 | 297460 |
| (z4,1) | 7119.1 | 7561.1 | 18353 | 49769 | 344420 |
| (z4,2) | 7115.6 | 7539.3 | 18358 | 48609 | 340930 |

TABLE II
VALUES OF THE SKEWNESS FOR REAL AND IMAGINARY TREES FOR DIFFERENT SUB-BANDS AND BAND-LIMITED SIGNALS

| Signals | Set A | Set B | Set C | Set D | Set E |
|---|---|---|---|---|---|
| Band-limited | -0.021 | 0.056 | -0.1517 | 0.078 | -0.06 |
| (y1,1) | -0.083 | -0.0064 | -0.4912 | -0.086 | -0.323 |
| (y1,2) | -0.017 | -0.0294 | -0.2625 | 0.224 | -0.366 |
| (y2,1) | 0.008 | -0.0653 | -0.0651 | -0.034 | 0.0331 |
| (y2,2) | -0.005 | 0.0146 | -0.0265 | -0.016 | 0.022 |
| (y3,1) | 0.005 | -0.0444 | -0.0185 | 0.114 | 0.029 |
| (y3,2) | 0.003 | 0.0084 | -0.0276 | -0.02 | -0.025 |
| (y4,1) | 0.024 | 0.0380 | 0.0574 | 0.012 | -0.02 |
| (y4,2) | -0.002 | -0.0066 | 0.0088 | -0.157 | 0.0008 |
| (z4,1) | -0.017 | 0.0146 | -0.1199 | 0.118 | -0.025 |
| (z4,2) | -0.015 | 0.0143 | -0.1468 | 0.051 | 0.034 |

TABLE III
VALUES OF THE KURTOSIS FOR REAL AND IMAGINARY TREES FOR DIFFERENT SUB-BANDS AND BAND-LIMITED SIGNALS

| Signals | Set A | Set B | Set C | Set D | Set E |
|---|---|---|---|---|---|
| Band-limited | 3.2278 | 3.2175 | 3.6478 | 4.3287 | 3.4372 |
| (y1,1) | 6.6088 | 11.8665 | 29.9735 | 42.023 | 40.635 |
| (y1,2) | 8.6096 | 11.7073 | 33.8681 | 54.524 | 37.607 |
| (y2,1) | 3.1780 | 3.3495 | 6.1513 | 11.231 | 6.784 |
| (y2,2) | 3.1664 | 3.3728 | 4.6434 | 9.8857 | 6.701 |
| (y3,1) | 3.3662 | 3.1536 | 3.9890 | 7.9607 | 4.411 |
| (y3,2) | 3.3029 | 3.1113 | 4.2934 | 7.7155 | 4.39 |
| (y4,1) | 3.3419 | 3.2872 | 3.6767 | 4.8281 | 3.185 |
| (y4,2) | 3.2537 | 3.3109 | 3.5491 | 4.9606 | 3.165 |
| (z4,1) | 3.2601 | 3.2634 | 3.6125 | 3.6594 | 2.569 |
| (z4,2) | 3.3009 | 3.1997 | 3.6846 | 3.6364 | 2.591 |

Tables I, II and III show the average values of the variance, skewness and kurtosis obtained from different sub-bands and for the band-limited signals from different sets. Note that the sub-bands are indicated as mentioned earlier; and the real and imaginary trees are indicated by tree 1 and 2, respectively. For instance, (y3, 2) means the imaginary tree of the y3 sub-band (8-15 Hz) or (z4, 1) means the real tree of the z4 sub-band (0-4 Hz).

The following observations are made from the tables:

- Differences among the values of the skewness are quite small for all the sub-bands for both of the trees for different Sets A, B, C, D and E.
- The values of the kurtosis are different from each other for different sets of EEG data. However, for the sub-bands y1 and y2, the differences are considerable, thus enabling easy discrimination among the various sets of EEG data.
- Also observe that the value of kurtosis is the highest for Set D among the different sets in the high frequency sub-bands; and is the lowest for Set E in the low-frequency sub-bands. The values of the kurtosis are smaller for the lower-frequency sub-bands than for the high-frequency sub-bands.
- Differences among the values of variance are large enough, signifying a good level of discrimination among the various types of EEG data.
- The value of variance is the highest for Set E and the lowest for Set C among the different sets in the high-frequency sub-bands. The values are larger in the low-frequency sub-bands than for the high-frequency sub-bands.
- Both the values of variance and kurtosis for Set E (ictal) are significantly different from those of Sets C and D (non-ictal).

The above observations are further confirmed from the p-values of the ANOVA analysis for the various sub-bands and band-limited signals. The p-values are given in table IV. The p-values for variance and kurtosis are quite small for all the sub-bands; indicating their ability to discriminate the different sets of EEG data. However, in comparison, the p-values obtained for skewness are relatively large, indicating a poor discrimination that concurs with our previous observations.

TABLE IV
VALUES OF THE P-VALUES OBTAINED FROM ANOVA ANALYSIS

| Signals | Variance | Skewness | Kurtosis |
|---|---|---|---|
| Band-limited | 4.8747e-70 | 0.0096 | 2.6432e-09 |
| (y1,1) | 3.2493e-31 | 0.5403 | 7.8511e-07 |
| (y1,2) | 2.0858e-31 | 0.4130 | 1.7699e-08 |
| (y2,1) | 4.9192e-30 | 0.1297 | 1.3875e-12 |
| (y2,2) | 3.5461e-28 | 0.6291 | 1.0839e-18 |
| (y3,1) | 8.3966e-38 | 0.0078 | 1.4361e-17 |
| (y3,2) | 6.2830e-38 | 0.9325 | 5.4842e-18 |
| (y4,1) | 2.4443e-61 | 0.5314 | 1.0696e-10 |
| (y4,2) | 1.3958e-59 | 0.0039 | 3.2135e-12 |
| (z4,1) | 4.7633e-54 | 0.0003 | 3.3696e-19 |
| (z4,2) | 3.3464e-55 | 0.0005 | 7.7344e-20 |

The higher-order moments of the data can be obtained using the underlying probability density function (pdf) that can model the statistics of the data accurately. Thus, it might be interesting to see whether the parameters of a pdf can be used to differentiate EEG data. In this paper, the normal inverse Gaussian (NIG) pdf is considered due to its ability to model a wide range of nonlinear data that include hydrophone data, financial data and economics data among others [21]-[24]. Also from Table III, it is seen that the values of the kurtosis

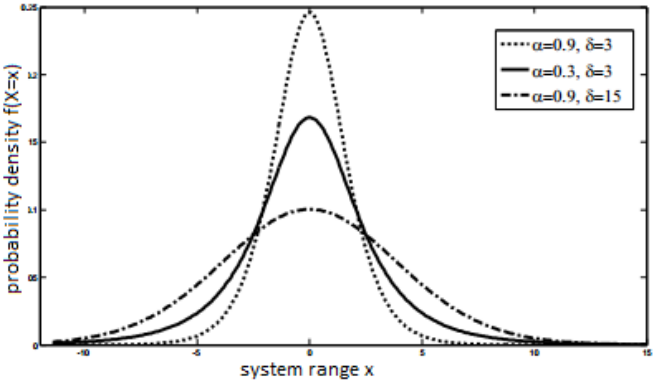



Fig. 4. Changes in shape of the NIG pdf for various values of its parameters

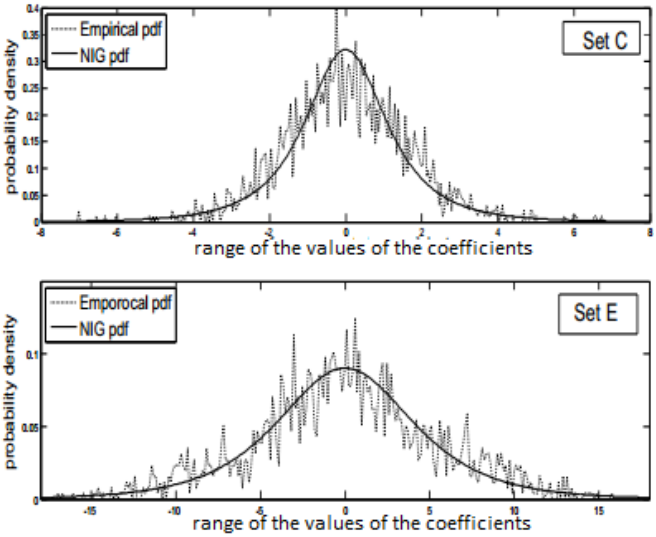



Fig. 5. Plots of the empirical pdf with the NIG pdf for the sub-band (y1,1) of two samples of EEG signals

become close to 3 (the kurtosis for a Gaussian) for lower-frequency sub-bands (e.g. z4). The values of kurtosis of the high-frequency sub-bands are comparatively higher especially for sets D and E, indicating a steeper pdf as compared to that of the low-frequency sub-bands. In this regard, the NIG pdf is interesting since it results in a Gaussian density of finite variance ($\frac{\delta}{\alpha}$)(where $\alpha$ and $\delta$ are the model parameters)as $\alpha \to \infty$ and $\delta \to \infty$, whereas a Cauchy pdf (known to be sharply peaked) is obtained for $\alpha \to 0$. The NIG pdf is expressed as:

$$P_{\alpha,\delta}(x) = \frac{A(\delta,\alpha)K_1(\alpha\sqrt{\delta^2+x^2})}{\sqrt{(\delta^2+x^2)}} \tag{4}$$

where, $K_1$(x)is the first ordered modified Bessel function of the second kind, and A($\delta,\alpha$)=$\frac{\delta\alpha}{\pi}$exp($\delta\alpha$). The steepness of the pdf is controlled by $\alpha$ in that, as it is increased, the pdf becomes steeper. The other parameter $\delta$ is a scale factor that controls its dispersion. Fig. 4 shows the plots of the pdf for various values of the parameters $\alpha$ and $\delta$.

For analysis, the values of the parameters $\alpha$ and $\delta$ are estimated [15] from the different EEG sub-bands. The parameters are obtained from the following equations-

$$\alpha = \sqrt{\frac{3K_x^2}{K_x^4}} \tag{5}$$

$$\delta = \alpha K_x^2 \tag{6}$$

where, $K_x^2$ and $K_x^4$ are the 2nd and 4th order cumulants respectively. Fig. 5 shows the plots of the empirical pdf and the corresponding NIG pdfs for the DT-CWT sub-band (y1, 1) of a sample EEG signal from Set C and Set E. It is seen that the NIG pdf fits the empirical one quite well.

The parameters $\alpha$ and $\delta$ are estimated from the DT-CWT sub-bands of the EEG signals. Table V and VI provide the averages of the estimated values of $\alpha$ and $\delta$ for different sub-bands and band-limited signals of EEG data. Furthermore, we also calculate the corresponding p-values that are shown in Table VII.

TABLE V
VALUES OF $\alpha$ FOR VARIOUS SIGNALS

| Signals | Set A | Set B | Set C | Set D | Set E |
|---|---|---|---|---|---|
| Band-limited | 0.1567 | 0.0806 | 0.0794 | 0.056 | 0.0108 |
| (y1,1) | 0.9407 | 0.5595 | 0.5858 | 0.423 | 0.0565 |
| (y1,2) | 0.6998 | 0.4840 | 0.3988 | 0.379 | 0.0529 |
| (y2,1) | 0.3957 | 0.2104 | 0.4776 | 0.378 | 0.0284 |
| (y2,2) | 0.4200 | 0.2205 | 0.5946 | 0.305 | 0.0220 |
| (y3,1) | 0.1215 | 0.0727 | 0.2125 | 0.097 | 0.0114 |
| (y3,2) | 0.1159 | 0.0655 | 0.1524 | 0.126 | 0.0093 |
| (y4,1) | 0.0698 | 0.0416 | 0.0758 | 0.056 | 0.0107 |
| (y4,2) | 0.0843 | 0.0360 | 0.0883 | 0.056 | 0.0100 |
| (z4,1) | 0.0418 | 0.0495 | 0.0254 | 0.031 | 0.0108 |
| (z4,2) | 0.0523 | 0.0594 | 0.0232 | 0.037 | 0.0129 |

TABLE VI
VALUES OF $\delta$ FOR VARIOUS SIGNALS

| Signals | Set A | Set B | Set C | Set D | Set E |
|---|---|---|---|---|---|
| Band-limited | 233.74 | 243.20 | 152.71 | 141.86 | 930.22 |
| (y1,1) | 9.1912 | 12.5290 | 2.7612 | 1.686 | 8.700 |
| (y1,2) | 8.0661 | 9.4721 | 2.1445 | 1.489 | 8.661 |
| (y2,1) | 69.169 | 64.4696 | 24.721 | 18.47 | 114.8 |
| (y2,2) | 75.203 | 70.3595 | 27.207 | 12.63 | 104.9 |
| (y3,1) | 143.07 | 304.6741 | 74.49 | 52.39 | 577.6 |
| (y3,2) | 162.80 | 268.9135 | 69.90 | 58.87 | 509.6 |
| (y4,1) | 173.70 | 303.4481 | 161.75 | 175.92 | 1572.3 |
| (y4,2) | 223.89 | 265.9313 | 176.72 | 227.41 | 1491.5 |
| (z4,1) | 285.77 | 326.1281 | 340.45 | 455.31 | 1535.4 |
| (z4,2) | 354.80 | 370.8992 | 323.41 | 467.92 | 3050.4 |

Now, from these tables we observe the following-

- Values of $\alpha$ and $\delta$ are different for different sets of EEG data, indicating their ability to discriminate EEG signals.
- The values of $\alpha$ are smaller for the ictal cases (Set E) than for the non-ictal ones (Sets C and D) and the healthy ones (Sets A and B).
- The values of $\delta$ for the ictal cases are larger than those for the non-ictal and healthy ones except the y1 sub-band only.

TABLE VII
P-VALUES FOR ANOVA ANALYSIS FOR THE NIG PDF

| Signals | p-values for $\alpha$ | p-values for $\delta$ |
|---|---|---|
| Band-limited | 2.5901e-05 | 1.9571e-05 |
| (y1,1) | 2.6342e-17 | 1.4899e-04 |
| (y1,2) | 5.8976e-14 | 3.4036e-07 |
| (y2,1) | 4.0702e-11 | 6.3342e-09 |
| (y2,2) | 3.1637e-12 | 1.1249e-11 |
| (y3,1) | 3.2703e-08 | 9.4881e-23 |
| (y3,2) | 4.4115e-11 | 4.6565e-22 |
| (y4,1) | 2.0494e-08 | 5.9246e-14 |
| (y4,2) | 3.2294e-04 | 9.2201e-13 |
| (z4,1) | 0.0064 | 1.8930e-06 |
| (z4,2) | 0.0099 | 2.1884e-07 |

- The p-values for $\alpha$ are smaller than those for $\delta$ in the case of high-frequency sub-bands such as y1 and y2. On the other hand, the p-values for $\delta$ are smaller than those for $\alpha$ in the case of low-frequency sub-bands such as y3, y4 and z4.
- Both the values of $\alpha$ and $\delta$ for Set E (ictal) are significantly different from those of Sets C and D (non-ictal).

## V. CONCLUSION

In this paper, statistical analysis (using higher-order moments) of EEG signals have been carried out in the dual tree complex wavelet transform (DT-CWT) domain. It has been shown that variance and kurtosis may be used to distinguish the sub-bands of the EEG signals. Since the EEG signals can be distinguished using these moments at sub-band levels, indicating a better representation of the underlying brain dynamics or processes, the discrimination could be better as compared to using moments from band-limited EEG signals. Furthermore, we have explored the suitability of using the parameters of a normal inverse Gaussian (NIG) pdf in discriminating EEG signals. It has been shown that the parameters of a NIG pdf can distinguish the various types of EEG data at the sub-band levels. Thus, these parameters may also be used to characterize and discriminate EEG signals. It should be mentioned that the parameters could equally characterize the band-limited signal directly. However, the discrimination cannot be guaranteed in this case because the underlying process can be better characterized at the sub-band levels [9]. Overall, a number of statistical quantities have been shown to be able to distinguish EEG signals at sub-band levels. Thus, these quantities may help researchers to develop effective classifiers for the diagnosis for the detection of epilepsy and seizure from EEG signals.

## REFERENCES


[1] Fisher RS, van Emde Boas W, Blume W, Elger C, Genton P, Lee P, Engel J Jr., "Epileptic Seizures And Epilepsy: Definitions Proposed by the *International League Against Epilepsy (ILAE) and The International Bureau for Epilepsy (IBE)*", vol. 46, Issue 4, page: 470-472, April 2005.

[2] Sheng-Fu Liang, Yi-Chun Chen, Yu-Lin Wang, Pin-Tzu Chen, Chia-Hsiang Yang and Herming Chiueh,"A hierarchical approach for online temporal lobe seizure detection in long-term intracranial EEG recordings," *Journal of Neural Engineering*, doi.10, May, 2013.

[3] Muhammad Tariqus Salam, Mohamad Sawan and Dang Khoa Nguyen,"Low-Power Implantable Device for Onset Detection and Subsequent Treatment of Epileptic Seizures," *Journal of Healthcare Engineering*, vol. 1, no. 5, pp 169-184, 2010.

[4] S M Shafiul Alam and M. I. H. Bhuiyan, "Detection of Seizure and Epilepsy using Higherorder Statistics in the EMD Domain," *Transactions on Information Technology in BioMedicine*, vol. 17, issue. 2, pp. 312-318, March, 2013.

[5] Rajeev Yadav , Rajeev Agarwal and M.N.S. Swamy , "A Novel Dual-Stage Classifier for Automatic Detection of Epileptic Seizures," proc. of *30th Annual International IEEE EMBS Conference*, pp. 911-914, August, 2008.

[6] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Automatic Seizure Detection Based on Time-Frequency Analysis and Artificial Neural Networks," *Computational Intelligence and Neuroscience, Hindawi Publishing Corporation*, vol. 2007, no. 18, August,2007.

[7] S. M. Shafiul Alam, M. I. H. Bhuiyan, Aurangozeb and S. T. Shahriar, "EEG Signal Discrimination using Non-linear Dynamics in the EMD Domain," in Proc. of *3rd International Conference on Signal Acquisition and Processing*, vol. 1, pp. 231-235, February 2011.

[8] Alexandros T. Tzallas, Markos G. Tsipouras and Dimitrios I. Fotiadis, "Epileptic Seizure Detection in EEGs Using Time-Frequency Analysis," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, September 2009.

[9] Hojjat Adeli, Samanwoy Ghosh-Dastidar and Nahid Dadmehr , "A Wavelet-Chaos Methodology for Analysis of EEGs and EEG Sub-Bands to Detect Seizure and Epilepsy," *IEEE Transactions on Biomedical Engineering*, vol:54 , Issue: 2 , pp: 205-211, Feb, 2007.

[10] R. Yadav, M.N.S.Swamy, and R. Agarwal, "Model-Based Seizure Detection for Intracranial EEG Recordings," *IEEE Transactions On Biomedical Engineering*. vol. 59, no. 5, pp. 1419-1428, May, 2012.

[11] S M Shafiul Alam and M. I. H. Bhuiyan, "Detection of Epileptic Seizures using Chaotic and Statistical Features in the EMD Domain," *Annual IEEE Conference, India (INDICON)*, December 2011.

[12] I. W. Selesnick, F. Shi, "Video De-noising using Oriented Complex Wavelet Transforms," proc. of *IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP)*,vol. 2, pp: 949-952, May 2004.

[13] W. Selesnick, R. G. Baraniuk, and N. Kingsbury, "The dual tree complex wavelet transform- A coherent framework for multiscale signal and image processing," *IEEE Signal Processing magazine*, vol: 22, no: 6, pp:123-151, November 2005.

[14] Rudra Pratap Singh Chauhan, Rajiva Dwivedi,Sandeep Negi, "Comparative Evaluation of DWT and DT-CWT for image fusion and de-noising," *International Journal of Applied Information Systems*, vol. 4, no. 2, 2012.

[15] M.I.H. Bhuiyan, M.O. Ahmad and M.N.S. Swamy, "Spatially adaptive thresholding in wavelet domain for despeckling of ultrasound images," *IET Image Processing*. pp. 147-162, vol. 03, no. 03, 2009.

[16] Ufuk Bal, "Dual tree complex wavelet transform based denoising of optical microscopy images," *Biomedical Optics Express*.vol. 3, no. 12, pp. 3231-3239, 2012.

[17] Huijuan Yang, Cuntai Guan, Kai Keng Ang, Chuan Chu Wang, Kok Soon Phua, Juanhong Yu, "Dynamic Initiation And Dual-tree Complex Wavelet feature-based Classification of Motor Imagery of Swallow EEG Signals," proc. of *International Joint Conference on Neural Networks (IJCNN)*, vol. 56, issue. 11, part. 2, June, 2012.

[18] EEG time series download page [Online] Available: http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3

[19] Ralph G. Andrzejak, Klaus Lehnertz, Florian Mormann, Christoph Rieke, Peter David and Christian E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," vol. 64, no. 6, November, 2001.

[20] G. Repovs, "Dealing with Noise in EEG Recording and Data Analysis," *Informatica Medica Slovenica*, vol. 15, no. 1, 2010.

[21] Barndorff-Nielsen, *Exponentially decreasing distributions for the logarithm of particle size-Proceedings of the Royal Society of London, 1977. Series A, Mathematical and Physical Sciences (The Royal Society)*, vol. 353, no. 1674, pp: 401-409.

[22] Alfred Hanssen and Tor Arne Oigard, "The normal inverse gaussian distribution as a flexible model or heavy-tailed processes," *IEEE International Conference* . vol.6, pp: 3985-3988,May 2001.

[23] Karsten Prause, "The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures," , *University of Freiburg.*1999.

[24] Arne Andresen, Steen Koekebakker, and Sjur Westgaard,"Modeling Electricity forward prices using the multivariate normal inverse Gaussian distribution," *Journal of Energy Markets*, vol. 03, no. 3, September 2010.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Channel Selection and Feature Extraction for Cognitive State Estimation with the Variation of Brain Signal

Monira Islam, Tazrin Ahmed, Md. Salah Uddin Yusuf, and Mohiuddin Ahmad*
Dept. of Electrical and Electronic Engineering, Khulna University of Engineering & Technology, Bangladesh
*E-mail: mohiuddin.ahmad@gmail.com

*Abstract*— **Cognitive state estimation with the variation of physiological signal has drawn extensive attention from disciplines such as psychology, cognitive science and engineering. In this paper, we present a cognitive state classification system to assess the subject's mental states based on EEG measurements. The cognitive state estimator is utilized in the context of an augmented cognition system that aims to enhance the cognitive performance of a human user through computer-mediated assistance based on EEG. This paper focuses on the channel selection of the BIOPAC automated EEG analysis and feature extraction based on spectral analysis. Different frequency components (i) real value; (ii) imaginary value; (iii) magnitude; (iv) phase angle and (v) power spectral density of the EEG data samples in different mental conditions are selected for the estimation of cognitive states. In this approach seven types of cognitive states such as – relax, mental task, memory related task, motor action, pleasant, fear, and enjoying music are selected for feature extraction in the three channels - EEG, Alpha, and Alpha RMS of BIOPAC EEG data acquisition system. After feature extraction the channel efficacy are evaluated by support vector machine (SVM) which is based on the classification rate in different cognitive states. From experimental results and classification accuracy, it can be determined that alpha channel can be selected for cognitive state estimation. The overall accuracy for alpha channel shows much improved result for power spectral density and the classification rate is 69.17% whereas for EEG and alpha RMS channel it is found 47.22% and 32.21% respectively.**

***Keywords— EEG, cognitive state, FFT, feature extraction, power spectral density, channel selection.***

## I. INTRODUCTION

The objective measures for inferring the human cognitive activities are hard to obtain as they occur inside the human mind [1]. It is quite difficult, however, to evaluate the effectiveness of the man machine interface, because human being (operator) is included as the system component interacting with the machine. Many projects had been introduced for human mental state estimation using EEG signal. The main difficulty lies in the fact that, it is very hard to uniquely map physiological patterns onto specific emotion types and the physiological data's are sensitive to artifacts and noises. Besides these limitations, it also has some substantial advantages. The information about the users' can be continuously congregated emotional changes through biosensors even in millisecond time resolution and available at low cost [2]. Cognitive states can be defined as the functional activity of brain and a concept of involving human expressions and biological stimulation for specific person for the specific task in different environmental conditions. Brain is one of the most important organs of the body and its structure and functions have become a great source for the research of emotion recognition. The electric potential measured at the scalp through a set of electrodes (channels) are rich in information about the brain activity. The recording of the brain's activity obtained by using electrodes is called electroencephalogram or EEG [3]. The electroencephalogram (EEG) is a complex and aperiodic time series, which is a sum over a very large number of neuronal membrane potentials [4]. The objective of this work is to study changes in physiological variables with respect to changes in human cognitive state during the problem solving process. For this reason two important aspects are held, (i) to propose a minimum redundancy maximum relevance method for extracting some salient spectral analysis based features for cognitive state estimation and (ii) to evaluate and compare the efficiency of the extracted features to compare the efficacy of different channels in different cognitive states using support vector machine. This method has the following two features. One is the elaboration of laboratory experiments (BIOPAC System) for obtaining physiological variables. Two is the utilization of BIOPAC Acknowledge Software for the classification of the physiological data. BIOPAC are adopted for the following reasons. First, inbuilt band pass and 50 Hz noise filters are available. Therefore, the signals obtained are free from noise and no need for further filtering. Second, BIOPAC Acknowledge Software has inbuilt modules for signal analysis [1]. Some of the important issues to be carefully discussed in the aspect of increasing classification accuracy of cognitive states are: (a) proper positioning of electrodes and depth of placement of electrodes on the scalp (b) time duration of video clips and (c) proper design of acquisition protocol [5]. In Fig. 1 the block diagram of our proposed method is shown.

## II. LITERATURE REVIEW

There are many works has been done in physiological signal analysis for cognitive state estimation. Thereby they find the relation between the changes in physiological signals for changing the emotions. In previous work some researchers [6] have discussed the major aspects of using EEG /ERP based

experiments including the recording of signals, removing noise, estimating ERP signals & signal analysis for better understanding of neural correlates of processes involved in creativity. Authors as in [7] have discussed the mental behavior detection from the EEG signal and also the findings of effective data recording from physiological signals, feature extraction through wavelet transform, data reduction, feature classification using various classification methods, real time applications. Authors in [8, 9] addressed an mental behavior evaluation system with temporal and spectral analysis based on different statistical and frequency domain features of EEG signal. Authors in [10] introduced an emotion recognition system based on time and frequency domain features of EEG signal to evaluate classification performance using k-nearest neighbor (kNN) algorithm, multilayer perceptron and support vector machines. Authors in [11] classified three different emotional states (boredom, pain, and surprise) from physiological signals using several machine learning algorithms k-NN (k-nearest neighbor algorithm), LDA (linear discriminant analysis), and SVM (support vector machine). Authors in [12] employed SVM (support vector machine) based classifier to detect epileptic seizure activity from background electroencephalographs (EEGs). In our previous work, emotion recognition system was developed through feature extraction using some salient global features such as amplitude (maximum and minimum, peak to peak value), mean, median, standard deviation, skew, kurtosis, frequency components and wavelet coefficients using FFT and DWT transformation. Authors in [13] classified two emotions: Pleasure and unpleasure using NN and SVM. Here Power at each frequency band and mean of raw signals are used as feature for emotion classification. Authors in [14] used Asymmetrical Alpha Power for classification of Four emotions (angry, sadness, pleasure, and joy) using MLP network. There are several studies proposed for channel selection on EEG signals for feature reduction. While collecting measurements from all EEG channels and then projecting their combined feature vector to a lower dimensional linear or nonlinear manifold would be desirable, the hardware limitations and the prohibitive cost of collecting and processing each additional EEG channel signal beyond the capacity of the hardware imposes us to focus on identifying the salient EEG channels that contain the most useful information for accurate estimation of the cognitive state in the design phase. The conventional feature selection methods were based on Principal Component Analysis (PCA), Independent Component Analysis (ICA), Fast Independent Component Analysis (FICA), and Moments Based Feature Reduction (MBFR) [15]. To the authors' knowledge, there is no work on channel selection for primary emotion recognition using EEG. Some authors used the Asymmetric Variance Ratio (AVR) and Amplitude Asymmetric Ratio (AAR) for selecting salient EEG Channels [16]. Asymmetric Ratios are used for determining the alpha band asymmetry in brain hemisphere studies on human cognition [17]. In this work, the ratio of variances between hemisphere channels is considered as a physiological indicator for assessing the region of brain and the channels which was responsible for detecting the emotions. In addition, ratios of spectral power between two hemispheres were used to accurately estimate the changes of electrical activity. Authors in [1] presented a classification system based on BIOPAC system, which robustly estimated the cognitive state using histograms of several physiological variables. Authors in [18] presented channel selection and feature projection for cognitive load estimation using ambulatory EEG signal. Authors in [16] proposed Asymmetric Ratio (AR) based channel selection for human emotion recognition using EEG where selection of channels reduced the feature size, computational load requirements and robustness of emotions classification. This crisis was addressed using Asymmetric Variance Ratio (AVR) and Amplitude Asymmetric Ratio (AAR) as new channel selection methods. When assessing human emotion using EEG classification, one of the critical problems was to deal with the very large number of features to be classified. Author in [19], addressed this problem using synchronization likelihood as a new channel selection method. Applying this method, the number of EEG channel to be used was significantly reduced in emotion assessment, with only slight (if any) loss of classification performance depending on the used feature.

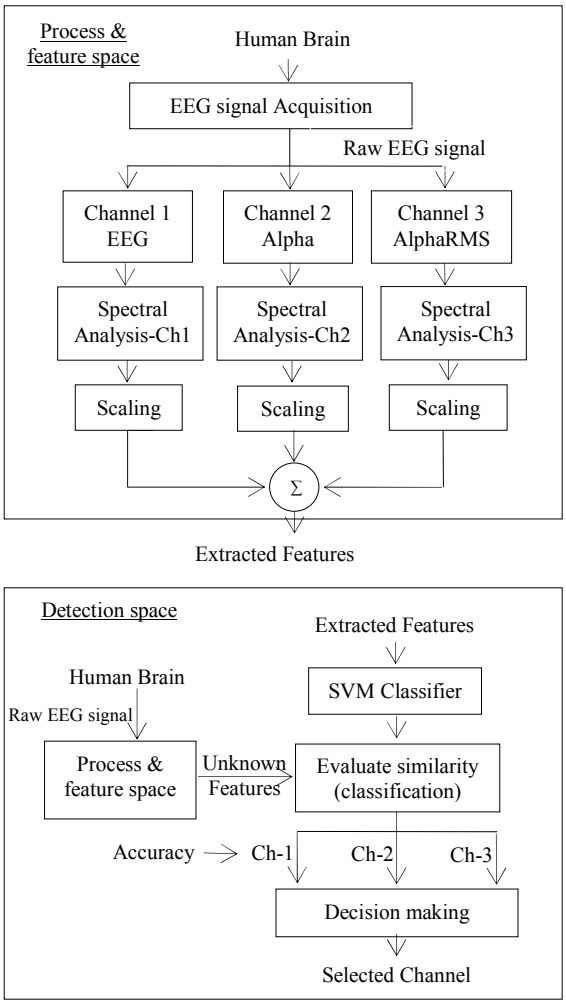



Figure 1. Block Diagram of the proposed method

## III. EEG RECORDING

### A. Categorization of cognitive states for each channel

The definition & the assumption for the categorization of the human cognitive states are quite important in the present study. The categorization should be based on the model of human mental activities: (a) Relax (RLX); (b) Mental task (TH); (c) Memory related (MR) state; (d) Motor Action (MA); (e) Emotional state (Pleasant); (f) Emotional state (Fear) and (g) Enjoying Music (EM). Figure 2 illustrates the seven categories of human state in this proposed work. The subjects were properly trained to perform the specific mental tasks during data collection.

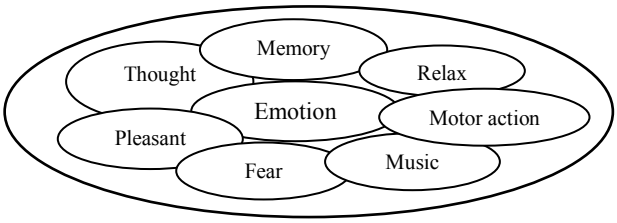



Figure 2. Categorization of different cognitive states.

### B. Subject Selection and preparation

For the purpose of recognizing the emotional states special methodical steps had been performed such as subject selection, data collection using hardware and finally analysis of the data using software. Different affective states were created by giving several instructions and training among people of different age, weight and height. The subject preparation for EEG recording according to the categorized cognitive states is shown in Table I. Consequently, the variation in EEG signal is evaluated for various states. The whole experimental data were collected from several subjects of this university. From them we have collected several data sets in various mental states. The whole experimental data were collected from several subjects of this university.

### C. Data Acquisition

For physiological variables of EEG measurement, the required equipments are BIOPAC electrode lead set (SS2L), BIOPAC disposable vinyl electrodes (EL503), BIOPAC data acquisition unit (MP36 and MP150) with cable and power. For EEG measurement, electrodes should be kept on one side (right or left) of the head. Electrode placement for EEG data extraction is shown in Fig. 3. The data acquisition method was also discussed in [8, 9]. In case of signal acquisition, the electrodes are placed in occipital lobe region which give mainly the variation in alpha amplitude of EEG with eyes closed and eyes opened. Alpha wave amplitudes vary with the subjects attention to mental tasks performed with eyes closed.

Figure 3. Pictorial view of EEG measurement.

## IV. CHANNEL SELECTION TECHNIQUE

In this work, spectral analysis based selection of salient homogeneous pair of electrodes for emotion detection is proposed. In our proposed method frequency based salient feature extraction is devised for proper selection of channel according to the classification rate of cognitive states. For this purpose several frequency components (i) real value; (ii) imaginary value; (iii) peak to peak magnitude; (iv) phase angle and (v) power spectral density were evaluated in case of cognitive state estimation.

TABLE I. SUBJECT PREPARATION FOR EEG RECORDING ACCORDING TO THE CATEGORIZED COGNITIVE STATES

| **Subject Description** | | | | **Subject Mental State** | **Subject Condition** | **Duration** |
|---|---|---|---|---|---|---|
| ***Gender*** | ***Age*** | ***Weight*** | ***Height*** | | | |
| M | 23 | 62 | 5′ 8″ | ***Relax*** | Lying with Eye closed | 30 sec |
| | | | | Take 2 minutes break (time for setup of next recording) | | |
| | | | | ***Thought*** | Asking some mathematical problems and intellectual questions | 1 min |
| | | | | Take 2 minutes break (time for setup of next recording) | | |
| | | | | ***Memory Related task*** | Showing pictures of some familiar things | 3 min |
| | | | | | Involving into other task( as a break) | 3 min |
| | | | | | Recalling the pictures | As required |
| | | | | Take 2 minutes break (time for setup of next recording) | | |
| | | | | ***Motor Action*** | Shaking hand | 20 sec |
| | | | | Take 2 minutes break (time for setup of next recording) | | |
| | | | | ***Pleasant*** | Making a funny environment | 1 min |
| | | | | Take 2 minutes break (time for setup of next recording) | | |
| | | | | ***Fear*** | Watching violent scenes | 1.5 min |
| | | | | Take 2 minutes break (time for setup of next recording) | | |
| | | | | ***Enjoying Music*** | Listening music | 2 min |

The steps in channel selection algorithm for the proposed method are given below:

i) The raw EEG signals from several subjects over seven discrete emotions are collected using three electrodes in BIOPAC system, which are placed on the scalp.

ii) This preprocessed signal is divided into different EEG frequency bands in different channel EEG, Alpha and Alpha RMS. The relative equations of extracted features in these channels are given in Equation (1-6). The Fourier transform of the EEG signal $x(n)$ is expressed as,

$$X(k)_{EEG} = \sum_{n=n_s}^{n_e-1} x(n) \exp\left( -j2\pi \frac{nk}{n_e - n_s} \right) \quad (1)$$

In Eq. (1), $k$ represents the harmonic number of frequency components, $n_s$ represents starting point and $n_e$ represents ending point of the sample of data and $n_e - n_s = N$ (total number of samples). It transforms the time domain data into frequency domain. Equation (1) can be expressed as,

$$X(k)_{EEG} = \sum_{n=n_s}^{n_e-1} x(n) \cos\left( 2\pi \frac{nk}{n_e - n_s} \right) - j \sum_{n=n_s}^{n_e-1} x(n) \sin\left( 2\pi \frac{nk}{n_e - n_s} \right) = R(k) - jI(k) \quad (2)$$

Finally, the magnitude, $|X(k)_{EEG}|$ of frequency transformed data is given by Eq. (3).

$$\left| X(k)_{EEG} \right| = \sqrt{R^2(k) + I^2(k)} \quad (3)$$

Here $R(k)$ and $I(k)$ gives the real and imaginary value of EEG signal in different cognitive states.

Different features for FFT analysis of alpha component of EEG signal is given by,

$$X(k)_{alpha} = \{X(0), \quad X(1), \quad \cdots \quad X(n-1)\} \quad (4)$$

In Eq. (4), $n$ represents the total number of frequency amplitude components and $k$ represents the harmonic number of frequency components. Here, $x(n)$ is the discrete time EEG signal in the time domain, and $X(k)_{EEG}$ is the discrete frequency domain EEG signal at different harmonic frequency. Alpha brain wave is one of five basic brain waves which makes up the EEG. These five brain waves are all oscillating electrical voltages in the brain. The Alpha brain waves oscillate about 10 times per second, and the range is 8-13 cycles per second. In these equations the EEG channel consists of all the five components (α, β, γ, θ, δ) of EEG signal and their specified frequencies whereas alpha channel consists of only the alpha frequency band which is very effective in BIOPAC electrode placement positions because the electrode placed in the occipital lobe that mainly gives the variations of alpha frequency band of EEG signal. In deep sleep and if someone is very highly aroused as in fear or anger there are virtually no alpha brain waves. Alpha brain waves are seen in wakefulness where there is a relaxed and effortless alertness whereas Alpha RMS channel consists of the RMS value of brain signal. Equation (5) represents the expression of phase angle, $\phi$ in frequency domain.

$$\phi = \tan^{-1} \frac{\mathrm{Im}_{X(k)_{EEG}}}{\mathrm{Re}_{X(k)_{EEG}}} \quad (5)$$

The power spectral density (PSD) divides up the total power of the EEG signal. To see this, it is integrated over its entire one-sided frequency domain (0, $F$):

$$\int_0^F PSD(k)dk = \int_0^F 2\left| X(k)_{EEG} \right|^2 / (t_2 - t_1)dk \quad (6)$$

Equation (6) represents the expression of power spectral density in frequency domain where the average power of the signal is in the time range ($t_1$, $t_2$).

The real (value), imaginary (value), magnitude (peak to peak), phase angle and power spectral density are calculated for the homogeneous pair of electrodes on each subject. The channel pairs of higher classification accuracy value among the cognitive states for these features are sorted as most significant channels for cognitive state estimation.

## V. Frequency Domain Based Feature Extraction

Various spatial approaches have been applied to extract features from the EEG signal. The main purpose of feature extraction is to derive the salient features which can map the EEG data into consequent emotional states [10]. For this study the real value, imaginary value, magnitude and phase of brain signal, power spectrum in the frequency domain are analyzed for feature extraction in case of cognitive state estimation and channel selection purpose. Frequency analysis shows the changes in spectral power and phase which can characterize the perturbations in the oscillatory dynamics of ongoing EEG. The frequency-domain features used in this paper are based on the spectrum analysis of each 128-point of EEG samples. Each epoch of the EEG data is processed with Hamming window by zero padding for 128-point fast Fourier transform. The detailed analysis was discussed in [8, 9]. Analysis of changes in spectral power and phase can characterize the perturbations in the oscillatory dynamics of ongoing EEG. The power spectrum of all the sub-epochs within each epoch is averaged to minimize the artifacts of the EEG in all sub-windows. Finally, EEG frequency spectrum features are extracted in different channels such as EEG, Alpha and Alpha RMS. After these operations, five kinds of frequency features are obtained. The dimension of each feature is 30, and the number of data sequence for each feature from each subject for each cognitive state is about 190.

## VI. Result and Discussion

In this paper the features are extracted using spectral analysis to represent the EEG signal which is particularly important for cognitive state estimation and channel selection purpose. In this work, the SVM algorithm LIBSVM 2.8 has

been applied to perform the classification. This kind of algorithms presents some advantages with respect to other classifiers, being one of the most important its high generalization capacity for a reduced number of training trials. A SVM is based on the idea of a hyper plane classifier that works by a separating surface (linear or nonlinear) in the input space of the data set. The SVM algorithm requires a training phase in which the feature vectors generated for a series of trials a provided, together with the corresponding class identifiers, or labels, in order to obtain a model which can separate the different classes considered. The train data is labeled of each action manually by an integer number. For example, RLX=1, TH=2, MR=3, MA=4, Pleasant=5, Fear=6 and EM=7. The MCSVM predicts the class label information for an arbitrary action. Table II represents the classification accuracy in each cognitive states for all the spectral features and it is found that the classification rate for the power spectral density much higher for Alpha channel and for relax (RLX) state it is 93.9% and for enjoying music (EM) state it is 92.1%.The alpha activity largely depends on subject's attention to mental task performed and there is an inverse relationship between alpha activity and mental task performed and the classification rate is also higher in the Alpha channel for individual cognitive states.

In applying SVM, the human emotions are classified into the defined classes. The learning and classification component consists of a training module and a classification module for different cognitive states. The training data of the defined features of the cognitive states are separated into defined classes manually. All the features are used as the input of SVM. Table III represents the training and testing data sequence for SVM classifier where each sample contains 128 data samples.

Table IV represents the overall classification accuracy in different channels for different spectral features. In case of overall accuracy of all cognitive states magnitude shows higher classification rate for Alpha channel and it is found 70% whereas for EEG channel it is 33.43% and Alpha RMS it is 34.89%. For power spectral density it is 69.17% whereas EEG channel it is 47.22% and Alpha RMS it is 32.21%.

Considering all the performances of all features power spectral density shows much improved result than other frequency based features. From the result it is shown that among all the features real value, imaginary value, magnitude (peak-peak), phase and power spectral density the classification rate is much higher for power spectral density for all cognitive states. So it is determined that in case of spectral analysis power spectrum density gives much improved result than other features.

TABLE II. ACCURACY OF COGNITIVE STATE (INDIVIDUAL) CLASSIFICATION USING SVM

| Mental states | Real value | | | Imaginary value | | | Magnitude | | | Phase | | | Power Spectral density | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *EEG* | *Alpha* | *Alpha-RMS* | *EEG* | *Alpha* | *Alpha-RMS* | *EEG* | *Alpha* | *Alpha-RMS* | *EEG* | *Alpha* | *Alpha-RMS* | *EEG* | *Alpha* | *Alpha-RMS* |
| ***RLX*** | 44.5% | 45% | 27.1% | 39.23% | 65.96% | 94.2% | 54.62% | 68% | 88.1% | 61.3% | 68% | 71.2% | 57.14% | 93.9% | 27.2% |
| ***TH*** | 33.35% | 60.46% | 12.1% | 62.3% | 44.12% | 11.6% | 22.1% | 33.2% | 12.2% | 47.05% | 49.6% | 38.57% | 65.0% | 83.33% | 23.5% |
| ***MR*** | 22.1% | 48.15% | 15.2% | 50.1% | 73.17% | 12.43% | 18.0% | 22.1% | 11.2% | 47.82% | 73.17% | 37.8% | 45.0% | 91.83% | 21.8% |
| ***MA*** | 24.5% | 45.6% | 22.0% | 43.06% | 47.1% | 12.7% | 25.0% | 12.1% | 13.2% | 52.94% | 47.06% | 22.82% | 36.2% | 56.0% | 18.1% |
| ***Pleasant*** | 47.1% | 53.19% | 37.5% | 61.36% | 71.05% | 28.16% | 32.1% | 84.5% | 18.0% | 47.92% | 59.18% | 43.02% | 38.42% | 86.5% | 27.64% |
| ***Fear*** | 14.3% | 45.0% | 38.2% | 47.05% | 46.34% | 10.1% | 27.3% | 87.2% | 12.0% | 52.27% | 50.0% | 42.37% | 47.3% | 77.0% | 33.03% |
| ***EM*** | 74.2% | 54.19% | 52.0% | 94.1% | 92.3% | 43.8% | 82.3% | 96.2% | 14.2% | 51.11% | 93.3% | 61.12% | 57.7% | 92.14% | 51.2% |

TABLE III. TRAINING AND TESTING DATA SEQUENCE OF SVM CLASSIFIER

| Data sequence in SVM | Real value | | | Imaginary value | | | Magnitude | | | Phase | | | Power | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *EEG* | *Alpha* | *Alpha-RMS* | *EEG* | *Alpha* | *Alpha-RMS* | *EEG* | *Alpha* | *Alpha-RMS* | *EEG* | *Alpha* | *Alpha-RMS* | *EEG* | *Alpha* | *Alpha-RMS* |
| ***Training data sequence*** | 180 | 172 | 178 | 142 | 142 | 142 | 154 | 153 | 146 | 142 | 144 | 142 | 150 | 153 | 150 |
| ***Testing data sequence*** | 175 | 165 | 162 | 140 | 140 | 141 | 135 | 135 | 135 | 141 | 144 | 141 | 140 | 136 | 137 |

TABLE IV. OVERALL ACCURACY OF ALL COGNITIVE STATES FOR IDENTICAL FEATURES USING SVM

| Frequency Components | *Channel 1* *EEG* | *Channel 2* *Alpha* | *Channel 3* *Alpha RMS* |
|---|---|---|---|
| ***Real value*** | 47.12% | 52.13% | 42.87% |
| ***Imaginary value*** | 37.73% | 43.40% | 34.89% |
| ***Magnitude*** | 33.43% | 70.00% | 54.33% |
| ***Phase*** | 35.66% | 38.18% | 36.08% |
| ***Power*** | 47.22% | 69.17% | 32.21% |

From experimental results and classification accuracy, it can be concluded that alpha channel can be selected for cognitive state estimation since classification accuracy is higher in individual cognitive states for different features. Alpha component occurred in the occipital region of brain and the electrode positions on the scalp was also in the occipital

lobe which makes alpha band more frequency is effective for channel selection. It is very hard to characterize the feature extraction and classification rate computational complexity of SVM classifier correctly. There are two complexities involved: at training time and at testing time. For kernel SVMs, at training time the support vectors must be selected and at testing time the complexity is linear on the number of the support vectors and linear on the number of features. In case of this study it is determined that the classification rate for all mental states are not satisfactory which varies with their training and testing data sets and different frequency band with their types of features.

## VII. CONCLUSION

This work described a framework based on mutual information maximization to solve the EEG feature/channel selection and dimensionality reduction problems in order to perform cognitive state classification. In EEG classification, one of the major problems is the huge number of features to be classified. In the present paper, channel selection method was developed to alleviate "the curse of dimensionality" in EEG classification. To evaluate the performance of channel selection, different kinds of frequency based features (real value, imaginary value, magnitude, phase angle, power spectral density) and SVM classifier were used to demonstrate the efficacy of cognitive states. The result of this study could help for cognitive states estimation and proper channel selection by using physiological brain signals and their classification rate with SVM classifier. For individual cognitive state it is found that the classification accuracy is higher for alpha channel than EEG channel or alpha RMS channel with small exception. For power spectral density (PSD) it was found 93.9% for RLX state, 91.83% for MR state and 92.14% for EM state. From overall accuracy all cognitive states magnitude showed 70.0% and PSD shows 69.17% accuracy. Considering the individual and overall the classification rate at a time it is determined that power spectral density is the most efficient feature in case of frequency based analysis than most other features and the alpha channel is the most significant channel in case of cognitive state estimation. So, proper selection of channel contributes to a significant decrease in computation time for emotion assessment. It can be concluded that due to significance of alpha band of frequency rhythm on emotion studies, only this band of frequency is effective for channel selection.

## ACKNOWLEDGMENT

The work is partly supported by UGC research grant, 6(77)/ugc/research/engg(9)/2009-10/3995 dated: 08/06/2011, Dhaka, Bangladesh.

## REFERENCES

[1] R. Singla and H. Singh, "Analysis of physiological measures for human cognitive state estimation," in *Proc. of the Int. Conf. on Bioinformatics & Computational Biology5, BIOCOMP'08*, Jul.2008, pp. 5126–5131.

[2] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: analysis of affective physiological state," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. *23*(10), pp. 1175-1191, 2001.

[3] Wikipedia, the free encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Electroencephalography.

[4] I. Guler and E. D. Ubeyli, "Multiclass support vector machines for EEG-signals classification," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 2, pp. 117–126, March 2007.

[5] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, D. Hazry, and I. Zunaidi, "Time-frequency analysis of EEG signals for human emotion detection," in *4th Int. Conf. on Biomedical Engineering*, Kuala Lumpur, Malaysia, Jun. 2008, pp. 262-265.

[6] N. Srinivasan, "Cognitive neuroscience of creativity: EEG based approaches," *Centre for Behavioural and Cognitive Sciences*, ELSIEVER, vol. 42, pp. 109–116, 2007.

[7] A. S. AlMejrad, "Human emotions detection using brain wave signals: a challenging." *European Journal of Scientific Research*, vol. 44, No. 4, pp. 640-659, 2010.

[8] M. Islam, T. Ahmed, S. S. Mostafa, M. S. U. Yusuf, and M. Ahmad, "Human emotion recognition using frequency & statistical measures of EEG signal," in International Conference on Informatics, Electronics & Vision, May 2013 © IEEE Xplore. doi: 10.1109/ICIEV.2013.6572658.

[9] T. Ahmed, M. Islam, M. S. U. Yusuf, and M. Ahmad, "Wavelet based analysis of EEG signal for evaluating mental behavior," in International Conference on Informatics, Electronics & Vision, May 2013 © IEEE Xplore. doi: 10.1109/ICIEV.2013.6572706.

[10] XW Wang, D Nie, and BL Lu, "EEG-based emotion recognition using frequency domain features and support vector machines," *Neural Information Processing,* Lecture Notes in Computer Science. Springer vol. 7062, pp. 734-743, 2011.

[11] EH. Jang, BJ. Park, SH. Kim, MA. Chung, MS. Park, and JH. Sohn, "Classification of human emotions from physiological signals using machine learning algorithms," *Int. Conf. on Advances in Computer-Human Interactions*, 2013, pp. 395-400.

[12] R. Panda, P. S. Khobragade, P. D. Jambhule, S. N. Jengthe, P. R. Pal, and T. K. Gandhi, "Classification of EEG Signal Using Wavelet Transform and Support Vector Machine for Epileptic Seizure Diction", in *Proc. Int. Conf. on System in Medicine and Biology*, Dec.2010, pp. 405–407.

[13] K. Takahashi and A. Tsukaguchi, "Remarks on emotion recognition from multi-modal bio-potential signals," *IEEE Transactions on Industrial Technology*, vol. 3, pp. 1654-1659, 2003.

[14] Y-P. Lin, C-H. Wang, T-L. Wu, S-K. Jeng, and J-H. Ch, "Multilayer perceptron for EEG signal classification during listening to emotional music," in *Proc. TENCON 2007 - 2007 IEEE Region 10 Conference*, Oct.-Nov. 2007 © IEEE Xplore doi: 10.1109/TENCON.2007.4428831.

[15] Carreño, I. Rodriguez, and M. Vuskovic, "Wavelet transform moments for feature extraction from temporal signals," in *Proc. Informatics in Control, Automation and Robotics II, Springer Netherlands*, 2007, pp. 235-242.

[16] M. Rizon, M. Murugappan, R. Nagarajan And S. Yaacob, "Asymmetric ratio and FCM based salient channel selection for human emotion detection using EEG." *WSEAS Transactions on Signal Processing,* vol. 4, no. 10, pp. 596-603, Oct. 2008.

[17] H. Ehrlichman, M. S. Wiener, "EEG asymmetry during covert mental activity", *Journal of Psychophysiology*, vol. 17, Issue 3, pages 228–235, May 1980.

[18] T. Lan, D. Erdogmus, A. Adami, S. Mathan, and Tian Lan, "Channel Selection and Feature Projection for Cognitive Load Estimation Using Ambulatory EEG," *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 74895, 12 pages, 2007. doi:10.1155/2007/74895.

[19] A-Asl. Karim, G. Chanel, and T. Pun. "A channel selection method for EEG classification in emotion assessment based on synchronization likelihood," in *Proc. 15th European Signal Processing Conference (Eusipco 2007),* Sep. 2007, pp. 1241-1245.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Feature Selection of EEG data with Neuro-Statistical Method

Md. Zakir Hossain[1,*], Md. Monirul Kabir[2] and Md. Shahjahan [1]
[1]Dept. of Electrical and Electronics Engineering, Khulna University of Engineering and Technology, Bangladesh
[2]Dept. of Electrical and Electronics Engineering, Dhaka University of Engineering and Technology, Bangladesh
[*] E-mail: zakireee62@yahoo.com

*Abstract*—**Feature selection (FS) of high dimensional electroencephalographic (EEG) data helps to identify and diagnose the brain conditions easily. Features can be selected with different ways where canonical correlation analysis (CCA) is one of them which are a statistical method. We employed neural network (NN) with CCA for salient features extraction of EEG data, called Neural Canonical Correlation Analysis (NCCA), which exhibits better result than individual CCA or NN. A NN classifier is used to test the classification of the selected features. The NN classifier shows remarkable result in terms of recognition rate.**

***Keywords—Electroencephalogram (EEG), Neural Canonical Correlation Analysis (NCCA), Clustering, Feature selection (FS), Neural network (NN).***

## I. INTRODUCTION

The electroencephalogram (EEG) is one kind of electrical signal which is generated from human brain and related to body functions. These signals can be measured with electrodes placed on the scalp, noninvasively. The EEG data is collected from brains using electrodes (channels) for a sufficient time. These are composed for several times for a subject called trials. There are different kinds of potential whose are generated from brain; Steady State Visual Evoked Potential (SSVEP) is one of them. It is evoked over occipital scalp areas, when subject focuses on the repetitive flicker of a visual stimulus [1]. The SSVEPs can be processed for different visual stimulation. There have no significant variability among trials for a subject at same stimulation [2]. When these data are added together for analysis, the data size is so high and every data is not important equally. The processing of such massive data is a great challenge for computational scientists. And for that reason, we want to find salient features of such high dimensional data with reduced size which will carry important information.

In this paper for analyzing EEG data, we execute canonical correlation analysis (CCA) using neural network (NN), since NN is well known for their powerful capacity [3]. Using of neural canonical correlation (NCCA) have several advantages such as i) high capacity machine is not required for NCCA, ii) exhibits superior correlation than standard statistical methods [3], iii) EEG data is entered sequentially in the network instead of complete data at a time.

It is difficult to analyze high dimensional datasets for any purpose. It contains very large feature sets, which cause learning to be more difficult and also degrade the generalization performance of the learned models. To simplify and improve the quality of dataset, it is needed to select the salient features of high-dimensional datasets. Generally for this purpose Feature Selection (FS) is used in machine learning. Ordinarily, spurious features are deleted from the original dataset using FS without sacrificing generalization performance. Normally the existence of the FS is essential in real-world problems for the following factors: (i) abundance of noise, (b) spurious information, and irrelevant and redundant features in the original feature set [4]. Accordingly, FS is an area of active research spreading into pattern recognition; data mining, text categorization, image mining and many others field [5].

Generally three types of FS approaches are used as wrapper, filter and hybrid [6]. Where predetermined learning model is assumed and features are selected that justify the learning performance of the particular learning model is called wrapper approach. In the filter approach, statistical analysis of the feature set is required, without utilizing any leaning model. The Complementary strengths of the wrapper and filter approaches are utilized in the hybrid approach. The hybrid model takes advantage of the two models by using their different evaluation criteria in different search stages.

There are different search processes for generating subsets. The sequential forward search (SFS) starts the search process with an empty set and successfully adds features where sequential backward search (SBS) option start with a full set and features are successfully removed. Another one is bidirectional selection [7] which is to start on both ends and add and remove features simultaneously. A fourth approach is to have a search process start with a randomly selected subset using bidirectional or sequential strategy. Due to involving computational complexity and partial search in the solution space, solution of optimal and near optimal is quite difficult for those search algorithms. For that reasons, recently the global search algorithms (or meta-heuristics) is used to find the solution on the basis of the activity of multi agents, which ultimately enhance to find very high-quality solution within a reasonable time [8]. They utilize local search suitability to find solution in the full search space by their global search capability.

In many pattern recognition applications, a large number of features are usually measured. However all of the features are not equally important for a specific task. The selection of feature subspaces can be found using a generalized CCA

framework and a minimum mean-square-error criterion [9]. In a variety of pattern recognition applications, Neural networks (NN) have proved themselves to be a powerful tool. Due to the ability to solve a task with a smaller number of features, NN is suitable for feature selection [10]. Constructive approach for FS based on the concept of the wrapper approach and sequential search strategy combines the FS with the architecture determination of the NN [11].

In this paper, we use a model for feature selection using CCA with NN for EEG data analysis. This shows greater flexibility to find the salient feature with reduced size. This algorithm finds the salient features on the basis of correlation. This process is a wrapper approach which finds the salient features with a backward search strategy. The samples are deleted from expected features according to correlation minimization. The reduced EEG dataset are obtained from that features those shows maximum correlation.

The rest of the papers organized as follows. Collection of EEG data and their characteristics are explored in section II. In section III, whole methodology is described in different subsections. Clustering using wavelet, selection of subsets, NCCA, extraction of features and classification accuracy measurement are described in A, B, C, D and E subsections respectively. In section IV, results and discussions are analyzed. Finally conclude the process in section V.

## II. Collection of EEG Data and Their Characteristics

EEG data are collected from SSVEP database (EEG) [12]. Firstly we describe about the database for the clarity. Acquisition of brain signal was performed at a sampling rate of 2048Hz by using 128 active electrodes or channels (Biosemi Inc., Amsterdam) [13]. Four healthy subjects were participated in this study and they were fully informed of all procedures and having no neurological disorders. For data collection, Subjects were seated 0.9m from a 21inch CRT computer display operated at a high vertical refresh rate. Before each experiment they were briefly tested for photo sensitive epilepsy, also they did not have any prior training except for a short practice run during the briefing.

There were used small reversing black and white checkerboards with 6×6 checks for SSVEP stimulation. The dimensions of the checkerboards were $1.8^o \times 1.8^o$ arc, so that the diameter ($2.5^o$ arc) would just cover the size of the fovea. Three frequencies were sequentially (8, 14 and 28Hz) used for stimulus a single small checkerboard and they covered each of the three SSVEP response regions (low, medium, and high frequency) [14]. To remove the interference, the EEG data was preprocessed using a demodulation procedure [15]. There were 5 trials of each subject for each frequency. Each subject had total 15 trials at 3 different frequencies of 8, 14 and 28Hz accordingly. Therefore, a total of 60 trials of four subjects were found in the database. There were 128 channels (rows) with more than 6000 columns (sampling points) for a trial. 128 EEG signals are generated at 128 channels for every trial. There are 6330 sampling points for every trial of subject 4 at 8Hz stimulation, but for convenient two EEG signals of last 100 samples are shown in Fig. 1. It was shown that there was no big difference within trial and trial [2]. For that reason, we wanted to search feature for depending on subject. There were more than 31500 columns (sampling points) with 128 rows (channels) for a subject.

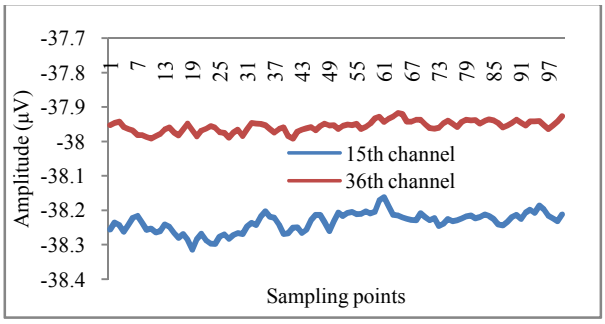



Figure 1. Samples of EEG signals from subject 4 at 8Hz stimulation.

## III. Methodology

We want to search salient features of EEG data with reduced size. SSVEP is one of most important EEG signal which detects the human brain condition at various modes such as sleep, reading, watching TV etc. But all of these data is not so important, also handing of such high dimensional is not an easy task. The correlated data is most important to find the brain conditions. For that motive, we search salient features of EEG data on the basis of NCCA. The correlation coefficient is computed from subset of EEG signals. The entire scheme of our work is exposed in Fig. 2. Every part of the method is briefly described in the following sub-sections.

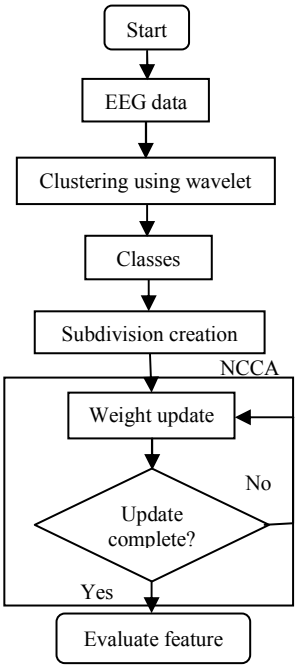



Figure 2. General overview of entire methodology.

### A. Clustering using wavelet

Clustering [16] is the process of organizing data or objects into groups whose members are similar in some way. It is the most important unsupervised learning problem that deals with finding a structure in a collection of unlabeled data. Though there have so many clustering algorithms, we use wavelet for

clustering EEG signals. When a small oscillatory wave concentrates its energy in time is called a wavelet. It is a suitable tool for transient, non-stationary or time-varying signals analysis, which is ability to allow simultaneous time and frequency analysis. Clustering signals using wavelet is proposed in [17], where a classical clustering strategy is applied to a suitably chosen set of wavelet coefficients that offers a useful tool to carry out both significant noise reduction and efficient compression. We use matlab environment for clustering of EEG data using wavelet. Multi-signal 1-D wavelet decomposition is used for EEG signal analysis. We exploit 7 levels near symmetric wavelet according to rows (128 channels). Universal threshold is utilized for signals de-noising, where level dependent estimation of level noise is used for rescaling. Signals are compressed using energy ratio with threshold '99'. Then three clusters are made using cell array that contains the list of EEG data to classify, also signals with coefficients of approximation at level '7' is used.

### *B. Data preparation*

There were used four subjects for SSVEP acquisition. EEG data were collected for every subject with three (8Hz, 14Hz and 28Hz) stimulus frequencies. For that reason, there are total three datasets for one subject and total of 12 datasets for four subjects. For every subject, there have more that 31500 sampling points with 128 channels. Class information was fixed for every rows found by wavelet. The class information of subject *i* are concatenated with EEG data of subject *i, where i=1, 2, 3 and 4 respectively*. Then for one set of EEG data, three subsections are created manually. For first subsection 1 to 10,500 columns, 10,501 to 21,000 columns for second and 21,001 columns to rest are used for third subsection.

### *C. NCCA*

There has a process for finding correlation between two datasets that is CCA [18]. It is a multivariable statistical method that is optimal with respect to correlations, but avoids nonlinear relationship between datasets. Implementation of NN with standard CCA can overcome this problem [5] and also search correlations among two or more datasets. For the sake of simplicity a brief description of neural CCA is presented here. Consider three different subsets of an EEG signal $x_1$, $x_2$ and $x_3$. Then we attempt to find the maximum correlation between the linear combinations of the signals as described in Fig. 3. Let

$$\mathbf{y_1} = \mathbf{w_1 x_1} = \sum_j w_{1j} x_{1j} \quad \ldots\ldots\ldots\ldots (1)$$
$$\mathbf{y_2} = \mathbf{w_2 x_2} = \sum_j w_{2j} x_{2j} \quad \ldots\ldots\ldots\ldots (2)$$
$$\mathbf{y_3} = \mathbf{w_3 x_3} = \sum_j w_{3j} x_{3j} \quad \ldots\ldots\ldots\ldots (3)$$

Where j is the number of column in every row, there were total 128 rows for 128 electrodes. Then we wish to find those values of $w_1$, $w_2$ and $w_3$ that maximize the correlation between $y_1$ and $y_2$, $y_2$ and $y_3$ & $y_3$ and $y_1$ respectively.

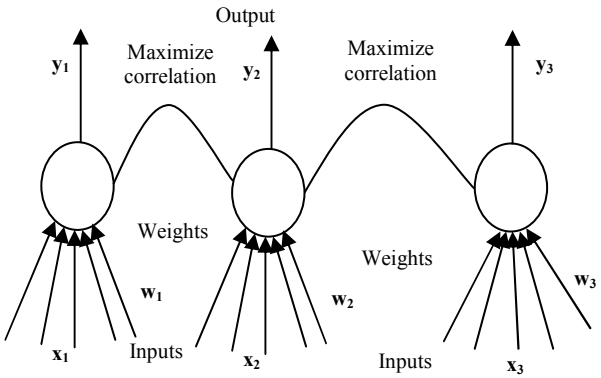



Figure 3. The NCCA network.

For three subsets, input data comprises three vectors $x_1$, $x_2$ and $x_3$. A complete column for a row of a particular subject is entered in the CCA network at a time as input ($x_1$, $x_2$ and $x_3$). In this way, every row is presented in the network sequentially. Activation is fed forward from each input to the corresponding output through the respective weights, $w_1$, $w_2$ and $w_3$. We use the following joint learning rules for linear correlation.

$$\Delta w_{1j} = \eta x_{1j}(y_2 - \lambda_1 y_1) \quad \ldots\ldots\ldots (4)$$
$$\Delta\lambda_1 = \eta_0(1 - y_1^2) \quad \ldots\ldots\ldots\ldots (5)$$
$$\Delta w_{2j} = \eta x_{2j}(y_3 - \lambda_2 y_2) \quad \ldots\ldots\ldots (6)$$
$$\Delta\lambda_2 = \eta_0(1 - y_2^2) \quad \ldots\ldots\ldots\ldots (7)$$
$$\Delta w_{3j} = \eta x_{3j}(y_1 - \lambda_3 y_3) \quad \ldots\ldots\ldots (8)$$
$$\Delta\lambda_3 = \eta_0(1 - y_3^2) \quad \ldots\ldots\ldots\ldots (9)$$

Where $\lambda_1$, $\lambda_2$ and $\lambda_3$ are Lagrange multipliers, $w_{1j}$ is the jth element of weight vector, $w_1$, etc. We choose $\eta_0$ =0.5 & $\eta$ = 0.001 and start at $\lambda_1$ = 0.015, $\lambda_2$ = 0.20 & $\lambda_3$ = 0.025 for representative result.

### *D. Feature extraction*

For SSVEP recognition, CCA works quite well and correlation is very important in order to assess the relationship between two time series. The joint weight update rules of Eq.1 to Eq.9 are applied to find the canonical correlations with different subjects. The correlations and weight maximization are made among most important features of a signal. Columns (samples) are deleted those weight are minimum. We want to keep only 15, 30 and 60 columns from more than 31,500 columns. For that reason, 5, 10 and 20 columns exist accordingly for every subset of a subject. It is done using backward search process i.e: whose correlation is lowest it is deleted firstly. In this way every samples are deleted without our expected features. Then outputs of every subset are added sequentially for select final feature of a subject. In Fig. 4 the whole process is described and the NCCA algorithm is outlined in A1 below.

A1: NCCA algorithm for feature selection

***Input***: $x_1$, $x_2$ *and* $x_3$ are three *subsections of EEG data of one subject corresponding to S sampling points.*
***Output***: *Correlation* $C_n$
***for*** *n=1 to S do*
*Random initialization for* $w_1$, $w_2$ *and* $w_3$
***repeat***
*Find* $w_1$, $w_2$ *and* $w_3$ *which maximize correlation between* $y_1$ *and* $y_2$, $y_2$ *and* $y_3$ *&* $y_3$ *and* $y_1$ *by the CCA*
***until*** *the maximum number of iteration is reached*
*Compute the optimized signals* $y_1$, $y_2$ *and* $y_3$
***end***
*Compute update weight* $w_1$, $w_2$ *and* $w_3$ *for correlation* $C_n$
*Select the features for weight maximization*

### *E. Classification accuracy measurement*

Neural network are used for accuracy measurement of classified selected features of EEG data. The Back-propagation (BP) algorithm is most effective learning method that use gradient descent to tune network parameters to best fit a training set of input-output pairs. The BP with feed forward NN model is shown in Fig. 4. NN generally refers to the interconnected groups of nodes in the different layers of each system, akin to the vast network of neurons in a brain. Three layers are used in this system. EEG signals are fed into the input neurons of input layer. There are 128 attributes for every (15, 30 and 60) EEG feature set. So these are fed into 128 input neurons of one input layer. Firstly 128 attributes of first feature is fed and accordingly all feature are fed to input layer one by one. These input neurons send data via synapses to the hidden layer of neurons. There are 7, 15 and 30 hidden neurons in one hidden layer for 15, 30 and 60 samples respectively and they send data via synapses to the output layer of output neurons. There are three output neurons in one output layer for three classes of EEG data. The synapses store weights to manipulate the data in the calculations. The sigmoid activation function is used here to convert a neuron's weighted input to its output activation. Each circular node in Fig. 4 represents an artificial neuron and an arrow represents a connection from the output of one neuron to the input of another.

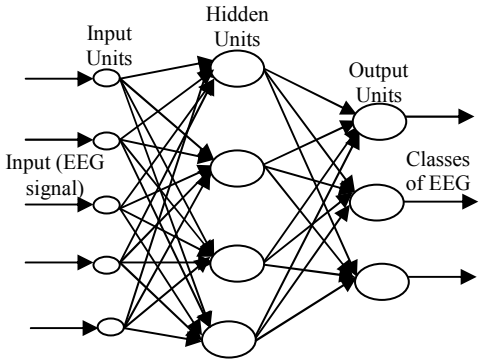



Figure 4. NN model for accuracy measurement of selected features.

## IV. RESULTS AND DISCUSSION

Here, wavelet is used for class determination of EEG data. For a subject, there have total of 128 attributes (rows) and more than 31500 sampling points (columns). We find three class cluster using wavelet. Seven levels, near symmetric wavelet are used for clustering and row is the direction of decomposition. There were four subjects and EEG data is collected for every subject at three different stimulus frequencies. So, there are total of twelve EEG dataset and wavelet is used for clustering every dataset differently. For understanding conveniently, clustering result of second subject whose is stimulated at 8Hz are given below in Fig. 5. For further analysis we use binary 001, 010, and 100 for 1, 2, and 3 respectively. Here Sub denotes subject.

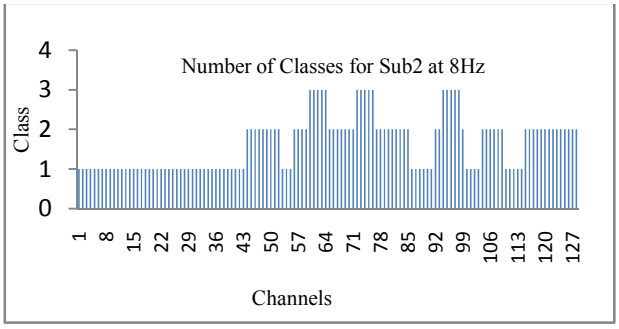



Figure 5. Classification of EEG data using wavelet.

Then we concatenate the class after EEG data of a subject. Three subsections are created conveniently from concatenated data for NCCA network. Data is entered into the network sequentially. Weight is updated for correlation maximization according to NCCA weight update rules shown in equations 1 to 9. Then features are selected according to SBS process. The columns are deleted one by one those shows minimum correlation. We accept three subsets for a subject at one stimulus frequency. We choose 15, 30 and 60 attributes (features) for three subsets of each subject. So there are nine subsets for a subject and total of 36 subsets. These subsets are analyzed using BP of NN. In this way, the accuracy is measured for the selected features. In BP network 7, 15 and 30 hidden nodes are used for 15, 30 and 60 features accordingly. TABLE I shows the accuracy for every subject at different stimulus frequency for various sizes of subsets. We choose 75% (96 rows) for training set and 25% (32 rows) for testing set. We take average accuracy of five runs for every feature set. It is shown from TABLE I that for subject 4 at 8Hz, classifier cannot classify nine attributes. Also classifier cannot detect one attributes for subject 1 and 2 at 14Hz. It is for noise contamination of EEG signals at that time when data is collected from brain. But for every other sector classifier shows 100 percent accuracy. For that reason, we can say that NCCA is a great tool for feature selection. This method is suitable for converting very high dimensional data into very low feature set which may contain almost same information as original set.

TABLE I. ACCURACY (%) FOR DIFFERENT SUBSETS OF EEG DATA. ST. FREQ., SUB, NO. OF ATTR. AND ACC INDICATE STIMULUS FREQUENCY, SUBJECTS, NUMBER OF FEATURES AND ACCURACY, RESPECTIVELY.

| *St. Freq.* | *Sub.* | *No. of attr.* | *Acc (%)* | *No. of attr* | *Acc (%)* | *No. of attr* | *Acc (%)* |
|---|---|---|---|---|---|---|---|
| 8Hz | 1 | 15 | 100 | 30 | 100 | 60 | 100 |
| | 2 | | 100 | | 100 | | 100 |
| | 3 | | 100 | | 100 | | 100 |
| | 4 | | 72 | | 72 | | 72 |
| 14Hz | 1 | 15 | 97 | 30 | 97 | 60 | 97 |
| | 2 | | 97 | | 97 | | 97 |
| | 3 | | 100 | | 100 | | 100 |
| | 4 | | 100 | | 100 | | 100 |
| 28Hz | 1 | 15 | 100 | 30 | 100 | 60 | 100 |
| | 2 | | 100 | | 100 | | 100 |
| | 3 | | 100 | | 100 | | 100 |
| | 4 | | 100 | | 100 | | 100 |

NN are also used for finding sum squares errors (SSE), as shown in eq. 10, where 'P' is numbers of patterns, 'O' is the numbers of output, 't' for target output and 'a' for actual output. We make 1000 iterations for a selected feature. The error curve with 100 iterations is shown in Fig. 6. We explore 8Hz_Sub2_30 for 30 features of second subject at 8Hz stimulation. In this way, other features are kept in the figure. For 1000th iteration, errors 0.001869, 0.002549 and 0.001869 are found for 8Hz_Sub2_30, 14Hz_Sub1_30 and 28Hz_Sub4_30 accordingly. Thus errors are less than 0.3% which shows performance of NCCA network is better.

$$SSE = \frac{1}{2}\sum_{i=1}^{P}\sum_{j=1}^{O}\left(t_{i,j} - o_{i,j}\right)^2 \quad \text{.....10}$$

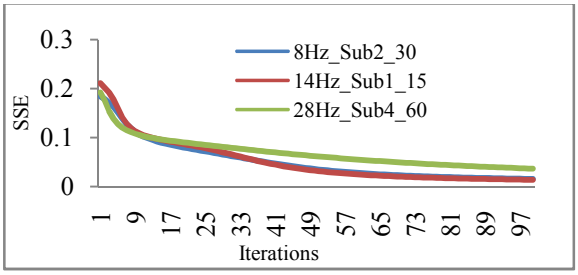



Figure 6. Iterations vs. error curve for three EEG features.

This EEG data contains brain conditions of human being. It may vary with subject's variations as well as with different stimulations. But though these are the time varying signals, they show coherency with one another on the basis of subjects and stimulations. Also, since NCCA finds out the features among the subdivision of entire EEG data via correlation maximization, there must have some kind of dependency among selected attributes. We observe this dependency with NN by varying the testing set with different subjects and stimulations. TABLE II shows dependency with different stimulus frequency as a confusion matrix for 15 attributes of subject 4. We have total 128 patterns (rows), 75% (96 patterns) of them are used as training set and rest (32 patterns) are used as a testing set. It is shown from TABLE II that there has no coincidence when test by 14 Hz stimulation but trained using 8 Hz stimulation; because all test data are misclassified with training set. But 56.25% test dataset of 28 Hz are coinciding with trained set of 8Hz.

We find recognition rate by considering different subjects as testing set. TABLE III shows these rates for 15 attributes of 28 Hz stimulation. In TABLE III, we denote subject1 as S1 and in this way for others. Firstly, we train the network using 96 patterns of S1, and then we test using 32 patterns of S1, S2, S3 and S4 accordingly, which shows that 32 patterns are classified for S1, 21 for S2, 13 for S3, but all are misclassified for S4. That's why we can say that 100% of S1, 65.63% of S2, 40.63% of S3 data are coincide with S1, but no coincidence data between S4 and S1. So, we can say that though brain signals are time varying quantity, it is possible to test the concidence between signals from selected features of NCCA network. It is compatible and reliable, because highest classified are found for own dataset. Also it shows low computational time.

TABLE II. CONFUSION MATRIX FOR 15 ATTRIBUTE OF SUBJECT 4

| | | **Training dataset** | | |
|---|---|---|---|---|
| | | ***8 Hz*** | ***14 Hz*** | ***28 Hz*** |
| **Test data set** | ***8 Hz*** | 23 | 4 | 8 |
| | ***14 Hz*** | 0 | 32 | 15 |
| | ***28 Hz*** | 18 | 13 | 32 |

TABLE III. CONFUSION MATRIX FOR 15 ATTRIBUTE OF 28 HZ STIMULATION

| | | **Training dataset** | | | |
|---|---|---|---|---|---|
| | | ***S1*** | ***S2*** | ***S3*** | ***S4*** |
| **Test data set** | ***S1*** | 32 | 16 | 21 | 09 |
| | ***S2*** | 21 | 32 | 23 | 01 |
| | ***S3*** | 13 | 10 | 32 | 17 |
| | ***S4*** | 0 | 09 | 14 | 32 |

In our EEG datasets, there have almost above 31,500 sampling points for a subject. When such a high dimensional data are feed into statistical CCA or NN seperately, it takes very long time to find features. On the other hand, we divide the whole dataset into three subdivisions and they are feed into NCCA network sequencially. From these subdivisions of data one can find out feature subsets of any desired length. It also shows computation only within a few seconds for find a feature subset. So, we may claim that NCCA shows better performance than that of individual CCA or NN. We performed the program by an Intel(R) Core(TM) i5-2450M CPU with 2.5GHz,4.00 GB of RAM, Operating system 64-bit computer.

## V. CONCLUSION

In this study, the NCCA approach is employed for FS of high dimensional EEG dataset. Firstly, every attribute of a subject is classified using wavelet clustering method. Then

class information is added to the original dataset. Expected subsections are created from whole database of a subject. NCCA is applied for finding correlation by weight maximization. Then features are deleted from original dataset according to weight minimization process. Thus only maximum weights features are remained. These features are tested according to BP rule, where training is done by 75% attributes and rests are used for testing. Finally, we get 100% accuracy mostly for every dataset except subject 4 for 8Hz stimulation due to higher level of noise. So, NCCA is a suitable choice for FS with a low computational cost**.** It is possible to recognize frequency of SSVEP to know the brain condition from selected EEG features with higher accuracy and lower computation cost, which will be our further study.

## REFERENCES


[1] D. H. Zhu, J. Bieger, G. G. Molina, R. M. Aarts, "A survey of stimulation methods used in SSVEP-based BCIs," *Comput. Intell. And Neurosci,*vol. 2010, doi: 10.1155/2010/702357, 2010

[2] M. Z. Hossain, M. J. A. Rabin, A. F. M. N. Uddin, M. Shahjahan, "Canonical correlation analysis with neural network for inter subject variability realization of EEG data," *in proc. IEEE of the 2nd Int. Conf. ICIEV,* Dhaka, 17-18 May, 2013, pp. 1-5.

[3] P. L. Lai and C. Fyfe, "A neural implementation of canonical correlation analysis," *Elsevier on Neural Networks,* vol. 12, pp. 1391-1397, 1999.

[4] M. M. Kabir., M. Shajahan., & M. Kazuyuki (2012). "A new hydrid ant colony optimization algorithm for feature selection," Elsevier on Expert Systems with Applications, vol. 39, pp. 3747-3763, 15 Feb. 2012.

[5] M. H. Aghdam, N. G. Aghaee & M. E. Basiri, "Text feature selection using ant colony optimization," *Elsevier on Expert Systems with Applications,* vol. 36, pp. 6843-6853, April 2009.

[6] H. Liu and L. Tu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Transactions on Knowledge and Data Engineering,* vol. 17, no. 4, pp. 491-502, 2005.

[7] R. Caruana, and D. Freitag, "Greedy atribute selection," in Proc. of the 11th Int. Conf. of machine learning, 1994, pp. 28-36

[8] P. Pudil, J. Novovico & J. Kittler, "Floating search methods in feature selection," *Pattern Recognition Letters,* vol. 15, pp. 1119-1125, 1994.

[9] B. Paskaleva, M. M. Hayat, Z. Wang, J. ScottTyo, and S. Krishna, "Canonical Correlation Feature Selection for Sensors with Overlapping Bands: Theory and Application," *IEEE transactions on geo science and remote sensing,* vol. 46, no. 10, 2010.

[10] R. Lotlikar, R. Kothari, "Bayes-optimality motivated-linear and multi layered perceptron-based dimansionality reduction," *IEEE Trans. On Neural Networks* vol. 11, pp. 452–463, 2000.

[11] M. M. Kabir, M. M. Islam and K. Murase, "A new wrapper feature selection approach using neural network," *Elsevier on Neurocomputing*, vol. 73, pp. 3273–3283, 2010.

[12] H. Bakardjian. (2010) Hovagim Bakardjian homepage. [Online] Available: http://www.bakardjian.com/work/ssvep_data_Bakardjian.html

[13] H. Bakardjiana, T. Tanakaa, A. Cichocki, "Optimization of SSVEP brain responses with application to eight-command Brain–Computer Interface," *Neuroscience Letters*, vol. 469, pp. 34–38, 2010.

[14] D. Regan, Steady-state evoked potentials, *J. Opt. Soc. Am.* Vol. 67, pp. 1475–1489, 1977.

[15] M. M. Müller, S. Hillyard, "Concurrent recording of steady-state and transient event related potentials as indices of visual-spatial selective attention," *Clin. Neurophysiol.* Vol. 111, pp. 1544–1552, 2000.

[16] R. B. Cattell, "The description of personality: Basic traits resolved into clusters," *Journal of Abnormal and Social Psychology*, vol. 38, pp. 476-506, 1943.

[17] M. Misiti, Y. Misiti, G. Oppenheim, J. M. Poggi, *"Clustering Signals Using Wavelets,"* Computational and Ambient Intelligence, Lecture Notes in Computer Science, vol. 4507, 2007, pp 514-521

[18] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321-377, 1936

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Numerical Investigation of Singularity at a Vertex in 3D Transversely Isotropic Piezoelectric Bonded Joints by FEM

Md. Shahidul Islam*, Mohiuddin Ahmed, and M. M. Kamal Uddin
Department of Mechanical Engineering, Khulna University of Engineering & Technology, Bangladesh
*E-mail: shahidulbitk@gmail.com

*Abstract*—**Stress singularity at the vertex of bonded joints is one of the main factors responsible for debonding under mechanical or electrical loading. The distribution of stress singularity field near the vertex of bonded joints is very important to maintain the reliability of intelligent materials. Piezoelectric materials are being widely used in the electronics industry, due to their high piezoelectric performance. In this paper, stress singularity at vertex in 3D transversely isotropic piezoelectric bonded joints is analyzed. Eigen analysis based on FEM is used for stress singularity field analysis of piezoelectric bonded joints. The Eigen equation is used for calculating the order of stress singularity, and the angular function of elastic displacement, electric potential, electric field, stress and electric displacement. The numerical result shows that the angular functions have large value near the interface edge than the inner portion of the joint. It was observed from the result, that the possibility of debonding and delamination at the interface edge of the piezoelectric bonded joints, due to the higher stress and electric displacement concentration at the free edge.**

***Keywords—Piezoelectric material, stress singularity, transversely isotropic material, smart structures, finite element method.***

## I. INTRODUCTION

Stress singularity field occurs at a vertex on an interface due to a discontinuity of materials. The stress singularity field near the vertex of two-phase materials is one of the main factors being responsible for debonding under mechanical or electrical loadings. It is very important for maintaining the reliability of intelligent materials to make clear the distribution of stress near the vertex. In recent years, intelligent or smart structures and systems have become an emerging new research area. Piezoelectric material, due to its characteristic direct-converse piezoelectric effect, has naturally received considerable attentions [1]. Piezoelectric materials are being widely used in the electronics industry due to their high functionality. Mechanical stress occurs in piezoelectric material for any electric input. The stress concentrations caused by mechanical or electric loads may lead to crack initiation and extension, and sometimes the stress concentrations may be high enough to fracture the material parts. Reliable service lifetime predictions of piezoelectric components demand a complete understanding of the debonding processes of these materials.

Sosa has suggested a general method of solving plane problems of piezoelectric media with defects [2]. Zak and Williams used eigen functions for analyzing stress singularity field at a crack tip perpendicular to a bimaterial interface. They found that a real part of eigen value is within the range of 0 to 1, and expressed a relationship between stress distribution and the order of stress singularity at the crack tip [3]. Aksentian determined eigen values and eigen vectors at the singular point in plane intersecting a free edge of the interface in three dimensional dissimilar joints [4]. Yamada and Okumura developed a finite element analysis for solving eigen value equation to determine directly the order of stress singularity and the angular variation of the stress and displacement fields [5]. Pageau and Biggers determined the order of stress singularity and the angular variation of the displacement and the stress fields around the singular points in plane intersecting a wedge front in the three-dimensional anisotropic material structures using the two dimensional displacement formulation under a plane strain assumption [6].

At present, no clear picture exists of the problem of singularity at a vertex for 3D transversely isotropic piezoelectric bonded joints. Therefore, the purpose of the present investigation is to obtain a better understanding of the problem of singularity at a vertex in a two-phase transversely isotropic piezoelectric material joint. In this study an eigen equation based on three-dimensional FEM was formulated for analyzing the singularity in piezoelectric bonded joints.

## II. THE BASIC FORMULA

In the absence of body forces and free charges, the equilibrium equations of piezoelectric materials are expressed as follows [7]:

$$\sigma_{ij,j} = 0 \tag{1}$$

$$d_{i,i} = 0 \tag{2}$$

where $\sigma_{ij}$ represents the stress and $d_i$ the electric displacement. The constitutive relations are shown as follows [7]:

$$\sigma_{ij} = c_{ijkl}\varepsilon_{kl} - e_{kij}E_k \tag{3}$$

$$d_i = e_{ikl}\varepsilon_{kl} + \chi_{ik}E_k \tag{4}$$

where $c_{ijkl}$ is the elastic constant, $e_{ikl}$ and $\chi_{ik}$ are the piezoelectric constant and electric permittivity (dielectric constant), respectively.

The elastic strain-displacement and electric field-potential relations are presented as follows [7]:

$$\varepsilon_{ij} = \frac{1}{2}\left(u_{j,i} + u_{i,j}\right) \tag{5}$$

$$E_i = -\psi_{,i} \tag{6}$$

where $\varepsilon_{ij}$ is the strain, and $E_i$ is the eclectic field. $u_i$ is the elastic displacement and $\psi$ is the electric potential.

Fig. 1 represents the geometry of a typical case where a singular stress state occurs at the point o. The region surrounding the singular point is divided into a number of quadratic pyramidal elements with a summit o, with each element being located in spherical co-ordinates $r$, $\theta$, and $\phi$ by its nodes 1 to 8.

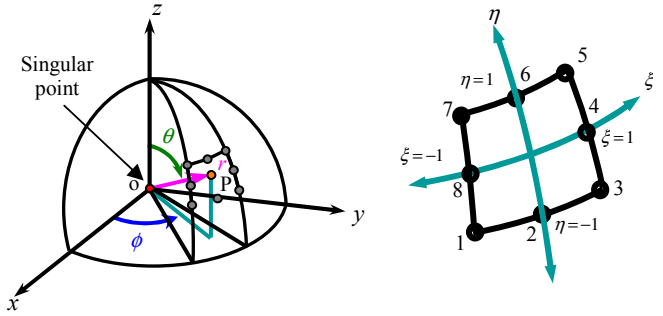



Figure1. Element geometry and natural co-ordinates at a free edge singular point

A point P in the element can be located using the singular transformation by the relations [7].

$$r = r_o\left(\frac{1+\alpha}{2}\right)^{1/P} \quad or, \quad \rho = \frac{r}{r_o}\left(\frac{1+\alpha}{2}\right)^{1/P}$$

$$\theta = \sum_{i=1}^{8} H_i\theta_i \quad and \quad \phi = \sum_{i=1}^{8} H_i\phi_i \tag{7}$$

where,

$$H_1 = -\frac{1}{4}(1-\eta)(1-\xi)(\eta+\xi+1) \qquad H_2 = \frac{1}{2}(1-\eta)\left(1-\xi^2\right)$$

$$H_3 = -\frac{1}{4}(1-\eta)(1+\xi)(-\eta+\xi-1) \qquad H_4 = \frac{1}{2}\left(1-\eta^2\right)(1+\xi)$$

$$H_5 = \frac{1}{4}(1+\eta)(1+\xi)(\eta+\xi-1) \qquad H_6 = \frac{1}{2}(1+\eta)\left(1-\xi^2\right)$$

$$H_7 = \frac{1}{4}(1+\eta)(1-\xi)(\eta-\xi-1) \qquad H_8 = \frac{1}{2}\left(1-\eta^2\right)(1-\xi)$$

$\theta$ and $\phi$ are the nodal values of the angular co-ordinates and $\alpha$, $\eta$, and $\xi$ are natural co-ordinates of the element whose ranges are shown in Fig. 1. $H$ interpolation function, $p$ Eigen value, $\rho = r/r_o$, $r$ the distance from the singular point. The elastic displacement and electric potential relations are:

$$\left(\bar{u} - \bar{u}_o\right) = \left(\frac{1+\alpha}{2}\right)\left[\sum_{i=1}^{8} H_i\left(\bar{u}_i - \bar{u}_o\right)\right] \tag{8}$$

$$\left(\bar{\psi} - \bar{\psi}_o\right) = \left(\frac{1+\alpha}{2}\right)\left[\sum_{i=1}^{8} H_i\left(\bar{\psi}_i - \bar{\psi}_o\right)\right] \tag{9}$$

where $\bar{u}_o$ and $\bar{u}$ is the elastic displacement at o and P, respectively, and $\bar{\psi}_o$ and $\bar{\psi}$ is the electric potential at o and P, respectively.

$$u = \left(\bar{u} - \bar{u}_o\right), \qquad u_i = \left(\bar{u}_i - \bar{u}_o\right) \tag{10}$$

$$\psi = \left(\bar{\psi} - \bar{\psi}_o\right), \qquad \psi_i = \left(\bar{\psi}_i - \bar{\psi}_o\right) \tag{11}$$

From the above equation, the elastic displacement and electric potential relations are expressed as follows [7]:

$$u_k = \rho^P\left[\sum_{i=1}^{8} H_i u_{ki}\right] \qquad (k = r, \theta, \phi) \tag{12}$$

$$\psi = \rho^P\left[\sum_{i=1}^{8} H_i\psi_i\right] \tag{13}$$

The Jacobian matrix relating the spherical coordinates to the natural coordinates is given below:

$$[\mathbf{J}] = \begin{bmatrix} \frac{\partial r}{\partial\alpha} & \frac{\partial\theta}{\partial\alpha} & \frac{\partial\phi}{\partial\alpha} \\ \frac{\partial r}{\partial\xi} & \frac{\partial\theta}{\partial\xi} & \frac{\partial\phi}{\partial\xi} \\ \frac{\partial r}{\partial\eta} & \frac{\partial\theta}{\partial\eta} & \frac{\partial\phi}{\partial\eta} \end{bmatrix} = \begin{bmatrix} \frac{r_o}{2P}\rho^{1-P} & 0 & 0 \\ 0 & \sum_{i=1}^{8} H_{i,\xi}\theta_i & \sum_{i=1}^{8} H_{i,\xi}\phi_i \\ 0 & \sum_{i=1}^{8} H_{i,\eta}\theta_i & \sum_{i=1}^{8} H_{i,\eta}\phi_i \end{bmatrix} \tag{14}$$

In (14) there is no dependence between the radial coordinate and the angular coordinate. A sub-matrix is extracted from the above equation as follows:

$$[\mathbf{J_1}] = \begin{bmatrix} \frac{\partial\theta}{\partial\xi} & \frac{\partial\phi}{\partial\xi} \\ \frac{\partial\theta}{\partial\eta} & \frac{\partial\phi}{\partial\eta} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{8} H_{i,\xi}\theta_i & \sum_{i=1}^{8} H_{i,\xi}\phi_i \\ \sum_{i=1}^{8} H_{i,\eta}\theta_i & \sum_{i=1}^{8} H_{i,\eta}\phi_i \end{bmatrix} \tag{15}$$

The strain and electric potential equation is obtained form (7) and (12), (13) and (15) by using the chain rule of differentiation. The strain and electric potential in a spherical coordinate system as follows [8]:

$$\left\{\boldsymbol{\varepsilon}^*\right\} = \sum_{i=1}^{8}[\mathbf{B}_i]\left\{\mathbf{u}_i^*\right\} = [\mathbf{B}]\left\{\mathbf{u}^*\right\} \tag{16}$$

Where

$$\left\{\boldsymbol{\varepsilon}^*\right\}^T = \left\{\varepsilon_{rr} \quad \varepsilon_{\theta\theta} \quad \varepsilon_{\phi\phi} \quad \varepsilon_{r\theta} \quad \varepsilon_{r\phi} \quad \varepsilon_{\theta\phi} \quad E_r \quad E_\theta \quad E_\phi\right\}$$

$$\left\{u_i^*\right\}^T = \left\{\bar{u}_{ri} \quad \bar{u}_{\theta i} \quad \bar{u}_{\phi i} \quad -\bar{\psi}_i\right\}$$

## III. THE EIGEN EQUATION

The eigen equation was formulated for determining the order of stress singularity as follows [9]:

$$\left(p^2[\mathbf{A}]+p[\mathbf{B}]+[\mathbf{C}]\right)\{\mathbf{U}\}=\{0\} \tag{17}$$

where

$$\{\mathbf{U}\}=\begin{Bmatrix}u_r\\u_\theta\\u_\phi\\\psi\end{Bmatrix}$$

$$[\mathbf{A}]=\sum_{\mathbf{s}}\left([\mathbf{k}_a-\mathbf{k}_{sa}]\right)$$

$$[\mathbf{B}]=\sum_{\mathbf{s}}\left([\mathbf{k}_b-\mathbf{k}_{sb}]\right)$$

$$[\mathbf{C}]=\sum_{\mathbf{s}}\left([\mathbf{k}_c-\mathbf{k}_{sc}]\right)$$

In (17) $p$ represents the characteristic root, which is related to the order of singularity, $\lambda$, as $\lambda = 1-p$. [**A**], [**B**] and [**C**] are matrices composed of material properties, and {**U**} represents the elastic displacement and electric potential vector.

$$[\mathbf{k}_a]=\int_{-1}^{1}\int_{-1}^{1}[\mathbf{B}_a]^T[\mathbf{D}][\mathbf{B}_a]\sin\theta|J_1|d\xi d\eta$$

$$[\mathbf{k}_b]=\int_{-1}^{1}\int_{-1}^{1}\left([\mathbf{B}_a]^T[\mathbf{D}][\mathbf{B}_a]+[\mathbf{B}_b]^T[\mathbf{D}][\mathbf{B}_a]\right)\sin\theta|J_1|d\xi d\eta$$

$$[\mathbf{k}_c]=\int_{-1}^{1}\int_{-1}^{1}[\mathbf{B}_b]^T[\mathbf{D}][\mathbf{B}_b]\sin\theta|J_1|d\xi d\eta$$

$$[\mathbf{k}_{sa}]=2\int_{-1}^{1}\int_{-1}^{1}[\mathbf{H}]^T[\mathbf{SD}][\mathbf{B}_a]\sin\theta|J_1|d\xi d\eta$$

$$[\mathbf{k}_{sb}]=\int_{-1}^{1}\int_{-1}^{1}2\left([\mathbf{H}]^T[\mathbf{SD}][\mathbf{B}_b]+[\mathbf{H}]^T[\mathbf{SD}][\mathbf{B}_a]\right)\sin\theta|J_1|d\xi d\eta$$

$$[\mathbf{k}_{sc}]=\int_{-1}^{1}\int_{-1}^{1}[\mathbf{H}]^T[\mathbf{SD}][\mathbf{B}_b]\sin\theta|J_1|d\xi d\eta \tag{18}$$

Equation (17) now expressed as follows:

$$\left(-p[\mathbf{B}]-[\mathbf{C}]\right)\{\mathbf{U}\}=p^2[\mathbf{A}]\{\mathbf{U}\} \tag{19}$$

Finally, letting $\{\mathbf{V}\}=p\{\mathbf{U}\}$, the characteristic equation can be transformed into the standard Eigen problem.

$$\begin{matrix}-[\mathbf{A}]^{-1}[\mathbf{B}] & -[\mathbf{A}]^{-1}[\mathbf{C}]\\ [\mathbf{I}] & [0]\end{matrix}\begin{Bmatrix}\{\mathbf{V}\}\\\{\mathbf{U}\}\end{Bmatrix}=p\begin{Bmatrix}\{\mathbf{V}\}\\\{\mathbf{U}\}\end{Bmatrix} \tag{20}$$

## IV. RESULT AND DISCUSSION

The elastic displacement and electric potential equation is expressed by the following equitation.

$$u_j(r,\theta,\phi)=b_j(\theta,\phi)r^{1-\lambda} \tag{21}$$

$$\psi(r,\theta,\phi)=q(\theta,\phi)r^{1-\lambda} \tag{22}$$

By differentiating the above two equations, get the angular function of strain and electric field equation respectively. The stress and electric displacement distribution equations in the stress singularity region can be expressed as follows.

$$\sigma_{ij}(r,\theta,\phi)=K_{ij}r^{-\lambda}f_{ij}(\theta,\phi) \tag{23}$$

$$d_i(r,\theta,\phi)=F_i r^{-\lambda}l_i(\theta,\phi) \tag{24}$$

Where $r$ represents the distance from the stress singular point, $b_j(\theta, \phi)$ the angular function of elastic displacement, $q(\theta, \phi)$ the angular function of electric potential, $f_{ij}(\theta, \phi)$ the angular function of stress distribution, $l_i(\theta, \phi)$ the angular function of electric displacement, $K_{ij}$ the intensity of singularity, $F_i$ the intensity of electric field, and $\lambda$ the order of stress singularity. Angular functions of stress and electric displacement components obtained from Eigen analysis in (17) are examined.

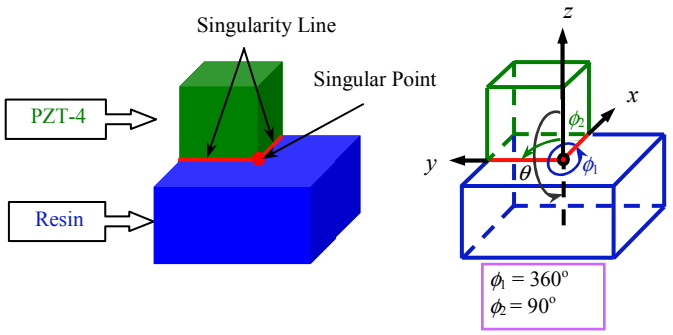



Figure 2. Singular point of 3D piezoelectric joint in $x$, $y$, $z$ plane

TABLE I. MATERIAL PROPERTY OF RESIN AND PZT-4

| Material | Elastic Constant, $10^{10}$ N/m$^2$ | | | | |
|---|---|---|---|---|---|
| | $c_{11}$ | $c_{12}$ | $c_{13}$ | $c_{33}$ | $c_{44}$ |
| Resin | 5.56 | 3.41 | 3.41 | 5.56 | 1.08 |
| PZT-4 | 13.9 | 7.78 | 7.43 | 11.3 | 2.56 |

| Material | Piezoelectric Constant, C/m$^2$ | | | Dielectric Constant, $10^{-10}$ C/Vm | |
|---|---|---|---|---|---|
| | $e_{31}$ | $e_{33}$ | $e_{15}$ | $\chi_{11}$ | $\chi_{33}$ |
| Resin | 0.0 | 0.0 | 0.0 | 37.9 | 37.9 |
| PZT-4 | -6.98 | 13.8 | 13.4 | 60.0 | 54.7 |

Solving eigen equation yields many roots $p$ and eigen vectors corresponding to each eigen value are obtained. However, if the root is within the range of $0 < p < 1$, this fact indicates that the stress field has singularity. The value of order of singularity at the singularity corner and line are shown in the following tables.

TABLE II. ORDER OF SINGULARITY FOR RESIN AND PZT-4

| Material | Order of singularity | | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 |
| Resin and PZT-4 | $\lambda_{line}$ | 0.3494 | 0.0895 | 0.0264 | ---- |
| | $\lambda_{vertex}$ | 0.4304 | 0.0291 | 0.0081 | 0.0023 |

The distributions of elastic displacement and electric potential are shown in fig. 3. It is found that, the angular function of elastic displacement and electric potential is

continuous at the interface of the joints and has larger value at the free edge than the inner portion. So there is a possibility to debond near the free edge of the joint. It is also found that the angular function of elastic displacement and electric potential at $\phi = 0°$ is agreed with that at $\phi = 90°$. The figure of $b_\theta$, $b_r$ and q is symmetry with respect to $\phi$ and the figure of $b_\phi$ is anti-symmetry with respect to $\phi$.

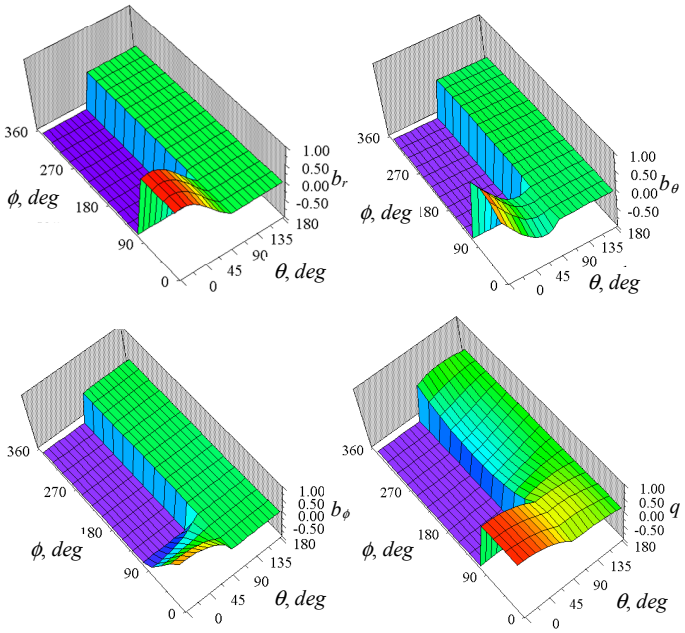
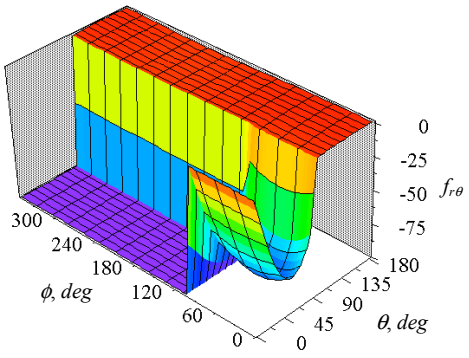



Figure 3. Distribution of $b_i$ & $q$ against angle, $\phi$ and $\theta$

The distributions of angular function of stress in $\phi$-$\theta$ plane for the piezoelectric bonded structure are shown in fig. 4. The angular function of stress $f_{r\theta}$ and $f_{\theta\theta}$ are symmetry and $f_{\phi\theta}$ is anti-symmetry at $\phi = 45°$ on interface ($\theta = 90°$). The value of angular function of stress $f_{\theta\theta}$ is larger than that of the $f_{r\theta}$ and $f_{\phi\theta}$. It is shown from the figure that the stresses have larger value near the free edge of the bonded joints. So there is another possibility to debond and delamination occurs near the free edge of the bonded joints.



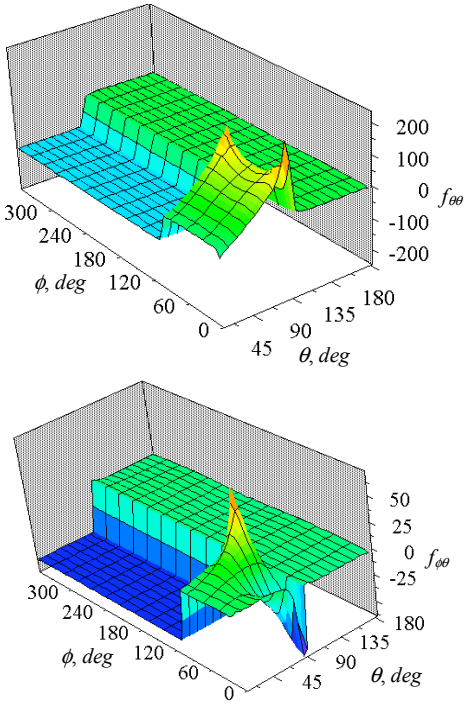
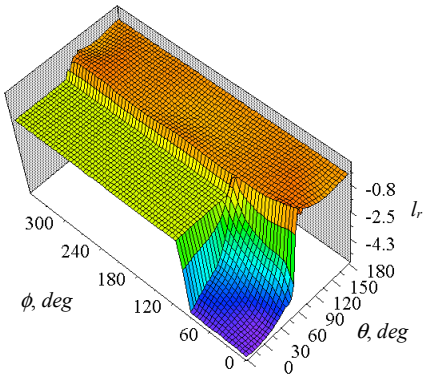




Figure 4. Distribution of stress $f_{ij}$ against angle $\phi$ and $\theta$

The distributions of angular function of electric displacement in $\phi$-$\theta$ plane for the piezoelectric bonded structure are shown in fig. 5. The angular function of electric displacement $l_r$ and $l_\theta$ are symmetry and $l_\phi$ is anti-symmetry at $\phi = 45°$ on interface ($\theta = 90°$). The value of angular function of electric displacement $l_\theta$ is larger than that of the $l_r$ and $l_\phi$. It is shown from the figure that the electric displacements have larger value near the free edge of the bonded joints. So there is a possibility to debond and delamination occurs near the free edge of the bonded joints.



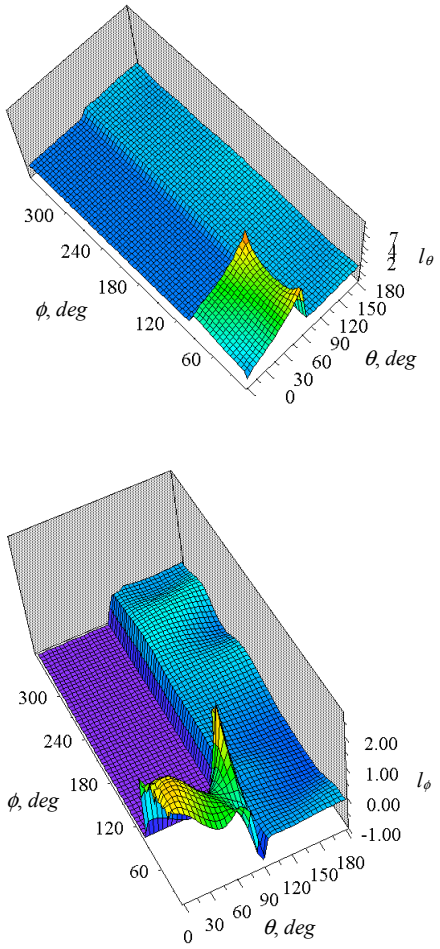



Figure 5. Distribution of electric displacement $l_i$ against angle $\phi$ and $\theta$

## V. CONCLUSION

An eigen equation formulation near the vertex of transversely isotropic piezoelectric bonded joint was presented. Angular functions for singularity corner were derived from eigen analysis using a finite element method. From the numerical results, the following conclusions can be drawn for the piezoelectric dissimilar joints.

(a) The value of the order of singularity at the singularity corner is larger than that at the line.

(b) Larger value of the angular function occurs at the interface edge in the material joint than the inner portion of the joint.

(c) It is suggested that delamination of the interface may occur at the interface edge of the piezoelectric material joints.

## REFERENCES


[1] H. Ding, and Chenbuo, "On the Green's functions for two-phase transversely isotropic piezoelectric media," *International Journal of Solids and Structure*, vol. 34, no. 23, pp. 3041-3057, 1997.

[2] H. A. Sosa, "Plane problems in piezoelectric media with defects," *International Journal of Solids and Structure*, vol. 28, pp. 491-505, 1991.

[3] A. R. Zak, and M. L. Williams, "Crack point stress singularities at a bi-material interface," *Journal of Applied Mechanics*, Brielf Notes, pp.142-143, 1963.

[4] O. K. Aksentian, "Singularities of the stress-strain state of a plate in the neighborhood of an edge," *PMM*, vol. 31, no. 1, pp.178-186, 1967.

[5] Y. Yamada, and H. Okumura, "Analysis of local stress in composite materials by 3-D finite element," *Proceeding of the Japan-U.S. Conference*, Tokyo, pp. 55-64, 1981.

[6] S. S. Pageau, and Jr. S. B. Biggers, "A finite element approach to three-dimensional singular stress states in anisotropic multi-material wedges and junctions," *International Journal of Solids and Structure*, vol. 33, pp. 33-47, 1996.

[7] K. M. Liew, and J. Liang, "Modeling of 3D piezoelectric and elastic bimaterials using the boundary element method," *Computational Mechanics*, vol. 29, pp. 151-162, 2002.

[8] M. S. Islam, and H. Koguchi, "Characteristics of singular stress distribution at a vertex in transversely isotropic piezoelectric dissimilar material joints," *JSME Journal of solid Mechanics and Material Engineering*, vol.4 no. 7, pp. 1011-1026, 2010.

[9] S. S. Pageau, P. F. Joseph, and Jr. S. B. Biggers, "Finite element analysis of anisotropic materials with singular inplane stress fields," *International Journal of Solids and Structure*, Vol. 32, pp. 571-591, 1995.

# Flexible Graphene Field Effect Transistor with Graphene Oxide Dielectric on Polyimide Substrate

Mohi Uddin Jewel*, Tanvir Ahamed Siddiquee, and Md. Rafiqul Islam
Department of Electrical and Electronic Engineering,
Khulna University of Engineering and Technology,
Khulna-9203, Bangladesh.
* E-mail: jewel2080@gmail.com

*Abstract*—**Graphene is an excellent material for flexible electronics due to its high carrier transport properties. We report a flexible graphene field effect transistor on polyimide substrate using graphene oxide as top-gate dielectric. Good current saturation and peak hole and electron mobilities of 496 $cm^2$/(V.s) and 164 $cm^2$/(V.s) are observed, respectively, for the proposed intrinsic RF device. A maximum transconductance of 0.42 mS and the intrinsic cutoff frequency of 117 GHz are achieved when the gate length is reduced up to 0.25 µm.**

***Keywords—Flexable GFET; Graphene Oxide; Polyimide substrate; Intrinsic cutoff frequency..***

## I. INTRODUCTION

Graphene is a monolayer of tightly bonded carbon atoms, having thickness of 0.34 nm, arranged in a two dimensional (2D) honeycomb lattice structure [1]. This, perfect two dimensional crystal has some unique properties. It is transparent, flexible, and stronger than structural steel, thinnest and lightest material ever known. However, electronic properties are much important for using this interesting material in electronic devices. Band structure of graphene was first realized by P. R. Wallace in 1947 [2]. Monolayer graphene shows a linear E-k relation around the K point of the first Brillouin zone. Intrinsic graphene is gapless and acts as a semimetal [3]. Electron mobility has exceeded 200000 cm2/(V.s) in suspended graphene which is higher than any conventional semiconductors [4].

For digital logic applications, a bandgap larger than 0.4 eV and $I_{on}/I_{off}$ ratio $10^4$-$10^6$ are required [5]. But, because of zero bandgap, graphene transistor never turns off and shows $I_{on}/I_{off}$ ratio only 2-20 [3]. As a result, zero bandgap graphene is not suitable for digital logic devices. While gapless nature hinders graphene's application in logic devices, it may find its places in radio-frequency (RF) electronics in which the transistor does not need to switch off.

The conventional semiconductors are rigid materials and not useful for flexible electronics as they are brittle in nature. Previously, this field was dominated by amorphous silicon and organic polymers, but they have very low field mobility [6]. On the other hand, graphene has higher mobility relative to organic semiconductors. So, graphene opens a new door to high frequency low voltage flexible and printable electronic applications [7].

A dielectric must make a good interface with the channel layer so that carrier transport does not get affected close to their interface. A high-*k* dielectric reduces Columbic scattering by screening charged impurities near graphene layer leading higher mobility. Ultrathin film is difficult to deposit over graphene surface due to absence of dangling bond. High-*k* dielectrics, such as $Al_2O_3$, $HfO_2$ and $ZrO_2$ cannot be grown in flexible plastic substrate as they require high growth temperature [8]. Although $Al_2O_3$ and $HfO_2$ films were deposited over graphene using atomic layer deposition technique, they make poor interface and are fragile in nature [9], [10]. Graphene oxide (GO), on the other hand, holding a good mechanical and optical properties and can be used for high performance flexible and transparent electronics. GO dielectric shows a dielectric constant of 5 at room temperature and very good environmental stability. Seoung-Ki Lee and co-workers were exhibited a good way to fabricate GO on graphene channel [11]. The performance of graphene field effect transistor (GFET) depends not only on the interface quality, but also on underlying substrate. GFET on poly-ethylene naphthalate (PEN) substrate melts down for drain-source voltage above 0.5 V [12]. Polyimide, a promising substrate for graphene based devices can be used as flexible substrate. Surface roughness of their interface can be reduced by using liquid polyimide PI-5878G and adhesion promoter VM651 [13].

In our work, we for the first time demonstrate an intrinsic monolayer GFET on flexible polyimide substrate with GO gate dielectric. We consider the effect of quantum capacitance ($C_q$) to calculate sheet carrier density ($\rho_{sh}$) and the effect of ionized impurity scattering, phonon scattering, and surface roughness scattering to calculate effective mobility and drain current. The proposed GFET shows a peak hole and electron mobilities 496 $cm^2$/(V.s) and 164 $cm^2$/(V.s), respectively, at $V_{ds}$= −0.1 V. Furthermore, an intrinsic cutoff frequency of 117 GHz is obtained for the channel length of 0.25 µm.

## II. DEVICE STRUCTURE

The proposed device structure is shown in Fig.1. The device consists of a large area monolayer graphene channel, graphene oxide (GO) gate dielectric, and flexible polyimide substrate. The device dimensions are taken in diffusive limit. 30 nm Ti/20 nm Au stacks are used as source and drain contacts. 10 nm thin GO is used as dielectric. Graphene

channel has a length of 2 μm and width of 1 μm. So far, we have neglected effect of source and drain contacts and analyzed the performance of an intrinsic GFET.

## III. DC CHARACTERISTICS

### A. Device Model

To understand the performance of the proposed GFET, it is important to investigate its dc characteristics. The model to simulate GFET we have used, was proposed in [14]. Quantum capacitance $C_q$, implies the change of charge carrier with respect to channel voltage.Quantum capacitance is given by [14] following expressions,

$$C_q = \frac{2q^2}{\pi}\frac{q|V_{ch}|}{(\hbar v_F)^2}, \quad (1)$$

where, $q$ is the elementary charge, $V_{ch}$ is the channel voltage, $\hbar$ the reduced Planck constant and $v_F$ is the Fermi velocity ($10^8$ cm/s). The channel voltage $V_{ch}$ can be represented by

$$V_{ch} = [V_{GS} - V_{DP} - V(x)]\frac{C_{ox}}{C_{ox} + \frac{1}{2}C_q}, \quad (2)$$

where $C_{ox}$ is the top gate oxide capacitance, $V_{DP}$ is the dirac point voltage and $V(x)$ is the voltage drop in the graphene channel and is zero at position, $x = 0$ and equal to $V_{DS}$ at $x = L$. Equation (1) and (2) are solved self-consistently to obtain $C_q$ and $V_{ch}$.. The sheet carrier density ($\rho_{sh}$) can be calculated by

$$\rho_{sh} = \left|-\frac{1}{2q}C_q V_{ch}\right| + \rho_{sho}, \quad (3)$$

where $\rho_{sho}$ is the residual carrier density. There exists residual charge in graphene at dirac point due to charge impurities close to the sheet on a substrate. Both GO dielectric and polyimide substrate cause p-type doping in graphene layer. Here $\rho_{sho} = 2 \times 10^{11}$ cm$^{-2}$ and $V_{DP} = 0.5$V are taken, respectively, for the simulation.

Mobility of an electronic device is limited by different scattering mechanisms. In this study, columbic, phonon, and surface scattering mechanisms are taken into consideration as prominent mechanisms. Thus, the total effective mobility can be calculated using the well known Matthiessen's rule, $\mu_{tot}^{-1} = \mu_{col}^{-1} + \mu_{ph}^{-1} + \mu_{surf}^{-1}$. The columbic, phonon, and surface roughness scattering dependent mobilities are given [15] by $\mu_{col} = A.(\rho_{sh})^{\alpha}$ , $\mu_{ph} = P.(\rho_{sh})^{\beta}(T/300)^{\delta}$, and $\mu_{surf} = \gamma/(\rho_{sh})^{\Delta}$ respectively, where A, P and γ are the fitting parameters, α is the degree of screening, T is the temperature with $\delta = 2$ and $\beta = -1$, and Δ is the degree of roughness

The saturation velocity ($v_{sat}$) can be expressed [14] by

$$v_{sat} = \frac{\Omega}{(\pi\rho_{sh})^{0.5+AV^2(x)}}, \quad (4)$$

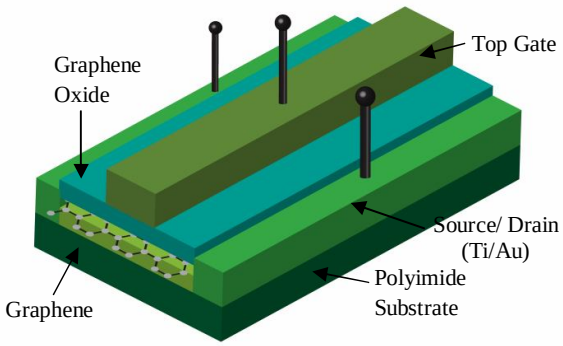



Figure 1. Schematic of top-gate graphene field effect transistor.

where, A is an empirical factor of the order $10^{-3}$ and Ω is the optical phonon frequency. Graphene on polyimide substrate shows optical phonon energy (ħΩ) equal to 53 meV [15]. Finally, with channel length $L$, width $W$, low field mobility $\mu$, drain-source voltage $V_{DS}$, and gate voltage $V_{GS}$, the drain current can be represented by [14],

$$I_d = q\mu W \frac{\int_0^{V_{DS}} \rho_{sh} dV}{L - \mu\int_0^{V_{DS}} \frac{1}{v_{sat}} dV}, \quad (5)$$

### B. Device Performance

Fig. 2 shows the transfer characteristics for different $V_{DS}$ where ambipolar conduction is clearly distinguished by a dirac point. The dirac point is shifted towards negative gate bias with increasing $V_{DS}$. The position of the Dirac point depends on several factors: the difference between the work functions of the gate material and the graphene, the type and density of the charges at the top and bottom interfaces of the channel, and doping of the graphene. The transfer characteristics indicate that the GFET never turns off. The value of residual charge at the dirac point increases with $V_{DS}$ as the channel potential depends not only on the $V_{GS}$ but also on the $V_{DS}$. Fig. 3 shows the output characteristics plotted for different negative gate bias $V_{GS}$. The output characteristics show a linear region, a weak saturation region and a second linear region for higher $V_{GS}$. Second linear region is disapeared for higher $V_{GS}$. Good current saturation is observed on output characteristics for higher $V_{GS}$. This saturation behavior is better than GFET of [11] and much similar to[13].

Hole mobility of 496 cm$^2$/V.s and electron mobility of 164 cm$^2$/V.s are obtained at $V_{DS} = -0.1$V. Hole mobility is higher and electron mobility is slightly lower than observed in [11]. Figure 4 (a) and (b) show the effective electron and hole mobilities along with individual component of mobility. The mobility due to ionized impurity scattering is so high due to oxygen absorption species and other residues; as a result, it does not have any contribution to effective mobility. It is well known that the polyimide plastic substrate cannot undergo high temperature treatment, so high impurity exists. As seen in Figs. 4(a) and (b) the $\mu_{tot}$ is firstly dominated by phonon scattering and as the carrier concentration increases surface

roughness scattering starts dominating for both hole and electron. But, hole trends higher surface roughness scattering compared to electron.

Fig. 5, clearly depicts the ambipolar conduction in GFET. As we go along the channel, the charge density decreases from source towards drain. When the potential of the channel is equal to potential corresponds to dirac point, the charge changes from p-type to n-type at x =1.5 µm. At this point, charge density is equal to residual charges ($\rho_{sho}$). This residual charge in responsible for GFET not to be switched off. After dirac point, the carrier density again increases with opposite type. This behavior is observed specifically in GFET, not any conventional FET. This is because of zero bandgap of graphene. Conventional FET shows unipolar characteristics and a pinch-off region near the drain terminal.

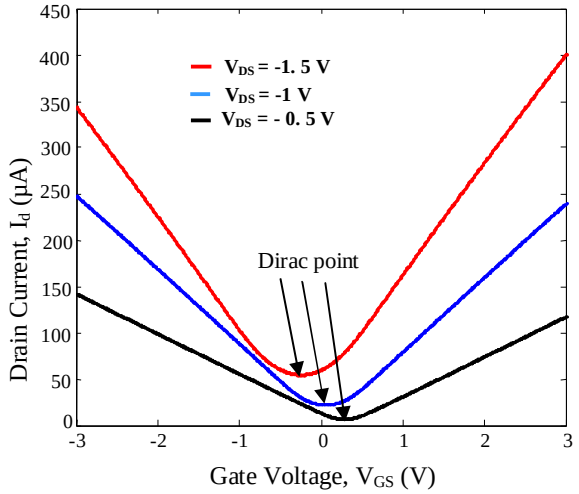



Figure 2. Transfer characteristics for different $V_{DS}$.

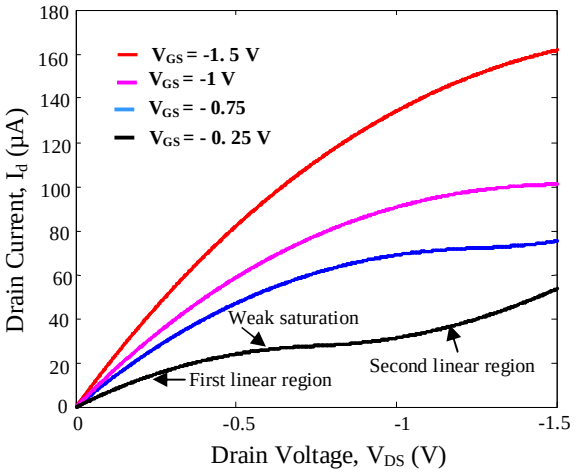



Figure 3. Output characteristics for different negative gate bias $V_{GS}$. The current conduction is taken place due to p-type carrier.

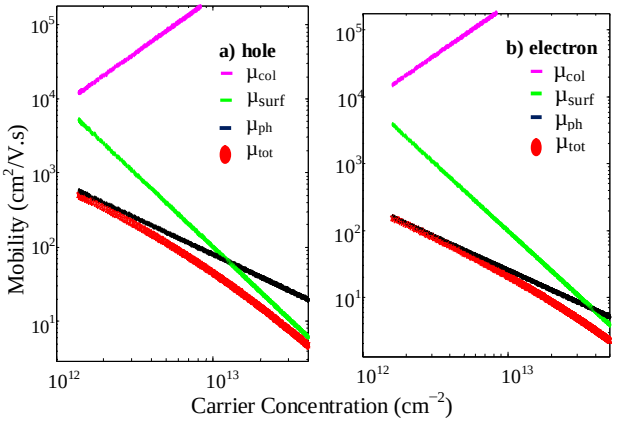



Figure 4. Carrier concentration dependent effective (a) hole, and (b) electron mobility for the proposed GFET.

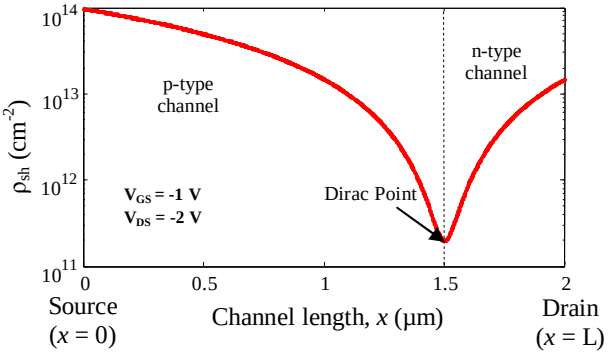



Figure 5. Charge density with respect to channel position. The dirac point indicates carrier switchover position in the channel.

## IV. RF PERFORMANCE

The transconductance ($g_m$) is an important parameter for understanding the RF performance of the GFET. Generally, high $g_m$ is desirable for high intrinsic gain and cutoff frequency. gm can be extracted from $I_d$-$V_{GS}$ characteristics of GFET. $g_m$ can be defined by,

$$g_m = \left.\frac{dI_d}{dV_{GS}}\right|_{V_{DS}=const.}, \qquad (6)$$

Fig. 6 shows $g_m$ as a function of gate voltage where the maximum $g_m$ of 0.42 mS is obtained for $V_{DS}$ kept constant at −1.5 V. $g_m$ takes a turn from negative to positive around dirac point due to change in charge carriers polarity. Cutoff frequency is another important parameter to characterize the GFET RF performance. It implies the frequency at which current gain becomes one. The intrinsic cutoff frequency ($f_{T.int}$) of a transistor is determined by charge carrier transit time ($\tau_t$) across the channel length ($L_{gate}$). The intrinsic cutoff frequency can be given [16] by

$$f_{T.\text{int}} = \frac{1}{2\pi\tau_t} = \frac{g_m}{2\pi C_g}, \qquad (7)$$

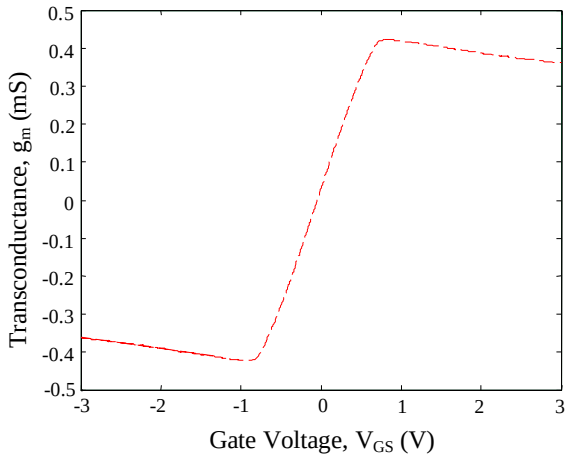



Figure 6. Transconductance as a function of gate voltage of GFET.

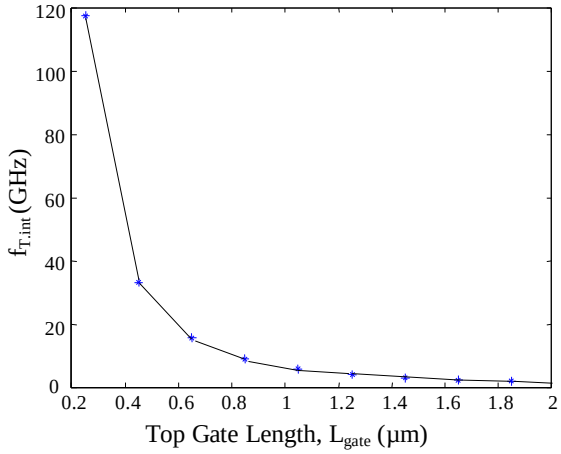



Figure 7. Variation of intrinsic cutoff frequency as a function of gate length.

The transit time can be given by

$$\tau_t = \frac{L_{gate}}{v_{drift}}, \tag{8}$$

The drift velocity $v_{drift}$ is accurately modeled by Thornber's equation [17]

$$v_{drift} = \left( \left( \frac{1}{\mu E_x} \right)^{\beta} + \left( \frac{1}{v_{sat}} \right)^{\beta} \right)^{(-1/\beta)}, \tag{9}$$

with $\beta = 2$, electric field $E_x = V_{DS}/ L_{gate}$.

Fig. 7 shows $f_{T.int}$ with respect to gate length. From Eqn. (8) it is cleared that the transit time decreases with gate length. As a result, cutoff frequency increases from 1 GHz to 117 GHz due to continuous channel length scaling from 2 μm to 0.25 μm.

## V. Conclusion

Performance of GFET on flexible polyimide substrate with graphene oxide gate dielectric is simulated. Soft saturation region is observed in the output characteristics. Good carrier mobility for both electron and hole are obtained by considering three prominent scattering mechanisms like columbic, phonon and surface roughness. Ambipolar nature of carrier conduction in graphene sheet is found with respect to position. It has been shown that significant increase in intrinsic cutoff frequency is possible by reducing the gate length up to 0.25 μm.

## Acknowledgment

The authors would like to thank Frank Schwierz his for fruitful discussions on the result of simulation.

## References

[1] Sasha Stankovich, Dmitriy A. Dikin, Richard D. Piner, Kevin A. Kohlhaas, Alfred Kleinhammes, Yuanyuan Jia, Yue Wu, SonBinh T. Nguyen, and Rodney S. Ruoff, "Synthesis of graphene-based nanosheets via chemical reduction of exfoliated graphite oxide," Carbon, vol. 45, pp. 1558–1565, June 2007.

[2] P. R. Wallace, "The band theory of graphite," Phys. Rev., vol. 71, pp. 622–634, 1947.

[3] Frank Schwierz,"Graphene transistors," Nature Nanotechnology, vol. 5, pp. 487–496, 2010.

[4] K. I. Bolotin, K. J. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. L. Stormer, "Ultrahigh electron mobility in suspended graphene," Solid State Communications, vol. 146, no. 9-10, pp. 351–355, 2008.

[5] Jeong Hun Mun and Byung Jin Cho, "Physical-gap-channel graphene field effect transistor with high on/off current ratio for digital logic applications," Appl. Phys. Lett., vol. 101, pp.143102, Oct. 2012.

[6] Zhenan Bao, Andrew J. Lovinger, and Ananth Dodabalapur, "Organic field-effect transistors with high mobility based on copper phthalocyanine," Appl. Phys. Lett., vol. 69, pp. 3066, Nov. 1996.

[7] Frank Schwierz, "Graphene transistors: status, prospects, and problems," Proceedings of the IEEE, vol. 101 , no. 7, pp. 1567 – 1584, July 2013.

[8] Beom Joon Kim, Houk Jang, Seoung-Ki Lee, Byung Hee Hong,Jong-Hyun Ahn, and Jeong Ho Cho, "High-performance flexible graphene field effect transistors with Ion Gel gate dielectrics," Nano Lett., vol.10 (9), pp 3464–3466, 2010.

[9] B. Fallahazad, K. Lee, G. Lian, S. Kim, C. M. Corbet, D. A. Ferrer, L. Colombo, and E. Tutuc, "Scaling of $Al_2O_3$ dielectric for graphene field-effect transistors," Appl. Phys. Lett. , vol.100, pp. 093112, 2012.

[10] Babak Fallahazad, Seyoung Kim, Luigi Colombo, and Emanuel Tutuc,"Dielectric thickness dependence of carrier mobility in graphene with HfO2 top dielectric," Appl. Phys. Lett., vol. 97, pp.123105, 2010.

[11] Seoung-Ki Lee, Ho Young Jang, Sukjae Jang, Euiyoung Choi, Byung Hee Hong, Jaichan Lee,*,Sungho Park and Jong-Hyun Ahn, "All graphene-based thin film transistors on flexible plastic substrates," Nano Lett., 2012, vol. 12 (7), pp 3472–3476.

[12] Nicholas Petrone, Inanc Meric, James Hone, and Kenneth L. Shepard, "Graphene field-e□ect-transistors with gigahertz-frequency power gain on flexible substrates." Nano Lett., vol. 13(1), pp. 121-125, 2013.

[13] Wenjuan Zhu, Damon B. Farmer, Keith A. Jenkins, Bruce Ek, Satoshi Oida, Xuesong Li, Jim Bucchignano, Simon Dawes, Elizabeth A. Duch, and Phaedon Avouris, "Graphene radio frequency devices on flexible substrate," APPLIED PHYSICS LETTERS, vol. 102, pp. 233102 ,2013.

[14] S. A. Thiele, J. A. Schaefer, and F. Schwierz, "Modeling of graphene metal-oxide-semiconductor field-effect transistors with gapless large-area graphene channels," JOURNAL OF APPLIED PHYSICS, vol. 107, pp. 094505, May 2010.

[15] Osama M. Nayfeh, “Graphene transistors on mechanically flexible polyimide incorporating atomic-layer-deposited gate dielectric,” Electron Device Letters, IEEE, Vol.32, no.10, pp. 1349-1351, October 2011.

[16] Lei Liao, Jingwei Bai, Rui Cheng, Yungchen Lin, Shan Jiang, Yongquan Qu, Yu Huang,, and Xiangfeng Duan, “Sub-100 nm channel length graphene transistors,” Nano Lett., vol. 10 (10), pp 3952–3956, 2010.

[17] Inanc Meric, Cory R. Dean, Andrea F. Young, Natalia Baklitskaya, Noah J. Tremblay, Colin Nuckolls, Philip Kim, and Kenneth L. Shepard, “Channel length scaling in graphene field-effect transistors studied with pulsed current-voltage measurements,” Nano Lett., vol. 11 (3), pp 1093–1097, 2011.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Vacancy Induced Phonon Properties of Hydrogen Passivated Graphene

Md. Sherajul Islam[1,*], Md. Tawabur Rahman [2], A. G. Bhuiyan [2], and Akihiro Hashimoto [1]
[1]University of Fukui, Japan
[2]Khulna Unversity of Engineering and Technonology, Bangladesh
[*] E-mail: sheraj_ruet@yahoo.com

*Abstract*—**The phonon properties of hydrogen passivated graphene with vacancy defects are studied using the forced vibrational method. The phonon density of states (PDOS) and typical mode patterns are calculated over a broad range of vacancies. We find that phonon properties of graphene strongly depend on the system size. We observe a broadening and softening of the PDOS peaks with the increase of vacancy concentrations. We find an increasing C-H stretching mode with the increase of defect density. Our numerical experiments reveal that the typical mode pattern for *K* point in-plane TO modes phonon show the spatial localized vibrations persuaded by vacancy defects, which are in conceptually good agreement with the experimental results of the large D band peak of the Raman spectra comes from the imperfections of crystal. The typical displacement pattern for C-H stretching mode shows a random displacement of H atoms in contrast to C atoms. Our simulation results show the significant impact of vacancy defects on the vibrational properties of graphene.**

***Keywords—Lattice vibration, graphene, vacancy defects, phonon density of states, hydrogen passivation.***

## I. INTRODUCTION

During the past few years, graphene and its related-materials have been in the scientific limelight due to their fascinating properties [1]-[3] and promising potential applications [4]-[7]. For any new material, the characterization and investigation of the fundamental and application oriented properties are a crucial prerequisite for tomorrow's applications. The study of fundamental properties such as vibrational modes in graphene-related materials have attracted considerable theoretical [8], [9] and experimental [10]-[12] impetus in the past few years. Vibration modes and frequencies are the basis for understanding the thermal, transport, and optical properties of solids. Graphene has been considered as a prototypical bench-top relativistic quantum system consist of a one-atom-thick planar sheet of $sp^2$-bonded carbon atom [2]. In $sp^2$ carbons, the phonons, like the electrons, depend on the atomic structure. As for most other materials, defects are unavoidable during the preparation of graphene. The disorder in $sp^2$ hybridized carbon systems leads to rich and intriguing phonon properties. For example, the effect of breaking the translational symmetry of crystal by introducing disorder into the lattice is the breakdown of momentum conservation, through the activation of phonons at interior *K* points of the Brillouin zone. Therefore, a strong D and features are observed in the Raman spectra which are generally inactive for perfect graphene.

The chemical modification of graphene is being the focus of another interest [13], [14]. Radicals such as oxygen, hydrogen, or fluorine atoms can be adsorbed on the surface of graphene. The adsorbed radicals can attach to the graphene layer in a random way, as is the case in graphene oxide [15] or they can form ordered patterns. It is challenging to estimate the vibrational properties of defective graphene, especially with complication due to chemical groups or dangling bonds at defective edge. Moreover, due to a strong electron phonon coupling in graphene-related materials [16] transport electrons have a significant effect on the lattice vibrations. The study of the vibrational properties of graphene with hydrogen passivation is thus of fundamental importance for the electron transport in electronic devices and of great general interest for the physical understanding of those structures.

In contrast to the intensive research performed on the phonon properties of pristine graphene [8]-[12], the phonon properties of defective or chemically functionalized graphene has not been well studied. When vacancy or other type of defects present in a system, they may induce the symmetry breakdown of elemental topological arrangements, which generate more complex lattice structures. Therefore, the dynamical matrix method requires huge computational resources. These long computational times and convergence problems in the dynamical matrix calculations limit the systems of interest to benchmark molecules. However, a significantly larger scale model is essential for an in-depth understanding of vibrational properties of defective atomic structures. It is thus very important to build a reliable model to measure the phonon properties of the defective graphene directly.

In this study, we have investigated systematically the phonon properties of graphene including vacancy defects and hydrogen passivation effects. We use the forced vibrational method, which is based on the mechanical resonance to extract the vibrational eigenmodes for very complex and large systems by numerical calculation to compute the phonon density of states. When defects present in a system, there may appear vibrational modes [17], [18] lying outside of the allowed frequency range of the perfect crystal. These are called localized vibrational modes or local modes because the mode energy is spatially concentrated near the defect site, which is similar to Anderson's localization for electronic wave functions

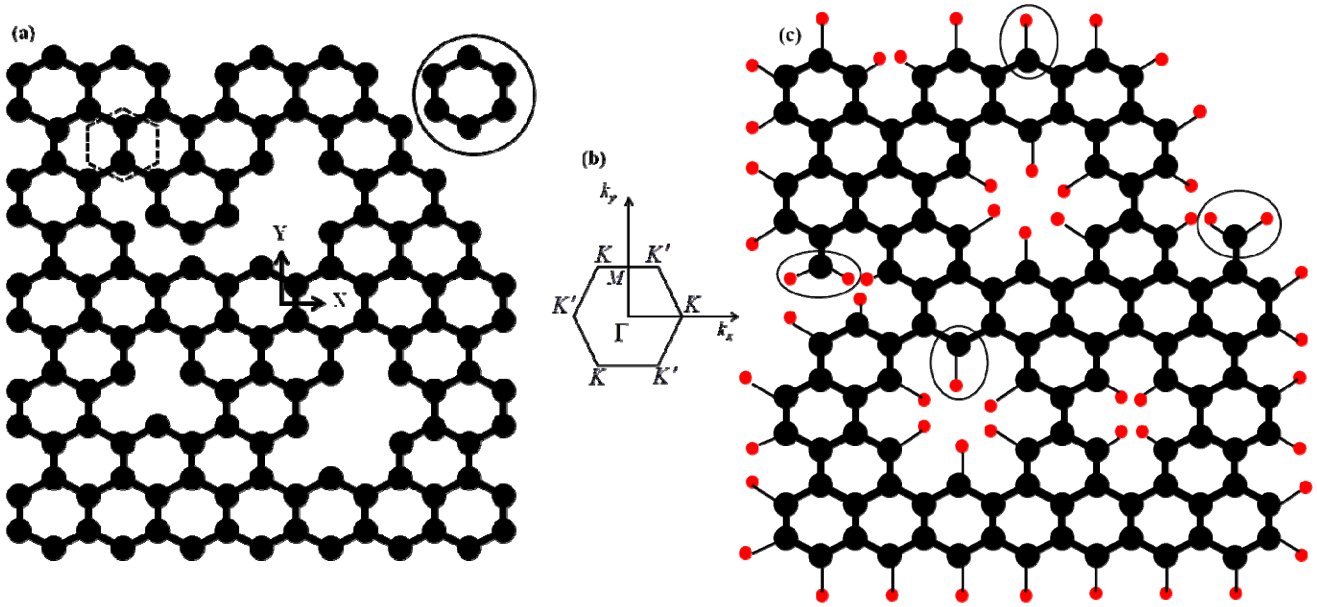



Figure 1. (a) Defective graphene with percolation network structures. (b) The Brillouin Zone of graphene with high symmetry points (c) H-passivated defective graphene. Black circles are C atoms and red circles are H atoms

in the disordered lattice. To show the nature of phonon states, we calculate the typical mode patterns in the presence of defects. We focus particularly on the in-plane optical phonon modes because of their importance in the Raman D band, the dominant feature in the Raman spectra. This information can be useful for the interpretation of future experiments on infrared, Raman, and neutron diffraction spectra of this defective material, as well as in the study of a wide variety of other physical properties such as specific heat, thermal expansion, heat conduction and electron phonon interaction.

## II. COMPUTATIONAL MODEL AND METHOD

The two dimensional honeycomb lattice structure of graphene introducing vacancy defects is shown in Fig. 1(a). The dashed hexagon indicates the unit cell of graphene. The Brillouin zone of the graphene with high-symmetry points such as $\Gamma$, $K$, and $M$ points are shown in Fig. 1 (b). The simulations are performed for a larger scaled systems with $N$=10, 500 atoms. The vacancies were introduced randomly into the graphene honeycomb lattice using site percolation procedure. It is well known that percolation threshold of site percolation network of honeycomb lattice is about 69%. Therefore, vacancy concentrations (i.e., the defect density) up to 30% are used in the present simulation. After the formation of vacancy defects in the networks, all dangling bonds are terminated by hydrogen atoms which is shown in Fig. 1(c). The force constant tensors up to the fourth nearest neighbors are used and taken from Jishi et al. [19]. Calculations reported in this paper have been carried out using the forced vibrational method. Here, we briefly elucidate the forced vibrational method relevant to treat physical structures with much large size, in which we obtain expressions that will be useful for the calculation phonon density of states and mode patterns.

The forced vibrational method introduced by Williams and Maris [20] is a numerical simulation technique based on mechanical resonance to extract eigenmodes. According to their formalism, let us consider a lattice consisting of $N$ atoms which are coupled together by linear springs. In general, the equation of motion of the systems with the scalar displacement of the $l$ th mass, $u_l(t)$, is

$$M_l \ddot{u}_l(t) + \sum_{l'} \phi_{ll'} u_{l'}(t) = 0 \qquad (1)$$

where $M_l$ is the mass of $l$ the atom and $\phi_{ll'}$ is the strength of the spring between the $l$ th and $l'$ th atoms. The FV method is based on the fact that if a periodic external force with frequency $\Omega$ is applied to the system, the response of the system will be dominated by the eigenmodes with the eigen frequency near to the $\Omega$. At time $t$ = 0, we prepare the lattice system with each atom at rest and with zero displacement. For $t > 0$ an external periodic force $F_l = F_0 \sqrt{M_l} \cos(\varphi_l) \cos(\Omega t)$ is applied to the each atom $l$, where $F_0$ is the independent of time and $M_l$ is the mass of $l$th atom. $\varphi_l$ is a random quantity. After a long time (i.e. for many cycles), the total energy of the system is

$$\langle E \rangle = \frac{F_0^2 \pi t}{8} \sum_{\lambda} \delta(\omega_\lambda - \Omega) = \frac{F_0^2 \pi t\, N\, g(\Omega)}{8} \qquad (2)$$

where $g(\Omega)$ is the phonon density of states. Therefore,

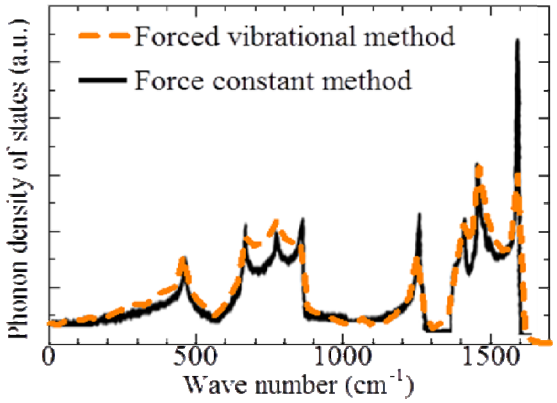



Figure 2. Phonon Density of states of pure grarphene in comparison to DOS by force constant method [19]

$$g(\Omega) = \frac{8\langle E \rangle}{\pi t F_0^2 N} \tag{3}$$

Thus, we can obtain the phonon density of states from the carrying out the time development in the presence of the periodic force. Mode patterns can also be obtained by applying iteratively the external periodic force proportional to the displacements of atoms at each step.

## III. Results and Discussion

Fig. 2 shows the calculated phonon density of states (PDOS) of the ideal graphene in the total of 10,500 lattice spaces. For comparison, we have also added the PDOS calculated by the force constant method [19]. The FV method reproduces the exact results of PDOS as can be found by the force constant method. The phonon properties greatly depend on the system size of the graphene. The PDOS of the perfect graphene in the total of 368 and 105,00 lattice spaces is shown in Fig. 3. It is well known that the exact PDOS of infinite two dimensional lattices just agree with the PDOS for large two dimensional finite lattices calculated by the FV method. Therefore, the violet line in Fig. 3 corresponds to the exact PDOS of the 2D honeycomb graphene lattice. As shown in Fig. 3, the PDOS of the 368 lattice space is considerable agree with the exact PDOS of the large two dimensional honeycomb lattice except for the large fluctuations. This result indicates that the fluctuation of the DOS in the case of finite lattice is caused by the rather small number of the carbon atoms.

Fig. 4 shows the PDOS of the defective graphene with 10%, 20% and 30% vacancy concentrations. The PDOS for perfect graphene is also shown in the inset of the Fig. 4. The PDOS of perfect graphene exhibits all the characteristic peaks that correspond to the $sp^2$ bonded carbon honeycomb lattice by force constant method [19]. However, the PDOS for the defective graphene strongly depend on the defect density. With the increase of vacancy concentration, we observe a broadening and softening of the PDOS peaks in the frequency range between 1591cm$^{-1}$ (which is the graphene $E_{2g}$ mode peaks calculated by FV method) to 1500 cm$^{-1}$. More importantly, we see that for vacancy concentrations of 10% and

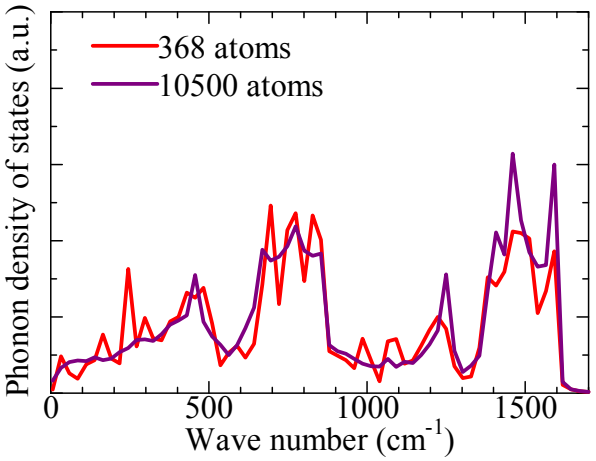



Figure 3. Phonon Density of states for the pure grarphene in the total of 368 and 10500 atoms

higher, the $E_{2g}$ mode peak has been reduced into a shoulder or it has completely disappeared. This disappearance of the peaks at high vacancy concentrations implies the defect-induced collapse of the long-range order in graphene at vacancy concentrations beyond a critical level. Furthermore, there may some fine peaks appear in PDOS curves at low frequency region with the increase of defect densities as shown in Fig. 4. The shift of the PDOS towards the low frequency region are caused by the increase in the numbers of the unsaturated C atom with some dangling bonds because the increase of the unsaturated C atoms in the networks cause to decrease of the density of high frequency modes.

The PDOS in the cases of all C dangling bonds are terminated by the H atoms for the defective graphene is shown in Fig. 5. In the case of H terminations, there appears high and low frequency region in the PDOS curves. The PDOS in the low frequency regions are correspond to the PDOS in the pristine graphene without H terminations as shown in Fig. 3 and 4. The High frequency modes near the 2900 cm$^{-1}$ are the localized vibrational modes which are dominantly as H modes,as can be expected from the C-H stretching modes.

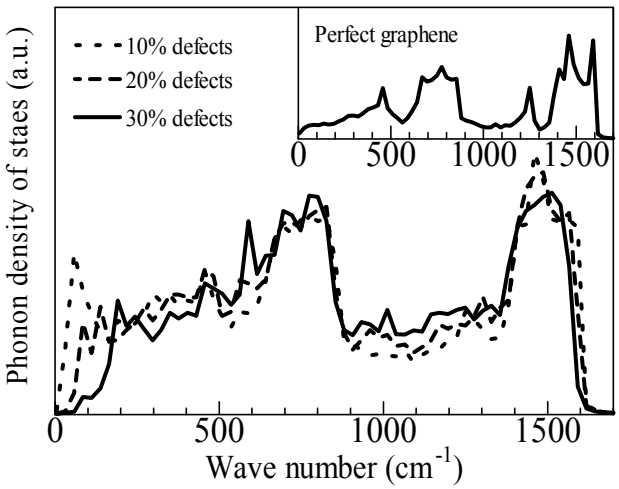



Figure 4. Phonon Density of states of graphene with different defect density

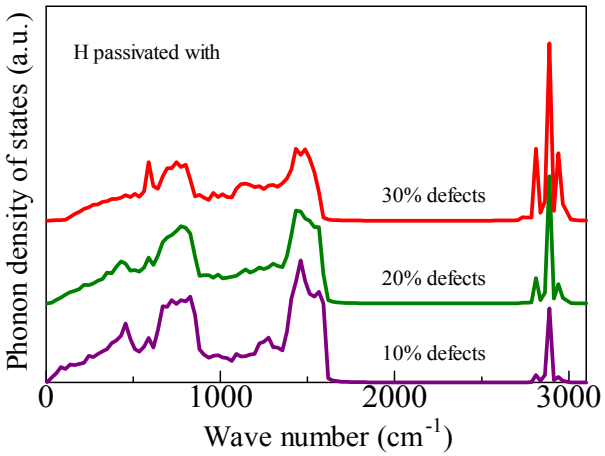



Figure 5. Phonon Density of states of H-passivated graphene with different vacancy concentrations

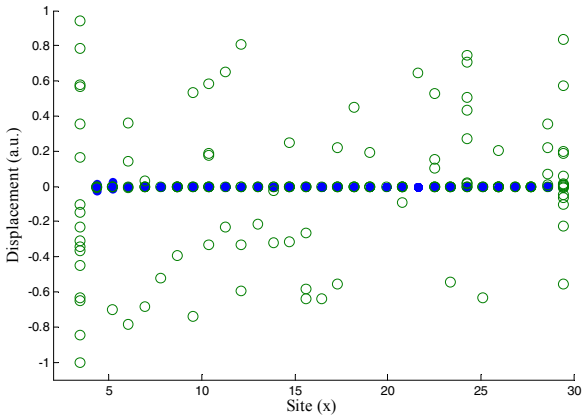
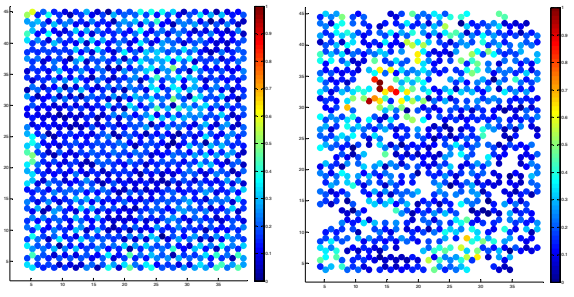



Figure 6. Dispalcement pattern for C-H stretching mode of H-passivated graphene with 20% vacancy concentrations

With the increase of defect density the peak in the high frequency region increases. The shift of the PDOS towards the high frequency region are caused by the increase of H atoms. Fig. 6 shows the typical displacement pattern for C-H stretching modes at 2900 $cm^{-1}$. The vertical and horizontal axes mean the displacement of the atoms and the site number of each atom, respectively. The filled circles stand for the displacement of C atoms and open circles stand for for the displacement H atoms. It is easily seen that the H atoms are large and randomly displaced from their equilibrium positions in contrast to C atoms. The random displacements of the H atoms come from the lack of any H-H interactions.

The propagation of phonon is restricted into a particular region of the solid due to the defect formations. Vacancy defects induce the symmetry breakdown of the lattice, and the wave vector is no longer a good quantum number. In this case, phonons are scattered into other phonon states and thus, it is expected that phonons are localized in the real space. In this paper, we mainly focus on the *K* point of the in-plane optical modes (iTO branch) phonon of the eigenmodes due to the dominant effect of these modes on the large D band features of the Raman spectra especially for defective states of the sample.



(a) (b)

Figure 7. The typical mode pattern of graphene at 1355 $cm^{-1}$ for (a) perfect and, (b) 20% vacancy defects. Each circle denotes an atom, the colors denote the displacement, which are linearly normalized by the maximum value.

The calculated typical mode patterns for *K* point iTO mode phonon of perfect and defective graphene are shown in Fig. 7 (a) and 7 (b), respectively. In both figures each circle denotes an atom, and the color denotes the displacement. The darker red colors indicate the larger displaced atoms. Here, we have shown the mode pattern for 1148 atoms only. As the modes are strongly localized within this range of lattice spaces, there is no need to study of the system with large number of lattice spaces. We can clearly observe in Fig. 7 (a) that all the modes are extended or no localized modes are found in perfect honeycomb lattice structure of graphene. However, if we introduce vacancies into the perfect structure, we observe a spatially localized eigenmodes. The carbon lattice vibrations in the percolation networks of graphene are strongly localized as shown in Fig 7 (b). Such localization effect on the lattice vibrations may be due to a random atomic arrangement in the percolation networks.

## IV. CONCLUSIONS

In conclusions, we present the results of a systematic analysis of defect induced phonon properties of H-passivated graphene. The forced vibrational method which is based on the mechanical response to extract the vibrational eigenmodes has been applied to calculate the phonon density of states and typical mode patterns of defective graphene. Our results show that the phonon properties of graphene strongly depend on the system size and defect density. The lattice vibrations in the defective graphene show the remarkable increase of the PDOS in the low frequency region induced by their defect formations. We find an increasing C-H stretching mode with the increase of defect density. Our calculated typical mode pattern show that *K* point in-plane TO modes phonon are strongly localized due to the presence of vacancy defects. The mode pattern for H-passivated graphene show a random displacement of H atoms from their equilibrium positions. This information can be useful for the interpretation of future experiments on infrared, Raman, and neutron diffraction spectra of these defective materials, as well as in the study of a wide variety of other physical properties such as specific heat, thermal expansion, heat conduction and electron phonon.

## ACKNOWLEDGMENT

This work is supported by a Grant-in-Aid of Basic Research (A) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

## REFERENCES

[1] A. K. Geim, K. S. Novoselov, "The rise of graphene," *Nature Mater.*, vol. 6, pp. 183-191, Mar. 2007.

[2] A. K. Geim, "Graphene: Status and prospects," *Science*, vol. 324, no. 5934, pp. 1530-1534, Jun. 2009.

[3] S. Mayorov, R. V. Gorbachev, S. V. Morozov, L. Britnell, R. Jalil, L. A. Ponomarenko, P. Blake, K. S. Novoselov, K. Watanabe, T. Taniguchi, A. K. Geim, "Micrometer-Scale Ballistic Transport in Encapsulated Graphene at Room Temperature," *Nano Lett.*, vol. 11, pp. 2396-2399, May 2011.

[4] Y. M. Lin, C. Dimitrakopoulos, K. A. Jenkins, D. B. Farmer, H. Y. Chiu, A. Grill, P. Avouris, "100-Ghz Transistors from Wafer-Scale Epitaxial Graphene," *Science*, vol. 327, no. 5966, pp. 662, Feb. 2010.

[5] K. Kim, J. Y. Choi, T. Kim, S. H. Cho, H. J. Chung, "A role for graphene in silicon-based semiconductor devices," *Nature*, vol. 479, no. 7373, pp. 338-344, Nov. 2011.

[6] F. N. Xia, T. Mueller, Y. M. Lin, A. Valdes-Garcia, P. Avouris, "Ultrafast graphene photodetector," *Nature Nano Technol*. vol. 4, no. 12, pp. 839-843, 2009.

[7] F. Rana, "Graphene Terahertz Plasmon Oscillators," *IEEE Trans. Nanotechnol*. vol. 7, no. 1, pp. 91-99, Jan. 2008.

[8] J. Yan, W. Y. Ruan, M. Y. Chou, "Phonon dispersions and vibrational properties of monolayer, bilayer, and trilayer graphene: Density-functional perturbation theory," *Phys. Rev. B*, vol. 77, pp. 125401, Mar. 2008.

[9] S. Piscanec, M. Lazzeri, F. Mauri, A. Ferrari, J. Robertson, "Kohn anomalies and electron-phonon interactions in graphite," *Phys. Rev. Lett*., vol. 93, no. 18, pp. 185503, Oct. 2004.

[10] A. Gruneis, J. Serrano, A. Bosak, M. Lazzeri, S.L. Molodtsov, L. Wirtz, C. Attaccalite, M. Krisch, A. Rubio, F. Mauri, T. Pichler, "Phonon surface mapping of graphite: disentangling quasi-degenerate phonon dispersions," *Phys. Rev. B*, vol. 80, pp. 085423, Apr. 2009.

[11] J. Maultzsch, S. Reich, C. Thomsen, H. Requardt, "Phonon Dispersion in Graphite," *Phys. Rev. Let*t., vol. 92, no. 7, pp. 075501, Feb. 2004.

[12] M. Mohr, J. Maultzsch, E. Dobardzic, S. Reich, I. Milosevic, M. Damnjanovic, A. Bosak, M. Krisch, C. Thomsen, "Phonon dispersion of graphite by inelastic x-ray scattering," *Phys. Rev. B,* vol. 76, pp. 035439, Jul. 2007.

[13] S. H. Cheng, K. Zou, F. Okino, H. R. Gutierrez, A. Gupta, N. Shen, P. C.Eklund, J. O. Sofo, And J. Zhu, "Reversible fluorination of graphene: Evidence of a two-dimensional wide bandgap semiconductor," *Phys. Rev. B*, vol. 81, no. 20, pp. 205435 , May 2010.

[14] F. Withers, M. Dubois, And A. K. Savchenko, "Electron properties of fluorinated single-layer graphene transistors," *Phys. Rev. B*, vol. 82, no. 7, pp. 073403, Aug. 2010.

[15] A. Dikin, S. Stankovich, E. J. Zimney, R. D. Piner, G. H. B. Dommett, G. Evmenenko, S. T. Nguyen, And R. S. Ruoff, "Preparation and characterization of graphene oxide paper," *Nature*, London, vol. 448, no. 7152, pp. 457-460, Jul. 2007.

[16] S. Pisana, M. Lazzeri, C. Casiraghi, K.S. Novoselov, A.K. Geim, A.C. Ferrari, F. Mauri, "Graphene calling," *Nat. Mater.* vol. 6, no. 3, pp. 198-201, Mar. 2007.

[17] A. A. Maradudin, E. W. Montroll, G. H. Weiss, *Theory of lattice dynamics in the harmonic approximation*, Academic Press, New York, 1963.

[18] A. A. Maradudin, Ed., *Theoritical and experimental aspects of the effects of point defects and disorder on the vibrations of crystal*, Solid State Physics. New York, USA: Academic Press, 1966. vol. 18.

[19] R. A. Jishi, L. Venkataraman, M. S. Dresselhaus, G. Dresselhaus, "Phonon modes in carbon nanotubules," *Chem. Phys. Lett*., vol. 209, no. 1-2, pp. 77-82, Jun. 1993.

[20] M. L. Williams, H. J. Maris, "Numerical study of phonon localization in disordered systems ," *Phys. Rev. B*, vol. 31, no. 7, pp. 4508-4515, Apr. 1985.

# Linear Asymmetric Pocket Profile Based Low Frequency Drain Current Flicker Noise Model for Pocket Implanted Nano Scale n-MOSFET

Muhibul Haque Bhuyan[1,*], and Quazi Deen Mohd Khosru[2]
[1]Department of Electrical and Electronic Engineering
Green University of Bangladesh, Dhaka, Bangladesh
[2]Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
[*]E-mail: muhibulhb@gmail.com

*Abstract*—**This paper presents an analytical drain current flicker noise model for the asymmetric pocket implanted nano scale n-MOSFET. The model is developed by assuming asymmetric linear pocket doping profile at the source edge only. The number of channel charges is found for the two regions and are incorporated in the unified flicker noise model developed by Hung et al. for the conventional metal oxide semiconductor field effect transistor (MOSFET). Simulation results for the various device as well as pocket profile parameters show that the derived drain current flicker noise model has a simple compact form that can be utilized to study and characterize the pocket implanted advanced ULSI devices.**

***Keywords—Linear Asymmetric Pocket Profile, n-MOSFET, Threshold Voltage, Flicker Noise, Drain Current.***

## I. INTRODUCTION

As the channel length of MOSFETs is scaled down to deep sub-micrometer or sub-100 nm regime, we observe the short channel effect (SCE) [1]. This effect can be reduced or even be reversed (then it is called reverse short channel effect or RSCE) by locally raising the channel doping near source and/or drain junctions. Lateral channel engineering utilizing halo or pocket implant [2-6] surrounding the source and drain regions is effective in suppressing SCE. In fact, this pocket implant technology is found to be very promising in the effort to tailor the short-channel performances of deep-submicron as well as sub-100 nm MOSFETs [3]. However, this pocket implantation can be either symmetric [3, 6-7] or asymmetric [8-9].

The increased doping concentration at the drain side increases the electric field strength and thus hot carrier reliability problems become more serious [10] and reduces the voltage gain. On the other hand, drive current improves in asymmetric devices with noise trade-offs [11].

In the low frequency region, flicker noise is dominant. Flicker noise affects the circuit performances. Hence it should be studied and modeled properly to reduce it. Already few papers have been published focusing on the degradation of drain current flicker noise due to pocket implantation in MOSFETs [12-18]. New pocket structures, such as, single pocket, asymmetric channel structure, [12, 14] and epitaxial channel MOSFETs [15, 16], were proposed to reduce the drain current flicker noise by elimination of pocket implantation. The low frequency noise in pocket implanted MOSFETs may result from additional oxide trap creation due to pocket implantation [16], but this was also not supported by the experiment [17]. In [17], it was shown that the non-uniform distribution of threshold voltage along the channel resulting from the pocket implantation is responsible for the low frequency drain current flicker noise degradation, but there step doping profiles are used in the pocket implanted region to model the drain current flicker noise. In this paper, an analytical drain current flicker noise model is developed taking into account the pocket doping effect. The model is developed by using asymmetric linear pocket doping profile only at the source side of the device along the channel. The pocket profile and device parameters as well as bias voltages are varied to investigate the asymmetric pocket implantation effect on low frequency drain current flicker noise in the nano scale n-MOSFET.

## II. POCKET DOPING PROFILE

The asymmetric pocket implanted at the source side n-MOSFET structure, shown in Fig. 1, is considered for this work. The assumed co-ordinate system is shown at the right side of the structure.

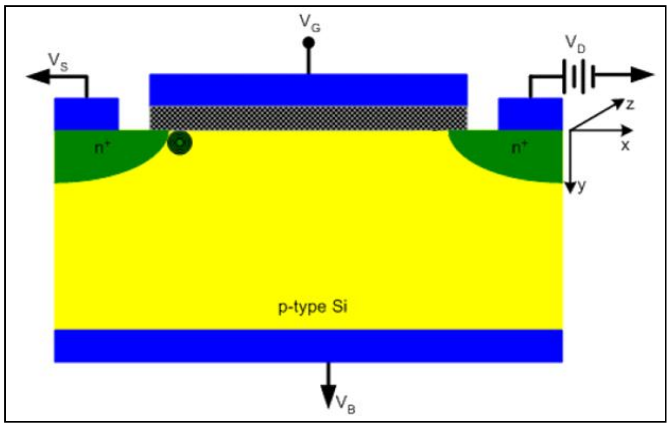



Figure 1. Pocket implanted n-MOSFET structure with asymmetric pocket at the source side only.

All the device dimensions are measured from the oxide-silicon interface. In the structure, the junction depth ($r_j$) is 25 nm. The oxide thickness ($t_{ox}$) is 2.5 nm, and it is $SiO_2$ with fixed oxide charge density of $10^{11}$ cm$^{-2}$. Uniformly doped p-type Si substrate is used with substrate doping concentration, $N_{sub} = 4.5\times10^{17}$ cm$^{-3}$, peak pocket doping concentration at the source side, $N_{pm} = 2.5\times10^{18}$ cm$^{-3}$, pocket length, $L_p = 25$ nm and source/drain doping concentration, $N_{sd} = 9.0\times10^{20}$ cm$^{-3}$.

To preserve the long channel threshold voltage behavior for the short channel device, pocket implantation, which causes the reverse short channel effect (RSCE), is done by adding donor atoms at the source edge. The peak pocket doping concentration ($N_{pm}$) gradually decreases, from the source side to the drain side along the channel, to the substrate level concentration ($N_{sub}$) with a pocket length ($L_p$). The basis of the model of the pocket doping is to assume a linear pocket doping profile across the channel as shown in Figs. 2-3. The pocket parameters, $N_{pm}$ and $L_p$, play important role in determining the RSCE.

At the source side, the pocket profile is given as:

$$N_s(x) = -\frac{N_{pm} - N_{sub}}{L_p}x + N_{pm}$$

$$\text{or, } N_s(x) = N_{sub}\frac{x}{L_p} + N_{pm}\left(1 - \frac{1}{L_p}x\right) \qquad (1)$$

,where $x$ represents the distance across the channel.

Since the pocket profile is due to the direct pocket implantation at the source side only, it is an asymmetric device. This conceptual pocket profile is integrated mathematically along the channel length ($L$) and then divided by $L$ to derive an effective concentration as in equation (2).

$$N_{eff} = \frac{1}{L}\int_0^L \left[N_s(x) + N_{sub}\right]dx$$

$$\text{or, } N_{eff} = \frac{1}{L}\int_0^{L_p} N_s(x)\,dx + \frac{1}{L}\int_{L_p}^{L} N_{sub}\,dx \qquad (2)$$

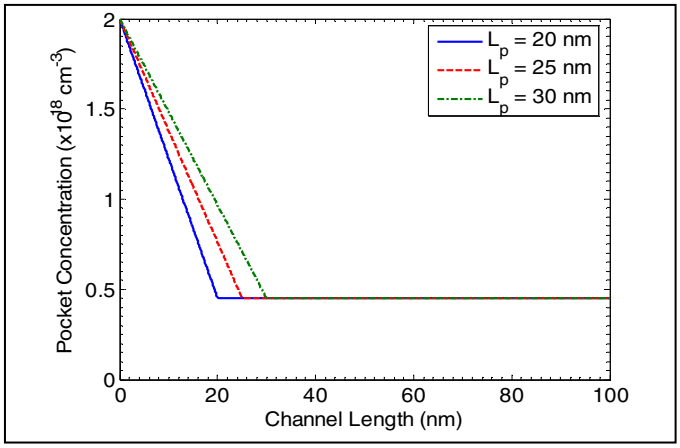



Figure 2. Simulated pocket profiles at the surface for different pocket lengths, $L_p = 20$, 25 and 30 nm, peak pocket doping concentration, $N_{pm} = 2.0\times10^{18}$ cm$^{-3}$.

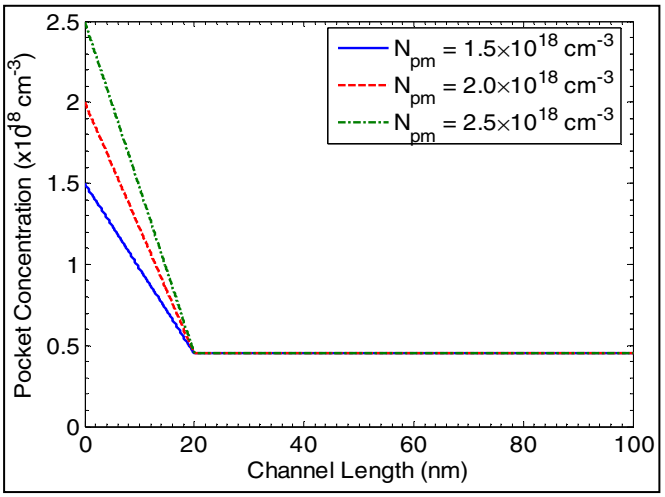



Figure 3. Simulated pocket profiles at the surface for various peak pocket concentrations, $N_{pm} = \{1.5, 2.0, 2.5\}\times10^{18}$ cm$^{-3}$, pocket length, $L_p = 20$ nm.

Putting the expression of $N_s(x)$ from equation (1) in to equation (2), the effective doping concentration is obtained as given in equation (3).

$$N_{eff} = N_{sub}\left(1 - \frac{L_p}{2L}\right) + \frac{N_{pm}L_p}{2L} \qquad (3)$$

When $L_p << L$ for long channel device then pocket profile has no effect on flicker noise, but when $L_p$ is comparable with $L$ then the pocket profile affects the flicker noise.

## III. Flicker Noise Model

According to the unified drain current flicker noise model [19], the normalized noise power spectrum density $\left(S_{id}/I_d^2\right)$ for the conventional MOS device is given by the analytic form at very low drain bias as in equation (4).

$$\frac{S_{id}}{I_d^2} = \frac{kT}{\gamma f W L^2}\int_0^L N_t\left(E_{fn}, x\right)\left[\frac{1}{N(x)} + \alpha\mu\right]^2 dx \qquad (4)$$

,where $\gamma = 10^8$ cm$^{-1}$ is the attenuation coefficient of the electron wave function in the oxide, $\alpha$ is the scattering coefficient [17], $N(x)$ is the number of channel carriers per unit area, and $N_t(E_{fn}, x)$ is the oxide trap density at the electron Fermi level $E_{fn}$.

At a very low gate overdrive bias, the mobility fluctuation term $\alpha\mu$ in equation (4) is smaller than $1/N(x)$ term. Thus the carrier density is uniform along the channel and is given by

$$qN = C_{ox}\left(V_{gs} - V_{th}\right) \qquad (5)$$

,where $V_{gs}$ is the gate bias and $V_{th}$ is the threshold voltage for the asymmetric pocket n-MOSFET as in equation (6) [20] and $C_{ox}$ is the oxide capacitance per unit area.

$$V_{th} = V_{th,L} + \gamma_B\left[\frac{N_{sub}}{N_{eff}}(2\varphi_F) - V_{BS}\right]^{\frac{1}{2}} - \gamma_A\frac{N_{sub}}{N_{eff}}(2\varphi_F)^{\frac{1}{2}} - \ldots$$

$$\ldots\frac{6t_{ox}}{d_1}\left[2(\varphi_{bi}-V_{BS})+V_{DS}\right]\exp\left(-\frac{\pi L}{4d_1}\right) \quad (6)$$

,where $V_{th,L}$ is the long channel threshold voltage, the second and the third parts include the threshold voltage due to both the substrate bias and effective doping concentration effects, the fourth part incorporates the drain bias and substrate bias as well as the short channel effects [21-22].

Since the oxide (interface) trap density is not affected by the pocket implantation process, it is assumed that the oxide trap density along the channel is uniform. From Fig. 1, it is evident that the channel region is divided into two distinct regions. The first region is the pocket implanted and the rest of the region is the uniformly doped substrate region. Hence, the noise model is also divided into two parts as in equation (7).

$$\frac{S_{id}}{I_d^2}=\frac{kT}{\gamma fWL^2}N_t\left[\int_0^{L_p}\frac{1}{N_s^{\,2}(x)}dx+\int_{L_p}^{L}\frac{1}{N_{sub}^{\,2}}dx\right] \quad (7)$$

Using equations (1), (3), (5) and (6), equation (7) can be written as in equation (8), which is our desired model.

$$\frac{S_{id}}{I_d^2}=\frac{kTq^2}{\gamma fWL^2C_{ox}^2}N_t\left[\frac{L_p}{(V_{gs}-V_{thp})^2}+\frac{L-L_p}{(V_{gs}-V_{th})^2}\right] \quad (8)$$

,where $V_{thp}$ is the threshold voltage in the pocket region found using equation (6) by using the effective doping concentration ($N_{effp}$) of the source side pocket region instead of using the effective doping concentration ($N_{eff}$).

## IV. Results and Discussions

Fig. 4 shows threshold voltage variation with gate lengths for three different pocket doses. It is observed that as the peak pocket concentration increases the reverse short channel effect increases and delays the threshold voltage roll-off. Thus, the threshold voltage roll-up happens at longer channel length. That is, the higher pocket dose eliminates the effect of SCE.

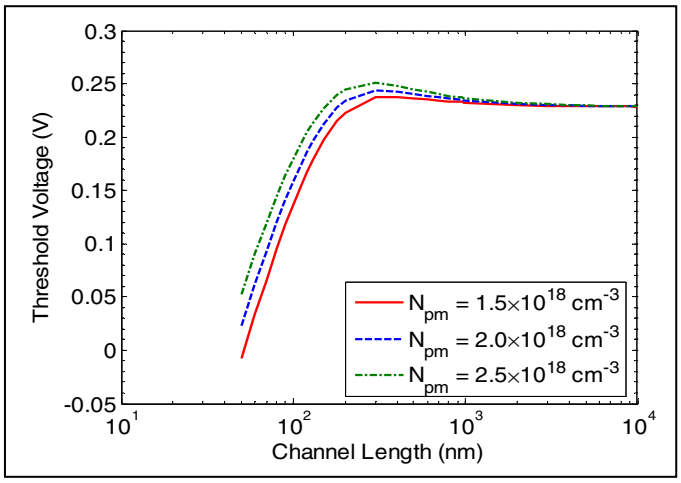



Figure 4. Threshold voltage vs. channel length curves for various peak pocket doping concentrations with $V_{DS}$ = 0.05 V and $V_{BS}$ = 0.0 V.

Fig. 5 shows the noise behavior for two different pocket lengths. Pocket lengths have little effect on noise behavior. Only for the gate bias lower than the threshold voltage, the noise increases when the pocket length is increased. But as shown in Fig. 6, the noise behavior is affected seriously by the peak pocket concentrations. The model is simulated for a very low frequency of 100 Hz and at a low drain bias of 0.2 V. As the gate bias is increased the noise is decreased. Also for the higher gate bias, the noise does not increased much for the increment of the pocket dose. It is also observed that for a particular gate bias noise behavior is worse for the higher pocket dose.

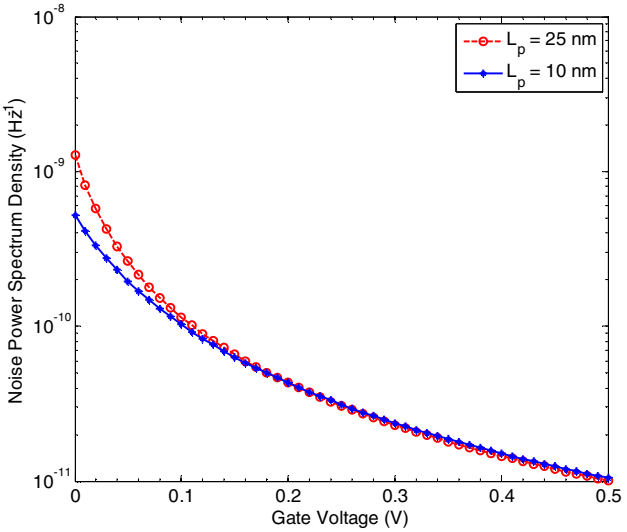



Figure 5. Noise power spectrum density vs. gate voltage for different pocket lengths with $L$ = 50 nm, $N_{pm}$ = 1.25×10$^{18}$ cm$^{-3}$, $V_{DS}$ = 0.2 V, $f$ = 100 Hz.

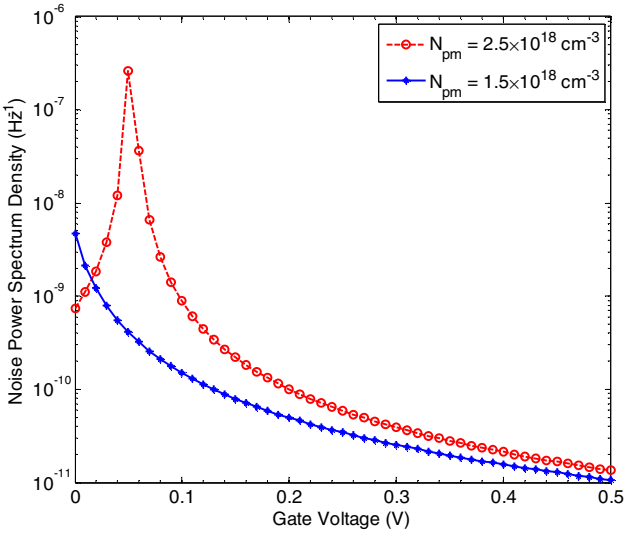



Figure 6. Noise power spectrum density vs. gate voltage for two pocket doses with $L$ = 50 nm, $L_p$ = 25 nm, $V_{DS}$ = 0.2 V, $f$ = 100 Hz.

Fig. 7 shows that as the drain bias is increased the noise is improved. At higher gate bias, the decrement of noise is smaller for the higher drain bias than that at the lower gate bias, but at lower gate bias as the drain bias is increased the

noise is improved at higher rate. As drain bias increases threshold voltage decreases and hence gate over drive voltage increases and finally the noise power spectrum density decreases as obtained from equation (8). This can be attributed to the increase of the lateral electric field with the increase of the drain bias thereby increasing the drift velocity of the electrons in the inversion channel and thus electrons get less time for thermal agitation in the channel.

Fig. 8 shows the noise power spectrum density variations with the gate length for two different pocket doses at gate and drain bias of 0.35 V (above threshold) and 2.0 V respectively. It is noted that the noise degradation is more significant in shorter channel length device as pocket implant region is comparable with the channel length of the device. For longer channel length, the noise is not changed with the increment of the pocket implantation dose.

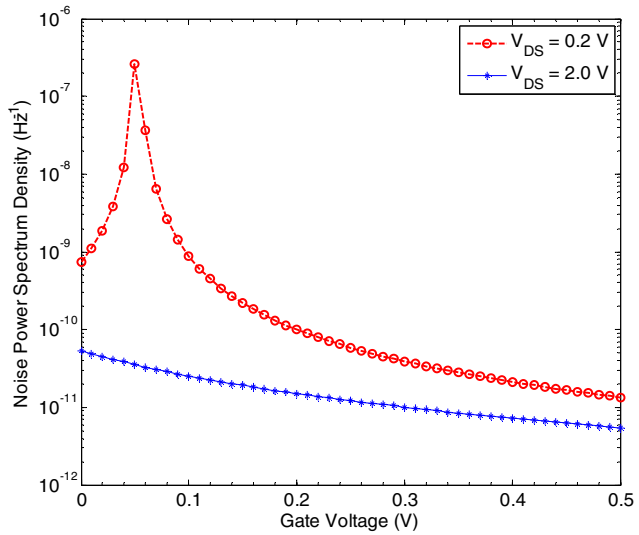



Figure 7. Noise power spectrum density vs. gate voltage for different drain biases with $L$ = 50 nm, $N_{pm}$ = 2.5×10$^{18}$ cm$^{-3}$, $L_p$ = 25 nm, $f$ = 100 Hz.

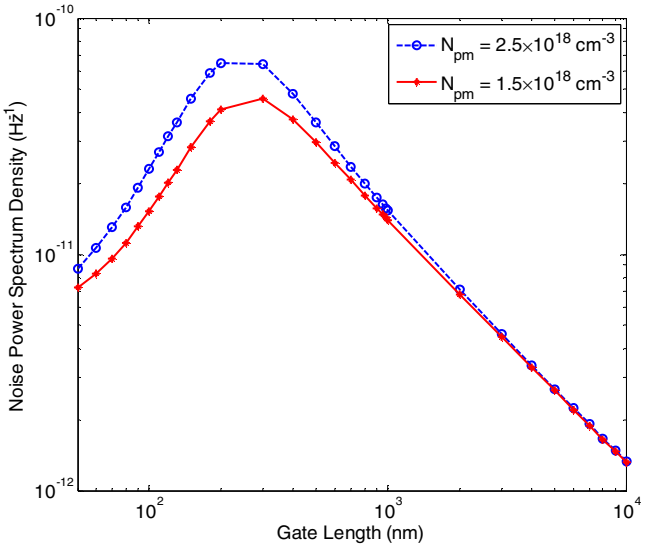



Figure 8. Noise power spectrum density vs. gate length for two pocket doses with $L_p$ = 25 nm, $V_{GS}$ = 0.35 V, $V_{DS}$ = 2.0 V, $f$ = 100 Hz.

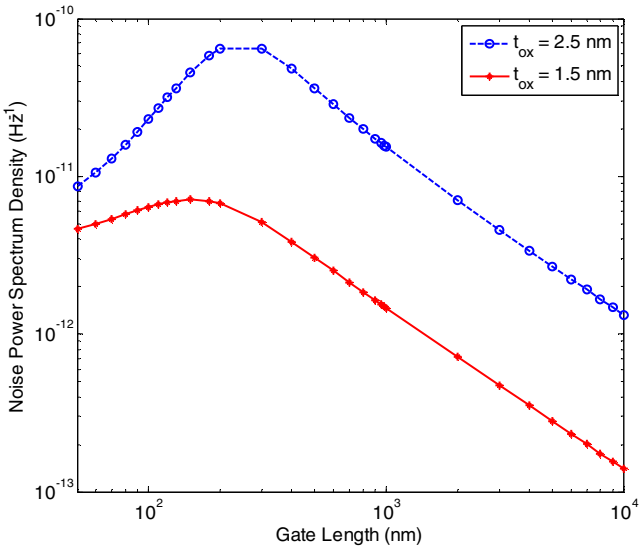



Figure 9. Noise power spectrum density vs. gate length for different oxide thicknesses with $N_{pm}$ = 2.5×10$^{18}$ cm$^{-3}$, $L_p$ = 25 nm, $V_{GS}$ = 0.35 V, $V_{DS}$ = 2.0 V and $f$ = 100 Hz.

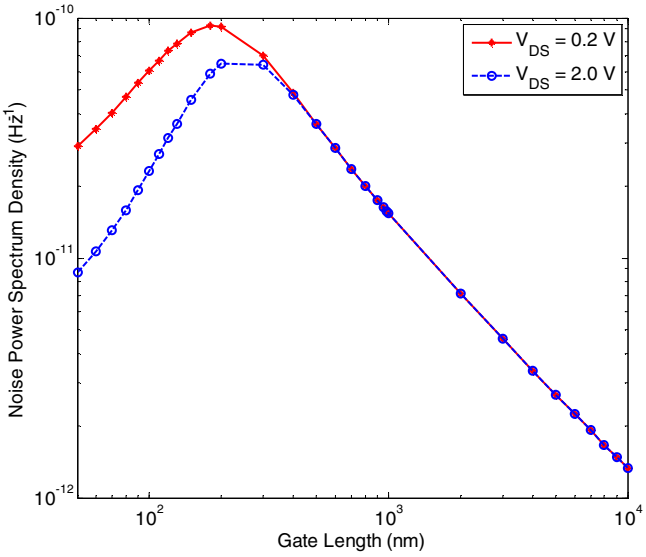



Figure 10. Noise power spectrum density vs. gate length for different drain biases with $N_{pm}$ = 2.5×10$^{18}$ cm$^{-3}$, $L_p$ = 25 nm, $V_{GS}$ = 0.35 V and $f$ = 100 Hz.

Fig. 9 shows the noise power spectrum density variations with the gate length for two different oxide thicknesses at gate and drain biases of 0.35 V (above threshold) and 2.0 V respectively. It is noted that the noise degradation is more significant when the oxide thickness is higher. Because when oxide thickness increases the oxide capacitance decreases as well as there are more interface trap charges in the oxide layer.

In Fig. 10, the noise behavior is shown with gate length for different drain biases. As the drain bias is increased for a particular gate length the noise behavior improves. At lower gate bias the noise decreases rapidly.

## V. CONCLUSIONS

A low frequency drain current flicker noise model for ultra thin oxide nano scale pocket implanted n-MOSFET has been developed based on the asymmetric linear pocket doping profile at the source edge of the device surface along the channel. The effects of changing the pocket profile and device parameters as well as bias voltages on the noise power spectrum density have been investigated through the analytical simulation of the proposed model. The simulated results show that the proposed model predicts the drain current flicker noise correctly for the channel length in the nano scale regime. Hence this model efficiently predicts the noise behavior of the asymmetric pocket implanted nano scale n-MOSFET.

## REFERENCES


[1] S. M. Sze, *Physics of Semiconductor Devices*, John Wiley & Sons, New York, USA, 1981.

[2] K. Y. Lim and X. Zhou, "Modeling of threshold voltage with non-uniform substrate doping," in *Proceedings of the IEEE International Conference on Semiconductor Electronics*, Malaysia, 1998, pp. 27-31.

[3] B. Yu, C. H. Wann, E. D. Nowak, K. Noda and C. Hu, "Short channel effect improved by lateral channel engineering in deep sub-micrometer MOSFETs," *IEEE Transactions on Electron Devices*, vol. 44, no. 4, pp. 627-633, Apr. 1997.

[4] B. Yu, H. Wang, O. Millic, Q. Xiang, W. Wang, J. X. An and M. R. Lin, "50 nm gate length CMOS transistor with super-halo: design, process and reliability," in *Proceedings of IEEE IEDM Technical Digest*, 1999, pp. 653-656.

[5] K. M. Cao, W. Liu, X. Jin, K. Vasant, K. Green, J. Krick, T. Vrotsos and C. Hu, "Modeling of pocket implanted MOSFETs for anomalous analog behavior," in *Proceedings of IEEE IEDM Technical Digest*, 1999, pp. 171-174.

[6] Y. S. Pang and J. R. Brews, "Models for subthreshold and above subthreshold currents in 0.1 μm pocket n-MOSFETs for low voltage applications," *IEEE Transactions on Electron Devices*, vol. 49, no. 5, pp. 832-839, May 2002.

[7] M. H. Bhuyan and Q. D. M. Khosru, "Linear pocket profile based threshold voltage model for sub-100 nm n-MOSFET," *International Journal of Electrical and Computer Engineering*, vol. 5, no. 5, pp. 310-315, May 2010.

[8] C. Bulucea, S. R. Bahi, W. D. French, Y. Jeng-Jiun, P. Francis, T. Harjono, V. Krishnamurthy, J. Tao and C. Parker, "Physics, technology and modeling of complimentary asymmetric MOSFETs," *IEEE Transactions on Electron Devices*, vol. 54, no. 11, pp. 2969-2974, Nov. 2007.

[9] J. P. Kim, W. Y. Choi, J. Y. Song, S. W. Kim, J. D. Lee and B.-G. Park, "Design and fabrication of asymmetric mosfets using a novel self-aligned structure," *IEEE Transactions on Electron Devices*, vol. 57, no. 10, pp. 2363-2380, Oct. 2010.

[10] A. Hokazono, S. Balasubramanian, K. Ishimaru, H. Ishiuchi, H. Chenming and L. T.-J. King, "MOSFET hot-carrier reliability improvement by forward-body bias," *IEEE Electron Device Letters*, vol. 27, no. 7, pp. 605-608, Jul. 2006.

[11] A. Cathignol, S. Bordez, A. Cros, K. Rochereau and G. Ghibaudo, "Abnormally high local electrical fluctuations in heavily pocket implanted bulk long MOSFET," *Solid State Electronics*, vol. 53, no. 2, pp. 127-133, Feb. 2009.

[12] A. Chatterjee, K. Vasanth, D. T. Crdier, M. Nandakumar, G. Pollack, R. Aggarwal, M. Rodder and H. Shichijo, "Transistor design issues in integrating analog functions with high performance digital CMOS," in *Proceedings of Symposium of VLSI Technical Digest*, 1999, pp. 147-148.

[13] M. H. Bhuyan and Q. D. M. Khosru, "Low frequency drain current flicker noise model for pocket implanted nano scale n-MOSFET," in *Proceedings of the IEEE and EDS Nano Materials and Device Conference* (NMDC) , CA, USA, Oct. 2010, pp. 295-299.

[14] H. V. Deshpande, B. Cheng and J. C. S. Woo, "Analog device design for low power mixed mode applications in deep sub-micrometer CMOS technology," *IEEE Electron Device Letters*, vol. 22, no. 7, pp. 588–590, Jul. 2001.

[15] T. Ohguro, H. Naruse, H. Sugaya, H. Kimijima, E. Morifuji, T. Yoshitomi, T. Morimoto, H. S. Momose, Y. Katsumata, and H. Iwai, "0.12 μm raised gate/source/drain epitaxial channel nMOS technology," in *Proceedings of IEEE IEDM Technical Digest*, 1998, pp. 927–930.

[16] T. Ohguro, R. Hasumi, T. Ishikawa, M. Nishigori, H. Oyamatsu, and F. Matsuoka, "An epitaxial channel MOSFET for improving flicker noise under low supply voltage," in *Proceedings of Symposium of VLSI Technical Digest*, 2000, pp. 160-161.

[17] J.-W. Wu, C.-C. Cheng, K.-L. Chiu, J.-C. Guo, W.-Y. Lien, C.-S. Chang, G.-W. Huang, and T. Wang, "Pocket implantation effect on drain current flicker noise in analog nMOSFET devices," *IEEE Transactions on Electron Devices*, vol. 51, no. 8, pp. 1262-1266, Aug. 2004.

[18] M. Miura-Mattausch, S. Hosokawa, D. Navarro, S. Matsumoto, H. Ueno, H. J. Mattausch, T. Ohguro, T. Iizuka, M. Taguchi, T. Kage, and S. Miyamoto, "Noise modeling with HiSIM based on self-consistent surface-potential description," *NSTI-Nanotech*, vol. 2, pp. 66-69, 2004.

[19] K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng, "A unified model for the flicker noise in metal–oxide–semiconductor field-effect transistors," *IEEE Transactions on Electron Devices*, vol. 37, no. 4, pp. 654–665, Apr. 1990.

[20] M. H. Bhuyan and Q. D. M. Khosru, "Linear asymmetric pocket profile based threshold voltage model for nano scale n-MOSFET," in *Proceedings of the International Conference on Electrical, Computer and Telecommunication Engineering* (ICECTE), Dec. 2012, pp. 300-303.

[21] K. N. Ratnakumar and J. D. Meindl, "Short-channel MOST threshold voltage model," *IEEE Journal of Solid State Circuits*, vol. 17, no. 10, pp. 937-948, Oct. 1982.

[22] M. H. Bhuyan and Q. D. M. Khosru, "Linear pocket profile based threshold voltage model for sub-100 nm n-MOSFET incorporating substrate and drain bias effects," in *Proceedings of the International Conference on Electrical and Computer Engineering* (ICECE), Dhaka, Dec. 2008, pp. 447-451.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Self Consistent Field Method with Damped Oscillation for Nano Device

Sudip Kumar Saha*, Ibnul Sanjid Iqbal, Md. Osman Goni
Department of Electronics & Communication Engineering
Khulna University of Engineering & Technology, Khulna-9203, Bangladesh
*Sudip_kuet08@yahoo.com

*Abstract*— **As CMOS technology progresses, device dimensions have been scaled down into the nanometer regime and classical physics failing to characterize the nanoscale devices as quantum mechanics start to play an important role in their characteristics. Modeling Quantum transport is being very significant for rigorous performance analysis of nano devices and simulation plays an important role in the field of nano device modeling. The most widely accepted method for simulation of nano devices is Self Consistent Field (SCF) method which uses Non-Equilibrium Green's Function (NEGF) formalism self consistently with Poisson solver equation. The NEGF formalism provides a rigorous description of quantum transport in nanoscale devices but heavy in computation. Now to reduce the computational burden a supervised algorithm in the updating procedure of SCF method is proposed in this paper. It reduces the computational burden drastically and to show the superiority a comparison in simulating DGMOSFET by traditional SCF and our proposed algorithm is provided.**

***Keywords—Quantum effect, DG-MOSFET, NEGF, SCF, Damped oscillation, Computational cost.***

## I. INTRODUCTION

Rapid device scaling pushes the dimensions of the field effect transistors to the nanometer regime [1]. Quantum effects play an important role in determining the device characteristics in this regime of operation. The main challenges arise with this advancement in feature size is Modeling of the complex physical phenomena which can't be rigorously explained by classical physics analysis based models where Non Equilibrium Green Function(NEGF) is required to rigorously describe the quantum phenomena that control the device behavior in the so extremely small dimensions [2]. To offer an additional insight into transport phenomena in these deeply scaled devices, simulation tools that consider quantum effect is highly appreciated [3]. In the field of Quantum Simulation Self Consistent Field (SCF) method is the most widely accepted method. It is an iterative process which solves the NEGF and Poisson's equation self consistently. As NEGF formalism is heavy in computation, simulation time of performance analysis has become an important factor [4]. We have provided a supervised algorithm in the updating procedure of SCF method to reduce iteration which in turn reduces the simulation time and we have considered Computational efficiency which is needed to make the self-consistent method suitable for device design and characteristic prediction. Our proposed supervised algorithm doesn't hamper the computational efficiency it just reduces the no. of iteration. For establishing the algorithm we have analyzed Double Gate (DG) MOSFET as it has emerged as promising devices for Very Large Scale Integration (VLSI) circuits due to their better scalability compared to bulk CMOS [5]. DG-MOSFETs characteristics are greatly affected by the quantum effects. These effects can be accurately predicted only using quantum mechanical based device simulation and the characteristics found by proposed algorithm shows exactly the same characteristics found by traditional SCF which proves the acceptance of the proposed method.

## II. SELF CONSISTENT FIELD METHOD

### A. Non Equilibrium Green Function & Poisson Solver

Device simulation based on NEGF is carried out using the so called self-consistent field method shown in Fig. 1. There are two main steps in SCF method, first one is quantum transport solver which is based on the NEGF formalism [6] and second one is Poisson's equation [7] solver. Electron ($n$) and hole ($p$) density will be found out from NEGF formalism and the electrical current ($I$) for a given potential $U$. From Poisson's equation, the electrostatic potential distribution ($U$) in the device for a given electron density ($n$) and hole density ($p$) can be calculated. The Poisson equation is solved by Finite Elements Method (FEM) [8] using the Partial Differential Equations (PDE) toolbox of MATLAB. The self-consistent method starts by assuming initial value for the potential $U_{old}$ and fed to the NEGF solver to calculate the $n$ and $p$ densities. The calculated densities are then fed to Poisson's solver to find the calculated potential $U_{new}$ in the device. We go forth and back between NEGF solver and Poisson's solver until the update in the potential drops below certain tolerance and then calculate the current.

### B. Updating Procedure of traditional SCF

To calculate the actual $U$ in the device, at first we start with an initial guess for $U$, calculate n from NEGF formalism, calculate an appropriate $U$ from the Poisson solver and

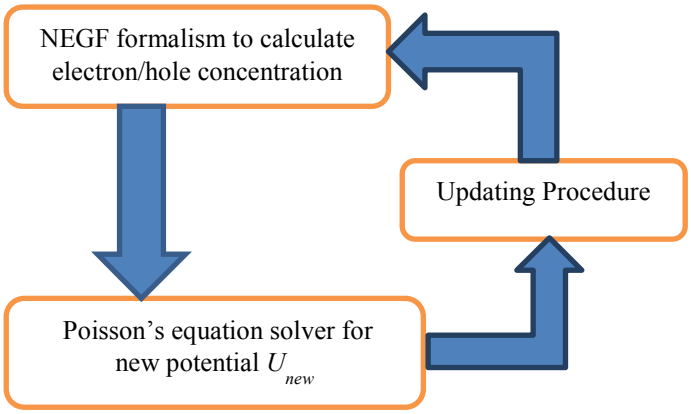



Figure 1. Flow chart illustrating the self-consistent method used in device simulation implementing the NEGF formalism.

compare with starting guess for $U$. If this new $U$ is not sufficiently close to the original guess, we revise our guess using a suitable algorithm; say something like [9]

$$U_{new} = U_{old} + \alpha\,(U_{calc} - U_{old}) \quad (1)$$

***where α is a positive number which is typically less than 1 that is adjusted to be as large as possible without causing the solution to diverge.*** The iterative process has to be repeated till we find a $U$ that yields an $n$ that leads to a new $U$ which is sufficiently close to the original value. Once a converged $U$ has been found, the current can be calculated.

### C. Supervised Updating Procedure

In supervised algorithm a new form of accelerator has been illustrated that is known as damped oscillation (DO). DO is produced by the following equations:

$$\mathrm{B}(i) = {}^{1}/_{(\mathrm{a}+i)} \quad (2)$$

$$C(i) = abs\left({}^{\sin(i*\frac{\pi}{6})}/_{10}\right) \quad (3)$$

$$\mathrm{DO}(i) = \mathrm{B}(i) \times \mathrm{C}(i) \quad (4)$$

Where 'a' is a user specified constant parameter & $i$ denotes the number of iteration.

In figure 2, Damped Oscillation is plotted. It is used instead of 'α' to accelerate the convergence and it achieves the convergence faster than 'α' as the initial value of DO is higher than α and also works well when reduction in error required without causing the divergence.

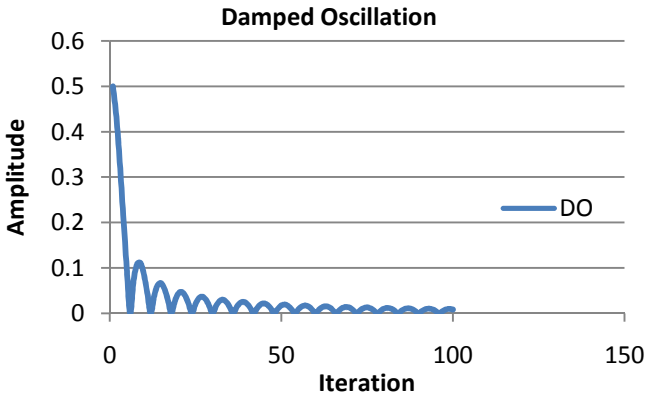



Figure 2. Damped Oscillation.

## III. Device Simulation

Here DG-MOSFET is chosen for simulation because it is the most widely used device for simulation which will help us to compare our result of simulation.

### A. Structure and parameters of DG-MOSFET

Fig. 3 shows a schematic diagram of DG-MOSFET. It is comprised of a conducting channel, surrounded by gate electrodes on either side. This ensures that no part of the channel is far away from a gate electrode.

In Fig. 4, we have tried to show how the channel is discretized in two dimensions. It means the 2-D spatial variation of potential along the channel. The potential is considered varying at every point of the co-ordinate.

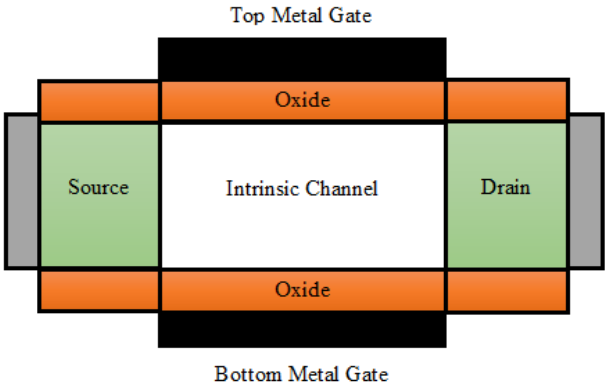



Figure 3. A General model of Double Gate MOSFET

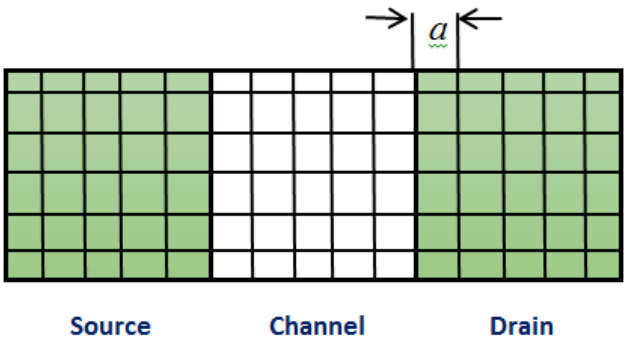



Figure 4. Geometrical view of Double Gate MOSFET.

Different parameters used in device simulation are listed below in Table I [10].

TABLE I. DIFFERENT PAREMETERS OF DG-MOSFET

| Parameter | Dimension |
|---|---|
| Donor Concentration, $N_D$ | $10^{20}$ cm$^{-3}$ |
| Channel Length, $L_g$ | 10 nm |
| Contact Length, $L_S$ *or* $L_D$ | 5 nm |
| Gate Work Function, $\square_m$ | 4.25 eV |
| Effective mass, $m^*$ | $m_x^*$=0.19$m_0$, $m_y^*$=0.91$m_0$ |
| Oxide thickness, $T_{ox}$ | 1 nm |
| Channel thickness, $T_{Si}$ | 2 nm |
| No. of grids along x-axis | 101 |
| No. of grids along y-axis | 9 |
| No. of bias points | 10 |
| Relative Permittivity, $\varepsilon_r$ | 11.7 |

*B. NEGF Formalism for DG-MOSFET*

The matrix form of Green function is given by

$$[\mathrm{G(E)}] = [EI - H - \sum_S - \sum_D]^{-1} \quad (5)$$

Where [G(E)] is a square matrix of order N.
*H*=Hamiltonian Matrix of order N.
*I* = Identity Matrix of order N.
$\sum_S$ = Self energy matrix of Source of order N.
$\sum_D$ = Self energy matrix of Drain of order N.

Formation of Hamiltonian for DGMOSFET:
The channel Hamiltonian *H* of DG-MOSFET is

$$H = U + \begin{bmatrix} \alpha & \beta & 0 & 0 & . & . & . & 0 \\ \beta^+ & \alpha & \beta & 0 & . & . & . & 0 \\ 0 & \beta^+ & \alpha & \beta & . & . & . & 0 \\ & & & . & . & . & & \\ & & & & . & . & . & \\ . & & & & & \beta^+ & \alpha & \beta \\ 0 & . & . & . & & 0 & \beta^+ & \alpha \end{bmatrix} \quad (6)$$

Where *U* is the self-consistent potential. Here *H* is an N × N matrix where N = $N_x \times N_y$.
Where,

$N_x$ = 101=number of grids along the x-axis
$N_y$ =9= number of grids along the y-axis

$$\alpha = \begin{bmatrix} 4t & -t & 0 & 0 & 0 & . & . & . & 0 \\ -t & 4t & -t & 0 & 0 & . & . & . & 0 \\ 0 & -t & 4t & -t & 0 & . & . & . & 0 \\ 0 & 0 & 0 & -t & 4t & -t & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & -t & 4t & -t \\ 0 & . & . & . & . & . & 0 & -t & 4t \end{bmatrix} \quad (7)$$

$$\& \quad \beta = \begin{bmatrix} -t & 0 & 0 & . & . & 0 \\ 0 & -t & 0 & . & . & 0 \\ 0 & 0 & -t & . & . & 0 \\ . & . & . & . & . & . \\ 0 & . & . & . & -t & 0 \\ 0 & . & . & . & 0 & -t \end{bmatrix} \quad (8)$$

α and β are $N_y \times N_y$ matrices.
Here $t = \hbar^2/2m^*a^2$, '*a*' is the grid spacing, '$m^*$' is effective mass [10].

Self Energy:

If we consider a large contact incorporated in the device Green's function using a self-energy ∑(E), then (5) can be written in

$$[\mathrm{G(E)}] = [EI - H - \sum(\mathrm{E})]^{-1} \quad (9)$$

$$\text{Where } \sum(\mathrm{E}) = \tau C^{-1} \tau^+ \quad (10)$$

is the self-energy function and $C^{-1}$is the contact part of the overall Green's function. Writing the above equation, we have ( $\equiv E + i0^+ - \alpha$ )

$$\sum = [-\beta \quad . \quad .] \begin{bmatrix} & -\beta & 0 & 0 & . & . \\ -\beta^+ & & -\beta & 0 & . & . \\ 0 & -\beta^+ & & -\beta & . & . \\ . & 0 & . & . & & \\ . & . & . & . & & \end{bmatrix}^{-1} \begin{bmatrix} -\beta^+ \\ . \\ . \end{bmatrix}$$

$$= \beta g \beta^+ \quad (11)$$

Where *g* surface Green's function, which is the matrix formed from the first $N_y \times N_y$ entries of $C^{-1}$. If we get *g* then we can find ∑. Now once *C* is a huge matrix, $C^{-1}$ can be written recursively [14] in terms of *C* as

$$C^{-1} = \begin{bmatrix} E + i0^+ - \alpha & -\beta \\ -\beta^+ & C \\ 0 & \\ . & \end{bmatrix}^{-1} \quad (12)$$

We need $g$, a small part of $C^{-1}$. So we use the following matrix identity: if

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Then $a = [A\text{-}BD^{-1}C]^{-1}$. Using this identity in (12), we get

$$g = [E + i0^+ - \alpha - \beta g \beta^+]^{-1} \quad (13)$$

Where, $0^+$ represents the positive infinitesimal.

Equation (13) may be solved in general using an iterative procedure [11]-[14] in which we start with an initial guess for $g$ and keep on refining it using (13) till we obtain the solution. But for our special case, since $\beta$ is a multiple of the identity matrix, the above equation may be written as a matrix quadratic equation as follows

$$\beta^2 g^2 + (\alpha - (E + i0^+)I)g + I = 0$$

It turns out that because all the matrices in the above equation viz. $\beta$, $\alpha$, $I$ commute with each other, the above equation may be solved just as if it were an ordinary scalar quadratic equation, and the solution is

$$g = \frac{-b+\sqrt{(b^2-4ac)}}{2a} \quad (14)$$

Where all the arithmetic operations in the above equation are to be carried out in a matrix sense, with $a=\beta^2$, $b = \alpha - (E + i0^+)I$ and $c=I$, the identity matrix.

If we get the surface green function, we can find out the self energy from (11) [10].
Now the density matrix $\rho$ is given by

$$\rho = \frac{1}{2\pi}\int_{-\infty}^{+\infty}[f_1(E)G\Gamma_S G^+ + f_2(E)G\Gamma_D G^+]dE \quad (15)$$

Where,
$f_1(E)$ = Fermi Function of Source
$f_2(E)$ = Fermi Function of Drain
$\Gamma_S = i[\sum_S - \sum_S^+]$
$\Gamma_D = i[\sum_D - \sum_D^+]$
$G^+$ = Conjugate Transpose of Green's Function, $G$
$\sum_S^+$= Conjugate Transpose of $\sum_S$
$\sum_D^+$= Conjugate Transpose of $\sum_D$

From which the total number of electrons n can be calculated by taking a trace

$$n = \text{Trace } [\rho] \quad (16)$$

The current is given by

$$I = \frac{2q}{h}\int_{-\infty}^{+\infty}[Trace(\Gamma_S G \Gamma_D G^+)(f_1(E) - f_2(E))]dE \quad (17)$$

Through (15) and (17), we can easily calculate electron density and current which is used to calculate the self-consistent potential for the next iteration.

### C. Poisson Solver

Using Poisson solver, we can solve Poisson equation. The Poisson equation is

$$\nabla^2 U = -\frac{q}{\varepsilon}(p - n + N_D - N_A) \quad (18)$$

Where,
$p$=hole distributions
$n$ =electron distributions
$N_D$=*doping concentration of donor*
$N_A$=*doping concentration of acceptor*
$q$= Charge of electron

Here we have ignored the hole density. Considering $p = 0$ & $N_A = 0$;

$$\nabla^2 U = -\frac{q}{\varepsilon}(-n + N_D) \quad (19)$$

Through (19) we can calculate the potential in the device or channel which in turn is used in calculating the current and electron density by NEGF formalism [15].

## IV. Result And Discussion

The linear and log scale I-V characteristics is shown in fig. 5 and 6. The off current is about 0.03μA and the on current is 1800μA. Here we used 11 bias points to draw the characteristics curve.
Here bias point indicates-
$V_{DS}$= [0 0.05 0.1 0.15 0.2 0.25 0.30 0.35 0.40 0.45 0.5]

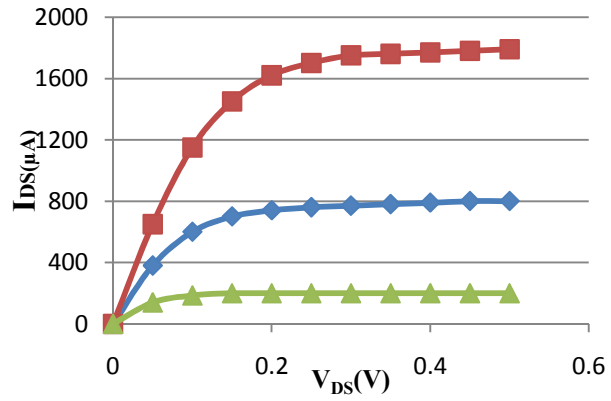



Figure 5. Drain current $I_{DS}$ vs. drain voltage $V_{DS}$ for different gate voltages

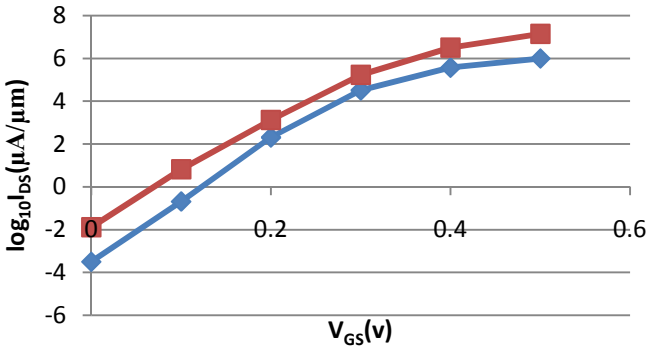



Figure 6. Drain current $I_{DS}$ Vs. gate voltage $V_{GS}$ for different drain voltages

Computational cost comparison between Traditional Updating Procedure of SCF Method and Supervised Updating Procedure SCF Method in 2D simulation of DGMOSFET is shown in Table II. In this table, total number of iteration and total time is reduced from the traditional updating procedure of SCF method.

TABLE II. TOTAL COMPUTATIONAL COST

| Parameters | SCF | SCFDO |
|---|---|---|
| Total No. of Bias point | 10 | 10 |
| Total Iteration | 872 | 342 |
| Total Time | 13.3 hours | 5.36 hours |

## CONCLUSION

In this paper, we have used traditional NEGF formalism for its rigorousness in Quantum analysis. In Quantum Simulation a supervised algorithm is proposed in the updating procedure of most widely accepted SCF method which is based on NEGF formalism. To establish the proposed algorithm Double Gate MOSFET is analyzed considering the 2-D spatial variation of potential along the channel with both traditional and proposed algorithm and shows the superiority of our proposed algorithm.

## REFERENCES

[1] International Technology Roadmap for Semiconductor (ITRS) [online]. Available: http://public.itrs.net

[2] J.A. Kenrow, "Characterization and Analysis of OFET Devices Based on TCAD Simulations," IEEE. Trans. Electron Devices, vol. 52, No. 9, pp. 2034-2041, Sept. 2005.

[3] Wanqiang Chen, Leonard F. Register and Sanjay K. Banerjee, "Scattering in a Nano-Scale MOSFET: A Quantum Transport Analysis," in *Proc. Third IEEE Conference on Nanotechnology,* Aug. 2003, pp. 32-35.

[4] Vasser M. Sabry, Mohammed T. Abdel-Hafez,Tarek M.Abdolkader,Wael Fikry Farouk "A Computationally Efficient Method for Quantum Transport Simulation of Double-Gate MOSFETs".

[5] Z. Ren, R. Venugopal, S. Goasguen, D. Datta and M. Lundstrom "nanoMOS 2.5: A Two-Dimensional Simulator for Quantum Transport in Double-Gate MOSFETs," IEEE Trans. Electron Devices, vol. 50, pp. 1914-1925, Sep. 2003

[6] S. Datta, "Nanoscale device modeling: the Green's function method," Superlatt. Microstruc., vol. 28, pp. 253-278, 2000

[7] Y. Taur and T. Ning, Fundamentals of Modern VLSI Devices,Cambridge University Press, Cambridge, UK, 1998.*FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.

[8] N. Ottosen, and H. Petterson, Introduction to the Finite Element Method. Prentice Hall,New York, 1992.A. Karnik, "Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP," M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.

[9] Supriyo Datta,Quantum Transport: Atom to transistor, Cambridge university press,2005.

[10] Prashant subhash damle,Nanoscale device modeling from MOSFETS to molecules, P.h.d thesis paper, purdue university,2003.

[11] R. Haydock, V. Heine, and M. Kelly, "Electronic structure based on the local atomic environment for tight-binding bands," J. Phys. C, vol. 5, pp. 2845-2858, 1972.

[12] E. Mele and J. Joannopoulos, "Electronic states at unrelaxed and relaxed GaAs surfaces," Phys. Rev. B, vol. 17, pp. 1816-1827, 1978.

[13] M. L. Sancho, J. L. Sancho, and J. Rubio, "Quick iterative scheme for the calcula¬tion of transfer matrices: application to Mo(100)," J. Phys. F, vol. 14, pp. 1205¬1215,

[14] M. Samanta, "Electronic conduction through organic molecules," Master's thesis, Purdue University, 1995.

[15] Abdolkader, Tarek M. , Fathi, M. , Fikry, W. , Omar, O.A.,"FETMOSS: a software tool for 2D simulation of double-gate MOSFET" 3rd International Conference on ITI,Dec,2005,pp-193-208

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Autocorrelation – MTM: A Compound Detector for Spectrum Hole Identification in Low SNR

Md. Mizanur Rahman[1], Chalie Charoenlarpnopparut[1,*], Prapun Suksompong[1], and Attaphongse Taparugssanagorn[2]
[1]Electronics and Communication Engineering, SIIT, Thammasat University, Thailand
[2]Wireless Information Security and Eco-Electronic Research Unit (WISRU), NECTEC, Thailand
[*] E-mail: chalie@siit.tu.ac.th

***Abstract*—After digital switch over of TV transmission FCC has reported that Cognitive Radios (CR) can opportunistically use the free TV band under certain constraint among which primary user (PU) detection at low SNR is of vital importance not to create interference to the TV receivers. Since then spectrum sensing at low SNR has become the prime research issue to find spatiotemporal information of the available spectrum hole into the TV band for CR applications. In this paper we have proposed an Autocorrelation-MTM for spectrum hole identification. It has been found that detection performance can be improved significantly in low SNR by the proposed Autocorrelation-MTM compared to MTM with added complexity.**

***Keywords—TV band, low SNR, Autocorrelation, MTM, discrete prolate spheroidal sequence.***

## I. INTRODUCTION

Demand for wireless services are increasing each day. But the entire spectrum suitable for wireless communication has been allocated to the primary wireless service providers. As a result there is no more space available into the spectrum to welcome new wireless service provides or the existing service providers with innovative wireless services. In [1, 2] J. Mitola has announced Cognitive Radio (CR), which is an advanced version of Software Defined Radio (SDR) to opportunistically use the idle primary spectrum as spatiotemporal basis. In a recent report FCC has stated that the allocated licensed spectrum are rarely utilized continuously across time and space [3]. Another report by McHenry shows low utilization of the frequency bands between 30 MHz and 3 GHz [4]. It is understood that the seemed spectrum scarcity problem appears because of the inefficient utilization of the allocated spectrum. To solve these issues after digital TV switch over FCC [5] and Ofcom [6] has published new rules to allow secondary users (SU) to opportunistically use the TV bands. TV bands have attracted the wireless service providers for its efficient long distance propagation characteristics compared to the high frequency spectrum.

Radius for TV transmission can be as high as 100 km, which causes low SNR at the coverage boundary. To avoid any interference with the primary TV receivers the principal task of a CR is to sense the spatiotemporal information of expected TV frequency band at low SNR to find any available spectrum hole for opportunistic usage. Existing well-known spectrum-sensing techniques are energy detection (ED) [7], matched filtering [8], Cyclostationary detector [8], and waveform-based sensing [9]. Cyclostationary feature detector and Matched filter detector are robust to noise power uncertainty and propagation channel, but needs priori information of primary signal. The design and detection process of a Cyclostationary feature detector or matched filter detector gets even more complicated as the number of signals to be detected increase [8]. Waveform based filtering also needs priori information of signals that is to be detected. Need of prior information of the signal to be detected is unrealistic for cognitive radio application. ED the simplest detection technique although does not need prior information of signal but suffers from performance degradation in low SNR and noise uncertainty situation. In addition ED is unable to discriminate primary signal and noise [8].

In [10] an Autocorrelation-based spectrum sensing has been proposed that exploits the distinctive features of signal (i.e., modulation schemes, bandwidth, RF channel guard bands). It has been shown that Autocorrelation-based detector provide improvement in the probability of detection over ED based sensing for moderate SNR scenarios where fast sensing is required. This Autocorrelation-based method is incapable to multi band detection.

In addition to the disadvantages mentioned for the detection techniques discussed above, they cannot distinguish multiband simultaneously. Non-parametric methods, like Multitaper method (MTM) [11, 12] and filter bank spectral estimator (FBSE) [13] have recently got attention for their accurate spectrum estimation without a priori knowledge of signal and noise. Multitaper method was first introduced for spectrum estimation in [14] by Thomson to improve the bias-variance dilemma of spectrum estimation. Later S. Haykin incorporated this spectrum estimation technique in CR for spectrum sensing [11]. Farhang Boroujeny represented the MTM as a FBSE in [13]. The basic difference between the MTM and FBSE is that FBSE uses only one prototype filter whereas in MTM different prototype filters are used where the filter coefficients are chosen from a orthogonal set of sequences called Discrete Prolate Spheroidal Sequence (DPSS) also known as Slepian sequences or Slepian tapers according to the name of its inventor David Slepian [15].

In this paper we propose to combine Autocorrelation and MTM to improve the detection performance at low SNR and incorporate the multiband detection capability of MTM. MATLAB simulation shows that multiband detection at significantly low SNR is possible with the proposed

compound detector that consists of an Autocorrelator and an MTM compared to MTM detector alone.

The rest of this paper is organized as follows. In section II a simple Autocorrelation method of signal detection from discrete time composite random signal is discussed. Spectrum estimation by MTM is presented in section III. Section IV represents the proposed Autocorrelation–MTM system of signal detection, which has been followed by complexity analysis in section V. Simulation and results, has been presented in section VI. Finally conclusion and possible future improvement of this work has been given in section VII and VIII accordingly.

## II. AUTOCORRELATION OF DISCRETE TIME COMPOSITE RANDOM SIGNAL

The simplest autocorrelation method as given in [16] of a discrete time random process is presented in brief as follows.

Let $y(n)$ be a received noise corrupted causal sequence of length $N$ represented as,

$$y(n) = x(n) + w(n), \tag{1}$$

where, $x(n)$ is a periodic sequence of period $N$ and $w(n)$ represents additive white Gaussian noise of length $N$. The biased autocorrelation can be given as in

$$\begin{aligned} R_{yy}(l) &= \tfrac{1}{N}\textstyle\sum_{n=0}^{N-l-1} y(n)y(n-l), \\ &= \tfrac{1}{N}\textstyle\sum_{n=0}^{N-1} x(n)x(n-l) \\ &+ \tfrac{1}{N}\textstyle\sum_{n=0}^{N-1}[x(n)w(n-l) + w(n)x(n-l)] \\ &+ \tfrac{1}{N}\textstyle\sum_{n=0}^{N-1} w(n)w(n-l) \\ &= R_{xx}(l)+R_{xw}(l) + R_{wx}(l) + R_{ww}(l) \end{aligned} \tag{2}$$

where, $R_{xx}(l)$ is the autocorrelation sequence of $x(n)$. Since $x(n)$ is periodic, its biased autocorrelation sequence exhibits the same periodicity and large peaks at $l = 0, N, 2N$ and so on. The cross correlation $R_{xw}(l)$ and $R_{wx}(l)$ between the signal $x(n)$ and the AWGN, and autocorrelation sequence $R_{ww}(l)$ has maximum value at $l = 0$ but decrease rapidly toward zero for other values of $l$. Consequently only $R_{xx}(l)$ is expected to have large peaks for $l > 0$. An autocorrelation-based detector has been shown in Fig. 1. Threshold of the Autocorrelator in Fig. 1 should be set by the *a priori* estimation of the correlated noise of Fig. 2(d). To explain the Autocorrelation process we have considered four sinusoidal frequencies 10 Hz apart starting from 50 Hz, the combination of which is presented in Fig. 2(a). The composition of the four sinusoids added to noise is presented in Fig. 2(b). In Fig. 2(d) autocorrelation of only the noise is presented. It is evident from Fig. 2(c) & 2(d) that autocorrelation sequences are significantly different when signal is present into the noise compared to the absence of signal. Thus the autocorrelation have the capability to identify the presence or absence of the signal in a band comparing the energy of the noisy signal calculated from autocorrelation sequence of Fig. 2(c) to the threshold (Th) which is calculated from the energy of only noise calculated from the autocorrelation sequence shown in Fig. 2(d). With this method if $N = 2W$ where $W$ is the maximum frequency of the received analog signal $r(t)$, then depending on the presence or absence of any signal with frequency less than $W$, decision is made in favor of $H_1$ or $H_0$ where $H_1$ is the hypothesis for signal detected and $H_0$ is the hypothesis for signal not detected. But detection decision over multiple frequency bands is not possible by this method.

## III. MULTITAPER ESTIMATOR

The FBSE realization of MTM can be presented as follows where coefficients of the prototype filters in different periodogram are DPSSs. The Slepian tapers expand the part of the time series in a fixed bandwidth $(f-W)$ to $(f+W)$, centered on some frequency $f$. MTM uses $K$ Slepian orthonormal tapers to obtain $K$ periodograms, which are averaged to achieve the MTM estimate. The process is explained as follows.

Fig. 3 shows FBSE realization of the MTM. Let for $n = 0, 1, \ldots\ldots\ldots, N-1$; $y(n)$ in (1) be the discrete time series of length $N$ whose power spectrum estimation (PSE) is required, $g_k(n)$, $k = 0,1, \ldots\ldots, K-1$, are the $K$ orthonormal Slepian tapers each of length $N$, then PSE of the discrete time signal $y(n)$ for each Slepian taper can be represented as in

$$\left(\hat{S}_{slep}(f\ )\right)_k = \textstyle\sum_{n=0}^{N-1}\left|g_k(n)y(n)e^{-j2\pi fn}\right|^2, \tag{3}$$

where, $k = 0, 1, \ldots\ldots\ldots, K-1$; $K = NW - 1$ and $NW$ is the time bandwidth product. Associated with each periodogram defined by (3) is an eigenvalue represented by $\lambda_k$, where, $= 0, 1, 2, \ldots\ldots, 2NW-1$. For simplification of analysis we consider $\lambda_k = 1$. In this paper the time-bandwidth product $NW$ has been considered to be 3.5 as well as ($K = NW-1= 6$ Slepian tapers for simulations.

Averaging the outputs of the K periodogram found by (3) the PSE of MTM is found as in

$$\hat{S}_{mtm}(f\ ) = \tfrac{1}{K}\textstyle\sum_{k=0}^{K}\left|\left(\hat{S}_{slep}(f\ )\right)_k\right|, \tag{4}$$

where, $k = 0, 1, \ldots\ldots\ldots, K-1$.

## IV. PROPOSED AUTOCORRELATION-MTM SYSTEM MODEL

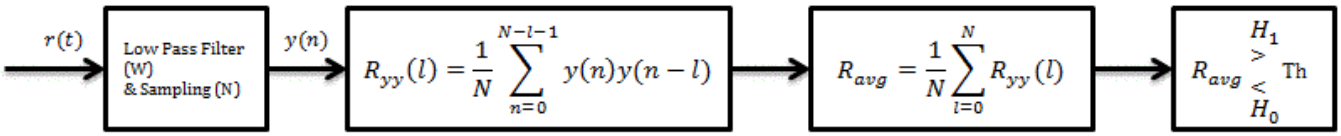



Figure 1. Autocorrelation-based detector.

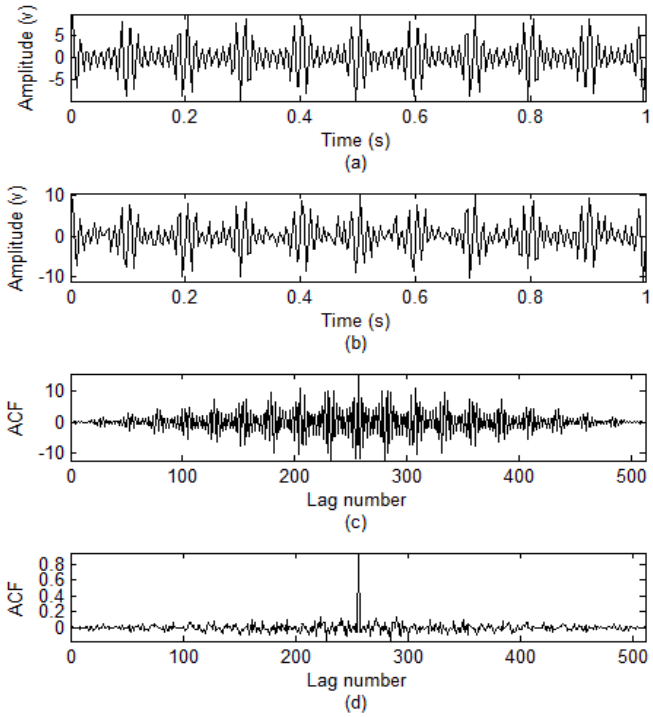



Figure 2. (a) Combination of four sinusoidal signals, (b) Noise of variance one and mean zero is added to the combined signal in (a), (c) Autocorrelation of the noisy signal of (b), (d) Autocorrelation of the noise.

The proposed detector in this paper consists of an Autocorrelator discussed in section II as the pre-processor of the discretized received signal to generate the autocorrelation sequence. To achieve the desired estimated spectrum an MTM estimator processes this autocorrelated signal in the second step.

To evaluate and compare the MTM and Autocorrelation-MTM detector performance we have estimated the AWGN channel noise considering the channel to be free of any signal and took the maximum noise level as threshold (Th). Fig. 4 shows the complete block diagram of the proposed Autocorrelation-MTM detector.

From $y(n)$ we get the autocorrelated sequence $R_{yy}(l)$, where, $l = 0, 1, 2, \ldots\ldots, 2N-1$. This autocorrelated sequence is then processed by the MTM module to find the estimated spectrum according to (5) and (6).

$$\left(\hat{S}_{slep-autocorr}(f\ )\right)_k = \sum_{l=0}^{2N-1}\left|g_k(l)R_{yy}(l)e^{-j2\pi fl}\right|^2 \quad (5)$$

where, $k = 0, 1, \ldots\ldots\ldots, K-1$; $K = NW - 1$, and $NW$ is the time bandwidth product.

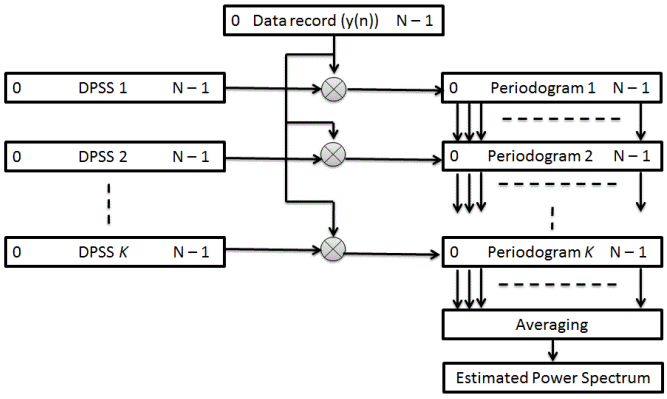



Figure 3. Realization of Multi-Taper Method as Filter Bank Spectrum Estimator.

Averaging the outputs of the $K$ peridograms found by (2) the PSE of the Autocorrelator-MTM is found as in

$$\hat{S}_{mtm-autocorr}(f\ ) = \frac{1}{K}\sum_{k=0}^{K}\left|\left(\hat{S}_{slep-autocorr}(f\ )\right)_k\right|. \quad (6)$$

## V. Complexity Analysis of MTM and MTM-Autocorrelation Methods

An MTM estimator requires a DPSS generator, $K$ periodogram generators, and a mean finder. As in this paper we focus into multi-frequency detection we compare the complexity of MTM to the proposed compound Autocorrelation-MTM methods. DPSS generator is common to both the methods. Although the DPSS complexity increases with the number of samples, for simplicity of analysis we do not consider the complexity of DPSS generator. We do not consider the complexity of serial to parallel converters also, as this is common to all types of estimators.

To calculate FFT of a discrete time sequence of length N, $\frac{N}{2}log_2N$ complex multiplications and $Nlog_2N$ additions are needed. One complex multiplication needs 4 real multiplications and 2 real additions. On the other hand 1 complex addition requires 2 real additions [17]. Thus complexity of a periodogram can be found to be $(2N + 2Nlog_2N)$ real multiplications and $(N + 3Nlog_2N)$ real additions. As MTM needs $K$ periodograms found by $K$ Slepian tapers the complexity becomes, $K(2N + 2Nlog_2N)$ real multiplications and $K(N + 3Nlog_2N)$ real additions. At the output of each periodogram we get $\left(\frac{N}{2}+1\right)$ frequency samples. To achieve MTM spectrum we need $\left(\frac{N}{2}+1\right)(K-1)$ additions and 1 multiplication by $\frac{1}{K}$ at the end. Thus the total complexity of MTM becomes, $K(2N + 2Nlog_2N) + 1)$ real multiplications and $(K\ (N + 3Nlog_2N) + \left(\frac{N}{2}+1\right)(K-1))$ real additions.


Figure 4. Proposed Autocorrelation-MTM detector.

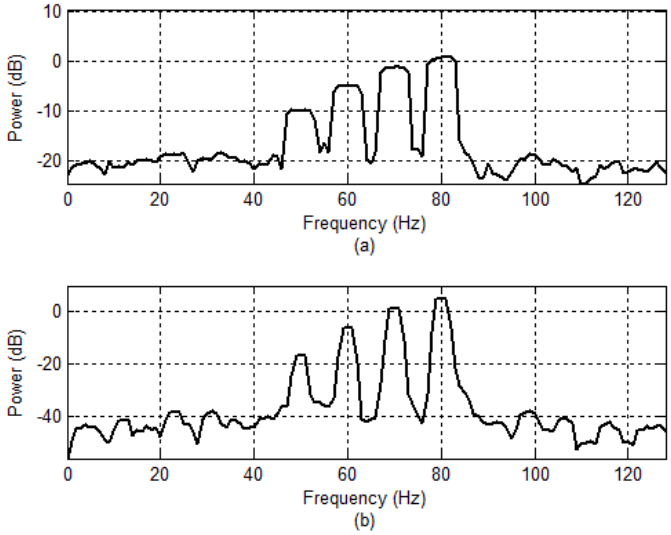



Figure 5. Estimated spectrum of four sinusoidal signals of frequency , , , , Hz in presence of noise shown in Fig. 1(b) by (a) MTM, and (b) Autocorrelation-MTM.

A Brute force method based on the definition of (1) for autocorrelation needs $N^2$ complex multiplications and $(N-1)^2$ additions. Again as one complex multiplication needs 4 real multiplications and 2 real additions for $N^2$ complex multiplications we need $4N^2$ real multiplications and $2N^2$ real additions. In addition $(N-1)^2$ complex additions needs $2(N-1)^2$ real additions. Thus an autocorrelator needs $4N^2$ real multiplications and $2(2N^2-2N+1)$ real additions.

The total complexity of Autocorrelation-MTM detector thus becomes $(4N^2 + K(2N + 2Nlog_2N) + 1))$ real multiplications and $(2(2N^2 - 2N + 1) + (K\,(N + 3Nlog_2N) + \left(\frac{N}{2} + 1\right)(K-1)))$ real additions.

## VI. Simulation and Results

Noisy signal of Fig. 2(b) was processed by the MTM estimator and autocorrelated sequence in Fig. 2(c) of the noisy signal shown in Fig. 2(b) was processed by the Autocorrelation-MTM estimator. Fig. 5 shows the simulation result of estimated spectrum. It can be seen that Autocorrelation-MTM has better resolution of estimated spectrum compared to MTM alone.

To analyze the detection capability of Autocorrelation–MTM method compared to MTM method we consider a sinusoidal signal of frequency 100.501 kHz, which was added to an AWGN of mean zero and variance one. The spectrum of the noise and the noisy signal both were estimated at every run and detection decision was made when the maximum signal level crossed the maximum noise level for both the methods mentioned above.

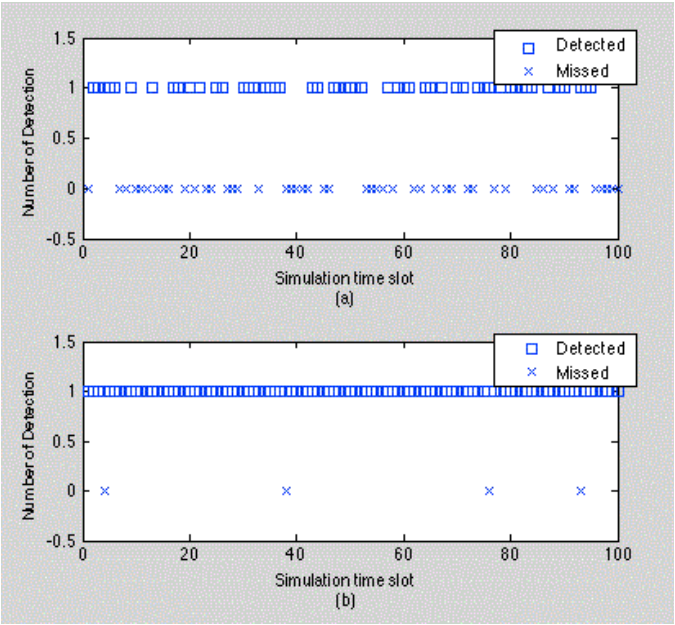



Figure 6. A sinusoidal signal detection at SNR of −56.5864 dB by (a) MTM, (b) Autocorrelation-MTM.

TABLE I Correct detections by MTM and Autocorrelation-MTM detectors at -56.5864 dB

| Number of Simulations | Number of correct detections by MTM detector | Number of correct detections by Autocorrelation-MTM detector |
|---|---|---|
| 10,000 | 5,108 | 9,635 |

Simulation was performed 10,000 times at an average SNR of −56.5864 dB (calculated by maximum signal to maximum noise ratio) at each time slot to detect the sinusoidal signal of 100.501 kHz emerged into the noise. The results represented in Fig. 6 and in table I show that Autocorrelation-MTM has higher detection capability compared to MTM detector at low SNR. Among 10,000 simulations the sinusoidal signal could be detected 5108 times by MTM and 9,635 times by Autocorrelation-MTM detector. This 45.27% boosted performance in detection by Autocorrelation-MTM at low SNR of −56.5864 dB was achieved with additional complexity of $4N^2$ real multiplications and $2(2N^2-2N+1)$ real additions over MTM, where $N$ represents number of samples. The detection information for first 100 simulation time slots for MTM and Autocorrelation-MTM has been presented in Fig. 6(a) and Fig. 6(b) accordingly. It is evident that Autocorrelation-MTM has significant detection improvement over MTM.

## VII. Conclusion

Autocorrelation-MTM, the compound detection method proposed in this paper can be used for CR applications in low

SNR. Moreover MTM method enriches the CR detection capability with its multiband detection capability. As was found from the simulation experiments that performing signal autocorrelation before MTM processor can detect multiple bands in low SNR for CRs. The choice of the methods depends on the acceptance level of the complexity and SNR requirement in specific situation.

## VIII. FUTURE WORK

DPSS implementation complexity was not considered in the paper. In future we hope to focus on DPSS implementation complexity and find approximate number of DPSSs for faster MTM performance. Experimenting the performance of the proposed method into the TV band is another future goal.

## ACKNOWLEDGMENT

The authors thank Thailand Office of Higher Education Commission (NRU Project) for the financial support.

## REFERENCE

[1] J. Mitola, "Software radios: survey, critical evaluation and future directions," *IEEE Aerospace and Electronic Systems Magazine*, vol. 8, no. 4, pp. 25-36, Apr. 1993.

[2] J Mitola, G.Q. Maguire, " Cognitive Radio: Making software radios more personal," *IEEE Personal Communications Magazine*, 6(4), pp.13-18, 1999.

[3] FCC, "Spectrum policy task force report (ET Docket no. 02-135)," Nov. 2002.

[4] M. A. McHenry, "NSF spectrum occupancy measurements project summary," shared spectrum co. report, Aug. 2005.

[5] "FCC adopts rules for unlicensed use of television whitespaces," Official announcement of FCC, Nov. 2008.

[6] (2009) Ofcom|Digital dividend: cognitive access. [Online]. Available: http://stakeholders.ofcom.org.uk/consultations/cognitive/statement/

[7] F. F. Digham, M. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," in *Proc. IEEE Intl. Conf. Commun.*, May 2003, pp. 3575–3579.

[8] A. Ghasemi and E. S. Sousa, "Spectrum sensing in Cognitive Radio networks: requirements, challenges and design trade-offs," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 32-39, Apr. 2008.

[9] A. Sahai, R. Tandra, S. M. Mishra, and N. Hoven, "Fundamental design tradeoffs in cognitive radio systems," in *Proc. of Int. Workshop on Technology and Policy for Accessing Spectrum*, Aug. 2006.

[10] R. K. Sharma and J. W. Wallace, "Improved Spectrum Sensing by Utilizing Signal Autocorrelation," in *VTC Spring, IEEE*, Apr. 2009, pp. 1-5.

[11] S. Haykin, "The Multitaper Method for Accurate Spectrum Sensing in Cognitive Radio Environments," *Forty-First Asilomar Conference Record on Signals, Systems and Computers*, Nov. 2007, pp. 436-439.

[12] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," Proc. IEEE, vol. 97, no. 5, pp. 849–877, May 2009.

[13] B. Farhang-Boroujeny, "Filter Bank Spectrum Sensing for Cognitive Radios," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1801-1811, May 2008.

[14] D. J. Thomson, "Spectrum Estimation and Harmonic Analysis," *Proceedings of the IEEE*, vol. 70, no. 9, pp. 1055-1096, Sept. 1982.

[15] D. Slepian, "Prolate spheroidal wave functions, fourier analysis, and uncertainty v: The discrete case," *Bell System Technical Journal*, vol. 57, no. 5 pp. 1371-1430, May-Jun. 1978.

[16] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing Principles, Algorithms, Applications*, Prentice-Hall International Inc., 1996.

[17] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing: A Practical Approach*, Pearson Education, 2002.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Balanced Energy and Coverage Guaranteed Protocol for Wireless Sensor Networks

Nam Tuan Le*, Nirzhar Saha, Ratan Kumar Mondal, Sunghun Chae, and Yeong Min Jang†
Dept of Electronics Engineering
Kookmin University, Seoul, Korea
E-mail: †yjang@kookmin.ac.kr

*Abstract*—**Wireless platforms are becoming less expensive and more powerful, enabling the promise of widespread use, everything from health monitoring to military sensing. In large scale applications like military surveillance, sensing coverage is essential for target detection and optimization of power consumption is important for lifetime extension of the network. Coverage and long lifetime network are two important factors and challenges for sensor networks. In this paper, we propose a new scheme, named as BECG, which remains the best coverage of the sensing area and energy balanced scheduling for all sensors. It allows nodes to determine when they can switch to sleep mode during operation. The energy efficiency of the proposed scheme is shown by intelligent decisions making. Each node in the network takes decision to turn on or turn off in a distributed manner that results a set of small number of active nodes throughout the lifetime of the network and covers sensing area of interest for target detection. It reduces redundancy, power consumption and increase the lifetime of a network. At first we applied our algorithm for a random topology and then we evaluate the performance of the network by simulation.**

***Keywords*— *Sensor Network, MAC protocol, Power efficiency, Coverage***.

## I. INTRODUCTION

Wireless sensor network is an emerging network technology in this century. Its emergence brings a new phase in the development and low cost sensor motes large scale applications of wireless sensor network such as monitoring of climate or animal/plant habitation, tracking of targets in the battlefield, and observing buildings or infrastructures for defense. In general, sensor nodes are operated by a small battery that has a limited amount of energy. Moreover, since the region where sensor nodes are deployed is usually dangerous and inhospitable terrain, the battery of the sensor node cannot be easily recharged or replaced. Therefore, in wireless sensor networks, reducing energy consumption of each sensor node is one of the important issues to prolong the network lifetime. Under such battery-based constraints with limited power supply, energy consumption is a key issue for designing sensor network applications. Some sensor products adopt an IEEE 802.11-like MAC protocol. However, IEEE 802.11MAC is not a good solution for sensor networks. S-MAC (sensor MAC) proposes enhanced schemes, such as periodic sleep and overhearing avoidance, to provide a better choice for sensor applications.

For military sensor network applications, power conservation and the quality of surveillance are two conflicting requirements of target detection in sensor networks. With unlimited power supply, a given area can be monitored perfectly with a set of sensor nodes that cover the entire area in terms of sensing. However, since the sensor nodes have limited power, the quality of monitoring becomes inversely proportional to the lifetime of the network. The coverage of a region of interest is very important for monitoring and target detection. Uniform deployment in the military field is impossible because of inaccessibility of human. So, sensors are deployed in the military field using unmanned vehicle or helicopter. It resembles to be randomly distributed. This distribution leads to overlapping sensing areas for some regions and some regions are out of coverage. To assure the coverage in the whole area of interest, it needs to deploy the nodes for two or three times or densely for one time. Besides fault-tolerance, the highly dense nodes can increase the information precision and collision. It decreases network's lifetime. In such dense networks, energy-efficient scheduling is a key factor to extend the functionality and the lifetime of the network. The fundamental challenge of scheduling is to maximize the number of sleeping nodes to conserve more energy while maintaining the coverage of the WSN.

There are several approaches which try to solve these problems. One important concern on network survivability is energy balance. If sensor nodes consume energy more equally, the chance that some nodes use up their energy much earlier and therefore the network is partitioned becomes lower. So, the question arises, if the network is made dense to get the coverage, how it is possible to find out a minimum set of nodes in a distributed manner to fulfill the coverage requirement and to turn off other nodes to increase the network lifetime. And how turning on and turning off scheduling can be made among the nodes so that all the nodes will maintain the coverage as well as balance the energy consumption among nodes.

In this paper we propose an energy balanced scheme for random deployment sensor network by reducing redundancy of sensors and sharing responsibility of working state to guarantee the coverage. The rest of this paper will further explain the scheme with the following sections. In section II, we summarize existing research works related to energy-conservation scheduling. In Section III we present our algorithm for Balanced Energy, Guaranteed Coverage and also analysis the mathematical model of the proposed scheme. Then the evaluation performances are in Section IV. Finally, Section V concludes the paper.

## II. RELATED WORKS

Since sensor nodes are equipped with a limited energy of batteries, reducing energy consumption is a critical issue for extending network lifetime. Besides the development of electronic design, node scheduling is necessary for making the network energy efficient. Redundancy minimization as well as coverage preservation in dense sensor networks is interested research for some years. This is done by turning off redundant nodes. Most of approaches are for the static case of network, they didn't consider the dynamic or harsh environment, the area of interest may have an adverse environment that has high degrees of humidity, temperature, or intentional destruction from malicious entities, in addition to node power depletion as nodes are not rechargeable; unexpected node failures are likely to become norms rather than exceptions. Some nodes may also shift from one place to another place due to storm or flood as the nodes are light weight.

In S-MAC, sensor nodes have a fixed duty cycle and a synchronized scheduling in a virtual cluster. Hence, in S-MAC, it is not easy to adapt to the variation of network environment. Moreover, due to the synchronized scheduling, transmission collisions will increase resulting in energy waste and low throughput. T-MAC [1] and DS-MAC [2] enhance S-MAC by adding a dynamic duty cycle feature. In T-MAC, if there is no activation event for a time threshold TA, the listen period ends immediately and the node goes to the sleep period, which results in further energy saving compared with S-MAC. In DS-MAC, the duty cycle of the sensor node is adjusted based on traffic condition, which results in improving latency performance compared with S-MAC. However, they also adopt synchronized listen-sleep scheduling and, thus, have the same problem as that of S-MAC. P-MAC [3] is another protocol that adaptively determines the listen-sleep schedule for a node. Each sensor node generates patterns indicating its tentative listen-sleep plan according to its data traffic to send and determines its actual listen-sleep schedule by comparing its own pattern with those of neighbor nodes. However, since the duty cycle of each node is determined based only on its own traffic amount to send, the receiver node may have a low duty cycle even though it has a lot of packets to receive, which results in a low duty cycle of the schedule for the transmission between transmitter and receiver nodes with heavy traffic load. Moreover, if a node has a packet to transmit, it always tries to transmit it in each time slot. Hence, P-MAC has the same collision problem as in S-MAC in heavy traffic and dense node deployment situations

If we consider about the power efficient in sensor network, there are two approach issues. Most of above research are designed for MAC layer. The work based on the scheduling of transmitting and receiving data by one or multi-hop. In the other way, the power efficient can be approached upper MAC layer. Di Tain et. al. [4] proposed a coverage preserving node scheduling scheme, which can reduce system overall energy consumption, therefore increasing system lifetime by turning off some redundant nodes. Random sleep schedule and coordinate sleep schedule are applied to get energy efficiency in [5]. In [6], they showed to find out the redundant nodes using Voronoi diagram to turn off them. But in this calculation, coordinates of nodes are necessary to know for this reason either GPS or some algorithm is used to know the co-ordinates of the nodes. In PEAS [7], they proposed a robust energy-efficient protocol for the harsh environment. They consider probing range but didn't consider sensing coverage. The network lifetime is scheduling by wakeup round robin in [8]. They proposed an approximate coverage base on triangle calculation and prediction threshold. And then scheduling for nodes which are necessary for coverage but not needed for connectivity, so they use periodically turn on and off for the radio of some nodes among them to save more energy. Their work has simple calculation but the coverage is not guaranteed. Our main contribution in this paper is improved their weak point and build a new protocol for energy balanced sensor network. We propose an energy efficient MAC protocol for wireless sensor networks. In this protocol, nodes balance their energy by going to working state based on their residual energy to keep the nodes alive to assure the coverage throughout the lifetime.

## III. PROPOSED SCHEME

In this paper the main contributions are (a) reducing redundancy by finding out small number of working nodes at a time and (b) sharing responsibility of working state to keep alive most of the nodes to assure coverage anytime throughout the lifetime of the network. The proposed protocol will make schedule for every sensor node based on the coverage guarantee calculation and power balance of network. Our protocol is an enhancement of control function for MAC layer. Extension scheduling functions were included based on MAC functionalities to control states of sensor board. Our protocol works up-layer of MAC layer so it can be dynamically in applying for any MAC protocol. Every sensor can change to "working state" base on two factors: energy level and coverage contribution. They are presented by Wake up scheduling and coverage calculation as following.

Our scheme operates based on some considerations as following: First, nodes are densely deployment. Secondly, nodes know the distance of the neighbors using some light protocol (Beacon node, Radio RSSI, Time difference of arrival (TDoA), Angle of arrival (AoA)).

### A. Wake up scheduling

The protocol uses three messages such as CHECK, REPLY and WAKE UP. It has three states to perform the operation such as SLEEPING, CHECKING and WORKING. We also use three types of message to control sensor nodes. These messages will be sent at the beginning of working state of MAC protocol to assure that all remaining node in-coverage can receive respectively.

**CHECK message:**

A node broadcast this message in the CHECKING state. It contains information of residual energy.

**REPLY message:**

The message is used by working node, it sends REPLY message when it gets CHECK message. It is a flag to indicate the CHECKING node to go to WORKING state and to go to SLEEPING state respectively.

**WAKE UP message:**

This message is sent by a working node when it is going to die for lacking of power.

The state diagram of scheduling process is showed in Figure 1. Sensor nodes will change to new state based on the current energy and coverage contribution. The network will operate in two phases. The first phase is initial coverage. All deployment sensor nodes will be in WORKING state. After they exchange "location information" together, every node will calculate its coverage contribution. If the coverage independent, it will broadcast REPLY messages to inform to neighbor nodes and go to SLEEPING state. After all nodes synchronize together we will have the maximum coverage with the minimum set of sensor.

Then our network will go to second phase, BEGC. Nodes which wakeup from SLEEPING state or WORKING state will change into CHECKING state periodically, step (1), (6). WORKING node will find chance to go SLEEPING mode. On the contrary, SLEEPING node will try exchange duty with current WORKING node. This is the way all sensors keep energy balanced through network. When nodes are in CHECKING state, they will broadcast CHECK message which includes current energy to all its neighbors. If nodes (SLEEPING state) receive CHECK message, they will compare its power strength and the guaranteed coverage. If its power level is higher than that one about 20% and it is full coverage by neighbor nodes, it will broadcast REPLY message and change into WORKING state. After that, node which is in WORKING state will check the coverage contribution after receiving REPLY message. If the coverage is full it will exchange duty and then go to SLEEPING state, step (4). When WORKING node remains less than the threshold to die (5%) energy level it will broadcast WAKEUP message to find new exchange node. Node which received WAKEUP message and in-coverage of sender will change into WORKING sate. During the WORKING state sensor node will operate the sensing duty using one MAC protocol to control data transmission and receiving.

### B. Coverage calculation

Main goal of our sensor network is the full coverage to guarantee all events in interesting area will be supervised. Before scheduling working or sleep for every sensor, our algorithm will select minimum nodes which achieving full coverage. The guaranteed coverage calculation must satisfy three conditions: perimeter-test center-test and distance-test [9]. In fact, this calculation takes some overhead in computation, energy and neighbor's location information. Assume that all sensor nodes have same sensing range $r_s$. The definition of coverage calculation is shown in Figure 2.

The perimeter-test checks whether there are enough coverage neighbors such that all points in the perimeter should be within a sensing range of a coverage neighbor. This is a necessary condition based on the assumption of densely deployed nodes. In Figure 2(a) node $C_0$ is covered perimeter by node $C_1$, $C_2$, $C_3$, and $C_4$.

In case one node passed the perimeter-test but the coverage is not also guaranteed for example in Figure 2(b). The center-test examined whether the center of a sensor's coverage can be covered by at least one of its neighbors. This can be shown in Figure 2(c) (node C2 covered the center of node C0). To satisfy this condition, there will be at least one node in perimeter-test list can cover the center. In Figure 2(c), we have $d(C2, C0) < r_s$.

The third condition is called distance-test. Those coverage neighbors must be close enough to the sensor, so that there may not be uncovered area inside the sensing region. When passing center-test, the best center node in center-test must reach to all perimeter-test nodes to make full coverage. In Figure 2(d) node C2 must satisfy: $d(C2,Ci) < d(C2, C0) + r_s$ where i =1, 3, and 4.

## IV. PERFORMANCE EVALUATION

In this section, we will evaluate the performance of our scheme based on one of the well known MAC protocols for sensor network, S-MAC [10]. And then make comparisons with "Redundancy Reduction Protocol with Sensing Coverage Assurance in Distributed Wireless Sensor Networks" [8]. As above explanation, all sensor nodes will operate on S-MAC when they are in "working state". S-MAC [10] is one of the most well known MAC protocols for energy efficiency in wireless sensor networks. It adopts a periodic listen-sleep cycle to reduce idle listening time. Each node turns off its radio transceiver in a sleep period. It wakes up in a listen period and can communicate with other nodes. This listen period is used for exchanging control packets such as SYNC, RTS, CTS, and ACK packets and data packets. Neighboring nodes form a virtual cluster and all nodes in the same virtual cluster set up a common synchronized listen-sleep schedule.

(1): Sleeping node wakes up and finds exchange node.
(2): Node will broadcast 3 times CHECK message.
(3): Node will goes to Working state if other node wants to exchange.
(4): Node goes back sleeping state
(5): Working node goes to sleep when it found exchange node.
(6): Working node goes to Checking state when it found exchange node.

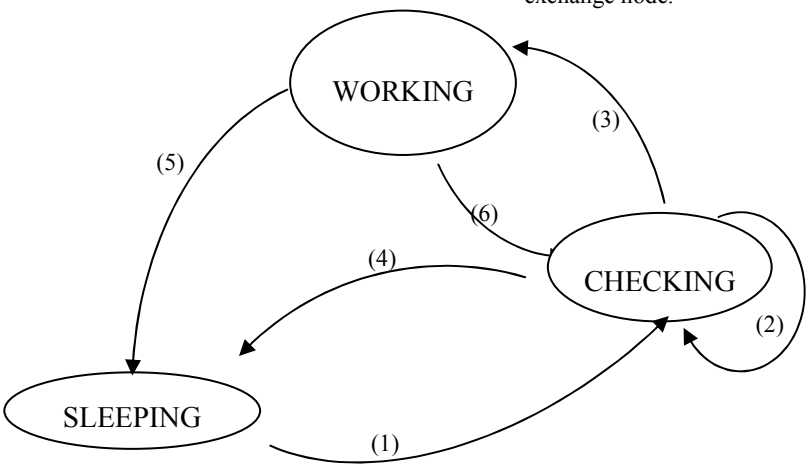



Figure 1. State diagram of proposed protocol

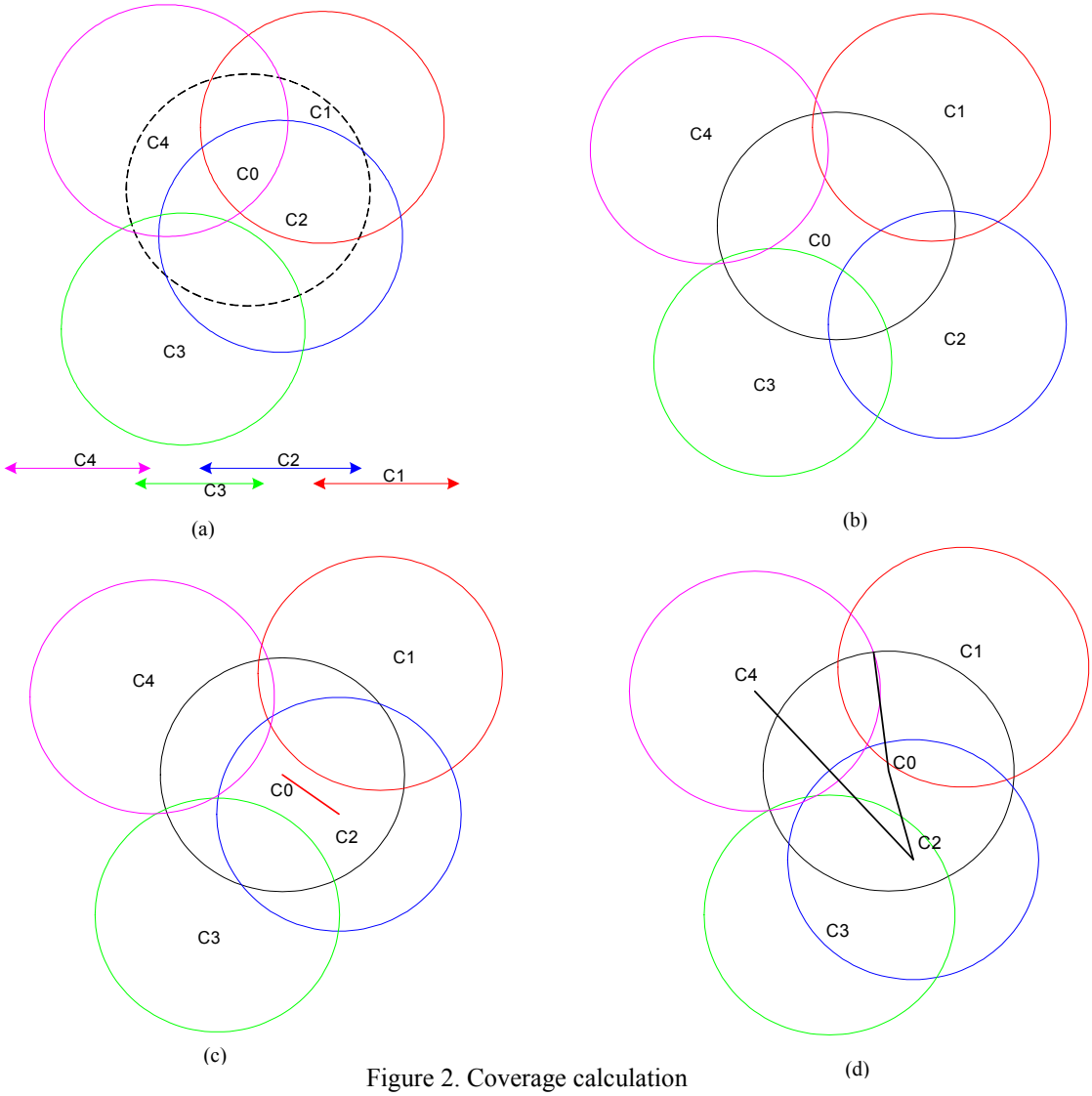



Figure 2. Coverage calculation

Though S-MAC introduces a sleep period to reduce idle listening due to the synchronized scheduling in the virtual cluster, a large number of collisions can occur in heavy traffic loads and dense node deployment situations, resulting in severe energy waste and lower throughput. Moreover, since each node has a fixed duty cycle, S-MAC cannot adapt to the variation of network environment.

We used C++ simulation for evaluation testing the performance of proposed scheme. Node configuration for communication and sensing is giving by table I. All nodes are deployed by random. Figure 3 shows the advantage of proposed scheme over S-MAC with original protocol. Our scheme can increase 40% network lifetime with full guarantee coverage. Number of working nodes for data sensing of network are optimized efficiently in Figure 4. Proposed scheme can remain the minimum number of sensor for best coverage.

Table I: Simulation configuration

| Parameter | Value |
|---|---|
| PHY data rate | 125 Kbps |
| MAC protocol | S-MAC |
| Data packet | 16 Byte |
| Initial power | 10 J |
| Power consumption for transmission | 6 mJ |
| Power consumption for reception | 3 mJ |
| Power consumption for sleep (idle listening) | 5 mJ |
| Transmission range | 10 m |
| Packet arrival rate | 0.1 s |
| Monitor area | 100 $m^2$ |

For evaluating the performance of the new scheme with one of related research, we make the comparison with "Redundancy Reduction Protocol". The protocol in [8] is message based where each node decides whether it will turn on or turn off in a distributed manner using the information in the messages. This protocol uses three messages such as CHECK, REPLY and WAKE UP. It has three states to perform the operation such as SLEEPING, CHECKING and WORKING. We introduced a new parameter λ which represents the percentage of the overlapped sensing area and it can be calculated by using two other parameters such as distance between nodes and sensing range. In the CHECKING state, each node calculates λ for each neighbor working node and comparing the sum of λ values with a predefined threshold value, the node can understand that its sensing range is already covered by other nodes or not. By using some flags and information in the message this protocol maintains a set of small number of nodes and balances the energy consumption among different nodes and keeping the coverage assured. Their scheme based on some assumption conditions as:

1. No node inside the sensing range of one working node can go to working state. So, the minimum distance between two working node is $r_s$.
2. A node calculates λ value for each working neighbor node around it and those working neighbor nodes will be considered whose distance is in between $r_s$ to $1.732r_s$.
3. The communication range of each sensor is twice or more than twice of the sensing range for connectivity throughout the network.

We test the sensing coverage that can be got by random distribution and the coverage after using our protocol and "Redundancy Reduction Protocol". We took the same topology and deploying a different number of nodes.

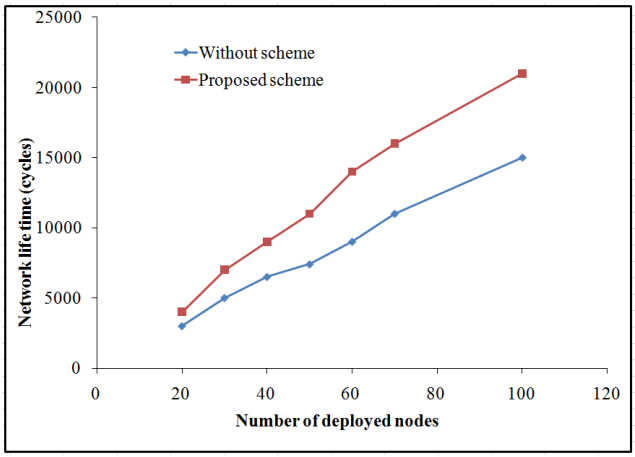



Figure 3. Network lifetime of proposed scheme over S-MAC

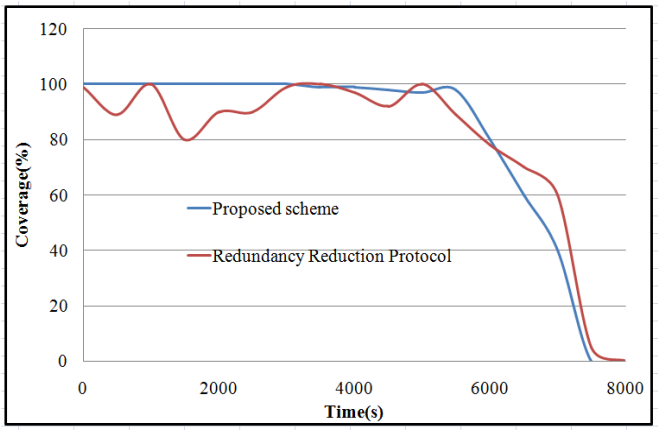



Figure 4. State diagram of proposed protocol

Figure 5 shows the coverage performance of our scheme and Redundancy Reduction Protocol of the first simulation. And Figure 6 the average result of 10 simulations of the coverage expectation through the network lifetime. It can be observed that our scheme can guarantee coverage better than Redundancy Reduction Protocol. This is because that Redundancy Reduction Protocol is based on the approximation of coverage based on λ coverage distribution.

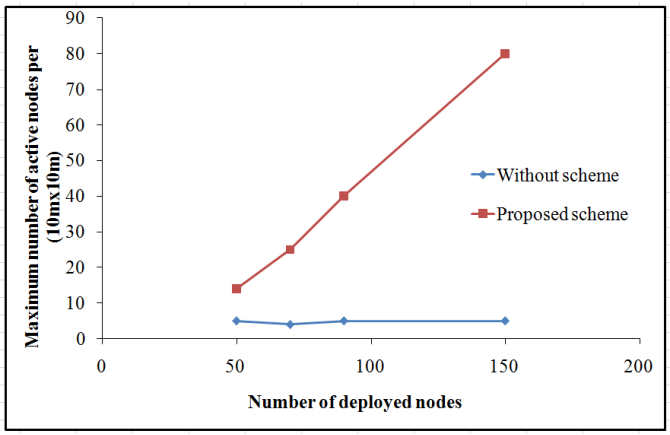



Figure 5. Coverage performance comparison

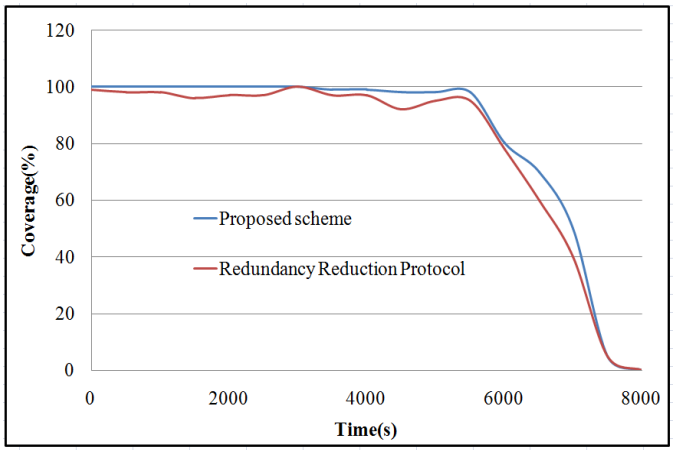



Figure 6. Coverage performance comparison

The protocol shows an excellent result to reduce the number of working nodes but preserving the coverage for dense network.

## V. CONCLUSION

The proposed scheme for energy-balanced and guaranteed coverage scheduling shows an efficient protocol for sensor network. The computation and process can be applied in sensor mote without too much complexity. Only using extended three messages the protocol finds out a small number of working nodes that can preserve the coverage and reduce the redundancy. This protocol shows better performance and keeps an energy balance through network.

## ACKNOWLEDGEMENTS

This research work is funded by the "MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013".

## References

[1] T. V. Dam and K. Lamgemdoen, "An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks," in ACM SenSys 2003, 2003, pp.171–180.

[2] C. Q. P. Lin and X. Wang, "Medium Access Control With A Dynamic Duty Cycle For Sensor Networks," in IEEE WCNC, vol. 3, 2004, pp. 1534–1539.

[3] S. R. T. Zheng and V. Sarangan, "PMAC:An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks," in IEEE IPDPS, 2005.

[4] Di Tian and Nicolas D. Geogranas, "A Coverage Preserving Node Scheduling Scheme for Large Wireless Sensor Networks," WSNA'02 Proc, 2002. pp. 32-41.

[5] Chih-fan Hsin and Mingyan Liu, "Network Coverage using Low Duty-Cycled Sensors: Random and Coordinated Sleep Algorithms," IPSN'04, April, 2004.

[6] Carbunar, A. Grama, J. Vitek, and Q. Carbunar, "Coverage preserving redundancy elimination in sensor networks," IEEE SECON, Oct. 2004.

[7] F. Ye, G. Zhong, J. Cheng, S. Lu, L. Zhang. "PEAS: A Robust Energy Conserving Protocol for Long-lived Sensor Networks," International Conference on Distributed Computing Systems, 2003, pp. 28-37.

[8] Noor Islam, Yeong Min Jang, Sunwoong Choi, Sangjoon Park, Hyeon Park, "Redundancy reduction protocol with sensing coverage assurance in distributed wireless sensor networks," IEEE ISCIT 2009, Sept. 2009.

[9] Nurcan Tezcan and Wenye Wang, "TTS: a two-tiered scheduling mechanism for energy conservation in wireless sensor networks," Int. J. Sensor Networks, Vol. 1, Nos. 3/4, 2006.

[10] J. H. W. Ye and D. Estrin, "Medium Access Control With Coordinated Adaptive Sleeping for Wireless Sensor Networks," IEEE/ACM Transaction on Networking, vol. 12, no. 3, June. 2004, pp. 493–506.

[11] Weihua Guo , Zhaoyu Liu, Guangbin Wu, "Poster Abstract: An Energy-Balanced Transmission Scheme for Sensor Networks," ACM 1-58813-707-9/03/0011.

[12] Cardei, M.; Thai, M.T.; Yingshu Li; Weili Wu," Energy-efficient target coverage in Wireless Sensor Network," IEE INFOCOM 2005.

[13] V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava. "Energy Aware Wireless Microsensor Networks," IEEE Signal Processing Magazine, Vol. 19, March 2002, pp. 40-50.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Realization of Low-Pass Filters From Arbitrarily Designed Nonuniform EBG Structures

S. M. Shakil Hassan[1,*], Mohammad Nurunnabi Mollah[2], S. M. Anayetullah[1], Khandkar Raihan Hossain[2], Md. Mahfuz Ahmed[1], Md. Abu-Al Shufian[1], and Md. Mehedi Hasan[1]
[1]Dept. of Electrical and Electronic Engineering, Eastern University, Bangladesh
[2]Khulna University of Engineering and Technology, Bangladesh
[*] E-mail: shakil.bq@gmail.com

*Abstract*—**Two patterns of electromagnetic bandgap structure (EBGS) have been proposed – one is dumbbell shape EBGS having circular slots as bigger slots and another one is rectangular EBGS. A sequence of arbitrary coefficients has been proposed to make nonuniform EBGSs that results better performance than that of nonuniform EBGSs realized by binomiall coefficients. A potential technique has also been imposed upon the nonuniform dumbbell shape EBG structure that provides more improved scattering parameter performance. An arbitrary nonuniform rectangular patterned EBGS has been investigated too that perform as a low-pass filter. The performances of both proposed designs have been compared to the relevant uniform and binomially distributed nonuniform EBG assisted designs. The results, of arbitrary nonuniform designs, are superior to the conventional form of noninform EBG elements.**

***Keywords—Electromagnetic bandgap structure, circular slotted dumbbell shape EBGS, nonuniform EBGS, rectangular EBG structure, low-pass filter.***

## I. INTRODUCTION

The use of electromagnetic bandgap (EBG) structures in designing filters for microwave and millimeter wave applications is increasing due to its compact size and capability of enhancing system performance compared to other existing filters. Electromagnetic bandgap structures (EBGSs) were formerly termed as photonic bandgap (PBG) structures [1], [2]. EBG structures have found potential applications not only in filters but also in antennas, phased array, amplifiers, waveguides, and many other microwave devices [2]-[4]. Hence, EBG engineering gained much attraction in the field of microwave technology and EBG structures have become very popular periodic structures. Extensive research has been conducted to optimize the size of EBG assisted microstrip filter and obtain better filtering performance. EBG, however, can be realized by etching, drilling or cutting of a deflected pattern from metallic ground plane in periodic manner. Periodic structures at microstrip transmission provide rejection property of certain frequencies of microwave and millimeter wave [5], [6]. Band rejection property, in fact, depends on too many designing parameters including size, shape and volume fraction of EBG as well as some parameters associated with the substrate properties such as dielectric constant and thickness. The use of uniform structure that is where each EBG element of the lattice is equal in size and shape is more common than non-uniform structures. However, recently it is reported that better filtering performance can be achieved by the use of nonuniform designs [7]-[10]. Various shapes of EBG structures have already been reported; among them dumbbell shape EBGS is most popular and has found wide range of applications [11]-[13]. Dumbbell shape EBGS is also known by defected ground structure (DGS).

In this paper, we have proposed a new pattern of dumbbell shape EBGS and have reported some investigations over nonuniform EBG structure assisted microstrip transmission line (T-line). Proposed dumbbell shape EBGS shown in Fig. 1(a) that consists of cicular slots instead of square slots of conventional dumbbell shape EBGS or DGS shown in Fig. 1(b). We have investigated a nonuniform design using cicular slotted dumbbell shape EBGS, where circular bigger slots are diminished by an arbitrary sequence and using binomial

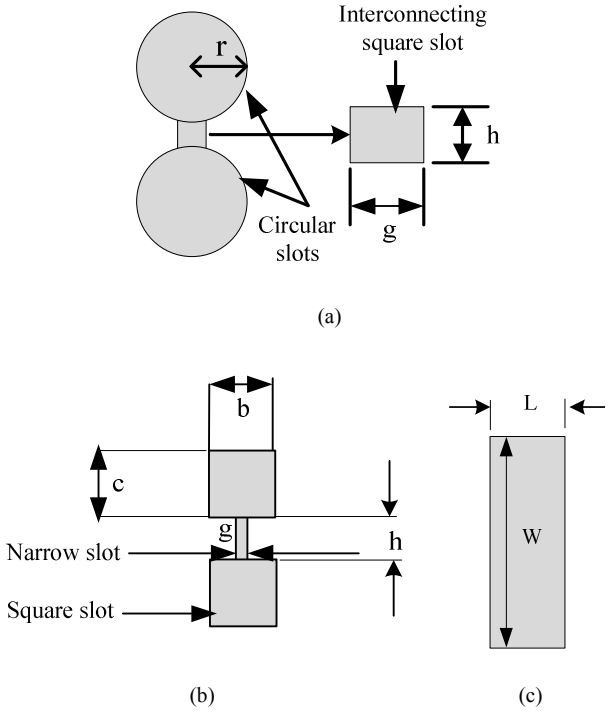



Figure 1. (a) conventional square slotted dumbbell shape EBG strucutre, (b) proposed circular slotted dumbbell shape EBG structure, and (c) proposed rectangular EBG structure.

distribution, on the other hand the narrow slots are rectangular having constant width while the lengths are extended from center to both sides using the arbitrary sequence as well. Results of above mentioned designs are compared with the results of corresponding uniform and binomially distributed designs. Also, we have proposed another pattern called rectangular EBG structure shown in Fig. 1(c) and we have investigated a nonuniform rectangular EBG assisted design that has shown very well formed low-pass performance.

## II. Theory of Binomial Distribution

Binomial distribution is a potential theory that is used to make electromagnetic bandgap structures nonuniform. However, it is observed that nonuniform EBGSs have few advantages over uniform structures [14].

The Pascal's triangle, shown in the Fig. 2, is found from binomial polynomial [14], [15]. In the Pascal's triangle *n* represents the number of the co-efficient that can be akin to the number of EBG elements in designing nonuniform EBG elements and the coefficients of the belonging *n* represents the relative amplitude of the EBG.

Calculation of amplitude coefficient is done in two steps. First step is the selection of number of EBG element that is n. In second step all the coefficients of $n^{th}$ row is divided by the highest value of the coefficient in that row to determine the coefficients of the radius of circular shaped bigger slots of dumbbell shape EBG elements. Hence the relative amplitudes of the periodic structures are proportional to relative values of the polynomials.

## III. EBGS Assisted T-Line

We have chosen dielectric constant ($\varepsilon_0$) and height of the substrate for all the designs are 2.45 and 31 mil or 0.7874 mm respectively. The width of the 50 Ω microstrip transmission line has been calculated using PCAAD is 2.2642 mm. However, architecture of a unit circular slotted dumbbell shape EBGS and a unit conventional dumbbell shape EBG are shown in Fig. 1(a) and in Fig. 1(b) with dimensions.

### A. Designs

In this section we will be discussing the designing specifications of our proposed arbitrarily nonuniform circular slotted dumbbell shape EBG structure with modified narrow slots. Also a binomially tapered design and a uniform design using the proposed standard (largest) circular slotted dumbbell shape EBG will be depicted.

n=1 1
n=2 1 1
n=3 1 2 1
n=4 1 3 3 1
n=5 1 4 6 4 1
n=6 1 5 10 10 5 1
n=7 1 6 15 20 15 6 1
n=8 1 7 21 35 35 21 7 1
n=9 1 8 28 56 70 56 28 8 1

Figure 2. Pascal's triangle.

*a) Proposed arbitrary design:* In this design both the bigger slot and narrow slot is made non- uniform. Bigger slots are circular shaped and narrow slots are rectangular shaped. Coefficients for tapering the bigger slots are 0.25, 0.50, 0.75, 1, 1, 0.75, 0.50, 0.25 and coefficients for modifing narrow slots are 4, 3, 2, 1, 1, 2, 3, 4 consequently from the left most EBG element. The radius of the circular slot and length of the narrow slot of the reference EBG element (biggest EBGS) are 3.651 mm and 2 mm respectively. The width of the narrow slot and inner element spacing are 1. 5 mm and 10.43 mm respectively and are kept unchanged shown in Fig. 3(a).

*b) Conventional binomially tapered design:* This time we have calculated tapering coefficients using pascal's triangle for $n$ = 8; coefficeiants are: 0.03, 0.2, 0.6, 1, 1, 0.6, 0.2, 0.03. The reference radius i.e radius of the biggest circular slot is 3.651 mm. The narrow slot is kept constant; $g$ = 2 mm and $h$ = 1.5 mm shown in Fig. 3(b).

*c) Uniform design:* Fig. 3(c) shows all the EBG elements are of same size: radius of circular slot is 3.651 mm and $g$ = 2 mm and $h$ = 1.5 mm of narrow slot.

Besides earlier designs more two designs have also been investigated to confirm the effect of modifying narrow slots of dumbbell shape EBGS. We have designed a new microstrip line using the coefficients of proposed design to taper only the circular slots keeping narrow slots identical to the reference EBG structure. On the other hand we have made another tapered design using binomial coefficients with modified narrow slots using the coefficients used in the proposed design. Basically, results of these designs have been compared to the corresponding earlier briefed designs to investigate the effect of narrow slot modification.

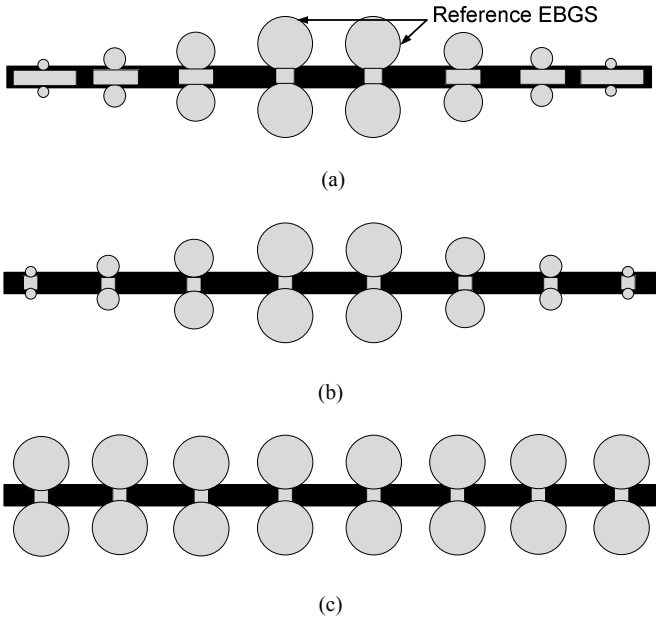
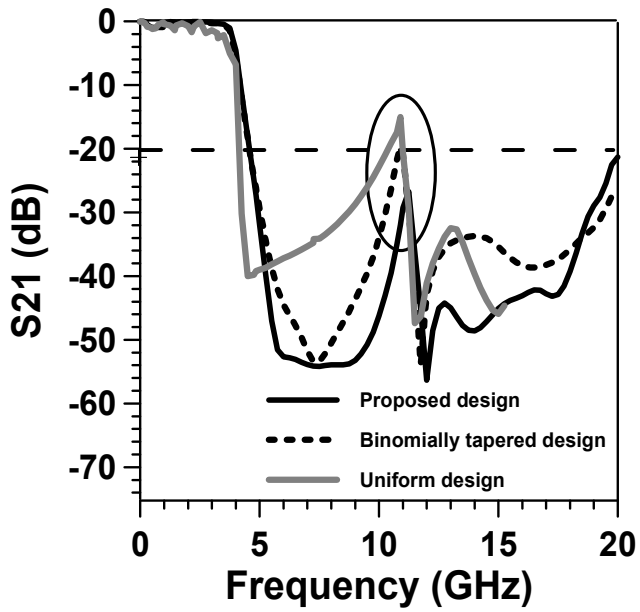



Figure 3. (a) proposed arbitrary design with modified narrow slots, (b) binomially distributed proposed dumbbell shape EBGS, and (c) design having uniform proposed dumbbell shape EBGS.

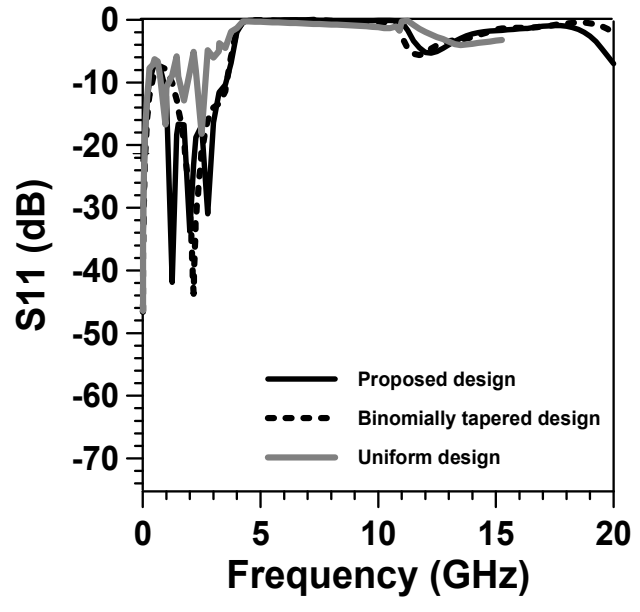
*B. Performances*

Scattering parameter (s-parameter) performances of the designs are shown in Fig. 4 and 5. Fig. 4 shows that the average -3dB cutoff frequency and starting -20dB band rejection of insertion loss curve (i.e. $S_{21}$ curve) are almost same for proposed arbitrary design, binomial design, and uniform design and that are 3.89 GHz and 4.57 GHz respectively. However, the supremacy of modifying narrow slots according to our proposed manner can be founds by the ascended peak of $S_{21}$ at approximately 11 GHz frequency; where the peak climbed beyond -20dB except for our proposed design. Very low ripples are observed using the proposed technique – maximum ripple height is observed is -0.9 dB. Therefore, proposed design has shown very nice low-pass performance.

Fig. 5 verifies the earlier mentioned effect of modified narrow slots by comparing the results of modified narrow slots and uniform smallest narrow slot of similar type of tapered design. Fig. 5(a) shows that modified narrow slot has descended the pick of $S_{21}$ curve than the uniform narrow slots

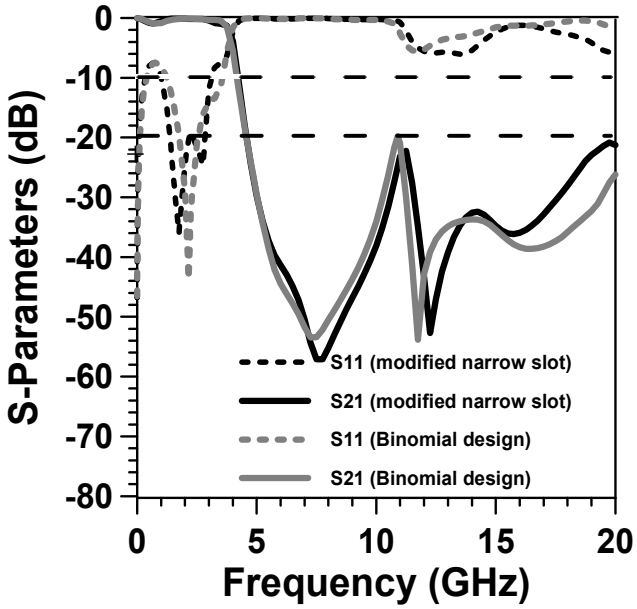



(a)

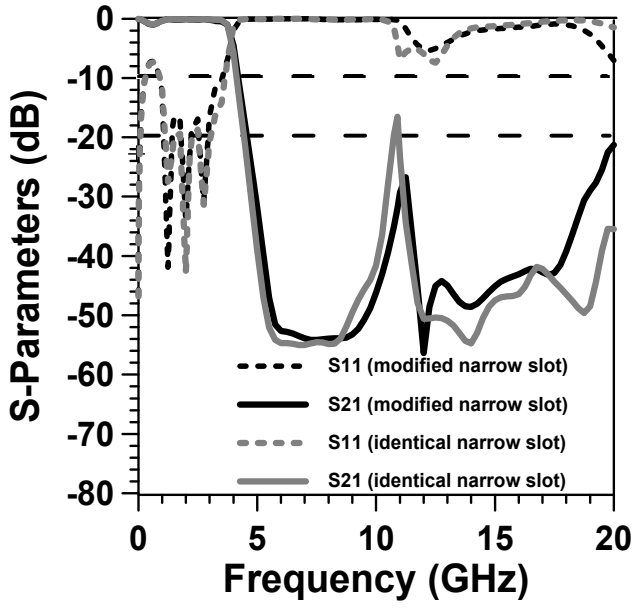



(b)

Figure 4. S-parameter performance of proposed arbitrary desing, binolially tapared design, and uniform design (a) $S_{21}$ and (b) $S_{11}$.




(a)




(b)

Figure 5. Comparison of s-parameter performance between identical and modified narrow slots of (a) binomially distributed circular slotted dumbbell shape and (b) proposed tapered design usng arbitrary coefficients.

of binomially nonuniformed design. Fig. 5(b) also verifies the same effect of modified narrow slot in case of tapered design using proposed arbitrary sequence. From the Fig. 5(a) and (b), moreover, it is visible that design tapered by the proposed arbitrary coefficients results better IL parameter that the binomial coefficients.

The proper reasoning of such better performance due to reversely increased narrow slots can be demonstrated by the equivalent circuit of dumbbell shape EBGS. Dal Ahn et al. [16] had reported for conventional square slotted dumbbell shape EBGS that two square slots and one connecting narrow slot correspond to the equivalently added inductance and capacitance respectively with characteristic impedance of T-line that yield a parallel *L-C* equivalent circuit [11], [16]. Akin to conventional dumbbell shape, similar equivalent circuit can be extracted from our proposed circular slotted dumbbell shape; but its modified narrow slots modifies the value of capacitances. Increased value of capacitance, indeed, is the main reason to give rise of better and smoother IL performance at the stopband frequency.

## IV. RECTEANGULAR EBGS

Rectangular electromagnetic bandgap structure is another kind of newly introduced EBG structure that is being proposed by investigating it into an arbitrarily nonuniformed design; a unit rectangular EBG element has been shown in Fig. 1(c).

### *A. Designs*

We have designed three designs: Fig. 6(a) is showing our proposed design, Fig. 6(b) is the binomially tapered design, and Fig. 6(c) is the design that has only smaller EBGSs of our proposed design. We have kept all the designing specifications same and the sizes of the rectangular EBGSs are $w$ = 30 mm for the center element and $w$ = 14 mm for the others, and $L$= 4 mm for all the EBG elements shown in Fig. 6(a). In case of nonuniform design we have chosen amplitude coefficients from the Pascal's triangle and the reference maximum value of the width of rectangular EBGS is 30 mm. Our proposed design has also resulted as a low-pass filter.

### *B. Performances*

The scattering parameter performances are shown in the Fig. 7. There we have shown the result of our proposed design compared to the binomially tapered design; where it is seen that the proposed results has shown a good LPF performance, whereas other one has shown somewhat stopband. However, -3dB cutoff frequency, -20dB starting frequency of LPF performance, and -10dB passband width of return loss curve ($S_{11}$) are 2.65 GHz, 4.225 GHz and 2.474 GHz respectively. The design where biggest EBG at the center is eliminated and smaller EBGSs are kept at same position also results a LPF with much larger ripple heights. In essence, smaller EBGSs have dominated the result. Binomial distribution, in contradiction, is not a potential technique to better off the result for this design.

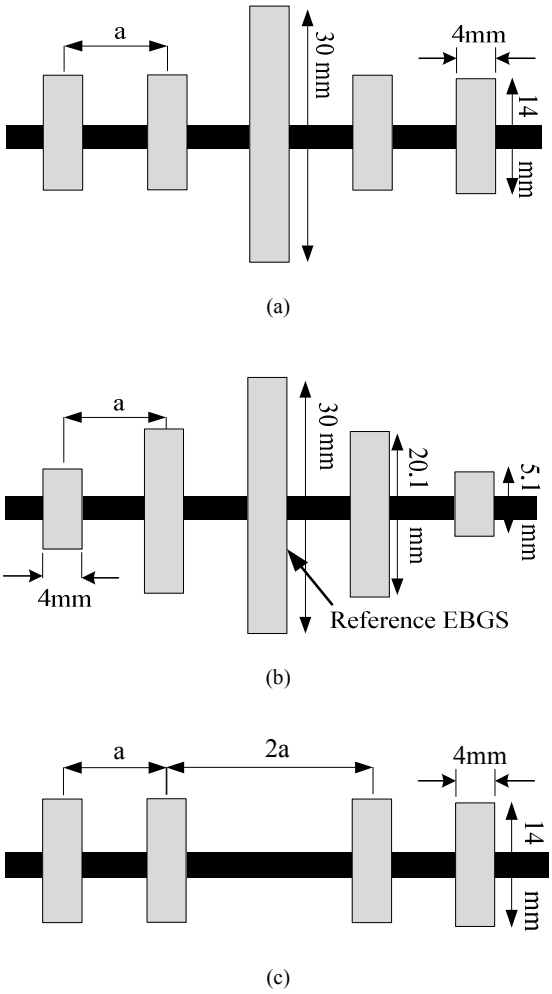
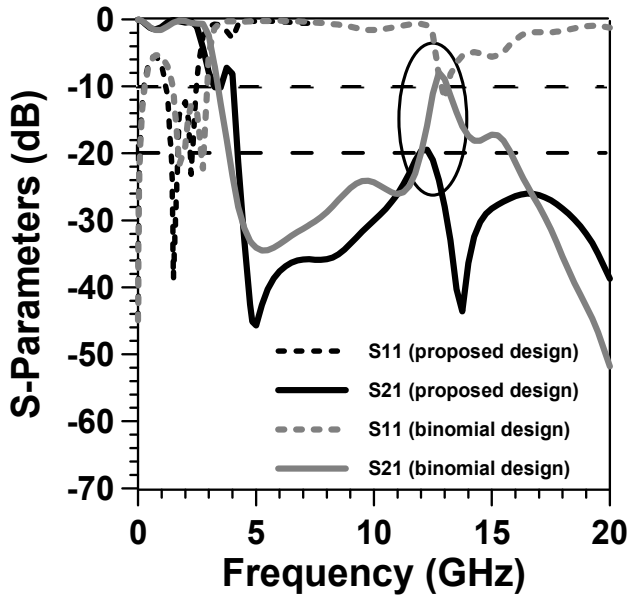


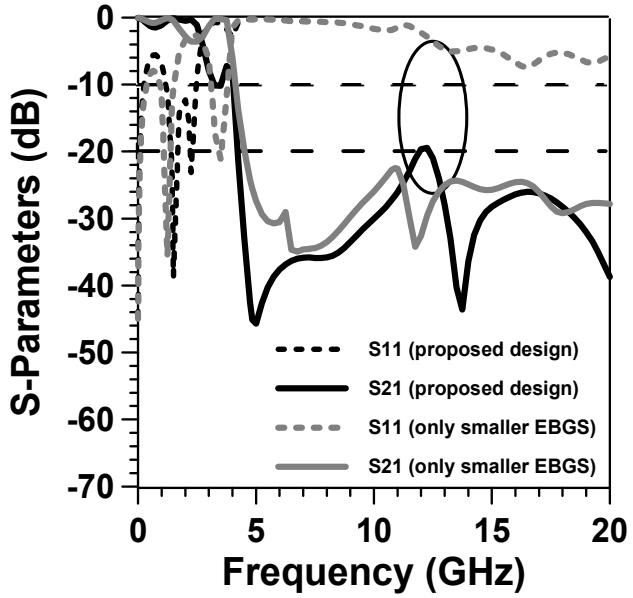

Figure 6. (a) proposed arbitrarily nouniform rectangular EBG structure (b) binomially distributed nonuniform rectangular EBGS and (c) only smaller EBG elements of proposed design.

## V. SIMULATED VS MEASURED RESULTS

We have used Zeland ie3d simulating software to design and simulate them. The results are agreed to be in well agreement with the measured results, have reported by researcher with required justification [8], [14]. In [11] a validation has also been provided by comparing a simulated result found by the EM software with a reported experimented result. The measured results may be sometimes erroneous or mismatched due to over etching. In this, however, we have designed a reported nonuniform dumbbell shape EBGS assisted design in [12] and simulated the result using Zeland ie3d software. The design and the result are shown in Fig. 8. In this case the dielectric constant and thickness of the substrate are 10 and 1.575 mm respectively. From the Fig. 8(b) it is seen that the simulation result is very congruent to the reported measured result. Simulation results of this paper, therefore, hopefully will find well agreement of the measured results.




(a)

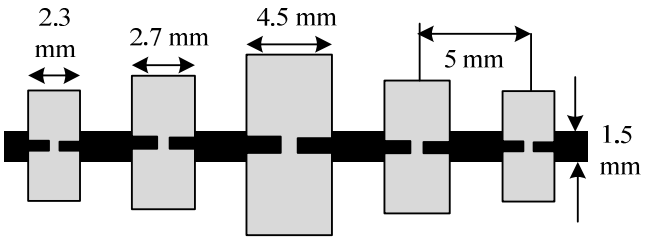



(b)

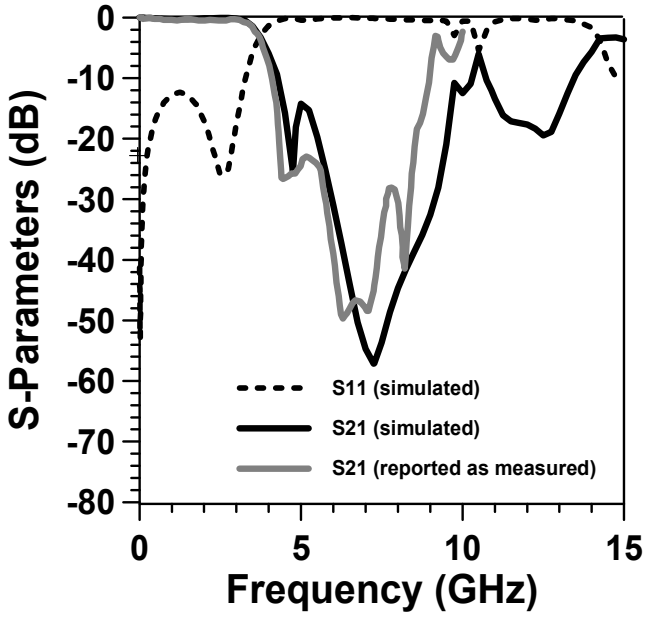
Figure 7. (a) comparison of proposed arbitrary design and binomial desin and (b) result of smaller EBG elements in contrast with proposed design.

## VI. CONCLUSION

Novel EBG patterns of circular bigger slotted dumbbell shape EBG structure and rectangular EBG structure have been introduced with its architecture. Two arbitrarily tapered designs using them have been reported with attractive low-pass performance. Through this work, we have shown the superiority of arbitrary coefficients of tapering over coefficients of binomial distribution. However, a potential




(a)




(b)

Figure 8. (a) a design reported in [12] and (b) comparison of measured $S_{21}$ performance (reported in [12]) with simulated performance acheived by using Zeland ie3d.

technique of modifying narrow slots has been proposed too. This technique has significant effect on s-parameter performance. It is observed that tapered bigger slots usually reduce the ripples and the modified narrow slots provide additional etched area, if it is increased as size of the bigger slots are reduced, that makes the results finer. It has been observed from proposed rectangular EBGS assisted design that there could be domination of some EBGSs of particular size in the performance. Therefore, arbitrarily tapered EBGS assisted design sometimes finds very effective use with greater performance.

## REFERENCES

[1] F. Falcone, T. Lopetegi, and M. Sorolla, "1-D and 2-D photonic bandgap microstrip structures," Microw. Opt. Technol. Lett., vol. 22, no. 6, pp. 411–412, Sep. 1999.

[2] V. Radisic, Y. Qian, and T. Itoh, "Broad-band power amplifier using dielectric photonic bandgap structure," IEEE Microw. Guided Wave Lett., vol. 8, no. 1, pp. 13–14, Jan. 1998.

[3] Nemai C Karmaker and Mohammad N Mollah, "Potential applications of PBG engineered structures in microwave engineering - part one," Microwave Journal, vol. 47, no. 7, pp. 22-44, July 2004.

[4] R. Gonzalo, P. D. Maagt, and M. Sorolla, "Enhanced patch-antenna performance by suppressing surface waves using photonic-bandgap substrates," IEEE Trans. Microw. Theory Techn., vol. 47, no. 11, pp.2131–2138, Nov. 1999.

[5] V. Radisic, Y. Qian, R. Coccioli, and T. Itoh, "Novel 2-D photonic bandgap structures for microstrip lines," IEEE Microwave and guided wave lett., vol. 8 no. 2, pp.69-71, Feb. 1998.

[6] S. Y. Huang, and Y. H. Lee, "Tapered dual-plane compact electromagnetic bandgap microstrip filter structures," IEEE Trans. Microw. Theory Techn., vol. 53, no. 9, pp.2656–2664, Sep. 2005.

[7] N. C. Karmakar, "Theoretical investigations into binomial distributions of photonic bandgaps in microstripline structures," Microw. Opt. Technol. Lett., vol.33, no. 3, pp.191–196, Jan. 2002.

[8] N. C. Karmakar, and M. N. Mollah, "Investigations into nonuniform photonic-bandgap microstripline low-pass filters," IEEE Trans. Microwave Theory Tech., Vol. 51, no. 2, pp. 564-572, Feb. 2003.

[9] S. Y. Huang, and Y. H. Lee, "Tapered dual-plane compact electromagnetic bandgap microstrip filter structures," IEEE Trans. Microw. Theory Techn., vol. 53, no. 9, pp.2656–2664, Sep. 2005.

[10] S. M. S. Hassan, S. M. Anayetullah, and M. N. Mollah, "Sinusoidal appearance of nonuniform dumbbell shape EBGSs in microstrip transmission line," in *Proc. Int. Conf. on Electrical Information and Communication Technology*, Dec. 2013, Khulna, Bangladesh.

[11] S. M. S. Hassan, M. N. Mollah, M. A. Rashid, N. H. Ramly, and M. Othman, "Dumbbell shape EBGS structure – worth to EBG assisted microwave filter designing," *2012 IEEE Asia-Pacific Conf. on Applied Electromagnetics*, Dec. 2012, Melaka, Malaysia, pp. 1-5.

[12] H. W. Liu, Z. F. Li, X. W. Sun, and J. F. Mao, "An improved 1-D periodic defected ground structure for microstrip line," IEEE Microwave and wireless components lett., vol. 14 no. 4, pp. 180-182, Apr. 2004.

[13] J. S. Lim, C. S. Kim, D. Ahn, Y. C. Jeong, and S. Nam, "Design of lowpass filters using defected ground structure," IEEE Trans. On Microwave Theory and Techniques, vol. 53, no. 8, pp. 2539-2545, Aug. 2005.

[14] N. C. Karmakar, M. N. Mollah, and S. K. Padhi, "Improved performance of a non-uniform ring patterned PBG assisted microstrip line," IEEE International Antennas and Propagation Symposium, Jun. 2002, vol. 2, pp.848–851, Texas, USA.

[15] C. A. Balanis, Antenna Theory Analysis and Design, 2 ed. New York: Wiley, 1997.

[16] D. Ahn, J.S. Park, C.S. Kim, J. Kim, Y. Qian and T. Itoh, "A design of the lowpass filter using the novel microstrip defected ground structure," *IEEE Trans. On Microwave Theory and Techniques*, vol. 49, no. 1, pp. 86-93, May 2001.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Gain and SAR Improvement of a Conventional Patch Antenna using a Novel Pi-shaped DNG Metamaterial

Anik Mallik*, Sanjoy Kundu and Md. Osman Goni
Department of Electronics & Communication Engineering (ECE)
Khulna University of Engineering & Technology (KUET)
Khulna, Bangladesh
*E-mail: anikmallik@yahoo.com

*Abstract*— **Conventional microstrip patch antennas have several advantages such as low profile, light weight and low cost over the others, but lag behind in some criterion such as low gain and high SAR, where using metamaterials can be a possible solution. In this paper, a, antenna system is proposed where an S-shaped planar antenna, operating at frequencies of 11 GHz and 13 GHz, uses a Pi-shaped double negative metamaterial array, designed in CST, to improve its gain and SAR. In this system, metamaterial array is applied not as an antenna, but as an associating device for the planar antenna. This array can be implemented on a printed circuit board while maintaining a specific distance from the patch antenna. The gain of the antenna system is enhanced by 6.063 dB at 11 GHz (Increase of 1088.5%) and 3.77 dB at 13 GHz (Increase of 83.4%). The metamaterial array is placed between the antenna and a human-head model which clearly reduces the SAR by 1.379 W/kg at 11 GHz (Decrease of 99.99%) and 1.774 W/kg at 13 GHz (Decrease of 99.66%). This system is applicable for any X and $K_u$ band applications such as satellite and Radar communications, where high gain and directivity is required. The negligible SAR value announces its capability of being used in any portable communication device where human-safety is a major concern.**

***Keywords—Double negative metamaterial, sub-wavelength structure, microstrip patch antenna, specific absorption rate, farfield gain.***

## I. INTRODUCTION

In various wireless communications systems, a microstrip patch array is greatly desired and widely used as it offers several advantages, such as low profile, light weight, and low cost, etc. [1]. A conventional microstrip patch antenna is usually mounted on a substrate which is backed by a conducting ground plane. The main drawbacks of a patch antenna are low gain, high SAR and narrow bandwidth, which is a great concern in communication technology. Recent theoretical studies and experimentations have showed that metamaterials provide an alternate design methodology to improve the performance of antenna system [2].

By definition, metamaterials are artificially engineered materials having some properties not found in nature. One important feature of these materials is that these materials gain their properties from their structures rather than chemical compositions. In 1968, Victor Veselago theoretically showed that materials with simultaneous negative permittivity ($\varepsilon<0$) and permeability ($\mu<0$) had some different properties as compared to ordinary materials. The peculiar properties of these materials were that the energy flow dictated by Poynting vector was in the opposite direction of the wave vector, negative refractive index, negative Vavilov–Cerenkov effect etc. [3]. A composite medium, based on a periodic array of inter-spaced conducting nonmagnetic Split Ring Resonators (SRRs) and continuous wires, works as a double negative material in a frequency range in the microwave regime, where Doppler effect, Cherenkov radiation, Snell's law are inverted [4]. In recent times, it has been easier to realize these novel electromagnetic properties by micro-structuring a material on a scale much less than the wavelength, which are called sub-wavelength structures [5].

Being the most important and promising device in this era of wireless communication, antennas have to ensure some safety of human body from its electromagnetic radiation. For two important reasons, it has been important to evaluate the human body interaction with such personal communications devices: 1) human health and safety concerns, 2) new product design and validation [6]. From the beginning of the year 1990, SAR calculation from the coupling between antennas and the users' body has snitched huge attention and achieved a significant priority [7], where this coupling, located at the antenna near-field zone, affects the antenna parameters and microwave energy is absorbed by the user's body [8]. On the other hand, modern communication system has now reached a level so high, that the antenna-human interaction should be thoroughly investigated in a wide range of probable scenarios, for portable devices and other communication units [9]. The maximum exposure of human tissues to electromagnetic fields has been defined in terms of Specific Absorption Rate (SAR), the value of which dictates the quantity of exposure to the electromagnetic waves. In Europe, the maximum allowed SAR (averaged over 10 g of tissue) is 2 W/kg (in case of mobile equipments) [10]. Again, the ICNIRP has restricted the maximum SAR value (averaged over 10 g of tissue) for the frequency range 10 MHz to 10 GHz as 10 W/kg (Occupational exposure) and 2 W/kg (General Public exposure) [9].

In 2004, Kivekas et. al. investigated that the antenna efficiency, bandwidth, and SAR is a function of a device's armature-associated parameters, such as length, thickness, width, and partition from the phantom [11]. Concerning the health issues and antenna performance it has been customary to

calculate the specific absorption rate (SAR) for any designed antennas especially which involves human interactions and devising a way to reduce it to a level that it would not affect the human body. Finite-difference time-domain (FDTD) method is extensively used for solving the electromagnetic problem and bio-heat equation for thermal modeling [12]. Different methods have been proposed over the last 20 years to reduce the SAR produced by emissions from antennas to levels below the current maximum exposure levels of the international standards. The ferrite sheet has proved to be a good material for reducing the SAR values [13], [14]. Recently, metamaterials, including electromagnetic bandgap (EBG) structures [15] are being used for this purpose. Many are interested in metamaterials with split ring resonator (SRR) structure that was proposed to reduce the SAR value [16]. In the year 2009, Manapati and Kshetrimayum showed that the SAR of an antenna system can be reduced by placing the metamaterial between the antenna and the human-body (head, in that case) [17].

In this paper, a novel Pi-shaped double negative (DNG) metamaterial has been designed on the basis of sub-wavelength structures, which works as an associating device with a conventional S-shaped patch antenna. The setup illustrated in this paper is designed in CST Microwave Studio 2011 [18], which can enhance the gain and reduce SAR efficiently from antenna radiation. A human head model is considered and the designed Pi-shaped metamaterial array is placed between the head model and the patch antenna for a specific spacing and the results have been analyzed. The constitutive parameters are extracted from the calculated S-parameter values with direct retrieval method as are done in [19]. The metamaterial array can be implanted on printed circuit boards with the distance maintained from the antenna.

## II. DESIGN & STRUCTURES

### A. S-shaped Planar Antenna

An S-shaped microstrip patch antenna is shown in Fig. 1. In this design, silicon is used as substrate having dimension of 30×30×1 mm$^3$. An S-shaped patch is placed over the substrate which is made of copper. This patch has an equal length and width of 20 mm and height of 1 mm. The width of the wire is 2 mm. Then, a PEC (Perfect Electric Conductor) is used as the ground plane whose dimension is 30×30×0.5 mm$^3$. The feeding line is fetched from the middle of the structure as co-axial feed. Its outer cylinder has a radius of 1.5 mm and the inner cylinder has radius of 0.5 mm. A wave-guide port is used to excite the antenna covering the whole feeding point.

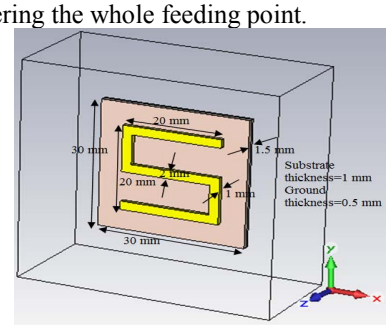



Figure 1. S-shaped planar antenna structure

### B. S-shaped Planar Antenna with Head model

The patch antenna, shown in Fig. 1, is placed in front of a human-head model. Fig. 2 shows the structure where the distance from head to antenna is 20 mm.

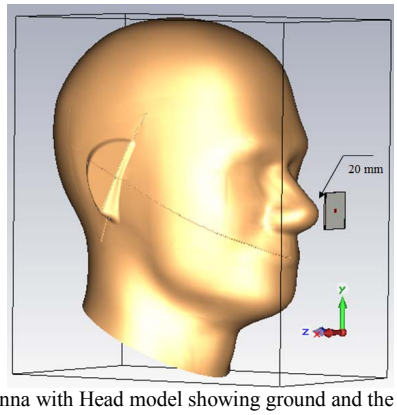



Figure 2. Antenna with Head model showing ground and the feeding line

### C. Pi-shaped metamaterial unit cell

A Pi-shaped metamaterial unit cell is designed as shown in Fig. 3. Here, FR-4 is used as the substrate which has a dimension of 10×10×0.5 mm$^3$. Then, silver is used to create the Pi-shape over the substrate, which has lengths of 8 mm and 7 mm, width of 1 mm and height of 0.5 mm.

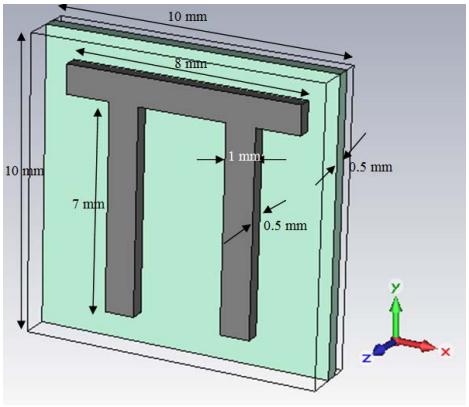



Figure 3. Pi-shaped metamaterial unit cell

### D. 3×3 array of the Pi-shaped metamaterial unit cell

The unit cell of the Pi-shaped metamaterial (shown in Fig. 3) is then used to create a 3×3 array. Fig. 4 shows the 3×3 array structure.

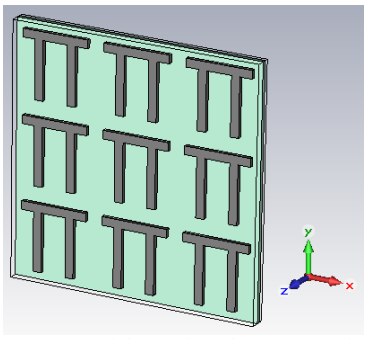


Figure 4. 3×3 array of the Pi-shaped metamaterial unit cell

### E. *Antenna with metamaterial and the head model*

The metamaterial array is placed between the antenna and the human-head model as shown Fig. 5. The distance between the antenna and the metamaterial array is 5 mm and the head model is 15 mm away from the metamaterial array.

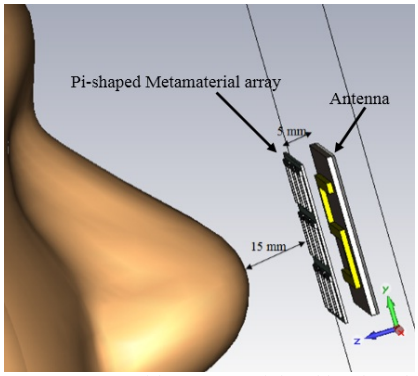



Figure 5. Antenna with metamaterial and head model

## III. RESULTS & DISCUSSIONS

### A. *Constitutive Paramters*

The S-parameter values from CST are used to extract the constitutive parameters (i.e. permittivity, permeability and refractive index) with the help of direct retrieval method.

*1) Metamaterial Unit Cell:* Fig. 6 shows the frequency vs constitutive parameters graph for the metamaterial unit cell. It can easily be seen that a DNG metamaterial has been successfully designed which has a resonant frequency at 13.11 GHz.

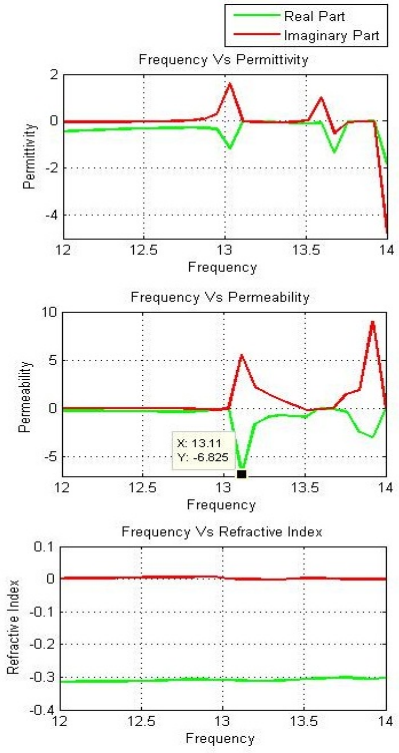



Figure 6. Frequency Vs Constitutive parameters (for unit cell)

*2) 3×3 Array of the Unict Cell:* Fig. 7 shows the frequency vs constitutive parameters graph for the 3×3 array of metamaterial unit cell. It has a resonant frequency at 6.323 GHz.

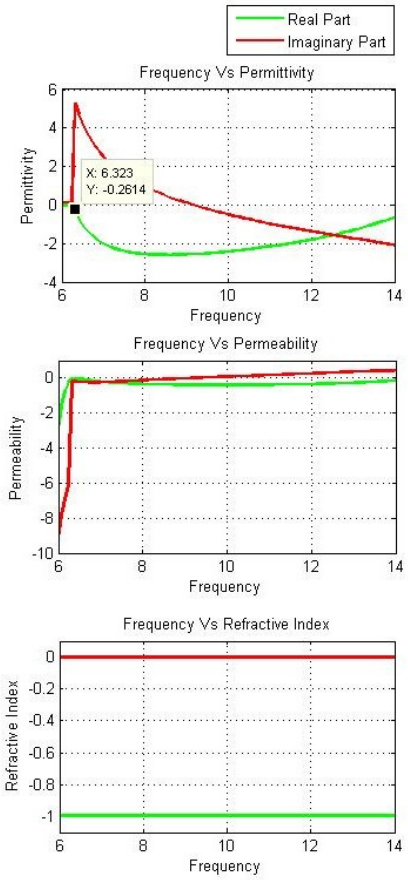



Figure 7. Frequency Vs Constitutive parameters (for 3×3 array)

### B. *Farfield Gain Pattern*

Farfield gain patterns are calculated in CST for each configuration of the antenna (antenna alone, with human-head model and with head using metamaterial).

*1) S-shaped antenna:* The farfield gain of the S-shaped planar antenna is shown in 3D view in Fig. 8 (at 11 GHz) and Fig. 9 (at 13 GHz). The maximum farfield gain of the antenna is 1.05 dB at 11 GHz and 5.11 dB at 13 GHz.

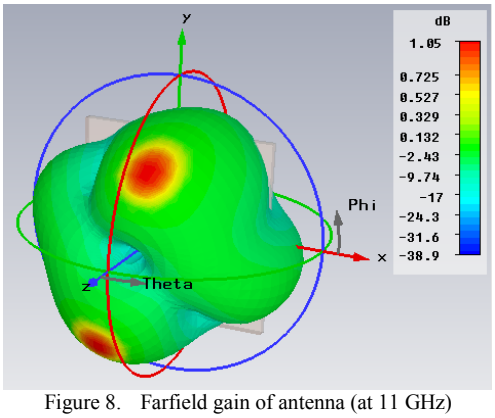



Figure 8. Farfield gain of antenna (at 11 GHz)

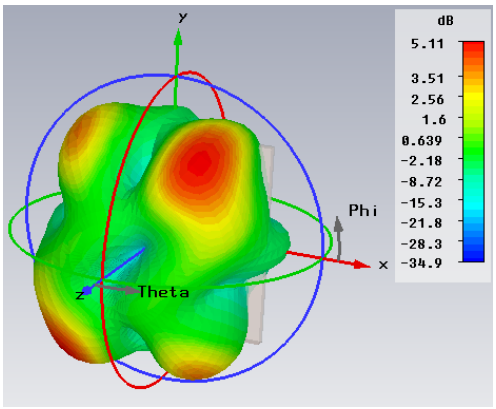



Figure 9. Farfield gain of antenna (at 13 GHz)

*2) Antenna with human-head model:* The maximum farfield gain of the S-shaped planar antenna with the human-head model is 0.557 dB at 11 GHz (Fig. 10) and 4.52 dB at 13 GHz (Fig. 11).

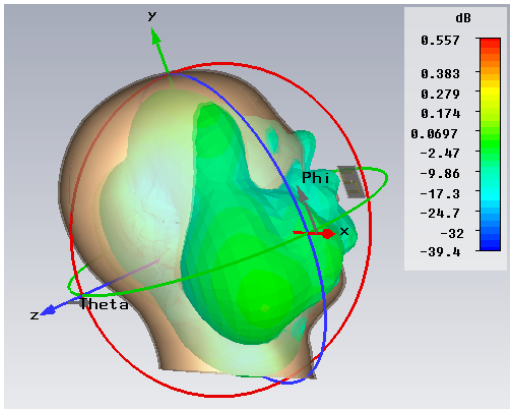



Figure 10. Farfield gain of antenna with head model (at 11 GHz)

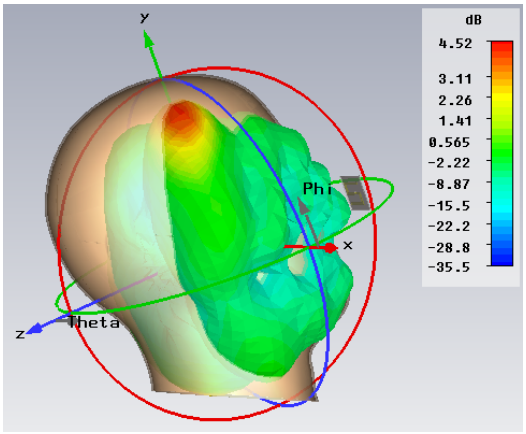



Figure 11. Farfield gain of antenna with head model (at 13 GHz)

*3) Antenna with metamaterial and human-head model:* The farfield gain of the antenna with metamaterial array and the human-head model is shown in this sub-section. The gain performance in this case is amazingly much better than those of the previous. The maximum farfield gain of the antenna is 6.62 dB at 11 GHz and 8.29 dB at 13 GHz as shown in Fig. 12 and Fig. 13 respectively.

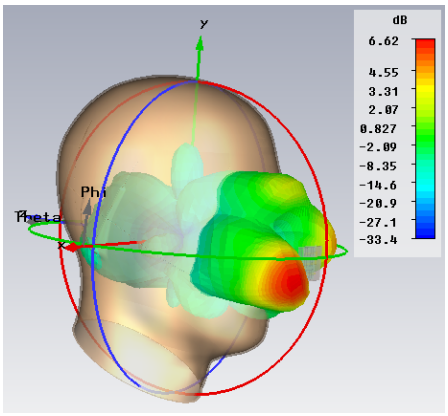



Figure 12. Farfield gain of antenna with meatamaterial and the head model (at 11 GHz)

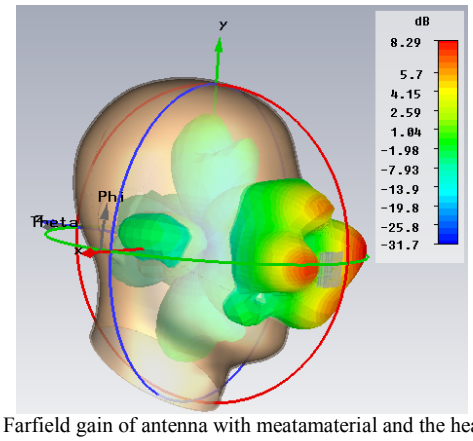



Figure 13. Farfield gain of antenna with meatamaterial and the head model (at 13 GHz)

## C. SAR calculation

SAR is calculated over an averaging mass of 10 g, which should be equal or less than 2 W/kg according to ICNIRP and CENELEC for general public exposure.

*1) Antenna with Head model:* For 11 GHz, the maximum SAR is calculated to be 1.38 W/kg (shown in Fig. 14) and for 13 GHz, the maximum SAR value is 1.78 W/kg (shown in Fig. 15).

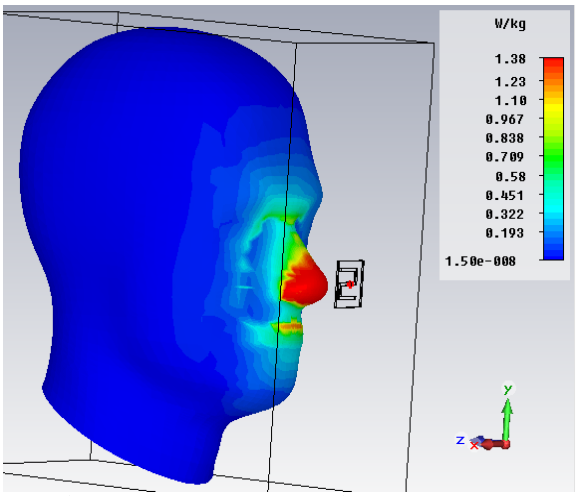



Figure 14. SAR over 10 g averaging mass (at 11 GHz)

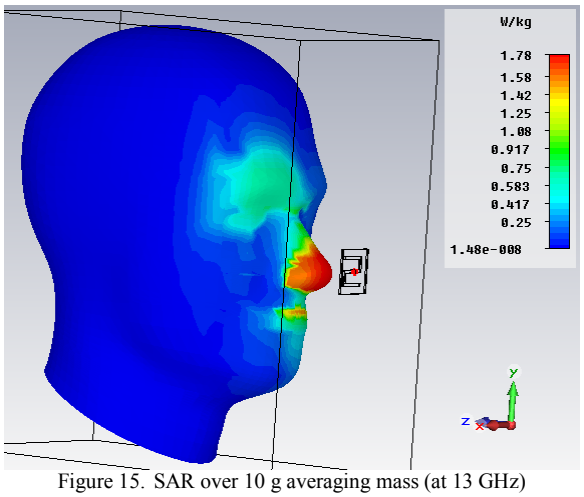


Figure 15. SAR over 10 g averaging mass (at 13 GHz)

*2) Antenna using Metamaterial Array:* For 11 GHz, the maximum SAR is calculated to be $1.8\times10^{-5}$ W/kg and for 13 GHz, the maximum SAR value is 0.00604 W/kg as shown in Fig. 16 and Fig. 17 respectively, which are much less than those of the previous.

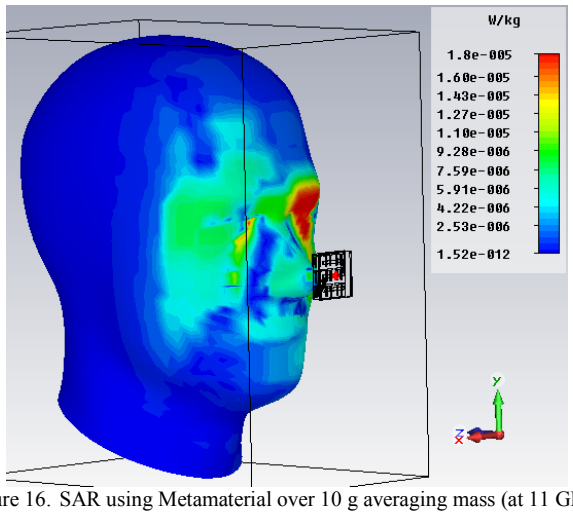


Figure 16. SAR using Metamaterial over 10 g averaging mass (at 11 GHz)

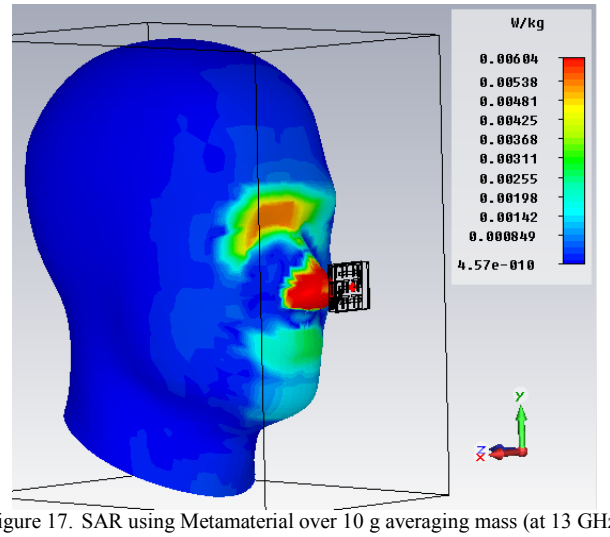


Figure 17. SAR using Metamaterial over 10 g averaging mass (at 13 GHz)

### *D. Summarized Result & Comparison*

The constitutive parameters of the metamaterial unit cell and its 3×3 array are given in the Table I. This table shows that the proposed structures exhibit the nature of DNG metamaterial, whose resonant frequency is at 13.11 GHz (unit cell) and 6.323 GHz (3×3 array).

TABLE I. CONSTITUTIVE PARAMETERS

| Constitutive parameters | Permittivity | Permeability | Refractive Index |
|---|---|---|---|
| Unit Cell (at 13.11 GHz) | $-0.008421$ $-j0.0069$ | $-6.825$ $+j5.579$ | $-0.309$ $-j0.0002$ |
| 3×3 array (at 6.323 GHz) | $-0.2614$ $+j5.315$ | $-0.0089$ $-j0.1863$ | $-0.9961$ $-j0.001$ |

The farfield gain shows a smaller value if a human-head is occupied with the antenna. But, using metamaterials, the gain is much increased and highly directive radiation pattern is achieved. The SAR from antenna with metamaterials and head model for both the frequencies (11 GHz and 13 GHz) remain far below 2 W/kg for an averaged mass of 10g. This values of SAR, can easily announce the S-shaped planar antenna with metamaterials as a safe electromagnetic device for human-health. The farfield gain and the SAR values for different frequencies are listed in Table II. We can see from the figures for gain patterns that the combined antenna system exhibits highly directional gain patterns, which should be much useful in satellite and Radar communications.

TABLE II. GAIN & SAR FOR DIFFERENT STATE OF ANTENNA

| Structures | Gain (dB) | | SAR over 10g of averaging mass (W/kg) | |
|---|---|---|---|---|
| | *At 11 GHz* | *At 13 GHz* | *At 11 GHz* | *At 13 GHz* |
| Antenna | 1.05 | 5.11 | ~ | ~ |
| Antenna with head model | 0.557 | 4.52 | 1.38 | 1.78 |
| Antenna with metamaterial and head model | 6.62 (Increase of 1088.5%) | 8.29 (Increase of 83.4%) | $1.8\times10^{-5}$ (Decrease of 99.99%) | 0.00604 (Decrease of 99.66%) |

## IV. CONCLUSION

In this investigation, a conventional S-shaped planar antenna is presented and its gain and SAR (over 10 g averaging mass of tissue) are observed. Then, a Pi-shaped metamaterial is designed as a sub-wavelength structure in CST and its 3×3 array is then created. The constitutive parameters for both the unit cell and its array are extracted from the S-parameter values with the help of direct retrieval method, which depicts that both the structures exhibit the properties of DNG metamaterial. Then, this array structure is used to

improve the gain and SAR performance. The metamaterial array is placed between the antenna and the human-head model, which acts not as an antenna but as an associating device. This configuration, where metamaterial is used between antenna and head-model, shows better performance in terms of gain and SAR. These improvements in gain and SAR are observed for 11 GHz and 13 GHz frequencies, where the gain is enhanced and the SAR is reduced by a large amount. The metamaterial array can be printed on a PCB while maintaining required distance from the antenna. This proposed configuration of antenna using metamaterial array would be very efficient where a huge gain and human-safety are required. Any X or $K_u$ band application including satellite and Radar communication can use this configuration. Again, as it produces a negligible SAR, this can be used in any portable communication device, such as mounted antennas on moving vehicles or antennas for domestic usages, at these operating frequencies. There exist many technologies in antenna fabrication such as PBG or EBG, but our proposed metamaterial antenna has better performance than the PBG antennas normally found in recent researches in case of Gain and SAR, which make this much more acceptable in the industry of communication and antenna technologies. We have a future plan to implant the metamaterial device beneath the skin of biological bodies (human or animal) which can be applied in RFID communications. Metamaterial cloaking, talk of the time nowadays, is also a concern of ours for future research.

## REFERENCES


[1] James, J. R. and P. S. Hall, *Handbook of Microstrip Antennas*, Peter Peregronic Ltd., London, 1989

[2] N. Engheta and R.W. Ziolkowski, "A positive future for double negative metamaterials", *IEEE Microwave Theory Tech.*, vol. 53, no. 4, pp. 1535–1556, 2005

[3] V. G. Veselago, "The electrodynamics of substances with simultaneously negative values of ε and μ", *Soviet Physics Usp.*, vol. 10, no. 4, pp. 509–514, January-February 1968

[4] John B. Pendry and David R. Smith, "Reversing light: negative refraction", *Physics Today*, December 2003

[5] J.B. Pendry, A.J. Holden, W. J. Stewart and I. Youngs, "Extremely low frequency plasmons in metallic mesostructures", *Phys. Rev. Lett.*, vol. 76, issue 25, pp. 4773–4776, June 1996

[6] Per O. Iversen, "Antenna Measurements of Wireless Communications Devices Including Human Body Interaction Effects", *European Co-operation in the Field of Scientific and Technical Research, COST 259, Wireless Flexible Personalised Communications*, April 2000

[7] Jensen, M. A. and Rahmat-Samii, Y., "EM interaction of handset antennas and a human in personal communications", *Proceedings of IEEE*, vol. 83, no. 1, pp. 7–17, 1995

[8] Jerzy Guterman, Antonio Moreiara, Custodio Peixeiro and Yahya Rahmat-Samii, "Antenna-User Interaction in MIMO-Enabled Laptops", *Radioengineering*, vol. 18, no. 4, December 2009

[9] *ICNIRP Guidelines for limiting exposure to time-varying Electric, Magnetic and Electromagnetic fields (up to 300 GHz),* International Commission on Non‐Ionizing Radiation Protection, Health Physics, vol. 74, no. 4, pp. 494–522, 1998

[10] *CENELEC, European Specification ES 59005, Considerations for the Evaluation of Human Exposure to Electromagnetic Fields (EMFs) from Mobile Telecommunication Equipment (MTE) in the Frequency Range 30 MHz-6 GHz,* Ref. no. ES–59005:1998 E, 1998

[11] Kivekas, Ollikainen, J. Lehtiniemi and T. Vainikainen, "Bandwidth, SAR, and efficiency of internal mobile phone antennas", *IEEE Transactions on Electromagnetic Compatibility*, vol. 46, no. 1, pp. 71–86, February 2004

[12] Faruque, M.R.I., Islam, M.T. and Misran N., "Effect of human head shapes for mobile phone exposure on electromagnetic absorption", *Informacije MIDEM*, vol. 40, no. 3, pp. 232–237, 2010

[13] Wang, J., Fujiwara, "FDTD computation of temperature rise in the human head for portable telephones", *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 8, pp. 1528–1534. 1999

[14] Islam, M.T., Faruque, M.R.I. and Misran N., "Design analysis of ferrite sheet attachment for SAR reduction in human head", *Progress in Electromagnetics Research (PIER)*, vol. 98, pp. 191–205, 2009

[15] Ikeuchi, R. and Hirata A., "Dipole antenna above EBG substrate for local SAR reduction", *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 904–906, 2011

[16] Pendry, J.B., Holden, A.J., Robbins, D.J. and Stewart, W.J. "Magnetism from conductors and enhanced nonlinear phenomena", *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, issue 11, pp. 2075–2084, November 1999

[17] M. B. Manapati and R. S. Kshetrimayum, "SAR Reduction In Human Head from Mobile Phone Radiation using Single Negative Metamaterials", *Journal of Electromagnetic Waves and Applications*, vol. 23, pp. 1385–1395, 2009

[18] CST Microwave Studio– Solve & Overview, CST Studio Suit 2011, CST-Computer Simulation Technology AG, website address: http://www.cst.com

[19] Anik Mallik, Sanjoy Kundu and Md. Osman Goni, "Design of a Novel Two-Rectangular U-Shaped Double Negative Metamaterial", *Proceedings of 2nd International Conference on Informatics, Electronics & Vision (ICIEV)*, Dhaka, Bangladesh, 2013

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Spectrum Hole Identification in Multiple TV bands by Adaptive Threshold Multi-Taper Spectrum Estimator for Cognitive Radio

Md. Mizanur Rahman[1], Chalie Charoenlarpnopparut[1,*], Prapun Suksompong[1], and Attaphongse Taparugssanagorn[2]
[1]Electronics and Communication Engineering, SIIT, Thammasat University, Thailand
[2]Wireless Information Security and Eco-Electronic Research Unit (WISRU), NECTEC, Thailand
[*] E-mail: chalie@siit.tu.ac.th

*Abstract*—**Cognitive radio (CR) is a wireless technology that facilitates unlicensed users to access the sensed spectrum without creating any interference to the primary user. Sensing the licensed spectrum to seek for spatiotemporal information of any possible spectrum hole is the first step needed to be performed by the CR devices. In this paper we propose an adaptive threshold technique to extract spectrum hole information from Multi-taper Spectrum Estimator (MSE) estimated spectrum. It has been found that when threshold is adapted depending on the change in noise level and presence of primary or secondary signals in a TV band, the detection performance improves compared to fixed threshold. A 400 kHz band was split into 20kHz sub bands to verify the possibility of splitting the TV bands into 200kHz bands from CR usage perspective. From analytical calculations and MATLAB simulations it has been shown that if TV band is viewed as 200 kHz sub-band channels from CR perspective; blocking probability for CR operation can be reduced. In addition it has also been shown that sensing multiple TV bands simultaneously can help CRs to find more opportunities with minor impact on system complexity.**

***Keywords—Multi-taper spectrum estimator, adaptive threshold, sub-band, TV band.***

## I. INTRODUCTION

Demand for wireless services is increasing immensely but no more spectrum is available to welcome new wireless services. Hopefully FCC [1] and Ofcom [2] has reported that allocated spectrum is inefficiently utilized. The report welcomes secondary users (SU) for efficient use of this invaluable spectrum opportunistically not crossing the interference temperature [3, 4] limit set by the regulatory bodies and the licensed primary users (PU) when the primary signal is present. "The interference temperature at a receiving antenna provides an accurate measure for the acceptable level of RF interference in the frequency band of interest. [4]" In contrast to noise temperature which quantifies the noise level, interference temperature quantifies the sources of interference in a radio environment. The interference temperature is measured in degrees Kelvin and the corresponding upper limit on permissible power spectral density in a frequency band of interest, is measured in joules per second or, equivalently, watts per hertz.

For efficient use of spectrum Jeorge Mitola in 1999 proposed a new wireless technology named cognitive radio (CR), which was warmly accepted by the wireless community [5].

Recently in North America, Europe and Australia TV channels have been switched and more countries have planned to switch their operation to digital transmission. Standard definition television SDTV in 4:3 aspect ratios has similar performance as regular analog TV (NTSC, PAL, SECAM) but uses quarter bandwidth [6]. As a result, some spectrum is free in the range of TV frequencies which can be used for additional services. Moreover TV spectrum has longer wavelength, more penetration power and can propagate longer distances which makes it attractive to the secondary users.

Secondary users can access spectrum in three approaches namely Underlay, Interweave and Overlay paradigms [7]. Underlay paradigm allows cognitive users to simultaneously transmit with the PUs but not crossing the interference temperature limit. The primary operator generally suggests interference temperature limit. Interweave paradigm focuses on identifying the available spectrum hole in a specific location and time. Whereas in Overlay paradigm, cognitive users keep knowledge of primary transmission and assigns part of their transmitting power to assist or relay PUs. In this paper we focus in Underlay and Interweave paradigms.

To provide the opportunity for the secondary users the biggest challenge for CRs is to sense the target spectrum as white or gray spectrum area. A way of segmenting the sensing methods is co-operative and non-cooperative. Cooperation results in large amounts of data, making it inefficient in terms of data throughput, delay sensitivity requirements and energy consumption [8]. In addition interference free data gathering in cooperative sensing requires dedicated channel. To avoid the drawbacks of co-operative scheme, non-cooperative spectrum-sensing technique is a good choice. Sensing may also be categorized into parametric and non-parametric. The major drawback for parametric methods is to keep modulation and coding information of all the channels of interest in its memory. In [9] it has also been mentioned that parametric methods such as Cyclostationary feature detector and Matched filter detector needs a priori information of both noise and primary signal. The design and detection process of a Cyclostationary feature detector or matched filter detector gets even more complicated as the number of signals to be detected increase.

On the other hand, non-parametric methods do not need any prior information therefore they are of good interest to CRs. The simplest non-parametric detector is energy detector (ED) which suffers from its detection inefficiency. ED needs good knowledge of the noise; large number of samples to achieve similar detection resolution compared to other methods, and cannot discriminate between primary and noise [9]. In [4] and [10] two non-parametric methods namely Multi-taper Spectrum Estimator (MSE) and Filter Bank Spectrum Estimator (FBSE) were proposed for spectrum estimation. As reported in [10] FBSE needs three times more samples and time to achieve similar accuracy compared to MSE. Moreover FBSE technique estimates the spectrum from filtered data sequence and needs two levels of signal processing. In the first step the filter bank performs the filtration and in the second step spectrum estimation is done. In contrast MSE estimates the spectrum in a single step by operating on the sampled version of the received signal from the antenna. As reported in [11] the MSE method is the best choice as of now for its accurate identification, estimation, quick computation, regularization, and signal classification.

In [4] MSE proposed by Haykin does not discuss on how spectrum hole information can be extracted from the estimated spectrum. A fixed threshold based detector for MSE has been proposed in [12] where the detector performance is likely to degrade due to changes in mean and variance of estimated spectrum. Mean and variance of estimated spectrum is changed frequently due to the changes in noise and signal strength in the channel from different CR transmitters. With the aid of (7.a), (7.b), (9.a) and (9.b) Fig. 1 (a) & (b) shows that threshold varies with noise variance when probability of false alarm, $P_{fa}$ is constant, and varies with noise variance when threshold is constant. In [13] it has been shown that "for the Multi-taper estimate, the probability of detection is increased when the SNR is increased." As a result SNR and detection performance changes when the noise variance is considered constant but in reality it actually varies. Thus threshold is needed to be adaptive. In this paper we have simulated an adaptive threshold method in MATLAB to extract channel activity information from MSE estimated spectrum and found better detection decision compared to fixed threshold method while noise and signals have been considered to be changing randomly at every simulation time slot.

A wideband spectrum sensing technique based on energy measurement in each sub-band has been proposed in [14] to improve the opportunistic throughput. In [15] another wideband spectrum-sensing technique were proposed, where each narrowband detection requires a down-converter, a periodogram estimator and an energy based threshold detector at different threshold level making the process complex. We propose to split a 6 MHz TV band into (6MHz / 200KHz =) 30 conventional wireless channels from CRs usage perspective so that CRs can operate in TV bands within the interference temperature limit. It was found that sub-band detection is effectively possible from MSE estimated spectrum by the proposed method in this paper. We also propose to sense multiple TV bands simultaneously by MSE to increase CR throughput with acceptable increase in additional complexity

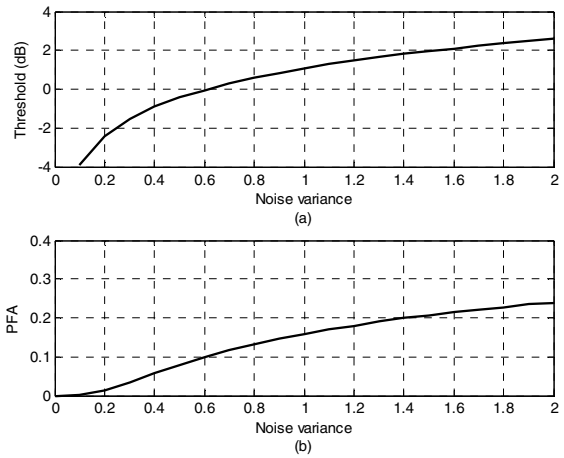



Figure 1. Impact of noise variance on (a) threshold, and (b) probability of false alarm.

for every additional 6 MHz channel. Every additional 6 MHz channel provides 30 additional sub bands and thereby offers less blocking probability to the CR devices. Moreover in every TV band on detection of the presence or absence of TV pilot frequencies CRs can switch their operation betwixt underlay or interweave paradigms.

The rest of the paper is organized as follows. Multi-taper spectrum estimator (MSE) has been explained in brief in section II. The system model for threshold adaptation technique, enhanced traffic handling by sub-band detection, and complexity rise due to multiple TV band sensing is presented in section III. Simulation results are discussed in section IV followed by conclusion in section V. Finally in section VI future work to improve the short comings of this paper has been mentioned.

## II. MULTI-TAPER SPECTRUM ESTIMATOR (MSE)

MSE is essentially a multi-periodogram estimator. A basic periodogram spectrum estimator obtains an estimate of a discrete time signal, $\{x(n), n = 0,1, \dots \dots, N-1\}$, where $N$ is the total number of samples, as in

$$\hat{S}_p(f) = \left|\sum_{n=0}^{N-1} w(n)x(n)e^{-j2\pi fn}\right|^2. \quad (1)$$

When, $w(n) = c\ (a\ constant)$, (1) is called a periodogram. When $w(n)$ is not constant (1) is called the modified periodogram and $w(n)$ is called the data window.

The basic modified periodogram suffers from bias-variance dilemma. MSE introduced by Thomson in [16] was implemented by Haykin in [4] to estimate spectrum for CR to improve this bias-variance dilemma. MSE is a Multi-periodogram spectrum estimator where window functions $w(n)$ in (1) are considered to be discrete prolate spheroidal sequences (DPSS) [17]. DPSS are also called Slepian tapers and have orthonormal property. The Slepian tapers expand the part of the time series in a fixed bandwidth $(f - \Delta f)$ to $(f + \Delta f)$, centered on some frequency $f$, where $\Delta f$ is signal bandwidth. MSE uses $K$ Slepian orthonormal tapers to obtain

$K$ periodograms, which are averaged to achieve the Multi-taper estimate. The process is as follows in brief:

Let, $\{x(n), n = 0, 1, \dots\dots\dots, N-1\}$, be the discrete time signal whose power spectrum estimation (PSE) is required, $\{g_k(n),\ k = 0,1, \dots\dots, K-1\}$, are the $K$ orthonormal Slepian tapers each of length $N$, then PSE for each Slepian taper can be represented as in

$$\left(\hat{S}_{slep}(f\ )\right)_k = \sum_{n=0}^{N-1}\left|g_k(n)x(n)e^{-j2\pi fn}\right|^2, \qquad (2)$$

where, $k = 0, 1, \dots\dots\dots, K-1$; $K = N\Delta f - 1$ and $N\Delta f$ is the time bandwidth product.

Averaging the outputs of the $K$ peridograms found by (2) the PSE of MSE is found by

$$\hat{S}_{mtm}(f\ ) = \frac{1}{K}\sum_{k=0}^{K}\left|\left(\hat{S}_{slep}(f\ )\right)_k\right|, \qquad (3)$$

where, $k = 0, 1, \dots\dots\dots, K-1$.

## III. System Model

In this section we propose and discuss the system model for threshold adaptation, and sub-band detection techniques. Block diagram of the adaptive threshold detector is depicted in Fig. 2. An analog to digital converter (ADC) convert the received analog signal by the antenna to discrete time signal to achieve $x(n)$. The DPSS generator in Fig. 2 generates the DPSS sequences to be used in the $K$ periodogram spectrum estimators. Mean of the $K$ periodograms are taken as threshold to adapt the threshold due to presence or variation in strength of noise and signals in the channel. The blocks indicated by $max(sub-band\ i)$, where $i = 1$ to $M$ and $M$ is the number of 200kHz sub-bands finds maximum signal level at each sub-band. Finally maximum signal level at each sub-band is compared to the adaptive threshold to find any free sub-band in the Observed TV band. Complexity and traffic analysis of sensing multiple TV bands simultaneously by the proposed model have also been presented. It can be seen from Fig. 2 that from the MSE the estimated spectrum is passed to the threshold block and $M$ sub-band detector blocks. A maximum value of each sub-band is compared in parallel by $M$ comparators depending on the calculated threshold to extract any possibly active primary pilot or secondary user activity in the sub bands.

### A. *Threshold adaptation and sub-band detection process for MSE*

We consider noise and signal to be Gaussian distributed and independent random process. Noise as well as its variance varies due to environmental disturbances and Tx-Rx electronics. Furthermore, number of simultaneously active channels also varies randomly and changes the mean and variance of composite signal and noise. As the estimated spectrum contains both signal and noise information, to find the available white or gray hole availability it is required to know the presence of the primary pilot and secondary signals into the TV sub-bands. Mean and variance of the output spectrum from MSE has been considered to estimate the

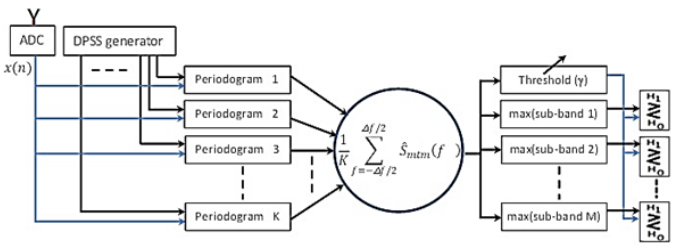



Figure 2. Block diagram of adaptive threshold detection from estimated spectrum of MSE.

threshold at every run. In this method the threshold will vary as the noise and signal status at every time slot in the channel varies.

Let us consider that we have $L$ samples of the estimated spectrum. The mean of the estimated spectrum can be found as in

$$\mu_{sn} = \frac{\sum_{l=0}^{L-1}\hat{S}(l)}{L}, \qquad (4)$$

where, $\mu_{sn} = \mu_s + \mu_n$, $\mu_s$ is the mean of the active signals in different sub bands and $\mu_n$ is the mean of noise.

The variance of the estimated spectrum can be found as in

$$\begin{aligned}\sigma_{sn}^2 &= \frac{1}{L}\sum_{l=0}^{L-1}\left(\hat{S}(l) - \frac{\sum_{l=0}^{L-1}\hat{S}(l)}{L}\right)^2 \\ &= \frac{1}{L}\sum_{l=0}^{L-1}\left(\hat{S}(l) - \frac{\mu_{sn}}{L}\right)^2, \end{aligned} \qquad (5)$$

where, $\sigma_{sn}^2 = \sigma_s^2 + \sigma_n^2$, $\sigma_s^2$ is the variance of the active signals in different sub bands and $\sigma_n^2$ is the noise variance.

If $X$ is a random variable representing the power at each frequency components of the spectrum and $\gamma$ is the threshold then the probability that a frequency component has more strength than the threshold is given by, $Pr[X > \gamma]$, where, $\gamma \in \mathcal{R}$ and $\mathcal{R}$ is a set of real values.

Let, $\eta = \frac{\gamma-\mu}{\sigma}$, then,

$$\begin{aligned} Pr(X > \gamma) &= Pr(X > \eta\gamma + \mu) \\ &= Pr\left(\frac{X-\mu}{\sigma} > \eta\right) \\ &= Pr(\eta) \\ &= Q\left(\frac{\gamma-\mu}{\sigma}\right), \end{aligned} \qquad (6)$$

where, $Q(.)$ is the $Q$ function given as, $Q(z) \triangleq \frac{1}{\sqrt{2\pi}}\int_z^\infty e^{-\frac{y^2}{2}}dy$.

Thus with the use of (6) probability of detection, $P_d$ and corresponding threshold, $\gamma$ is given as follows in

$$\mathrm{P_d} = \mathrm{Q}\left(\frac{\gamma-\mu_{\mathrm{sn}}}{\sigma_{\mathrm{sn}}}\right), \qquad (7.a)$$

$$\gamma = \sigma_{sn}.Q^{-1}(P_d) + \mu_{sn}. \qquad (7.b)$$

The probability of missed detection is given as follows in

$$P_{md} = 1 - Q\left(\frac{\gamma-\mu_{sn}}{\sigma_{sn}}\right). \quad (8)$$

And the probability of false alarm, $P_{fa}$ and the corresponding threshold, $\gamma$ is given as follows in

$$P_{fa} = \frac{\gamma-\mu_n}{\sigma_n}, \quad (9.a)$$

$$\gamma = \sigma_n Q^{-1}(P_d) + \mu_n, \quad (9.b)$$

where, $Q^{-1}(.)$ is the inverse $Q$ function.

Using (4) and (5) the $\mu_n, \mu_{sn}, \sigma_n, and\ \sigma_{sn}$ are updated to adapt the threshold $\gamma = \sigma_{sn}Q^{-1}(P_d) + \mu_{sn}$ as given in (7.b).

*B. Enhanced traffic handling capability of sub band detection*

As general voice and data services use 200 kHz band each, a 6MHz TV channel can be thought of as a combination of thirty 200kHz sub-bands from secondary user perspective so that on detection of any free sub-band CRs can switch their operation between underlay and interweave paradigms depending on the presence or absence of the primary pilot. Since this method will detect the presence of the signal in every sub-band, there is less probability to create interference to the primary not knowing its presence as the detection of primary or any secondary at all the sub-bands will force the CRs not to cross the interference temperature limit while any other CR or wireless microphone working in the area of observation.

We consider that all the served and non-served CRs at the present time slot will continue its search for the possible free channel in the next time slot also and wish to access the channel again. Thus the process can be considered as a Binomial distribution. The probability that exactly *m* channels are found free and exactly *m* CRs get access among the total of *M* CRs is given by the probability mass function:

$$Pr(X = m) = \binom{M}{m} p^m (1-p)^{M-m}$$

for, *m* = 0, 1, 2, ..., *M*, where, $\binom{M}{m} = \frac{M!}{m!(n-k)!}$.

When spectrum holes are found some CR devices get access and others are dropped. In the next time slot if all the CRs bid with equal probability then the process is considered to have constant arrival rate. Erlang B formula can be used to analyze the blocking probability when a TV band is considered to be split into sub-band from CR perspective. Total numbers of sources are assumed to be infinite as served CRs keep bidding. For fair use policy holding time $h$ is equal to slot time. Blocking probability in such a lossy CR network is given by

$$P_b = B(E, m) = \frac{\frac{E^m}{m!}}{\sum_{i=0}^{m}\frac{E^i}{i!}}, \quad (10)$$

where, $m$ is the number of free sub-bands, $E = \lambda h$ Erlangs, λ = arrival rate and $1/h$, is service rate.

*C. Complexity of Sensing Multiple TV band by MSE*

Multiple TV bands can be sensed simultaneously to increase CR throughput with minor increase in complexity by choosing only the suitable sampling frequency $f_s$ and thereby number of samples $N$. As serial to parallel converter is common in all types of estimator we do not consider the complexity of serial to parallel converter. An FFT block requires $\frac{N}{2}log_2N$ complex multiplication, where one complex multiplication needs 4 real multiplications and 2 real additions. On the other hand 1 complex addition requires 2 real additions [18]. Thus complexity of each periodogram is $2N + 2Nlog_2N$ real multiplications and $N + 3Nlog_2N$ real additions. For $K$ orthonormal Slepian tapers we need $K$ periodograms and the complexity becomes, $K(2N + 2Nlog_2N)$ real multiplications and $K(N + 3Nlog_2N)$ real additions. Each periodogram provides $\left(\frac{N}{2}+1\right)$ samples at their outputs and as we have $K$ periodograms, $(K-1)$ additions are required. So, $\left(\frac{N}{2}+1\right)(K-1)$ additions and one multiplication by $\frac{1}{K}$ are required to achieve the MSE estimation. Thus the total complexity of MSE to sense multiple TV band is given as, $K(2N + 2Nlog_2N) + 1)$ real multiplications and $(K\,(N + 3Nlog_2N) + \left(\frac{N}{2}+1\right)(K-1))$ real additions.

## IV. Simulation and Results

Simulation was performed 10,000 times where each simulation indicates a time slot; a 400 kHz band whose lower and upper side frequencies are 101kHz and 500kHz was split into 20 sub bands each having 20kHz bandwidth. Multiple sub-bands were considered to be activated randomly and simultaneously. Detection process was performed for all the 10,000 time slots. Probability of false alarm $p_{fa}$ and probability of detection $p_d$ is considered to be 0.1 and 0.9 respectively. It can be seen in table I that among 10,000 simulations accurate detection is 76.63%, false detection is 18.27%, and missed detection is 5.10%. Detection performance at 100 simulation time slots has been shown in Fig. 3. From Fig. 3(a) it can be seen that for fixed threshold at different time slots, number of sub-bands detected busy are less than actually active sub-bands. Fig. 3(b) shows that the detection performance of adaptive threshold has better detection capability compared to fixed threshold method. Due to shortage of simulation resources a real TV channel cannot be taken for experiment but a 400 kHz band was chosen. In the experiment missed detection occurred 510 times with most of the active sub band channels detected, whereas only 1 or 2 sub-bands were miss detected in all the aforementioned 510 time slots as it can be seen in Fig. 4. As most of the active sub bands are detected correctly CRs will operate in the sub-bands with such transmission power to operate in underlay mode. On detection of absence of primary pilot CRs can switch their operation to interweave paradigm.

TABLE I. CORRECT DETECTIONS, FALSE DETECTIONS, AND MISSED DETECTIONS IN 10,000 OBSERVATION TIME SLOTS.

| Detections | Number of Observation Time slots | In Percent |
|---|---|---|
| Correct | 7,663 | 76.63% |
| False | 1,827 | 18.27% |
| Missed | 510 | 5.10% |
| Total | 10,000 | 100.00% |

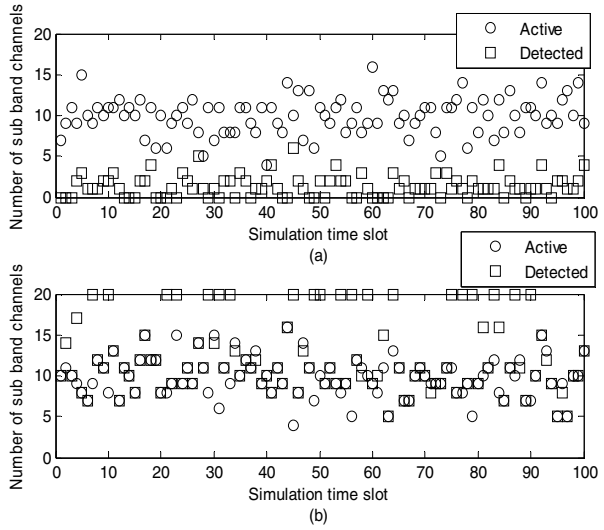



Figure 3. Detection performance at 100 simulation time slots for (a) fixed threshold, and (b) adaptive threshold.

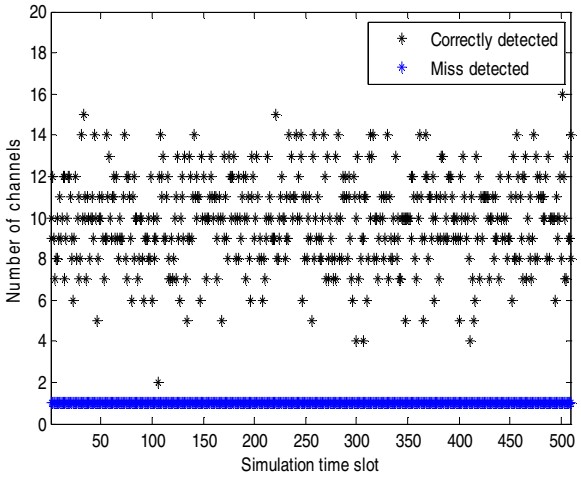



Figure 4. Active and miss detected sub bands at every time slot.

From Fig. 5 it is evident that the blocking probability for CR decreases when a whole TV spectrum is seen as split into different sub-bands from CR perspective. Let us consider the sub-bands are to serve the CRs. A total TV band will be serving 30 CRs by 30 sub-bands simultaneously where CRs are considered to be infinite. (As served CRs may also request for a service in the next time slot.) By Erlang blocking probability formula we can see that if we view a TV band as a single channel from CR perspective then blocking probability is very high. Whereas blocking probability decrease as we spl-

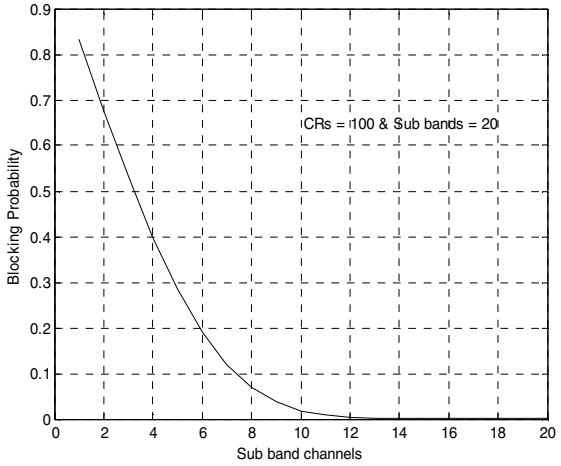



Figure 5. Blocking probability decreases as the number of sub-bands to support CRs increases.

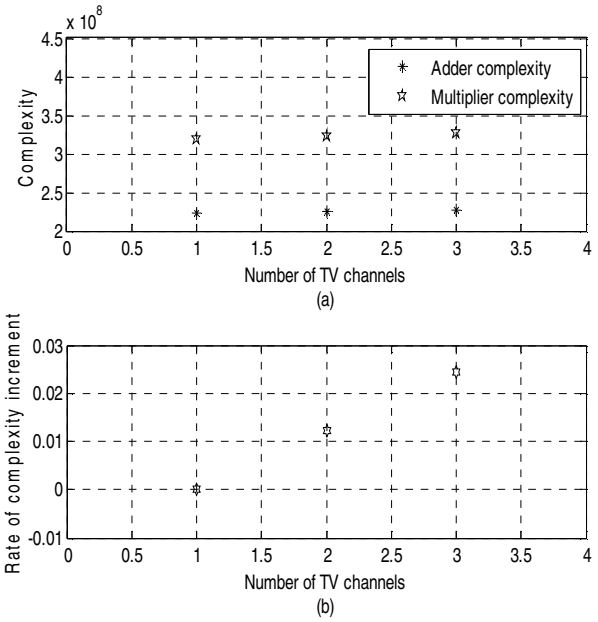



Figure 6. Complexity analysis of sensing multiple 6MHz TV band simultaneously. (a) Complexity to sense multiple TV bands, (b) rate of increase in complexity.

it it into sub channels to be shared by 30 CRs simultaneously.

Fig. 6 shows the complexity analysis of sensing multiple TV channels to increase CR throughput further by finding more spectrum resources simultaneously. We can see that complexity increases at a rate of 1.22% to sense every additional 6MHz TV band (compared to single TV band). We can increase throughput and decrease blocking to support CRs operation into the TV bands if we accept this minor increment in complexity.

## V. CONCLUSION

In this paper adaptive threshold method has been employed to extract sub-band utilization information from MSE estimated spectrum. It has been found that threshold

adaptation improves the detection performance significantly in terms of probability of false alarm, probability of detection and probability of missed detection. It has also been shown that sub-band detection process increases the throughput and decreases the 'blocking probability' and the 'probability of interference' to the primary while providing more opportunities to CRs. In some time-slots most of the sub-bands were correctly detected, however a few sub-bands were miss-detected or false-detected assuring the PU with less possible interference. Sensing multiple TV band simultaneously has been seen to be efficient in terms of complexity. Furthermore with sub band detection technique white or gray space can be detected. Pilot detection facilitates the CR to switch its operation between underlay or interweave paradigms.

## VI. FUTURE WORK

Due to hardware limitation, real TV band processing algorithm may not be feasible. In future we plan to experiment the detection model for real TV bands. Optimal distribution of the detected free sub-bands from multiple TV channels also needs to be looked into.

## ACKNOWLEDGMENT

The authors thank Thailand Office of Higher Education Commission (NRU Project) for the financial support.

## REFERENCES

[1] (2002) Federal Communications Commission. [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-228542A1.pdf

[2] (2009) Ofcom|Digital dividend: cognitive access. [Online]. Available: http://stakeholders.ofcom.org.uk/consultations/cognitive/statement/

[3] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201-220, Feb. 2005.

[4] S. Haykin, "The Multitaper method for accurate apectrum aensing in Cognitive Radio environments," *Forty-First Asilomar Conference Record on Signals, Systems and Computers*, Nov. 2007, pp. 436-439.

[5] J. Mitola III, "Software radios: survey, critical evaluation and future directions," *IEEE Aerospace and Electronic Systems Magazine*, vol. 8, no. 4, pp. 25-36, Apr, 1993.

[6] (2013) International Cablemakers Federation. [Online]. Available: http://www.icf.at/en/6000/how_much_bandwidth.html

[7] S. Srinivasa and S. A. Jafar, "Cognitive radios for dynamic spectrum access-the throughput potential of cognitive radio: A theoretical perspective," *IEEE Communications Magazine*, vol. 45, no. 5, pp. 73-79, May 2007.

[8] M. Subhedar and G. Birajdar, "Spectrum sensing techniques in Cognitive Radio networks: A survey," *International Journal of Next-Generation Networks (IJNGN)*, vol. 3, no. 2, pp. 37-51, Jun. 2011.

[9] A. Ghasemi and E. S. Sousa, "Spectrum sensing in Cognitive Radio networks: requirements, challenges and design trade-offs," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 32-39, Apr. 2008.

[10] B. Farhang-Boroujeny, "Filter Bank spectrum sensing for Cognitive Radios," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1801-1811, May 2008.

[11] A. Al-Mamun and M. R. Ullah, "Cognitive Radio for short range systems based on ultra-wideband," M. Eng. Thesis, Blekinge Institute of Technology, Sweden, Jun. 2011.

[12] C. Tsung-Wei, L. Jung-Mao and M. His-Pin, "Optimal detector for Multitaper spectrum estimator in Cognitive Radios," *IEEE Global Telecommunications Conference*, Nov. 30 – Dec. 4, 2009, pp. 1–6.

[13] E. H. Gismalla and E. Alsusa, "*New and accurate results on the performance of the Multitaper-based detector,"* in *IEEE International Conference on Communications,* Jun. 2012, pp. 1609–1613.

[14] S. Srinu, S. L. Sabat, and S. K. Udgata, "Wideband spectrum sensing based on energy detection for Cognitive Radio network," in *World Congress on Information and Communication Technologies (WICT)*, Dec. 2011, pp. 651-656.

[15] Y. Pei, Y. C. Liang, K. C. Teh, and K. H. Li, "How much time is needed for wideband spectrum sensing?," *IEEE Transactions on Wireless Communications*, vol. 8, no. 11, pp. 5466-5471, Nov. 2009.

[16] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, vol. 70, no. 9, pp. 1055-1096, Sept. 1982.

[17] D. Slepian, "Prolate spheroidal wave functions, fourier analysis, and uncertainty v: The discrete case," *Bell System Technical Journal*, vol. 57, no. 5 pp. 1371-1430, May-Jun. 1978.

[18] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing: A Practical Approach, 2nd Edition,* Pearson Education, 2002.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Measuring Security for Cloud Service Provider : A Third Party Approach

Md Whaiduzzaman[1*], Abdullah Gani[2]
[1]Member IEEE, [2]Senior Member IEEE
[1,2] Mobile Cloud Computing Research Lab
Faculty of Computer Science and Information Technology
University of Malaya
Malaysia
[*] E-mail: wzaman110054@siswa.um.edu.my

*Abstract*— **Cloud Computing (CC) is a new paradigm of utility computing and enormously growing phenomenon in the present IT industry hype. CC leverages low cost investment opportunity for the new business entrepreneur as well as business avenues for cloud service providers. As the number of the new Cloud Service Customer (CSC) increases, users require a secure, reliable and trustworthy Cloud Service Provider (CSP) from the market to store confidential data. However, a number of shortcomings in reliable monitoring and identifying security risks, threats are an immense concern in choosing the highly secure CSP for the wider cloud community. The secure CSP ranking system is currently a challenging aspect to gauge trust, privacy and security. In this paper, a Trusted Third Party (TTP) like credit rating agency is introduced for security ranking by identifying current assessable security risks. We propose an automated software scripting model by penetration testing for TTP to run on CSP side and identify the vulnerability and check security strength and fault tolerance capacity of the CSP. Using the results, several non-measurable metrics are added and provide the ranking system of secured trustworthy CSP ranking systems. Moreover, we propose a conceptual model for monitoring and maintaining such TTP cloud ranking providers worldwide called federated third party approach. Hence the model of federated third party cloud ranking and monitoring system assures and boosts up the confidence to make a feasible secure and trustworthy market of CSPs.**

***Keywords— Cloud computing, cloud service provider, trusted third party, cloud security ranking.***

## I. INTRODUCTION

Nowadays, mobile devices are rapidly becoming a mandatory part of human life as this is the most convenient and easy way of communication tools. Mobile users experienced rich applications sense of various services from mobile applications, which run on the devices and on remote servers via wireless networks [1]. Virtualization, elasticity, on-demand, instant service and pay as you go through the internet are the main characteristics of CC. Mobile Cloud Computing (MCC) can be defined as a rich mobile computing technology that leverages unified elastic resources of varied clouds and network technologies toward unrestricted functionality, storage, and mobility. Hence, it serves a multitude of mobile devices anywhere, anytime through the channel on the Internet regardless of heterogeneous environments and platforms based on the pay-as-you-use principle [2, 3]. MCC is an amalgam of three foundations, namely cloud computing, mobile computing, and networking [4].

MCC is a current IT buzzword and becomes a powerful trend in the development of IT technology as well as commerce and industry fields. For such business benefits offered by mobile cloud computing, many organizations have started building applications on the Cloud infrastructure and making their businesses agile by using flexible and elastic Cloud services. But moving applications and data into the Cloud is not very easy task. Many challenges exist to leverage the full potential that Cloud computing promises. These challenges are often related to the fact that existing applications have specific requirements and characteristics [5]. However, it is difficult for a new cloud customer to get the best secured CSP from the market. For some company owner, it is particularly important to find out the secured Cloud service provider from the market. This is a major problem for every user, especially those who are more concerned about data security and privacy from CSP. Our motivation is to help the new Cloud customer to find the most reliable and secured CSP in terms of security and trust through a secured ranking system.

Due to the massive diversity in the available Cloud services, from the customer's point of view, it has become challenging to select whose services they should use and what is the base of their choice. Presently, there is a lack of frameworks that can permit customers to evaluate Cloud offerings and rank them based on their ability to meet the user's Quality of Service (QoS) and security requirements. In this work, a secure Cloud service provider ranking system and a mechanism that measure the secured Cloud services are proposed which can make a major impact and will craft a healthy competition among Cloud providers to satisfy their Service Level Agreement (SLA) and improve their QoS and trustworthiness [6]. Hence, we feel strong requirement of a ranking system by which a new cloud customer can identify his/her needs and take the calculated risk of business data before handover to some unknown cloud service provider in cyber space. Our objective is to run the cloud service provider and the new cloud customer and maintain the smooth trust and provide them a tool that can verify and ranked the Cloud

service provider. With these ranked results new cloud customers justifies the business needs in terms of security and reliability which cloud service provider is the best option. This ranked system which provided by the trusted third party, will provide more confidence and validity among the mobile cloud market.

New cloud customers can independently make decisions without any cloud broker which is a significant feature of this model. Moreover, making this TTP model more synchronized and robustly in worldwide a TTP monitoring system is proposed that provide higher confidence and wider acceptability of ranking system homogeneously and impartially worldwide. We advocate these proposed tools and model have a viable business needs and the proposed TTP monitoring model is a feasible business solutions model for mobile cloud computing community.

This paper highlights CSP security aspects and how to ensure of these in mobile cloud user's perspective. We propose a conceptual security vulnerability measuring automated model for ranked the cloud provider authenticity and reliability such as issuing certificates which help the new cloud customer to evaluate the best CSP in the market. The conceptual model contains a monitoring of TTP to protect cloud users rights and cloud provider's security. The regions are divided and employed federated monitoring approach globally. Federated TTP monitoring assures the same service and more interoperable among TTP worldwide.

This paper is organized as follows: Section 2 discusses related work, Section 3 introduces the proposed model with flow charts, and explains ranking system and finally Section 4 provides the conclusion.

## II. RELATED WORK

A personalized cloud component ranking for different designers of cloud applications, and proposed a QoS driven component ranking framework for cloud applications by taking advantage of the past component usage experiences of different component users. Again several resource provisioning policies can be used to extend the capacity of a local cluster by leveraging external resource providers, as well as reduce the cost by using the Spot Market. Using the indicator as one of the main SLA parameters to determine who is responsible for the violation of the revenue or profit parameters were proposed and explained by M. Alhamad et al. [7]. A proposal of a set of cloud computing specific performance and quality of service (QoS) attributes, an information collection mechanism and the analytic algorithm based on Singular Value Decomposition Technique (SVD) to determine the best service provider for a user application with a specific set of requirements were proposed by H. Chan et al. [8]. The SMICloud model was proposed by S. K. Garg et al. [9] which let users compare different Cloud offerings, according to their priorities and along several dimensions, and select whatever is appropriate to their needs. An analytical Hierarchical Process (AHP) based ranking appliance was proposed which can calculate the Cloud services based on different applications.

## III. CONCEPTUAL MODEL OF FEDERATED CSP RANKING SYSTEM (CMFCSPRS)

We consider the following scenario for explaining our model.

Consider a scenario of a new cloud customer; say a company owner or manager is considering adopting cloud facility for the company. Main priority and mandatory condition is to protect company data security and privacy. The manager can see lots of cloud service provider in the market but not adequate guidelines to adopt the best secured cloud service provider for an organization. New cloud customer needs the security and trust certificate or report of these providers for making a decision to choose the right provider in terms of reliability, security and trustworthiness. So, clearly security issues are the most significant issue which is impeding the growth of mobile cloud computing [10, 11]. However, few ranked systems are available in service provisioning or performance issues but not adequate cloud service provider security ranking system is currently available.

Mobile Cloud Service Providers: Mobile Cloud service providers are the entities who own the cloud infrastructure and provide cloud services for consumers. The design and implementation of cloud provider infrastructure and price models are outside the scope of this paper.

Security Metrics: To know what to measure, how to measure and communicate those metrics which can help us to improve security's efficiency, effectiveness and standing in the business perspective. The generated metrics will provide an initial baseline to ensure achieving the targeted goals, which will evolve over time according to particular business needs and information security risk aspect.

Attack Vectors: Attack vectors are normally routes or methods used to get into computer systems, usually for evil purposes. They take advantage of known weak spots to gain entry. Many attack vectors take advantage of the human element in the system, because that's often the weakest link. Mainly, it refers to any methods of attacks chosen by hackers to identify weak points or vulnerability on the client or server end of a network for engineering defects in the user system in order to infect or achieve control over system resources.

In fact, by considering several security issues [12-15] we found obvious to need some sort of monitoring, assurance and trust which not only come from the Cloud Service provider but also from a trusted third party. So, Trusted Third Party and security ensuring features together provide trust among the cloud community as a whole. Figure 1 shows the conceptual model of CSP ranking system.

### *A. Assumptions*

In this conceptual model, several assumptions should be considered as follows:

- Like credit Rating company, TTP must maintain the trust and reliability.
- TTP should have enough resources to provide for processing and executing their own work.

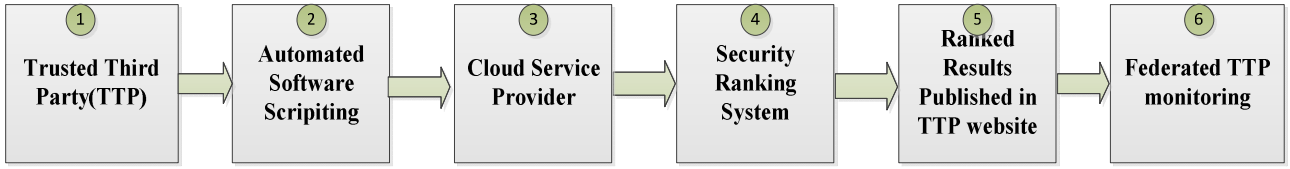



Figure 1. Conceptual Model of CSP ranking system

- TTP must be maintained and regulated by strict laws, regulations and transparent policies.
- Both TTP and CSP mutually agree before executing the software penetration test.
- Considered as CSP provide SaaS, PaaS and IaaS of its own.
- TTP is responsible for collecting non measurable metrics from trusted source and process this information for ranking results.
- A new cloud customer looking for security and trust certificate should pay to the third party to see the ranked results and use their services.

### *B. Conceptual Model of Federated CSP ranking system*

In this section, our model is explained and provides each step explanations. Firstly, a TTP engages to provide a ranking system of secure cloud service providers. A conceptual software scripting model which is executed by TTP in the CSP side to collect measurable security metrics.By using attack vectors every attempt of security breach is counted as a metric and weighted into numbers. Then the TTP is responsible for accumulating several non measurable metrics from different reliable sources and weighted it as a corresponding number. After software scripting, TTP collects several measurable security metrics and send back to the TTP authority and by adding both measurable and non measurable metrics, TTP provides a security ranking system. The Ranking systems publish on the TTP website with self-explanatory guidelines and ready to use for new cloud customer. Paying a prescribed fee new cloud customer can easily make the decision to choose the best secure cloud service provider for the company from available CSPs of the market. In addition, finally, in order to increase the authenticity and acceptability of TTP globally, TTP conceptual worldwide federated monitoring system is introduced for monitoring and justify the ranking system acceptability.

### *C. Automated security checker software penetration by the TTP to CSP*

For addressing security issues, a TTP is engaged which is accountable for the reliable, safe and secured cloud service provider and ensured the privacy and security.

Hence, a conceptual software scripting model is proposed for the TTP to help to monitor and check the security strength of the CSP side security system. In addition, we advocate a simple methodology that includes top threats like denial of services, side channel attack, encryption, log on, and other measurable required parameters embedded into the software to run into the cloud provider side with both party mutual agreements.

However, a trusted third party will test several predefined security and vulnerability criteria of the Cloud Service provider by automated software scripting. This security element of measurable criteria or factors are mentioned security metrics attack vectors (S). This software will run into Cloud Service provider end and equipped with the latest security issues and try to penetrate the service provider security defense. Then, the return results of successful or unsuccessful attempts will analyze by measurable statistical methods and provide the ranking system of the security strength of CSP to the TTP.

Moreover, TTP should collect other non-measurable factors such as: previous down time, user feedback, provide input of other non-measurable criteria or metric (F) as a value or metric to give the result in a number or numerically. In addition, influencing factors such as the location of CSP, platform, hardware, infrastructure, elasticity, service provisioning, hiring human resources, company policies, maintaining standards of SLA and previous downtime, customer satisfaction records are should be considered. A detailed flowchart is provided in Fig.2

Furthermore, automated TTP security checker software run in the cloud service provider side and provide a result. All sorts of securities will be trying to break by the TTP attack vectors automated software scripting. The security metrics help the TTP to issue certificates of the secured cloud service provider. Finally TTP calculate a security ranking by adding S+ F and provide a numerical value and publish to the TTP website. The website should be self-explanatory about ranking system and provide the numerical result descriptions regarding ranked results. Hence, new cloud customer easily interprets the ranked results and understands the security and vulnerability of a specific CSP security strength.

The interaction among the cloud customer, TTP and CSP is provided with a sequence diagram as shown in Figure 3.

### *D. Algorithms for Ranking Systems*

We provide two algorithms and both are are explained here. Pseudo code 1 explains the calculating procedure of security metrics S and Pseudocode 2 describes the final calculation of ranking results, R. Figure 2 shows the flow chart of the system.

***Pseudo code 1: Calculating Security metrices S***

*1 /* Calculate Security metrics S */*

*2 Initially S=0;*

*3 /* Negotiate with CSP side from TTP to network and connection setup */*

*4 While connection setup =0 do*

*5 connection setup=1;*

*6 end ;*

*7 /* Software scripting try to break the security of*

*8 CSP side . Here, I is the several numbers of top*

*9 threats defined and included in TTP software*

*9 Scripting */*

*10 While I ! = 0 do*

*11 /* Exucucute the specific software coding to test the strength or defense mechanism of CSP */*

*10 If successful to break the specific security, S=S+1;*

*12 endif ;*

*13 end ;*

*14 /* Finally send the Security metrics S to the TTP*/*

*15 Send S to TTP*

**Pseudo code 2: Calculating Final Rank R**

*1 /* Collect the non-measurable metrics (F) from reliable sources and input by the TTP*/*

*2 Get the input F;*

*3 Get the input R /* As collected by employing Pseudo code 1 by TTP */*

*4 Final Rank, R=S+F;*

*5 Published the Rank result, or to the TTP website*

*E. Federated TTP Monitoring Model*

We propose a conceptual design of Trusted Third Party monitoring system model by considering dividing the whole world in several regions. In every region several TTP will engage to provide ranking system and each region is monitored and coordinated by a Regional Third Party Monitoring Agency (RTPMA) and they will monitor all the Trusted Third Party in that specific region. For ensuring more transparency, trust, reliability and homogeneity another layer of monitoring is introduced. This layer of monitoring is named as International Cloud Monitoring Agency (ICMA) which will monitor these RTPMA. The total system, as a whole is named as Federated Third party Cloud Monitoring System.

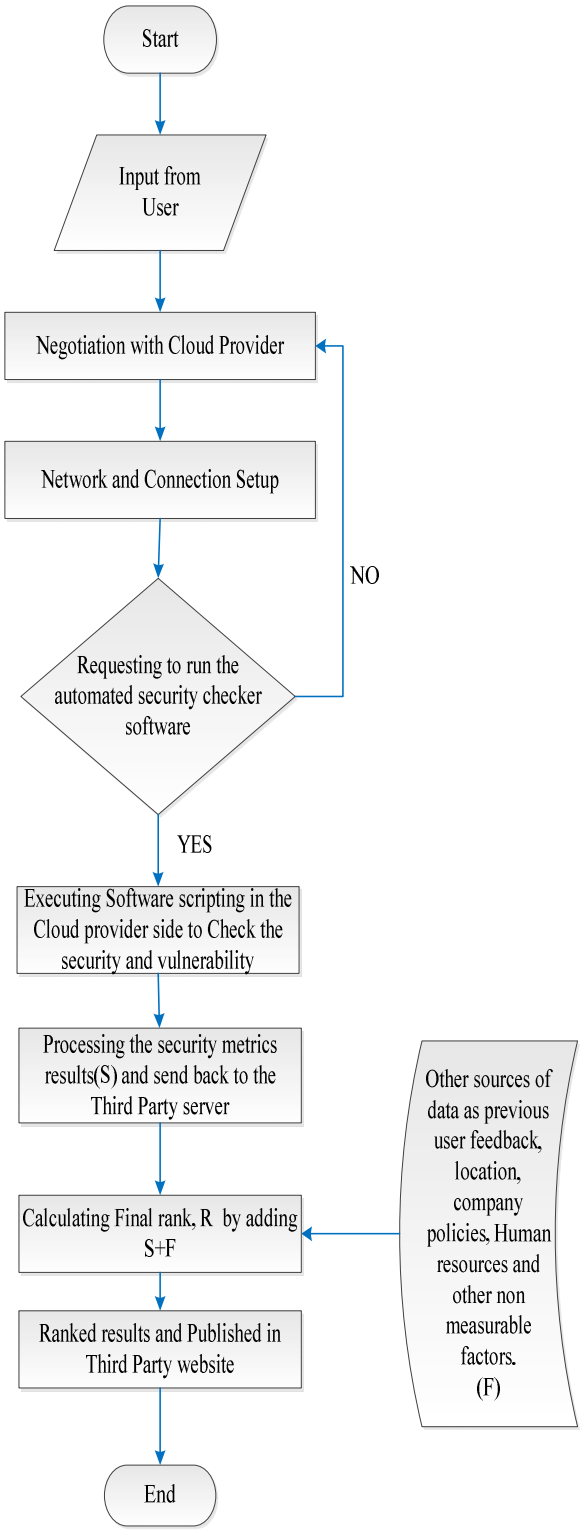



Figure 2. Flow Chart of Security Ranking system

So, the entire Regional cloud agency will be centrally monitored by ICMA which can be named as Federated cloud monitoring and this can be initiated and formed by IEEE or other technological or business consortium to provide trust, security and homogeneous cloud service facilities globally.

Conceptual model of federated Third party cloud monitoring is shown in Fig. 4.

*F. Advantages:*

In this section, the benefits and advantages of CMFCSPRS model are explained:

- Our proposed model provides an easy and convenient way to find a secure CSP by rank system
- The interpretation of ranking system is easy so customer can evaluate the ranked result for TTP website and interpreted themselves.
- No need of any brokerage or other help to find desired CSP.
- This rank system conceptually monitored by different layer of supervision and monitoring which ensures the homogeneity of rank system worldwide.
- Unavailability of one TTP will not be a problem since multiple TTPs will work in one region.
- Our proposed model ensures a new customer to adopt best secured cloud from the market.

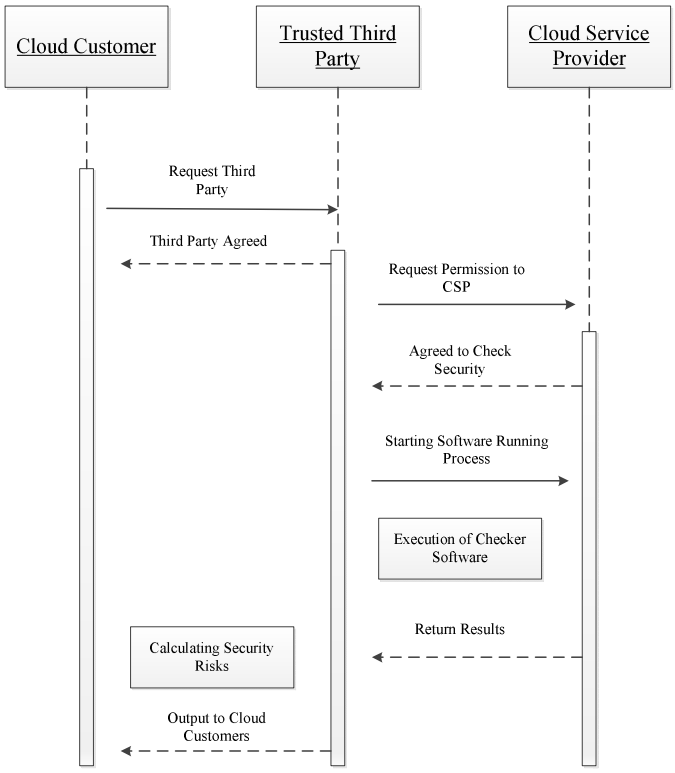



Figure 3. Sequence diagram of CSP security checking

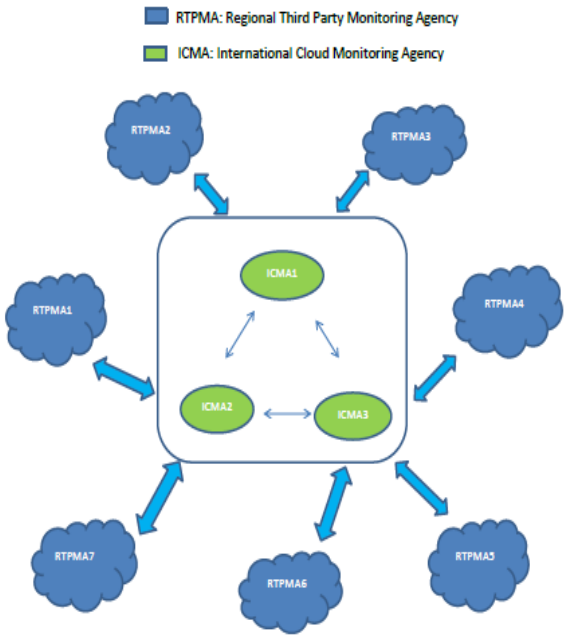



Figure 4. Conceptual model of Federated TTP monitoring

## IV. CONCLUSION

In this paper, we identify and highlight the CSP side security issues and tolerance of security strength by employing and introducing TTP which provide us CSP ranking system. To the best of our knowledge, using attack vectors to protect and ensure customer interest and confidence by issuing security ranking systems to select secure CSP is the first time in MCC. First, TTP uses automated software scripting to check security vulnerabilities in CSP side by running software scripting to break the security strength of the CSP. Therefore, considering several non-measurable metrics such as customer satisfaction, previous down time, location etc. factors, TTP announce a secured CSP ranked system in their website. We compare this TTP Cloud provider ranking system like as a credit rating agency. The deployment of TTPs is divided by region wise and then again extends up to worldwide or globally, which can be called as federated TTP cloud monitoring system. Moreover, TTPs are monitored by RTPMA and RTPMA are monitored by ICMA. Furthermore, our proposed conceptual model provides confidence to the new cloud customer to select the best cloud service provider in terms of security and trust. So, we envision a strong business feasibility of this model. Thus proposed Conceptual Model of Federated CSP Ranking System (CMFCSPRS) ensures enhanced and additional security in the MCC arena.


## ACKNOWLEDGMENT

This work is fully funded by the Malaysian Ministry of Higher Education under the University of Malaya High Impact Research Grant UM.C/HIR/MOHE/FCSIT/03

## REFERENCES


[1] Z. Sanaei, S. Abolfazli, A. Gani, and M. Shiraz, "SAMI: Service-based arbitrated multi-tier infrastructure for Mobile Cloud Computing," in *Communications in China Workshops (ICCC), 2012 1st IEEE International Conference on*, 2012, pp. 14-19.

[2] M. Shiraz, M. Whaiduzzaman, and A. Gani, "A Study on Anatomy of Smartphone," *Computer Communication & Collaboration,* vol. 1, 2013.

[3] M K Nasir and M. Whaiduzzaman, "Use of Cell Phone Density for Intelligent Transportation Systems(ITS) in in Bangladesh," *Journal Of Information Technology, Jahangirnagar University,* vol. 1, 2012.

[4] S. K. Garg, S. Versteeg, and R. Buyya, "A framework for ranking of cloud computing services," *Future Generation Computer Systems,* vol. 29, pp. 1012-1023, 6// 2013.

[5] M. Whaiduzzaman, M. Sookhak, A. Gani, and R. Buyya, "A survey on vehicular cloud computing," *Journal of Network and Computer Applications*.

[6] Z. Zibin, Z. Yilei, and M. R. Lyu, "CloudRank: A QoS-Driven Component Ranking Framework for Cloud Computing," in *Reliable Distributed Systems, 2010 29th IEEE Symposium on*, 2010, pp. 184-193.

[7] M. Alhamad, T. Dillon, and E. Chang, "SLA-Based Trust Model for Cloud Computing," in *Network-Based Information Systems (NBiS), 2010 13th International Conference on*, 2010, pp. 321-324.

[8] C. Hoi and C. Trieu, "Ranking and mapping of applications to cloud computing services by SVD," in *Network Operations and Management Symposium Workshops (NOMS Wksps), 2010 IEEE/IFIP*, 2010, pp. 362-369.

[9] S. K. Garg, S. Versteeg, and R. Buyya, "SMICloud: A Framework for Comparing and Ranking Cloud Services," in *Utility and Cloud Computing (UCC), 2011 Fourth IEEE International Conference on*, 2011, pp. 210-218.

[10] M. T. Khorshed, A. B. M. S. Ali, and S. A. Wasimi, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing," *Future Generation Computer Systems,* vol. 28, pp. 833-851, 6// 2012.

[11] R. Buyya, Y. Chee Shin, and S. Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities," in *High Performance Computing and Communications, 2008. HPCC '08. 10th IEEE International Conference on*, 2008, pp. 5-13.

[12] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," *Future Generation Computer Systems,* vol. 25, pp. 599-616, 6// 2009.

[13] A. T. Monfared and M. G. Jaatun, "Monitoring Intrusions and Security Breaches in Highly Distributed Cloud Environments," in *Cloud Computing Technology and Science (CloudCom), 2011 IEEE Third International Conference on*, 2011, pp. 772-777.

[14] J. L. Garcia, R. Langenberg, and N. Suri, "Benchmarking cloud security level agreements using quantitative policy trees," presented at the Proceedings of the 2012 ACM Workshop on Cloud computing security workshop, Raleigh, North Carolina, USA, 2012.

[15] S. M. Habib, S. Ries, and M. Muhlhauser, "Towards a Trust Management System for Cloud Computing," in *Trust, Security and Privacy in Computing and Communications (TrustCom), 2011 IEEE 10th International Conference on*, 2011, pp. 933-939.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Road Structure Analysis using GPS Information

Md. Kamrul Hasan[1], Javed Iqbal Khan[2], Raquib Ahmed[3], Mollah Md Awlad Hossain[4] and Md.Nur-Us-Shams[5]
[1]Dept. of Computer and Communication Engineering Patuakhali Science and Technology University, Bangladesh
[2]Department of Computer Science, Kent State University, USA
[3]Geography and Environmental Studies University of Rajshahi, Bangladesh
[4]GIS Division Center for Environmental and Geographic Information Services (CGIS), Bangladesh
[5]Local Government Engineering Department LGED, Dhaka-1207, Bangladesh
Email:[1]kamrul@pstu.ac.bd,[2]javed@kent.edu,
[3]raquib_ahmed@yahoo.com,[4]mmawlad@gmail.com and [5]shams.mnu@gmail.com

***Abstract*—Geospatial information is fast becoming a key data resource for a locality or country around the world. In recent years there have been major innovations in location based applications. Geographic Information Systems (GIS) are helping all spheres of science. It is spawning new public services, helping developmental planning with unprecedented precision and playing vital role in national resource management. Unfortunately however, until now there is no good publicly available comprehensive Global Positioning System (GPS) data for Bangladesh. This is the case also for many under developed localities around the world. Obtaining comprehensive GPS data is expensive. One interesting phenomena in many of these localities is the proliferation of smart phone devices, many with GPS sensors due to 3G network expansion in recent years. In this paper, we present a pilot study where these low cost devices. It can be used to potentially build GIS databases at national scale. We have designed and developed a mobile application for GPS data collection with the help of crowd. Indeed if such data is publicly available many advanced and comprehensive application are possible. Therefore we have recorded the GPS data of Patuakhali municipality road as a test case. Then we have further demonstrated its application in analyzing the road structure of Patuakhali municipality using this comprehensive GPS data. The analysis will help to predict the difficulty and costing for similar type of area mapping.**

***Keywords*—GPS information, crowdsourcing, location based application, GPS sensor, smart phone.***

## I. Introduction

A modern GIS is a system designed to capture, archive, manipulate, analyze, manage, and present all types of d ata in association with geographical location. Though it started with cartography with main application in geography, modern database and computing technology has transformed it enormously. Under the broader domain of geo-informatics, GIS along with GPS supports and delivers information to application developer, environmental managers and the public [1]. It is a powerful software technology that links unlimited amount and type of information to a geographic location. Today one of the main thrust of GIS is to map the road network objects.

The GIS data can be collected in many ways [2] ranging from control of the field survey, aerial photogrammetric, digital mapping, conversion of existing maps to digital files, selection of data and attributes, digital ortophotos, remote sensing, verification and correction of existing data, modification of data obtained by others sources. Numerous GIS software such as ArcGIS [3], GRASS (Geographic Resources Analysis Support System) [4], MapWindow GIS [5] are used today for geospatial data management and analysis, image processing, graphics and maps production, spatial modeling, and visualization.

The essential technology that helps GIS information to be correlated with graphical location data is the satellite-based GPS. The US system is made up of a network of 24 satellites placed into orbit by the U.S. Department of Defense launched initially for navigation. In 1980s, the US government made the system available for civilian use [6]. Using GPS receiver device, one can find out his three dimensional position, velocity on land, sea, and airborne round the clock in all weather, anywhere in the world. It can work with high precision and has accuracy better than other radio-navigation systems available today. Todays GPS receivers are extremely accurate due to their parallel multi-channel design. Russia operates the GLONASS (Global navigation satellite systems) satellite network for GIS information.

There are however other radio methods. Localization can be measured either via multilateration of radio signals between three or more radio towers of the network and the phone. Cell tower triangulation, for instance, is the form of geolocation for mobile device. These alternate location tracking technology often provides better results in urban areas and less precise in rural areas.

Today, there are almost unlimited types and varieties of GIS applications. The innovative applications includes zoning and land use planning, crime prevention, natural resource management, demographic analysis, disaster management planning and recovery, public health and healthcare access, tax assessment and collection, parks and open space planning, transportation and public transit planning, and historic preservation etc. In addition, location-based applications can provide interactive services to their customers based on the knowledge of their current positions. Examples of these services include continuous live traffic reports, point of interest finder (Where is my nearest fast food restaurant), and location-based advertising (Send e-coupons to all cars that are within two miles of my gas station) [7]. The growth rate of the location based application users is shown in Fig. 1.

Availability, quality of GIS information and their users vary drastically depending on localities and countries. A December 2010 study [8] shows that, 70% users are from highly developed countries (US-71%, UK-81%, Germany-69%, Canada-62%, and Japan-67%) are using the GPS based applications

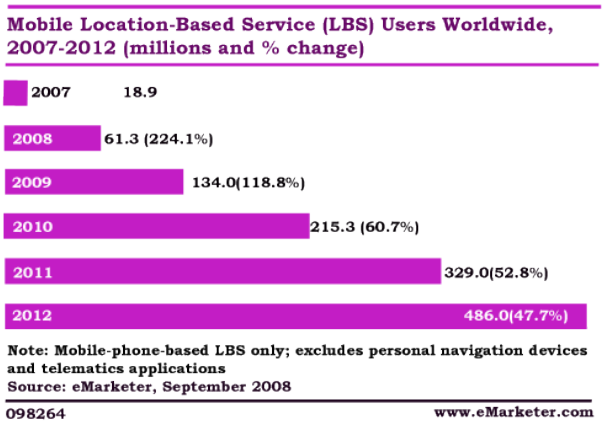



Fig. 1. Mobile location-Based Service (LBS) Users Worldwide.

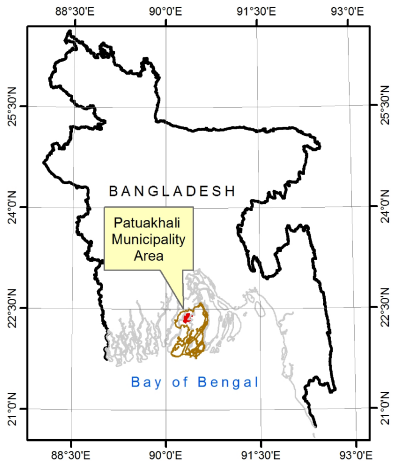



Fig. 2. Location of Patuakhali district in Bangladesh.

for traveling. On the other hand, GPS data of rural areas are not publicly available by online map providers [9],[10] in many parts of the world. India has street view GPS map nearly for 4,787 cities collected by a private venture. Not only that, even where some rudimentary coverage is available, details of the divisional cities are not well mapped yet. Often few of the main roads are mapped using satellite image digitization and GPS survey. Many are proprietary and thus this data are not available to the application developers under any easy term. In reality, GPS data is a major obstacle.

This focus of this pilot project is to explore how the complexity and various costs can be minimized to develop the GIS database using mobile devices. As part of this, we have developed an android based mobile application (app) for GPS data collection with crowdsourcing. The crowdsourcing approach has the potential to allow a locality to populate GPS data of vast area in short time with manageable cost. Consequently, online community can be involved to upload the GPS data from their place of interests (POI).

As part of experiment, we have mapped the municipality of Patuakhali, southern district of Bangladesh, using GPS data. The area of the municipality is 32.36 Sq.Km. (Core Area-19.41 Sq.Km & Fringe Area-12.95 Sq.Km) shown in Fig. 2. We have recorded the road types, length, and width as tagged featured data. Then we have analyzed the road length, type, distribution, and density based on the GPS data. Different type of road distribution require specific surveyor, data accuracy and travel charges. This pilot analysis can help predicting mapping issues for other areas of similar in future.

The rest of this paper is organized as following. The Section II provides the state of the art. The mobile application development process and GPS data collection methodology is illustrated in the Section III. The analysis on the collected GPS data is illustrated in the Section IV and finally, Section V summarizes the project findings.

## II. Problem Statement

Indeed, there are few large GPS road network databases available online. Third party like Google, Bing and Yahoo are providing some online maps. But their mapping coverage is extremely sparse for developing countries including Bangladesh. These maps only have the main roads of urban areas. The branch roads and lanes are utterly insufficient. The rural area roads are completely absent, whereas countries like Bangladesh are predominantly rural. Also, the feature data of the road structure is inadequate. But there is enormous potential use of GIS applications by villagers. As an illustration, GIS applications can be envisioned for crop disease detection and immediate response. In the remote location, if the farmer observes the crop infection then they can capture the picture of the diseased crop and upload it to specialist website. The experts then not only can advise the farmer but various control measures can also be launched. If several farmers from same vicinity are facing similar crop contamination then the respective authority and researchers can be notified for immediate response. It is also extremely useful in development planning and maintenance of rural infrastructure. It is also useful by many non-governmental organizations (NGO) who are involved in various development services.

### A. Related works

In Bangladesh, the first GPS based GIS mapping reportedly started in 1991 following the major floods of 1987 and 1988 as part of Flood Action Plan (FAP) project. It made limited use of GPS data mainly for geo-referencing and registration of satellite images with existing maps at national levels. Such pioneer mapping projects however didn't cover the road structure.

In Bangladesh, Local Government Engineering Department (LGED) [12], [11] is believed to have the most detailed map base for roads in civil domain. LGED maps are on geographic coordinate system. LGED has attempted to map local road of Upazila and Union level since 1998 using traditional survey. They have used old system of surveyors armed with GPS devices. Unfortunately, the dataset is riddled with incompleteness and inaccuracies. During previous data collection, there were many situations where GPS surveyors had cancelled visiting of many regions below the level of Union and Upazila (administrative units in Bangladesh). These would be due to weather condition, vehicle problem or to cut the expenditure or due to lack of time, etc. In many such instances, surveyors inserted planned roads by drawing

it directly on the computer map editor. Finally these were uploaded onto LGED-GIS archive. LGED has also collected road data from Roads and Highways Department (RHD) of Bangladesh. Significant part of its GIS data for old roads have been derived using digitizer and sometimes drawing in ArcView software.

Todays LGED field engineers are facing these anomalies while trying to use the data. Often they observe mismatch between actual road length and coordinates from the archived GPS data. LGED has undertaken recent initiatives to address some of these problems. In the current GIS archive, LGED too has limited number of branch roads, lanes in the villages. In addition, tracks through the forest, cultivation land, and shortcut communication ways are very few in their map. They couldn't cover all the municipality and city corporation area map. Again, these LGED data despite their insufficiencies are not accessible for other application developers. Only some raw image maps are provided into public domain. This prevents any possibility of advanced geographic analysis or GIS application. Generally speaking, no comprehensive GPS mapping data is available for Bangladesh. Few other agencies have partial data specific to their need.

Among the international initiatives Open Street Map (OSM) [9], a notable attempt, is trying to give solution for free mapping data. Many projects have been benefited by the OpenStreetMap in the recent year. For example, the MapKibera project [14] and WikiProject Haiti [13] that shows conspicuous success during emergency.

In Bangladesh, there is already demand and user awareness for the location based application regarding public health sector, disaster response and management, helping the disabled people, transportation, pathfinder for the volunteers, tourism, and research [15]. In 2010, OpenStreetMap foundation donated few GPS devices to BUET students. Later few more devices are sent to CUET students. BUET uploaded road traces of their campus and few different places in the capital city of Dhaka and few locations outside Dhaka district in Bangladesh. Their mapping interests are hindered by verities problem like GPS device unattainability, lack of technical knowledge of GPS devices, problem of GPS data transmission, and deficiency of location based application development. The system we propose in this paper seems to be one of the first mobile platform based GPS information collection demonstration at scale.

### *B. Research challenges*

Mapping of vast rural and semi-urban areas of Bangladesh including all types of roads to achieve the goal, in particular with specialized GPS device and conventional surveyor team will face several challenges. Some of these are illustrated in the following.

(1) Manpower: The manpower should be motivated, trained and managed for data acquisition. GPS data recording, POI selection, rout creation, data transfer, and data testing are the fundamental steps in this management system.

(2) Privacy and security: There are many concerns about privacy during data collection. Neighborhood people are often apprehensive about the effort and ask around the GPS data collector with many peculiar inquiries that makes the surveying awkward in the street.

(3) Data accumulation: The surveyors may have to travel different remote places for data collection. So, the GPS data should be recorded and saved safely. In addition, data should be accumulated from each regional user perfectly.

(4) Waypoint mapping: Recording the waypoints names and other meta-data can be cumbersome. This method increases the time and irritates maximum number of GPS users.

(5) Data accuracy: The accumulated data can be wrong due to the lack of GPS users technical experience, GPS signal strength, data transfer problem, and wrong identification of location name etc.

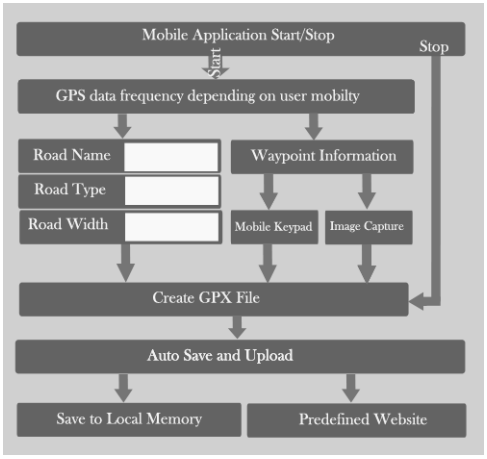



Fig. 3. Mobile application development processes.

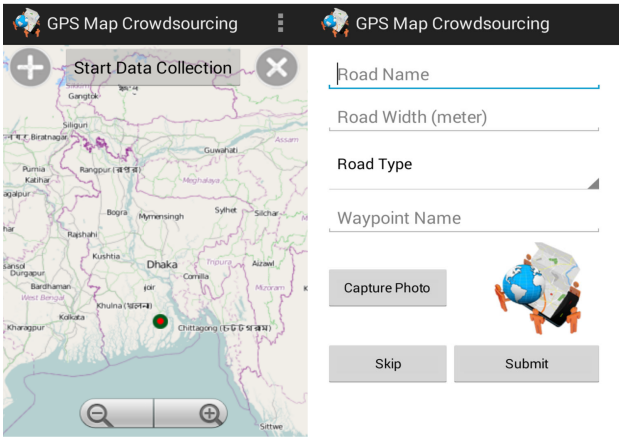



Fig. 4. Android based mobile application screenshot.

## III. GPS DATA COLLECTION METHODOLOGY

Todays most smart phone devices have GPS chip and allow one to trace an area using satellite signal. These phones also allow programmable apps. We have designed a crowdsourcing focused mobile application. The architecture of the system is shown in Fig. 3. The developed mobile app based on this flowchart is shown in Fig. 4. There are select few configuration parameters. If activated, the GPS data collection frequency is

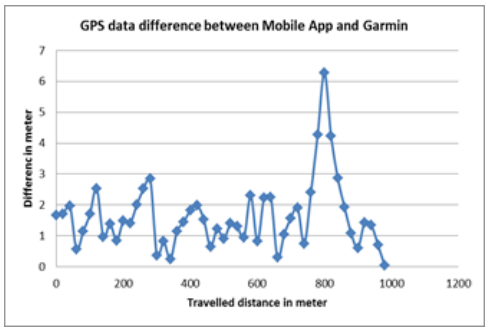



Fig. 5. Distance between two abut GPS traces recorded by Garmin and mobile application.

taken as input from user. The smart mobile phone record GPS data at very high frequently when it is turned on. For this reason, it needs to be configured at start for power and data efficiency. The users are allowed to choose one option from low, medium and high frequency. GPS data will be recorded fast when the user is moving on motor vehicle. Again, the user walking on foot should select low frequency to collect GPS data moderately. Afterwards, the value regarding name, type and width of the road can be inserted with waypoint information in this application. For quick data collection, user can capture image of waypoint without writing the POI information in the mobile application. The GPS data of road and waypoints are saved in the mobile memory pressing submit button. In addition, the GPS data are uploaded automatically from the users mobile to the predefined website when certain amount of data are recorded. If the internet is unavailable during GPS data transmission then the GPS exchange format (GPX) file is kept inside the memory to upload later.

This mobile application can perform almost all major tasks of a GPS device [16] and indeed can support additional functionalities. On trial basis, the mobile application was handed out to the smart phone users after a training and motivation session. To try out the crowdsourcing approach, the mobile app is used by the students of Patuakhali Science and Technology University (PSTU) and more importantly few local residents of Patuakhali districts. Inspired by the advantages of mobile app they have collected their own regional GPS data using the mobile application. To verify the technical accuracy of the mobile GPS traces, we have tested out the data with traces captured by Garmin GPS device. Fig. 5 plots the differences between these two traces. Plotting both GPS traces, we found that the average difference to be +/- 1.62 meter, variance is 1.21 meter. The pick distance recorded was 6.28 meter. Maximum difference is recorded between Mobile and Garmin traces when the GPS signal strength is not adequate. This seems to be the well within the margin of typical human errors. Now we describe the advanced and the customized road structure analysis that was possible using the traces.

## IV. PATUAKHALI ROAD STRUCTURE ANALYSIS USING GPS INFORMATION

We have recorded five types of roads during the mapping of Patuakhali municipality. The roads are classified by Government as Bituminous Carpeting (BC) and Roller-Compacted Concrete (RCC), Haring Bone Bond (HBB), Brick Flat Soling

TABLE I. ROAD DISTRIBUTION OF PATUAKHALI MUNICIPALITY.

| SL | Road Type | Length (Km) |
|---|---|---|
| 1 | Dirt Road | 2.75 |
| 2 | BFS Road | 1.12 |
| 3 | HBB Road | 6.89 |
| 4 | RCC Road | 31.54 |
| 5 | BC Road | 42.36 |
| Total: | | 84.67 Km |

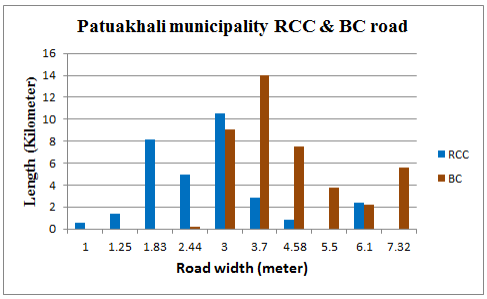



Fig. 6. RCC and BC road distributions in the Patuakhali municipality area.

(BFS) and dirt road. Launch ghat is identified as the city center (central business district) since this riverine municipality grew based on this location. Extensive commercial places are situated near the city center. Schools, colleges, administrative offices, clinic, hospitals are situated further away. Residential area, public and private organizations are limited in close proximity of city center. Wider paved roads are connecting these establishments to the center point. Comparatively narrow paved roads are used as branch road and lane. The branch roads and lane have connected the residential areas to the main communication way of the municipality.

We also observed that different types of streets have various width and lengths (Table I) in the municipality according to the recorded GPS data. The captured data is presented in the Table I. About 85 Kilometer GPS road data is traced in Patuakhali municipality. It is interesting to note that the widths of some the roads are found to be non-standard. Normally roads are supposed to follow government standard and municipality recognized width. For example, 7.32, 5.50, 3.7, 3.0 meters are standard sizes mentioned by LGED and RHD. However, 6.10, 4.58, 2.44, 1.25 and 1.00 meter spacious roads are also visible in the municipal area. It seems local municipality has indeed extended the width of urban and sub-urban area roads further especially for the newer roads- possibly based on availability of fund and deter encroachment.

In the municipal area, maximum length of road is made of bituminous carpeting (42.36 Km) having divergent size. Secondly, many RCC roads are also found (31.54 Km). In addition, 10.76 Km dirt, BFS, and HBB roads have been recorded. 3.00 meter is a pivot width because RCC roads are plentiful up to this range. Again, adequate numbers of BC roads have 3.00 meter or more wide (shown in Fig. 6). Current municipal authority is rebuilding and developing new RCC road instead of BC streets to withstand the local weather. Patuakhali has high rainfall as well as salty climate. For this reason, more of RCC roads are caught in sight here. During this same conversion roads are also widened. These are the 6.10 meter wide roads noted- which were not in record in previous databases. We also found almost no roads below

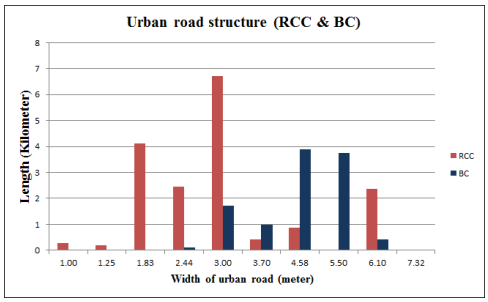



Fig. 7. RCC and BC road distributions in the urban area of Patuakhali.

TABLE II. DIFFERENT ROAD TYPES-LENGTH(METER) IN URBAN AREA.

| Width(meter) | Dirt | BFS | HBB | RCC | BC |
|---|---|---|---|---|---|
| 1.00 | 47.5 | 0.0 | 0.0 | 271.4 | 0.0 |
| 1.25 | 0.0 | 0.0 | 0.0 | 192.8 | 0.0 |
| 1.83 | 330.9 | 289.9 | 319.1 | 4101.0 | 0.0 |
| 2.44 | 20.0 | 0.0 | 0.0 | 2447.0 | 116.0 |
| 3.00 | 0.0 | 0.0 | 30.0 | 6702.0 | 1726.0 |
| 3.70 | 0.0 | 0.0 | 0.0 | 416.0 | 973.7 |
| 4.58 | 0.0 | 0.0 | 0.0 | 869.3 | 3891.0 |
| 5.50 | 0.0 | 0.0 | 0.0 | 0.0 | 3734.0 |
| 6.10 | 0.0 | 0.0 | 0.0 | 2366.0 | 421.0 |

1.83 meter. Presently, the municipal authority does not allow building RCC roads below 1.83 meters. So, 4.58 meter width RCC streets are noticeable in the municipal urban area. But, short distance of 1.00 and 1.25 meter wide RCC roads are still alive in Patuakhali since these were built a decade ago.

Patuakhali municipality area is subdivided into three regions for our analysis. Launch ghat (river port) is taken as the city-center. Two circles were drawn and approximately the halfcircles (of the city side of the river were zoned). Zone within a radius of 1.25 km is defined as urban area. The zone from 1.25- 2.6km radius is considered as sub-urban area. Rest of the municipality region is identified as rural area. Patuakhali municipality area is subdivided into three regions for our analysis. Launch ghat (river port) is taken as the city-center. Two circles were drawn and approximately the halfcircles (of the city side of the river were zoned). Zone within a radius of 1.25 km is defined as urban area. The zone from 1.25- 2.6km radius is considered as sub-urban area. Rest of the municipality region is identified as rural area.

Fig. 7, 8 and 9 shows the road classifications of these areas. In the urban section, RCC and BC roads have been expanded highly than other types. The length of RCC road is recorded as 17.37 Km. In addition, 10.86 Km bituminous carpeting roads are established in the same area shown in Fig. 7. RCC roads are plenty in length than BC road because they are rebuilt on BC road and made it wider in urban area. Very limited number of brick made (HBB, BFS) road is recorded here. In addition, BC road below 2.44 meter breadth were not found in the GPS description shown in Table II.

In the sub-urban area, dirt road, BFS, HBB, RCC and carpeting roads are recorded but in different ratio. Here we have observed many BC roads than the RCC roads. Apparently the municipal authority didn't change the material when it rebuilt the street. Here many roads are newly developed using bricks that are presented in the Fig. 8. Again, dirt roads are

TABLE III. DIFFERENT ROAD TYPES-LENGTH(METER) IN SUB-URBAN AREA.

| Width(meter) | Dirt | BFS | HBB | RCC | BC |
|---|---|---|---|---|---|
| 1.00 | 1074.0 | 21.5 | 0.0 | 50.0 | 0.0 |
| 1.25 | 176.9 | 0.0 | 0.0 | 705.5 | 0.0 |
| 1.83 | 141.8 | 212.9 | 0.0 | 4067.8 | 0.0 |
| 2.44 | 365.3 | 404.6 | 0.0 | 1303.9 | 101.0 |
| 3.00 | 0.0 | 0.0 | 5610.0 | 3814.3 | 6007.0 |
| 3.70 | 0.0 | 0.0 | 0.0 | 2292.9 | 6566.0 |
| 4.58 | 0.0 | 0.0 | 0.0 | 0.0 | 2132.0 |
| 6.10 | 0.0 | 0.0 | 0.0 | 0.0 | 1300.0 |
| 7.32 | 0.0 | 0.0 | 0.0 | 0.0 | 1760.0 |

TABLE IV. DIFFERENT ROAD TYPES-LENGTH(METER) IN RURAL AREA.

| Width(meter) | Dirt | BFS | HBB | RCC | BC |
|---|---|---|---|---|---|
| 1.00 | 0.0 | 0.0 | 0.0 | 188.0 | 0.0 |
| 1.25 | 268.0 | 0.0 | 0.0 | 453.0 | 0.0 |
| 1.83 | 341.2 | 0.0 | 0.0 | 0.0 | 0.0 |
| 2.44 | 0.0 | 200.0 | 930.0 | 1198.0 | 0.0 |
| 3.00 | 0.0 | 0.0 | 0.0 | 0.0 | 1350.0 |
| 3.70 | 0.0 | 0.0 | 0.0 | 100.0 | 6472.0 |
| 4.58 | 0.0 | 0.0 | 0.0 | 0.0 | 1462.0 |
| 5.50 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 6.10 | 0.0 | 0.0 | 0.0 | 0.0 | 513.0 |
| 7.32 | 0.0 | 0.0 | 0.0 | 0.0 | 4040.0 |

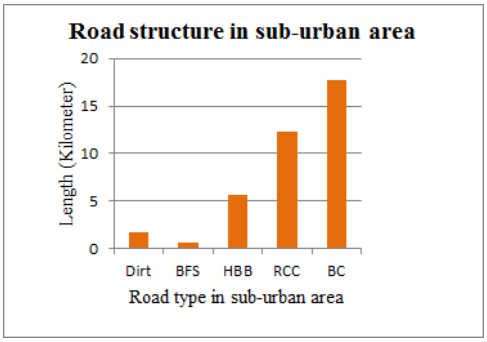



Fig. 8. Road distributions in the sub-urban area of Patuakhali.

created by mud, sand, silt, and rubbish to make it useable year round. Here, dirt roads as well as brick roads are made up to 3.00 meter wide. One A-category (7.32 meter wide) road was found in the sub-urban area. It has connected the city to the national highway. The length of different road types of sub-urban region is shown in Table III. We found there is not much narrow BC roads nor wide-expansive RCC roads are in sub-urban areas. In the sub-urban area, dirt road, BFS, HBB, RCC and carpeting roads are recorded but in different ratio. Here we have observed many BC roads than the RCC roads. Apparently the municipal authority didnt change the material when it rebuilt the street. Here many roads are newly developed using bricks that are presented in the Fig. 8. Again, dirt roads are created by mud, sand, silt, and rubbish to make it useable year round. Here, dirt roads as well as brick roads are made up to 3.00 meter wide. One A-category (7.32 meter wide) road was found in the sub-urban area. It has connected the city to the national highway. The length of different road types of sub- urban region is shown in Table III. We found there is not much narrow BC roads nor wide-expansive RCC roads are in sub-urban areas.

National highway has passed through the rural region of Patuakhali municipality and it has gone towards Kuakata,

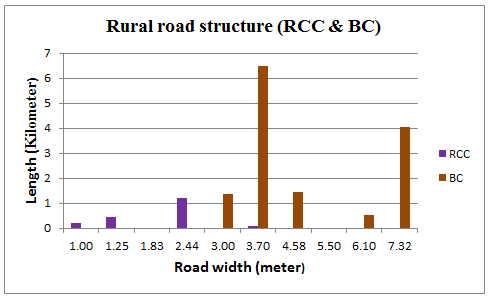



Fig. 9. RCC and BC road distributions in the rural area of Patuakhali.

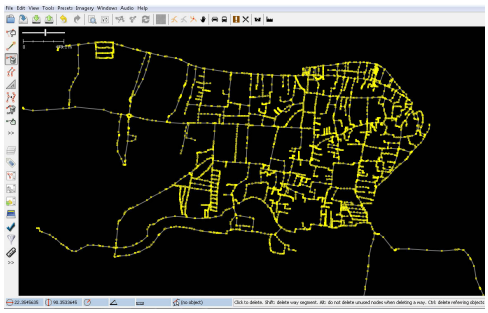


Fig. 10. Patuakhali municipality map using GPS raw data on JOSM editor.

the famous southern tourist spot of Bangladesh. Category-A, different width of BC, RCC, BFS, HBB and dirt roads recorded in rural area are presented in Table IV. According to Fig. 9, 3.70 and 7.32 meter wide roads are found ample in rural area.

$$\beta = \frac{\sum e}{\sum v} \tag{1}$$

In Patuakhali municipality, the road density (beta) index is measured by (1). Where, $\beta$=beta index, $\sum e$= total number of roads and $\sum v$= total number of POI. The beta index value is calculated 1.35, 1.15 and 1.10 for urban, sub-urban and rural area respectively. This infers that the proliferation of road in the urban area is more than the other region of the same district. Again, there are a lot of crossings coalesced by three, four or five roads in this municipality. Among these crossings, 41 crowded crossroads are recorded as active junction where 29 busy intersections are in urban area, 10 intersections are in sub-urban area and 2 connections are in rural area. The maximum diameter (from city center to most far distance location) of the Patuakhali municipality is measured as 4.50 Km. The GPS map data of Patuakhali municipality is shown in Fig. 10. Here, the yellow points are the GPS coordinates of each path presented on the Java Open Street Map (JOSM) editor [17].

## V. CONCLUSION

In this test project we have collected comprehensive road network data for Patuakhali town. The data had been collected using relatively simple and inexpensive effort. In contrast, conventional method for GPS data acquisition is complex and costly. At the heart of this success lies the wide availability of smart mobile phones and crowdsourcing approach. We have also shown a sample analysis of the township based on the road network. This is not only novel for this locality but possibly one of the first such attempts in Bangladesh. We have now made this complete Patuakhali road network database available for researchers. Indeed, through this database Patuakhali municipality has now been reintroduced to the researcher, developers, planner, administrator and all inviting numerous other beneficial analysis and applications potentially serving the locality. We believe other communities in developing world can also benefit from such effort.

## REFERENCES


[1] Awange, Joseph L., and John B. Kyalo Kiema. "Environmental geoinformatics: monitoring and management." (2013).

[2] Lwin, Ko Ko, Ronald C. Estoque, and Yuji Murayama. "Data Collection, Processing, and Applications for Geospatial Analysis." *Progress in Geospatial Analysis*. Springer Japan, 2012. 29-48.

[3] Kennedy, Michael D. *Introducing Geographic Information Systems with ArcGIS: A Workbook Approach to Learning GIS.* Wiley.com, 2013.

[4] Neteler, Markus, et al. "GRASS GIS: a multi-purpose Open Source GIS." *Environmental Modelling & Software* 31 (2012): 124-130.

[5] Ames, Daniel P., et al. "MapWindow GIS." *Encyclopedia of GIS.* Springer US, 2008. 633-634.

[6] Huang, J. Y., C. H. Tsai, and S. T. Huang. "The next generation of GPS navigation systems." *Communications of the ACM* 55.3 (2012): 84-93.

[7] Chaffey, Dave, and Paul Russell Smith. *eMarketing eXcellence: Planning and optimizing your digital marketing.* Routledge, 2013.

[8] Mobile Marketing Association. "Mobile Location Based Services Marketing Whitepaper." *New York* (2011).

[9] Neis, Pascal, and Alexander Zipf. "Analyzing the contributor activity of a volunteered geographic information projectThe case of OpenStreetMap." *ISPRS International Journal of Geo-Information* 1.2 (2012): 146-165.

[10] Fuquan, Peng, et al. "A City Modeling and Simulation Platform Based on Google Map API." *Proceedings of the 2011, International Conference on Informatics, Cybernetics, and Computer Engineering (ICCE2011) November 19-20, 2011, Melbourne, Australia.* Springer Berlin Heidelberg, 2012.

[11] Fujita, Yasuo. "What Makes the Bangladesh Local Government Engineering Department (LGED) So Effective?." (2011).

[12] LGED (2012). About GIS Unit - LGED, Local Government Engineering Department (LGED). [Online]. Available at [http://www.lged.gov.bd/UnitAbout.aspx?UnitID=5]. Retrieved on 12.06.2012.

[13] (2011) Project Haiti. [Online]. Available: http://wiki.openstreetmap.org/wiki/WikiProject Haiti

[14] Hagen, Erica. "Mapping Change: Community Information Empowerment in Kibera (Innovations Case Narrative: Map Kibera)." *innovations* 6.1 (2011): 69-94.

[15] Ridwan, Sohaila Binte, Hasan Shahid Ferdous, and Syed Ishtiaque Ahmed. "The challenges and prospect of OpenStreetMap in Bangladesh." *Computer and Information Technology (ICCIT), 2011 14th International Conference on.* IEEE, 2011.

[16] Hess, Basil, et al. "Evaluation of fine-granular GPS tracking on smartphones." *Proceedings of the First ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems.* ACM, 2012.

[17] Beiglboeck, Kristof. *Programming Gps and Openstreetmap Applications with Java: The Realobject Application Framework.* CRC Press, 2012.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Performance Improvement Techniques For RSSI Based Localization Methods

Md Al Shayokh[1], Ugur Alkasi[2] and Hakan P Partal[1,2]
[1]Department of ECE, Yildiz Technical University, Istanbul, Turkey
[2]Department of EECS, Syracuse University, Syracuse, NY, USA
shayokh_2055r@yahoo.com, uguralkasi@hotmail.com, hpartal@ieee.org

***Abstract*—Performance improvement techniques for two popular RSSI based indoor localization methods have been studied experimentally by using Wi-Fi modems. The improvement of RF Fingerprinting and RSSI Multi-lateration methods have been suggested from different aspects for both line-of-sight (LoS) and non-line-of-sight (nLoS) medium in an indoor environment. Various testing scenarios have been examined for comparison of the two methods, as the performance level in RF Fingerprinting is mainly depend on the number of modems, as well as the density of training data and, the multilateration method is mainly depend on correctly modelling of the path loss exponent. Optimizing and defining a unique path loss exponent for each of the wireless transmitter modems, testing in a LoS and a nLoS medium, changing the number of transmitters, etc, have been tried and performance plots have been shown for comparison purposes.**

***Keywords*— *LBS, Indoor localization, RSS(Receive Signal Strength), RF fingerprinting, lateration technique, WiFi based localization,Signal strength to distance conversion.***

## I. INTRODUCTION

As computing and communications equipment becomes ever more powerful and cheaply available they are starting to pervade all areas of life. In order to make applications more useable products are being developed that have the ability to detect and respond to the context in which they are being used, for example the location of the user. Localization using radio signals was first introduced in the World War II to locate soldiers in emergency situation. During the war in Vietnam the Global Positioning System (GPS) was introduced and became available for commercial applications in the 90s of the last century. [1]Although it is the most popular positioning system for open outdoor environments, there is an unmet need for a reliable positioning system that can work indoors, where the microwave radio signals used by the GPS are greatly attenuated.[2]

The need for LBS has been increasing everyday with several applications such as monitoring services for customers and helping them find their desired products, discounts and promotional offers in shopping malls, airports and other indoor environments.[3] Applications also include information services in museums, navigation services for indoor locations and shopping mall optimizations. There are various methods available to achieve a precise positioning. The time of arrival (TOA), angle of arrival (AOA), time difference of arrival (TDOA), and received signal strength (RSS) methods are well known reported techniques. Among these, TOA, AOA and TDOA based system offer high accuracy of determining a mobile user's position. The angle of arrival (AOA) is the common metric used in direction-based systems, where additional hardware is needed in order to measure the angle of incidence of the received signal. The received signal strength (RSS) methods use the signal propagation models in estimation of distance of transmitter and receiver. Time of arrival (TOA) and time difference of arrival (TDOA) of the received signal are the metrics used for estimation of distance between transmitter and receiver both require precise clock synchronization and expensive infrastructure.[4]Due to the necessity of expensive hardware system additionally needs to be installed, these methods are may not be suggested. Researchers are going on to make this service less cost effective. It has been found that Wi-Fi RSS based localization techniques offer no additional hardware, since usual Wi-Fi modems are usually available for most indoor shopping malls.[5] Comparing to techniques like TDOA or TOA the performance rate in RSS based methods remain insufficient which makes researcher to develop new techniques and improve the performance. Previous study in literature has been proved that these techniques could be improved by various research techniques. Some experimental comparison study has been performed to find out the better solution.[6] In this work, the RSS measurements are used for improving the performance of two different techniques which are RF fingerprinting and multilateration techniques. The purpose of this paper is to present some techniques for the improvement of these two methods for finding a better accuracy rate and low error rate from a cost effective aspects in an indoor environment.

## II. TEST BED AND EQUIPMENTS

In this study, test bed has been selected at the School of Electrical and Electronics Engineering hallways in Yildiz Technical University. The test area has a dimension of about 6m by 24 m area of a ground floor. Normally, 5 modems can be detected in this area. During the experiment,

Wi-Fi scanner tool on a Macintosh Computer with a high sensitivity Wi-Fi receiver antenna is used for data collection.

## III. METHODOLOGY AND PERFORMANCE

*RF Fingerprinting method:*

A common approach for the localization of a handheld terminal or mobile device by means of Wi-Fi is based on measurements of RSS (Receive Signal Strength) of the Wi-Fi signals from the surrounding access points at the terminal. This information is available due to the beacon broadcast multiple times a second by every access points. An estimate of the location of the terminal is then obtained on the basis of these measurements and a signal propagation model inside the indoor environment. The propagation model can be obtained using simulations or with prior calibration measurements at certain locations. In the second case, the measured signal strengths values at a certain location in the test bed are compared with the signal strengths values of calibrated points stored in a database.[7] RF fingerprinting method is constructed in two steps which are offline and online phases. In the offline phase, certain amount of locations will be chosen, depending on the size and layout of the test bed. At each of these locations, a number of calibration measurements will be performed. This is due to the fact that the orientation of the user affects the RSS value measured by the Wi-Fi device. For example, if the user's physical location is between the access point and the mobile device, the measured signal strength will probably be smaller compared to the situation where the user positions itself on the opposite side of the device. This is due to the fact that the signal is attenuated by the human body. Then a database is created by these recorded the received signal strength values transmitted from the Wi-Fi modems. These constituted a signal strength map or the RF fingerprints. Figure 2 illustrates the RSS values of 4 Wi-Fi modems at different points in the test bed. In this experiment, at first 4 Wi-Fi modems have been taken which located in four lab rooms of a rectangular hallway where the measurement intervals are set 1.2 meter and the total number of reference points are 126.

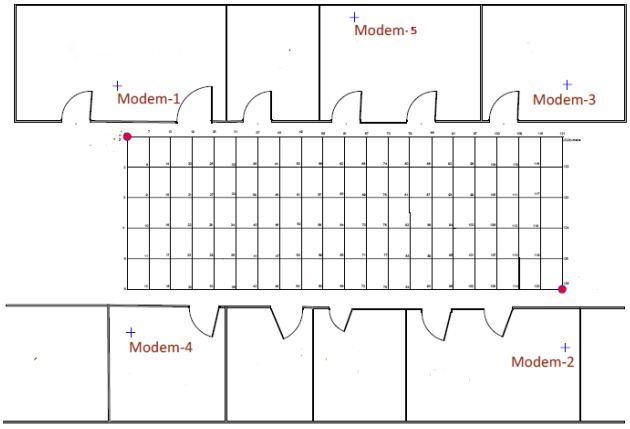



Figure 1: The test location layout with positions of 5 modems

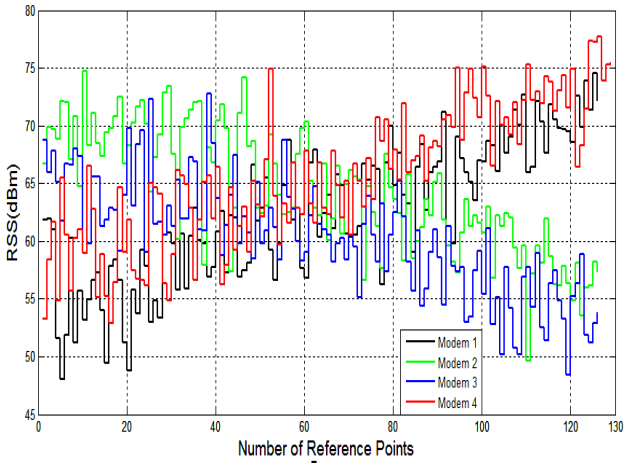



Figure 2: RSS value of 4 Wi-Fi Modems at different reference points

For each grid point, 12 RSS data collected with a time interval of 5 seconds and recorded in a database. Total a number of 6048 data was collected and recorded for constructing a radio map. The online phase is the phase where the calculation software (Wi-Fi Scanner) periodically receives measurements from one or more mobile devices. In other words, real time RSS values were measured and these data were compared against the previously collected offline RF fingerprints. Then, the mobile user's location is estimated by using some classification algorithms in Matlab. The History based K-NN algorithm is used which provides low estimation error with the help of the nearest neighboring location data. K-NN algorithm is a non parametric learning method for classifying objects based on closest training examples in the feature space.

In this experiment, firstly collected real time data from the mobile user are compared against the offline database. With the help of a tracking system based on K-NN algorithms, RF Fingerprinting shows a result of 1.5 meter accuracy with a 55% a probability and 4.8 meter accuracy with 96% probability which is relatively fair. Later, another modem has been added to justify the performance in same condition. This time it shows a better

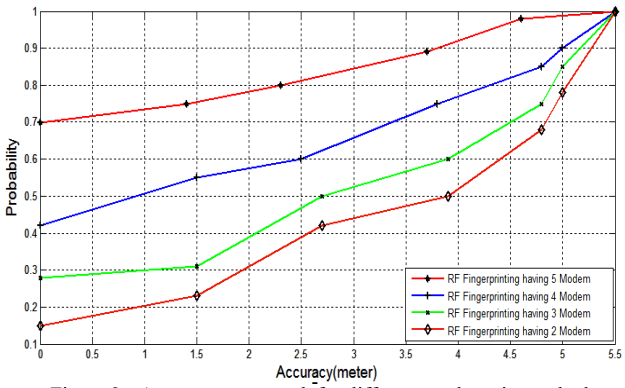



Figure 3: Accuracy rate graph for different modems in test bed having a step interval of 1.2 meter

accuracy rate with a 75% probability in 1.4 meter accuracy and 98% probability in 4.6 meter accuracy. Same experiment has been performed with the presence of 3 and 2 modems respectively. This time it has been noticed that the accuracy rate becomes relatively low than previous which led the importance of modem number for finding better accuracy. A Matlab graph shows the results of RF Fingerprinting methods in Figure 3.

In second stage, the step interval considered as 1 meter to increase the density of training data in test bed. This time the density of collected database increased significantly. Later, having 4 modems real time data compared with the collected database data. This time a noticeable accuracy rate has been found. The accuracy rates go higher having 1.2 meter accuracy with 70% probability and 4.2 meter accuracy with 95% probability. Having 5 modems in test beds the result shows better result than the previous one with 1 meter accuracy with 78% probability and 4 meter accuracy with 97% probability. Same experiments have been done with the presence of 3 and 2 modems respectively. A Matlab graph in Figure 4 illustrates the results.

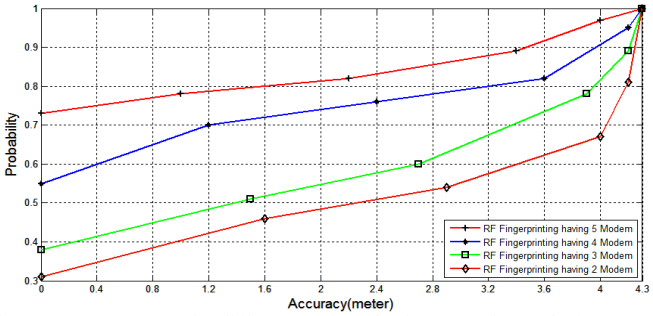



Figure 4: Accuracy rate for different modems having a step interval of 1 Meter

This graph led to a new solution for improving the performance of RF Fingerprinting in an environment. It illustrates that having a maximum number of modes and a great density of data in a database which refers a relatively small step interval in RF Fingerprinting methods could give a preferably better solution than past.

*RF Multilateration Method:*

RF Multilateration is one of those methods which are the focus of much research that is to use the received signal strength from Wi-Fi modems or access points. Multilateration estimates the position of the device of interest by using the strength of signals received from several non-collocated, non collinear transmitters. In this method, a conversation from SS to distance is needed which plays a vital role to find the exact position in an indoor environment. Different experiments have been done before on multilateration method in an indoor system in a LoS medium. In this experiment, this is the first time when multilateration methods was done in an nLoS medium for improving and justify the performance of this method. Firstly, the test has been done in a LoS condition. Signals are recorded from 1 meter to 30 meter with 1 meter steps having a time interval of 5 seconds. In order to integrate Wi-Fi positioning determination with other location techniques not only on the coordinate level it is necessary to convert the measured signal strength values at one location to a range or distance to an access point. Then, it would be possible to perform a multilateration using distances to several access points or radio transmitters. In multilateration, path loss exponent plays an important role as well as SS to distance conversion. In a LoS system, modeled path loss exponent is as same as air. After modeling the path loss exponent for a LoS environment, RF multilateration methods have been performed in the same test bed with 4 modems in a LoS environment. Performance rate in accuracy shows poor result by using this technique. Using this method in a fairly longer test bed accuracy in 5 meter with a 5% probability and accuracy in 9 meter with a 25% has obtained.

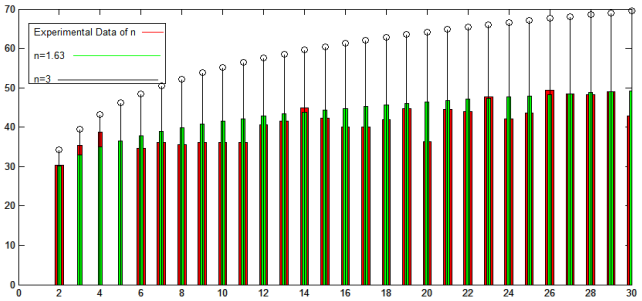



Figure 5: Path Loss exponent modeling for LoS environment

After modeling the path loss exponent for a LoS environment, RF multilateration methods have been performed in the same test bed with 4 modems in a LoS environment. Performance rate in accuracy shows poor result by using this technique. Using this method in a fairly longer test bed accuracy in 5 meter with a 5% probability and accuracy in 9 meter with a 25% has obtained. Later, same techniques have been performed using 5 modems. This time result in accuracy shows better but not satisfactory result. Using 5 modems accuracy of 4 meters with a 11% probability has been obtained. Same techniques have been performed using 3 and 2 modems respectively. This time the performance rate shows very poor result than previous one. This led us to a decision of using more modems for getting a better solution in an indoor environment to perform this method. Results have been shown by a graph in Figure 6.

This graph illustrates and points out that the improvement of RF Multilateration could be improvised by increasing the modem numbers. Later, a smaller area of test bed has been selected to perform the same

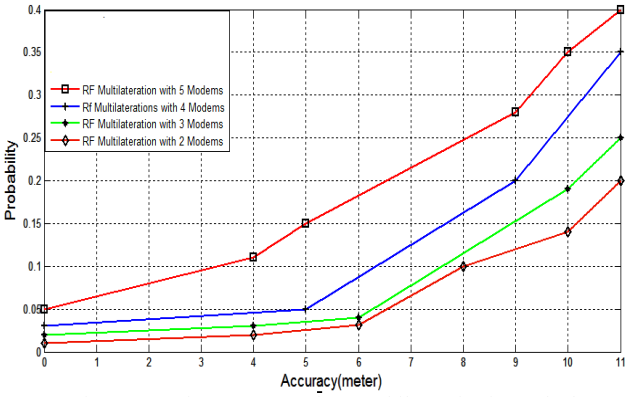



Figure 6: Performance test of RF Multilateration in test bed

methods. This time the test bed dimensions considered 5.2 meter by 10.4 meter in a classroom of Electrical Electronics Faculty at Yildiz Technical University. In this stage, better accuracy results have been obtained than previous test by using 5 modems. This time the interval of distance has taken smaller than previous test. Accuracy performance shows a better result with an accuracy of 0.5 meter with an accuracy of 25%. By decreasing the modem in the experiment it has been found that the performance is relatively lower in this method. Results have been shown in Figure 7.

This graph indicates that a better performance could be obtained in RF Multilateration by confining the test bed and step interval relatively small. Later to justify the performance in a nLoS system, same techniques has been performed with a nLoS environment considering different obstacles like walls, windows, human body reflection etc. At some specific or marked place which is especially behind the obstacles has been chosen and data has been measured. In this experiment most of the obstacles are considered as thick walls. Considering this condition, it has been noticed that the collected RSS value in a LoS medium and an nLoS medium varies significantly. Figure 8 shows Matlab graph of the RSS value in two different mediums.

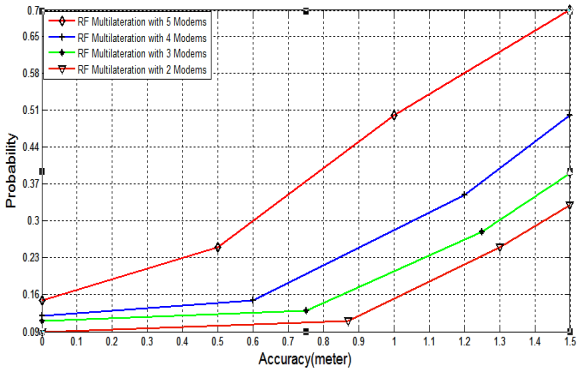



Figure 7: Performance test of RF Multilateration in test bed with a small area and small step interval

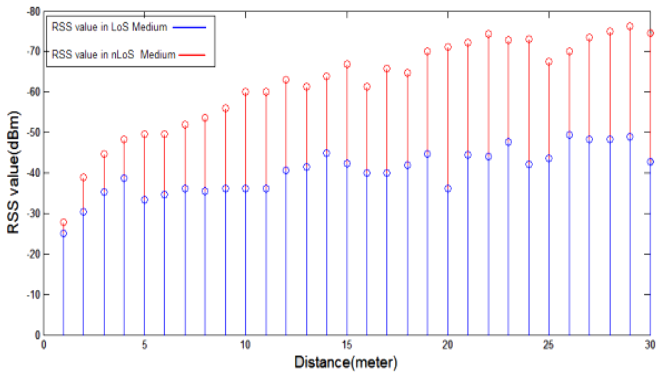



Figure 8: RSS value in Los and nLoS medium

This difference in RSS led us to model a new path loss exponent for an nLoS medium as well as to find a new value of *n* separately for every modem. Previously, in literature same path loss exponent has been used to find out accurate position using RF Multilateration method in an nLoS system. It has been found that to model path loss exponent for every modem is quite difficult work. To model path loss exponent, measurements has been taken for each modem separately in a LoS and an nLoS system. Later, data has been calculated and used in a formula for getting the path loss exponent. Formula for modeling path loss exponent has given below:

$$-P_r(d) = P_L(d_0) + 10n \text{ Log } (d/d_0) \quad (1)$$

where

$P_r(d)$: received signal strength at distance d

$P_L(d_0)$: offset loss at reference point

*n:* path loss exponential

$d_0$: distance from the transmitter to the receiver reference

d: distance between transmitter and receiver

In this stage, using this formula different path loss exponent for different modems has been found. As an example, to find out the path loss exponent using this formula, the value obtained as 2.10 which is shown in Figure 9.

By using the same techniques path loss exponent for other modes are obtained which is shown in Table-1.To justify the performance, at first path loss exponent is modeled in general for all modems in an nLoS medium. The path loss exponent *n* is considered in an nLoS medium with a value of 3.4 (*n*=3.4). Figure 10 shows Matlab graph of the obtained path loss exponent in an nLoS medium.

Table-1: Values of $n$ for different modems

| | Test-2 |
|---|---|
| Modem1 | $n_1$=2.10 |
| Modem2 | $n_2$=3.15 |
| Modem3 | $n_3$=2.75 |
| Modem4 | $n_4$=2.35 |

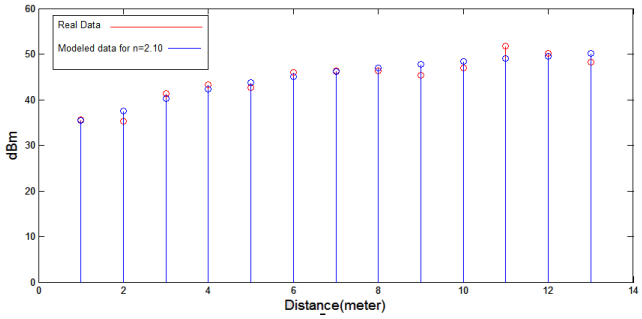



Figure 9: Path loss exponent modeling for Modem 1 in nLoS environment

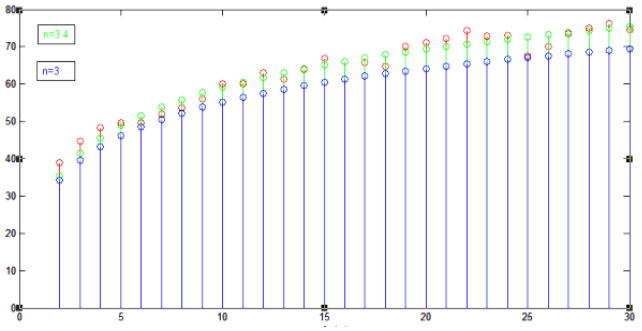



Figure 10: Path loss exponent modeling in nLoS environment (distance [m] vs. RSS[-dBm])

Later, performance test has been done by using this obtained value. It has been found that the obtained result was not satisfactory. By using same path loss exponent, an accuracy of 2 meter with a 3% probability has been obtained. On other hand, using different path loss exponent in RF Multilateration the performance improves significantly with an accuracy of 2 meter having a 25% probability rate. Results have been shown in Figure.11. Figure 11 indicates the improvement of performance in RF Multilateration by using different value of $n$ in an nLoS system. It should be mentioned that high error rate in RF Multilateration could be occurred because of multipath propagation, scattering, reflection and diffraction of substance. Moreover, it has been found that RF multilateration is suitable for small test bed area.

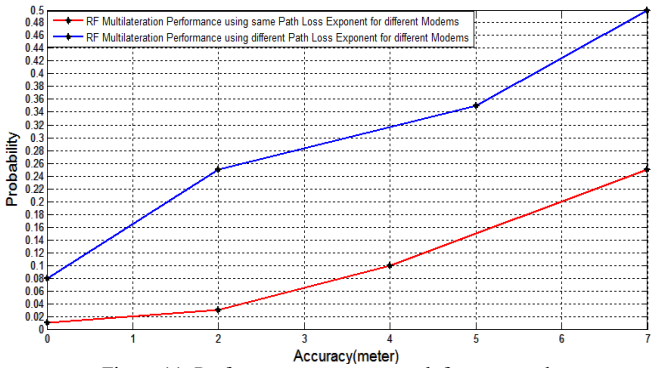



Figure 11: Performance accuracy graph for same and different path loss exponent in nLoS system

## IV. CONCLUDING REMARKS

RF fingerprinting and multilateration use received wireless signal strength to determine the exact location of a mobile user in an indoor environment in different ways.In this paper, two popular methods of indoor based localization systems have been studied and performance improvement aspects have been suggested.

To improve RF Fingerprinting methods performance

(i) Offline measurements intervals must be smaller.

(ii) The number of modem or APs should be increased in number.

(iii) The number of offline (training) data must be increased.

To improve RF multilateration methods performance

(i) For a LoS system test bed area should be smaller.

(ii) Number of modems must be increased in a LoS medium.

(iii) For an nLoS system, path loss exponent for each modem must be modeled.

(iv) Number of obstacles such as walls, windows etc. should be reduced in an nLoS system.

## REFERENCES


[1] J. Schiller and A.Voisard, “Location Based Services”, Morgan Kaufmann Puv,2004.

[2] NICULESCU, D., BADRI, N. Ad hoc positioning system (APS) using AOA. In *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2003.* San Francisco (USA), 2003, p. 1734 – 1743.

[3] Hasan Buyruk, A.Kenan Keskin,Şeyma Şendil,Hasari Çelebi,Hakan P.Partal,Salih Ergut,Engin Zeydan,and omer Ileri, “RF Fingerprinting Based GSM Indoor Localization”, Signal Processing and Communications Applications Conference,SIU,2013.

[4] A Ward, A Jones, A Hopper, “A New Location Technique for Active Office”, IEEE Personal Communications, Vol 4, No 5, Oct 1997.

[5] SAYED, A. H., TARIGHAT, A., KHAJEHNOURI, N. Network based wireless location: challenges faced in developing techniques for accurate wireless location information. *IEEE Signal Processing Magazine*, 2005, vol. 22, no. 4, p. 24 – 40.

[6] Ugur Alkasi, Md Al Shayokh , Hakan P. Partal, “An experimental comparison study on indoor localization: RF fingerprinting and multilateration methods,” *10th international Conference on electronics, Computer and Computation*, Ankara, Turkey, Nov. 7-9, 2013.

[7] Li,B.,Kam,J.,Lui,I.,& Dempstar,A.(2007). Use of Directional Information in Wireless LAN based indoor positioning. *Proceedings of IGNSS(International Global Navigation Satellite Systems Society)* Symposium.Taipei:IEEE.

# Devising a Solar Powered Standalone Vehicle using GSM Communication Network

A.S.M. Ashraf Ahmed, Labina Alamgir, Abu Nayeem, Devzani Sharma and Bishwajit Banik Pathik
Department of Electrical and Electronic Engineering, American International University-Bangladesh, Banani, Dhaka
E-mail: ashraf_ahmed786@yahoo.com

*Abstract*—**This paper presents the technical construction of a standalone vehicle controlled by GSM communication network. The designed GSM based solar powered vehicle could be operated from almost anywhere under GSM network which is powered by solar energy using 5 watt photo voltaic (PV) panel, stored in 3 similar 4V rechargeable batteries. The operation commences with a call generated from a cell phone which is auto received by another phone stalked in the vehicle motor driver. In the course of a call, if any of the buttons, 2, 4, 6 or 8, is pressed a tone corresponding to the button pressed is heard at the other end of the transmission which is called Dual Tone Multiple Frequency (DTMF) tone. The received tone in the cell phone at vehicle end is processed by a set of relays. These relayed signals are sent to the motor driver IC (L293D) which drives the motor forward, reverse, right or left. Most importantly as the car will be running by solar energy, so the vehicle can be sent to a long distance not worrying about the charge of the battery, since it accumulates the greater portion of the energy required from the external PV panel that absorbs and converts sunlight to generate the driving power, though there will be DC battery as a backup.**

***Keywords—Solar vehicle, remote controlled transport, solar robot, GSM based remote.***

## I. INTRODUCTION

A remote control vehicle is typically defined as any mobile device that is controlled by a means that does not restrict its motion with an origin external to the device. This is often a radio control device, cable between control and vehicle, or an infrared controller. A remote control vehicle (RCV) differs from a robot in that the RCV is always controlled by a human and takes no positive action autonomously [2]. One of the key technologies which underpin this field is that of remote vehicle control. It is vital that a vehicle should be capable of proceeding accurately to a target area; maneuvering within that area to fulfill its mission and returning equally accurately and safely to base [1].

The first general use of radio control systems in models started in the late 1940s with single channel self-built equipment; commercial equipment came soon thereafter. Initially remote control systems used escapement, (often rubber driven) mechanical actuation in the model [23]. Commercial sets often used ground standing transmitters, long whip antennas with separate ground poles and single vacuum tube receivers [5]. The first kits had dual tubes for more selectivity. Such early systems were invariably super regenerative circuits, which meant that two controllers used in close proximity would interfere with one another [6].

## II. DESIGN AND CONSTRUCTION

In this project the vehicle is attached with a mobile phone under GSM communication network which is controlled by a user mobile phone. With the help of user mobile phone we can move the vehicle in desired direction as per our requirement. This project is constructed from a very compact dual tone multi-frequency (DTMF) based decoder, and the GSM network controlled vehicle organizes the switching from the decoded and power switching device for controlling the motor drive of the vehicle using two cell phones.

We know RC (Remote Controlled) cars or vehicle do not have a high range of wireless network. This means that the operator has to be in touching distance to the receiver of the vehicle. Thus it is clear that a remote controlled vehicle cannot be applied for an array of duty due to its lacking of controlling range. This is where GSM controlled vehicle steps in. Using two GSM able phones we can create a controlling mechanism for the vehicle. Here we do not have to worry about the range for operation, if sensors such as IR sensors and camera or 3G enabled mobile phones are used, as most of the world is under the assortment of GSM network [9]. By using this prospect we can take this vehicle and turn it for human benefits. These vehicles can be used as firefighting robots, battle vehicles or applied in vast places where it's not possible or dangerous for any human being to go.

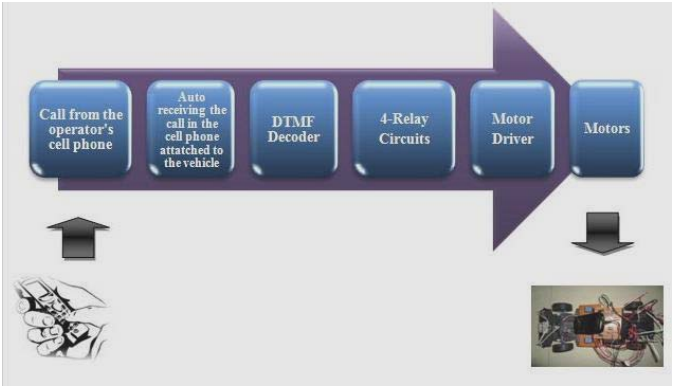



Figure 1. Block diagram of GSM network controlled vehicle.

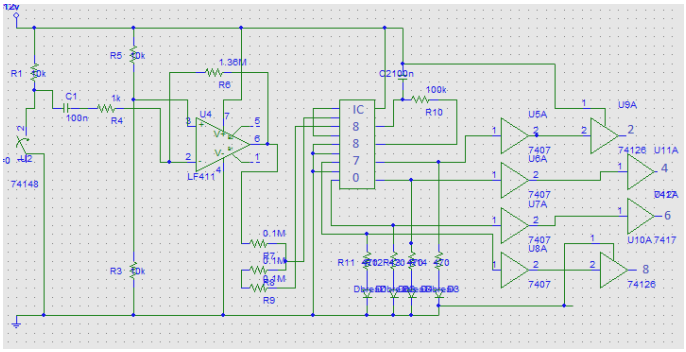


Figure 2. Circuit diagram of the DTMF operation

The solar powered standalone vehicle was controlled by a mobile phone that made calls to the mobile phone attached to the vehicle. In the course of the call if any button was pressed, pulse sound corresponding to the pressed button was heard at the other end of the call. This tone is called dual tone multi frequency (DTMF) [3]. The vehicle received this DTMF tone with the help of phone stacked in the vehicle. The connection between the cellphone and the decoder is made with the help of a universal 3.5mm audio jack. The received tone was processed by the relays Q1, Q2, Q3, and Q4. The relays are wired such that for a particular pulse from the DTMF voltage will pass through only one relay and the other three relays are closed [4].

TABLE I. LOGIC TABLE FOR TURN ON AND TURN OFF FOR THE SET OF RELAYS [26]

| When no buttons is pressed | Button '1' is pressed and held | Release button '1' | Button '2' is pressed and held | Release button '2' | Button '0' is pressed and held |
|---|---|---|---|---|---|
| Logic 0 | Logic 1 | Logic 0 | Logic 1 | Logic 0 | Logic 1 |
| $Q_4$: 0 | $Q_4$: 0 | $Q_4$: 0 | $Q_4$: 0 | $Q_4$: 0 | $Q_4$: 1 |
| $Q_3$: 0 | $Q_3$: 0 | $Q_3$: 0 | $Q_3$: 0 | $Q_3$: 0 | $Q_3$: 0 |
| $Q_2$: 0 | $Q_2$: 0 | $Q_2$: 0 | $Q_2$: 1 | $Q_2$: 1 | $Q_2$: 1 |
| $Q_1$: 0 | $Q_1$: 1 | $Q_1$: 1 | $Q_1$: 0 | $Q_1$: 0 | $Q_1$: 0 |

Thus it's possible for the motor drives to drive the motors for forward or backward motion or make a turn. The mobile that makes a call to the mobile phone stacked in the vehicle acts as a remote [8]. The DTMF decoder and the switching circuit is designed to permit a digital signal processing device control high power external loads by issuing commands encoded as audio DTMF signals. The relays direct the overall operation of the DTMF decoder to perform the actual DTMF audio tone pair decoding. When a valid tone pair is detected by the DTMF decoder, an interrupt is signaled the tone pair code from the decoder and places the symbol in an internal quell for further processing [5]. DTMF signaling is used for telephone signaling over the line in the voice-frequency band to the call switching center [5]. The version of DTMF used for telephone dialing is known as touch-tone. DTMF assigns a frequency (consisting of two separate tones) to each key so that it can easily be identified by the electronic circuit. The signal generated by the DTMF encoder is the direct algebraic summation, in real time, of the amplitudes of two sine (cosine) waves of different frequencies, i.e., pressing '5' will send a tone made by adding 1336 Hz and 770 Hz to the other end of the mobile [1]. The tones and assignments in a DTMF system are shown in Table II.

TABLE II. DTMF DATA OUTPUT [24]

| Digit | Low Frequency (in Hz) | High Frequency (in Hz) | $D_4$ | $D_3$ | $D_2$ | $D_1$ | $D_0$ |
|---|---|---|---|---|---|---|---|
| 1 | 697 | 1209 | H | L | L | L | H |
| 2 | 697 | 1336 | H | L | L | H | L |
| 3 | 697 | 1477 | H | L | L | H | H |
| 4 | 770 | 1209 | H | L | H | L | L |
| 5 | 770 | 1336 | H | L | H | L | H |
| 6 | 770 | 1477 | H | L | H | H | L |
| 7 | 852 | 1209 | H | L | H | H | H |
| 8 | 852 | 1336 | H | H | L | L | L |
| 9 | 852 | 1477 | H | H | L | L | H |
| 0 | 941 | 1209 | H | H | L | H | L |
| * | 941 | 1366 | H | H | L | H | H |
| # | 941 | 1477 | H | H | H | L | L |
| A | 697 | 1633 | H | H | H | L | H |
| B | 770 | 1633 | H | H | H | H | L |
| C | 852 | 1633 | H | H | H | H | H |
| D | 941 | 1633 | H | L | L | L | L |
| Any | … | … | L | 0 | 0 | 0 | 0 |

L: Low (Logic 0)
H: High (Logic 1)

Here, the relays are switches that open and close circuits electronically. Relays control one electrical circuit by opening and closing contacts in another circuit. When a relay contact is normally open (NO), there is an open contact when the relay is not energized. When a relay contact is Normally Closed (NC), there is a closed contact when the relay is not energized.

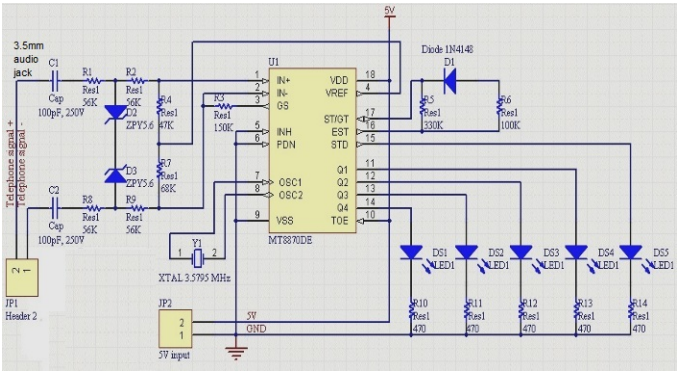


Figure 3. Connection diagram of DTMF circuit

In either case, applying electrical current to the contacts will change their state.

Relays are generally used to switch smaller currents in a control circuit and do not usually control power consuming devices except for small motors and Solenoids that draw low amps. Nonetheless, relays can "control" larger voltages and amperes by having an amplifying effect because a small voltage applied to a relays coil can result in a large voltage being switched by the contacts [21].

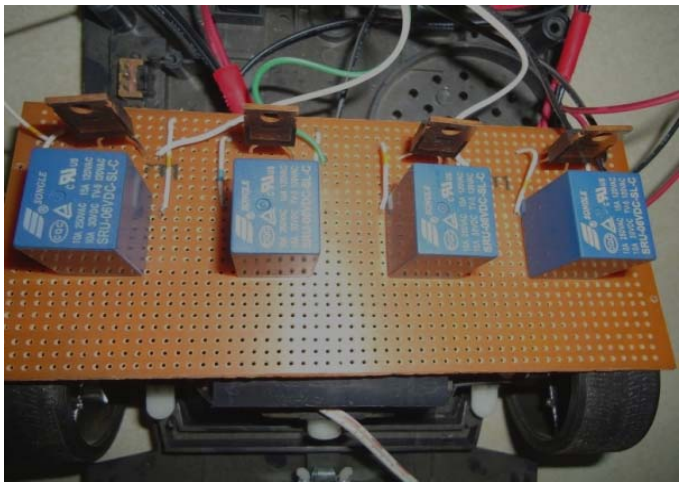


Figure 4. Relays implemented on vero board

To charge the cells that used in the whole project an external solar charged controlled was used. The cell that was used was a mono 5watt power cell with rated voltage of 17V. A charge controller, or charge regulator is basically a voltage and/or current regulator to keep batteries from overcharging [22]. It regulates the voltage and current coming from the solar panels going to the battery. Most "12 volt" panels put out about 16 to 20 volts [13]. Most batteries need around 14 to 14.5 volts to get fully charged [12].

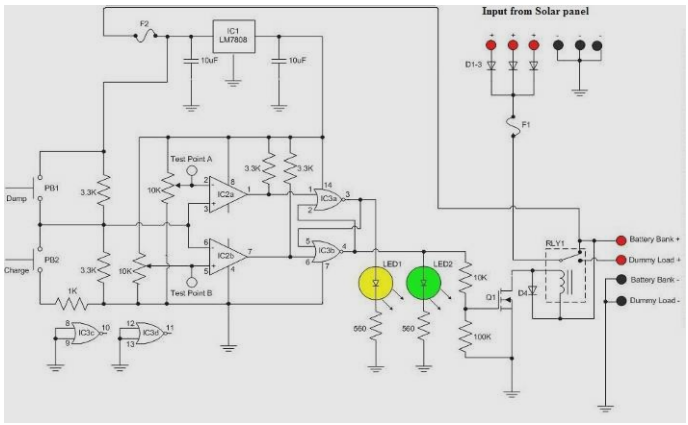



Figure 5. Solar charge controller circuit diagram

A photovoltaic cell consists of a light absorbing material which is connected to an external circuit in an asymmetric manner. Charge carriers are generated in the material by the absorption of photons of light, and are driven towards one or other of the contacts by the built-in spatial asymmetry. This light driven charge separation establishes a photo-voltage at open circuit, and generates a photocurrent at short circuit. When a load is connected to the external circuit, the cell produces both current and voltage and can do any electrical work [14].

TABLE III. TABLE FOR SPECIFICATIONS OF THE SOLAR CELL [27]

| **Module Type- Mono 5W** | |
|---|---|
| Dimensions (mm) | 244 X 288 X 17 |
| Maximum Power ($P_{max}$) | 5W |
| Tolerance of ($P_{max}$) | 0±3% |
| Rated Voltage ($U_{mpp}$) | 17V |
| Rated Current ($I_{mpp}$) | 0.29A |
| Open Circuit Voltage ($U_oC$) | 21.6V |
| Short Circuit Voltage ($V_{sc}$) | 0.34V |
| Maximum System Voltage | 600V |
| Weight (Kg) | 0.75 |

After numerous testing and running of the batteries a table was constructed of the time required for charging and discharging of the batteries using the solar charged panel. The three separate 4V batteries were connected to our solar charge controller circuit and were left to charge on the roof of a high rise building during day time. This was done for few days and the corresponding charging time for the battery to get fully charged was recorded at different time of the day. To record the discharge reading of the batteries when all our components were connected the system was turned on and was used until the batteries ran out of full charge. The time required for the batteries to lose full charge was also recorded. All these recorded values were placed in Table IV.

TABLE IV. RECORDED CHARGING AND DISCHARGING DURATIONS

| **Test Run** | **Charging (3 separate 4V batteries)** | **Discharging Running Condition** | **Time of Day** |
|---|---|---|---|
| 1 | 130 Min | 120 Min | 9:00 |
| 2 | 125 Min | 125 Min | 11:00 |
| 3 | 115 Min | 122 Min | 14:00 |
| 4 | 170 Min | 120 Min | 16:30 |
| 5 | 210 Min | 122 Min | 18:00 |

Outputs from the relays are fed to inputs of motor drivers respectively to drive two geared DC motors. The relay sends voltage to drive the DC motors. Drivers are required for motor rotation. The L293D is a quad, high-current, half-H driver designed to provide bidirectional drive currents of up to 600mA at voltages from 4.5 V to 36V [20].

TABLE V. OPEARATION FROM DTMF DECODER TO RELAY [25]

| Button pressed | Output of decoder | Input to Relay | Output from Relay | Action Performed |
|---|---|---|---|---|
| 2 | 0X02<br>00000010 | 0XFD<br>11111101 | 0X09<br>00001001 | Forward Motion |
| 4 | 0X04<br>0000100 | 0XFB<br>11111011 | 0X05<br>00000101 | Left Turn<br>Right Motor Forward<br>Left Motor Back Warded |
| 6 | 0X06<br>00000110 | 0XF9<br>11111001 | 0X0A<br>00001010 | Right Turn<br>Right Motor Back Warded<br>Left Motor Forwarded |
| 8 | 0X08<br>00001000 | 0XF7<br>11110111 | 0X06<br>00000110 | Backward motion |

The driver makes it easier to drive the DC motors. The L293D consists of four drivers. Pins IN1 through IN4 and OUT1 through OUT4 are input and output pins, respectively, of Driver 1(D1) through Driver 4(D4).drivers 1 and 2, and drivers 3 and 4are enabled by enable pin 1 (EN1) and pin9 (EN2), respectively. When enable input EN1 (pin1) is high, drivers1 and 2 are enabled and the outputs corresponding to their inputs are active. Similarly, enable input EN2 (pin9) enables drivers 3 and 4.

## III. METHOD OF STUDY

The information about designing the circuits were collected from many sources i.e. books, papers, websites etc and was studied well to get idea [18]. We have studied on different kind of remote controlled vehicle and also the use of GSM communication network. Photovoltaic panel was implemented to recharge the rechargeable battery. And also a solar charge controller is designed and implemented to regulate the power flowing from a photovoltaic panel into a rechargeable battery. The solar charge controller features easy setup with one potentiometer for the float voltage adjustment and an equalize function for periodic overcharging.

The steps involved to finish this project are listed below:

Step 1 Collecting the information about the topic from many sources like books, papers, websites, etc

Step 2 Choosing equipments available for the study and developing idea about cell phones to act as a remote controlling device

Step 3 Designing the circuit using circuit stimulating software like PSpice

Step 4 Implementing the circuit in breadboard, and then in vero board to minimize the size of the circuit

Step 5 Observing the output signal by varying the input signal and recording the charging and discharging time

Step 6 Reducing the error

The performance of the vehicle was observed carefully. Thus the study of the project was successfully completed overcoming the limitations to some extent.

## IV. COMPLETED SYSTEM & SUGGESTED MODIFICATION

Previously some researches involved vehicles which could be remotely controlled, but it should have been kept under a range where it could be supervised. Here as we are willing to make it cell phone based remote control vehicle it can be operated almost everywhere if GSM network exists. Moreover it can charge its battery by its own by the use of solar panels so it is itself a standalone system. Since the car will be running

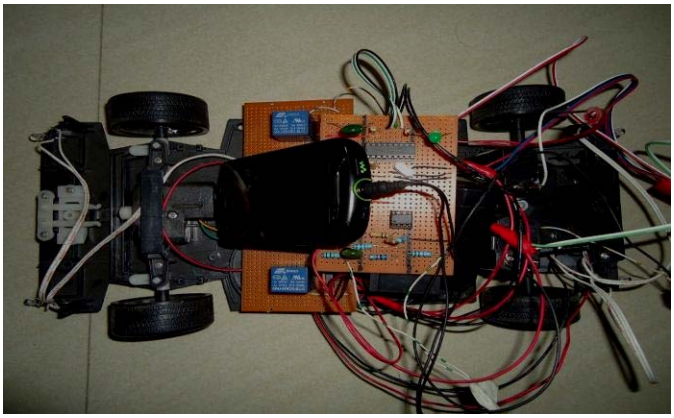


Figure 6. Implemented GSM controlled vehicle

by solar energy, the vehicle can be sent to a long distance not worrying about the charge of the battery, since it will try to gather most of the energy by solar power, though there will be a DC battery as backup [17]. As it will be dependent on solar energy it is quite obvious that it is an eco-friendly project. In addition the charge controller will make the project even more efficient.

On the other hand, for the operation of this system always two cell phones are required so every time the remote has to dependent on another cell phone that has to connected and stacked to the chassis of the vehicle and by default it was assumed that the cell phone's number is a secured one, which is only known by the system.

A prototype of the remote system controlling has been implemented in this project. Although the implementation is just a simple automatic vehicle utilizing renewable energy as its power source, it may be a pathway for more such researches.

Evaluating this project and thesis paper, it is clearly noticeable that this project has opened the window for enormous future researches in this field for the next researchers.

*A) Substituting the 2G GSM cell phone with 3G handset*

3rd Generation or 3G is the generic term used for the next generation of mobile communications systems. 3G technology is commonly used in smart phones, where a strong emphasis is put on internet and multimedia services while its predecessor, second generation or 2G technology emphases mostly on voice applications like talking, call waiting, etc. 3G technology has two major advantages over 2G which enables always connectivity to internet [16]. Hence finally it can be stated/covered that replacing the existing handset with a 3G one will not only extends operation of the developed circuit but will also enable some more additional features to be employed alongside the present one .

*B) Modification in the System Design*

This project can also be made perfect by means of conducting three simple modifications in this existing circuit:

i. *Replacing the DTMF Decoder with DTMF Transceiver:* Future researchers can implement this assignment by substituting the DTMF Decoder IC 8870 by a DTMF Transceiver IC 8880, allowing the system to generate a DTMF tone by itself [7]. If an additional alarm circuit along with sensors is implemented along with the existing one, the system will then be able to notify the user when an alarm initiates via calling a fixed number.

ii. *Password Protection:* In order to prevent unauthorized access of this robot, the project can be employed by means of password protection. In case of one interested in implementing the present circuit, the cell phone connected should be password protected [10].

iii. *Deploying a Camera:* Installing a camera with the current system will enable the vehicle to operate in difficult territories which are either out of range of human reach or are hazardous for human life [11].

*C) Implementation of PWM Charge Controller*

Future researchers can work out to determine the possibility of implementing a Pulse Width Modulation (PWM) based controller circuit for managing the charging and discharging of the battery not only using photovoltaic energy but also wind energy [15].

*D) Modifications in the System Operation*

The operation of the system can easily be modified and this vehicle can be used for variety of purposes. Conducting some adjustments in the system design and body will enable this vehicle to serve as remote control robot performing other wide range of operations. Some of such operations are highlighted below:

i. Remote Control Racing Vehicle: Robotic Race Cars can easily be designed via some simple modifications in this design. The cars will travel in a pre-designed track and users will be in command of the navigation of the car.

ii. Bomb Detector Vehicle: Future researchers can easily modify this vehicle and use it as bomb detector robot. It can be done efficiently by implementing a program to trace the exact positions of bombs on a pre-designed map. The robot will be designed to sense bombs in a remotely controlled way, and user will navigate the robot and locate the landmines and update information into the program's database.

iii. Remote Control Fire Fighter Device: This project can be modified easily and be implemented as a prototype model for fire fighter robot. The microcontroller operated robot will move through a structure, detect fire and then extinguish it with the help of blower.

*E) Replacing the Relay Logic Circuit with a Micro-Controller*

This circuit is constructed using a set of relays whose function solely is to transfer the DTMF tone from the cell phone to the motor driver IC, L293D. Upcoming researchers can execute this project replacing the relays with a microcontroller which is a small integrated circuit with a microprocessor, memory and programmable input / output support [19].

## V. CONCLUSION

The key purpose was to develop a circuit that can drive an electric vehicle in any directions using GSM based cell phones as a distant controller, and the trial approached has been a success. This system utilizes a renewable energy based battery management system and a GSM technologically operated mobile phone for its operations. The second part of this project highlights on deploying a battery management system using renewable photovoltaic energy as its power source from which the system can charge its batteries using solar panels as a standalone system. This system can be a test-bed for any future projects and or appliances interested to work with both renewable energy and remote control communication technology together.

## REFERENCES


[1] T. M. Ladwa, S. M. Ladwa, R. S. Kaarthik, A. R. Dhara and N. Dalei, "Control of Remote domestic System using DTMF", International *Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), Bandung, Indonesia, 2009*, pp. 1-6.

[2] D. Heß, C. Röhrig. "Remote Controlling Technical Systems Using Mobile Devices", *IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, Rende (Cosenza), Italy 21-23 September 2009*, pp.625-628.

[3] M. Callahan Jr, "Integrated DTMF Receiver," *IEEE Transactions on communications, vol. 27, pp. 343-348, February, 1979.*

[4] Y. C. Cho and J. W. Jeon, "Remote Robot control System based on DTMF of Mobile Phone" *IEEE International Conference INDIN 2008,* July 2008.

[5] R. Sharma, K. Kumar, and S. Viq, "DTMF Based Remote Control System," *IEEE International Conference ICIT 2006,* pp. 2380-2383, December 2006.

[6] S. Selman, R. Paramesran, "Comparative Analysis of Methods Used in the Design of DTMF Tone Detectors" *IEEE International Conference on Telecommunications and Malaysia International Conference on Communications,* 14-17 May 2007, Penang, Malaysia.

[7] R. C. Luo, T. M. Chen, and C. C. Yih, "Intelligent autonomous mobile robot control through the Internet," *IEEE International Symposium ISIE 2000,* vol. 1, pp. 611, December, 2000.

[8] C. K. Das, M. Sanaullah, H. M. G. Sarower and M. M. Hassan, "Development of a cell phone based remote control system: an

effective switching system for controlling home and office appliances", *International Journal of Electrical & Computer Sciences IJECS,* vol. 9 No: 10 pp3743, 2010

[9] V. Hachenburg, B. Holm and J. Smith, "Data signaling functions for a cellular mobile telephone system", *IEEE Trans Vehicular Technology*, vol. 26, #1 p. 82, 1977

[10] Lagura, J. L. R., Pascual, M. C. G., Rabadan, R. J. R., & Tabac, A. G. (2004) "Microcontroller- based aid for the blind in reading mobile phone messages using braille system", *Undergraduate Thesis. De La Salle University Manila.*

[11] H. Haldun GÖKTAS, Nihat, "A Cellular Phone Based Home / Office Controller &Alarm System" *DALDAL Gazi University Technical Education Faculty, 06500, Besevler, Ankara, TURKEY*

[12] Nelson, Jenny, "The Physics of Solar Cells", London: Imperial College, 2003. Print

[13] N. C. Bhowmik, Bangladesh Renewable Energy Report, *APCTT-UNESCAP*, 2012

[14] R. Subarkah, Belyamin, "Improved Efficiency Solar Cells with Forced Cooling Water", retrieved on 16/10/2011.

[15] A. Z. Alkar, U. Buhur, "An Internet Based Wireless Home Automation System for Multifunctional Devices," *IEEE Consumer Electronics*, vol. 51, no. 4, pp. 1169-1174, 2005.

[16] R. Shahriyar, E. Hoque, S. M. Sohan, I. Naim, M. M. Akbar, and M. K. Khan, "Remote Controlling of Home Appliances using Mobile Telephony", *International Journal of Smart Home*, Vol. 2, No. 3, July, 2008.

[17] K. Reddy, F. Althobeti, Dr Md A. Hussain and R. Reddy, "GSM-Controllable Power Switch System for Industrial Power Management", *International Journal of Engineering Trends and Technology*, August 2011

[18] T. Kubik and M. Sugisaka, "Use of a Cellular Phone in Mobile Robot Voice Control", Proceedings of the *40th SICE Annual Conference*. International Session Papers, Nagoya, 2001, pp.106-111.

[19] http://www.newworldencyclopedia.org/entry/Remote_control [May 5, 2013]

[20] http://www.extremeelectronics.co.in/avrtutorials/pdf/avr_tutorial7---motor-control.pdf [May 15, 2012]

[21] http://electronics.howstuffworks.com/relay1.htm [Sep 5, 2012]

[22] http://www.solar-electric.com/solar-charge-controller-basics.html [Sep 15, 2012]

[23] http://www.studymode.com/subjects/historical-background-of-the-ntc-radio-control-office-page2.html [July 11, 2012]

[24] http://www.datasheetarchive.com/MT8870-datasheet.html [May 7, 2012]

[25] http://www.dattalo.com/technical/theory/dtmf.html [June 19, 2012]

[26] http://www.omega.com/manuals/manualpdf/M1906.pdf [May 19, 2012]

[27] http://www1.eere.energy.gov/buildings/residential/pdfs/rerh_pv_guide.pdf [Sep4, 2012]

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Design and Implementation of Fast FPGA Based Architecture for Reversible Watermarking

Sudip Ghosh[1*], Bijoy Kundu[2], Debopam Datta[3], Santi P Maity,[4] and Hafizur Rahaman[1,4]
[1] School of VLSI Technology (Bengal Engineering and Science University at Shibpur, India)
[2] Dept. of Electronics and Telecommunication (Bengal Engineering and Science University at Shibpur, India)
[3] Dept. of Electrical and Computer Engineering (University of Illinois at Chicago, USA)
[4] Dept. of Information Technology (Bengal Engineering and Science University at Shibpur, India)
[*] E-mail: sudip_etc@yahoo.co.in

*Abstract*—**There are diverse hardware realization for digital watermarking of multimedia proposed in the literature. This paper focuses on the design and implementation of a fast FPGA(Field Programmable Gate Array) based architecture using reversible contrast mapping (RCM) based image watermarking algorithm. The specialty of this architecture attracts to the fact of clock-less encoder design and implementation which makes the design faster. The encoder module response time is independent of clock frequency, so the embedding of the watermark is possible as soon as the input is fetched. The schematic based design and implementation of the VLSI architecture have been done with Xilinx 14.1 on Spartan 3E FPGA family. The encoder requires 528 4-input LUTs and 303 slices. On the contrary, the decoder requires 613 LUTs and 347 slices. The maximum clock frequency of the decoder is 45 MHz. The results show the viability of low cost, high speed real-time use of the proposed VLSI architecture.**

***Keywords- VLSI Architecture,Reversible Watermarking, FPGA .***

## I. INTRODUCTION

Digital watermarking [1] is an efficient tool to prevent unauthenticated use of data. Digital watermarks may be used to verify the authenticity or integrity of the original data. Nowadays, it is prominently used for tracing copyright infringements and for banknote authentication. Digital watermarking is broadly classified depending on the type of signal like audio watermarking, image watermarking, video watermarking, and database watermarking etc. The present work is focused on image watermarking.

In *image watermarking*, the digital information (like a digital image, a digital signature or a random sequence of binary numbers) is embedded into an image. The embedded information may or may not be perceptible after watermarking and therefore falls into the category of *visible or invisible watermarking* respectively. Depending on the robustness of the watermark, it can also be categorized as *robust or fragile watermarking*[2].

One limitation of watermarking-based authentication schemes is the distortion inflicted on the host media by the embedding process. Although the distortion is often insignificant, it may not be acceptable for some applications, especially in the areas of medical imaging and military applications. Therefore, watermarking scheme capable of removing the distortion and recovering the original media after passing the authentication is desirable. Schemes with this capability are often referred to as reversible watermarking schemes [3].Various Reversible Watermarking techniques have been proposed with different type of algorithm [4]-[5]. Popular techniques of reversible watermarking are: i) Difference Expansion, ii) Histogram bin Shifting, iii) Data hiding using Integer Wavelet Transform, iv) Contrast Mapping, and v) Integer Discrete Cosine Transform.Usually, a reversible scheme performs some type of lossless compression operation on the host media in order to make space for hiding the compressed data and the Message Authentication Code (MAC) (e.g., hash, signature, or some other feature derived from the media) used as the watermark [6]. To authenticate the received media, the hidden information is extracted and the compressed data is decompressed to reveal the possible original media. MAC is then derived from the possible original media. If the newly derived MAC matches the extracted one, the possible original media is deemed authentic/original.

However, in this paper a reversible watermarking technique is implemented using a specific transform reported by Coltuc *et. al.* in [5]. The choice of this technique includes its low computational complexity and robustness. The primary goal of the proposed design is to achieve high speed hardware efficient VLSI architecture. The RCM technique was first implemented in Matlab to verify the algorithm and analyze various design constraints. Later the desired architecture is established in FPGA using Xilinx. The paper is organized by starting with an abstract followed by section I with an introduction. Section II describes the related works. Next section III reports the proposed VLSI architecture of reversible watermarking followed by the analysis and experimental results in section IV, finally the work is concluded in section V with references.

## II. RELATED WORKS

In the scheme proposed by Fridrich *et al.* [1], Discrete Cosine Transform (DCT) technique has been implemented. 128-bit hash of all the DCT coefficients is used as the watermark. The extracted compressed bit-stream is used for verification;however, the hash contains only the signature of the image, with no local information.Therefore, despite its simplicity and ability to detect inauthenticity, this technique is unable to locate the position where the tampering has been done. Van Leest *et.al.*[3] proposed another reversible watermarking scheme based on a transformation function that introduces “gaps” in the image histogram of image blocks.

One drawback of this scheme is its need for the overhead information and the protocol to be hidden in the image. Moreover, a potential security loophole in the scheme is that given the fact that the computational cost for extracting the watermark is insignificant; an attacker can defeat the scheme by exhausting all the 256 possible gray level assuming that the gray level being tried is the gap. In [5], Coltuc *et al*. proposed a Reversible Contrast Mapping (RCM) based algorithm of reversible watermarking in the spatial domain. It provides a high data embedding bit-rate at a very low mathematical complexity. The proposed scheme does not need any additional data compression but is able to recover the original image even after alterations in the encoded data.

Over the last decade, a lot of research is performed on Reversible Watermarking, however, VLSI implementation of RCM based approach is still an area to be explored. In this paper, the advantages of RCM based watermarking technique have been explored and implemented. Major concentration is given on developing a low cost, high speed VLSI architecture that can be used for real-time applications. Some significant hardware implementations of digital watermarking include the work [8], [9], [10], [11]. Mohanty *et al.* [9]concentrated on a spatial-domain invisible-fragile watermarking and their architecture. But these designs are seriously constrained due to their hardware complexities. In [12], a hardware architecture that can insert two visible watermarks in images in the spatial domain is introduced. The main objective of the proposed architecture was to decrease the hardware complexity keeping the performance intact. Employing the advantages of RCM technique, a low cost hardware efficient VLSI implementation of RCM based RW has been presented in this paper.

## III. Proposed VLSI Architecture of Reversible Watermarking

The implementation of the watermarking algorithm is done using the ISE Design Suite of Xilinx for Spartan 3E FPGA family. FPGA, because of its advantages like re-configurability, low cost and simpler design process, is used for the hardware implementation. The entire watermarking architecture design involved construction of two main blocks, the encoder and the decoder. Each of the blocks is further divided into three sub-blocks named as module 1, module 2, and module 3. Each of these modules is designed individually through modularization and later interfaced with each other. The encoder and decoder were designed and simulated separately.
Both the encoder and decoder designs are described in detail with their respective modules in the following subsections 1 and 2.

### 1. ENCODER:

In the proposed architecture, the encoder part is designed in three stages as given in Fig. 1.

*Image Acquisition and Pixel Transform:*

The proposed architecture is implemented and optimized for 8 bit gray image. The source to the encoder, which is basically a device providing image pixel as input, can be a storage device like a RAM or direct external input by the user in 8 bit

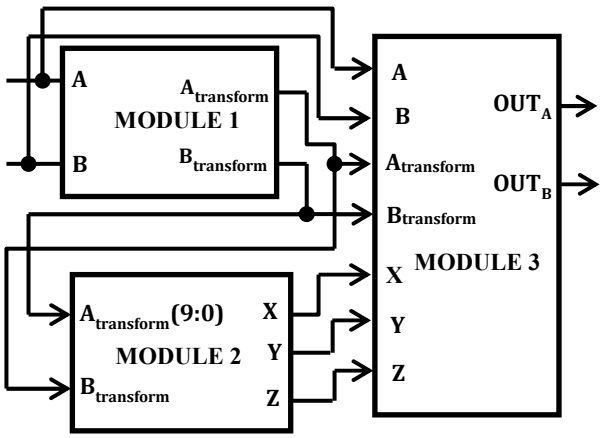



Fig. 1. Data flow path in encoder

digital form. In the proposed architecture, original image data is stored in a 256 byte RAM (eight 32-word by 8-bit SRAM). As discussed in [5], a specific transformation technique is performed on the image involving a pair of pixels. Among various ways of acquiring these pixels from the source, sequential column wise fetching (each element of a particular row and column is an 8-bit pixel value represented by an 8-bit address) from the memory is carried out in this design. The 8-bit pixel value read from the memory is then converted to a 10-bit data (adding zero at the 9th and 10th bit position) to provide the correct form of input for pixel transformation. The transformation technique [5] is mathematically given by,

$$A_{transform} = 2 * A - B \quad (1)$$
$$B_{transform} = 2 * B - A \quad (2)$$

Where A, B are the pair of input pixels of the original image and $A_{transform}$, $B_{transform}$ are the pair of transformed pixels. This transformation technique allows error free transmission and

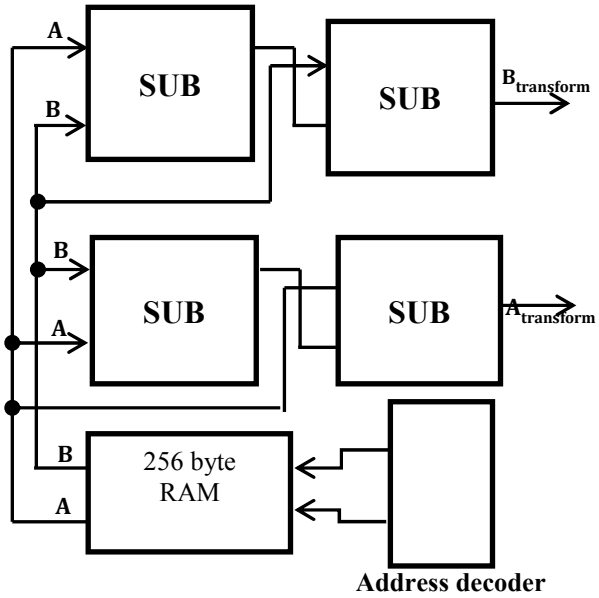



Fig. 2. Data flow path in image acquisition and pixel transform module

detection of image both at the transmitter and receiver end respectively [5].

Fig. 2 shows the data flow path in image acquisition and pixel transform module gives the pixel transform module implemented by using only two subtractor modules for each pixel. 10 bit subtractor ensured signed subtraction using twos complement logic and also prevented overflow.

*Control Signal:*

As mentioned earlier, the watermarking (embedding watermark image data into original image) algorithm is performed on image pixels constrained to a particular domain, $D_c$, of the transformed pairs [5]. Domain $D_c$ of transformed pixels of the original image is defined such that the pair of transformed pixels, $A_{transform}$& $B_{transform}$, belong to [0, L] where L takes values from 0 to 254 leaving 1. The domain $D_c$ prevents underflow and overflow as well as removes ambiguous pairs. It also ensures robust error free transmission of the watermarked image. This module generates control signals that are essential to carry out data embedding process which include determining $D_c$ along with other essential control signals. As mentioned by Coltuc *et. al.* in [5], three distinct groups are made partially depending on $D_c$ which are determined distinctly by three control signals (X, Y, and Z) in Fig. 3.

The generation of these control signals is briefed below.

- A low logic level, '0', of X is generated when pair $A_{transform}$ and $B_{transform}$ belongs to $D_c$ and each of them is even.
- A high logic level, '1', of Y is generated when pair $A_{transform}$ and $B_{transform}$ belongs to $D_c$ and is odd.
- A high logic level, '1', of Z is generated when the pair does not belong to $D_c$.

These control signals are generated completely using logical gates as shown on Fig. 3. The $10^{th}$ bit determines the polarity (either positive or negative) of the transformed pair while the $9^{th}$ bit determines if the transformed pair is below 255. The LSB determines whether it is even or odd.

*Data Embedding:*

The circuit generating performing the task of watermark image embedding in the original image is given in Fig. 4. This module uses the control signals as input to selectively perform watermarking depending on the control signals. The control signals determine whether the pixels are to be transformed before embedding the watermark sequence into the original image. The LSB of the $2^{nd}$ pixel among the pair is used to embed the watermark image while the transformation information (whether pixels are transformed or not) is embedded into the LSB of the $1^{st}$ pixel. The watermarking algorithm in terms of the control signals is given in pseudo code as follows.

Watermarking based on control signals:

- When X='0', pass pair $A_{transform}$ and $B_{transform}$ for watermarking. Set LSB of $A_{transform}$ to '1' and LSB of $B_{transform}$ replaced by watermark image.
- When Y='1', pass pair A and B for watermarking. Set LSB of A to '0' and LSB of B replaced by watermark image.
- When Z='1', watermarking step is skipped. Set LSB of A to '0' and the original image pixels are transmitted.

2. DECODER:

The decoder block is structured similarly like the encoder block. It is comprised of the three modules, the signal generation block, the inverse transform block, and the image and watermark extraction block. The entire decoder architecture is given in Fig. 5. Following sections from (i)-(iii) give a detailed hardware description of the individual modules of the decoder.

*Control Signal:*

Similar to the encoder part, the control signals are generated from the 8-bit input data received from the transmitter i.e. the encoder. The watermark image data as well as the transformation information has been embedded into the LSB of the transmitted pairs. Therefore, the preliminary task of this module is to extract LSB of both the received pairs. The LSB of the $WI_A$ received signal contained the transformation information and $WI_B$ contained the watermark data. The

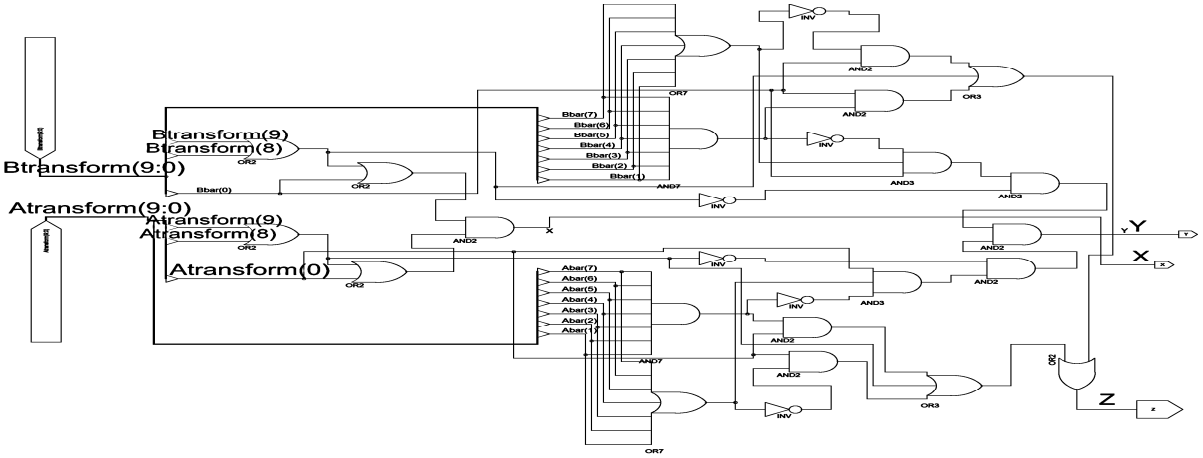



Fig. 3. Circuit diagram of the control signal module of Encoder

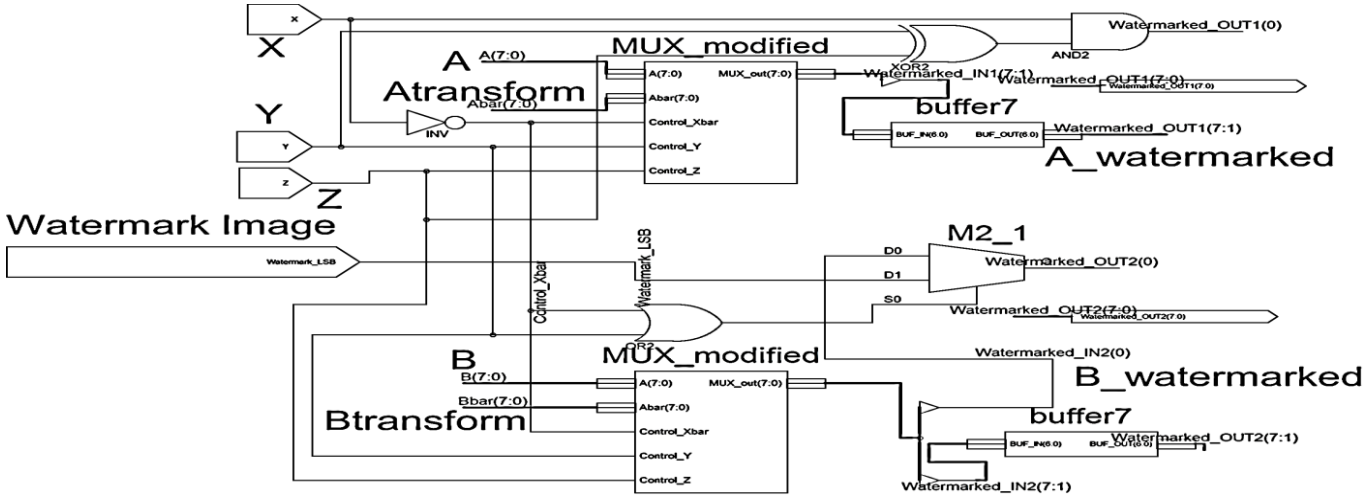



Fig. 4. Circuit diagram of the watermark data embedding module of encoder

primary task of this module is the generation of the signal labeled as $D_c$ which checks if the corresponding signal in the encoder input belonged to the domain $D_c$. The image transform module, previously used in encoder, followed by the $D_c$ check block and few logical blocks generate this signal. The LSB of $WI_A$ determines if the received pair was transformed. If the LSB, $WI_A(0)$, is equal to logical '1' then the pair was transformed. Consequently the pair is fed to the inverse transform module; else it is passed on to the Dc check block. This control is achieved by the 8-bit 2-1MUX and DEMUX pair. If the generated signal $D_c$ is satisfied, then the pair corresponds to one of the odd pairs transmitted by the encoder. It is then passed forward for further processing.

*Inverse Transform:*

This module is the most important part of the decoder and consumes major processing time. The received transformed pairs are performed inverse transform to get original image pixel. As mentioned in [5], the inverse transform is achieved by the mathematical expressions as given below:

$$A = \left\lceil \frac{2}{3} * A_{transform} + \frac{1}{3} * B_{transform} \right\rceil = \left\lceil \frac{1}{3} * (2 * A_{transform} + B_{transform}) \right\rceil \quad (3)$$

$$B = \left\lceil \frac{2}{3} * B_{transform} + \frac{1}{3} * A_{transform} \right\rceil = \left\lceil \frac{1}{3} * (2 * B_{transform} + A_{transform}) \right\rceil \quad (4)$$

where, $\lceil x \rceil$ denotes the ceil function (the smallest integer greater than or equal to x).

From (3) & (4), the above inverse transform can be executed by addition and division without using multiplier. Addition is performed twice followed by division by 3. All of these tasks are executed by the inverse transform block in Fig. 5. Keeping the cost constraint in mind, the division is performed by repetitive subtraction method limiting the overall decoding speed. However, the delay due to other combinatorial blocks is low enough to facilitate a high frequency clock. The divider used in this module had to be 10-bit because the upper limit of $A_{transform}$ and $B_{transform}$ (inputs WIA and WIB at the decoder) is 255 which could result a 10 bit input at the divider. The ceil function is achieved by a check operation performed on the two LSBs, $0^{th}$ and $1^{st}$ bit, followed by adder block. In Fig. 5, the check is performed by a single OR gate, and the output is added with the counter output, quotient, of the divider block.

*Image and Watermark Extraction:*

As mentioned earlier, watermark extraction isentirely performed depending on the control signals. The watermark image, embedded into the LSB of $WI_B$, is extracted by using the signal labeled Watermark_seq_sig in Fig. 5. This signal determines the $WI_B$ that contains the embedded watermark sequence. The watermark sequence is stored in a 128 byte RAM (four 32-word by 8-bit SRAM). The size of the storage device depends on the size of the watermark image. The signal labeled Watermark_seq_sig is applied to the Write Enable (WE) input to the RAM as shown in Fig. 5. The A and B output signals from the 8 bit MUX forms the extracted image. The select line of this MUX is generated from the LSB of the $WI_A$ and carry out ($C_{out}$) of the inverse transform block as shown in Fig. 5.

## IV. Analysis And Experimental Results

The simulation and implementation of the entire architecture is carried out in ISE Design suite and other tools of Xilinx. The hardware is optimized in terms of hardware cost. A $(32 \times 16)$ size binary watermark image is used to perform the watermark embedding and its extraction. The watermarking is performed on an 8 bit gray image of size$(32 \times 32)$. The experimental results and their analysis are summarized in following part of this section.

### 1. Encoder Results :

As discussed earlier, the watermark encoding process is carried by important blocks like the transform block, control signal block and the final watermark image embedding block. Considering the hardware complexity of these blocks, they were designed and implemented separately and then integrated

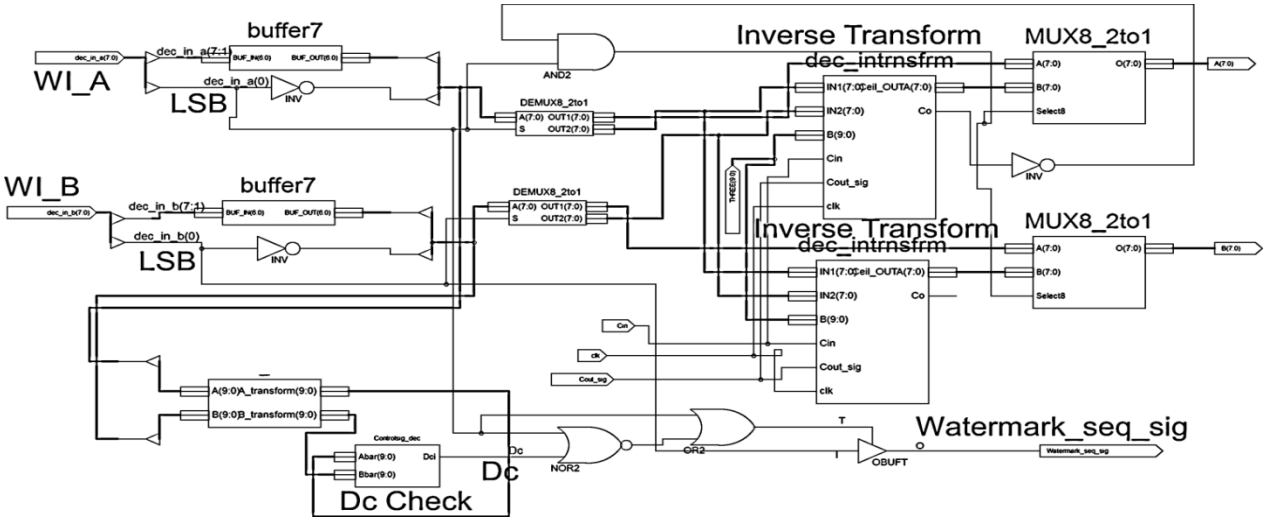



Fig. 5. Circuit diagram of the decoder comprised of all necessary modules

to perform the desired operation. The multiplier requirement of the transform operation of the pair of pixels is achieved using four subtractors in order to maintain hardware efficiency. The complex watermark embedding operation is very efficiently performed by using the control signals generated from combinational logical gates. Finally, the watermark embedding is realized using customized multiplexers and logical blocks. The implementation of the entire encoder required 303 slices and 528 four input LUTs. The hardware utilization of encoder along with its sub-blocks is given in Table I.

TABLE I: DEVICE UTILIZATION SUMMARY (ESTIMATED VALUES)

| Logic Utilization | Different modules | | | |
|---|---|---|---|---|
| | Pixel transform | Control signal | Watermark embedding | Watermarking encoder |
| Number of 4 input LUTs | 432 | 17 | 19 | 528 |
| Number of occupied Slices | 268 | 9 | 10 | 303 |
| Number of bonded IOBs | 152 | 23 | 50 | 230 |

The encoder module is practically devoid of any clock signals as the module is realized only using combinatorial and logical elements. As a result, the watermarking process is fast although some intensive operations are being performed. The clock less architecture can be extended by incorporating parallel processing and pipelining, enabling the system to be highly effective in real time applications like in digital cameras, printers, medical and military applications etc.

A concern of the implemented encoder is the combinational path delay whose maximum value is found out to be 31.642ns. However, there is a high scope of reducing this delay by employing pipeline architecture.

2. Decoder results:

The decoder module is having a higher complexity as compared to the encoder mainly because of the pixel inverse transform and the recovery of the original pixels. As a result, the hardware requirement is also higher that the encoder part which is detailed in Table II. The divider (division by 3) used for pixel inverse transformation is application specific and is of subtraction followed by right shifting type. It require only 613, 4-input LUTs, 347 slices and 56 slice flip-flops.This transform module also involved 4 adders, 1 subtractor, 8 bit counter and multiplexers.

TABLE II: DEVICE UTILIZATION SUMMARY (ESTIMATED VALUES)

| Logic Utilization | Different modules | | |
|---|---|---|---|
| | pixel inverse transform | control signal | Watermark extraction (decoder) |
| Number of Slice Flip Flops | 48 | 8 | 56 |
| Number of 4 input LUTs | 217 | 13 | 613 |
| Number of occupied Slices | 125 | 8 | 347 |
| Number of bonded IOBs | 42 | 20 | 37 |

Because of the hardware complexity, the implemented architecture is prone to lower response time. The incorporation of the pipelined architecture facilitated in reducing the delay to a minimum of 22.663 ns. The maximum clock frequency of the watermark extraction module is 45 MHz.

## V. CONCLUSION

This paper focuses on the design and implementation of a fast FPGA(Field Programmable Gate Array) based architecture using reversible contrast mapping (RCM) based image watermarking algorithm. To the best of our knowledge, prior research on RCM based watermarking algorithm with its VLSI implementation is very shallow. This limited the comparison of this hardware implementation with others and hence sole significance has been summarized. The encoder requires528 4 input LUTs and 303 slices. On the contrary, the decoder requires 613 LUTs and 347 slices. The encoder module is practically independent of clock, so the embedding of the watermark is possible as soon as the input is fetched. This feature along with low hardware cost facilitates the

prospect of its use in real-time applications like digital cameras, medical and military applications. The hardware complexity of the decoder module is higher compared to the encoder module because of the division followed by the ceil function in inverse transform module. The maximum clock frequency of the decoder is 45 MHz.The design is fast, low cost and easily implementable for real time watermarking.

## REFERENCES


[1] Fridrich, J., Goljan, M., & Du, R. (2001). Invertible authentication watermark for JPEG images. *Proceeding of the IEEE International Conferenceon Information Technology,* 223–227.

[2] Cox, I., Miller, M., & Jeffrey, B. (2002). *Digital watermarking: Principlesand practice*. Morgan Kaufmann.

[3] Van Leest, A., Van der Veen, M., & Bruekers, F. (2003). Reversible image watermarking. *Proceedings of the IEEE International Conference onImage Processing*, *II*, 731–734.

[4] J. Tian. Wavelet-based reversible watermarking for authentication. In E. J. Delp III and P. W. Wong, editors, Security and Watermarking of Multimedia Contents volume 4675 of Proc. of SPIE, pages 679-690, Jan. 2002

[5] *Coltuc*, D., Chassery, J.M.: *Very Fast Watermarking* by Reversible Contrast Mapping. IEEE Signal Processing Letters 14, 255–258 (2007).

[6] Juergen Seitz. Digital Watermarking for digital media, University of Cooperative Education Heidenheim, Germany, 2005

[7] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber. Reversible data hiding. In Proc. of International Conference on Image Processing, volume II, pages 157-160, Sept. 2002.

[8] Mohanty SP, Ranganathan N, Namballa RK. VLSI implementation of invisible digital watermarking algorithms towards the development of a secureJPEG encoder. In: Proceedings of the IEEE workshop on signal processing systems; 2003. p. 183–8.

[9] Mohanty SP, Kougianos E, Ranganathan N. VLSI architecture and chip for combined invisible robust and fragile watermarking. IETComput Digital Tech(CDT) 2007;1(5):600–11.

[10] Mohanty SP, Nayak S. FPGA based implementation of an invisible-robust image watermarking encoder. In: Lecture notes in computer science, vol.3356; 2004. p. 344–53.

[11] A. Garimella, M. V. V. Satyanarayan, R. S. Kumar, P. S. Murugesh, and U. C. Niranjan, "VLSI Impementation of Online Digital Watermarking Techniques with Difference Encoding for the 8-bit Gray Scale Images," in Proceedings of the International Conference on VLSI Design, 2003, pp. 283–288.

[12] S. P. Mohanty, N. Ranganathan, and R. K. Namballa, "A VLSI Architecture for Visible Watermarking in a Secure Still Digital Camera ($S^2$DC) Design," IEEE Transactions on Very Large Scale Integration Systems, vol. 13, no. 8, pp. 1002–1012, August 2005.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Design and Implementation of a BIST Embedded Inter-Integrated Circuit Bus Protocol over FPGA

Shumit Saha, Md. Ashikur Rahman, Amit Thakur
Department of Electronics & Communication Engineering
Khulna University of Engineering & Technology, Khulna, Bangladesh
shumit.ece.kuet@gmail.com, ashik_ece_kuet@yahoo.com, amit_kuet2k8@yahoo.com

***Abstract*— The I²C (Inter-Integrated Circuit) protocol is used to attach two devices for communicating with each other in a fast way excluding data losses. With the fast development of Integrated Circuits (ICs) technology, the complication of the circuits has also raised. Therefore, the complexity of the circuit requires self-testability in hardware to palliate the product failure. Built-in-self-test (BIST) is such a technique which can meet the necessity of self-testability with an effective solution over pricy circuit testing system. This paper represents designing and implementation of Inter-Integrated Circuit (I²C) protocol with self-testing ability. The need of programming for setting up a network with two devices is no longer needed in this proposed system. In order to attain compact, stable and reliable data transmission, the I²C is designed with Verilog HDL language and synthesized on Spartan 2 FPGA. An EEPROM and FPGA Spartan 2 are used for the communication testing where the FPGA is master and EEPROM is a Slave.**

***Keywords*— Inter-Integrated Circuit; Embedded built-in-self-test architecture; Verilog HDL; FPGA**

## I. INTRODUCTION

I2C or I square C or I²C – commonly known as Inter Integrated Circuit, is a bus protocol which enables faster devices to communicate with slower devices without any data loss. I²C (Inter IC) protocol was invented by Philips Semiconductors in January 2000. I²C protocol is considered as one of the very best among the systems that are connected to a number of devices and make the communication smooth and fast [1].

Several works have been done using VHDL in designing I²C. An FPGA based interface model is designed and implemented for scale-free network using I²C bus protocol on Quartus II 6.0 in [2]. On that paper, a generic design on an FPGA platform is presented and the entire design has been coded in VHDL and verified using Quartus II 6.0. A prototype of On-chip I²C module is designed for FPGA Spartan 3A series using Verilog HDL in [3].

This paper emphasizes on a new approach of designing I²C with BIST using Field Programmable Gate Array (FPGA) technology. Testing of a circuit has become increasingly tough as the scale of integration grows. I²C with the BIST capability provides the specified testability requisites and lowest-price with the highest performance implementation. Much lesser blocks and modules are used to design this I²C so that the testing complexity can be reduced. This system can be fabricated into a single chip. Verilog HDL is used for the coding of this system & designed, tested and evaluated using the ISE 6.0 tool of Xilinx and VeriloggerPro 6.5. For the design implementation the Xilinx Spartan-2 FPGA (XC2S150) is used [7].

The I²C protocol architecture, implementation technique of the system, circuit schematic and simulation results will be discussed briefly in the following sections. The system demands of high integration, low bit error rate and low cost can be satisfied by this I²C.

## II. I²C PROTOCOL ARCHITECHTURE

In this paper, the I²C protocol implementation uses two main buses: clock and data bus. Clock is a unidirectional bus which fed into the slave devices. Data is bidirectional bus which can transfer data from slave to master and vice versa. Master device first initiates a data transfer and generates the clock. Slave device then give acknowledge signal back. To transfer a data from master device to slave, there are three types of data formats required. Fig. 1 shows the data format of I²C protocol.

### *A. Slave Address*

It is 8 bit data format. By this slave address, the master determines the address of the slave device where it transmits the data.

### *B. Word Address*

It is also 8 bit data format. By this word address, the master tells the slave device the address of data. So, word address is actually like the address bus of the data.

### *C. Data Value*

Data values are 8 bit long. These are the values which transmit from master to slave.

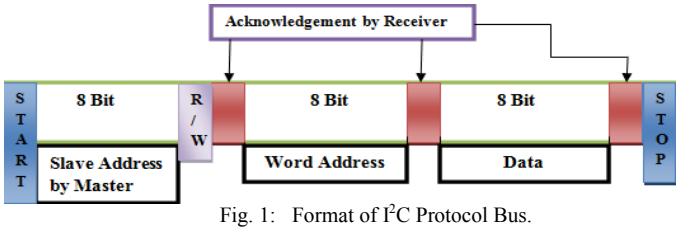



Fig. 1: Format of I²C Protocol Bus.

## III. PROPOSED ARCHITECHTURE

The Proposed structure consists of two modes. One is BIST mode where the $I^2C$ test itself. Another is normal mode. In normal mode the device works like usual $I^2C$ device.

### *A. BIST Module*

The design technique that grants a circuit to test itself is called Built-In-Self-Test (BIST). Combinational and sequential logic, memories, multipliers, and other embedded logic blocks can test themselves with this technique. Many cores are integrated into a single chip in the modern System-on-a-Chip (SoC) design. From outside of the chip cannot be accessed directly because most of them are embedded. Such SoC designs make the test of these embedded cores a great challenge [4].

Fig. 2 shows the structure of the $I^2C$ with BIST. The BIST control signal controls the BIST module. In the BIST module, there are four sub blocks. They are three random pattern generators and a comparator.

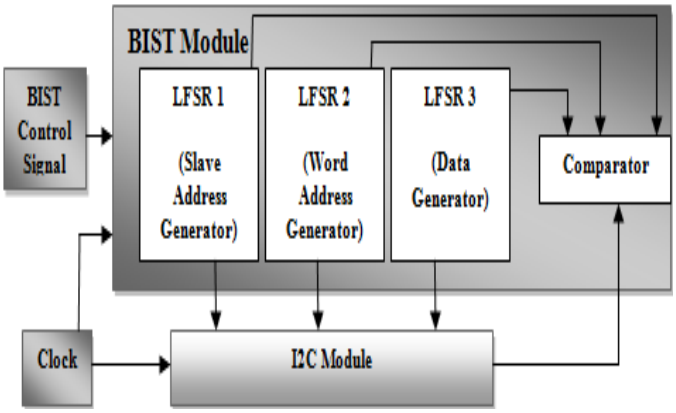



Fig. 2. BIST Structure.

*1) Random Pattern Generators (RPG):* Random Pattern Generator (RPG) generates random patterns which can be used for the verification of device like $I^2C$. The RPG is a part of the BIST in the verification of the circuits. Many methods have been proposed for the BIST equipment design [6], [8]. To produce bytes to test the circuit the method of a random pattern generator (RPG) is used.

This RPG consists of three LFSRs. LFSR 1 is used to generate the slave address. LFSR 2 produces word address. LFSR 3 gives the data. The generated bytes are used directly in the $I^2C$ to obtain better fault coverage. A comparator evaluates the response of the $I^2C$ with these bytes.

*2) Comparator:* This is a comparator which is used to compare the received and transmitted bit pattern. And then it gives the value of error. If the comparator gives bit stream of 111 then the device is perfect and running good. If it gives 101 then there are some faults occur in the protocol.

### *B. $I^2C$ Structure*

Fig. 3 shows the $I^2C$ flow chart. Table I demonstrates the operating modes with reset and reset_n signal. In the table it is seen that, BIST mode is on when reset pin is set low i.e. 0 & reset_n is 1 and vice versa.

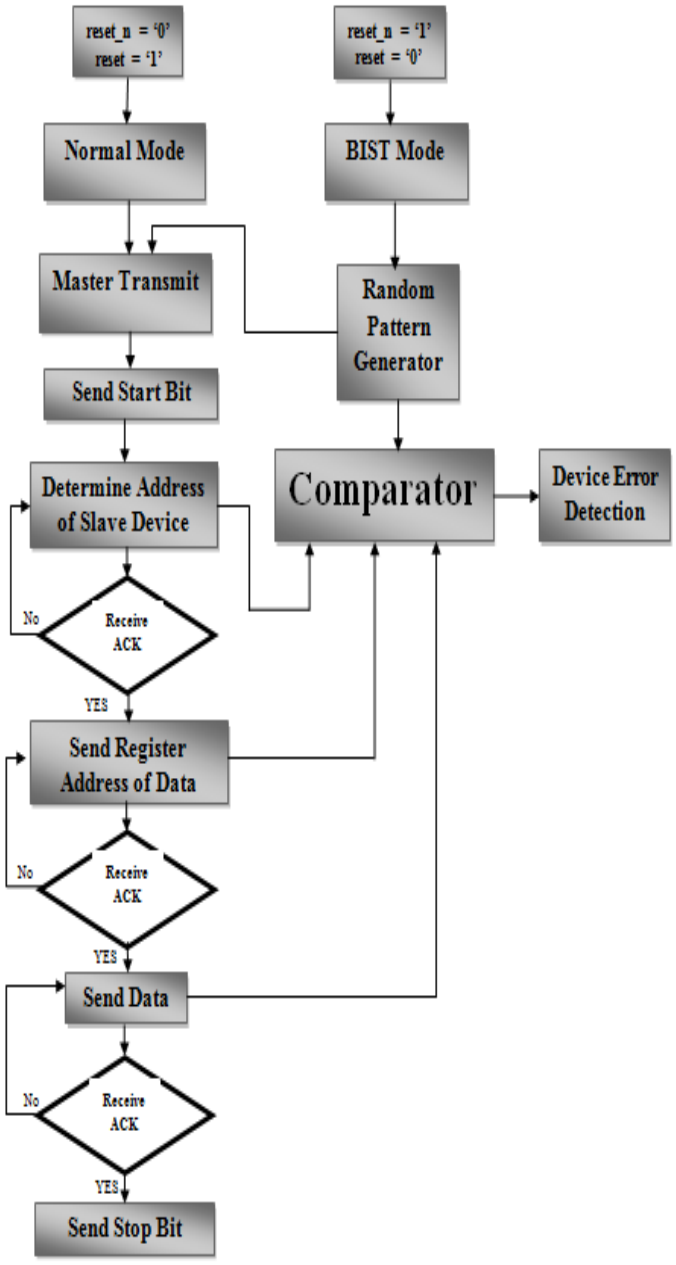



Fig. 3. $I^2C$ Flow Chart with BIST Mode.

TABLE I. OPERATING MODES OF $I^2C$

| reset | reset_n | BIST Mode | NORMAL Mode |
|---|---|---|---|
| **1** | **0** | OFF | ON |
| **0** | **1** | ON | OFF |

From Fig. 3, it is shown that when reset =1 and reset_n=0, then normal mode begins and master starts to communicate with its slave devices.

*a) Step 1:* Master sends the start bit at first to initiate transmission. The start condition is when the data line goes low and immediately after the clock line goes low.

*b) Step 2:* Master send the adress of the slave device.

*c) Step 3:* Slave give a acknowledgement signal to the master device.

*d) Step 4:* After receiving the acknowledgement signal, the master then sends the address of the register where to save the data.

*e) Step 5:* Slave again sends an acknowledgement signal to master device about the confirmation of word address reception.

*f) Step 6:* Master then sends the data to the slave using the data bus.

*g) Step 7:* After the reception of data the slave again gives an acknowledgement signal.

*h) Step 8:* Master then send the stop bit. When the clock line goes high from low and immediately after the data line goes high then this indicates a stop bit.

## IV. SYSTEM SYNTHESIZE & IMPLEMENTATION

### A. Circuit Schematic

Fig. 4 & Fig. 5 show the pin diagram and top level diagram of the I²C with BIST capability circuit respectively. Table II shows the pin descriptions of the top level schematics of the Verilog HDL implementation shown in the Fig. 4. In this table, the input and output pins are also specified.

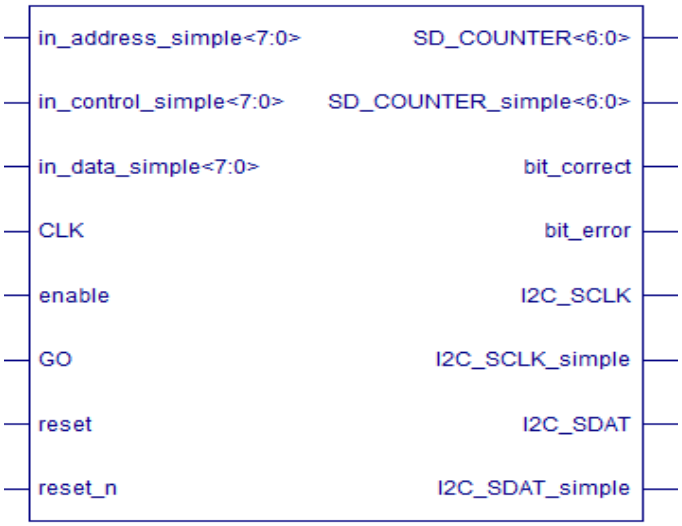



Fig. 4. Pin Diagram of I²C.

In Fig. 6, the top level schematic of BIST module is shown. As described earlier, BIST module consists of three LFSRs and one comparator is depicted in the Fig. 6.

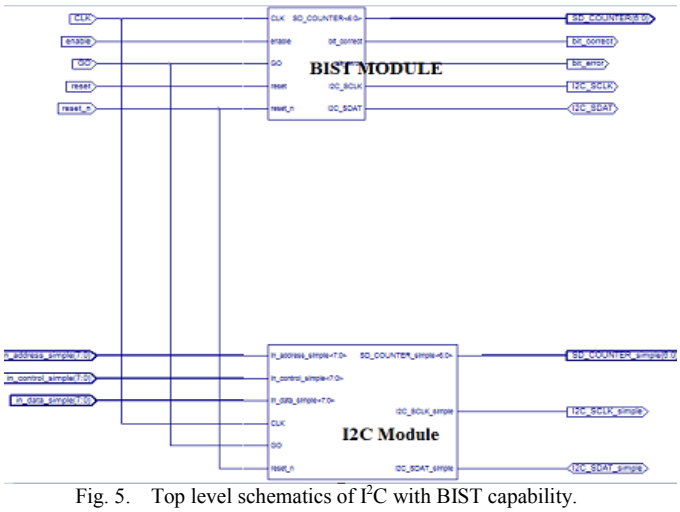



Fig. 5. Top level schematics of I²C with BIST capability.

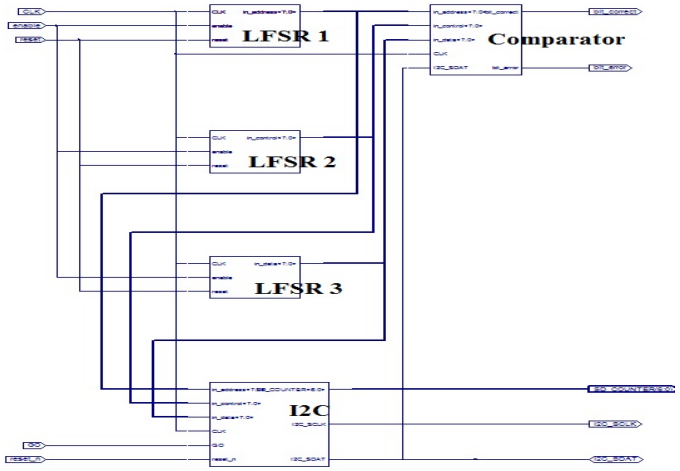



Fig. 6. Top level schematics of BIST Module.

TABLE II. I²C PIN DESCRIPTION

| Pin | IN/OUT | Description |
|---|---|---|
| CLK | IN | Clock generator |
| reset | IN | Bit for controlling BIST mode |
| reset_n | IN | Bit for controlling NORMAL mode |
| enable | IN | Enables the Random Pattern Generator |
| GO | IN | Control bit of the I2C |
| in_data_simple | IN | Input data byte of the I2C |
| in_control_simple | IN | Input control byte of the I2C |
| in_address_simple | IN | Input word address of the I2C |
| SD_COUNTER | OUT | CLK pulse counter for BIST Mode |
| SD_COUNTER_simple | OUT | CLK pulse counter for NORMAL Mode |
| I2C_SCLK | OUT | Output pin for I2C CLK for BIST Mode |
| I2C_SCLK_simple | OUT | Output pin for I2C CLK for BIST Mode |
| I2C_SDAT | OUT | Output data bus for BIST Mode |
| I2C_SDAT_simple | OUT | Output data bus for NORMAL Mode |
| bit_correct | OUT | Output pin of Comparator for correct bits |
| bit_error | OUT | Output pin of Comparator for wrong bits |

## B. Simulation Results

The timing diagram achieved from Testbencher showed that the received data in the receiver. Then the design is tested in the Xilinx FPGA where it also gave the correct output. The 8 bits of outputs are converted here into 2-digits Hexadecimal numbers.

### 1) Simulation Results for BIST Mode

#### a) 8-bit Random Bit Pattern Generator

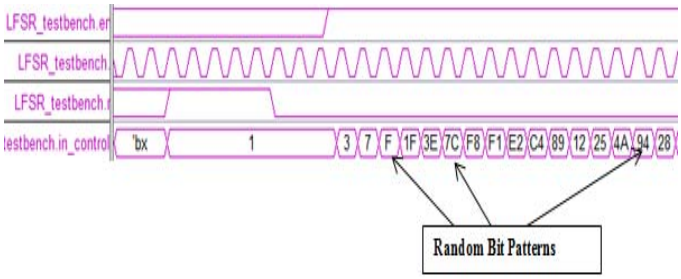



Fig. 7. 8-bit random bit pattern.

#### b) Output Data Format

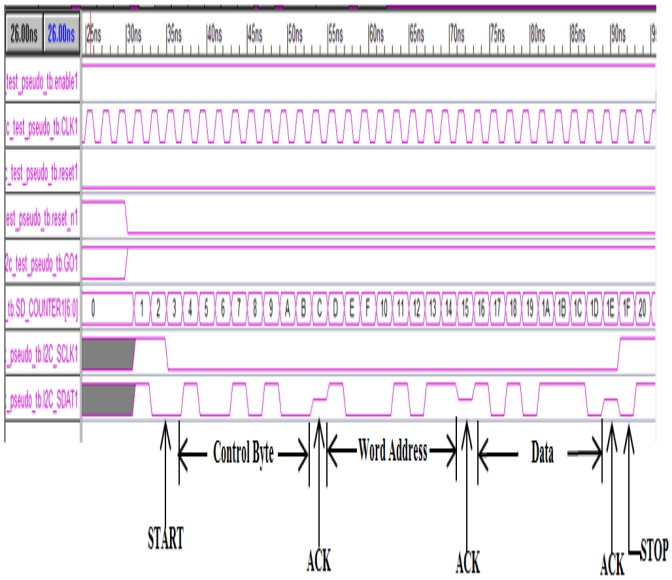



Fig. 8. BIST mode Output.

Fig. 7 demonstrates the output of the LFSR 1 in random pattern generator. Same type of outputs also come from the LFSR 2 and 3. These ouputs from LFSRs are directly goes into the I$^2$C module.

Fig. 8 shows the timing diagram of the I$^2$C output in the BIST mode. Three Random Pattern Generators generate random bits for the 8-bit control byte, 8-bit word address and 8-bit data. When the "reset_n" is low and "GO" is high then the control bytes, word address and data bits are passed depending on the "ACK" from the slave. When the "START" bit is passed, then the control byte start passing through the I2C bus. After that, an acknowledgement (ACK) is given by the slave. The next 8-bit is the word address or the memory address which is passed. Then, an "ACK" from the slave is given and 8-bit data is passed. Then, the "ACK" is given from slave and at last, the stop bit ends the operation. The process continues. And in this process, the efficiency and bit error rate of the I$^2$C can be tested.

### 2) Simulation Results NORMAL Mode

Fig. 9 depicts the output of the I$^2$C bus at "NORMAL Mode". The Control Byte, Word Address and Data in hexadecimal are "CA", "00" and "AA" respectively which are found in the I$^2$C bus as "11001010", "00000000" and "10101010" respectively when the "reset_n" is low and "GO" is high. The after receiving each byte (i.e., control byte, word address and data) an acknowledgement (ACK) is given by the slave. With the "START" bit the operation starts and with the "STOP" bit the operation ends.

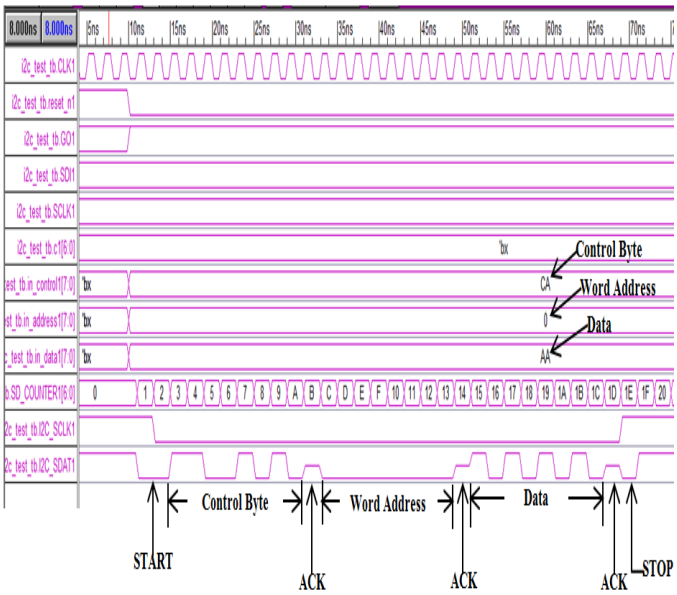



Fig. 9. Normal Mode Output Stream.

This output is justified by Table III.

Table III. RECEIVER OUTPUT FORMAT

| Input Type | Binary | Hexadecimal |
|---|---|---|
| Control Byte | 11001010 | CA |
| Word Address | 00000000 | 00 |
| Data | 10101010 | AA |

## C. FPGA Implementation

FPGAs can be used to implement any logical function that can be performed by an ASIC, and it should also be noted that FPGA design is more cost-effective than that of ASIC. They have lots of advantages over microcontrollers, such as greater speed, number of I/O ports and performance.

The proposed design is implemented on Xilinx Spartan-2 FPGA (XC2S150). So here, the master device is Xilinx Spartan-2 FPGA. The slave device used here is EEPROM in FPGA. Here, EEPROM means electrically erasable programmable read only memory. EEPROM is a non-volatile memory. It is used as a slave device to store small amounts of configuration information. Fig. 10 shows the FPGA implementation with Clock and Data bus.

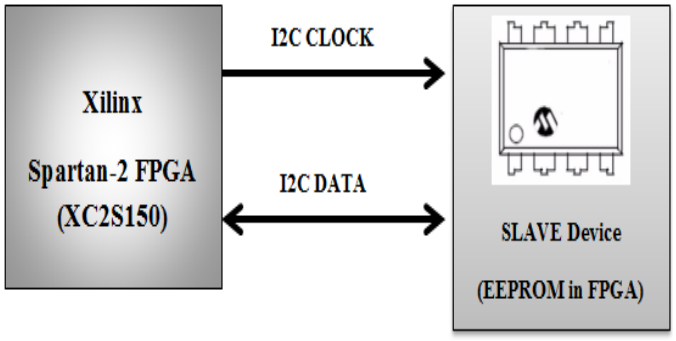



Fig. 10. FPGA Implementation of I2C Bus Protocol.

The Device utilization summary and timing summary are given in Table IV and Table V respectively. From the table it is seen that only 21% if the flip flops are used. In case of input output buffers it is 28%. So that, there are many more feathers can be added with the proposed architecture. In the timing summary, it is seen that the maximum delay is 5.936ns which is much lesser than conventional UARTs.

Table IV. DEVICE UTILIZATION SUMMARY

| Name | Used Blocks | Percentages (%) |
|---|---|---|
| Number of Slices | 41 out of 192 | 21 |
| Number of Slice Flip Flops | 44 out of 384 (FDRE:35 FDSE :3 LD: 2, LD_1: 4) | 11 |
| Number of 4 input LUTs | 72 out of 384 | 18 |
| Number of bonded IOBs | 26 out of 90 (IBUF :6 OBUF : 18 OBUFT: 2) | 28 |
| Number of GCLKs | 2 out of 4 | 50 |

Table V. TIMING SUMMARY

| Parameters | Seconds |
|---|---|
| Minimum period | 6.016ns |
| Minimum input arrival time before clock | 8.641ns |
| Maximum output required time after clock | 8.183ns |
| Maximum delay | 5.936ns |

## V. CONCLUSION

In this paper, an FPGA based implementation of $I^2C$ with BIST capability is presented. Here all the modules are designed and simulated with Verilog HDL. Then the system is downloaded in the Xilinx Spartan-2 FPGA (XC2S150). This $I^2C$ is much more flexible, speedy, low cost, and stable with respect to conventional one. This $I^2C$ control bus architecture can enable the industrial fabrication of chip in a way where only a pressing of one switch can test itself. So that, it would save valuable time and cost of testing circuits significantly.

## References


[1] *I2C Bus Specification, Version 2.1,* Philips Semiconductors, 2000.

[2] P. Venkateswaran, M. Mukherjee, A. Sanyal, S. Das, and R. Nandi, "Design and Implementation of FPGA Based Interface Model for Scale-Free Network using I2C Bus Protocol on Quartus II 6.0," in *Proc. 4th International Conference on Computers and Devices for Communication*, Dec. 2009, pp. 1-4.

[3] R. Singh, and N. Sharma, "Prototyping of On-chip I2C Module for FPGA Spartan 3A series using Verilog," *International Journal of Computer Applications*, vol. 68, no. 16, April 2013.

[4] M. Bushnell and V.D. Agarwal, *Essentials of electronic testing for digital, memory and mixed-signal VLSI circuits*, Kluwer Academic Publishers, 2000.

[5] S. Jamuna and Dr. V.K. Agrawal, "Implementation of BIST structure using VHDL for VLSI circuits," *International Journal of Engineering Science and Technology,* vol. 3, no. 6, pp. 5041-5048, June 2011.

[6] V.K. Agrawal, C.R. Kime and K.K., Saluja, "A tutorial on BIST, part 1: Principles," *IEEE Design & Test of Computers*, vol. 10, No.1, pp.73-83, March 1993.

[7] J. Bhasker, *Verilog® HDL synthesis: a practical primer*, Star Galaxy Publishing, 1998.

[8] M.Y.I. Idris, M. Yaacob, "A VHDL implementation of BIST technique in UART design," in *Proc. Conference on Convergent Technologies for the Asia-Pacific Region (TENCON),* Oct. 2003, pp.1450-1454

[9] S. Saha, M. A. Rahman, A. Thakur, "Design and Implementation of a BIST Embedded High Speed RS-422 Utilized UART over FPGA," in *Proc. of 4th IEEE International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Jul. 2013.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# A New Assignment of Free Links in Midimew Connected Mesh Network

Md. Rabiul Awal*, M.M. Hafizur Rahman*, Rizal Bin Mohd Nor*,
Tengku Mohd Bin Tengku Sembok†, and M.A.H. Akhand‡

*Dept. of Computer Science, KICT, IIUM, Kuala Lumpur 50728, Malaysia
E-mail: rabiulawal1@gmail.com, hafizur@iium.edu.my, rizalmohdnor@iium.edu.my
†Cyber Security Center, National Defense University Malaysia, Kuala Lumpur 57000, Malaysia
E-mail: tmtsembok@gmail.com
‡Dept. of Computer Science and Engineering, KUET, Khulna - 9203, Bangladesh
E-mail: akhandkuet@yahoo.com,

***Abstract*— A Midimew connected Mesh Network (MMN) is a MInimal DIstance MEsh with Wrap-around links network of multiple basic modules, in which the basic modules are 2D-mesh networks that are hierarchically interconnected for higher-level networks. In this paper, we present a new assignment of free links to form the higher level networks of the MMN (called HMMN), addressing of node, routing of message, and evaluate the static network performance of the HMMN, its counter rival VMMN, TESH, mesh, and torus networks. It is shown that the proposed HMMN possesses several attractive features, including constant degree, small diameter, low cost, small average distance, moderate bisection width, and same fault tolerant performance than that of other conventional and hierarchical interconnection networks. It is also shown that with the same node degree, arc connectivity, bisection width, and wiring complexity, the diameter and average distance of the HMMN is lower than that of the TESH network and trivially higher to that of VMMN.**

***Index Terms*— Massively Parallel Computers, Interconnection Network, Free Links, MMN, and Static Network Performance.**

## I. Introduction

The demand of computation power will never stop, it is increasing day by day. However, sequential computer steadily increases their speed to meet the computation demand, and it has already been reached saturated. Thus, the only way to meet the increasing demand of computation power to solve the grand challenge problems is to use parallel computers. Parallel computers provide computationally efficient techniques for the solution of large and difficult problem in a reasonable time [1]. To solve the grand challenge problems in many areas such as development of new materials and sources of energy, development of new medicines and improved health care, strategies for disaster prevention and mitigation, weather forecasting, and for scientific research including the origins of matter and the universe. This makes the current supercomputer changes into massively parallel computer (MPC) systems with thousands of node (Tianhe-2, Titan, Sequia, Kei, Cray XT5-HE [2]), that satisfy the insatiable demand of computing power. In near future, we will need computer systems capable of computing at the tens of petaflops or exaflops level, or even more. To achieve this level of performance, we need MPC with tens of thousands or even millions of nodes [3]. It is predicted that the parallel systems of the next decade will contain 10 to 100 millions of nodes [4], [5].

One crucial step on designing such a MPC systems is to determine the topology of the interconnection network, because the overall performance is significantly affected by the network topology [6]–[8]. Many recent experimental and commercial parallel computers use direct networks for low latency and high bandwidth of inter-node communication. For future MPC with millions of nodes, the large diameter of conventional topologies is completely infeasible. The hierarchical interconnection network (HIN) provides a cost-effective way in which several network topology can be integrated together [9]. Therefore, HIN is a plausible alternative way to interconnect the future MPC [9] systems. A variety of hypercube based HINs found in the literature, however, its huge number of physical links make it difficult to implement. To alleviate this problem, several $k$-ary $n$-cube based HIN have been proposed [10], [11]. However, the performance of these networks does not yields any obvious choice of a network for MPC. No one is clear winner in all aspect of MPC design.

Since reducing the diameter is likely to improve the performance of an interconnection network, the problem of designing interconnection network with low diameter with scalability of network size is still desirable [11], [12]. A Tori connected mESH (TESH) network [13], [14] is an HIN aiming for large-scale 3D MPC systems, consisting of multiple basic modules (BMs) which are 2D-mesh networks. The BMs are hierarchically interconnected by a 2D-torus etwork to build higher level networks. The topology of the contemporary massively parallel computers use $k$-ary $k$-cube networks. No HIN draws the potential attention from the industry community yet. However, TESH network is a suitable network because of constant & small node degree and scalability for modular expansion.

The main objective of this paper is to find a network

which is suitable for interconnecting a large number of nodes while keeping small diameter. It has already been shown that a **MI**nimal **DI**stance **ME**sh with **W**rap-around links **(midimew)** network is an optimal topology in the sense that there is no direct symmetric network of degree 4 with lower diameter or average distance [16]. To fulfill our objective, with this key motivation, we have replaced the higher level 2D-torus of a TESH network by a 2D midimew network. To use the free ports in the periphery of the 2D-mesh network for higher level interconnection, we kept the basic module as 2D-mesh network same as TESH network. This modified TESH network is a 2D-midimew networks of multiple BMs, in which the BMs are 2D-mesh networks that are hierarchically interconnected for higher-level networks. Analogous to the TESH network, we called it **M**idimew-connected **M**esh **N**etwork (**MMN**). MMN is a hierarchical interconnection network, thus allowing exploitation of computation locality, as well as providing scalability up to a million of nodes.

In this paper, we address a new assignment of the free links of the BM to interconnect the higher level network of the MMN and evaluate its static network performance. The static network performance will be evaluated in terms of node degree, network diameter, cost, average distance, bisection width, and arc connectivity.

The remainder of the paper is organized as follows. In Section II, we briefly describe the new assignment of the free links of the MMN. Addressing of nodes of the MMN and the routing of messages are discussed in Section III and Section IV, respectively. The static network performance of the proposed MMN is discussed in Section V. Finally, in Section VI, we conclude this paper.

## II. Interconnection of the MMN

The **M**idimew connected **M**esh **N**etwork (**MMN**) is a hierarchical interconnection network consisting of multiple basic modules (BM) that are hierarchically interconnected to form a higher level network. A $(2^m \times 2^m)$ BM consists of a 2D-mesh network of $2^{2m}$ processing elements (PE) having $2^m$ rows and $2^m$ columns, where $m$ is a positive integer. Considering $m = 2$, a BM of size $(4 \times 4)$ is portrayed in Fig. 1. Each BM has $2^{m+2}$ free ports at the contours for higher level interconnection. All ports of the interior nodes are used for intra-BM connections. All free ports of the exterior nodes, either one or two, are used for inter-BM connections to form higher level networks. In this paper, BM refers to a Level-1 network.

Successive higher level networks are built by recursively interconnecting $(2^{2m})$ immediate lower level sub-networks in a $(2^m \times 2^m)$ *(midimew)* network. In a midimew network, one direction (either horizontal or vertical) is symmetric tori connected and other direction is diagonally wrap-around connected. In our previous study [17], we have assigned the horizontal free links of the BM for symmetric tori connection and vertical free links are used for diagonal wrap-around connection. To the thirst of more efficient way to interconnect the higher level networks, we reverse the connection, i.e., vertical free links of the BM for symmetric tori connection and horizontal free links are used for diagonal wrap-around connection. For clarification, let us call the former one as vertical MMN (VMMN) and this proposed one is called horizontal MMN (HMMN).

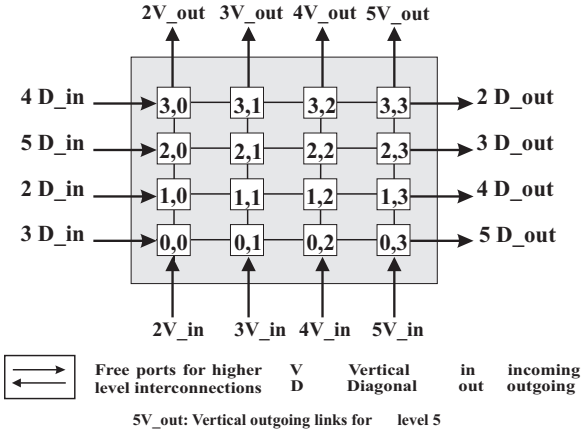



Figure 1. Basic module of the MMN

As depicted in Fig. 2, considering $(m = 2)$ a Level-2 HMMN can be formed by interconnecting $2^{2\times2} = 16$ BMs. Similarly, a Level-3 network can be formed by interconnecting 16 Level-2 sub-networks, and so on. Each BM is connected to its logically adjacent BMs. It is useful to note that for each higher level interconnection, a BM uses $4 \times (2^q) = 2^{q+2}$ of its free links, $2(2^q)$ free links for diagonal interconnections and $2(2^q)$ free links for horizontal interconnections. Here, $q \in \{0, 1, ....., m\}$, is the inter-level connectivity. $q = 0$ leads to minimal inter-level connectivity, while $q = m$ leads to maximum inter-level connectivity. For example the $(4\times4)$ BM has $2^{2+2} = 16$ free ports as shown in Figure 1. If we chose $q = 0$, then $4(2^0) = 4$ of the free ports and their associated links are used for each higher level interconnection, 2 for horizontal and 2 for diagonal interconnection. Among these 2 links, one is used for incoming link and another one for used for outgoing link, i.e., a single links is used for diagonal_in, diagonal_out, horizontal_in, and horizontal_out.

A HMMN$(m, L, q)$ is constructed using $2^m \times 2^m$ BMs, has $L$ levels of hierarchy with inter-level connectivity $q$. In principle, $m$ could be any positive integer value. However, if $m = 1$, then the network degenerates to a hypercube network. Hypercube is not a suitable network, because its node degree increases along with the increase of network size. If $m = 2$, then it is considered the most interesting case, because it has better granularity than the large BMs. If $m \geq 3$, the granularity of the family of networks is coarse. If $m = 3$, then the size of the BM becomes $(8 \times 8)$ with $64$ nodes. Correspondingly, the Level-2 network would have $64$ BMs. In this case, the total number of nodes in a Level-2 network is $N = 2^{2\times3\times2} = 4096$ nodes, and Level-3 network would have $262144$ nodes. Clearly, the granularity of the family of networks is rather coarse.

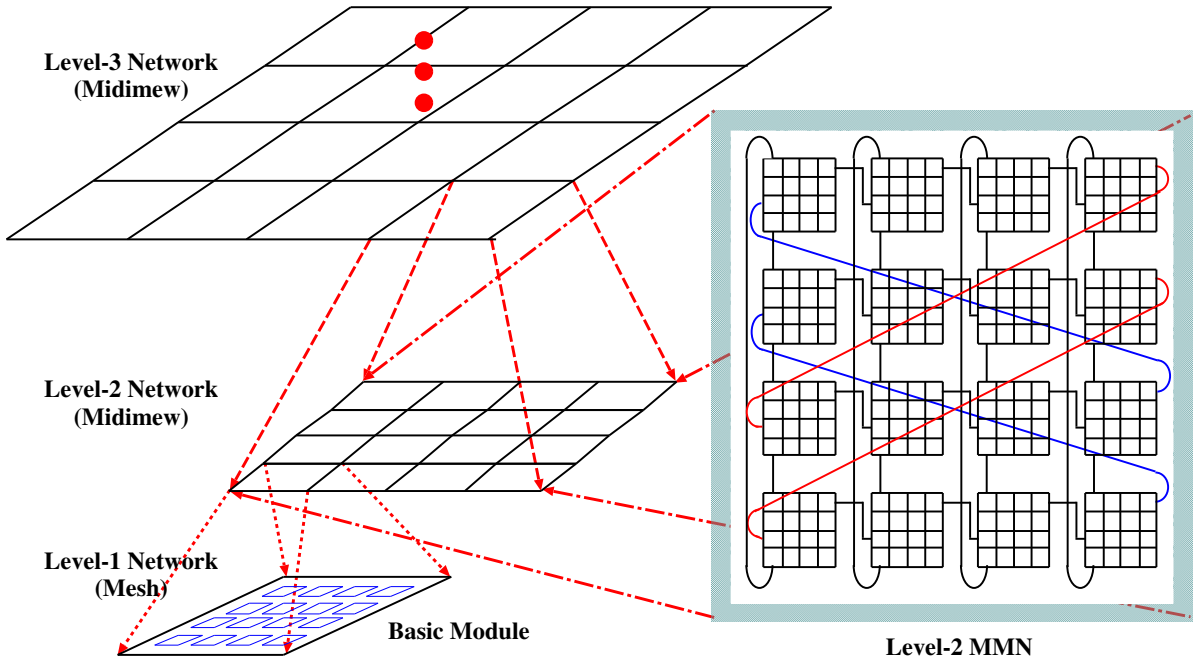



Figure 2. Higher level network of a MMN

In the rest of this paper we consider $m = 2$, therefore, we focus on a class of HMMN$(2, L, q)$ networks.

The highest level network which can be built from $(2^m \times 2^m)$ BM is $L_{max} = 2^{m-q} + 1$. With $q = 0$ and $m = 2$, $L_{max} = 5$, Level-5 is the highest possible level. The total number of nodes in a HMMN having $(2^m \times 2^m)$ BMs is $N = 2^{2mL}$. Using maximum level of hierarchy, $L_{max} = (2^{m-q} + 1)$, the maximum number of nodes which can be interconnected by a HMMN$(m, L, q)$ is $N = 2^{2m(2^{m-q}+1)}$. For the case of $(4 \times 4)$ BM with $q = 0$, a network consists of 1 million nodes.

## III. Addressing of Nodes

Base-$2^m$ numbers are used for convenience of address representation. As seen in Figure 1, nodes in the BM are addressed by two digits, the first representing the row index and the next representing the column index. More generally, in a Level-$L$ HMMN, the node address is represented by:

$$\begin{aligned} A &= A^L A^{L-1} A^{L-2} \ \ldots\ \ldots\ \ldots\ A^2 A^1 \\ &= a_{n-1} a_{n-2} a_{n-3} \ \ldots\ \ldots\ \ldots\ a_2 a_1 a_0 \\ &= a_{2L-1} a_{2L-2}\ a_{2L-3} a_{2L-4} \ \ldots\ \ldots\ \ldots\ a_3 a_2\ a_1 a_0 \\ &= (a_{2L-1}\ a_{2L-2})\ (a_{2L-3}\ a_{2L-4})\ \ldots\ \ldots \\ &\quad \ldots\ \ldots\ \ldots (a_3\ a_2)\,(a_1\ a_0) \end{aligned} \tag{1}$$

Here, the total number of digits is $n = 2L$, where $L$ is the level number. $A^L$ is the address of level $L$ and $(a_{2L-1} a_{2L-2})$ is the co-ordinate position of Level-$(L-1)$ for Level-$L$ network. Pairs of digits run from group number 1 for Level-1, i.e., the BM, to group number $L$ for the $L$-th level. Specifically, $l$-th group $(a_{2l-1}\ a_{2l-2})$ indicates the location of a Level-$(l-1)$ sub-network within the $l$-th group to which the node belongs; $1 \le l \le L$. In a two-level network the address becomes $A = (a_4\ a_3)\ (a_1\ a_0)$. The first pair of digits $(a_4\ a_3)$ identifies the BM to which the node belongs, and the last pair of digits $(a_1\ a_0)$ identifies the node within that BM.

The assignment of inter-level ports for the higher level networks has been done quite carefully so as to minimize the higher level traffic through the BM. The address of a node $n^1$ encompasses in BM$_1$ is represented as $n^1 = \left(a^1_{2L-1} a^1_{2L-2} \ \ldots\ \ldots\ \ldots\ a^1_3 a^1_2\ a^1_1 a^1_0\right)$. The address of a node $n^2$ encompasses in BM$_2$ is represented as $n^2 = \left(a^2_{2L-1} a^2_{2L-2} \ \ldots\ \ldots\ \ldots\ a^2_3 a^2_2\ a^2_1 a^2_0\right)$. The node $n^1$ in BM$_1$ and $n^2$ in BM$_2$ are connected by a link if the following condition is satisfied.

$\exists i\{a^1_i = (a^2_i \pm 1) \text{mod } 2^m\ \wedge \forall j(j \neq i\ \rightarrow\ a^1_j = a^2_j)\}$
where $i\%2 = 0$, $i, j \ge 2$
$\exists i\{a^1_i = (a^2_i \pm 1)\ \wedge \forall j(j \neq i\ \rightarrow\ a^1_j = a^2_j)\}$
where $a^1_i = a^2_i$, $a^1_j < 3$
$\exists i\{a^1_i = (a^2_i \pm 1) \text{mod } 2^m \wedge \forall j(j \neq i \rightarrow a^1_j = a^2_j + 2)$
$\%(2^m - 1)\}$ where $i\%2 = 1$, $i, j \ge 2$

## IV. Routing Algorithm for MMN

Routing of messages in the MMN is performed from top to bottom as in TESH network [14], [15]. That is, it is first done at the highest level network; then, after the packet reaches its highest level sub-destination, routing continues within the sub-network to the next lower level sub-destination. This process is repeated until the packet arrives at its final destination. When a packet is generated at a source node, the node checks its destination. If the packets destination is the current BM, the routing is

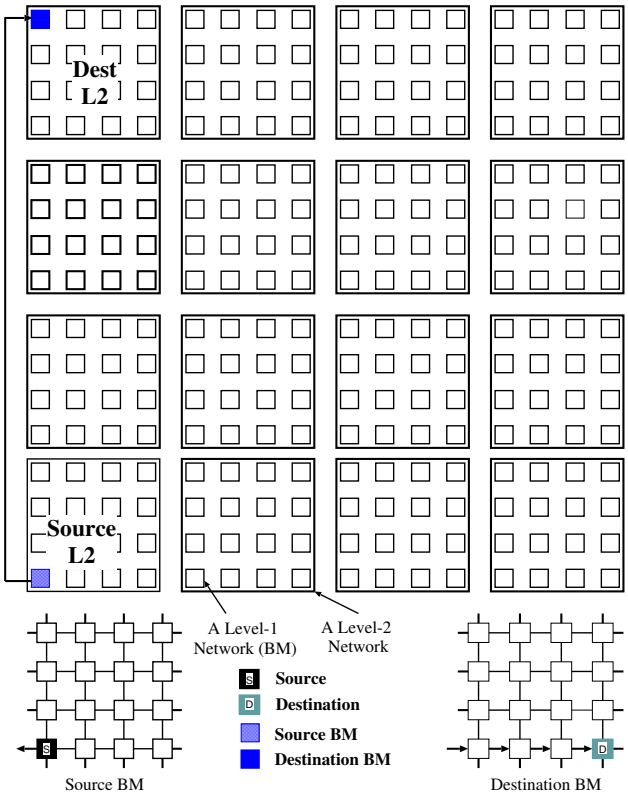



Figure 3. An example of routing algorithm of the MMN

performed within the BM only. If the packet is addressed to another BM, the source node sends the packet to the outlet node which connects the BM to the level at which the routing is performed. At each level, horizontal routing is performed first when the packet reaches the destined column, then diagonal routing start to reach the packet in the destination.

Let us consider an example in which a packet is to be routed from source node 000000 to destination node 130303. In this case, routing is to be done at Level-3, therefore the source node sends the packet to the outlet node of Level-3, 00 30 30, whereupon the packet is routed at Level-3, as shown in Figure 3. To avoid clutter diagonal wrap-around links are not shown in the figure. After the packet reaches the Level-2_$(1,1)$ network, then routing within that network continues until the packet reaches the $\mathrm{BM}(1,0)$. Finally, the packet is routed to its destination.

## V. Static Network Performance

Although the actual performance of a network depends on many technological and implementation issues, several topological properties and performance metrics can be used to evaluate and compare different interconnection networks in a technology-independent manner. Most of these properties are derived from the graph theoretic model of the topology. In this section, we discuss some graph theory metrics that characterize the cost and performance of an interconnection network. These parameters have a direct impact on network performance. For the performance evaluation, we have considered mesh, torus, TESH network, and proposed HMMN along with VMMN.

### A. Node Degree

The *node degree* is defined as the maximum number of physical links emanating from a node. Constant degree networks are easy to expand, and the network interface cost of a node remains unchanged with increasing network size. The I/O interface cost of a particular node is proportional to its degree. For fair comparison, we have consider degree $4$ network. It is shown in Table I that the degree of the mesh, torus, TESH, VMMN, and HMMN are equal, it is $4$ are independent of network size.

### B. Diameter

The distance between two adjacent nodes is unity. The *diameter* of a network is the maximum inter-node distance, i.e., the maximum number of links that must be traversed to send a message to any node along the shortest path. Diameter is the maximum distance among all distinct pairs of nodes along the shortest path. The diameter is commonly used to describe and compare the static network performance of the interconnection network topology. Networks with small diameters are preferable. The smaller the diameter of a network, the shorter the time to send a message from one node to the node farthest away from it. Also, network diameter determines the highest latency.

We have evaluated the diameter of the TESH, VMMN and HMMN network by simulation and mesh and torus network by their static formula and the results are tabulated in Table I. Clearly, the HMMN has a much smaller diameter than that of TESH and mesh networks; and a slightly high diameter than that of VMMN and torus networks.

### C. Cost

Inter-node distance, message traffic density, and fault-tolerance are dependent on the diameter and the node degree. The product $(diameter \times node\ degree)$ is a good criterion for measuring the relationship between cost and performance of a parallel computer systems [18]. An interconnection network with a large diameter has a very low message passing bandwidth, and a network with a high node degree is very expensive. In addition, a network should be easily scalable; there should be no changes in the basic node configuration as we increase the number of nodes. The cost of different networks is plotted in Table I, and it is shown that the cost of HMMN is far lower than that of mesh and TESH networks, exactly equal to that of VMMN, and slightly higher than that of torus network.

### D. Average Distance

The *average distance* is the mean distance between all distinct pairs of nodes in a network. Average distance is

TABLE I.
COMPARISON OF STATIC NETWORK PERFORMANCE OF VARIOUS NETWORKS

| | Node Degree | Diameter | Cost | Average Distance | Arc Connectivity | Bisection Width | Wiring Complexity |
|---|---|---|---|---|---|---|---|
| 2D-Mesh | 4 | 30 | 120 | 10.67 | 2 | 16 | 480 |
| 2D-Torus | 4 | 16 | 64 | 8.00 | 4 | 32 | 512 |
| TESH(2,2,0) | 4 | 21 | 84 | 10.47 | 2 | 8 | 416 |
| VMMN(2,2,0) | 4 | 17 | 68 | 9.07 | 2 | 8 | 416 |
| HMMN(2,2,0) | 4 | 17 | 68 | 9.33 | 2 | 8 | 416 |

reflected on the average network latency. A small average distance results small communication latency, especially for distance-sensitive routing, such as store and forward. But it is also crucial for distance-insensitive routing, such as wormhole routing, since short distances imply the use of fewer links and buffers, and therefore less communication contention. We have evaluated the average distances for HMMN, VMMN, and TESH network by simulation and mesh and torus networks by their corresponding formulas and the results are tabulated in Table I. It is shown that the average distance of HMMN is lower than that of mesh and TESH networks, and slightly higher than that of VMMN and torus networks.

Although the dynamic communication performance of a program on a parallel computer depends on the actual times taken for data transfer, a smaller average distance and diameter of an interconnection network yields a smaller communication latency of that network.

### E. Bisection Width

The *Bisection Width (BW)* of a network is defined as the minimum number of links that must be removed to partition the network into two equal halves. Many problems can be solved in parallel using *binary divide-and-conquer:* split the input data set into two halves, and solve them recursively on both halves of the interconnection network in parallel, then merge the results from both halves into the final result. Small bisection width implies low bandwidth between the two halves, and it can slow down the final merging phase. On the other hand, a large bisection width is undesirable for the VLSI design of the interconnection network, since it implies a lot of *extra chip wires*.

Bisection bandwidth reflects the communication capacity between any two halves of the network when dealing with uniform, non-uniform, local, and non-local traffic. It is calculated by counting the number of links that need to be removed to partition the Level-$L$ HMMN. We have calculated the average distance of HMMN, VMMN, TESH, mesh, and torus networks by their respective static formula and they are shown in Table I. It is shown that the the bisection width of the HMMN is exactly equal to that of the VMMN and TESH network; and it is lower than that of mesh and torus network.

### F. Arc Connectivity

*Arc Connectivity* measures the robustness of a network and the multiplicity of paths between nodes. Arc connectivity is the minimum number of links that must be removed in order to break the network into two disjoint parts. High arc connectivity improves performance during normal operation by avoiding link congestion, and also improves fault tolerance. The ratio between arc connectivity and the degree of a node gives a measure of static fault tolerance performance. A network is maximally fault-tolerant if its connectivity is equal to the degree of the network. The arc connectivity of various networks are tabulated in Table I. Clearly, the arc connectivity of the HMMN is exactly equal to that of VMMN and TESH network. However, the arc connectivity of the torus network is exactly equal to its degree. Thus, torus is more fault tolerant than all the networks. HMMN is exactly equal fault tolerant to that of VMMN, mesh, and TESH networks.

### G. Wiring Complexity

The *wiring complexity* of an interconnection network refers to the total number of links required to form the network. It has a direct correlation to hardware cost and complexity. A $(16 \times 16)$ 2D-mesh and 2D-torus networks have $\{N_x \times (N_y - 1) + (N_x - 1) \times N_y\} = 16 \times (16 - 1) + (16 - 1) \times 16 = 480$ and $(2 \times N_x \times N_y = 2 \times 16 \times 16) = 512$ links, respectively. $N_i$ represents the number of nodes in the $i^{th}$ dimension. The wiring complexity of a Level-$L$ HMMN, VMMN, and TESH networks is $\left[\#\text{ of links in a BM} \times k^{2(L-1)} + \sum_{x=2}^{L} 2(2^q) \times k^{2(L-1)}\right]$. Considering, $m = 2$, a BM of HMMN, VMMN, and TESH network have 24 links. Hence the total number of links of a Level-2 HMMN, VMMN, and TESH are 416. Table I compares the wiring complexity of a HMMN with that of several other networks. It is shown that the total number of links of HMMN is lower than that of mesh and torus network and exactly equal to that of VMMN and TESH network.

### H. Some Generalization

The question may arise, whether we need HMMN over torus, where torus network results lower diameter and average distance to that of HMMN as shown in Table I. The answer is 'yes'. The torus network has a large number of long wrap-around links. The operating speed of a network is limited by the physical length of its links. The length of the longest wire may become more important than the diameter of the network.

This is the very first stage of research on MMN. We tried in different way to find the best suitable assignment of free links for higher level interconnection of MMN. As mentioned earlier in our previous study [17], we used the horizontal free links as tori connected and vertical free links as diagonal wrap-around connected, called VMMN. In this research we reverse the connection. We used the vertical free links as tori connected and horizontal free links as diagonal wrap-around connected, called HMMN.

It is seen from Table I that the static network performance of the HMMN is better than that of mesh and TESH network. And almost equal to that of VMMN. Only the average distance of the HMMN is slightly higher than that of VMMN. However, the difference is trivial.

## VI. Conclusion

A new hierarchical interconnection network, called **HMMN**, is proposed for the future generation MPC systems. The architecture of the HMMN, addressing of nodes, and routing of message have been discussed in detail. We have evaluated the static network performance of the HMMN, as well as that of several other networks. From the static network performance, it has been shown that the HMMN possesses several attractive features, including constant node degree, small diameter, low cost, small average distance, and better bisection width. We have seen that with the same degree, arc-connectivity, bisection width, and wiring complexity, the diameter and average distance of the HMMN is lower than that of the TESH network. Also, the diameter and average distance of the HMMN is lower than that of mesh network and slightly higher to that of HMMN. The HMMN yields better static network performance with reasonable cost, which are indispensable for next generation MPC systems.

This paper focused on the architectural structure and static network performance. Issues for future work include the following: (1) evaluation of static network performance considering different value of $m$, $L$, & $q$ and (2) evaluation of dynamic communication performance using dimension order routing algorithm.

## Acknowledgment

This work is partly supported by FRGS13-065-0306, Ministry of Education, Government of Malaysia. The authors would like to thank the anonymous reviewers for their constructive comments and suggestions on the paper which have helped to improve the quality of the paper.

## References

[1] B.S.P. Mishra and S. Dehuri, Parallel Computing Environments: A Review, *IETE Technical Review*, vol. 28, no. 3, pp. 240-247, 2011.

[2] http://top500.org/blog/lists/2013/06/press-release/

[3] V. Puente, J.A. Gregorio, R. Beivide and F. Vallejo, A Low Cost Fault Tolerant Packet Routing for Parallel Computers, *Proceedings of the 17$^{th}$ IEEE/ACM IPDPS*, April, 2000.

[4] P. Beckman, Looking Toward Exascale Computing, *Keynote Spech. International Conference PDCAT*, Dunedin, New Zealand, December 2008.

[5] Wolfgang E. Nagel, From TERA- to PETA- to EXA-Scale Computing: What does that mean for our Community?, *Keynote speech in the 10$^{th}$ IASTED Int'l Conf. PDCN*, Innsbruck, Austria, 2011.

[6] Y. Yang, A. Funahashi, Akiya Jouraku, H. Nishi, H. Amano, T. Sueyoshi, Recursive Diagonal Torus: An Interconnection Network for Massively Parallel Computers, *IEEE Transactions on Parallel and Distributed Systems*, vol. 12, no. 7, pp. 701715, 2001.

[7] M.M. Hafizur Rahman and S Horiguchi, Network Performance of Hierarchical Torus Network *Proc. of the 3$^{rd}$ PDCAT*, pp. 122–129, Kanazawa, Japan, 2002.

[8] M.M. Hafizur Rahman, X Jiang, MSA Masud, S Horiguchi, Network Performance of Pruned Hierarchical Torus Network *Proc. of the 6$^{th}$ IFIP Int'l Conference NPC*, pp. 9–15, Australia, 2009.

[9] M. Abd-El-Barr and T.F. Al-Somani, Topological Properties of Hierarchical Interconnection Networks: A Review and Comparison, *Journal of Electrical and Computer Engineering*, Hindawi Publishing Corporation, Vol. 2011, 12 pages.

[10] P.L. Lai, H.C. Hsu, C.H. Tsai, I.A. Stewart, A class of hierarchical graphs as topologies for interconnection networks, *Theoretical Computer Science, Elsevier*, Vo. 411, pp. 2912–2924, 2010.

[11] Youyao Liu, Cuijin Li, and Jungang Han, RTTM: A New Hierarchical Interconnection Network for Massively Parallel Computing, *Proc. of the HPCA*, LNCS 5938, pp. 264–271, 2010.

[12] Cristobal Camarero, Carmen Martnez, and Ramon Beivide, L-Networks: A Topological Model for Regular 2D Interconnection Networks, *IEEE Transactions on Computers*, vol.67, no.7, pp. 1362 – 1375, July, 2013.

[13] V.K. Jain, T. Ghirmai, and S. Horiguchi, "TESH: A new hierarchical interconnection network for massively parallel computing", *IEICE Trans. on Inf. & Syst.*, vol.E80-D, no.9, pp.837-846, 1997.

[14] M.M. Hafizur Rahman, Y. Inoguchi, Y. Sato, Y. Miura, and S. Horiguchi, "Dynamic Communication Performance of the TESH Network under Nonuniform Traffic Patterns", *Journal of Networks*, vol.4, no. 10, pp. 941 - 951, 2009.

[15] Faiz Al Faisal and M.M. Hafizur Rahman, "Symmetric Tori Connected Torus Network", *Proc. of the 12$^{th}$ ICCIT*, pp. 174-179, Dhaka, Bangladesh, 2009.

[16] V. Puente, C. Izu, J.A. Gregorio, R. Beivide, J. M. Prellezo and F. Vallejo, Improving parallel system performance by changing the arrangement of the network links, *Proceedings of the 14$^{th}$ International Conference on Supercomputing, 2000*, pp. 44 - 53, Santa Fe, New Mexico, USA, 2000.

[17] Md. Rabiul Awal, M.M. Hafizur Rahman, and M. A. H. Akhand, A New Hierarchical Interconnection Network for Future Generation Parallel Computer *Proc. of the 16$^{th}$ ICCIT*, Khulna, Bangladesh, 2013.

[18] J.M. Kumar and L.M. Patnaik, "Extended Hypercube: A Hierarchical Interconnection Network of Hypercube," IEEE Trans. on Parallel and Distributed Systems, vol. 3, no. 1, pp. 45-57, 1992.

# Power, Delay and Area Optimization in a Full-Adder Circuit using FinFETs

Imam-Uz-Zaman, Tonmoy Roy*, and M. S. Islam*
Bangladesh University of Engineering and Technology, Bangladesh
* E-mail: tonmoy_roy@live.com, islams@eee.buet.ac.bd

*Abstract*—**This paper proposes power, area and delay optimization for one bit full adder circuit using FinFETs. Double gate FinFET currently is the most attractive choice among the multi-gate transistor structure and has emerged as one of the most likely successors to the classical planar MOSFET. FinFET is more versatile than planar single-gate FET because it has two gates that can be biased to control the drive current and threshold voltage. By controlling the back gate static power, dynamic power, delay, area of a circuit can be improved. Moreover FinFETs' second gate can be utilized to implement circuit with fewer transistors which are important when area efficiency is necessary. This paper mainly deals with various logic design styles to obtain low leakage, minimum delay and area efficient circuits through the judicious use of FinFETs. Simulation shows various modes of circuits constructed using FinFETs have trade-offs in terms of leakage power, dynamic power, propagation delay and area.**

***Keywords—FinFET, VLSI, Full-Adder, Static Power.***

## I. INTRODUCTION

Continuous demand for high performance and smaller electronic equipment has forced the fabrication technology of CMOS to a significant reduction in the possible minimum feature size of the transistor to the nanometer regime. This reduction also scaled down other device parameters such as gate oxide thickness, threshold voltage and supply voltage to maintain stability rules [1]. The main barrier to this continuous reduction is the leakage phenomenon. This leakage becomes a major concern in the sub-32nm CMOS technology due to poor channel electrostatic potential which leads to a degraded short-channel behavior and high leakage current [2]. Recently double-gate FET specially, FinFET (Fig. 1) has become a suitable candidate for the future generation of CMOS technology.

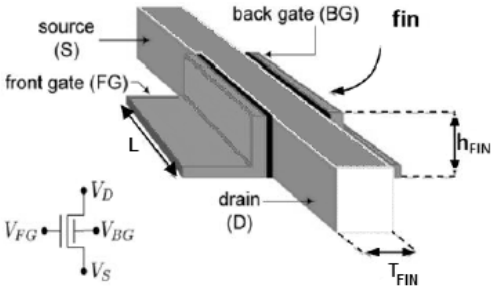



Figure 1. Four Terminal FinFET Device [4].

The motivation of this paper is to develop low leakage, low delay, area efficient Full-Adder circuits utilizing FinFETs. Design and simulation of different modes of FinFET circuits were conducted. Simulation of similar CMOS circuits was done for comparison as well.

The rest of the paper is organized as follows, a concise review of four terminal FinFET technology, design methodology and simulation results.

## II. THE FINFET TECHNOLOGY

Contrast to planar single and double gate device, the FinFET effective channel width is perpendicular to the surface. That's why increasing of effective channel width and on current is possible by varying fin height. FinFET transistors overcome the static power dissipation problem by providing effective control over the channel potential by using two gates wrapped around the fin [3]. A FinFET can be used to replace two parallel transistors with one if the gate signals of the two transistors are applied to it its two independent gates. This makes it useful for low power area efficient circuits. FinFET device parameters used throughout this paper are listed in Table I. For traditional bulk CMOS simulation, channel length of 32nm, channel width for the NMOS of 100nm and for 250nm for PMOS channel width was chosen. These parameters were selected to make the power and speed of CMOS comparable to the FinFET technology being used.

TABLE I. DEVICE PARAMETERS FOR THE FINFET MODEL

| Parameter | Value |
|---|---|
| Length of the channel (L) | 32nm |
| Thickness of front gate/back gate oxide | 1.4nm |
| Thickness of the fin ($T_{FIN}$) | 20nm |
| Height of the fin ($h_{FIN}$) | 15nm |
| Power supply (VDD) | 1.0V |
| Channel Doping | $2\times10^{16}cm^{-3}$ |

At the sub-32nm technology, it is evaluated that leakage power might account for as much as 40-50% of total power consumption in CMOS technology [5]. This problem can be significantly mitigated by the use of FinFET. Independent control of front gate and back gate of FinFET can be effectively used to reduce power consumption and to improve circuit performance. FinFET also provides much less sub-threshold leakage.

## III. Design Methodology

Four modes have been proposed for design of circuits using FinFETs [6]. All of these modes have been studied previously for simple logic circuits like NAND gates [7]. The first mode is the short gate (SG) mode. In this mode, the front gate and back gate of a device is shorted together. Hence, both the gates get biased to the same voltage. This gives better drive strength and better control over the channel [6]. The second mode is the independent gate (IG) mode [8] [9]. In this mode, the front and back gate of the device is driven from different sources and are biased at different voltages. In properly designed logic circuits this can decrease the required area considerably. The third mode is the low power (LP) mode [10]. In this mode, the back gates of all the n-channel transistors are connected to the ground and the back gates of all the p-channel transistors are connected to the logic high voltage. This increases the threshold voltage of the device and decreases leakage current and hence the amount of static power consumed gets decreased significantly. The final mode is the IG-LP mode which is the combination of the IG and LP modes. This mode offers low power circuits which also require fewer transistors as well.

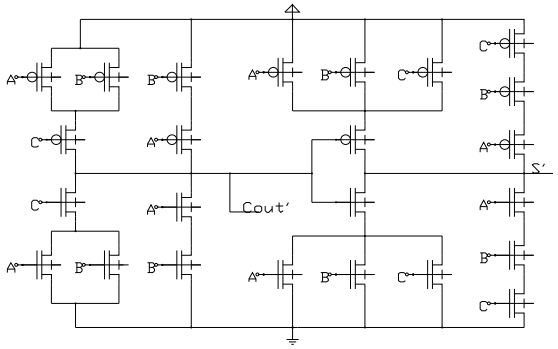



Figure 2. The 1-Bit Full Adder Short Gate (SG) Mode Design.

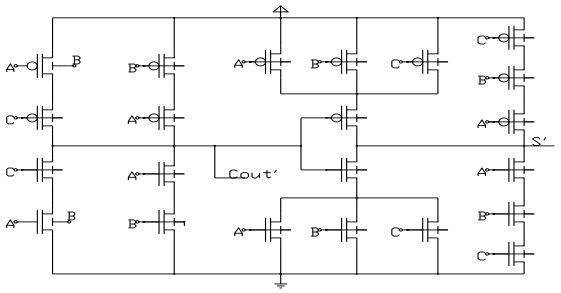



Figure 3. The 1-Bit Full Adder Independent Gate (IG) Mode Design.

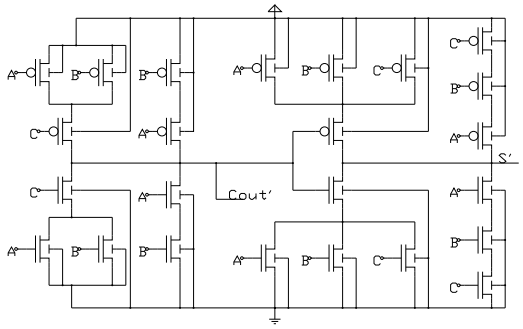



Figure 4. The 1-Bit Full Adder Low Power (LP) Mode Design

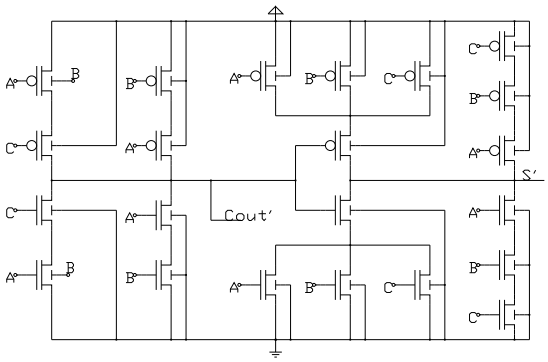



Figure 5. The 1-Bit Full Adder Independent Gate-Low Power (IG-LP) Mode Design.

TABLE II. Required Area for the Circuits of Different Modes

| Circuit Mode | Number of Transistors | Required Area ($m^2$) |
|---|---|---|
| Short Gate (SG) | 12 | $2.4192\times10^{-14}$ |
| Low Power (LP) | 12 | $2.4192\times10^{-14}$ |
| Independent Gate (IG) | 10 | $2.016\times10^{-14}$ |
| Independent Gate-Low Power (IG-LP) | 10 | $2.016\times10^{-14}$ |

In this paper we have shown design and simulation of a 1-bit Full Adder circuit utilizing FinFETs in the discussed four modes as well as with traditional CMOS technology for comparison. The full adder circuit used was first designed using FinFETs in SG mode. This is shown in Fig 2. Here the front and back gates of all the transistors are shorted, giving better drive current and consequently higher speeds which make it best suited for high performance operations. The full adder circuit using FinFETs in IG mode is shown in Fig 3. In

this circuit the front gate and back gates of the transistors are operated at different bias voltages and can be used for separate inputs. Every two traditional transistors connected in parallel are replaced by one FinFET with the inputs of the two transistors connected to the two gates of the FinFET. This reduces the number of transistors required and hence reduces the circuit area. Reduced number of transistors also means lower power consumption. However this mode suffers from long propagation delays for some input configurations. The LP mode of the circuit is shown in Fig 4. In this circuit all the back gates of the n-channel FinFETs are connected to the ground and the back gates of the n-channel FinFETs are connected to the logical high voltage. This configuration increases the threshold voltage of each transistor. Increased threshold voltage reduces leakage current and power consumption in static condition. However this mode suffers from increased propagation delay and does not provide any area benefit either. Finally, the IG-LP mode is shown in Fig 5. This circuit is a combination of both the IG and the LP mode. In this circuit all the parallel transistors are replaced by single IG mode transistor. On the other hand, all the other transistors are connected in LP mode with the n-channel transistors' back gates and the p-channel transistors' back gates connected to the ground and the logic high respectively. This circuit gives reduced area and consumes less power, but suffers from slightly increased propagation delays.

The effective channel width of the FinFET is quantized by the fin number. Our experiments showed that 2 fins for the p-channel transistor and 1 fin for n-channel transistor produces the minimum difference between rise time and fall time at the output for the circuits in consideration. This configuration was used in the simulation of the circuits.

Table II shows the chip area required for the full adder circuits operating in different modes. It can be seen that the short gate mode and the low power modes require more area. This is expected as they require more transistors. The independent gate mode and the independent gate low power modes on the other hand require relatively lower area as they are made with fewer transistors.

## IV. Simulation Results

In this section, the simulation results for power and circuit of the 1-bit full adder delay using FinFET is presented. The simulation was done on Synopsys HSPICE [11] using Arizona State University Predictive Technologies Model (ASU-PTM) Double Gate model [12]. The model parameters used in the simulation is given in Table I. The effective width per fin of the device was calculated from the fin width and fin height using the following equation [13]:

$$W_{eff} \cong 2 \times h_{FIN} + W_{FIN} \quad (1)$$

For CMOS simulation, Predictive Technologies Model [12] high-performance model (PTM-HP) was used with parameters taken from International Technology Roadmap for Semiconductors (ITRS) 2008 for 32nm Technology. All the aforementioned circuits were loaded with fan out of four inverters to simulate heavy load. For static power calculations, uniform distribution of the input vector was assumed. For dynamic power calculation, transient analysis was made with the inputs stimulated with signals of 1GHz frequency. The propagation delay was calculated from transient analysis of the circuits. The maximum of all the delays of all the inputs to output paths was taken into consideration.

TABLE III. Simulation Results for the Circuits of Different Modes

| Circuit Mode | Static Power (Watt) | Dynamic Power (Watt) | Propagation Delay (sec) |
|---|---|---|---|
| Short Gate (SG) FinFET | $7.55\times10^{-08}$ | $2.01\times10^{-05}$ | $3.53\times10^{-11}$ |
| Low Power (LP) FinFET | $6.63\times10^{-08}$ | $1.36\times10^{-05}$ | $5.28\times10^{-11}$ |
| Independent Gate (IG) FinFET | $6.73\times10^{-08}$ | $9.07\times10^{-06}$ | $4.25\times10^{-11}$ |
| Independent Gate-Low Power (IG-LP) FinFET | $6.01\times10^{-08}$ | $9.12\times10^{-06}$ | $4.16\times10^{-11}$ |
| CMOS | $1.10\times10^{-07}$ | $5.80\times10^{-05}$ | $3.99\times10^{-11}$ |

The results for static power, dynamic power and propagation delay obtained from the simulations are given in Table III. It is evident that the results are comparable to the

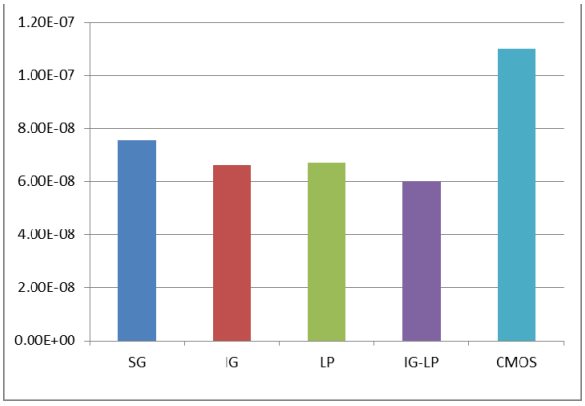



Figure 6. Static Power (W) for the Circuits of Different Modes

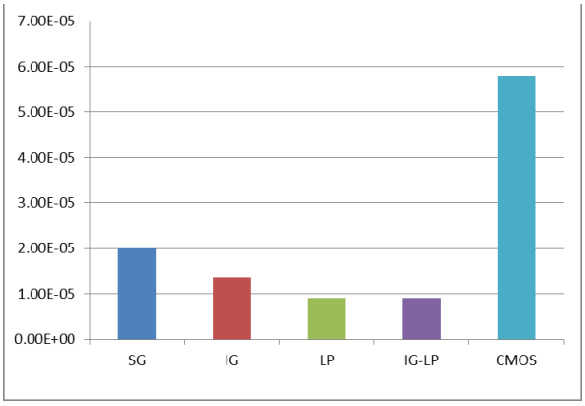



Figure 7. Dynamic Power (W) for the Circuits of Different Modes.

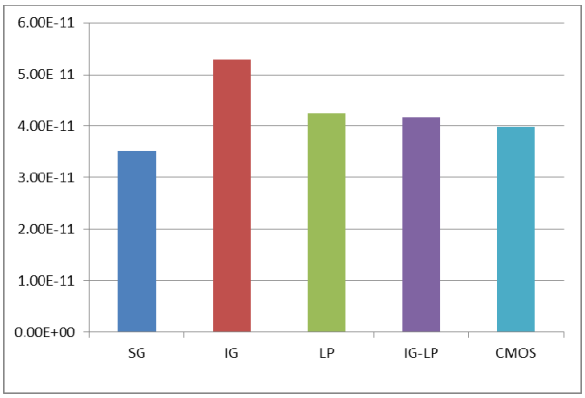



Figure 8. Propagation Delay (s) for the Circuits of Differnet Modes.

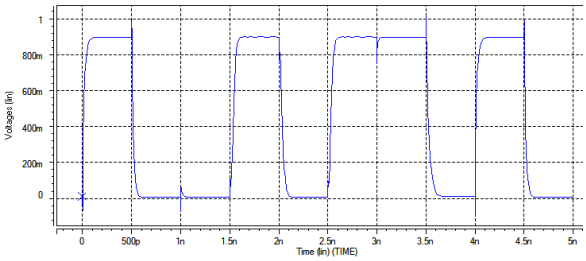



Figure 9. Sample Transient Response for the SG Mode Circuit.

expected values. While the SG mode circuit consumes the most static power, the IG-LP mode consumes the lowest static power (Fig. 6) due to lower leakage loss and a lower number of transistors. From the dynamic power data (Fig. 7), it can be observed that the SG mode consumes the most dynamic power as well while the LP and IG-LP modes consume the least amount of dynamic power. The IG-LP mode gives comparable propagation delay to other FinFET modes (Fig. 8). Normal CMOS circuits exhibit the highest static and dynamic powers (Figs. 6 and 7), as expected.

Fig 9 shows samples of the transient response used for propagation delay calculation. All the input to output paths are considered and the longest delays for each circuit are given in Table III. As expected, the SG mode wields the lowest delay because of the higher drive current. The delay for IG mode was found to be the worst among the FinFET circuits. The estimated area required for each circuit is given in Table I in the previous section. The SG and LP modes require more area as compare to the IG and IG-LP mode as they require more transistors. This suggests a trade-off of power, delay and area among these modes of the circuit.

## V. CONCLUSIONS

In this paper, we have showed the design and simulation of a Full Adder circuit using double gate (DG) FinFET device as well as traditional bulk CMOS for comparison. The back gates of the FinFETs were used to control the threshold voltage of the front gate. This ability to control the threshold voltage of the front gate dynamically provides the design opportunity to construct digital circuits in various modes to optimize for leakage, power, area or delay according to the needs. From the simulation results it was evident that the SG, while consuming the most power is the fastest among the modes studied. It was also observed that the LP and IG-LP modes consumed the least static and dynamic powers whereas the IG and IG-LP modes require fewer transistors, reducing the area requirement of the circuits. All the FinFET circuit modes were superior compared to the traditional bulk CMOS in terms of both static and dynamic power and the propagation delay was comparable to the delay of CMOS.

## REFERENCES

[1] S. Cayouette, *Static Power Dissipataion in Arithmetic Circuits: The Nanometer Domain*, Royal Military College of Canada, 2007.

[2] JBhoj, A.N.; Jha, N.K., "Design of ultra-low-leakage logic gates and flip-flops in high-performance FinFET technology," *Quality Electronic Design (ISQED), 2011 12th International Symposium on* , vol., no., pp.1,8, 14-16 March 2011.

[3] Agostinelli, M.; Alioto, M.; Esseni, D.; Selmi, L., "Leakage–Delay Tradeoff in FinFET Logic Circuits: A Comparative Analysis With Bulk Technology," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on* , vol.18, no.2, pp.232,245, Feb. 2010.

[4] Alioto, M., "Comparative Evaluation of Layout Density in 3T, 4T, and MT FinFET Standard Cells," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on* , vol.19, no.5, pp.751,762, May 2011.

[5] Weimin Zhang; Fossum, J.G.; Mathew, L.; Du, Y., "Physical insights regarding design and performance of independent-gate FinFETs," *Electron Devices, IEEE Transactions on* , vol.52, no.10, pp.2198,2206, Oct. 2005.

[6] Mishra, Prateek and Muttreja, Anish and Jha, Niraj K., "FinFET Circuit Design", *Nanoelectronic Circuit Design*, Springer New York, pp. 23-54, 2011.

[7] V Narendar, Wanjul Dattatray R, Sanjeev Rai and R A Mishra. Article: Design of High-performance Digital Logic Circuits based on FinFET Technology. *International Journal of Computer Applications* vol. 41, no. 20, pp. 40-44, March 2012.

[8] Muttreja, A.; Agarwal, N.; Jha, N.K., "CMOS logic design with independent-gate FinFETs," *Computer Design, 2007. ICCD 2007. 25th International Conference on* , vol., no., pp.560,567, 7-10 Oct. 2007.

[9] Rostami, M.; Mohanram, K., "Novel dual-Vth independent-gate FinFET circuits," *Design Automation Conference (ASP-DAC), 2010 15th Asia and South Pacific* , vol., no., pp.867,872, 18-21 Jan. 2010.

[10] Tawfik, S.A.; Kursun, V., "Characterization of New Static Independent-Gate-Biased FinFET Latches and Flip-Flops under Process Variations," *Quality Electronic Design, 2008. ISQED 2008. 9th International Symposium on* , vol., no., pp.311,316, 17-19 March 2008.

[11] Synopsys Hspice, reference, "http://www.synopsys.com"

[12] Arizona State University Predictive Technologies [Online].

[13] Shrivastava, M.; Baghini, M.S.; Sharma, D.K.; Ramgopal Rao, V., "A Novel Bottom Spacer FinFET Structure for Improved Short-Channel, Power-Delay, and Thermal Performance," *Electron Devices, IEEE Transactions on* , vol.57, no.6, pp.1287,1294, June 2010.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# VANET Topology Based Routing Protocols & Performance of AODV, DSR Routing Protocols in Random Waypoint Scenarios

Bijan Paul[1], Aditi Roy[2,*]
University of Asia Pacific, Bangladesh
Shahjalal University of Science & Technology, Bangladesh
[*] E-mail: bijan.paul@uap-bd.edu, aditiroy1710@gmail.com

*Abstract*—**Vehicular Ad-hoc Network is a new technology in the modern era and due to road accident daily occurrence which has taken enormous attention in the recent years. Because of rapid topology changing and frequent disconnection makes it difficult to design an efficient routing protocol for routing data among vehicles, called V2V or vehicle to vehicle communication and vehicle to road side infrastructure, called V2I.To design of an efficient routing protocol has taken significant attention because existing routing protocols for VANET are not efficient to meet every traffic scenarios. For this reason it is very necessary to identify which protocol is better. By using simulation of protocols we can understand existing routing protocols behavior. In this research paper, we focus on VANET topology based routing protocols and also measure the performance of two on-demand routing protocols AODV & DSR in random waypoint scenario.**

***Keywords—VANET, AODV, DSR, TCP, CBR, PDR, E-2-E Delay, LPR***

## I. INTRODUCTION

VANET (Vehicular adhoc network) is a special form of MANET which is an autonomous & self-organizing wireless communication network, where nodes in VANET involve themselves as servers and/or clients for exchanging & sharing information. Due to new technology government has taken huge attention on it. There are many research projects around the world which are related with VANET such as COMCAR [1], DRIVE [2], FleetNet [3] and NoW (Network on Wheels) [4], CarTALK 2000 [5], CarNet [6]. There are several VANET applications such as Vehicle collision warning, Security distance warning, Driver assistance, Cooperative driving, Cooperative cruise control, Dissemination of road information, Internet access, Map location, Automatic parking, and Driverless vehicles.

In this paper, we mainly focus on VANET topology based routing protocols and we also have evaluated performance of AODV and DSR based on random waypoint model. The remainder of the paper is organized as follows: Section 2 describes the VANET characteristics. Section 3 describes shortly about VANET topology based routing protocols. Section 4 discusses briefly about two on demand routing protocols AODV and DSR procedure. Section 5 describes connection types like TCP and CBR. Section 6 presents performance metrics and the network parameters. Section 7 presents our implementation. Section 8 presents experimental analysis. Section 9 presents our decisions. We conclude in Section 10 and with the references at the end.

## II. VANET TOPOLOGY BASED ROUTING PROTOCOLS

Topology based routing protocols use link's information within the network to send the data packets from source to destination. Topology based routing approach can be further categorized into proactive (table-driven) and reactive (on-demand) routing. Figure1 shows a summary of unicast VANEL Routing protocols:

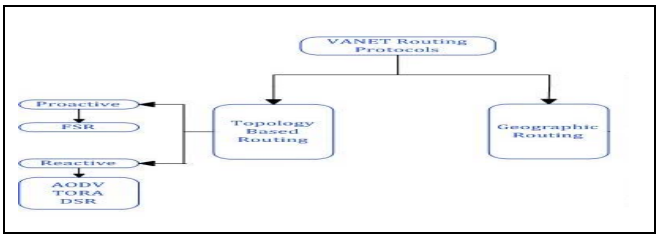



Figure 1. Unicast routing protocols in VANET

### A. Proactive (table-driven)

Proactive routing protocols are mostly based on shortest path algorithms. They keep information of all connected nodes in form of tables because these protocols are table based.

*1) Fisheye State Routing*

FSR [7] is a proactive or table driven routing protocol where the information of every node collects from the neighboring nodes. Then calculate the routing table. It is based on the link state routing & an improvement of Global State Routing.

### B. Reactive (On Demand)

Reactive routing protocol is called on demand routing because it starts route discovery when a node needs to communicate with another node thus it reduces network traffic.

*1) AODV*

Ad Hoc On Demand Distance Vector routing protocol [8] is a reactive routing protocol which establish a route when a node requires to send data packets.

*2) DSR*

The Dynamic Source Routing (DSR) protocol presented in [9] which utilize source routing & maintain active routes. It has two phases route discovery & route maintenance.

*3) Temporally Ordered Routing Protocol (TORA)*

Temporally Ordered Routing Protocol [10] is based on the link reversal algorithm that creates a direct acyclic graph towards the destination where source node acts as a root of the tree.

## III. Procedure of AODV & DSR Protocol

A routing protocol is necessary to forward a packet from source node to destination in adhoc network. Thus for packet forwarding numerous routing protocols have been proposed. In this paper we focus on two important protocols AODV [8] & DSR [9]. AODV & DSR are reactive routing protocols which are also called on demand routing protocol because routes are created only when source node wants to send packet to destination.

### *A. AODV*

AODV routing protocol mechanism is that if a source node wants to send a packet to a destination node at first the entries in routing table are checked to confirm that whether current route exists to that destination node or not. If exists the data packet will forward to the next hop otherwise the route discovery process is initiated. For Route discovery process using Route Request (RREQ) & Route Reply (RREP). The route is established when receiving RREP message from destination to source. This procedure is called backward learning. Figure 2 and Figure 3 shows the RREQ and RREP for the AODV route discovery.

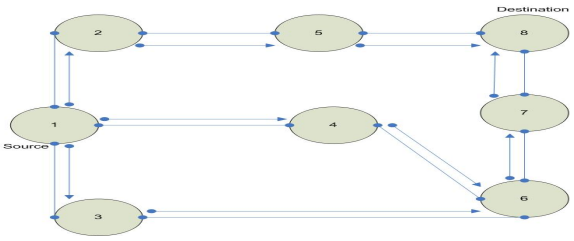



Figure 2. Propagation of the RREQ for AODV route discovery

### *B. DSR*

The Dynamic Source Routing (DSR) [9] protocol has two phases route discovery & route maintenance. It does not use periodic routing message. All the intermediate nodes ID are stored in the packet header of DSR that it has traversed. Thus when destination occurs then from the query packet it retrieves the entire path information which is used to respond to the source.

As a result, establish a path from source node to the destination node. If there has multiple paths to go to the destination DSR stores multiple path of its routing information. If any route breaks occur in that case an alternative route can be used in DSR. It will generate an error message if there is any link failure. DSR route discovery process shown in Figure 4.

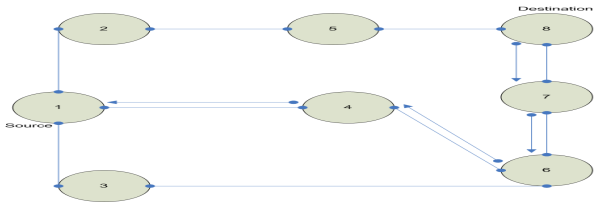



Figure 3. Path of the RREP to the Source for AODV route discovery

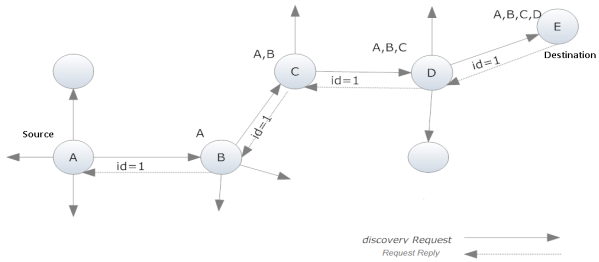



Figure 4. DSR Route Discovery

## IV. Connection Types

There are several types of connection pattern in VANET. For our simulation purpose we have used CBR and TCP connection pattern.

### *A. Constant Bit Rate (CBR)*

In CBR, data packets are sent with fixed size and fixed interval between each data packets. Establishment phase of connection between nodes is not required here.

### *B. Transmission Control Protocol (TCP)*

TCP is a connection oriented and reliable transport protocol. To ensure reliable data transfer TCP uses acknowledgement, time outs and retransmission.

## V. Performance Metrics & Network Parameters

For network simulation, there are several performance metrics which is used to evaluate the performance. In our simulation purpose we have used the below three performance metrics equations.

**i. Packet Delivery Ratio (PDR)** = ***(no of received packets / no of sent packets) * 100***

**ii. Average end-to-end delay (E-2-E Delay)** = ***(time packet received - time packet sent) / total no of packets received * 100***

**iii. Loss Packet Ratio (LPR)** = ***(no of sent packets – no of received packets) / no of sent packets * 100***

## VI. Our Implementation

For implementing our simulation we used Network Simulator NS-2.34 [11, 12]. Also we used random waypoint mobility model for our simulation. To measure the performance of AODV and DSR same scenario is used for both the protocols.

### *A. Simulation Parameters*

| Parameter | Value |
|---|---|
| Protocols | AODV, DSR |
| Simulation Time | 200 s |
| Number of Nodes | 30, 90, 150 |
| Simulation Area | 840 m x 840 m |

| | |
|---|---|
| Speed Time | 5 , 10 , 15 , 20 , 25 m / s |
| Pause Time | 50 ,100, 150, 200, 250 s |
| Traffic Type | CBR , TCP |
| Mobility Model | Random Waypoint |
| Network Simulator | NS 2.34 |

### *B. Performance Measure Script*

Generally, in NS-2 [12] when we execute a program there creates two types of file trace file and nam file where nam file is used to visualize the simulation and trace file keep records of various interesting quantities such as each individual packets as its arrives, departs or is dropped at a link or queue by which we can measure a protocol performance.

Awk script is required to analysis trace file for performance measure. To measure packet delivery ratio, loss packet ratio & average end-to-end delay of AODV and DSR we make two awk scripts. The scripts sudo codes are given below.

**PDR and LPR Measure AWK Script**

```
START
//initialization
SET nSentPackets to 0
SET nReceivedPackets to 0
IF $1 =  "s" AND  $4 = "AGT" THEN
   INCREMENT nSentPackets
ENDIF
IF $1 =  "r" AND  $4 = "AGT" THEN
   INCREMENT nReceivedPackets
ENDIF
COMPUTE rPacketDeliveryRatio as nReceivedPackets / nSentPackets * 100
COMPUTE  lpr as ((nSentPackets-nReceivedPackets) /  nSentPackets) * 100
PRINT  nSentPackets
PRINT  nReceivedPackets
PRINT  rPacketDeliveryRatio
PRINT  lpr
END
```

**END-TO-END DELAY AWK Script**

```
START
//initialization
SET seqno to -1
SET count to 0
IF $4 = "AGT" AND $1 = "s" AND seqno < $6 THEN
   COMPUTE seqno as $6
ENDIF
IF $4 = "AGT" AND  $1 == "s" THEN
   COMPUTE  start_time[$6] as $2
ELSE IF $7 = "tcp" AND $1 = "r" THEN
   COMPUTE  end_time[$6]  as $2
ELSE IF $1 = "D" AND $7 = "tcp" THEN
   COMPUTE  end_time[$6] as -1
ENDIF
FOR   X = 1 to seqno
  IF  end_time[X] > 0 THEN
   COMPUTE delay[X] as end_time[X] - start_time[X]
          INCREMENT count
  ELSE
    COMPUTE  delay[i] as  -1
       ENDIF
END FOR
FOR   X = 1 to seqno
 IF  delay[X] > 0 THEN
  COMPUTE n_to_n_delay as n_to_n_delay + delay[X]
 ENDIF
END FOR
COMPUTE   n_to_n_delay as  n_to_n_delay / count * 1000
PRINT   n_to_n_delay
END
```

### *C. Standard for Analysis*

For analyzing our experiment we define a standard for simulation results. We consider 30 nodes as low density, 90 nodes as average density and 150 nodes as high density. We also consider 5 m/s as low speed, 15 m/s as average speed and 25 m/s as high speed.

**The standard for PDR values (approx.) defines below**

**For speed & pause time:** High: >=98%,  Average: 96% to 97%, Low: <=95%

**The standard for E-2-E values (approx.) defines below**

**For pause time:** High: >=351ms, Average: 151ms to 350ms, Low: <=150ms

**For speed time:** High: >=150%, Average: 51% to 150%, Low: <=50%

**The standard for LPR values (approx.) define below**

**For pause time:** High: > 2%, Average: 1% to 2%, Low: < 1%

**For speed time:** High: > 3%, Average: 1.5% to 3%, Low: < 1.5%

### *D. Simulation Results*

The performance of AODV & DSR has been analyzed with varying pause time 50s to 250s and speed time 5 to 25 m/s for number of nodes 30, 90, 150 under TCP & CBR connection. We measure the packet delivery ratio, loss packet ratio & average end-to-end delay of AODV and DSR. Based on the simulation result we have generated the graph which shows the differences between AODV and DSR. The graphs are given below.

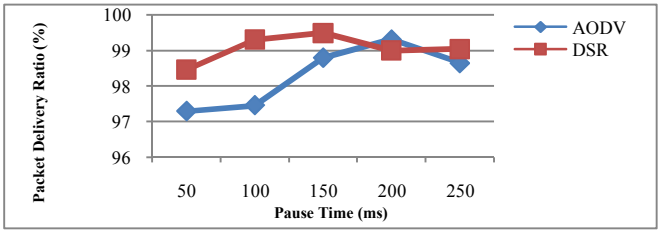



Figure 5. PDR (w.r.t. Pause) of 30 nodes using TCP

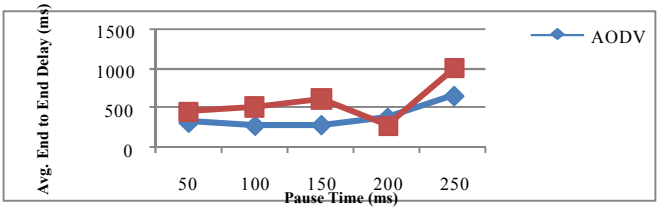



Figure 6. Avg.E2E delay (w.r.t. Pause) of 30 nodes using TCP

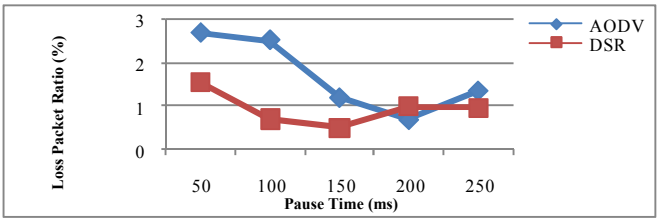



Figure 7. LPR (w.r.t. Pause) of 30 nodes using TCP

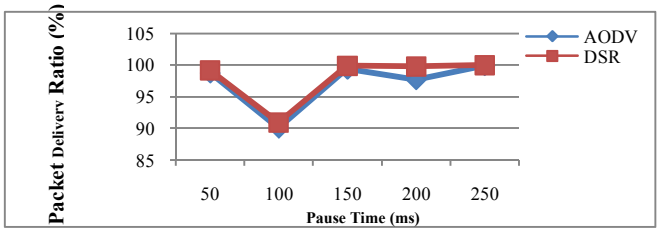



Figure 8. PDR (w.r.t. Pause) of 30 nodes using CBR

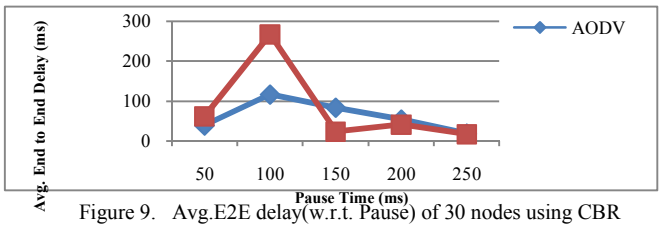



Figure 9. Avg.E2E delay(w.r.t. Pause) of 30 nodes using CBR

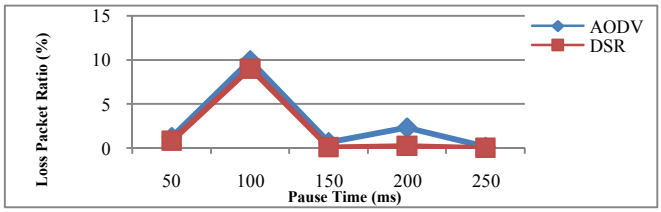



Figure 10. LPR (w.r.t. Pause) of 30 nodes using CBR

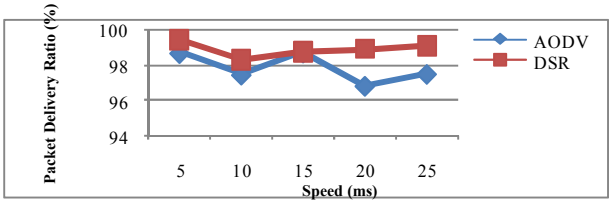



Figure 11. PDR (w.r.t. Speed) of 30 nodes using TCP

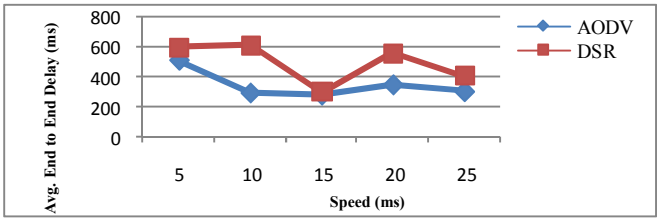



Figure 12. Avg.E2E delay (w.r.t. Speed) of 30 nodes using TCP

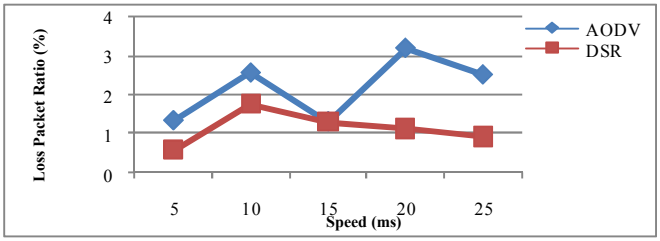



Figure 13. LPR (w.r.t. Speed) of 30 nodes using TCP

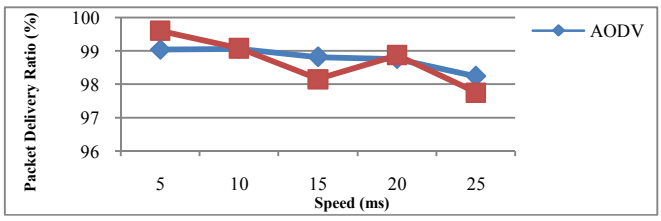



Figure 14. PDR (w.r.t. Speed) of 30 nodes using CBR

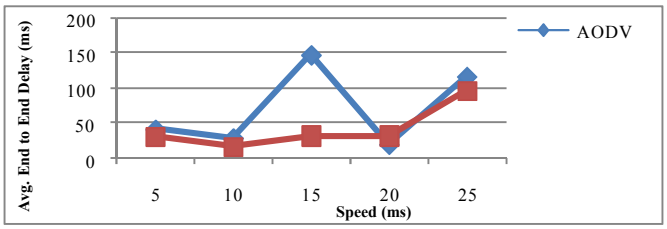



Figure 15. Avg.E2E delay (w.r.t. Speed) of 30 nodes using CBR

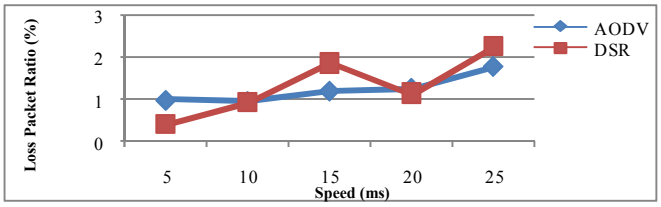



Figure 16. LPR (w.r.t. Speed) of 30 nodes using CBR

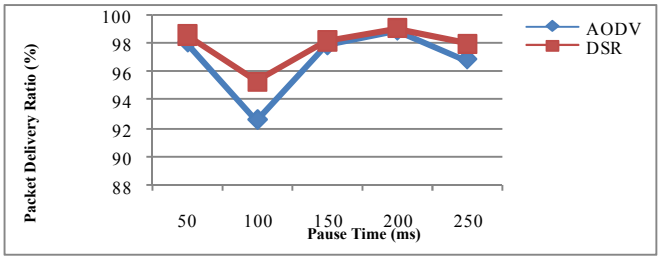



Figure 17. PDR (w.r.t. Pause) of 150 nodes using TCP

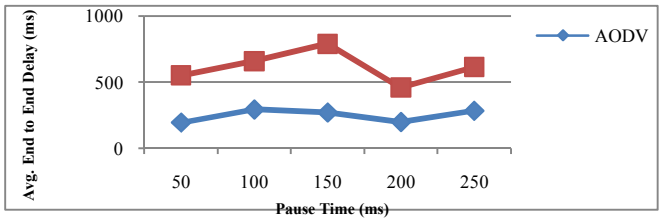



Figure 18. Avg.E2E delay (w.r.t. Pause) of 150 nodes using TCP

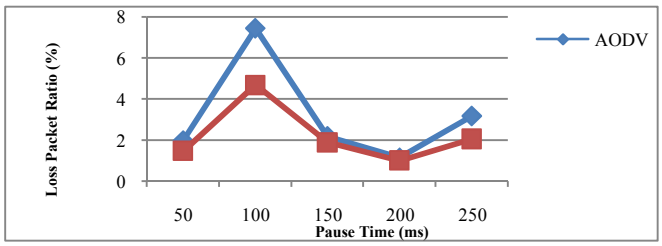



Figure 19. LPR (w.r.t. Pause) of 150 nodes using TCP

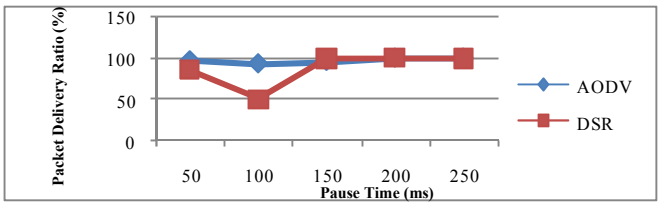



Figure 20. PDR (w.r.t. Pause) of 150 nodes using CBR

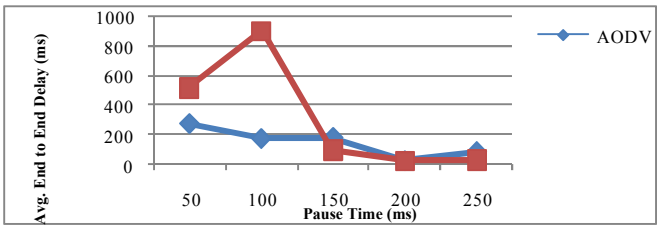



Figure 21. Avg.E2E delay (w.r.t. Pause) of 150 nodes using CBR

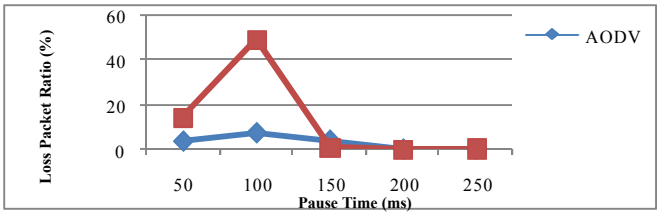



Figure 22. LPR (w.r.t. Pause) of 150 nodes using CBR

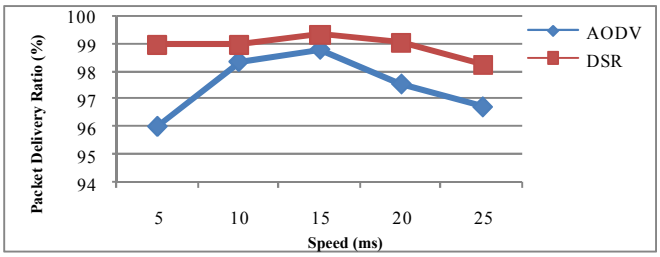



Figure 23. PDR (w.r.t. Speed) of 150 nodes using TCP

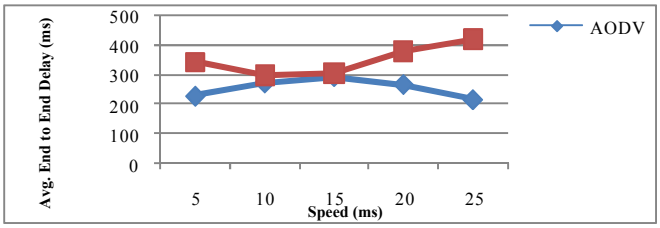



Figure 24. Avg.E2E delay (w.r.t. Speed) of 150 nodes using TCP

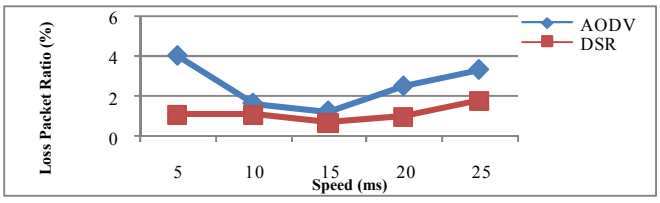



Figure 25. LPR (w.r.t. Speed) of 150 nodes using TCP

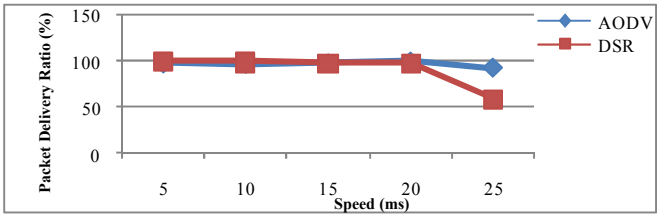



Figure 26. PDR (w.r.t. Speed) of 150 nodes using CBR

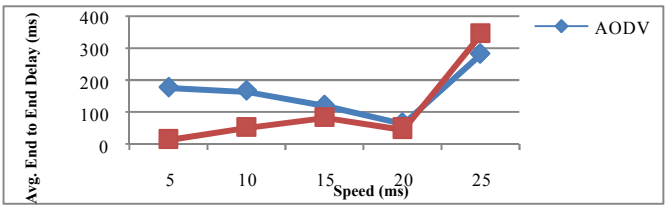



Figure 27. Avg.E2E delay (w.r.t. Speed) of 150 nodes using CBR

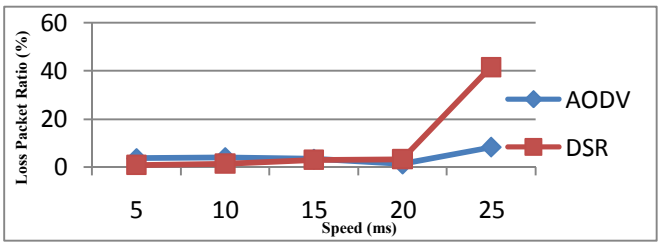



Figure 28. LPR (w.r.t. Speed) of 150 nodes using CBR

## VII. Experimental Analysis

Our simulation area considered is 840 × 840 and simulation run time is 200 seconds. Speed has been varied from 5m/s to 25 m/s. Pause time has been varied from 50s to 250s. Based on our standard in section 7.2 we can summarize the following differences between AODV and DSR based on our estimated parameters.

### *A. Pattern analysis of 30 nodes using TCP connection*

From our experimental analysis we observe that for TCP connection using pause time as a parameter in low mobility low pause time the packet delivery ratio (PDR) is average for AODV and high for DSR. In that scenario average end to end delay (E-2-E Delay) is average for AODV and high for DSR. The loss packet ratio for TCP connection is high for AODV and average for DSR. If the pause time is high the PDR for both routing protocols is high and E-2-E Delay for both protocols is high. LPR of DSR is low but for AODV it is average.

On the other hand, using speed as a parameter in low mobility low speed the packet delivery ratio for both protocols is high. In that scenario average E-2-E Delay is high, the loss packet ratio is low for both routing protocol. But in low mobility high speed, the PDR for AODV is average but high for DSR. E-2-E Delay for both protocols is high. LPR of AODV is average. But for DSR it is low.

### *B. Pattern analysis of 30 nodes using CBR connection*

We observe that for CBR connection using pause time as a parameter in low mobility low pause time the PDR of CBR for both routing protocols is high. In that scenario average E-2-E Delay is low for both protocols .The loss packet ratio is average for AODV and low for DSR. If the pause time is high the PDR for both routing protocols is high and E-2-E Delay is low for both routing protocols. LPR of DSR is low but for AODV it is low. On the other hand, using speed as a parameter in low mobility low speed the packet delivery ratio for both protocols is high. In that scenario average E-2-E

Delay and the loss packet ratio is low for both routing protocol. But in low mobility high speed, the PDR for AODV is high but average for DSR. E-To-E for both protocols is low. LPR is average for both routing protocols.

### C. *Pattern analysis of 150 nodes using TCP connection*

Pause time as a parameter in high mobility low pause time PDR for both protocols is high. In that scenario average E-2-E Delay is average for AODV and high for DSR. The LPR is average for both protocols. If the pause time is high the PDR for both routing protocols is average and E-2-E Delay is average for AODV and high for DSR. LPR is high for AODV and DSR.

On the other hand, using speed as a parameter in high mobility low speed, PDR of AODV is average but high for DSR. Though, E-2-E for AODV & DSR is high. LPR is low for DSR and high for AODV. If the speed is high AODV performs average and DSR performs high. E-2-E is high for both routing protocol. LPR of AODV is high but for DSR it is average.

### D. *Pattern analysis of 150 nodes using CBR connection*

We observe that for CBR connection using pause time as a parameter in high mobility low pause time the PDR of CBR it is average for AODV and low for DSR. E-2-E for AODV is average but it is high for DSR. The loss packet ratio is high for both protocols. If the pause time is high the PDR for AODV and DSR using CBR is high. E-2-E and LPR is low for both routing protocols.

On the other hand, using speed as a parameter in high mobility low speed the packet delivery ratio for AODV is average but high for DSR, Though E-2-E and LPR for AODV is high but low for DSR. If the speed is high the PDR for AODV and DSR is low. E-2-E is high for both routing protocol. LPR of AODV and DSR is high for CBR connection.

## VIII. OUR DECISIONS

After analyzing the performance of AODV & DSR we now summarize our decisions in the below decision tables based on our standard defined in section VII(C). In figure29 we have shown the performance of PDR, E-2-E and LPR for TCP and CBR connections where speed and node density is varied. In figure30 we have shown the performance of PDR, E-2-E and LPR for TCP and CBR connections where pause time and node density is varied.

## IX. CONCLUSION

In this paper we mainly focus on VANET topology based routing protocols. At first, we describe about VANET topology based routing protocols. We choose two on demand routing protocols AODV & DSR on the basis of packet delivery ratio, average End-to-End delay and Loss packet ratio for analysis their performance. We analysis the performance of AODV & DSR using decision Tables & observe that the performance of AODV and DSR depends based on scenarios. For further development of these protocols the performance evaluation should shed some light in near future.

| Nodes Density | Packet Delivery Ratio | | | | Avg.End to End Delay | | | | Loss Packet Ratio | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | TCP | | CBR | | TCP | | CBR | | TCP | | CBR | |
| **Low Density** | AODV | DSR | AODV | DSR | AODV | DSR | AODV | DSR | AODV | DSR | AODV | DSR |
| Low Speed | High | High | High | High | High | High | Low | Low | Low | Low | Low | Low |
| Avg.Speed | High | High | High | High | High | High | Avg | High | Low | Low | Low | Avg |
| High Speed | Avg | High | High | Avg | High | High | Avg | Avg | Avg | Low | Avg | Avg |
| **Avg. Density** | | | | | | | | | | | | |
| Low Speed | High | High | High | High | High | High | Avg | Low | Avg | Low | Low | Low |
| Avg.Speed | Avg | Avg | Low | Avg | High | High | Avg | High | High | Avg | High | Avg |
| High Speed | Avg | Avg | Avg | Low | High | High | Avg | High | High | Avg | High | High |
| **High Density** | | | | | | | | | | | | |
| Low Speed | Avg | High | Avg | High | High | High | High | Low | High | Low | High | Low |
| Avg.Speed | High | High | Avg | Avg | High | High | Avg | Low | Low | Low | High | Avg |
| High Speed | Avg | High | Low | Low | High | High | High | High | High | Avg | High | High |

Figure 29. PDR, E-2-E and LPR for TCP & CBR connections where varying speed & node density

| Nodes Density | Packet Delivery Ratio | | | | Avg.End to End Delay | | | | Loss Packet Ratio | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | TCP | | CBR | | TCP | | CBR | | TCP | | CBR | |
| **Low Density** | AODV | DSR | AODV | DSR | AODV | DSR | AODV | DSR | AODV | DSR | AODV | DSR |
| Low Pause Time | Avg | High | High | High | Avg | High | Low | Low | High | Avg | Avg | Low |
| Avg Pause Time | High | High | High | High | Avg | High | Low | Low | Avg | Low | Low | Low |
| High Pause Time | High | High | High | High | High | High | Low | Low | Avg | Low | Low | Low |
| **Avg. Density** | | | | | | | | | | | | |
| Low Pause Time | Avg | Avg | Avg | High | High | High | Low | Low | High | High | High | Avg |
| Avg Pause Time | Avg | High | High | Avg | Avg | High | Low | Low | High | Avg | Low | Avg |
| High Pause Time | Low | High | High | High | Avg | High | Low | Low | High | Avg | Low | Low |
| **High Density** | | | | | | | | | | | | |
| Low Pause Time | High | High | Avg | Low | Avg | High | Avg | High | Avg | Avg | High | High |
| Avg Pause Time | Avg | High | Avg | High | Avg | High | Avg | Low | High | Avg | High | Low |
| High Pause Time | Avg | Avg | High | High | Avg | High | Low | Low | High | High | Low | Low |

Figure 30. PDR, E-2-E and LPR for TCP & CBR connections where varying pause time & node density

## REFERENCES


[1] Ericson, "Communication and Mobility by Cellular Advanced Radio", ComCar project, www.comcar.de, 2002.

[2] Online, http://www.ist-drive.org/index2.html.

[3] W. Franz, H. Hartenstein, and M. Mauve, Eds., *Inter-Vehicle-Communications Based on Ad Hoc Networking Principles-The Fleet Net Project.* Karlshue, Germany: Universitatverlag Karlsuhe,November 2005.

[4] A. Festag, et. al., "NoW-Network on Wheels: Project Objectives,Technology and Achievements", Proceedings of 6th InternationalWorkshop on Intelligent Transportations (WIT), Hamburg, Germany,March 2008.

[5] Reichardt D., Miglietta M., Moretti L., Morsink P., and Schulz W.,"CarTALK 2000 – safe and comfortable driving based upon inter-vehicle-communication," in Proc. IEEE IV'02.

[6] Morris R., Jannotti J., Kaashoek F., Li J., Decouto D., "CarNet: A scalable ad hoc wireless network system," 9th ACM SIGOPS European Workshop, Kolding, Denmark, Sept. 2000.

[7] Pei, G., Gerla, M., and Chen, T.-W. (2000), "Fisheye State Routing: A Routing Scheme for Ad Hoc Wireless Networks," Proc. ICC 2000, New Orleans, LA, June 2000.

[8] Perkins, C.; Belding-Royer, E.; Das, S. (July 2003)"Ad hoc On-Demand Distance Vector (AODV) Routing".

[9] Johnson, D. B. and Maltz, D. A. (1996), "Dynamic Source Routing in Ad Hoc Wireless Networks," Mobile Computing, T. Imielinski and H. Korth, Eds., Ch. 5, Kluwer, 1996, pp. 153–81.

[10] Park, V.D., Corson, M.S. (1997), "A highly adaptive distributed routing algorithm for mobile wireless networks," INFOCOM '97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE , vol.3, no., pp.1405-1413 vol.3, 7-12 Apr 1997.

[11] http://www.isi.edu/nsnam/ns/

[12] http://www.isi.edu/nsnam/ns/tutorial/

# Resource Allocation in Hybrid Access Control Femtocell Network Targeting Inter-cell Interference Reduction

Afaz Uddin Ahmed[1,2,*], Fayeem Bin Aziz[1], Taufiq Mahmud Masum[1], Md. Sorwar Jahan[1], Mohammad Mahbubur Rahman[1], C.M.F.S. Reza[1], Abdullah Al Mamoon[1], Papan Day[1] and M. R. Zaman[2]
[1]Chittagong University of Engineering and Technology, Bangladesh
[2]Universiti Kebangsaan Malaysia, Malaysia
[*] E-mail: afazbd@gmail.com

*Abstract*— **Inter-cell interference is the most challenging issue in femtocell deployment within the coverage of Macro base station (MBS). In this paper, we have explored the role of femtocell in LTE network and proposed a Dynamic Resource allocation management algorithm (DRAMA) for hybrid access control in spectrum shared OFDMA network to optimize the interference and increase the quality of service (QoS). The user gets the privilege to assign the femtocell service level ensuring a minimum level of QoS. A simulation setting is developed to study the performance of DRAMA in hybrid access control femtocell network and compared it to closed and open access control.**

***Keywords—Femtocell network, inter-cell interference, co- channel interference, resource allocation, interference avoidance.***

## I. INTRODUCTION

Femtocell, which is also known as femtocell access point (FAP), is a promising technology that fills up the coverage and capacity gaps in a heterogeneous network. As the growth rate of the internet users increasing rapidly, new technologies that requires huge amount of traffic (video sharing, video conference, online storage, gaming, medical imaging etc.) are on peak demands. The number of mobile phone users is increasing rapidly and it is anticipated that by the end of 2013 the number of mobile phone users worldwide will be 5.63 billion. The data traffic in mobile phones during the years 2008-2012 had increased by a factor of 30 while at the same time the numbers internet users has grown up to 1.6 billion [1]. In figure 1, a chart of Voice, Data, Average Revenue per User (ARPU), and Mobile subscriber growth-rate is illustrated from the duration 2007 to 2013.

Femtocell makes it feasible to extend the radio properties of the macro cellular network while exploiting on the terrestrial broadband backhaul system. They are the customized small-scale mobile phone access points that customers can conveniently install inside their residence and office. Femtocell can offload large quantity of traffic (up to 80%) from macrocell's [3]. It can also act as a true enabler for high

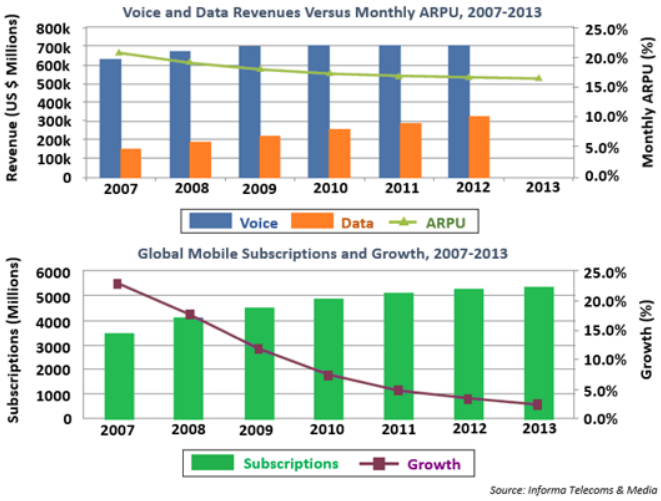



Figure 1. Global internet and mobile growth usage [2]

efficiency since 1957, due to the spatial reuse provided by reducing cell sizes and transmission distances. With further reduction in cell size, and additional isolation arising from in-building operation in many femtocell scenarios, this trend is expected to continue, with great benefits for high capacity services such as multimedia.

Due to dense deployment of FAP, interference among the existing macro-channels is a regular alarming issue regarding the performance of FAP. There are two types of interference scenario in two-tier networks. Co-tier interference occurs between the same elements of the network e.g. interference among femtocells. Cross-tier interference occurs between different elements of the network e.g. interference between macrocell and femtocell. The major interference occurrence scenarios in a two-tier femtocell network are listed below.

- Femtocells interfering with Macrocell on the same spectrum.
- Macrocell interfering with femtocells on the same spectrum.
- Neighboring femtocells that are close, interfering with one another.

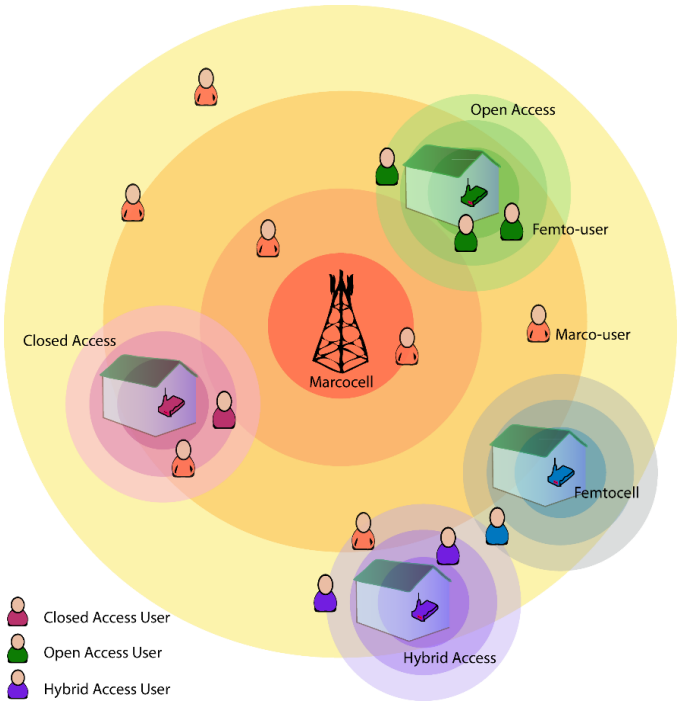



Figure 2. FAP network with different access method

- User hand-equipment transmitting with a sufficiently high power may reach the macrocell that will have an impact on the level of noises received by the macrocell.

The access control mechanism for FAP is one of the most effective methods for reducing cross-tier interferences. The performance of FAP greatly depends upon the mechanism that decides the user connection with a particular cell. Open access allows all the subscribers to connect the network within the range of a FAP. In close access, FAP denies the access to the subscriber who does not belong to a particular group even if it provides the best signal quality. Hybrid access reaches a compromise between the impact on the performance of subscribers and the level of access granted to non-subscribers allowing them to access to a limited amount of FAP resources [5]. Figure 2 shows briefly the access control mechanism in existing femtocell network.

Resource partitioning hybrid access mechanism ensures a quality service to the high priority subscribers and utilize the rest for unregistered subscribers. Recent studies have considered the performance of closed access [4-6], open access [7-9] and resource allocation management [10-12] in a heterogeneous network. In [13] a resource allocated OFDMA femtocell network is designed considering instantaneous FUEs power control for a target SINR. Joint resource allocation and admission control design for dynamic frequency sharing was discussed in [13-16]. In this paper, a dynamic resource allocation scheme- DRAMA is proposed that serves two different level of users on priority basis to limit the cross-tier interference in the two-tier network. It also ensures a higher network performance by offloading the excess macrocell traffic.

## II. System Analysis and Modelling

OFDMA in the downlink is robust against multi-path interference and frequency selectivity. It facilitates frequency domain scheduling, co-channel deployment and advanced MIMO techniques [17]. Within the coverage of the FAP, the user gives the highest priority to that FAP for better signal quality. MBS is relaxed to execute mobility towards FAP to reduce user congestion and signaling overhead. When FAP does not allow access, than the user tries to stay connected to the MBS to avoid call drop or network disconnect. In this paper, a spectrum sharing two-tier OFDMA downlink network is considered in an area of 400m x 400m consisting of one MBS in the middle. Random number of FAPs are distributed in this area using homogeneous Poisson Point Distribution (PPP). FAPs are in the middle of a rectangular house. We assumed that the FAPs do not overlap with each other or with the MBS. During the selection, users are considered to be at least 1 m apart from the MBS and the FAPs.

The path-loss among the users and the cells are calculated by using 3GPP-LTE standard HATA small/large city model [18]. For outdoor users:

$$\begin{aligned} PL_{hata}(dB) &= 69.55 + 26.16\log_{10}(f_c[MHz]) \\ &-13.82\log_{10}(h_{MBS}) - [1.1\log_{10}(f_c[MHz]) - 0.7]h_m \\ &-[1.56\log_{10}(f_c[MHz]) - 0.8] \\ &+[44.9 - 6.55\log_{10}(h_{MBS})]\log_{10} d_{Km} \end{aligned} \quad (1)$$

where $f_c$, $h_{MBS}$, $h_m$ and $d_{Km}$ are frequency in MHz, MBS antenna height, mobile equipment height and distance between MBS and mobile equipment in Km, respectively.

Indoor path-loss scenarios for both line-of-sight (LOS) and non-LOS, WINNER II channel models is considered [19].

$$PL_{LOS}(dB) = 18.7\log_{10} d_m + 46.8 + 20\log_{10}\left(\frac{f_c[GHz]}{5.0}\right) \quad (2)$$

$$PL_{NLOS}(dB) = 20\log_{10} d_m + 46.4 + 20\log_{10}\left(\frac{f_c[GHz]}{5.0}\right) + wL_w \quad (3)$$

where $f_c$, $d_m$, $w$, $L_w$ are frequency in GHz, distance between FAP and user, number of penetrated walls and penetration loss for each wall, respectively.

For $m$ number of macro users (outdoor users) and $k$ number of sub-carriers, the expression of SINR is:

$$SINR_{m,k} = \frac{P_{M,k}G_{m,M,k}}{N_0\Delta f + \sum_F P_{F,k}G_{m,F,k}} \quad (4)$$

where $N_0$, $\Delta f$, $P_{F,k}$, $P_{M,k}$, $G_{m,M,k}$ and $G_{m,F,k}$ are white noise power spectrum density, sub-carrier spacing, gain (depends on path-loss), transmitting power of neighboring FAP, transmitting power of serving MBS, channel gain between the macro user and their serving MBS on subcarrier $k$ and channel of macro user and neighboring FAP on sub-carrier $k$, respectively [20].

For $f$ number of femto-users and $k$ number of sub-carriers, the expression of SINR is:

$$SINR_{f,k} = \frac{P_{F,k}G_{f,F,k}}{N_0\Delta f + P_{M,k}G_{f,M,k} + \sum_F P_{F',k}G_{f,F',k}} \quad (5)$$

where $P_{F',k}$ and $G_{f,F',k}$ are the transmitting power of neighboring FAP and channel gain between the femto user and neighboring FAP, respectively.

In OFDMA network, FAP has the advantage of allowing the allocation of orthogonal frequency/time resources. In the simulation, the whole bandwidth is divided into certain resource blocks and certain numbers of sub-carriers are allotted to each of the resource blocks. These resource block assumption is adopted from the 3GPP-LTE standard concept defined in [21]. The randomly deployed FAP operates in the same bandwidth and same carrier frequency. A frequency reuse factor of 1 is proposed so that each cell has the full system bandwidth [22].

## III. Resource Allocation Management Algorithm (DRAMA) based hybrid access

The algorithm is a combination of both centralized and distributed resource management approach. Radio Network Controller (RNC) controls the network centrally and FAP controls the rest locally. FAP connects to the RNC through the backhaul connection. The central controller determines the relative position of the FAP and MBS. FAP synchronizes the strongest signal with the desired macro signal.

$$C_r \propto C_f, \quad \{N_R \in \Re\} \quad (6)$$

where $C_f$, $C_r$ and $N_R$ are the assigned capacity of FAP owner, random users and number of random user under that FAP, respectively.

Assuming a hybrid resource distribution constant $K_r$

$$K_r = \frac{C_r}{C_f}, \ \{x \in K_r : 0 \le x \le \infty\} \quad (7)$$

If $K_r = \infty$ all the resources of that FAP will be, assign to the Random users.

If $K_r = 0$, the FAP will act like a Closed access FAP

And if $K_r = 1$, the FAP will act like an Open access FAP

For Hybrid access: $0 < K_r < 1$

The FAP allows $N_R$ number of random users to ensure the minimum assign capacity for both type of users. If the assigned bandwidth for the random users is not higher enough, the user might experience a lower service even if they get a better coverage. Any random user, who gets better service from MBS, will switch stations.

Whenever a user is active within the range of the FAP, it will check the user according to the enlisted owner id. The subscribed owner id for each FAP is different. FAP reserves $C_f$ for each owner and $C_r$ for any other user within the range. Detecting the total number of users, FAP selects $N_R$ and calculates owners and random users' throughput. MBS also calculate the available service for random users. If the reserved throughput for the random from FAP is not higher than the MBS, the FAP reduces $N_R$ by 1 each time to calculate the higher service before executing.

TABLE I. Algorithm

***Algorithm: Dynamic Resource Allocation Management Algorithm***

```
% both MBS and FAP are active in the system
for n=1 to N
    n number of users gets active and looks for the strongest cell coverage
    if MBS has the strongest coverage than it connects to the MBS
    else
        if the user is FAP owner, C_f assigned by FAP
        else the user is Random user, C_r assigned by FAP
        end if
        calculate N_R
        i=0
        while i=0
            calculate C_r and C_f
            calculate MBS C_r
            if FAP C_r > MBS C_r
                i=1
            else N_R = N_R - 1
            end if
        end while
    end if
end for
```

## IV. Simulation result

To simulate the performance of the network, a simulation environment is developed in MATLAB©. The simulations are event-based and according to 3GPP standards. The plotted values are an average of 100 independent simulations.

TABLE II. System Parameters

| System Parameters | Value/Range |
|---|---|
| MBS | 1 |
| FAP | 1-15 |
| Active user | 200-600 |
| Active owners in FAP | 3-5 |
| Range of MBS | 400 m |
| Range of FAP | 20 m |
| MBS Antenna height | 30 m |
| FAP Antenna height | 1 m |
| User equipment height | 1 m |
| Frequency | 2 GHz |
| Bandwidth | 10 MHz |
| Sub-carrier spacing | 15 KHz |
| MBS transmission power | 46 dBm |
| Macro antenna gain | 13 dB |
| FAP transmission power | 20 dBm |
| Distribution time interval | 400 |
| Shadow fading std | 6 dB |
| White noise power density | -174 dBm/Hz |
| Modulation Scheme | 64-QAM |
| Number of Resource block | 50 |
| Sub-carrier per resource block | 12 |
| Resource block size | 180 Khz |
| BER | $10^{-6}$ |

The standard length of the cyclic prefix in LTE is 4.69 µs [23]. This enables the system to tolerate path variations of up to 1.4 km with the symbol length set to 66.7 µs. Each sub-carrier can carry maximum data rate of 15 Ksps (kilo-symbols per second). Modulation 64-QAM represents 6 bits per symbol [24]. Therefore, 10 MHz can provide a raw symbol rate of 9 Msps or 54 Mbps. This enables the system to compartmentalize the data across standard numbers of subcarriers.

Figure 3 shows a sample layout of the simulation.

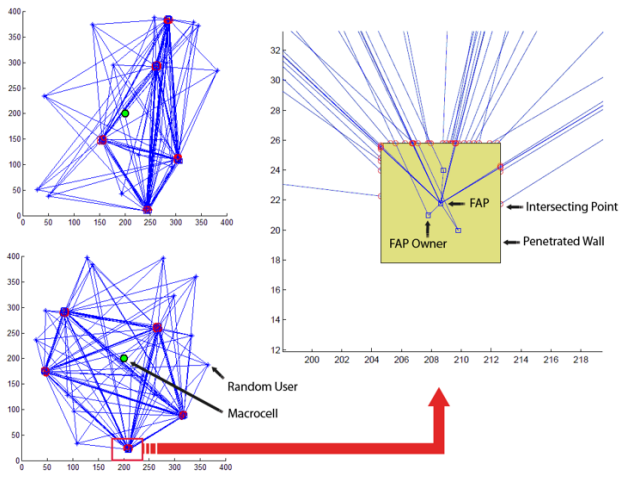



Figure 3. Two layouts of the simulation taking 5 FAPs and 30 UEs

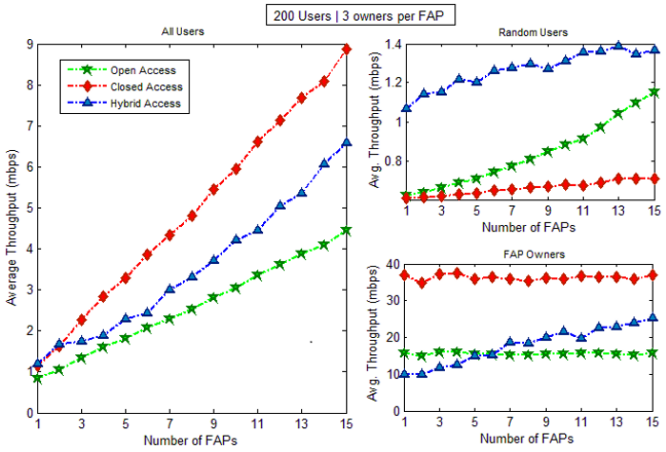



Figure 4. Avg. Throughput vs. number of FAPs for 200 users and 3 owners per FAP

It the figure 4, 5, and 6, the performance of MBS and overall system is increasing along with the number of active FAP. The throughput of FAP owners is always high in the closed access. This high throughput rose the average throughput of all the users even though the throughput of the random users is comparatively low. In the case of open access, the average throughput of the random users grew gradually with the number of FAP as higher number of random users got better service from nearby FAP. The hybrid access average throughput line shows a better performance comparing with the

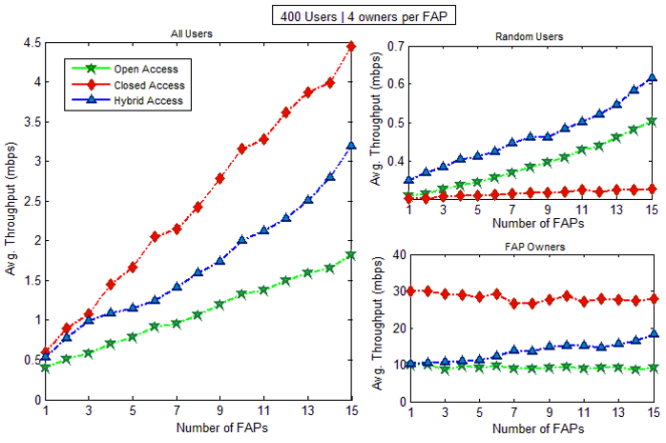



Figure 5. Avg. Throughput vs. number of FAPs for 400 users and 4 owners per FAP

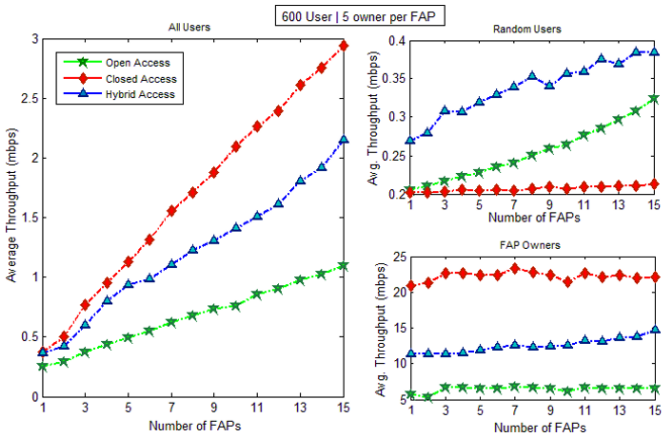



Figure 6. Avg. Throughput vs. number of FAPs for 600 users and 5 owners per FAP

open access. Random users got a boost-up in average throughput as FAPs assigned the unoccupied resources to them. However, the ratio of the unoccupied resources can vary in different FAPs. Floating users who gets access to FAP service possess better throughput than the MBS users. The number of users for each FAP is high enough for minimizing the cross-tier interference and low enough to ensure quality service with less congestions.

## V. CONCLUSION

The paper represents an algorithm for hybrid access mechanism in femtocell network. The network can offload a certain amount of excess MBS load by diverting them to FAPs. It utilizes the unused resources and ensures a superior level of service for the roaming users. Based on the capacity selection, it ensures the minimum resources for the FAP owners and rest to assist MBS offloading. It minimizes macro-femto interferences and suggestively increases the network performance that is suitable for higher deployment densities of femtocell. Future research will focus on the spectrum leasing features to different tier of users.

## REFERENCES


[1] Accenture. (2012) Mobile Web Watch. [Online]. Available: http://www.accenture.com/us-en/Pages/insight-mobile-web-watch-2012-mobile-internet.aspx

[2] N.Q.Logic. Broadband goes mobile [Online]. Available: http://www.nqlogic.com

[3] J. Zhang and G. De la Roche, *Femtocells: technologies and deployment*, Wiley Online Library, 2010.

[4] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *Communications Magazine, IEEE.*, vol. 46, pp. 59-67, 2008.

[5] P. Xia, V. Chandrasekhar, and J. G. Andrews, "Open vs. closed access femtocells in the uplink," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 3798-3809, 2010.

[6] H. Widiarti, S.-Y. Pyun, and D.-H. Cho, "Interference mitigation based on femtocells grouping in low duty operation," in *Vehicular Technology Conference Fall (VTC 2010-Fall),* 2010 IEEE 72nd, 2010, pp. 1-5.

[7] P. Tarasak, T. Q. Quek, and F. Chin, "Closed access OFDMA femtocells under timing misalignment," in *Global Telecommunications Conference (GLOBECOM 2010)*, 2010 IEEE, 2010, pp. 1-5.

[8] H.-S. Jo, P. Xia, and J. G. Andrews, "Open, closed, and shared access femtocells in the downlink," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, pp. 1-16, 2012.

[9] H. A. Mahmoud, I. Guvenc, and F. Watanabe, "Performance of open access femtocell networks with different cell-selection methods," in *Vehicular Technology Conference (VTC 2010-Spring)*, 2010 IEEE 71st, 2010, pp. 1-5.

[10] I. Demirdogen, I. Guvenç, and H. Arslan, "A simulation study of performance trade-offs in open access femtocell networks," in *21st International Symposium Personal, Indoor and Mobile Radio Communications Workshops (PIMRC Workshops),* 2010 IEEE, 2010, pp. 151-156.

[11] L. Li, C. Xu, and M. Tao, "Resource Allocation in Open Access OFDMA Femtocell Networks," *IEEE Wireless Communications Letters,* vol. 1, pp. 625 - 628 2012.

[12] X. Chu, Y. Wu, L. Benmesbah, and W.-K. Ling, "Resource allocation in hybrid macro/femto networks," in *Wireless Communications and Networking Conference Workshops (WCNCW),* 2010 IEEE, pp. 1-5.

[13] F. Tariq, L. S. Dooley, A. S. Poulton, and Y. Ji, "Dynamic fractional frequency reuse based hybrid resource management for femtocell networks," in *Wireless Communications and Mobile Computing Conference (IWCMC),* 2011 7th International, 2011, pp. 272-277.

[14] L. Le, D. Niyato, E. Hossain, D. Kim, and D. Hoang, "QoS-Aware and Energy-Efficient Resource Management in OFDMA Femtocells," *IEEE Trans. Wireless Commun.*, vol. 12, 2013, pp. 180 – 194.

[15] J.-H. Yun and K. G. Shin, "Adaptive interference management of OFDMA femtocells for co-channel deployment," *Selected Areas in Communications, IEEE Journal on,* vol. 29, pp. 1225-1241.

[16] D. Fooladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, 2013, pp. 248 - 257.

[17] X. Chu, Y. Wu, D. López-Pérez, and X. Tao, "On providing downlink services in collocated spectrum-sharing macro and femto networks," *IEEE Trans. Wireless Commun.*, vol. 10, 2011, pp. 4306-4315.

[18] W. Debus and L. Axonn, "RF Path Loss & Transmission Distance Calculations," Axonn, LLC, 2006.

[19] P. Kyosti, J. Meinila, L. Hentila, X. Zhao, T. Jamsa, C. Schneider, et al., "WINNER II channel models," European Commission, Deliverable IST-WINNER D, vol. 1, 2007.

[20] C. Bouras, V. Kokkinos, K. Kontodimas, and A. Papazois, "A simulation framework for LTE-A systems with femtocell overlays," in *Proceedings of the 7th ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks,* 2012, pp. 85-90.

[21] 3GPP Release 8. [Online]. Available: http://www.3gpp.org/Release-8

[22] M. Rahman and H. Yanikomeroglu, "Enhancing cell-edge performance: a downlink dynamic interference avoidance scheme with inter-cell coordination," *IEEE Trans. Wireless Commun.*, vol. 9, 2010, pp. 1414-1425.

[23] A. Ghosh, R. Ratasuk, W. Xiao, B. Classon, V. Nangia, R. Love, et al., "Uplink control channel design for 3GPP LTE," in *Personal, Indoor and Mobile Radio Communications (PIMRC),* 2007. IEEE 18th International Symposium, 2007, pp. 1-5.

[24] A. Pande, V. Ramamurthi, and P. Mohapatra, "Quality-oriented Video delivery over LTE using Adaptive Modulation and Coding," in *Global Telecommunications Conference (GLOBECOM 2011),* IEEE, 2011, pp. 1-5.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Security Enhancement of Public Cloud by Parity Encryption through Two-dimensional Parity Scheme

Md. Nafiur Rahman Protik*, Fatema Khatun
Dept. of Telecommunication and Electronic Engineering, Hajee Mohammad Danesh Science and Technology University, Dinajpur, Bangladesh
*E-mail: protik69@hotmail.com

*Abstract*—**Cloud computing concept is a reality now due to the advancement and transformation of computer generations. However, the implementation of this concept as a technology is still at infancy stage due to several issues such as policy, standards, users, security etc. This paper features the basic concepts of cloud computing, security issues and proposes an encryption and decryption process which is the parity encryption and decryption. The parity encryption and decryption algorithm proposed for the cloud environment integrating Kerberos bit conversion procedure and two-dimentional parity scheme. It is introduced because that can enhance the security in public cloud through several stages encryption. Stages are determined by authenticated user id and password. The novelty of parity encryption is that for "n" stages encryption it increases in every message blocks $2^n$-1 bits in encrypted bit stream comparing to 1st stage message blocks. Furthermore, the proposed encryption provides concrete cryptography.**

***Keywords—Cloud service model, public cloud system, parity encryption, security issues, two-dimensional parity scheme.***

## I. INTRODUCTION

Cloud computing is a paradigm shift against traditional IT resources setup which offers dynamically scalable resources provisioned as services over the internet [1], [3], [4]. It is believed that cloud computing offers tremendous economic benefits. In addition, it allows service becomes convenient and users can demand network access to a shared pool of configurable computing resources (e.g., servers, networks, storage, applications, and services). It can also rapidly provision, deploy, and release with the minimal management effort. The cloud computing has some essential characteristics, such as resource pooling, on-demand self-service, broad network access, measured service, and rapid elasticity [2]-[5], which means users access information easily from anywhere.

Based on the cloud service model, the security responsibility will be separated to 2 parts: one is the service providers and another one is the consumer. The service provider responsibilities are for security up to hypervisor, including physical security, environmental security, and virtualization security. The consumer is responsible for security controls that relate to the IT system, including the operation system, application, and data.

Irrespective of the service mentioned above cloud computing is commonly deployed in four models [6] based upon customers' requirement. They are *Private Cloud, Community Cloud, Public Cloud,* and *Hybrid Cloud.* Public cloud computing model is adopted in this study because provide increased user mobility, mass energy reduction and cost saving. Public Cloud is a setup which is open for use by the general public [7]. This setup could be controlled, maintained and manipulated by different government organizations or corporate organizations or academic institutions, or any combination of them permitted by the Cloud Service Provider (CSP). On the other hand, public cloud exposes users, their assets in cloud and data to extreme danger since they can be publicly accessed through the Internet. Due to these reasons, users and organizations are not confident in adapting to public cloud. This is why the other three models are more favorable to public cloud model. Integration of security, authentication and functionality is vital in maintaining an efficient and effective working environment for the business. One of the essential tasks in maintaining a secure environment is the control of user identities. Creating and deleting identities, monitoring them, enforcing password policies are routines that are important IT operational tasks. This does not change with the public cloud. Using strong passwords and changing them on regular basis are crucial in a public environment. Even though identity federation ensures these tasks but these identities may be easily stolen and imitated.

The key benefits of cloud computing is the ability to instantly expand or contract computational capacity as required and this make it more possible to be attacked. The attackers could precisely map where a target's data is physically within the cloud and then use various tricks to gather intelligence. Sitting on the same physical machine is still the problem for the hackers to hack in even creates the unreachable virtual walls between the virtual machines. It can also detect the victim's passwords through a so-called keystroke attack or from a victim's previously idle virtual machine. If an attacker sits on the same server as his victim, a conventional denial-of-service attack becomes easier by amplifying up his resource usage [7]. Therefore, the cloud security becomes one of the most important subjects. On this

paper, we will focus on the proposed algorithm which can enhance the public cloud security matter.

The rest of the paper is divided into the following sections. Section II describes the three major cloud computing service models, section III describes employment model and section IV is for security issues of cloud. The main thrust of this paper – to introduce with an algorithm that is described in section V and VI which enhance the security matter of public cloud system. The proposed algorithm is "Parity encryption and decryption" which follow the rules of two-dimensional parity scheme.

## II. Cloud Computing Service Model

Cloud computing consists of three service models [3]-[7]. Infrastructure-as-a-Service (IaaS) is the bottom most layer. It provides basic computing infrastructure components such as storage, CPU, memory offered by Citrix, 3tera and VMware who are among the IaaS vendors. These resources are offered to users over the internet ondemand and rental basis. The second layer which serves as a platform to deploy and dynamically scalable web applications such as Google App Engine is known as Platform-as-a-Service (PaaS). The top most layer is called as Software-as-a-Service (SaaS) [8]. It provides users with ready to use applications such as Zoho, Google Docs and Microsoft CRM. In addition these services are supported by two main technologies, web services and web browsers to access these services [9]. Web services are commonly used to provide access to IaaS services and Web Browsers are used to access SaaS applications. Both approaches can be found in PaaS. The cloud computing service model is showed in Fig. 1.

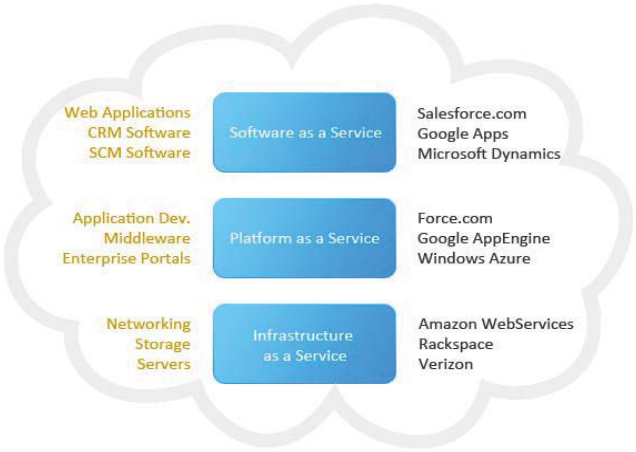



Figure 1. Cloud computing service model

## III. Public Cloud Employment Model

Depending on infrastructure ownership, there are four deployment models of cloud computing each with its merits and demerits. This is where the security issues start.

The Public Cloud - this is the traditional view of cloud computing in every day lingua. It is usually owned by a large organization (e.g., Amazon's EC2, Google's AppEngine and Microsoft's Azure). The owner-organization makes its infrastructure available to the general public via a multi-tenant model on a self-service basis delivered over the Internet. Public cloud deployment models and issues are shown in Table I. This is the most cost-effective model leading to substantial savings for the user, albeit with attendant privacy and security issues since the physical location of the provider's infrastructure usually traverses numerous national boundaries.

TABLE I. Public Cloud Deployment Models and Issues

| Models | Cost issues | Security Issues | Control Issues |
|---|---|---|---|
| **Public** | Setup: Highest<br>Usage: lowest | Least secure | Least control |

## IV. Public Cloud Securyty Issues in Cloud Computing

### A. Availability

This is one of the prime concerns of mission and safety critical organizations. Availability concerns also extend to the need to migrate to another provider, uptime periods of current provider or long-term viability of the cloud provider [10].

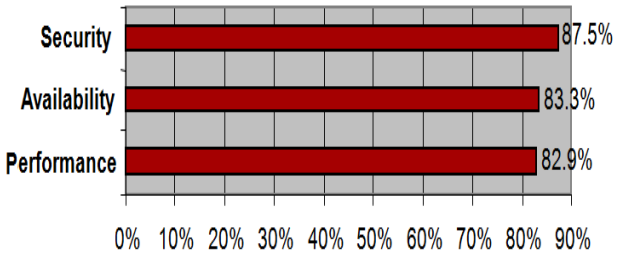
### B. Security

Security has always been the main issue for IT Executives when it comes to cloud adoption. In two surveys carried out by IDC in 2009, security came top on the list (see Fig. 2). However, cloud computing is an agglomeration of technologies, operating systems, storage, networking, virtualization, each fraught with inherent security issues. For example, browser based attacks, denial of service attacks and network intrusion become carry over risks into cloud computing. There are potentials for a new wave of large-scale attacks via the virtualization platform. Chow et al. described the "Fear of the Cloud" by categorizing security concerns into three traditional concerns, availability and third party data control. Research firm Gartner posited seven security risks ranging from data location and segregation to recovery and long-term viability. The European Network and Information Security Agency also published a list of 35 issues in cloud computing in 4 categories. Organizations such as ISACA and Cloud Security Alliance publish guidelines and best practices to mitigate the security issues in the cloud [10]-[11].

### C. Data Security

This risk stems primarily from loss of physical, personnel and logical control of data. Issues include virtualization vulnerabilities, SaaS vulnerabilities (e.g., a case in which Google Docs exposed private user files) [11], phishing scams and other potential data breaches. Other data security risks mentioned in include data leakage and interception, economic and distributed denial of service and loss of encryption keys. Unique risks also arise due to the multi-tenancy and resource-sharing models as pointed out in [12]. The inability to fully segregate data or isolate separate users can lead to undesired

exposure of confidential data in the investigation of a situation involving co-tenants. Hypervisor vulnerabilities can also be leveraged to launch attacks across tenant accounts. Data containing social and national insurance details, health data and financial information raise issues about authorization, rights management, authentication and access controls




(Scale: 1 = Not at all concerned 5 = Very concerned)
Source: IDC Enterprise Panel, 3Q09, n = 263, September 2009

Figure 2. Cloud User Surveys - Challenges

## V. Parity Encryption and Its Potential in the Cloud

Data in the cloud typically resides in a shared environment, but the data owner should have full control over who has the right to use the data and what they are allowed to do with it once they gain access. To provide this data control in the cloud, a standard-based heterogeneous data-centric security approach is an essential element that shifts data protection from systems and applications. In this approach, documents must be self-describing and defending regardless of their environments [13], [14].

Cryptographic approaches and usage policy rules must be considered. When someone wants to access data, the system should check its policy rules and reveal it only if the policies are satisfied [14]-[17]. Existing cryptographic techniques can be utilized for data security, but privacy protection and outsourced computation need significant attention—both are relatively new research directions. Data provenance issues have just begun to be addressed in the literature. In some cases, information related to a particular hardware component (storage, processing, or communication) must be associated with a piece of data.

In our method, we used the concept of even parity by using the parity checker. As we already know that even parity means that the data value contains even number of "1". We inserted "1" at a data value if it had odd parity otherwise "0". In case, if even parity is not present at that location then we made even parity over that location by adding "1". In this way we can insert "0" or "1" at specific location. The insertion process is shown in Fig. 3.

For Retrieval of message, again we used the parity checker. If even parity is present at the selected location then "0" is message bit, else message bit is "1". Retrieval process was repeated for all locations where message bits were hidden. In this way, we retrieved the message bits from all the locations where the message bit were inserted. Fig. 3 shows a two-dimensional generalization of the single bit parity scheme. A parity value is computed for each row and for each column. With this two dimensional parity scheme, the parity of both columns and the row containing the flipped bit will be in error. The receiver can thus not only detect the fact that a single bit error has occurred, but can use the column and row indices of the column and row with parity errors to actually identify the bit was corrupted and correct that error. Two-dimensional parity can also detect any combination of the two errors in a packet.

Example:

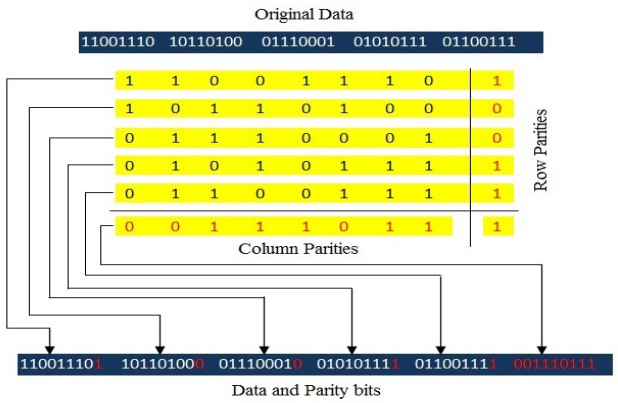



Figure 3. Example for two dimensional even parity scheme

## VI. Proposed Algorithm for Parity Encryption and Decryption for Public Cloud

Larger block sizes mean greater security (all other things being equal) but reduced encryption/decryption speed for a given algorithm. The greater security is achieved by greater diffusion. A block size of 64 bits has been considered a reasonable tadeoff and was nearly universal in block cipher design. For this 64 bits block size is used here for our total encryption process. The essence of this proposed algorithm is that a single round offers inadequate security but the multiple rounds or stages offer increasing security. The basic parity encryption process showed in Fig. 4.

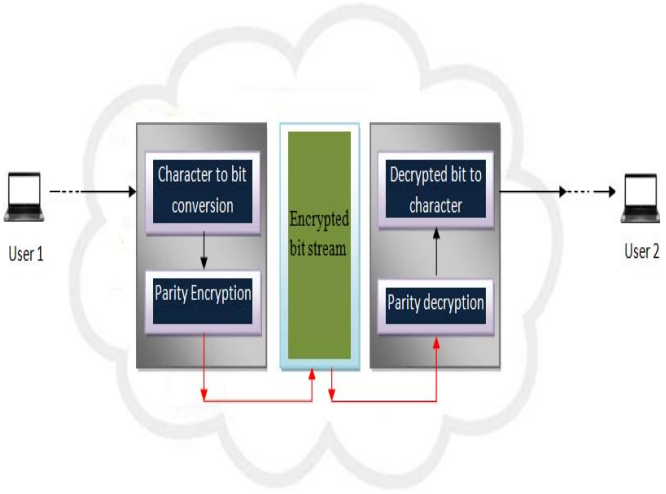



Figure 4. Summary of proposed parity encryption and decryption

### A. *Total Stage Determination for Encrypting and Decrypting Process*

When a user is requesting to connect to the cloud to access his resources

Let,

***u*** = id for the authenticated user;

**pass** = Authenticated user password;

Doing **AND** operation among **u** and **pass**. After completing the operation, the data then convert into digital value which is considered as total stage number to perform encryption as well as decryption event. For bit conversion here "Kerberos" bit conversion process is applied. By this process data are represented in a 7-bit ASCII format. Data of arbitrary length is converted into bit stream which is then assign as below following Kerberos process. This bit stream is used for encryption process. Kerberos process also follow for the conversion bit to data in decryption section. Karberos bit conversion process is showed in Fig. 5.

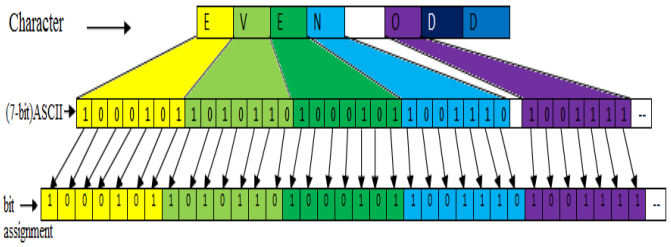



Figure 5. Bit conversion process

### B. *Algorithm for $n^{th}$ Stage Parity Encryption*

Steps are:

i. Takes bits to encrypt $\mathbf{2^n}$ (example for stage 1 then $2^1$=2 bits for each message block from original message).For $n^{th}$ stage original message divided as message block-

Original message block = $m_1 + m_2 + m_3 + -- + m_n$;

ii. Add extra bits"0" in last message block if ($m_n < 2^n$).

iii. Add parity result to construct cipher text. For $n^{th}$ stage message block then represent as-

Cipher text = $m_1p_1 + m_2p_2 + m_3p_3 + -- + m_np_n + (c_1 + c_2 + c_3 + --- + c_n + C)$; [see Fig. 6, 7 and Fig. 8]

So we can rewrite Cipher text as-

Cipher text= a + b + c + d + f;

Where,

a= $m_1p_1$;
b= $m_2p_2$;
c= $m_3p_3$;
- - - - - - - - -
- - - - - - - - -
d= $m_np_n$;
f= $c_1+c_2+c_3+---+c_n+C$;

iv. After successful encryption some information need to assign:

- $P_{RE}= p_1 + p_2 + \ldots\ldots\ldots + p_n$;
- $P_{CE}= c_1 + c_2 + \ldots\ldots\ldots + c_n$;
- $E_e$ , $S_e = n^{th}$;

- If extra bit added

$B_{extra}$= {added bit "0", number of bits }

where "+"=bit by bit added

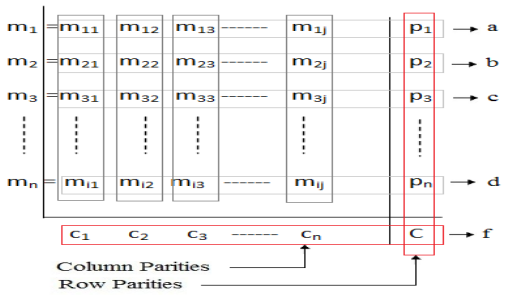



Figure 6. Two-dimensional parity encryption process

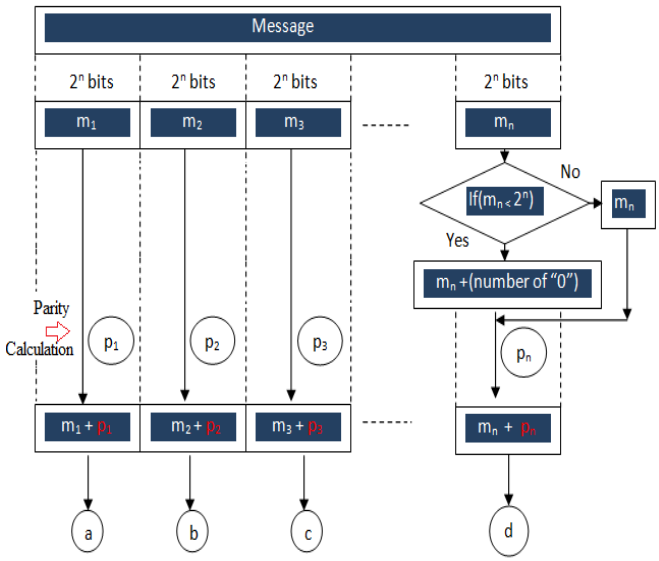



Figure 7. Parity encryption process (row parities)

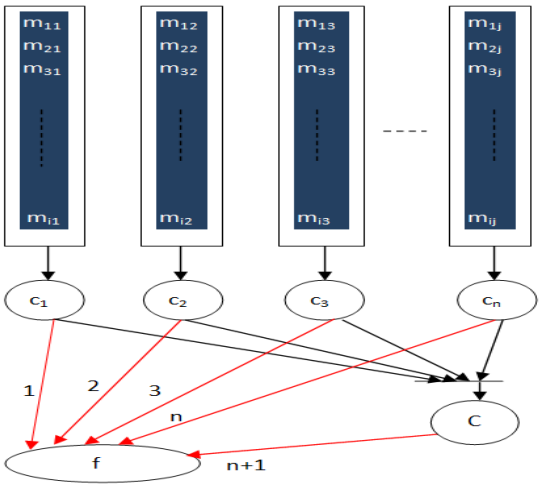


Figure 8. Parity encryption process (column parities)

A desirable property of any encryption algorithm is that a small change in plaintext should produce a significant change in the cipher text. In particular, a change in one bit of the plaintext should produce a change in many bits of the cipher text. Table II shows that after just 4 stages 15 bits increased in a block from the main block. As brief, for the 1st stage every message blocks contain 2 bits ($2^1$=2). After parity encryption every message blocks become 3 bits. So bit increased occur here 1 (3-2). According to the previous manner 2nd stage contains 5 bits after encryption which is increased about 3 bits from stage 1 original message blocks. Table II shows the avalanche effect in parity encryption in every stage's message block with respect to stage 1 message blocks. Also parameters are showed in Table III.

TABLE II. AVALANCHE EFFECT IN PARITY ENCRYPTION

| Change in plaintext block | |
|---|---|
| **Round/stage** | **Number of bits that increased** |
| 1 | 1 |
| 2 | 3 |
| 3 | 7 |
| 4 | 15 |
| ------ | ------- |
| n | $2^n$-1 |

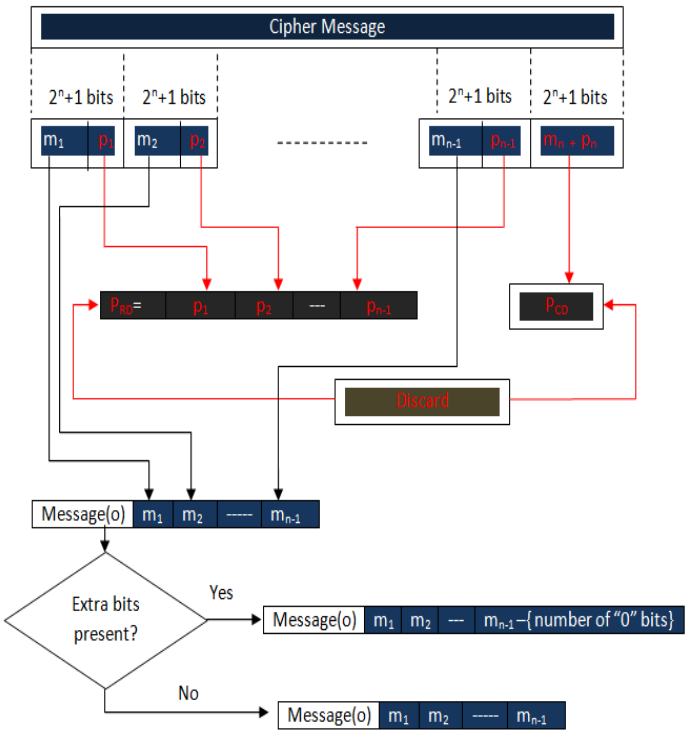



Figure 9. Parity decryption process

### C. *Algorithm for $n^{th}$ Stage Parity Decryption*

Steps are opposite of parity encryption except some matter. Steps are:

i. Takes bits to encrypt $2^n$+1 (example for stage 1 then $2^1$+1=3 bits for each message block from cipher message);
Cipher message block = $m_1p_1 + m_2p_2 + m_3p_3 + \text{-- } + m_{n-1}p_{n-1} + m_np_n$ ; [see Fig. 9]

ii. Last bit from message block {1 to (n-1)} which are row parity and last message block ($n^{th}$ block) assign as column parity:
   - $P_{RD} = p_1 + p_2 + \ldots\ldots\ldots + p_{n-1}$;
   - $P_{CD} = m_np_n$;

iii. **For**($S_d$=$n^{th}$ (initializing) to stage 1)
{ **For**($S_e$=$n^{th}$(initializing) to stage 1)
{ **If**($S_d^{th}$ = = $S_e^{th}$)
{ **If**($P_{RD}$ ==$P_{RE}$ && $P_{CD}$== $P_{CE}$)
{
$P_{RD}$, $P_{CD}$ are discarded;
Then message=$m_1$+ $m_1$+------ + $m_{n-1}$;
**If**(extra bits present==true)
{ Original message = $m_1$+ $m_1$+------ +$m_{n-1}$;}
**else**
{
Original message
=$m_1 + m_1 +$ ------ + $m_{n-1} - B_{extra}$;
}
}
}
Else{data retrieve by parity checking;}
}end **For**;
}end **For**;

iv. Get final original message;

TABLE III. PARITY ALGORITHM PARAMETERS

| | **For encryption** | **For decryption** |
|---|---|---|
| **Block size(bits)** | $2^n$ | $2^n$+1 |
| **Message block** | $m_1 + m_2 + m_3 + \text{-- } + m_n$ | $m_1p_1 + m_2p_2 + m_3p_3 + \text{-- } + m_{n-1}p_{n-1} + m_np_n$ |
| **Parity scheme** | Two-dimensional parity | Two-dimensional parity |
| **Number of event** | Same as total stage | Same as total stage |
| **Output** | $m_1p_1 + m_2p_2 + m_3p_3 + \text{-- } + m_np_n + (c_1 + c_2 + c_3 + \text{--- } + c_n + C)$ | $m_1 + m_1 +$ ------ $+ m_{n-1}$ or $m_1 + m_1 +$ ------ $+ m_{n-1} - B_{extra}$ |
| **Parity(Row ,column)** | $P_{RE}$,$P_{CE}$ | $P_{RD}$,$P_{CD}$ |
| **Stage** | $S_e^{th}$ | $S_d^{th}$ |

## VII. Conclusion and Future Work

In this study, parity encryption and decryption algorithm is proposed for public cloud computing model. Public cloud computation model contains least secure. So, proposed model is enhanced the security with its simply coding model for the cloud system. The main superiority aspect of this algorithm is there is a great benefit in making the algorithm easy to analyze. It is easier to analyze as well as difficult for cryptanalyst. Because it is impossible to get all the parameters which are stored in cloud system in every encryption stages. For that encryption stage is not accessible or imaginable by others. Also another matter is that if anything wrong occur in bit stream then that can be find out and retrieve the original bit streams, which is the another best feature of two-dimensional parity scheme.

The cloud environment involves the scalable, expandable, and virtualized. It makes security become more complex. To have more understanding, more experiments are needed. There are still several areas which can enhance the security matter that are not addressed in this paper. These areas can be ventured in further works of this research. The proposed algorithm presented in this paper is the initial step and needs more modifications; however it can provide the basis for the deeper research on security deployment of cloud computing for the research community working in the field of Cloud Computing.

## Acknowledgment

Authors are cordially grateful to the Department of Telecommunication and Electronic Engineering, Hajee Mohammad Danesh Science and Technology University, Dinajpur to conduct this research. The authors would also like to thank Md. Mehedi Islam of HSTU, dinajpur for his most constructive remarks and suggestions.

## References

[1] Rohit Bhadauria, R. C. (2012). A Survey on Security Issues in Cloud Computing. International Journal of Computer Applications , 47 (18), 47-66.

[2] Allan A.Friedman, D. M. (2010, October). Privacy and Security in Cloud Computing. Issues in Technology Innovation , pp. 1-13.

[3] Jenson Meiko, J. S. (2009). On Technical Security Issues in Cloud Computing. International Conference on Cloud Computing. 2, pp. 109-116. IEEE.

[4] John C.Roberts II, W. A.-H. (2011). Who Can You Trust in the Cloud?A Review of Security Issues Within Cloud Computing. Information Security Curriculum Development Conference (pp. 15-19). Kennesaw, GA, USA: ACM.

[5] Soren Bleikertz, M. S. (2010). Security Audits of Multi- tier Virtual Infrastructures in Public Infrastructure Clouds. ACM, 93-102.

[6] W.Webb, K. (2004). Biometric Security Solution. In Biometrics for Network Security (p. 376). IEEE computer society.

[7] Wu, Z. (2008). Biometrics Authentication System on Open Network and Security Analysis. International Symposium on Electronic Commerce and Security.

[8] F.A. Alvi, B. C. (2013). A review on cloud computing security issues & challanges. International Journal of Cloud Computing and Services Science(IJ- CLOSER) , 2 (1).

[9] Doina Bein, W. B. (2009). The Impact of Cloud Computing on Web 2.0. Economy Informatics Journal , 9 (1), 5-12.

[10] Aderemi A. Atayero, Oluwaseyi Feyisetan ,"Security Issues in Cloud Computing:*The Potentials of Homomorphic Encryption",* Journal of Emerging Trends in Computing and Information Sciences(CIS) VOL. 2, NO. 10, October 2011

[11] Kuyoro S.O, I. F. (2011). Cloud Computing Security Issues and Challanges. International Journal of Computer Networks (IJCN) , 3 (5), 247-255.

[12] Madhan Kumar Srinivasan, K. S. (2012). State- of- the -art Cloud Computing Security Taxonomies- A classification of security challenges in the present cloud computing environment. International Conference on Advances in Computing, Communications and Informatics (ICAACI-2012) (pp. 470-476). Chennai: ACM.

[13] Swarnpreet Singh, R. B. (2012). Architecture of Mobile Application,Security Issues and services involved in Mobile Cloud Computing Environment. International Journal of Computer & Electronics Research , 1 (2), 58-67.

[14] Kevin Hamlen, M. K. (2010). Security Issues for Cloud Computing. International Journal of Information Security and Privacy , 4 (2), 39-51.

[15] L. B. Jivanadham**,** A.K.M. Muzahidul Islam Yoshiaki Katayama **,** Shozo Komaki**,** S. Baharun **"**Cloud Cognitive Authenticator (CCA)- A Public Cloud Computing Authentication Mechanism"2nd International Conference on Informatics, Electronics & Vision (ICIEV-2013), page-71, 17-18 May, 2013 *(IEEE Xplore)*

[16] I. Menken, and G.. Blokdijk, Cloud Computing Virtualization Specialist Complete Certification Kit - Study Guide Book and Online Course, Emereo Pty Ltd, 2009.

[17] M. Vouk, "Cloud Computing—Issues, Research, and Implementations," *Proc. 30th Int'l Conf. Information Technology Interfaces*, Univ. Computing Centre, Zagreb, Croatia, 2008, pp. 235–246.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Analysis of LTE Radio Parameters in Different Environments and Transmission Modes

Nafiz Imtiaz Bin Hamid*, Nafiu Salele, Mugumya Twarik Harouna, Rammah Muhammad
Department of Electrical and Electronic Engineering,
Islamic University of Technology (IUT), Board Bazar, Gazipur-1704, Bangladesh.
*E-mail: nimtiaz@iut-dhaka.edu

*Abstract*— **The ever-growing need for higher data transmission capacity drives the network service providers to build cellular Long term evolution (LTE) networks in urban areas. 3rd Generation Partnership Project (3GPP) LTE is the evolution of the UMTS in response to the ever-increasing demand for high speed data and high quality multimedia broadcast services. LTE promises to deliver an unrivalled user experience with ultrafast broadband and very low latency and at the same time, a very compelling business proposition for operators with flexible spectrum bandwidth, smooth migration and the ability to deliver low cost per bit voice and data services. LTE is designed to have wider channels up to 20 MHz, with low latency and packet optimized radio access technology. The peak data rate envisaged for LTE is 100 Mbps in downlink and 50 Mbps in the uplink. LTE has many promising features like bandwidth scalability. It is developed to support both the time division duplex as well as frequency division duplex mode. This paper provides analyses of the performance of radio parameters required for efficient LTE radio planning through numerous simulations in different transmission modes and radio environments. It mainly highlights the throughput and Blok Error Rate (BLER) with respect to Signal-to-Noise Ratio (SNR) on the physical layer and in network context through different simulation environments.**

***Keywords— LTE, BLER, SNR, CQI, Throughput***

## I. INTRODUCTION

LTE is a 3GPP standard considered a major advancement in wireless technology. It is expected to be the mobile broadband platform for services in innovation for the foreseeable future [1]. LTE is a fourth generation technology envisaged to provide a peak data rate of 100Mbps in the downlink and 50Mbps in the uplink. It is designed to have wider channels up to 20MHz, with packet optimized radio access technology. LTE has very promising features for example high scalability, Frequency Division Duplex (FDD) and Time Division Duplex (TDD) duplexing mode. To meet the user's expectations, LTE aims at better spectral flexibility, higher data rates, low latency, improved coverage and better battery lifetime. To achieve these targets, mainly three enabling technologies are employed namely; Orthogonal Frequency Division Multiple Access (OFDMA), Single Carrier Frequency Division Multiple Access (SC-FDMA) and Multiple Input Multiple Output (MIMO). LTE employs OFDMA in the downlink direction and SC-FDMA in the uplink data transmissions [2],[3]. To substantially enhance the air interface, MIMO employs multiple transmit and receive antennas, for higher data rates and fight against multi path fading.

The remainder of this paper is organized as follows: Section II contains the brief summary of related works. In Section III an overview of transmission modes has been given. Afterwards, the uses of the link level and system level simulations are presented in Section IV. In Section V, link and system level simulation results and their analyses have been given.

## II. RELATED WORKS

Similar works using link level results include: SNR to Channel Quality Indicator (CQI) mapping for different MIMO settings [4], limiting downlink Hybrid Automatic Repeat Request (HARQ) retransmission in poor link [5]. Radio network planning for Dhaka city- coverage and capacity analysis approach has been suitably presented in [6], [7]. An open-source framework is presented to provide a complete performance verification of LTE networks in [8]. But none of those had the clear motive to thoroughly investigate the LTE radio parameters in different transmission modes and environments running simulation [9] with numerous different settings. So, this has been chosen as the focus of this paper as it will further improve the network planning issue of LTE.

## III. TRANSMISSION MODES

During dynamic resource scheduling, suitable transmission mode can be adapted semi-statically according to various channel conditions. Physical Downlink Shared Channel (PDSCH) channel employs different transmission modes utilizing multiple antennas in both transmitting and receiving sides. Till now nine transmission modes have been released but only first four have been implemented [2]. The nine transmission modes are:

1. Single antenna; port 0,
2. Transmit diversity,
3. Open loop spatial multiplexing,
4. Closed loop spatial multiplexing,
5. MU-MIMO,
6. Closed loop rank=1 precoding,
7. Single antenna; port 5,
8. Dual layer transmission; port 7 and 8 and
9. Up to 8 layer transmission; port 7-14.

Among these transmission modes, single antenna, transmit diversity and spatial multiplexing will likely be the point of interest based on the implementations.

### A. SISO

SISO is used in transmission mode 1. It uses single antenna at the eNodeB. The data rate is the lowest compared to other transmission modes.

### B. Transmit Diversity

Transmit Diversity (TxD) is used in transmission mode 2. Transmit diversity increases the SNR at the receiver instead of directly increasing the data rate. Each transmit antenna transmits essentially the same stream of data and so the receiver gets replicas of the same signal. It improves the cell edge user data rate and coverage range. The transmit diversity is an open-loop scheme and feedback from the UE is not required. Transmit diversity is only defined for 2 and 4 transmit antennas and one data stream. The number of layers is equal to the number of antenna ports.

### C. Spatial Multiplexing

Spatial multiplexing allows multiple antennas to transmit multiple independent streams. So it is sometimes referred to as the true MIMO technique.

- Open-Loop Spatial Multiplexing (OLSM) is used in transmission mode 3. It makes use of the spatial dimension of the propagation channel and transmits multiple data streams on the same resource blocks. The feedback from the UE indicates only the rank of the channel using Rank Indication (RI) and not a preferred precoding matrix and hence, it is termed as open-loop.
- Closed-Loop Spatial Multiplexing (CLSM) is used in transmission mode 4. In this case, the UE estimates the radio channel and selects the most desirable entry from a predefined codebook. Then the UE sends a feedback to the eNodeB and hence, it is termed as closed-loop.

## IV. Simulations for Performance Analysis

For efficient deployment of LTE, performance analyses of different radio parameters are worth investigating. Simulations are necessary to test and optimize algorithms and procedures. These have to be carried out on both the physical layer and in the network context. LTE physical layer is important for conveying both data and control information between an eNodeB and UE. To enable reproducibility and to increase credibility of our results, simulation of the physical layer is done using a link level simulator [9],[10] and in the network using a system level simulator [9],[11].

### A. Link level Simulation

Link level simulations allow for the investigation of channel estimation, tracking, and prediction algorithms, synchronization algorithms, MIMO gains, Adaptive Modulation and Coding (AMC) and feedback. Furthermore, receiver structures, modeling of channel encoding and decoding, physical layer modeling crucial for system level simulations and alike are typically analyzed on link level [7].

### B. System level simulation

System level simulations analyze the performance of a whole network, It focuses more on network-related issues, such as resource allocation and scheduling, multi-user handling, mobility management, admission control, interference management, and network planning optimization [12].

## V. Simulation Results and Analysis

### A. Link Level

The analysis with link level simulations was carried out using parameters stated in Table I. The focus was to analyze throughput and BLER values with the change of SNR. Number of subframes was varied from 100 to 1000 to visualize the effect. Results of throughput vs SNR were obtained and shown in Fig. 1 and 2 respectively. Again, the results of BLER vs SNR are presented in Fig. 3 for 100 subframes, and in Fig. 4 for 1000 subframes. With different transmission modes: SISO, TxD 2×1, TxD 4×2 and OLSM 4x2 simulation of both throughput and BLER were performed taking Pedestrian B (PedB) and Flat Rayleigh channel using CQI value 7.

TABLE I. Basic Settings Used For Link Level Simulator.

| Parameter | Settings |
|---|---|
| Number of UEs | 1 |
| Bandwidth | 1.4MHz |
| Retransmissions | 0 and 3 |
| Channel type | Flat Rayleigh, PedB uncorrelated |
| Filtering | Block Fading |
| Receiver | Soft Sphere Decoder |
| Simulation length | 100,1000 subframes |
| Transmit modes | SISO, TxD (2x1 and 4x2) and OLSM (4x2) |

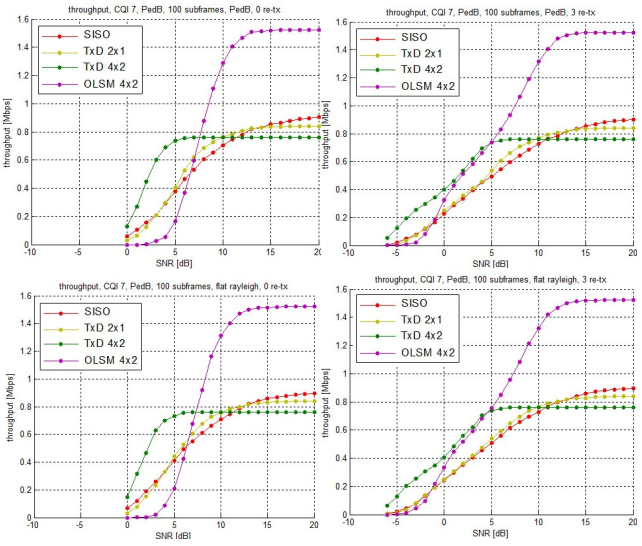



Figure 1. Throughput vs SNR Results For 100 Subframes.

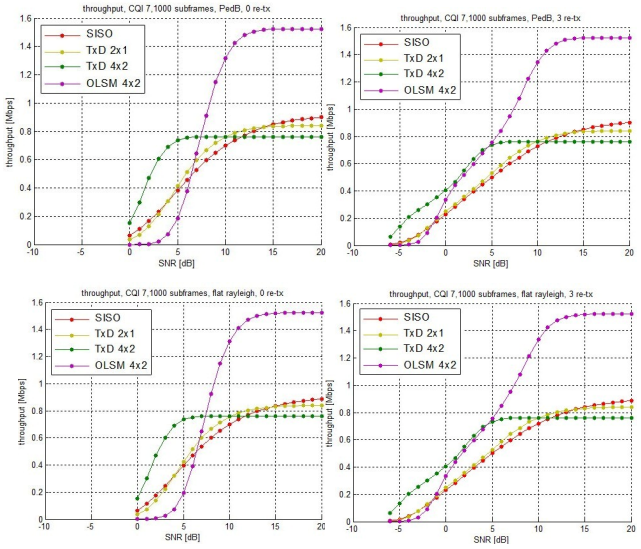



Figure 2. Throughput vs SNR for 1000 Subframes.

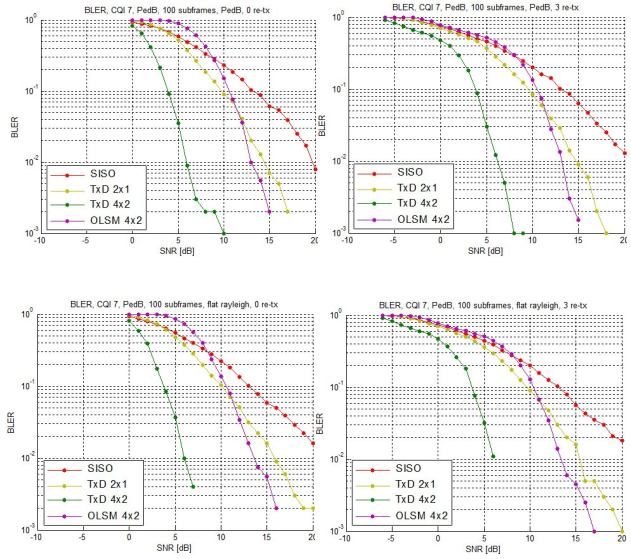



Figure 3. BLER vs SNR results for 100 subframes

It is quite clear from Fig. 1 to 4 that as the number of subframes was increased, the plots became smoother and more realistic. Again, plots obtained using Flat Rayleigh channel look almost similar to those of PedB channel.

*1) Throughput analysis:* Considering the case of 1000 subframes from Fig. 2, if SNR requirement is fixed at 15dB, the maximum throughput is found as 1.52Mbps achieved with OLSM 4x2 and the least throughput is about 0.78Mbps obtained with TxD 4x2.

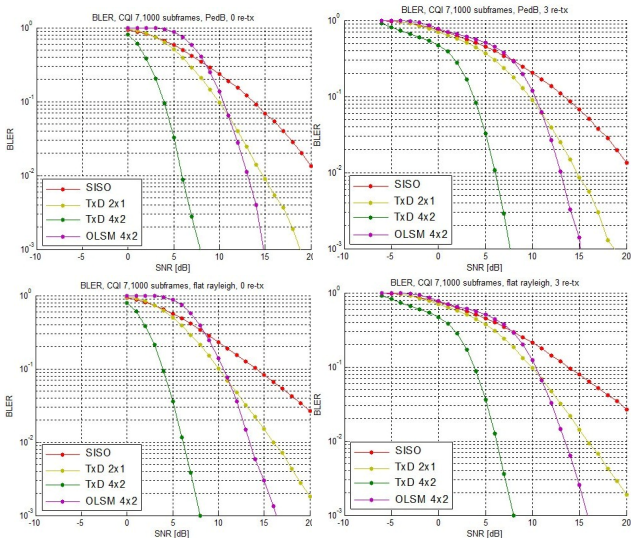



Figure 4. BLER vs SNR results for 1000 subframes.

*2) BLER analysis:* BLER is defined as the ratio of the number of erroneous blocks received to the total number of blocks sent. An erroneous block is defined as a Transport Block, the cyclic redundancy check (CRC) of which is wrong. A 0% BLER is not always necessary or practical, due to the extra time it takes to resend blocks with errors. In LTE, adaptive modulation and coding has to ensure a BLER value smaller than 10 % [10]. If the case of 1000 subframes is considered as per Fig. 4 and BLER value is limited at maximum $10^{-1}$ (10% of the max.); at least a SNR of about 4 dB is required to reach this target BLER. It is achieved through TxD 4x2. This means less signal power is needed with this transmission mode- TxD 4x2 for minimum possible BLER. The same BLER can also be achieved with a maximum SNR of 14dB given by SISO and this implies that more signal power has to be given using that scheme. So, SISO is supposedly not a good choice for BLER sensitive environment because of its higher power requirement.

*3) Limitations and future work:* The link level simulator used [9] for this work was implemented for MIMO modes: Transmit diversity, OLSM only. But CLSM was out of the scope of it, and the simulator could only support one UE per eNodeB, multiple UEs could not be simulated. So, these limitations should be kept as considerations for improvement in future radio planning work.

## *B. System Level*

To determine the level at which predicted link level gains impact network performance, system level simulations were performed [9],[12] and results shown in Fig. 5-6. Parameters set for the simulator were 21 cells which form the region of interest. A simulation length of 50 TTIs was used. Scheduler: Proportional fair, 2 transmitting and 2 receiving antennas, and MIMO Transmit mode was CLSM. Table II-VI show case studies carried out through numerous simulations taking consecutively 10 to 30 user equipments per cell.

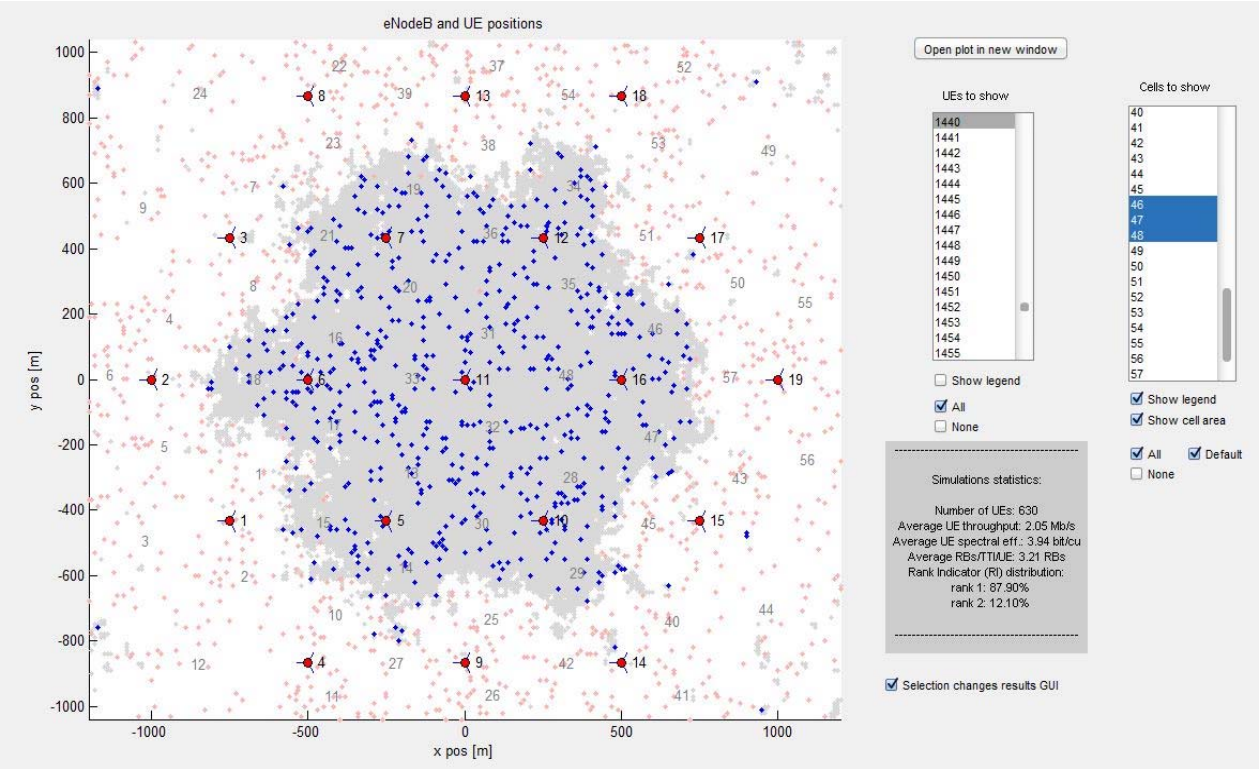



Figure 5. Region of Interest, eNodeB-UE Distribution for 30 UEs per cell

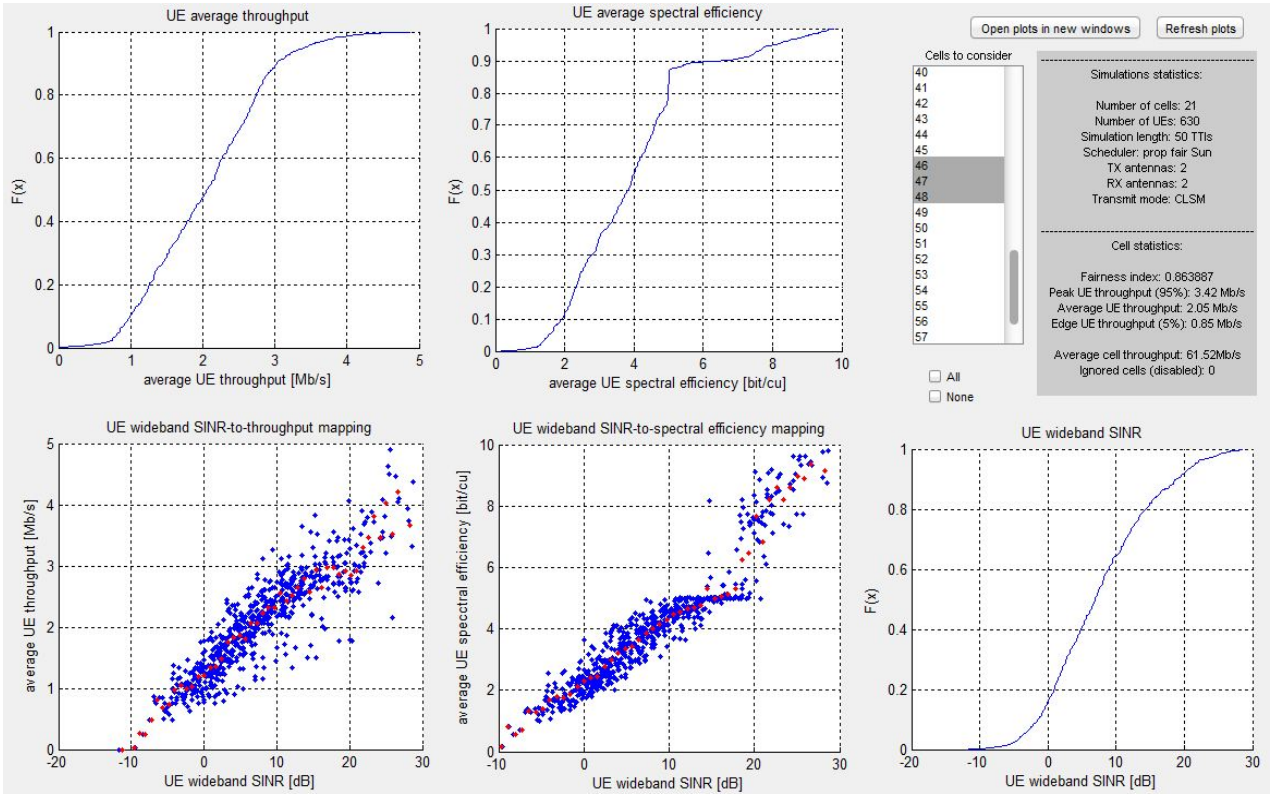



Figure 6. Throughput and aggregate results

TABLE II. CASE STUDY WITH 10UES PER CELL

| User Equipment | Region | Average Throughput (Mb/s) |
|---|---|---|
| UE 311 | Close | 7.17 |
| UE 315 | Intermediate | 7.83 |
| UE 320 | Far | 5.68 |

TABLE III. CASE STUDY WITH 15UES PER CELL

| User Equipment | Region | Average Throughput (Mb/s) |
|---|---|---|
| UE 474 | Close | 5.66 |
| UE 467 | Intermediate | 5.43 |
| UE 472 | Far | 4.16 |

TABLE IV. CASE STUDY WITH 20UES PER CELL

| User Equipment | Region | Average Throughput (Mb/s) |
|---|---|---|
| UE 631 | Close | 3.61 |
| UE 656 | Intermediate | 2.12 |
| UE 651 | Far | 1.38 |

TABLE V. CASE STUDY WITH 25UES PER CELL

| User Equipment | Region | Average Throughput (Mb/s) |
|---|---|---|
| UE 790 | Close | 3.70 |
| UE 792 | Intermediate | 3.23 |
| UE 795 | Far | 2.33 |

TABLE VI. CASE STUDY WITH 30UES PER CELL

| User Equipment | Region | Average Throughput (Mb/s) |
|---|---|---|
| UE 978 | Close | 0.84 |
| UE 973 | Intermediate | 2.79 |
| UE 988 | Far | 2.33 |

*1) Throughput analysis:* From the simulation results, with 21 cells as region of interest in all the cases average throughput of 5.87Mb/s, 4.27Mb/s, 3.00Mb/s, 2.45Mb/s and 2.05Mb/s were attained for 10UEs, 15UEs, 20UEs, 25UEs and 30 UEs per cell respectively. This indicates that the average throughput deteriorates with the increase of UEs per cell. Further analysis at close, far and intermediate regions also implies that UE throughput fades as they move away from the eNodeB. However some UEs at close region were found with low throughput which might result from other factors such as fading, scattering, interference or other phenomenon. Table II – VI show the distributed UEs at different regions with their corresponding throughputs.

*2) Spectral efficiency analysis:* Spectrum efficiency is the optimized use of spectrum or bandwidth so that the maximum amount of data can be transmitted with the fewest transmission errors. It equates to the maximum number of users per cell that can be provided while maintaining an acceptable quality of service (QoS). Here a spectral efficiency of 3.73bit/cu, 4.00bit/cu, 3.79bit/cu, 3.74bit/cu and 3.94bit/cu were also attained for 10UEs, 15UEs, 20UEs, 25UEs and 30 UEs respectively.

*3) Aggregate UE results:* Fig. 6 shows the aggregate UE results, as well as some cell-related statistics. For the UE-related results, only active UEs from the selected cells were used. Results of the Empirical Cumulative Distribution Function (ECDF) of the UE average throughput, ECDF of the UE average spectral efficiency and UE wideband SINR were obtained. In the case of 30 UEs per cell, the average throughput was 2.05 Mb/s, its corresponding CDF was 0.5 as seen in Fig. 6, and this same average throughput had a wideband signal to interference and noise ratio (SINR) of about 7 dB. Average spectral efficiency of 3.8 bit/cu was also obtained at the same CDF.

*4) Limitations and future considerations:* The system level simulator could support maximum 30 user equipments per cell. So, the obtained radio parameter values involve the effect of this limitation along with those of link level simulator. But as broader arrays of variations were made while creating simulation environments, these limitations aren't likely to create any noteworthy negative impact. But to get a more accurate radio network planning these limitations should be overcome and thus all those fall under possible future works.

## VI. CONCLUSION

From the simulation, it is observed that the highest throughput is achieved with MIMO scheme: Open Loop Spatial Multiplexing (OLSM) 4x2 mode, while the suitable BLER is achieved with the transmit diversity (TxD) 4x2 Mode. Besides these, effect of changed number of subframe on throughput; spectrum efficiency for different UE/cell, BLER and aggregate UE parameters involving throughput and CDF were also evident and analyzed with different case-studies. In short, this paper should help guiding the LTE radio network planning work with more precision.

## REFERENCES


[1] (2013) 3GPP: Mobile Broadband Standard website. [Online]. Available: http://www.3gpp.org/

[2] H. Holma, A. Toskala, *LTE for UMTS - OFDMA and SC-FDMA Based Radio Access*, John Wiley & Sons Ltd, 2009.

[3] S. Sesia, I. Toufik, M. Baker, *LTE - The UMTS Long Term Evolution: From Theory to Practice*, John Wiley & Sons Ltd, 2009.

[4] M. T. Kawser, N. I. B. Hamid, M. N. Hasan, M. S. Alam, and M. M. Rahman, "Downlink SNR to CQI Mapping for Different Multiple Antenna Techniques in LTE," *International Journal of Information and Electronics Engineering*, vol. 2, no. 5, pp. 757-760, Sep. 2012.

[5] M. T. Kawser, N. I. B. Hamid, M. N. Hasan, M. S. Alam, and M. M. Rahman, "Limiting HARQ Retransmissions in Downlink for Poor Radio Link in LTE," *International Journal of Information and Electronics Engineering*, vol. 2, no. 5, pp. 707-709, Sep. 2012.

[6] N. I. B. Hamid, M. T. Kawser, M. A. Hoque, "Coverage and Capacity Analysis of LTE Radio Network Planning considering Dhaka City," *International Journal of Computer Applications*, vol. 46, no. 15, pp. 49-56, May 2012.

[7] N. I. B. Hamid, M. A. Hoque, K. K. Islam, "Nominal and Detailed LTE Radio Network Planning considering Future Deployment in Dhaka City," *International Journal of Computer Applications*, vol. 50, no. 17, pp. 37-44, July 2012.

[8] G. Piro, L. Grieco, G. Boggia, F. Capozzi, and P. Camarda, "Simulating LTE cellular systems: An open-source framework," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 498 - 513, Feb. 2011.

[9] (2013) LTE Simulators homepage on Institute of Telecommunications, TUWIEN. [Online]. Available: http://www.nt.tuwien.ac.at/ltesimulator/

[10] C. Mehlfuhrer, M. Wrulich, J. C. Ikuno, D. Bosanska, and M. Rupp, "Simulating the long term evolution physical layer," in *Proc. the 17th European Signal Processing Conference (EUSIPCO 2009)*, Aug. 2009, pp. 1471-1478.

[11] J. C. Ikuno, M. Wrulich, and M. Rupp, "System level simulation of LTE networks," in *Proc. IEEE 71st Vehicular Technology Conference (VTC-Spring)*, May 2010, pp. 1-5.

[12] C. Mehlfuhrer, J. C. Ikuno, M. Simko, S. Schwarz, M. Wrulich, and M. Rupp, "The Vienna LTE simulators - enabling reproducibility in wireless communications research," *EURASIP Journal of Advances in Signal Processing*, vol. 2011, no. 29, 2011.

# A Generalized Model using Genetic Algorithm for Optimization of Material Gain of the Active layer of a MQW Edge Emitting Laser with Unequal Well Width

Md. Mobarak Hossain Polash[1,*], Md Imrul Kayes[2]
[1] Bangladesh University of Engineering and Technology, Bangladesh
[2] Bangladesh University of Engineering and Technology, Bangladesh
[*] E-mail: polash066@gmail.com

*Abstract*— **In this work, a computational model for optimization of material gain of the active region of a multiple quantum well (MQW) edge emitting laser (EEL) using genetic algorithm has been developed. Through this optimization procedure, the values of the width of quantum wells (QW), width of barriers, lattice temperature, injection carrier density and number of QWs which are related to material gain of the active region are optimized for optimizing the design of a MQW EEL. For the above mentioned optimization the numerical simulation of the optical gain expression of MQW EEL along with the solution of the Schrodinger's equation has been performed using MATLAB. The developed optimization based design technique has been applied for the design of (i) a 1550nm $In_{0.72}Ga_{0.28}As_{0.82}P_{0.18}$ / $In_{0.42}Ga_{0.58}As_{0.82}P_{0.18}$ MQW EEL having 3 QWs and also (ii) for a 1550nm $In_{0.72}Ga_{0.28}As_{0.82}P_{0.18}$ / $In_{0.42}Ga_{0.58}As_{0.82}P_{0.18}$ MQW EEL having 5 QWs for testing the validity. The computation indicates clearly that the optimization based computational model works well and can be easily used for the design of MQW EELs.**

***Keywords— MQW Edge Emitting Laser, Material Gain, Optimization.***

## I. INTRODUCTION

The active region material of a laser is usually chosen depending on some of the important properties of laser e.g. (i) laser emission wavelength, (ii) output optical power, (iii) method of laser excitation whether by optical pumping or electrical current injection, (iv) laser power consumption and efficiency, (v) high speed modulation or short pulse generation ability, (vi) wavelength tunability, (vii) output beam quality and (viii) device size [1], [2]. Moreover, the types of application of the laser also limit the choice of active region materials. To obtain a compromise between the required properties of a laser and the types of applications, the choice of semiconductor material for the active region is of significant importance in the design of a laser. Semiconductor diode laser is popular in modern optical fiber communication because such a laser can be fabricated for the wavelengths within the range of around 900–1700 nm [3], [4]. It is also easy to couple a semiconductor diode laser to single mode optical fiber. On the other hand, solid state laser and fiber laser can only emit discrete values of wavelengths restricted by the electronic transitions in ions [1].

Among the semiconductor materials which are commonly used in laser structure, quaternary compounds are mostly used now a days because using these compound semiconductors one can adjust simultaneously the band gap as well as the lattice constant. InGaAsP is one of the commonly used quaternary semiconductor material to achieve the long wavelengths within the range of 1100nm -1650 nm [4]. 1550 nm wavelength is significantly used for long distance fiber-optic communication for its low path loss (0.15 dB/km) [4].

The material gain of the active region material is defined as a fractional increase in photons per unit length [5]. It is an amplification factor that a gain material imparts to the stimulated emission produced by radiative recombination of the conduction sub-band electron and the valence sub-band hole. The amplification of the output optical power happens after overcoming the loss component of the laser [6]. This material gain spectrum contains the information of optimizing operating characteristics like threshold current, lasing wavelength, gain bandwidth etc [7][8][9]. For these reasons, gain optimization is an important issue for the design of laser structure. In this work, the laser structure is an asymmetric MQW EEL where the active region material is chosen to be $In_{0.72}Ga_{0.28}As_{0.82}P_{0.18}$ and the barrier and the separate confinement hetero-structure (SCH) materials are chosen to be $In_{0.42}Ga_{0.58}As_{0.82}P_{0.18}$ which is lattice matched.

To obtain the optimization of the material gain, genetic algorithm is one of the most popular algorithms. The genetic algorithm performs the optimization process by using the idea of the principles of biological evolution, repeatedly modifying a population of individual points using rules modeled on gene combinations in biological reproduction. Due to its random nature, the genetic algorithm possesses high probability of finding a global solution. It enables to solve unconstrained, bound-constrained, and general optimization problems, and it does not require the functions to be differentiable or continuous [10].

## II. COMPUTATIONAL MODEL

### A. Energy sub-band calculation

In quantum well region, from the solution of the well known Schrödinger's equation, it is observed that some sub-band energy levels are created above the conduction band edge and below the valence band edge. The actual carrier transitions take place between these conduction sub-bands to valence sub-bands. For the calculation of discrete energy levels in the quantum well region for both conduction and valence band the

time independent Schrodinger equation under effective mass approximations is used because of parabolic band nature [11].

Considering the effective mass approximations and the parabolic band nature, the time independent Schrodinger's equation [11] is written as.

$$-\frac{\hbar^2}{2m^*}\nabla^2\Psi + V\Psi = E\Psi \quad (1)$$

Where, $\Psi$ is the particle wavefunction, $\hbar$ is the Planck's constant divided by $2\pi$, $m^*$is the effective mass of the carrier, $V$ is band potential and $E$ is energy levels of conduction or valence band.

*B. Optical gain calculation*

The optical gain of a material is written as a function of photon energy to emphasize that the optical gain experienced by an incoming photon is dependent on the photon's energy [5]. The optical gain for a material is achievable over a narrow band of photon's energy. Gain calculation of quantum well is more complicated than the gain calculation of bulk material due to many quantized energy levels between which transitions can occur [12]. Among the all possible transitions between the conduction(C) sub-band and valence (V) sub-band, some transition are allowed and some are forbidden [5] according to the calculation of the transition occurrence probability. The allowed and the forbidden transitions are shown in figure 1.

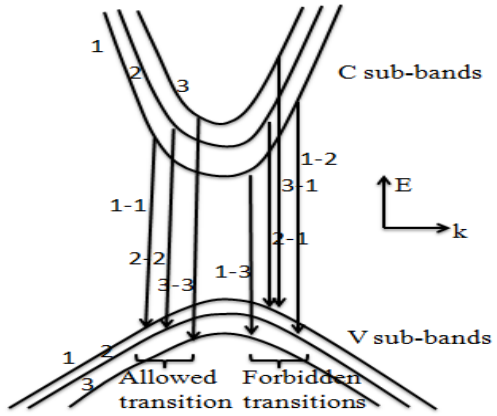



Figure 1: Allowed and forbidden energy transitions between the C sub-bands to V sub-bands of the quantum well.

For the calculation of gain spectra of a MQW semiconductor laser, two expressions of optical gain are used. For the first expression, spectral broadening effect by intraband scattering is not considered. The first gain expression is [5]

$$g(E) = g_0(E)|M_b|^2 \sum_{j,n} \rho_{r,jn}(E)(f_c - f_v)_{E,jn} \quad (2)$$

Here, $g_0$ is gain prefactor, $M_b$ is the average, energy independent, momentum transition matrix element for the dipole transition in the bulk semiconductor; $\rho_{r,jn}$ is the volume density states, $f_c$ and $f_v$ are the electron quasi-Fermi level functions in the conduction and valence bands respectively, $j$ and $n$ are for C sub-band and V sub-band numbers. The second gain expression is presented later in this section.

However, for the first gain expression the photon energy dependent gain prefactor and momentum transition matrix element for dipole transition in the bulk semiconductor are expressed as [5]

$$g_0(E) = \frac{\pi\hbar q^2}{m_0^2 n_{eff} E c \varepsilon_0} \quad (3)$$

And,

$$|M_b|^2 = \frac{2}{3}\left(\frac{m_0}{m_c} - 1\right)\left[\frac{E_g + \Delta}{2(E_g + 2/3\Delta)}\right] m_0 E_g \quad (4)$$

Where, $q$ is electron charge, $m_0$ is the mass of free electron, $m_c$ is the effective mass of electron for active region, $E_g$ is the energy gap between the conduction and valence band edge, $\Delta$ is split off band energy, $n_{eff}$ is the effective refractive index of the laser structure, $E$ is photon energy, $c$ is the vacuum speed of light and $\varepsilon_0$ is the free space permittivity. The expression of volume density of state is as presented below [5]

$$\rho_{r,jn}(E) = \frac{m_{r,j}}{2\pi\hbar^2 L_z} H(E - E_{g,jn}) \quad (5)$$

Here, [7]

$$m_{r,j} = \frac{m_c m_{v,j}}{m_c + m_{v,j}} \quad (6)$$

$m_{r,j}$ is the spatially weighted reduced effective mass for transition where $m_c$ and $m_v$ are conduction and valence band effective masses respectively and $j$ is for denoting the heavy and light holes of valence band. $L_z$ is the width of a QW. $H(E - E_{g,jn})$ is the Heaviside unit step function where, [5]

$$E_{g,jn} = E_g + E_{c,n} + E_{v,jn} \quad (7)$$

$E_{g,jn}$ is the energy gap between the sub-band pairs conduction and valence band where, $E_{c,n}$ is the energy gap between C band edge to $n$ C sub-band and $E_{v,jn}$ is the energy gap between V band edge to $n$ V heavy hole sub-band or light hole sub-band. These energy levels are shown in figure 2 with an arbitrary unit.

The Fermi-Dirac functions $f_c$ and $f_v$ for the QW semiconductor laser are evaluated with the respective quasi-Fermi levels being fixed by the injection carrier density. The expressions for the $f_c$ and $f_v$ are [5]

$$[f_c]_{E,jn} = [1 + \exp\{(E'_{jn} + E_{c,n} - F_c)/k_b T\}]^{-1} \quad (8a)$$

$$[f_v]_{E,jn} = [1 + \exp\{-(E''_{jn} + E_{v,jn} - F_v)/k_b T\}]^{-1} \quad (8b)$$

Where, $F_c$ and $F_v$ are the quasi-Fermi levels measured from the respective band edges of the bulk semiconductor and $k_b$ is Boltzmann's constant. From the parabolic band approximation, for constant effective masses the energies in the expressions of $f_c$ and $f_v$ are calculated from following relationships [5].

$$E'_{jn} = (E - E_{g,jn})m_{r,j}/m_c \quad (9)$$

$$E''_{jn} = (E - E_{g,jn})m_{r,j}/m_{v,j} \quad (10)$$

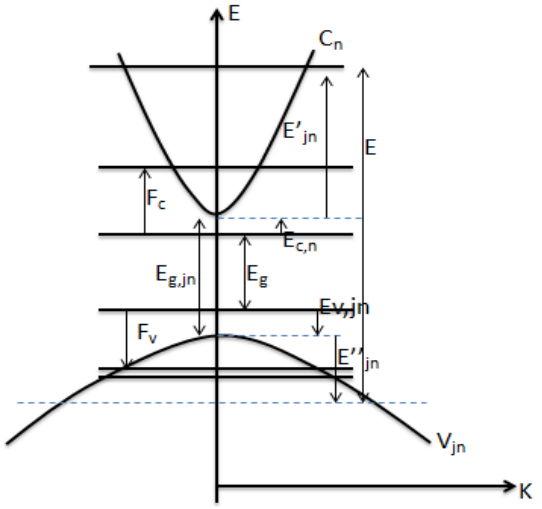



Figure 2: E versus K energy band diagram for a particular sub-band pair showing all the energy level including energy band gap, Fermi-level and the transition energy levels.

For accurate calculation of gain spectra, energy and density dependent broadening [8][9][13] is considered in the second gain expression which is given below [5][14].

$$g(E) = g_0(E)|M_b|^2$$
$$\sum_{j,n} \int_{E_{g,jn}}^{E_{g,b}} \rho_{r,jn}(\hat{E}) A_{ij} C_{ij} (f_c - f_v)_{\hat{E},jn} L(E - \hat{E}) d\hat{E} \quad (11)$$

Where,

$$L(E - \hat{E}) = (\delta E_{in}/\pi)[(E - \hat{E})^2 + (\delta E_{in})^2]^{-1} \quad (12)$$

$L(E - \hat{E})$ is the normalize Lorentzian lineshape function where $\delta E_{in} = \hbar/\tau_{in}$, is the intraband energy broadening factor. $A_{ij}$ is angular anisotropy factor [15] and $C_{ij}$ is the spatial overlap factor between the states i and j [15]. For this second gain expression the photon energy dependent gain prefactor, momentum transition matrix element for dipole transition in the bulk semiconductor, quasi Fermi level equation, equations for the reduced density states are same as equations (3)- (10).

In this work, for the active region optimization of an MQW EEL at 1550 nm, Genetic algorithm has been applied. This global search heuristic algorithm searches for the global optimum for a multi-variable problem. From a batch of initial population (i.e., probable solutions) the genetic algorithm repeatedly modifies the population of individual solutions at each step (or generation). Over successive generations, the population evolves toward optimal solutions [16]. In this work, maximum value of gain is taken as the fitness function for the Genetic algorithm (GA). Minimization of the fitness function is done using the Genetic algorithm; therefore, negative value of the target functions G (material gain) is used.

For evaluating the fitness of the solutions the Pareto ranking technique is used. The implemented Genetic algorithm uses tournament type selection function with tournament size of 2 and adaptive feasible mutation function that generates adaptive directions depending on previous generations. The intermediate crossover function is used for reproduction with crossover fraction equal to 0.8 and crossover ratio equal to 1. After every 20 generations, forward migration is used with a migration fraction of 0.2. The initial populations are crated randomly within the given boundary conditions. For this optimization problem, the input vector consists of the physical device parameters such as layer thicknesses and temperature. Layer thicknesses are only allowed to have integer values in angstroms and kept within a reasonable boundary. The impact of these boundary conditions is discussed in the following section.

## III. Calculations and Results

For solving the Schrodinger's equation, shooting method in [11], has been used. Optimization of gain spectra has been performed using genetic algorithm toolbox of MATLAB taking the six variables mentioned in abstract. With 4 parallel schedulers of configurations manager of MATLAB, the whole simulation procedure took around a time range of 6 hours to 8 hours to get a satisfactory results. For the calculation of the quaternary parameter of $In_{0.72}Ga_{0.28}As_{0.82}P_{0.18}$, the following interpolation scheme is used [17].

$$Q(x,y) = xyB_{GaAs} + x(1-y)B_{GaP} + (1-x)yB_{InAs} + (1-x)(1-y)B_{InP} \quad (13)$$

For the calculation, corresponding binary material parameter are listed in following table [14][17][18].

TABLE I. PARAMETRS OF THE BINARIES

| Materials | $m_c/m_0$ | $m_{hh}/m_0$ | $m_{lh}/m_0$ | $a_0(A^0)$ |
|---|---|---|---|---|
| GaAs | 0.067 | 0.55 | 0.083 | 5.6533 |
| GaP | 0.114 | 0.52 | 0.17 | 5.4508 |
| InAs | 0.024 | 0.36 | 0.026 | 6.0583 |
| InP | 0.0793 | 0.69 | 0.11 | 5.8690 |

To calculate the band gap of the chosen InGaAsP material, the following equation has been used [19].

$$E_g = 1.35 + 0.642x - 1.101y + 0.758x^2 + 0.101y^2 - 0.159xy - 0.28x^2y + 0.109xy^2 \quad (14)$$

Calculation of the optimization process using the computational model has been performed considering 6 variables for asymmetric laser structure including 3 QW's widths, 2 barrier widths and temperature for fixed injection carrier density and laser cavity. As the number of obtained bands from shooting method is not fixed, for gain calculation 2-4 energy sub-bands are considered during gain calculation process. The optimization process has been performed using the genetic algorithm for 3 and 5 Qws for various injection carrier density based on [6], [15], [19].

From the 1st and 2nd gain model, a gain spectrum is calculated for 3 QWs ridge waveguide with $8\times10^{18}$ $cm^{-3}$ injection carrier density. In this performance, 3 conduction sub band has been considered. In Figure 3, the discontinuity of the

gain plot indicates the energy sub bands existence without considering the energy broadening. From the calculation, it has found that the first conduction sub band and the first valence sub band has been created at 40.35 meV above the conduction band edge and 3.73 meV (for heavy hole) and 19.6 meV (for light hole) below the valence band edge. The laser structure parameters for this calculation are taken from Table II.

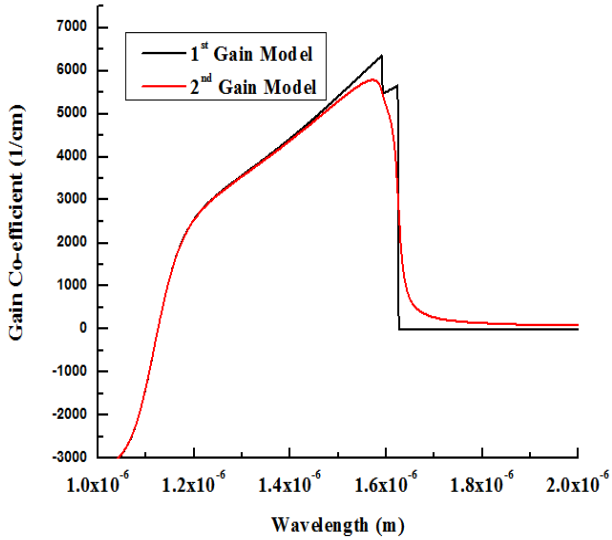



Figure 3: Material gain spectral of InGaAsP MQW EEL with 3 QWs for $8\times10^{18}$ $cm^{-3}$ carrier density. Spectra is calculated for 3 conduction sub band.

Next, from the 1st and 2nd gain model, another gain spectrum is obtained for 5 QWs ridge waveguide with same injection carrier density. In this calculation, 3 conduction sub band has been considered. In Figure 4, the discontinuity of the gain plot indicates the energy sub bands existence without considering the energy broadening. From the calculation, it has found that the first conduction sub band and the first valence sub band has been created at 44.75 meV above the conduction band edge and 24.58 meV (for heavy hole) and 19.85 meV (for light hole) below the valence band edge. The laser structure parameters for this calculation are taken from Table III.

The optimized results for the optimum material gain of the active region material $In_{0.72}Ga_{0.28}As_{0.82}P_{0.18}$ have been presented in the tables II and III. All the values are taken from the output of the genetic algorithm process for same QW material and barrier material. Optimized parameters values of an MQW EEL for 3 QWs are presented in table II for 3 different injection carrier densities.

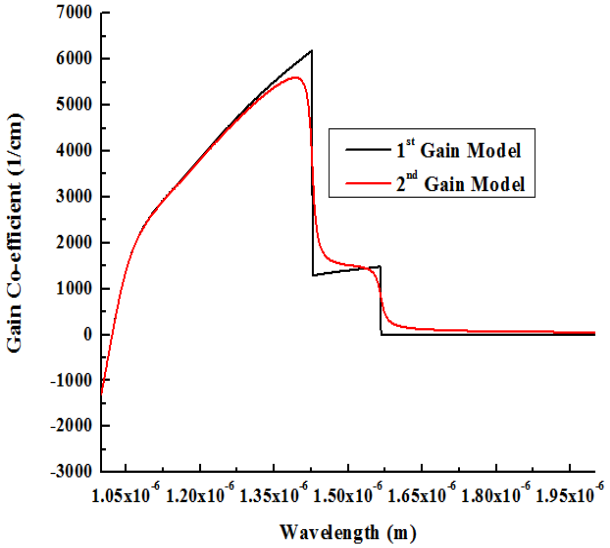



Figure 4: Material gain spectral of InGaAsP MQW EEL with 5 QWs for $8\times10^{18}$ $cm^{-3}$ carrier density. Spectra is calculated for 5 conduction sub band.

TABLE II. LASER PARAMETERS VALUE FOR OPTIMIZED GAIN FOR 3 QUANTUM WELLS

| **Laser parameters for 3 QWs** | | | |
|---|---|---|---|
| Injection Carrier Density ($cm^{-3}$) | $3\times10^{18}$ | $5\times10^{18}$ | $8\times10^{18}$ |
| Well width 1 (nm) | 3 | 2.8 | 4.1 |
| Well width 2 (nm) | 4.2 | 4.8 | 8.1 |
| Well width 3 (nm) | 6.9 | 8.4 | 5.7 |
| Barrier width 1 (nm) | 12.4 | 13.5 | 13.0 |
| Barrier width 2 (nm) | 12.4 | 13.6 | 13.1 |
| Temperature (K) | 276 | 278 | 276 |
| Optimized gain ($cm^{-1}$) | $3.1177\times10^{3}$ | $4.6141\times10^{3}$ | $5.9952\times10^{3}$ |

Table III represents the optimized values of the optimized design parameters of an MQW EEL for 5 QWs using the same well and barrier material for 3 different injection carrier densities.

TABLE III. LASER PARAMETERS VALUE FOR OPTIMIZED GAIN FOR 5 QUANTUM WELLS

| **Laser parameters for 5 QWs** | | | |
|---|---|---|---|
| Injection Carrier Density ($cm^{-3}$) | $3\times10^{18}$ | $5\times10^{18}$ | $8\times10^{18}$ |
| Well width 1 (nm) | 6.4 | 6.8 | 7.5 |
| Well width 2 (nm) | 6.7 | 7.3 | 6.5 |
| Well width 3 (nm) | 7.6 | 7.2 | 7.1 |
| Well width 4 (nm) | 5.7 | 7.1 | 6.9 |
| Well width 5 (nm) | 5.4 | 5.4 | 3.4 |
| Barrier width 1 (nm) | 11.7 | 11.2 | 13.0 |
| Barrier width 2 (nm) | 11.5 | 11.4 | 12.1 |
| Barrier width 3 (nm) | 11.7 | 11.4 | 13.0 |
| Barrier width 4 (nm) | 11.5 | 11.3 | 12.1 |
| Temperature (K) | 276 | 287 | 282 |
| Optimized gain ($cm^{-1}$) | $3.7668\times10^{3}$ | $4.854\times10^{3}$ | $5.8524\times10^{3}$ |

The optimized gain spectra for 3 QWs with different injection carrier density are illustrated in figure 5. It is observed from the figure that with high injection the gain co-efficient and the gain band width increase. This gain spectrum is calculated considering TE mode in laser cavity. With the increase in injection carrier density, threshold gain point is also shifted towards lower values.

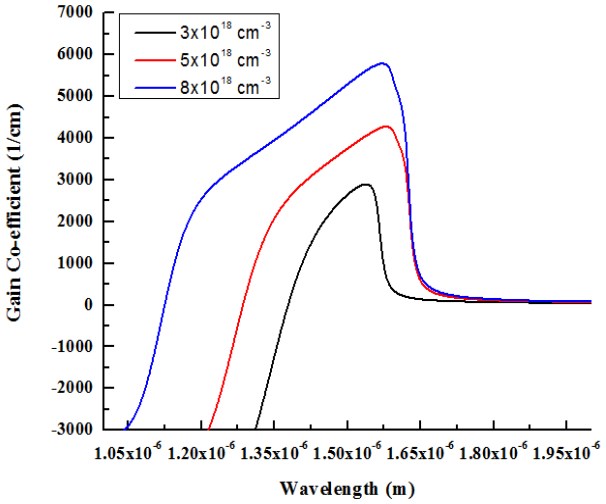



Figure 5: Gain spectra for different injection carrier density for 3 QWs with the optimized laser design parameters.

In case of 5 QWs, variations in gain properties are not regular as can be seen from the presented plots obtained after optimization which have been presented in figure 6. For a value of injection carrier density of $5\times10^{18}$/cm3, gain spectrum shows better results. For other carrier densities, the wavelengths of the design structures are shifted because of the lower transition energy. Gain bandwidths and threshold gain points show irregularity. Overall, the results obtained using this gain optimization method has been found to be satisfactory.

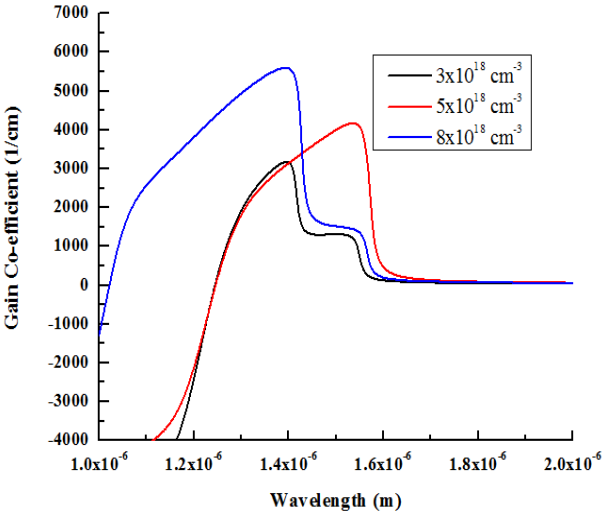



Figure 6: Gain spectra for different injection carrier density for 5 QWs with the optimized laser design parameters obtained from genetic algorithm.

## IV. CONCLUSIONS

For application in quantum wells of an edge emitting laser, an MATLAB based optimization technique using genetic algorithm has been presented which works well for the material gain of any quaternary compound semiconductor. In this technique, the optimization has been carried out for well width, barrier width and temperature. The obtained optimized parameters can be utilized in designing an MQW Edge Emitting semiconductor laser. This technique is made for unequal widths of quantum wells and barriers. The above mentioned computational model for the optimization work has been designed on the basis of well known analytical expression of material gain and the Schrodinger's equation. The computational model for optimization has been applied for designing an EEL using InGaAsP as the well and barrier materials for unequal widths of quantum wells and barriers. For the presented design, optimization has been performed considering 6 variables for 3 QWs at first for a 3 different injection carrier densities by using genetic algorithm. Next, the optimization has been performed considering 10 variables for 5 QWs with 3 different injection carrier density values also by using the same genetic algorithm. This optimization shows that with the increase of injection carrier density, material gain increases. In this work, the effects of threshold current and heat profile is avoided for simplicity. These effects limit the choice of optimum injection current density. Overall, the results obtained using this gain optimization method has been found to be satisfactory. The outcome of the design indicates clearly that the program works well.

## REFERENCES


[1] O. G. Okhotnikov, *Semiconductor Disk Lasers: Physics and Technology*, John Wiley & sons, New Jersey, 2010, pp. 1-5.

[2] G. Hunziker, W. Knop and C. Harder "Gain Measurement on One, Two, and Three Strained GaInP Quantum Well Laser Diodes", *Transactions of Quantum Electron*, vol. 30, no. 10, pp. 652-658, Oct. 1994.

[3] E. Yablonivitch and E. O. Kane, "Band Structure Engineering of Semiconductor Lasers for Optical Communications", *Journal of Lightwave Technology*, vol. 6, no. 8, pp. 581-587, Aug. 1988.

[4] J. Hecht, *The Laser Guidebook*, McGraw-Hill Companies, New York, 1993, pp. 355-360.

[5] P. S. Zory, *Quantum Well Lasers*, Academic Press Limited, London, 1993, pp. 28-41.

[6] L. A. Coldren and S. W. Corzine, *Diode Lasers and Photonic Integrated Circuits*, John Wiley & Sons, New Jersey, 1995, pp. 1-60, 79-84, 185-240.

[7] G. Hunziker, W. Knop, P. Unger and C. Harder, "Gain, Refractive index, Linewidth Enchancement Factor from Spontaneous Emission of Strained GaInP Quantum-Well Lasers", *IEEE Journal of Quantum Electronics*, vol. 31, no. 4, pp. 643-646, Apr. 1995.

[8] R. S. Quimby, *Photonics and Lasers An Introduction*, John Wiley & Sons, New Jersey, 2006, pp. 307-360.

[9] S. Lien Chuang, *Physics of Optoelectronic Devices*, Wiley Interscience , New York, 1995, pp. 394-470.

[10] S. DeLand (2011) Global Optimization Toolbox-MATLAB. [Online]. Available: http://www.mathworks.com/products/global-optimization

[11] P. Harrison, *Quantum Wells, Wires and Dots: Theoretical and Computational Physics of Semiconductor Nanostructures*, John Wiley & sons, West Sussex, 2009, pp. 17-70, 73-82.

[12] E. Zielinski, H. Schweizer, S. Hausser, R. Stuber, M. Pilkuhn, G. Weimann, "Systematics of Laser Operation in GaAs/AlGaAs Multiquantum Well Heterostructures", *IEEE Journal of Quantum Electronics*, vol. 23 , no. 6, pp. 969 – 976, 1987.

[13] R. W. Martin, H. L. Stormer, "On the Low Energy Tail of the Electron-Hole Drop Recombination Spectrum", *Solid State Communications*, vol. 22, no. 8, pp. 523-526, 1977.

[14] Sandra R. Selmic, Tso-Min Chou, Jiehping Sih, Jay B. Kirk, Art Mantie, Jerome K. Butler, David Bour and Gary A. Evans, "Design and characterization of 1.3-µm AlGaInAs-InP Multiple Quantum Well Lasers", *Journal of Quantum Electronics*, vol. 7, no. 2, pp. 340-349, March/April 2001.

[15] M. J. Connelly, *Semiconductor Optical Amplifiers*, Kluwer Academic Publisher, 2002, pp. 56-65.

[16] M. Mitchell, *An Introduction to Genetic Algorithms*, MIT Press, Massachusetts, 1998.

[17] S. Adachi, *Properties of Semiconductor Alloys Group-IV, III-V, II-VI Semiconductor*, John Wiley & Sons, New Jersey, 2009, pp. 133-214, 229-253, 277-286, 307-332.

[18] I. Vurgaftman, J. R. Meyer and L. R. Ram-Mohan, "Band Parameters for III-V Compound Semiconductors and Their Alloys", *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815-5875, Jun. 2001.

[19] J. C. L. Yong, J. M. Rorison and Ian H. White, "1.3- m Quantum-Well InGaAsP, AlGaInAs, and InGaAsN Laser Material Gain: A Theoretical Study", *IEEE Journal of Quantum Electronics*, vol. 38, no. 12, pp. 1553-1564, Dec. 2002.

[20] Y. Arakawa and A. Yariv, "Quantum Well Lasers - Gain Spectra and Dynamics", *Journal of Quantum Electronics*, vol. QE28, no. 09, pp. 1887-1899, Sep. 1986.

[21] D. Ahn and S. Chuang, "Optical Gain in a Strained-Layer Quantum Well Laser", *Journal of Quantum Electronics*, vol. 24, no. 12, pp. 2400-2406, Dec. 1988.

[22] H. Wada, K. Takemasa, T. Munakata, M. Kobayashi and T. Kamijoh, "Effects of Well Number on Temperature Characteristics in 1.3-µm AlGaInAs–InP Quantum-Well Lasers," *Journal of Quantum Electron*, vol. 5, no. 3, pp. 420-427, 1999.

[23] K. Prosyk, J. G. Simmons and J. D. Evans, " Well Number, Length, and Temperature Dependence of Efficiency and Loss in InGaAsP–InP Compressively Strained MQW Ridge Waveguide Lasers at 1.3 m", *IEEE Journal of Quantum Electronics*, vol. 33, no. 8, pp. 1360-1368, Aug. 1997.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Doping Dependency on Absorption Spectrum of Intraband Transition Based Photodetector

Sumit Narayan Saurov
Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka 1000, Bangladesh
sumit.saurov@gmail.com

*Abstract*—**Intraband transition based quantum cascaded photodetector is designed using GaN/AlN system. Carrier density distribution and dipole matrix element is calculated to determine the strength of optical transition. Doping effect on absorption spectrum has been investigated numerically.**

***Keywords—Photodetector, intersubband, dipole matrix element, absorption coefficient, photo-excitation.***

## I. INTRODUCTION

Photodetector is a special type of electrical device that converts optical signal directly into electrical signal. So, if light is given as an input to the photodetector than current or voltage is produced as an output. When radiating energy is absorbed by the electrons, they gain power and move into the upward energy levels. These excited charge carriers can be collected by placing electrodes across the device and then can be used to drive external loads. Photodetectors can play a vital role in renewable energy field as they can utilize solar energy or any other heat radiation. Day by day, demand of power is increasing beyond imagination. Fossil fuels are the major sources of power in present time. If this conventional energy source is exploited in this rate, there is a much higher probability of unavailability of this long time dependable energy source in the next century. So, it is high time to take step of using renewable energy sources as a substitute of conventional energy sources. But even today, technology regarding the extraction of renewable energy is not developed so much. Only an insignificant percentage of available renewable energy sources is used to generate electricity. Special effort of developing technology regarding this field should be taken as early as possible to meet the challenge of increasing demand.

Photodetector can be classified in two groups. One of these is interband photodetector. In this detector, when light is absorbed by the valence band electrons, they move into conduction band. That means transfer occurs between two bands. As a result free electron-hole pairs are created by photon absorption i.e., electron in the conduction band and hole in the valence band [1]. Another type of photodetector is intraband photodetector. Here, transition of electron occurs inside the same energy band. Holes have much higher effective mass than electrons. In case of a photodetector, faster carrier movement process is required to ensure larger current flow from the device. For this reason, only conduction band is important from design perspective [2]. Intraband i.e., intersubband transition based photodetectors are normally designed to operate at infrared region of electromagnetic spectrum. Till date, different matured semiconductor systems are used to operate as an intersubband photodetector. Due to low conduction band offset, they operate at far-infrared or at best at mid-infrared region [3]. Systems having larger conduction band offset can be used to design detector which is capable of operating in near-infrared range. In present work, we are mainly concentrating on the group-III nitride based quantum cascaded photodetector heterostructure.

Doping is a very important parameter for intersubband photodetector in a view of design corner. Transition of electrons between different eigen energy levels depends on the available carrier density in those corresponding energy levels. Doping effect should be analyzed carefully to maximize absorption capability as well as to ensure faster relaxation process.

The organization of this paper is as follows: In Sec. II, we will present the design procedure of quantum cascaded heterostructure based photodetector. In Sec. III, we will analyze and calculate carrier density of different energy levels. In Sec. IV, we will investigate the doping dependency on the absorption capability numerically. In Sec. V, we will draw conclusions.

## II. PHOTODETECTOR HETEROSTRUCTURE DESIGN

In this work, we have assumed that the GaN/AlN heterostructure is grown on Sapphire ($Al_2O_3$) substrate. As GaN/AlN epitaxial layers are not fully lattice matched with Sapphire substrate, the designed structure is strain compensated. Though AlN and GaN can exist in the cubic zincblende phase, only the hexagonal wurtzite phase is thermodynamically stable. In the wurtzite phase, group-III nitrides form a continuous alloy system with direct bandgaps [4]. So, the total conduction band offset of the material system can be exploited for intersubband transitions. In this work, we have considered the wurtzite phase of nitrides. Here, GaN serves as a well material and AlN serves as a barrier material. These samples can be created by using either plasma assisted molecular beam epitaxy (PAMBE) or metal organic vapor phase epitaxy (MOVPE) process. The different parameter values that are important to design GaN/AlN heterostructures are given in Table I [4]–[8].

TABLE I. MATERIAL PARAMETERS OF THE SYSTEM

| Parameters [unit] | GaN | AlN |
|---|---|---|
| Bandgap, $E_g$ [eV] | 3.438 | 6.25 |
| Effective mass, $m^*/m_0$ | 0.18 | 0.30 |
| Refractive index, $n$ | 2.335 | 2.154 |
| High Permittivity, $\varepsilon_\infty$ | 5.31 | 4.35 |
| Static Permittivity, $\varepsilon_0$ | 10.2 | 9.32 |
| LO-phonon energy, $\hbar\omega_{LO}$ [meV] | 91.2 | 99 |

The valley conduction band diagram and the associated wavefunctions of designed quantum cascade structure have been shown in Fig. 1. We adjust the thickness of the barrier and well materials so that a broadband light absorption is possible. Here, we have doped in three quantum wells (QW). Optical transition will take place inside these three quantum wells. As a result, absorption of energy at different range is possible.

To increase the current and hence the efficiency of the quantum cascade structures, as the carriers are photo-excited to the upper states, they should be available at the lower states of the next stage at a very short time. The fast transport of the excited carriers can be achieved by employing scattering due to Longitudinal Optical (LO) phonons. We design the structures so that the path through which the excited carriers relax is divided into steps of energy levels which differ in energy comparable to the LO phonon energy of the material system. The LO phonon energy of GaN material is 91.2 meV. In this structure, the energy spacing between any two consecutive levels in the relaxation path is between 80 to 110 meV.

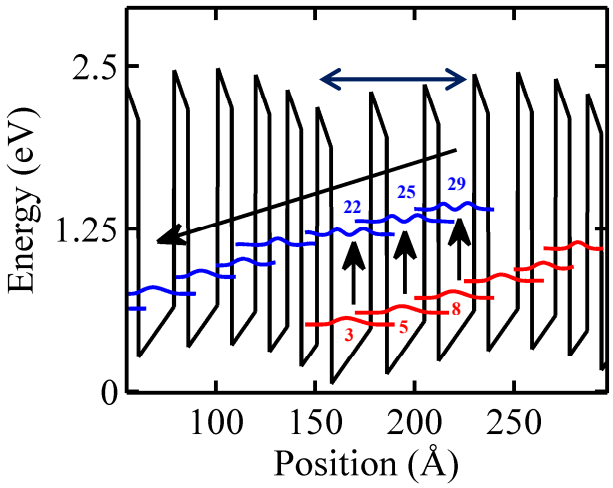



Figure 1. Energy wavefunctions with squared envelope functions of the designed structure. The layer thicknesses (shown in Angstrom unit) of this designed structure are **7**/20/**7**/19/**7**/18/**7**/15/**7**/12/**7**/**9**/**7**/**8**/**7**. The numbers in bold fonts are AlN layers (barrier material) and the numbers in normal fonts are GaN (well material) layers. The underlined layer has been n-type doped. The photo-excitation of the carriers from one period (red coloured) to the next period (blue coloured) i.e., photon absorption process has been shown by the vertical arrow. The carrier relaxation from the upper energy levels to lower energy levels through subsequent stages i.e., photo-carrier collection process has been shown by the inclined arrow. The electric field F = 50 kV/cm across the structure corresponds to the operating photovoltage point. Doped Quantum Wells (DQW) are shown by a two-headed arrow and corresponding eigen energy levels related to charge carrier transition process are shown.

Therefore, carriers relax from the upper states of an active region to the lower states of the next active region at a faster rate.

We have designed the structure in such a way, that higher energy transition is possible for excited charge carriers. We have also ensured stronger transition after absorbing photons. Strength of an optical transition between two subbands depends on the dipole matrix element between them. Dipole matrix element does not depend on the doping. It depends on the electric dipole moment associated with the two states. In general it is a complex vector quantity that includes the phase factors associated with the two states. Its direction gives the polarization of the transition, which determines how the system will interact with an electromagnetic wave of a given polarization. The dipole matrix element between two states $i$ (initial) and $f$ (final) is given by [9]

$$z_{ij} = \frac{\hbar}{2(E_j - E_i)} \langle \psi_j \mid p_z \frac{1}{m^*(E_j, z)} + \frac{1}{m^*(E_i, z)} p_z \mid \psi_i \rangle , \quad (1)$$

where $m^*$ is the energy dependent effective mass and $p_z$ is the momentum operator, $E$ is the eigen energy level and $\Psi$ is the probability function describing the availability of carrier at the corresponding state. Calculated dipole matrix element of different energy levels for these three doped wells of our designed structure has been shown in Table II.

Higher value of dipole matrix element denotes the stronger interaction between subbands. So, stronger transition is possible between the lower energy levels and upper energy levels for our designed structure as dipole matrix elements are stronger for these corresponding energy levels. As a result, stronger absorption of optical illumination is possible.

## III. CARRIER DENSITY CALCULATION

Number of available carriers in a definite energy state depends on the doping as well as density of states. In case of a quantum well, there are only two degrees of freedom. As a result a two-dimensional electron gas is created inside the well. Density of states for a single subband in a quantum well can be calculated numerically using [10]

$$\rho^{2D}(E) = \sum_{i=1}^{n} \frac{m^*}{\pi\hbar^2} \Theta(E - E_i), \quad (2)$$

where, $\Theta$ is the unit step function and $n$ is total confined states within the quantum well system.

We have assumed that, before photo-excitation system is in thermal equilibrium. So, carriers are distributed among

TABLE II. DIPOLE MATRIX ELEMENT

| Doped QW | Initial Energy Level, $i$ | Final Energy Level, $f$ | Dipole Matrix element, $z_{ij}$ (Å) |
|---|---|---|---|
| 1 | 3 | 22 | 1.2761 |
| 2 | 5 | 25 | 0.3628 |
| 3 | 8 | 29 | 0.3525 |

different energy levels according to Fermi-Dirac probability function as [10]

$$f(\varepsilon) = \frac{1}{e^{\frac{\varepsilon - E_F}{k_B T}} + 1}, \tag{3}$$

where $k_B$ is the Boltzmann constant, $T$ is the device temperature. As the photodetector operates in the room temperature, $T$ is taken as 300 K. Here, $E_F$ is the quasi-Fermi energy level which describes the carrier population within a subband. Quasi-Fermi distribution of carriers depends on the doping concentration. For a quantum well, we need to know the two-dimensional carrier density which is calculated by multiplying the doping density of the well by the thickness of the doped wells. If $n_{2D}$ is the total carrier density than quasi-Fermi level, $E_F$ can be calculated as [10]

$$\begin{aligned} n_{2D} = \frac{m^* k_B T}{\pi \hbar^2} \{ [\frac{E_{\max} - E_F}{k_B T} - \ln(1 + e^{\frac{E_{\max} - E_F}{k_B T}})] \\ - [\frac{E_{\min} - E_F}{k_B T} - \ln(1 + e^{\frac{E_{\min} - E_F}{k_B T}})] \}, \end{aligned} \tag{4}$$

where $E_{min}$ is the subband minima and $E_{max}$ is the top energy level of the quantum well.

Carrier density of energy levels where optical transition takes place plays a vital role to determine the absorption characteristics. As photon is absorbed by the electrons of the lower energy levels, there should be higher availability of electrons to utilize this optical power. After photon absorption, these electrons move to the upper energy level. So, there should be available empty states where photo-excited carriers will be moved. Lower carrier density at upper energy levels before photo-excitation is required to fulfill this requirement. As a result, stronger absorption directly depends on the carrier density difference between the two energy levels taking part in transition. At higher energy levels, there is less number of available carriers as carriers are mostly distributed in the lower energy levels. In our designed structure, absorption takes place in higher energy range. So, it can absorb electromagnetic radiation effectively and strongly. Two-dimensional carrier density of a state $i$ with energy $E_i$ can be calculated as [11]

$$N_i = \frac{m^* k_B T}{\pi \hbar^2} \ln(1 + e^{\frac{E_F - E_i}{k_B T}}). \tag{5}$$

To calculate the effect of doping on carrier density distribution, we have considered two different doping. One is lower doping of $N_{dop}$= $10^{17}$/m$^3$ and other one is high doping with $N_{dop}$= $10^{26}$/m$^3$. In present work, we will study and analyze the effect of these two different doping levels on the carrier density distribution as well as on absorption capability. Two-dimensional carrier density at different energy levels of three doped well in designed structure has been shown in Table III.

So, by higher doping, larger carrier density difference between the energy levels can be achieved. But excess heavy doping can cause random tunneling through barrier. Different secondary leakage mechanisms will be significant resulting in uncertain behavior of photo-detection process. Moreover, excess carrier density requires greater time to relax to lower energy levels resulting slower carrier relaxation process. As a result, higher power generation process will be interrupted. So, for effective design moderate doping should be taken. Normally group-III nitride based quantum well heterostructure is doped at a range of $1\times10^{23}$/m$^3$ to $1\times10^{26}$/m$^3$ [4].

TABLE III. CARRIER DENSITIES

| Eigen Energy Level | 2D Carrier Density (no. / m$^2$) | |
|---|---|---|
| | $N_{dop} = 10^{17}/m^3$ | $N_{dop} = 10^{26}/m^3$ |
| 3 | $4.60 \times 10^8$ | $2.89 \times 10^{17}$ |
| 5 | $2.48 \times 10^7$ | $2.32 \times 10^{17}$ |
| 8 | $4.26 \times 10^5$ | $1.52 \times 10^{17}$ |
| 22 | 0 | $1.15 \times 10^{10}$ |
| 25 | 0 | $2.69 \times 10^8$ |
| 29 | 0 | $6.68 \times 10^5$ |

## IV. ABSORPTION SPECTRUM

Absorption coefficient ($\alpha$) is the measurement of how well a device absorbs energy from the optical source. Higher value of this coefficient denotes higher absorption ability of the structure. In our work, we have considered two-dimensional absorption coefficient ($\alpha_{2D}$) that means absorption coefficient is normalized to the stage inverse thickness. The absorption spectrum can be calculated using [12]

$$\begin{aligned} \alpha_{2D}(\hbar\omega) = \sum_{i \in gl} \sum_{j \in el} \frac{q^2 z_{ij}^2 \omega}{nc\varepsilon_0} (N_i - N_j) \\ \cdot \frac{\Gamma_{ij}/2}{(E_j - E_i - \hbar\omega)^2 + (\Gamma_{ij}/2)^2}, \end{aligned} \tag{6}$$

where "*gl*" denotes the ground energy levels and "*el*" denotes the upper excited energy levels. The ground energy levels are at the bottom of vertical arrow and the upper excited energy levels are on the top of the vertical arrows in Fig. 1. Here, $\omega$ is the radian frequency, $n$ is the refractive index, $c$ is the velocity of the light, $q$ is the charge of an electron, $\varepsilon_0$ is the static dielectric constant of quantum well material, $z_{ij}$ is the dipole matrix element, $E_j$ is the energy value of upper level and $E_i$ is the energy value of lower level. The line width of absorption spectrum, $\Gamma_{ij}$, has been assumed to be 11% of the transition energy $E_j - E_i$ [4].

So, absorption coefficient is directly proportional to the carrier density difference, $N_i - N_j$ as shown in Eq. (6). Stronger absorption coefficient along with wider bandwidth denotes the capability of absorbing electromagnetic optical radiation effectively by the photodetector structure. These requirements can be obtained by creating carrier density difference between the corresponding energy levels i.e., by changing doping level.

Absorption coefficient is plotted against absorbed photon energy for $N_{dop}$ = $10^{26}$/m$^3$ in Fig. 2. There are two major peaks in the absorption spectrum. They are at 596 meV and 704 meV. Peak absorption coefficient is almost $3\times10^{-5}$ which is at

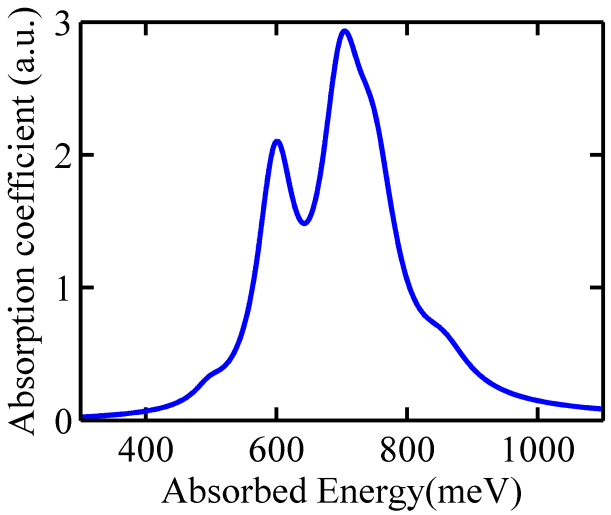



Figure 2. Absorption spectrum for $N_{dop} = 10^{26}/m^3$. Absorption coefficient is normalized with respect to $1\times10^{-5}$.

704 meV. So, our designed structure can absorb from almost 550 meV to 850 meV of available solar energy as shown in Fig. 2. Linewidth i.e., bandwidth of absorption spectrum is considered as Full-Width at Half-Maximum (FWHM). Effective bandwidth is almost 180 meV for this doping level.

Absorption coefficient is plotted against absorbed photon energy for $N_{dop} = 10^{17}/m^3$ in Fig. 3. Absorption peaks are in the same energy level. It does not depend on the doping concentration as peak absorption coefficient occurs at the energy difference of two levels where transition takes place as shown in Eq. (6). But, peak absorption coefficient is almost $2.5\times10^{-15}$ for this low doped structure. So, absorption capability decreases in a large scale. To absorb photons, enough electrons should be available at lower energy levels. Otherwise, only a small portion of available photon energy will be used. In case of low doping, there are very few charge carriers available in ground states. So, energy transferred by these charge carriers is also small. Linewidth for this doping is

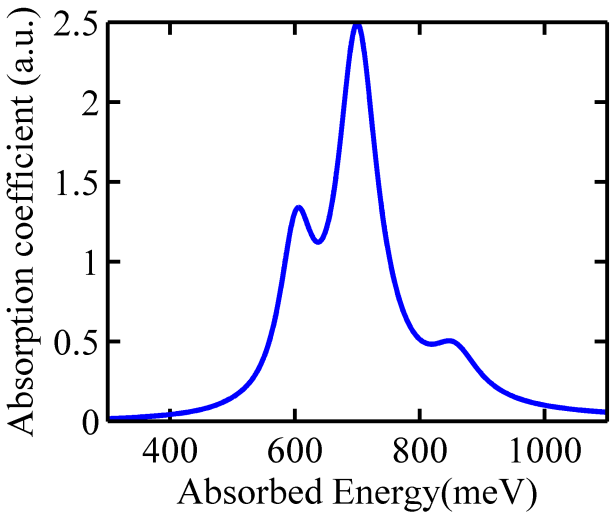



Figure 3. Absorption spectrum for $N_{dop} = 10^{17}/m^3$. Absorption coefficient is normalized with respect to $1\times10^{-15}$.

only about 90 meV. So, bandwidth also decreases for low doping. As bandwidth is a measurement of effective absorbed energy range, lower doping results in ineffective design of photodetector. Absorbed optical power leads to the generation of photocurrent which is used to drive external loads. So, narrow bandwidth of a structure denotes lower generation of current. As a result, less power will be generated from this designed structure. So, low doping leads to lower efficiency quantum well photodetector. So, doping should be increased at a certain higher level to absorb more electro-magnetic radiation as well as to avoid secondary leakage effects.

## V. Conclusions

In summary, we have shown that, absorption spectrum of a quantum well cascaded heterostructure is strongly dependent on the doping density. Higher doping leads to stronger absorption coefficient and wider bandwidth in comparison to lower doping. As excess heavy doping leads to the rise of different non-linear secondary effects and random tunneling probability, moderate doping should be introduced to design an effective photodetector.

## Acknowledgment

The author would like to acknowledge the support from the department of Electrical and Electronic Engineering (EEE) of Bangladesh University Engineering and Technology (BUET) in carrying out this work.

## References

[1] S. O. Kasap, *Optoelectronics and Photonics: Principles and Practices*, New Delhi: Pearson Education Inc., 2009.

[2] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed., New Jersy: John Wiley & Sons Inc., 2007.

[3] R. Paiella, *Intersubband Transitions in Quantum Structures*, New York: McGraw-Hill, 2006.

[4] E. Baumann, "Near infrared intersubband absorption and photovoltaic detection in GaN/AlN multi quantum well structures," Ph.D. dissertation, Department of Physics, University of Neuchátel, Neuchátel, Switzerland, Sep. 2007.

[5] M. E. Levinshtein, S. L. Rumyantsev, and M. S. Shur, *Properties of advanced semiconductor materials: GaN, AlN, InN, BN, SiC, SiGe*, New York: John Wiley and Sons, 2001.

[6] I. Vurgaftman and J. R. Meyer, "Band parameters for nitrogen-containing semiconductors," *J. Appl. Phys.*, vol. 94, pp. 3675-3696, Jun. 2003.

[7] W. J. Moore, J. A. Freitas, R. T. Holm, O. Kovalenkov, and V. Dmitriev, "Infrared dielectric function of wurtzite Aluminum Nitride," *Appl. Phys. Lett.*, vol. 86, pp. 141912-1—141912-3, Apr. 2005.

[8] T. Azuhata, T. Sota, K. Suzuki, and S. Nakamura, "Polarized Raman-spectra in GaN," *J.Phys.: Condens. Matter*, vol. 7, pp. 129-133, Mar. 1995.

[9] H. C. Liu and F. Capasso, *Intersubband Transitions in Quantum Wells: Semiconductors and Semimetals*, San Diego: Academic Press, 2000.

[10] P. Harrison, *Quantum Wells, Wires and Dots: Theoretical and Computational Physics* , Chichester: John Wiley and Sons, 1999.

[11] D. Jena, "Bandstructure in Low-Dimensional Structures," Department of Electrical Engineering, University of Notre Dame, 2004.

[12] M. Helm, *Intersubband transitions in quantum wells: physics and device applications I*, San Diego: Academic Press, 2000.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Effects of Interlayers in Threading Dislocation Reduction of Step-graded InGaN Heteroepitaxy

Shifa Khatun, Syeda Arza Sanober, Md. Arafat Hossain*, and Md. Rafiqul Islam
[1]Department of Electrical and Electronic Engineering, Khulna University of Engineering & Technology, Khulna 9203, Bangladesh
* E-mail: arafat_kuet_eee@yahoo.com

*Abstract*—**This paper investigates the effects of interlayers to reduce the threading dislocation density at the top surface of the epilayer in step-graded InGaN heteroepitaxy. The reaction kinetic coefficients are considered as key parameter in the dislocation reduction and calculated analytically. The reaction model has been solved numerically with different number of interlayer. A significant improvement of epilayer quality with extremely low threading dislocation densities have been evaluated with increasing the interlayer up to 4 where 8% In composition difference used for each step-graded interlayers.**

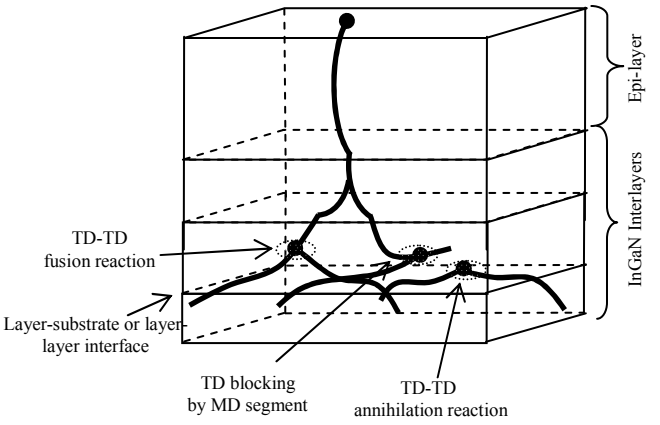
***Keywords—Threading dislocation, reaction kinetic coeficient, InGaN, step-graded interlayer.***

## I. INTRODUCTION

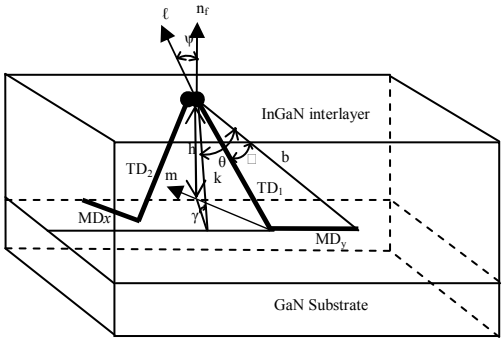
In recent years group-III Nitrides with their related alloys have been intensively studied with remarkable breakthrough in the growth of multilayer structures for device applications. Among them InN-based alloys are predicted to show high mobility, high absorption coefficient, long lifetime of charge carriers, low effective mass of electrons and holes and superior resistance against ingredient damages [1], [2]. Especially, the ternary nitride alloy InGaN is predicted as a promising candidate for multi-junction (MJ) tandem solar cells with high conversion efficiency [3], [4]. In order to realize a tandem solar cell based on InGaN, it is essential to grow InGaN alloy with high indium (*In*) content [4], [5]. As a result of increasing the *In* composition, it poses many challenges in controlling defect densities due to large difference in lattice constant between InN and GaN (11%). Therefore, despite the tremendous success, these technologies still suffer mostly from the lack of perfect substrates and have to cope with strongly mismatched heteroepitaxial growth. Consequently, a large lattice and thermal expansion coefficient mismatches between the layer and substrate results a large number of threading dislocations (TDs) in the epilayer. These dislocations act as scattering centres and midband gap states, which function as recombination centres and reducing minority carrier lifetime [6]. Hence the reduction of dislocation density is a budding issue for InGaN heteroepitaxial devices. Efforts have been carried out in several ways to reduce the TDs [7]. The compositional grading is considered as one of the useful techniques which can be done in several ways such as step grading, linear or nonlinear grading etc [8], [9]. In almost every works it has been reported that, the key strategy to reduce threading dislocation density is to initiate annihilation and fusion reaction among them during their glide process. In this process, the inclinations of TDs at layer-substrate or layer-layer interface are responsible for interaction between them and hence results in reduction by reaction [10]. Consequently, linear or nonlinear grading shows poor performance for reducing the TDs in the epilayer due to the less interaction probability. In step-graded heteroepitaxy growth of InGaN, due to the step increase of In composition in the epitaxial layers, a multiple step inclination of TDs at each interface will occur. This promotes the reaction among the TDs at each interlayer. Few experimental works have been carried out on step-graded interlayer used for different material system such as InGaN/GaN, SiGe/Si, InGaAs/GaAs [8], [11]. The results confirmed that, this technology will be a promising solution for reduction of threading dislocation during growth of heteroepitaxy. A reaction model developed in our previous work for step-graded InGaN heteroepitaxy has also predicted the superiority of this technique [12]. In order to realize the reduction of dislocation it is necessary to properly calculate the values of reaction kinetic coefficient at each interlayer. A proper evaluation of kinetic coefficients at each interlayers need to carry out for a detail investigation of reaction equations of step-graded InGaN heteroepitaxy. A one step more analysis of the model with different number of interlayer is urgently needed for further optimization of the structure.

## II. MATHEMATICAL MODELING

A step-graded interlayer technique is one, where the *In* composition in InGaN increases step wise in each interlayer up to the epilayer. During the growth of each InGaN interlayer, the TDs generated from the misfit dislocations (MDs) or propagated from the previous layer experience a misfit strain due to the lattice and thermal expansion coefficient mismatch with substrate or previous interlayer. As a result, the TDs bend at each interlayer depending on the applied strain and so the *In* composition. Due to this bending, the TDs interact with each other resulting reaction once reaching the interaction radius. The different strategy of threading dislocation density reduction such as annihilation, fusion and scattering reactions are shown in Fig. 1(a) for a step-graded structure. The geometry of two TDs (type *x* and *y*) connected with their corresponding MD segments and their relative motion with film thickness in a particular interlayer is shown in Fig. 1(b). All parameters indicated in figure are defined elsewhere [12], [13]. Some of the physical parameters are important to evaluate the




(a)




(b)

Figure 1. Geometry of dislocations in step-graded structure. (a) Different strategy of reaction and (b) Different parameters of two growing up threading dislocation connected with MD segments.

reaction kinetic coefficient, $K_{xy}$ between these two types of dislocations and evaluated from the geometry according to the following approach.

$$\psi = \cos^{-1}\left(\ell.\mathrm{n_f}\right) \tag{01}$$

$$\gamma = \cos^{-1}\left(\mathrm{n_g}.\mathrm{n_f}\right) \tag{02}$$

$$\theta = \cos^{-1}\left(\frac{\mathrm{b}.\mathrm{n_f}}{|\mathrm{b}|\sin\gamma}\right) \tag{03}$$

Now using these angles for a particular interlayer $K_{xy}$ can be evaluated using the following equation.

$$\mathrm{K_{xy}} = 2\mathrm{r_I}\left|\frac{\sqrt{1-\left(\ell_{\mathrm{x}}.\mathrm{n_f}\right)^2}}{\left|\ell_{\mathrm{x}}.\mathrm{n_f}\right|}\frac{\left|\ell_{\mathrm{x}}-\left(\ell_{\mathrm{x}}.\mathrm{n_f}\right)\mathrm{n_f}\right|}{\left|\ell_{\mathrm{x}}-\left(\ell_{\mathrm{x}}.\mathrm{n_f}\right)\mathrm{n_f}\right|} - \frac{\sqrt{1-\left(\ell_{\mathrm{y}}.\mathrm{n_f}\right)^2}}{\left|\ell_{\mathrm{y}}.\mathrm{n_f}\right|}\frac{\left|\ell_{\mathrm{y}}-\left(\ell_{\mathrm{y}}.\mathrm{n_f}\right)\mathrm{n_f}\right|}{\left|\ell_{\mathrm{y}}-\left(\ell_{\mathrm{y}}.\mathrm{n_f}\right)\mathrm{n_f}\right|}\right| \tag{04}$$

Where $l$ and $n_f$ represent the dislocation line direction and normal to the film-substrate or layer-layer interface.

$$\ell = \frac{\sin\left(\theta-\varphi\right)}{\sin\theta}\frac{\mathrm{b}}{|\mathrm{b}|} + \frac{\sin\varphi}{\sin\theta}\mathrm{k} \tag{05}$$

$$\mathrm{k} = \frac{\mathrm{n_g}\times\left(\mathrm{n_f}\times\mathrm{n_g}\right)}{\sin\gamma} \tag{06}$$

Now these coefficients are used to develop a generalized governing equation of TD reduction considering all possible reaction for the step-graded structure [12].

$$\left(\frac{d\rho_x}{dh}\right)_{i^{th}\text{layer}} = \left(-\sum_{y}K_{xy}\rho_x\rho_y + \sum_{l}\sum_{n}K_{ln}\rho_l\rho_n - \sum_{m}\rho_x\rho_m\right)_{i^{th}\text{laye}} \tag{07}$$

Here $i$ is the index of interlayer. The possible annihilation, fusion reactions between each pair of TDs in each step-graded InGaN interlayer, with characteristic radii $r_A$ and $r_F$ have been considered for $r_I$. There are 20 unique families of dislocations in InGaN according to their burger vector, b and line directions [14]. Among them, two for screw dislocation ($+c$ and $-c$), six different edge dislocations ($+a_1$, $-a_1$, $+a_2$, $-a_2$, $+a_3$ and $-a_3$) and twelve for different mixed character dislocations ($+a_1+c$, $+a_1-c$, $-a_1+c$, $-a_1-c$, $+a_2+c$, $+a_2-c$, $-a_2+c$, $-a_2-c$, $+a_3+c$, $+a_3-c$, $-a_3+c$, and $-a_3-c$). According to the reaction table developed for wurzite InGaN, TDs from family 1 can only have annihilation and fusion reactions with TDs from family 9, 10, 11, 12, 15, and 16 that reduce the TDs. The first term of Eq. (07) can be expressed for the $i^{th}$ InGaN interlayer as

$$\left(\frac{d\rho_1}{dh}\right)_{i^{th}\text{layer}} = \begin{pmatrix}-K_{1,9}\rho_1\rho_9 - K_{1,10}\rho_1\rho_{10} - K_{1,11}\rho_1\rho_{11} - K_{1,12}\rho_1\rho_{12}\\ -K_{1,15}\rho_1\rho_{15} - K_{1,16}\rho_1\rho_{16}\end{pmatrix}_{i^{th}\text{layer}} \tag{08}$$

On the other hand it follows from the analysis of the table that, TDs from family 1 may be generated within the interlayer only as the product of fusion reactions between TDs from the following pairs of families 7-8, 7-14, 7-18, 7-20, 8-13, 8-17, 8-19, 13-18 and 14-17. Therefore, for production of TDs from family 1 as a result of fusion, one can write

$$\left(\frac{d\rho_1}{dh}\right)_{i^{th}\text{layer}} = \begin{pmatrix}+2K_{7,8}\rho_7\rho_8 + K_{7,14}\rho_7\rho_{14} + K_{7,18}\rho_7\rho_{18} + K_{7,20}\rho_7\rho_{20} + K_{8,13}\rho_8\rho_{13}\\ +K_{8,17}\rho_8\rho_{17} + K_{8,19}\rho_8\rho_{19} + K_{13,18}\rho_{13}\rho_{18} + K_{14,17}\rho_{14}\rho_{17}\end{pmatrix}_{i^{th}\text{laye}} \tag{09}$$

Finally, the threading dislocation of family 1 could be blocked by edge type MD in the interlayer. Therefore, TD reduction for type 1 due to this effect can be written as

$$\left(\frac{d\rho_1}{dh}\right)_{i^{th}\text{layer}} = \left(-\rho_1\rho_m\right)_{i^{th}\text{layer}} \tag{10}$$

Combining (08), (09) and (10) we can write the governing differential equation for the density of TD family 1 in the $i^{th}$ interlayer of step-graded InGaN. Similarly, 20 nonlinear first order differential equations can be formed with corresponding burger vector and possible reaction for the step-graded structure.

## III. SIMULATION RESULTS AND DISCUSSION

The set nonlinear differential equations regarding the TD densities have been solved by using Euler's method. The initial

values of the TDs are set as the boundary conditions in the simulation and found elsewhere [15]. In our previous work we have set the misfit dislocation density as the initial value which was more reasonable for a complete model considering both dislocations. Here we have evaluated only the TD densities which are generally propagated to the epilayer are most crucial for optoelectronic devices and our main objective is to optimize the model for step-graded structure. The calculations have been done for a step-graded InGaN with 4 interlayers each of 300 nm and a final epilayer of 1.5 μm structure with final In composition, $x$ = 0.4. A multiple values of reaction kinetic coefficients among the TDs are evaluated considering the parameters at each interlayer and the annihilation and fusion reaction radius of 1000 Å. The $K_{xy}$ values at the starting interlayer in six different structures with different number of interlayer have been listed in Table I. The calculation have been done by considering the angle between surface normal and glide plane normal, the angle between surface normal and burger vector.

TABLE I. REACTION KINETIC COEFFICIENTS FOR FIRST INTERLAYER

| $K_{xy}$ Type | Values at different interlayer number | | | | | |
|---|---|---|---|---|---|---|
| | *1* | *2* | *3* | *4* | *5* | *6* |
| Edge-mixed | 0.0212 | 0.0856 | 0.1971 | 0.3697 | 0.6569 | 1.1898 |
| Mixed-mixed | 0.0019 | 0.0082 | 0.0202 | 0.0427 | 0.0984 | 0.2401 |
| Screw-mixed | 0.0193 | 0.0774 | 0.1769 | 0.3270 | 0.5585 | 1.1308 |

A multiple step-inclination at each interlayer in the step-graded structure increases the angle and subsequently the reaction probability among the TDs. As a result the values of $K_{xy}$ increase with interlayer number and the values indicate that tendency of mixed types interactions with others is higher than by themselves.

Though we have investigated the behaviour of each type of TDs, Fig. 2 only presents the behaviour of mixed type TD

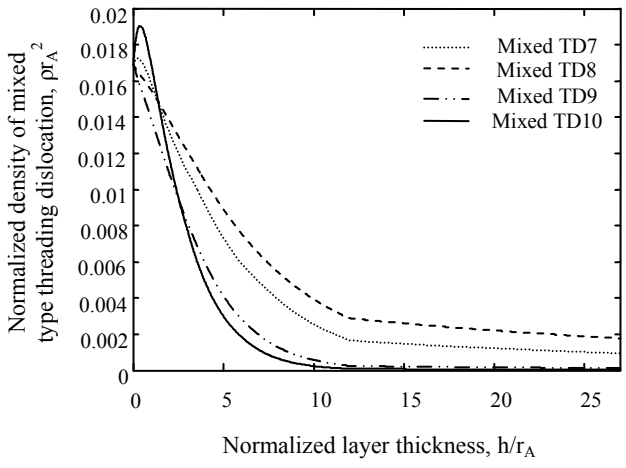



Figure 2: The variation of mixed TD density with layer thickness in step-graded InGaN heteroepitaxy

density with some specific numbers for 2.7 μm step-graded InGaN using 4 interlayers. The mixed type TD decreases monotonically with increasing film thickness and these decreases are more quickly than any other types. The higher rate of decreasing mixed type TDs is due to the more relative motion of these TDs even without any glide motion. The nature of screw type TDs are slightly different from the others. Both type of TDs increase with considerable thickness. The reason is, though the screw dislocations are considered very low initial values they can be the product of other reaction and allowed to participate in reaction with other TDs. As a result, both the screw-mixed reaction producing vertical edge dislocation and the edge-mixed reaction producing vertical screw dislocations. Therefore the mixed TD density becomes so low as to prevent the further interaction among TDs, with the end result of saturation in TD density.

The improvement in TD density in the top surface of the epilayer has been understood from the comparison in Fig. 3 which shows the total TD densities for the step-graded and without graded structure and evaluated for the same initial density. The TD density has extremely lower values at the top surface of the epilayer in the step-graded structure than that of the without graded layer. At the top surface of the epilayer a total TD density of $1.75\times10^{10}$ $cm^{-2}$ and $4.07\times10^{10}$ $cm^{-2}$ can be found for the step-graded and without graded structure with a 58.4% and 3.28% reduction respectively. The step inclination of TD at each interface promotes the reaction between them in case of step-graded layers. Furthermore, a reuse of previous dislocation relax the strain in the upper layer instead of generation of new dislocation in the step-graded structure is also responsible for improved epilayer quality of the graded structure.

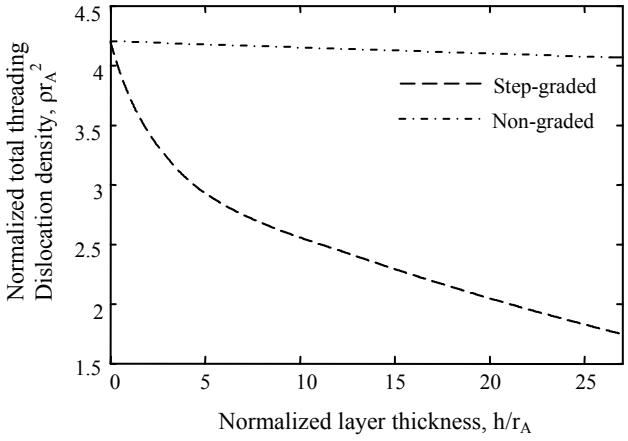
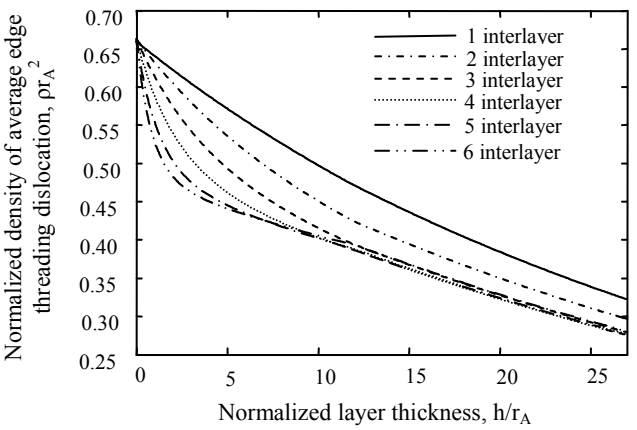
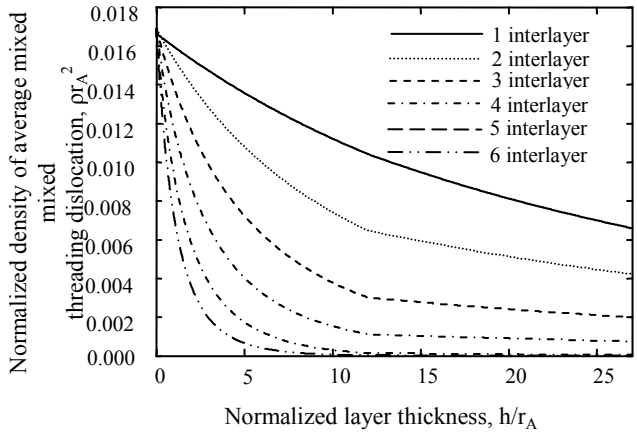



Figure 3. Comparison the total TD densities for the step graded and without graded structure with normalized values.

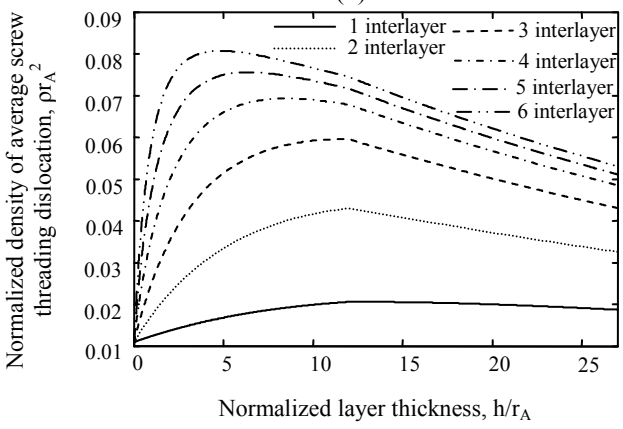
In order to realize the effects of interlayer on different types of TDs reduction, we have simulated the model with increasing the number of interlayers up to 6 where 5.7% *In* composition difference have been used shown in Fig 4. It is clear from the figures that the mixed type TDs decreases monotonically with




(a)




(b)




(c)

Figure 4. The effects of increasing number of interlayers in average dislocations densities of- (a) edge (b) mixed and (c) screw type.

a continuous addition of interlayer instead of being saturation as observed in other types. This is due to its increase of $K_{xy}$ values with interlayers. Since Fig. 4(a)-(c) report the significant reduction of average TDs up to different number of interlayer we are probably unable to select the optimal number which could be suggested for the proposed technique. A combination of all these results may indicate the optimal point. In this

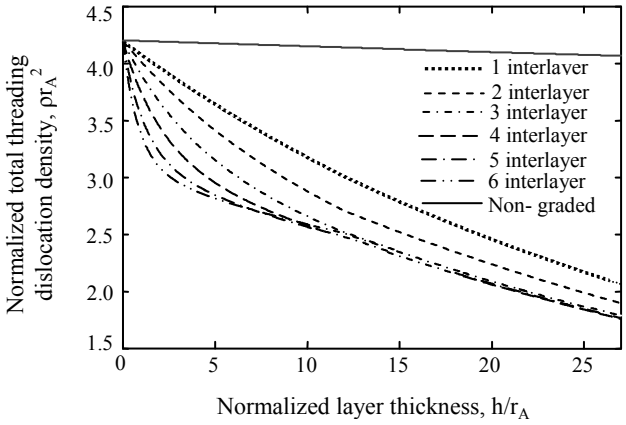



Figure 5. The effects of increasing number of interlayers on total threading dislocations density of a step-graded structure.

purpose, we have evaluated the total TDs densities with different number of interlayers and shown in Fig. 5. Now, it is clear from the figure that, the threading dislocation densities are significantly decreased with the insertion of interlayers up to 4 where 8% *In* composition is used for each interlayer. However, the epilayer is further improved for 5 interlayers with very low increasing tendency. The insertion of more and more interlayer increases the experimental complexity and there will introduce interfacial dislocations in each layer. These results are consistent with the results of improved InGaN epilayer quality using step-graded interlayer [8].

## IV. CONCLUSION

From the above realization it can be summarized that, step-graded interlayer technique with an optimized structure can be a promising solution for the improved epilayer quality with extremely low threading dislocation density. In this work, the step-graded interlayer technique for wurzite InGaN heteroepitaxy has been investigated and optimized through mathematical modeling and numerical simulation. The model verification has been confirmed by the judgment with published experimental works. Another optimistic site of this simulation methodology is its applicability in other wurzite heteroepitaxial material for future generation high performance devices.

## REFERENCES


[1] Y. Nanishi, Y. Saito and T. Yamaguchi, "RF-Molecular beam epitaxy growth and properties of InN and related alloys," *Jpn J. Appl. Phys.*, vol. 42, no. 5A, pp. 2549-2559, May 2003.

[2] J. Wu, W. Walukiewicz, K. M. Yu, W. Shan, J. W. Ager III, E. E. Haller, H. Lu, W. J. Schaff, W. K. Metzger and S. Kurtz, "Superior radiation resistance of $In_{1-x}Ga_xN$ alloys: Full-solar-spectrum photovoltaic material system," *J. App. Phys*, vol. 94, no. 10, pp. 6477-6482, Nov. 2003.

[3] C. J. Neufeld, N. G. Toledo, S. C. Cruz, M. Lza, S. P. DenBaars and U. K. Mishra, "High quantum efficiency InGaN/GaN solar cells with 2.95 eV band gap," *App. Phys. Lett.*, vol. 93, no. 14, pp. 143502-3, Oct. 2008.

[4] A. Yamamoto, M. R. Islam, T. T. Kang and A. Hashimoto, "Recent advances in InN-based solar cells: status and challenges in InGaN and InAlN solar cells," *Phys. Stat. Sol. C,* vol. 7, no. 5, pp. 1309-1316, Mar. 2010.

[5] J. F. Geisz, S. R. Kurtz, M. W. Wanlas, J. S Ward, A. Duda, D. J. Friedman, M. Olson, W. E. MeMhan, J. T. Kiehl, M. J. Romero, A. G. Norman and K. M. Jones, ''Inverted GaInP/(In)GaAs/InGaAs Triple-Junction Solar Cells With low-stress metamorphic bottom junctions,'' in *Proc. of 33rd IEEE Photovoltaic Specialist Conference*, San Diego May 11-16, 2008, pp. 1-5.

[6] M. Song, Z. Wu, Y. Fang, R. Xiang, Y. Sun, H. Wang, C. Yu, H. Xiong, J. Dai and C. Chen, ''Improved photovoltaic performance of InGaN single junction solar cells by using n-on-p type device structure,'' *J. Optoelectronics and Advanced Materials*, vol. 12, no. 7, pp. 1452-1456, Jul. 2010.

[7] E. Kasper and K. Lyutovich, ''Strain adjustment with thin virtual substrates,'' *Solid-State Electronics*, vol. 48, no. 8, pp. 1257-1263, Aug. 2004.

[8] M. R. Islam, Y. Ohmura, A. Hashimoto, A. Yamamoto, K. Kinoshita and Y. Koji, "Step-graded interlayers for the improvement of MOVPE $In_xGa_{1-x}N$ ($x \sim 0.4$) epi-layer quality," *Phys. Stat. Sol. C,* vol. 7, no. 7–8, pp. 2097-2100, July 2010.

[9] B. Bertoli, D. Sidoti, S. Xhurxhi, T. Kujofsa, S. Cheruku, J. P. Correa, P. B. Rago, E. N. Suarez, F. C. Jain and J. E. Ayers, "Equilibrium strain and dislocation density in exponentially graded $Si_{1-x}Ge_x$/Si (001)," *J. Appl. Phys.* vol. 108, no. 11, pp. 113525/5, Dec. 2010.

[10] A. E. Romanov, W. Pompe, S. Mathis, G. E Beltz and J. S. Speck," Threading dislocation reduction in strained layers," *J. Appl. Phys.*, vol. 85, no. 1, pp. 182-192, Dec. 1999.

[11] G. Macpherson, R. Beanland and J. P. Goodhew, "A novel design method for the suppression of edge dislocation formation in step-graded InGaAs/GaAs layers," *Philosophical Magazine A*, vol. 73, no. 05, pp. 1439-1450, 1996.

[12] M. A. Hossain, M. R. Islam, M. M. Hasan, A. Yamamoto and A. Hashimoto, "A mathematical modeling of dislocations reduction in $In_xGa_{1-x}N$/GaN heteroepitaxy using step-graded interlayers," in *Proc. of 7th International Conference on Electrical and Computer Engineering*, Dhaka, Bangladesh, Dec. 20-22, 2012.

[13] A. E. Romanov, W. Pompe, G. Beltz and J. S. Speck, "Modeling of threading dislocation density reduction in heteroepitaxial layer: Part I. Geometry and Crystallography," *Phys. Stat. Sol. (b)*, vol. 198, no. 2, pp. 599-613, Dec. 1996.

[14] S. K. Mathis, A. E. Romanov, L. F. Chen, G. Beltz, W. Pompe and J. S. Speck, "Modeling of threading dislocation reduction in growing GaN layers," *Phys. Stat. Sol. (a)*, vol. 179, no. 1, pp. 125-145, May 2000.

[15] Y.S. Chen, W.Y. Shiao, T.Y. Tang, W.M. Chang, C.H. Liao, C.H. Lin, K.C. Shen, C. C. Yang, M.C.Hsu, J.H. Yeh, and T.C. Hsu, "Threading dislocation evolution in patterned GaN nanocolumn growth and coalescence overgrowth," *J. Appl. Phys.,* vol. 106, no. 2, pp. 023521-6, Jul. 2009.

# Effect of QD Size and Band-offsets on Confinement Energy in InN QD Heterostructure

Udoy Paul*, Mahmudul Hasan, Md. Tawabur Rahman and Ashraful G. Bhuiyan
Department of Electrical and Electronic Engineering
Khulna University of Engineering & Technology
Khulna-9203, Bangladesh
* E-mail: udoy.eee@gmail.com

*Abstract*—**Detailed theoretical analysis of how QD size variation and band-offset affects the confinement energy of InN QD is presented. Low dimensional structures show a strong quantum confinement effect, which results in shifting the ground state away from the band edge and discrete eigen-states. Graphically solving 1D Schrödinger ground quantized energy levels of electrons were computed and using Luttinger-Khon 4×4 Hamiltonian matrix ground quantized energy level of holes were determined. Our results allow us to tune dot size and band-offset to obtain required bandgap for InN based low dimensional device design.**

***Keywords—Indium Nitride, Confinement energy, Quantum dots (QD) .***

## I. INTRODUCTION

Advancements in the field of III-nitride (InN, GaN, and AlN) semiconductors have been remarkable due to their highly attractive inherent properties. [1]. With the recent discovery that the bandgap of InN is about ~0.64eV [2], InN based III-nitrides became materials of interest for the researchers especially because of their application in most of the near infrared optical and optoelectronic applications [3]. InN could be very interesting for high performance heterostructure laser of 1.55 μm which is the most promising communication window for long distance optical fiber communication. Recent research is devoted to the fabrication of low dimensional optoelectronic devices, preeminently QD optoelectronic devices as they offer ultra-high temperature stability, narrow chirp characteristics and higher performance [4].

The necessity of determining confinement energy for QD based device design is very important as it is the most significant property of QDs. Confinement energy has fairly large impact in the understanding of the physics of QDs and their technological application. Two most significant phenomena attributed to the quantum confinement effect in QDs are: (a) the ground state shifts upward and (b) the eigen energies are discrete [5].The ground quantized energy of electron and holes add up to the bulk bandgap of the QD material, i.e., the total energy of transition is actually the summation of bulk bandgap of the QD material ($E_g$) and ground quantized energy levels of electron ($E_e$) and holes ($E_h$). So it is very important to determine these energies very precisely. Instead of predicting confinement energy from statistical data or neglecting their effect for the simplification of design, as was in most of the previous work on InN and InN based III-nitride optoelectronic device design [6], [14] it is more important to theoretically calculate the confinement energy of InN and InN based III-nitride QDs in order to have a more practical and hence more acceptable design.

In this study, we theoretically computed the confinement energies for electron and holes in InN QD heterostructure and their dependence on the size of the dot as well as the band-offset. 1D Schrödinger was graphically solved to determine ground quantized energy level of electrons, on the other hand ground quantized energy level of holes were determined by computing the eigen values of Luttinger-Khon 4×4 Hamiltonian matrix.

## II. CALCULATING CONFINEMENT ENERGY

In InN quantum dot heterostructure, where different barrier materials are combined to form a heterostructure, the shift in the band structure induces a step like potential profile as shown in figure 1; the height of the discontinuity is called band-offsets [7]. The conduction band-offset confines the electrons whereas the valence band-offset confines the holes.

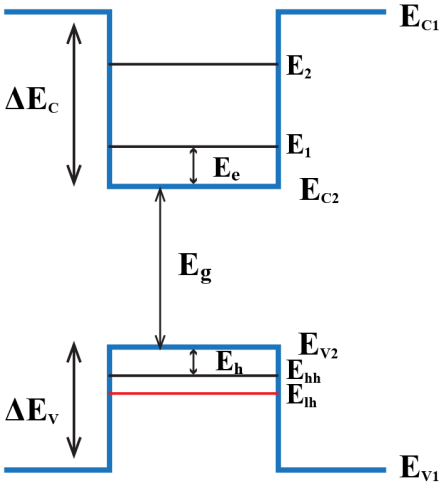



Figure 1. The schematic diagram of a quantum structure. The conduction, heavy-hole and light-hole valence band edge profiles, as well as electron and hole confined states are shown.

In figure 1, $E_{C1}$ and $E_{V1}$ refer to conduction band and valance band profiles of the barrier material respectively. $E_{C2}$ and $E_{V2}$ is the conduction band and valance band profiles of the dot material respectively. $E_1$ and $E_2$ represent different eigen states of the conduction band of the dot material. $E_{hh}$ and $E_{lh}$ are heavy-hole and light-hole valence band edge profiles. $\Delta E_C$ and $\Delta E_V$ represents conduction band offset and valance band offset respectively.

Semiconductor with electron and hole confinement offer a distinct method to study electrical and optical phenomena as a function of variation in dot size and band-offset provided by the barrier material [8]. The shift in the ground state energy of electrons and holes from their respective band edge has been computed using following techniques:

### *A. Ground quantized energy level of electrons*

The one dimensional Schrödinger equation for conduction band is:

$$\frac{-\hbar^2}{2m^*_e}\frac{d^2}{dx^2}\Psi(x) + V_c(x)\,\Psi(x) = E_e\,\Psi(x) \qquad (1)$$

Here $\Psi(x)$ is the conduction band wave function. $E_e$ represents electron energy level and $V_c(x)$ for conduction band is actually the band-offset ($\Delta E_C$) in the conduction band. The nontrivial solution of equation (1) has the form [9]:

$$\xi\,(\tan\xi - \cot\xi) = \sqrt{(\alpha^2 - \xi^2)} \qquad (2)$$

In equation (2) $\xi$ is a function of $E_e$ [9], effective mass ($m_e^*$) of electron and base length of the QD whereas $\alpha$ is a function of band-offset, effective mass of electron and QD base length. Graphically solving the above mentioned transcendental equation the value of $\xi$ was determined. We calculated the value of $E_e$ by knowing the value of $\xi$ for a particular value of base length of InN QD for a known band-offset $\Delta E_C$.

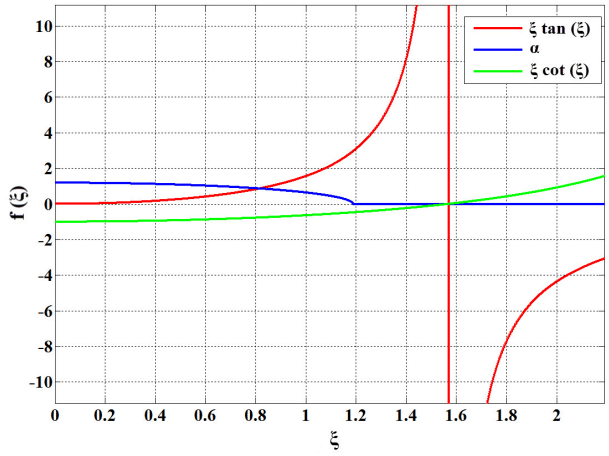



Figure 2. Graphical solution of the transcendental equation for $\xi$ to determine the value of $E_e$

Figure 2, is a plot of equation (2) for some particular value of effective mass of electron, base length of dot and band-offset. Taking effective mass of electron in InN as $0.04m_e$ [12], [13] for dot base length of 6 nm and band-offset of 0.15 eV from figure 2, we graphically calculated the value of $\xi$ as 0.8156 and we get the value of $E_e$ as 0.0705 eV.

### *B. Ground quantized energy level of holes*

For valance band; to determine the confinement energy of holes, we start with effective mass theory on quantum wells using the Luttinger-Khon 4×4 Hamiltonian matrix [10]. The Luttinger-Khon 4×4 Hamiltonian matrix can be expressed by:

$$H_0\Psi_h(x) = E_h\Psi_h \qquad (3)$$

The components of the resultant Hamiltonian matrix are given below:

$$H_0 = \begin{pmatrix} H_{hh} & b & c & 0 \\ b^* & H_{lh} & 0 & c \\ c^* & 0 & H_{lh} & -b \\ 0 & c^* & -b^* & H_{hh} \end{pmatrix} \qquad (4)$$

Here,

$$H_{hh} = \left\{ \frac{-\hbar^2}{2m_0}\left[(\gamma_1 + \gamma_2)k_1^2 - (\gamma_1 - 2\gamma_2)\frac{\partial^2}{\partial x^2}\right] + V_h(x) \right\} \qquad (5)$$

$$H_{lh} = \left\{ \frac{-\hbar^2}{2m_0}\left[(\gamma_1 - \gamma_2)k_1^2 - (\gamma_1 + 2\gamma_2)\frac{\partial^2}{\partial x^2}\right] + V_h(x) \right\} \qquad (6)$$

$$b = \frac{i\sqrt{3}\hbar^2}{m_0}(-k_2 - ik_1)\,\gamma_3\,\frac{\partial}{\partial x} \qquad (7)$$

and,

$$c = \frac{\sqrt{3}\hbar^2}{2m_0}\left[\gamma_2(k_1^2 - k_2^2) - 2i\,\gamma_3 k_1 k_2\right] \qquad (8)$$

$$k_n = \frac{n\pi}{L}\;; \quad \text{for, } n = 1,2,3,4,\ldots\ldots \qquad (9)$$

Here, $E_h$ is the energy of heavy and light holes and $m_0$ is free electron mass. $H_{hh}$, $H_{lh}$, b and c are matrix elements and are functions of Luttinger parameters $\gamma_1$, $\gamma_2$ and $\gamma_3$ and components of the transverse wave vector $k_1$ and $k_2$. $V_c(x)$ in equation (5) and (6) is actually the band-offset ($\Delta E_V$) in the valance band. L in equation (9) represents the base length of InN quantum dot. From the Hamiltonian matrix $H_0$, the eigen-energies and eigen-vectors are calculated by knowing the values of $\gamma_1$, $\gamma_2$, $\gamma_3$, $k_1$ and $k_2$ for InN, which gave the value of ground quantized energy level of holes (both for heavy and light holes) for a particular value of InN dot size and $\Delta E_V$. For InN values of Luttinger parameter considered are 3.72, 1.26 and 1.63 [11].

## III. RESULTS AND DISCUSSION

We have calculated confinement energies for InN QDs using methods mentioned in the previous section for different dot size and for different band-offset provided by the barrier material. Computation and analysis carried out for determining confinement energy of electrons are graphically represented as follows.

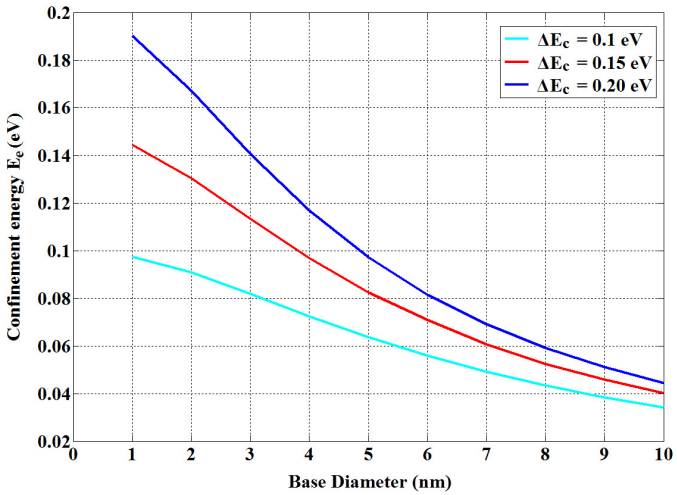



Figure 3. Change in confinement energy with the variation of QDs base length (for conduction band).

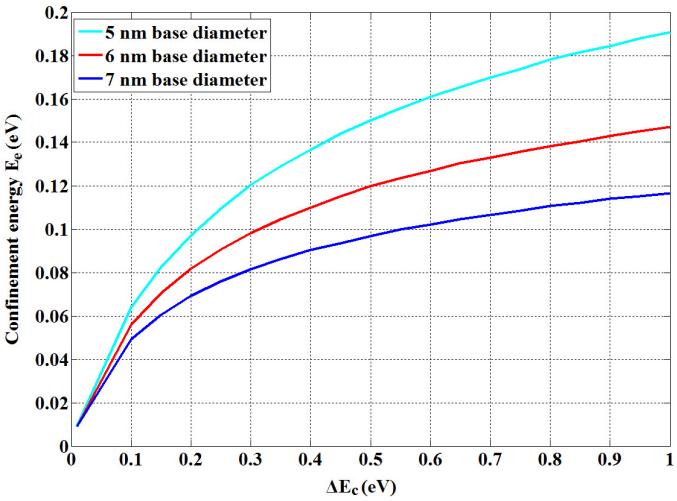



Figure 4. Change in confinement energy with the variation of band offset (for conduction band).

Turning now to the computational evidence of dependency of confinement energy on dot size; figure 3, shows how confinement energy of electron changes with the change in base length of InN QDs, for some fixed values of $\Delta E_C$ (0.1 eV, 0.15 eV and 0.20 eV). With the rise in base length of quantum dot, confinement energy of electrons decreases. Another important finding was that at a very low value of dot base length the confinement potential is almost equal to the band offset ($\Delta E_C$).

For 3 distinctive value of QD base length (5 nm, 6 nm and 7 nm) the dependency of confinement energy of electrons in the conduction band on band-offset $\Delta E_C$ is shown in figure 4. Ground quantized energy level of electrons shifts up with the increase in $\Delta E_C$ and therefore the confinement energy increases. Confinement energy becomes very prominent when the band-offset is large. Also the smallest dot provides highest confinement at any particular value of $\Delta E_C$.

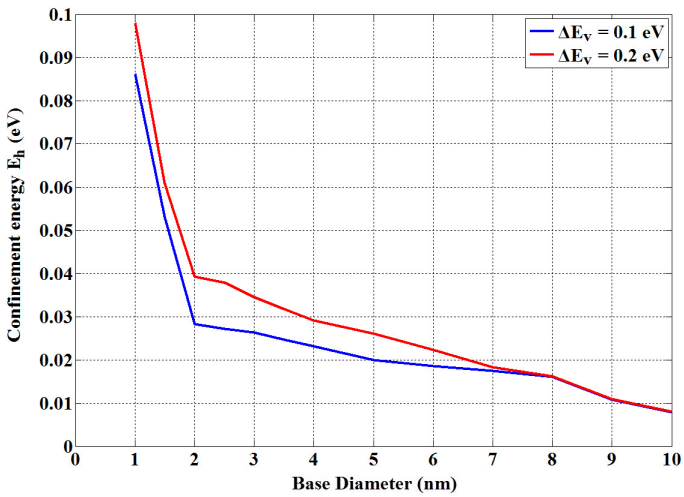



Figure 5. Change in confinement energy with the variation of QDs base length (for valance band).

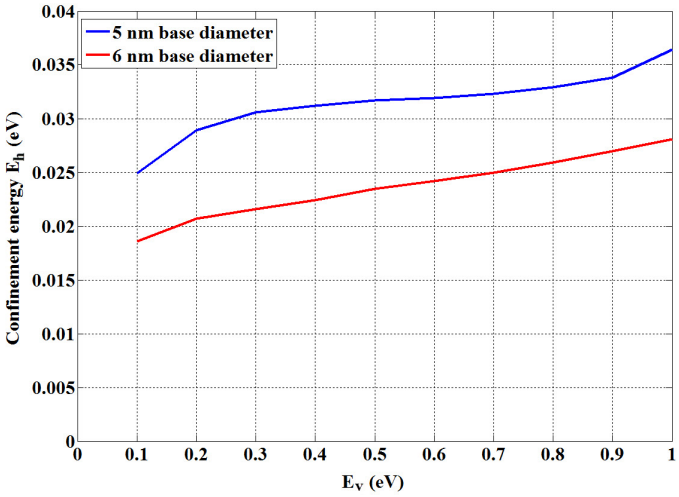



Figure 6. Change in confinement energy with the variation of band offset (for valance band).

On the other hand calculated values of ground quantized energy level of holes with the alteration of dot size and $\Delta E_V$ are graphically represented above. Though in almost all previous studies it was the most neglecting parameter because of computation complexity and least contribution to the total energy of transition, we computed confinement potential for holes and plotted it against base diameter of QDs in figure 5. Ground quantized energy level shifts away from the band edge with decreasing value of base length and become more prominent for smaller dot size.

In figure 6, for 2 different value of QD base length (5 nm and 6nm); computed confinement energy was plotted against the variation in band-offset, $\Delta E_V$ in valance band. Here also we found that increasing value of $\Delta E_V$ increases the confinement energy. The variation in confinement energy with the change in band-offset is not very sharp here as was in conduction band, which indicates that confinement energy of holes is relatively less dependent on the band-offset.

The evidence from this study suggests that, the total energy of confinement for heterostructures composed of same dot material and barrier material may not be unique, as it is a function of dot size. Another striking finding was the dependency of confinement energy on band-offsets, which clarifies that changing barrier material or its composition would also affect the total energy of transition. It is therefore likely that to obtain the actual energy of transition and hence the wavelength of the light; confinement energies should be determined very precisely. Design considerations of optoelectronic devices especially lasers and LEDs used in optical fiber communication, where wavelength has to be exact to ensure lower attenuation and dispersion, should include the effect of confinement properly.

## IV. CONCLUSION

A convenient method of determining ground quantized energy level of electrons and holes was studied. Theoretical computation method has been applied to determine the confinement potential. In addition, confinement energy of InN quantum dot heterostructure was calculated for various dot size and band-offsets. Variation of confinement energy as a function of dot size, $\Delta E_C$ and $\Delta E_V$ was demonstrated, which is important for the designing of InN and InN based III-nitride QD optoelectronic devices. It was done for the first time with an approach to highlight their contribution and to suggest their inclusion in design consideration. Thus the overall theoretical analysis presented, helped precisely determining the confinement energy and taking its effect for more practical device designing.

## REFERENCES


[1] A. G. Bhuiyan, A. Hashimoto, and A. Yamamoto, "Indium nitride (InN): A review on growth, characterization, and properties," *Appl. Phys. Lett.,* Vol. 94, No. 5, pp. 2779–2808, Sep. 2003.

[2] C.-L. Hsiao, L.-W. Tu, M. Chen, Z.-W. Jiang, N.-W. Fan, Y.-J. Tu and K.-R. Wang, "Polycrystalline to Single-Crystalline InN Grown on Si(111) Substrates by Plasma-Assisted Molecular-Beam Epitaxy," *Jpn. J. Appl. Phys.,* Vol. 44, No. 34, pp.L1076–L1079, Aug. 2005.

[3] B.Monemar, P.P.Paskov, A. Kasic, "Review Optical properties of InN –the bandgap question," *Superlattices and Microstructures*, Vol. 38, No. 1, pp. 38–56, Jul. 2005.

[4] A. Salhi, G. Rain`o, L. Fortunato, V. Tasco, G. Visimberga, L. Martiradonna, M. T. Todaro, M. D. Giorgi, R. Cingolani, A. Trampert, M. D. Vittorio, and A. Passaseo "Enhanced Performances of Quantum Dot Lasers Operating at 1.3 μm" ", *IEEE Journal on Selected Topics in Quantum Electronics,* Vol. 14, No. 4, pp. 1188–1196, Aug. 2008.

[5] S.-W. Feng, J. Han," Quantum-confinement effect on recombination dynamics and carrier localization in cubic InN and $In_xGa_{(1-x)}N$ quantum boxes," *Thin sold films.,* Vol. 517, No. 11, pp. 3315–3319, Apr. 2009.

[6] M. M. Hossain, M. A. A. Humayun, M. T. Hasan, A. G. Bhuiyan, A. Hashimoto, and A. Yamamoto, "Proposal of high performance 1.55 μm quantum dot heterostructure laser using InN," *IEICE Transactions on Electronics,* vol. E95-C, No. 2, pp. 255-261, Feb. 2012.

[7] F. Capasso, *Physics of Quantum Electron Devices*, Springer-Verlag Berlin Heidelberg, 1990

[8] D. R. Khanal, Joanne W.L. Yim, W Walukiewicz, and J. Wu, "Effect of Quantum Confinement on the Doping Limit of Semiconductor Nanowires," *Nano Lett.,* Vol. 7, No. 5, pp. 1186–1190, Mar. 2007.

[9] A. Ghatak, S. Lokanathan, *Quantum Mechanics: Theory And Applications,* Dordrecht, The Netherlands, 2004**.**

[10] S.L. Chuang, *Physics of optoelectronic devices*, John wiely & sons,1995

[11] H. Morkoç, *Handbook of nitride semiconductor and devics Series: Electronic and Optical Processes in Nitrides*, John wiely & sons, 2008, vol. 2.

[12] S. Wang, H. Liu, B. Gao, and H. Cai, "Monte Carlo calculation of electron diffusion coefficient in wurtzite indium nitride," *Appl. Phys. Lett.,* Vol. 100, No. 14, pp. 142105 - 142105-3, 5 Apr. 2012.

[13] F. M. Abouel-Ela and B. M. El-Assy, "Electron transport in wurtzite InN," *Pramana – J. Phys.,* Vol. 79, No. 1, pp. 125–136, Jul. 2012.

[14] U. Hohenester, R. D. Felice, and E. Molinari, "Optical spectra of nitride quantum dots: Quantum confinement and electron–hole coupling," *Appl. Phys. Lett.,* Vol. 75, No. 22, pp. 3449–3451, Nov. 1999.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# The effect of quantum dot size, interdot distance and indium content on $In_xGa_{1-x}N$/GaN QD–IBSC

Md. Mafizul Islam, Md. Touhidul Islam Bhuiyan*, Md. Tawabur Rahman and Ashraful G. Bhuiyan
Department of Electrical and Electronic Engineering
Khulna University of Engineering and Technology
Khulna–9203, Bangladesh
* E–mail: anik0803010@gmail.com

*Abstract*—**$In_xGa_{1-x}N$/GaN quantum dot intermediate band solar cell (QD–IBSC) is a promising candidate for the purpose of efficiency improvement of solar cells. In this work, the influences of interdot distance, quantum dot size and indium content of $In_xGa_{1-x}N$/GaN QD–IBSC on the position and width of the intermediate bands are investigated by solving Schrödinger equation using Kronig–Penney model. Finally the effects of intermediate band position and width on energy conversion efficiency are analyzed. The results reveal that the $In_xGa_{1-x}N$/GaN quantum dot intermediate band solar cell manifests much larger power conversion efficiency than that of conventional solar cells. The maximum efficiency occurs when the intermediate band is at the middle position of barrier material bandgap and bandwidth is wide enough to increase absorption of photons keeping carrier recombination negligible by tuning interdot distance, quantum dot size and indium content.**

***Keywords—Quantum dot, intermediate band, $In_xGa_{1-x}N$***

## I. INTRODUCTION

The intermediate band solar cell (IBSC) is an advanced approach with the potential for exceeding the performance of conventional single gap solar cell [1]. Quantum dot (QD) approach seems to be the quickest experimental way to manufacture prototypes in which intermediate band (IB) concepts could be investigated [2]. In this approach, the IB is formed by the confined electron energy states introduced by an array of QDs embedded in high bandgap barrier region. The maximum conversion efficiency of quantum dot intermediate band solar cell (QD–IBSC) is higher as the IB formed by the QDs increases absorption of longer wavelength sunlight and diminishes nonradiative recombination [3]. The maximum attainable conversion efficiency of QD–IBSC is predicted to be about 66% [4].

The III–nitride material system is promising for integration into high–efficiency PV systems [5]. Indium Gallium Nitride ($In_xGa_{1-x}N$) material is now widely used in various optoelectronic devices and is being studied increasingly as a prospective material for solar cells as the bandgap energy can be engineered from 0.64 eV to 3.4 eV by varying the indium composition ($x$), which covers most of the solar spectrum from ultraviolet to infrared region with characteristics of direct bandgap, high absorption coefficients, and excellent irradiance resistance. Particularly, the $In_xGa_{1-x}N$ material system presents excellent properties in low–dimensional semiconductor devices [6]. It is therefore important to concentrate on $In_xGa_{1-x}N$–based QD solar cells.

At present the $In_xGa_{1-x}N$–based QD research is limited only in the growth and characterization, there is no report on the fabrication of QD solar cells using nitride materials. Recently, the $In_xGa_{1-x}N$–based QD solar cell has been investigated theoretically. Deng et al. theoretically investigated the $In_xGa_{1-x}N$/GaN QD–IBSC and $In_xGa_{1-x}N$ /InN QD–IBSC by tuning the size of QDs and the interdot distance [7]. These works do not clarify about the effects of QD size, interdot distance and indium content on IB position and width which are very important factors for the realization of the performance of a QD–IBSC. In our work, we have investigated the effect of quantum dot size, interdot distance and indium content on intermediate band position and width, and also on solar cell performance which may open the door for future experiment.

## II. THEORY AND MODEL

An isolated QD is able to create discrete energy levels for electrons. A periodic array of QDs results in IB due to coupling of wave functions between adjacent QDs. For simplicity, the orthorhombic symmetry is used for the QDs as shown in Fig. 1. The behavior of an electron in such a system is described by the Schrödinger equation [8]:

$$\left[-\frac{\hbar^2}{2}\nabla_r\frac{1}{m^*(r)}\nabla_r+V(r)\right]\varphi(r)=E\varphi(r) \quad (1)$$

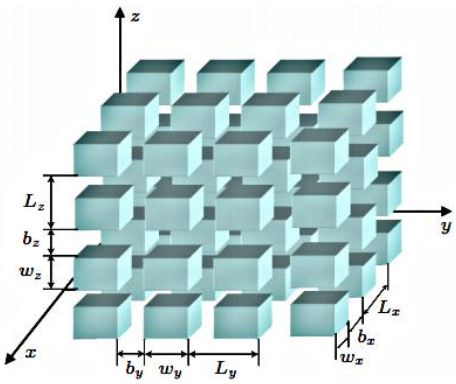



Figure 1. Schematic of orthorhombic QDs.

where $m^*$ is the effective mass of electron. The potential $V(r)$ corresponds to an infinite sequence of QDs of the sizes $w_x$, $w_y$ and $w_z$ separated by the barriers of thicknesses $b_x$, $b_y$ and $b_z$ as shown in Fig. 1. Here $L_x$, $L_y$ and $L_z$ represent the periods of QDs in the coordinates $x$, $y$ and $z$ respectively. The size of QDs and interdot distance are assumed to be equal ($w_x = w_y = w_z$ and $b_x = b_y = b_z$).

According to the Kronig–Penney model, solution of the Schrödinger equation has already been accomplished by following the Lazarenkova and Balandin envelope function approximation, which has the following form [8]:

$$\cos\left(L_\zeta q_\zeta\right) = \cos\left(k_\zeta^w w_\zeta\right)\cos\left(k_\zeta^b b_\zeta\right) - 0.5\left(\frac{k_\zeta^b m_w^*}{k_\zeta^w m_b^*} + \frac{k_\zeta^w m_b^*}{k_\zeta^b m_w^*}\right) \times \sin\left(k_\zeta^w w_\zeta\right)\sin\left(k_\zeta^b b_\zeta\right), \quad (2)$$

$$\text{if } E_\zeta \geq V_0$$

$$\cos\left(L_\zeta q_\zeta\right) = \cos\left(k_\zeta^w w_\zeta\right)\cosh\left(k_\zeta^b b_\zeta\right) - 0.5\left(\frac{k_\zeta^w m_b^*}{k_\zeta^b m_w^*} - \frac{k_\zeta^b m_w^*}{k_\zeta^w m_b^*}\right) \times \sin\left(k_\zeta^w w_\zeta\right)\sinh\left(k_\zeta^b b_\zeta\right), \quad (3)$$

$$\text{if } 0 < E_\zeta < V_0$$

where $k_\zeta^b = \frac{1}{\hbar}\sqrt{2m_b^*\left|E_\zeta - V_0\right|}$, $k_\zeta^w = \frac{1}{\hbar}\sqrt{2m_w^*\left|E_\zeta\right|}$, $m_b^*$ and $m_w^*$ are the effective masses of electron in the barrier region and QD region, respectively. $\hbar$ is the Dirac constant, $E_\zeta$ and $q_\zeta$ are the single electron energy and wave vector along the particular coordinate axis, respectively and $V_0$ is the potential barrier. The bandgap and the effective mass of $In_xGa_{1-x}N$ composite material are calculated using the data of InN (0.64 eV, $0.05m_0$) and GaN (3.4 eV, $0.2m_0$). The bowing parameter for $In_xGa_{1-x}N$ is 1.43 [9].

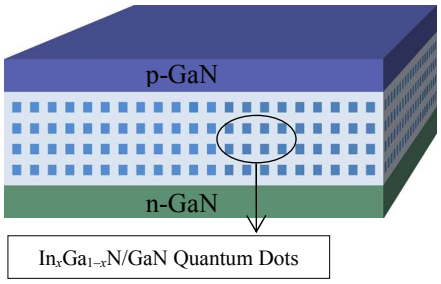



Figure 2. Schematic of $In_xGa_{1-x}N$/GaN quantum dot solar cell.

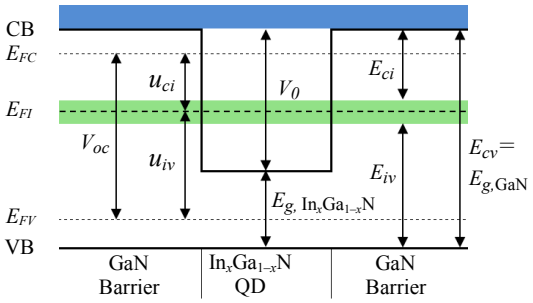



Figure 3. Energy band diagram of $In_xGa_{1-x}N$/GaN QD–IBSC along the [[1 0 0]] direction.

$$E_g\left(In_xGa_{1-x}N\right) = 0.64x + 3.4(1-x) - 1.43x(1-x) \quad (4)$$

$$m^*\left(In_xGa_{1-x}N\right) = x \cdot m^*\left(InN\right) + (1-x) \cdot m^*\left(GaN\right) \quad (5)$$

For a convenient study, the [[1 0 0]] direction is investigated as it defines the charge carrier transport in the vertical direction from n–type to p–type layer [10]. This direction is related to the ordering of the QDs, but not associated with the coordinate axis directions. In this work the IB in the potential well and the performance of the solar cell are investigated as a function of $In_xGa_{1-x}N$ QD size, interdot distance and indium content for $In_xGa_{1-x}N$/GaN QD–IBSC. The calculations are done by following the standard assumptions of the ideal IBSC [1]. The cell is considered to be thick enough to ensure full absorption of photons for the transition of electrons as in Fig. 2 and to restrict nonradiative transitions [1]. The photo induced current for an ideal IBSC strongly depends on the number of photons absorbed by the device and the number of photons emitted from the device. According to Roosbroeck–Shockley formula the photon flux absorbed by or emitted from the device is defined as [11]:

$$\dot{N}\left(E_a, E_b, T, u\right) = \frac{2\pi}{h^3c^2}\int_{E_a}^{E_b} \frac{E^2 dE}{e^{(E-u)/kT} - 1} \quad (6)$$

where $E_a$ and $E_b$ are the lower and upper energy limits of the photon flux, respectively. $T$, $h$, $c$, $k$ and $u$ are the temperature, Planck constant, light speed in vacuum, Boltzmann constant and chemical potential of the transition, respectively. The short circuit current density ($J_{sc}$) resulting from the photon absorption can be expressed as [1]:

$$\frac{J_{sc}}{q} = \left[\dot{N}\left(E_{cv}, \infty, T_s, 0\right) - \dot{N}\left(E_{cv}, \infty, T_0, u_{cv}\right)\right] + \left[\dot{N}\left(E_{ci}, E_{cv}, T_s, 0\right) - \dot{N}\left(E_{ci}, E_{cv}, T_0, u_{ci}\right)\right] \quad (7)$$

$$\frac{J_{sc}}{q} = \left[\dot{N}\left(E_{cv}, \infty, T_s, 0\right) - \dot{N}\left(E_{cv}, \infty, T_0, u_{cv}\right)\right] + \left[\dot{N}\left(E_{iv}, E_{cv}, T_s, 0\right) - \dot{N}\left(E_{iv}, E_{cv}, T_0, u_{iv}\right)\right] \quad (8)$$

where $T_s$ is the sun temperature (6000K), $T_0$ the temperature of the solar cell (300K), $q$ the electron charge, $u_{cv}$, $u_{ci}$ and $u_{iv}$ are the chemical potentials between the conduction band (CB) and valence band (VB), the CB and IB, and the IB and VB, respectively. $E_{cv}$, $E_{iv}$ and $E_{ci}$ are illustrated in Fig. 3 and they are given according to the solution of (3). $E_{FV}$, $E_{FC}$ and $E_{FI}$ are the quasi–Fermi levels for VB, CB and IB. As IB is formed by coupling of the adjacent QDs, so the IB is not stable state and partially filled by electrons [7]. If $E_{ci} > E_{iv}$, $J_{sc}$ is calculated using (7), otherwise using (8). The open circuit voltage ($V_{oc}$) of the cell is given by the summation of chemical potentials between CB and VB, i.e.,

$$qV_{oc} = u_{cv} = u_{ci} + u_{iv} \quad (9)$$

The efficiency of the solar cell is given as

$$\eta = \frac{V_{oc} \times J_{sc} \times F}{P_{in}} \tag{10}$$

where, $F$ is the fill factor ($FF$) and $P_{in}$ is the incident power per unit area, $P_{in} = \sigma T_s^4$, where σ is the Stefan–Boltzmann constant. In this work, $FF$ is assumed to be unity.

## III. Result and Discussion

The Important factors for improving the QD–IBSC performance are the number of intermediate bands, their positions and width of each intermediate band (IB). QD size, indium content and interdot distance have significant effects on the position and width of the IBs. IB formation using QDs and the effect of QD parameters are investigated. The effect of IB position and width on efficiency of the cell is also discussed in this section.

### *A. Effect of interdot distance on IB*

The width of IB has a significant influence on absorption of photons and on carrier recombination. It is challenging to keep an optimum width of the IB for maximum number of photon absorption and minimum number of carrier recombination. The IB width can be tuned by changing interdot distance as shown in Fig. 4. IBs have wider bandwidth at smaller interdot distance due to stronger wavefunction overlap. Increase in interdot distance causes reduction in bandwidths and finally transforms the bands into discrete energy levels.

### *B. Effect of QD size on IB*

The number of IBs and their position inside the barrier material bandgap is a vital issue for the performance of QD–IBSC. QD size has a significant effect on these parameters which is observed from Fig. 5. Larger size of QDs generates more intermediate bands. Increase in QD size increases the wavelength of the wavefunctions hence lowers the position of the intermediate bands and reduces the bandwidths to some extent giving space for additional bands. In the upper energy levels carrier accumulation is more resulting in higher number of quantized energy states, so the upper bands are wider than lower ones.

### *C. Effect of Indium content on IB*

Since $In_xGa_{1-x}N$ is used as QD–material, there is an effect of indium content on IB which is shown in Fig. 6. With increase in indium content, bandgap of $In_xGa_{1-x}N$ decreases parabolically which in turn increases the height of the potential barrier $V_0$. With the increase in $V_0$ downwards in the barrier material bandgap, the energy levels go in higher position in the potential well but lower in the barrier material bandgap. So the bandwidth of the IBs becomes narrow. However, if the increase in energy in the potential well overcomes the effect of narrowing of the dot material bandgap, then the energy bands go up in the barrier material making the bands wider.

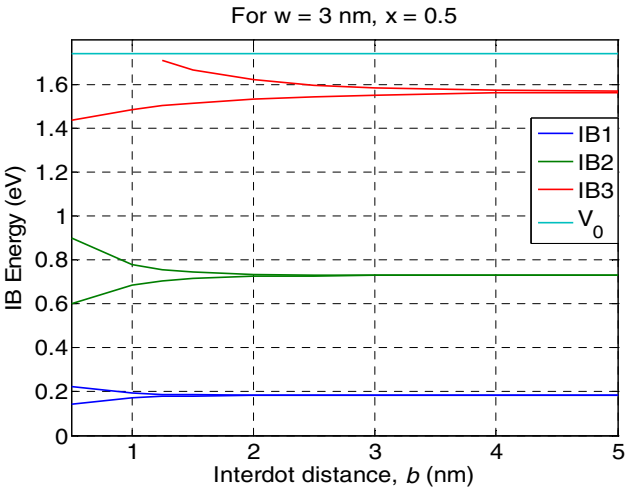



Figure 4. IB width as a function of interdot distance in the potential well.

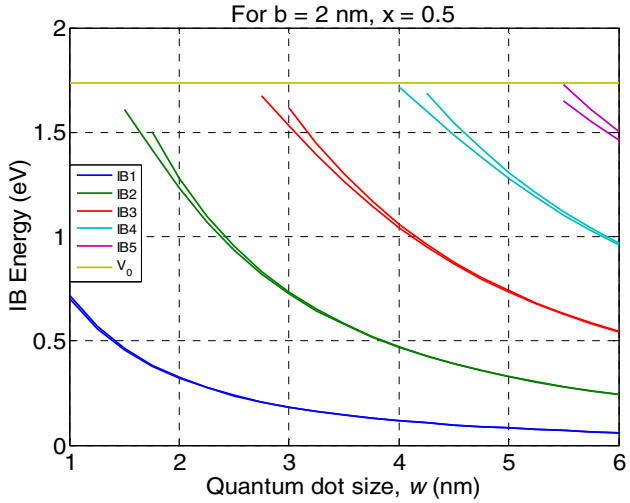



Figure 5. IB position and width as a function of quantum dot size in the potential well; IB1, IB2, IB3, IB4, IB5 are the intermediate bands from lower position to upper position, respectively and $V_0$ is potential barrier height.

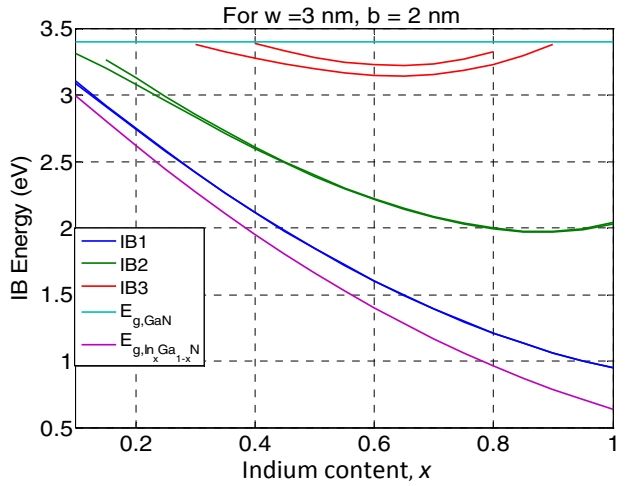



Figure 6. IB position and width as a function of indium content in $In_xGa_{1-x}N$ inside the previously forbidden bandgap region. $E_{g,GaN}$ is the bandgap of GaN, the bandgap of $In_xGa_{1-x}N$ as a function of indium content, $x$ is also shown.

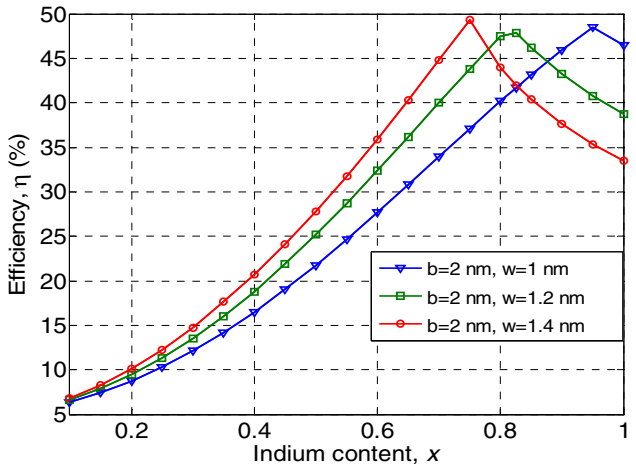



Figure 7. Dependences of indium content and size of each QD on energy conversion efficiency in the $In_xGa_{1-x}N/GaN$ QD–IBSC.

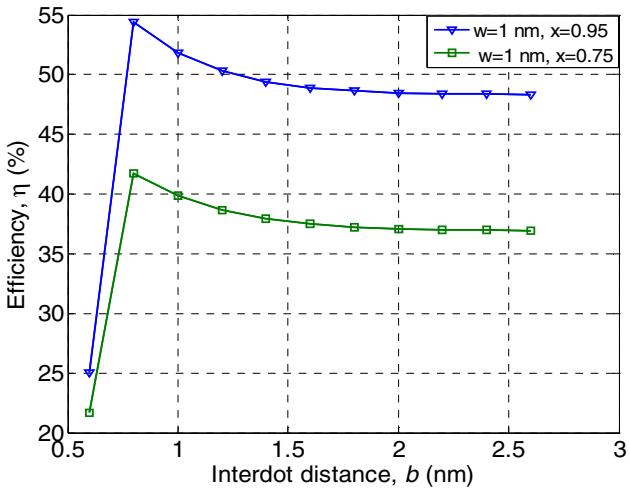



Figure 8. Dependences of interdot distance on energy conversion efficiency in the $In_xGa_{1-x}N/GaN$ QD–IBSC.

### *D. Effect of QD size, indium content and interdot distance on energy conversion efficiency*

Higher range of solar spectrum can be converted into electrical energy forming IB inside the barrier material bandgap. The IB formation and the variation of IB parameters are realized using QDs. It is found that for quantum dot size ranging from 1 to 1.4 nm and indium content from 0.1 to 1 only one intermediate band forms in the potential well of $In_xGa_{1-x}N/GaN$ QD–IBSC. As the value of sub bandgap $E_{ci}$ approaches to the value of $E_{iv}$, the energy range for photon absorption increases. This increases the carrier generation and the efficiency of the cell. Efficiency is maximum when $E_{ci}$ and $E_{iv}$ are equal i.e., IB is exactly at the middle position of the barrier material bandgap. From Fig. 7 efficiencies 48.48, 47.82 and 49.27% are obtained at $x$ = 0.95, 0.825 and 0.75 for $w$ = 1, 1.2 and 1.4 nm, respectively. Hence, if we increase QD size, maximum efficiency for that size requires reduced indium content to locate the intermediate band midway between the barrier material bandgap. It is already known that decreasing interdot distance causes wider bandwidth which increases the range of photon absorption. If the interdot distance is very small then photon emission due to recombination becomes dominant which decreases the efficiency and degrades solar cell performance. In Fig. 8 for $w$ = 1 nm and $x$ = 0.75 efficiency is further increased from 48.48 to 54.41% by decreasing interdot distance to increase bandwidth till radiative recombination becomes dominant. The maximum efficiency will be 62.79% for two bands in the potential well.

## IV. CONCLUSION

In summary, the effects of quantum dot (QD) size, interdot distance and indium content on $In_xGa_{1-x}N/GaN$ quantum dot intermediate band solar cell has been investigated. The position of intermediate band (IB) strongly depends on QD size and indium content whereas the bandwidth is strongly influenced by interdot distance. The maximum efficiency has been obtained when the IB is positioned midway between the bandgap of the barrier material and bandwidth is increased until radiative recombination starts being dominant. The maximum efficiency for one IB is 54.41 % and it can be promoted to 62.79 % if two bands are formed in the potential well.

## REFERENCES


[1] A. Luque, A. Marti, "Increasing the efficiency of ideal solar cells by photon induced transitions at intermediate levels," *Phys. Rev. Lett.*, vol. 78, no. 26, pp. 5014–5017, 1997.

[2] L. Cuadra, A. Marti, A. Luque, "Present status of intermediate band solar cell research," *Thin Solid Films*, vol. 451–452, pp. 593–599, 2004.

[3] A. Marti, L. Cuadra, A. Luque, "Partial filling of a quantum dot intermediate band for solar cells," *IEEE Trans. Electron Devices*, vol. 48, no. 10, pp. 2394–2399, Oct. 2001.

[4] A. J. Nozik, "Quantum dot solar cells," *Physica E*, vol. 14, no. 1–2, pp. 115–120, Apr. 2002.

[5] Na Lu and Ian Ferguson, "III–nitrides for energy production: photovoltaic and thermoelectric applications," *Semiconductor Science and Technology*, vol. 28, no. 7, pp. 4023–4033, 2013.

[6] Ashraful Ghani Bhuiyan, Kenichi Sugita, Akihiro Hashimoto, and Akio Yamamoto, "InGaN Solar Cells: Present State of the Art and Important Challenges," *IEEE Journal of Photovoltaics*, vol. 2, no. 3, pp. 276–293, July 2012.

[7] Qingwen Deng, Xiaoliang Wang, Cuibai Yang, Hongling Xiao, Cuimei Wang, Haibo Yin, Qifeng Hou, Jinmin Li, Zhanguo Wang, Xun Hou "Computational Investigation of $In_xGa_{1-x}N/InN$ Quantum–Dot Intermediate–Band Solar Cell," *Chin. Phys. Lett.*, vol. 28, no. 1, pp. 8401–8404, 2011.

[8] O. L. Lazarenkova, A. A. Balandin, "Miniband formation in a quantum dot crystal," *J. Appl. Phys.*, vol. 89, no. 10, pp. 5509–5515, 2001.

[9] J. Wu, W. Walukiewicz, K.M. Yu, J.W. Ager III, E.E. Haller, H. Lu, W.J. Schaff, "Small bandgap bowing in $In_{1-x}Ga_xN$ alloys," *Appl. Phys. Lett.*, vol. 80, no. 25, pp. 4741–4743, 2002.

[10] Q. Shao and A. A. Balandin, A. I. Fedoseyev and M. Turowski, "Intermediate–band solar cells based on quantum dot supercrystals," *Appl. Phys. Lett.*, vol. 91, no. 16, pp. 3503–3505, Oct. 2007.

[11] W. Van Roosbroeck and W. Shockley, "Photon–Radiative Recombination of Electrons and Holes in Germanium," *Phys. Rev.*, vol. 94, no. 6, pp. 1558–1560, 1954.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Discrimination Analysis of EEG Signals at Eye Open and Eye Close Condition for ECS Switching System

Fayeem Aziz*, Mohammad Mahbubur Rahman, Tanvir Ahmad, Md. Sorwar Jahan, Toalha Mohammad Tosrif, Mohammad Mojammal Huq, C.M.F.S. Reza, Afaz Uddin Ahmed, Papan Day, Shahriar Badsha, Abdullah Al Mamoon and S.M. Mithun Hasan
Chittagong University of Engineering and Technology, Bangladesh
* E-mail: imtarit@gmail.com

*Abstract*—**Dependable operation of brain computer interface (BCI) needs accurate classification of EEG. Application based on environment control system for disable people need comfortable and simple switching modes. This paper studies about eye close (EC) and eye open (EO) conditions of EEG for communication tool for severely disable users. For this purpose, Multivariate Gaussian Distribution analysis is applied to check the discrimination of two classes. Power Spectral Density and Central Tendency Measurement are used as features. These features used in Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Fisher's Linear Discriminant Analysis (FLDA) to observe the classification performance in linear and nonlinear environment. The experiment results show that EEG signals can be a reliable media for a switching paradigm of ECS.**

***Keywords—Brain Computer Interface (BCI), Electroencophelogram (EEG), Multivariate Gaussian Analysis, Linear Discrimination Analysis (LDA), Quadratic Discrimination Analysis (QDA).***

## I. INTRODUCTION

Disable people need an assistive Environment Control System (ECS) for their daily comfort living. Human Machine Interface (HMI) has brought many great and ease way for their communication [1]. Hand or body muscle movement is a good media for ECS for the patients can move hand or any other muscle [2]. Patients who have only eye and brain control activity like Amyotrophic lateral Sclerosis (ALS) patients, those systems are not sufficient [3].Several eye controlled HMI system has been developed for them. Electrooculography (EOG) and Video Oculography (VOG) are most popular ways of eye controlled HMI systems [4]. The video based eye tracking systems are the finest method to track the Point of Gaze (POG). However; the system accuracy is still lack the level for real-time systems. The electrode placing method EOG achieved good accuracy. These systems are applicable in mobile application like wheel chair drive system. But the cumbersome and disturbing electrode setups around the eye reduce the usability of these systems. The dipolar potential of the human eye can have a great impact on the EEG signal [5]. However, Electroencephalography (EEG) is found less invasive and more ease to use as control signal than EOG. The detection and analysis of potential over scalp by using electrodes pressed against the scalp to sense the signals, is generally referred as electroencephalography (EEG) [6]. Generally EEG is acquired by the standard 10 to 20 system. There are several methods of recording brain signals, such as: Electroencephalography (EEG), Magnetic Resonance Imaging (MRI) or Magneto encephalography (MEG). EEG provides high temporal resolution and is comparatively easy to apply [7]. In previous studies it is shown that the alpha response of EEG is impacted by the relaxation state of brain [8]. The spectral amplitude of alpha response is found higher while the eye is closed than the eye is opened. Switching mechanisms for ECS using eye open (EO) and eye close (EC) has been proposed in several previous studies. Noticeable discrimination has been observed in fractal dimension (FD) of alpha response [9]. Power spectral density (PSD) and Central Tendency Measurement (CTM) also have been found as significant feature to classify EC and EO condition [10]. However, the classification of those systems was depended on a threshold level between the two classes. Individual setup for each trial was necessary for that threshold, which decreased the robustness of the mechanism. Disable people deserve less complicated and more comfortable system without any calibration at the beginning of every starting. Moreover, the accuracy of those methods was not sufficient for real-time application on ECS. The accuracy of the switching was found 68.7 and 94% accuracy on switching time less than 5 and 10 seconds [10]. Ease use of ECS system required 100% accuracy of switching paradigm.

In this paper we have described the acquisition of EEG data of EC and EO. Discrimination of PSD and CTM of EEG data is analyzed using Multivariate Gaussian Distribution analysis. LDA, QDA and FLDA are applied to observe the classification performance of those features. The data have been tested in both of the single dimension and 2D feature spaces of PSD and CTM feature vectors. From those classification result we concluded with better classification procedure and switching mechanism for future real-time application.

## II. METHODS

### A. International 10-20 system:

International 10-20 system is a standard for acquire EEG signal. The positions of electrodes in this system are determined as follows: Reference points are nasion, level with the eyes; and inion. From these points, the skull perimeters are measured in the transverse and median planes. Fig. 1 shows the electrode locations which are determined by dividing these

(a)

(b)

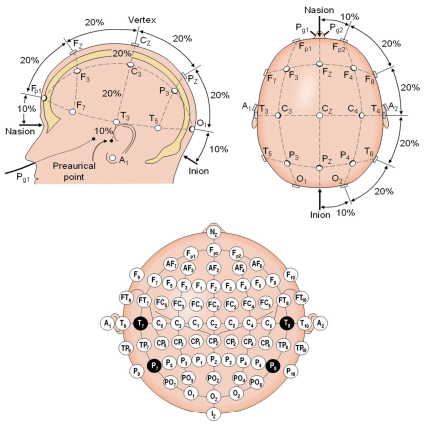



Figure 1: (a) International 10-20 system for EEG signal acquisition. (b) Location of Channels. From [11]

perimeters into 10% and 20% intervals. Three other electrodes are placed on each side equidistant from the neighboring points. In this system the electrodes are placed over scull with a cap. The electrode locations are called channels. The channels are named using F, T, P and O for Frontal, parietal, temporal and occipital area of brain.

### *B. Power Spectral Tendency (PSD):*

Change of relaxation state of brain has great impact on PSD of EEG signal. Vigilance researches showed that human vigilance level is mainly reflected by Power Spectral Density (PSD) of EEG signals [12]. Discrete Short Time Fourier Transform (STFT) is applied to extract the PSD from EEG signals and the PSD bellow 50Hz was taken as the feature information with frequency resolution l Hz.

$$PSD = \frac{1}{n} STFT(Y(k)), k = 1,2,\cdots,n. \tag{1}$$

$$Y(\omega) = \sum_{n} Y(i) e^{-j\omega n} \tag{2}$$

STFT is the DFT of short lengthen signal. The signal is broken into several blocks consists of same number of sample. DFT is applied for each block. $\omega$ is the discrete frequency of DFT.

### *C. Central Tendency Measurement (CTM):*

The CTM quantifies the variability seen in the second order difference plot. It is computed by selecting a circular region of radius (r) around the origin, counting the number of points that fall within the radius and dividing by the total number of points. Let n be the total number of points

$$CTM = \frac{1}{n-2} \sum_{i=1}^{n-2} \delta(d_i) \tag{3}$$

$$\delta(d_i) = \begin{cases} 1 \ if \ \left( [x(i+2) - x(i+1)]^2 + [x(i+1) - x(i)]^2 \right)^5 < r \\ 0 \ otherwise \end{cases} \tag{4}$$

For each radius r, CTM provides the fraction of the total number of points that lie within it. Sample values, which are within a small radius r, correspond to successive rates that are small. These reflect the low frequency components or the slow ascending, descending part of the function in the time series. Sample values that correspond to high r are those of high frequency components or the rapid ascending, descending part of the function in the time series.

### *D. Multivariate Gassian distribution*

The criterion of the data can be described through the Gaussian distribution. This distribution shows the most probable conditions of a random variable. A multivariate Gaussian distribution is a probability distribution of higher dimensional random variables. From the single dimensional normal distribution we have the following equation

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{5}$$

where, $x$ is the random variable , $f(x)$ is the distribution value, $\sigma$ is standard deviation and μ is the mean of x. Distribution of higher directional data is defined by the mutual covariance matrix of the vectors of random variable instead of standard deviation

$$f_x(x_1, x_2, \cdots x_n) = \frac{1}{\sqrt{(2\pi)^n |S|}} e^{\left(-\frac{1}{2}(x-\mu)^T S^{-1}(x-\mu)\right.} \tag{6}$$

where, $S$ is the covariance of variable of n dimension. The exponential part of the equation $\frac{1}{2}(x-\mu)^T S^{-1}(x-\mu)$ is a quadratic term which forms a parabola. The peak of the parabola is the mean of the variable $x$.

### *E. Linear Discrimination Analysis (LDA):*

The famous Naïve Bayes rule stated that for multidimensional feature vector, each class has separate feature density. Therefore, to estimate the class density, the density for every dimension can be estimated separately and then multiply to get joint density. In LDA density for feature space for every class is estimated using Gaussian distribution formula as given in (2). LDA assumes that the normal density function for all classes have the same covariance. Say there are $K$ classes. Let $x_k$ be a $p \times N_k$ matrix of $N_k$ samples, having $p$-dimensional feature from class $k$. Denoted that the prior probabilities $\pi_k$, means $\mu_k$ of each class and the common covariance matrix $S$, a new data point $x$ is classified by maximizing the probability from the normal density function $f_k(x)$.

$$Class(y) = \arg\max_k \left[ f_x(x) \right] \tag{7}$$

This maximization of distribution is subjected to

$$Class(y) = \arg\max_k \left[ x^T S^{-1} \mu_k - \frac{1}{2} \mu_k^T S^{-1} \mu_k + \log(\pi_k) \right] \tag{8}$$

The resulting LDA decision boundaries between classes of data are linear.

### *F. Quadratic Discriminant Analysis:*

Same as LDA, Bayes rule of classification is followed by QDA. The probability of the feature is derived from the

distribution and same as LDA the normal distribution theory is applied. In LDA, as we mentioned, you simply assume for different k that the covariance matrix is identical. By making this assumption, classifier becomes linear. The only difference in quadratic discriminant analysis is that we do not assume the covariance matrix is identical for different classes.

$$S_i \neq S_j \text{ for some } i \neq j; i, j \in k$$

Quadratic discriminant analysis calculates a Quadratic Score Function which looks like this:

$$\delta_k(x) = -\frac{1}{2}\log|S_k| - \frac{1}{2}(x-\mu_k)_k^T S^{-1}(x-\mu_k) + \log(\pi_k) \quad (9)$$

The quadratic term of the function $-\frac{1}{2}(x-\mu_k)_k^T \Sigma^{-1}(x-\mu_k)$ cannot be avoided. The decision boundary is determined by the quadratic function. The decision rule of the classifier is:

$$Class(y) = \arg\max_k [\delta_k(x)] \quad (10)$$

### G. Fisher's Linear Discriminant Classifier

Fisher's linear discriminant is a classification method that projects high-dimensional data onto a line and performs classification in this one-dimensional space. The discrimination in feature space of different classes can be increased by maximizing between-class to within-class scatter ratio. The goal of this maximization is to find the vector on which the projections of the samples are to be spread as much as possible. The projection maximizes the distance between the means of the classes while minimizing the variance within each class. The ratio of between-class to within-class scatter matrixes is known as fisher's criterion function [13]. Equation 11 shows the mathematical expression of fisher's criterion function.

$$J(u) = \frac{u^T S_B u}{u^T S_W u} \quad (11)$$

where, $S_B$ and $S_W$ are respectively between-class and within-class scatter matrixes. The Eigen decomposition of ratio of the scatter matrixes gives the vector $u$ on which the projections of the samples have the larger discrimination than the feature space. This vector $u$ is called the fisher's linear vector.

## III. Experiment

### A. EEG Signal Recordings:

EEG signals are acquired through standard 10-20 system. We choose only O2 channel because of the following reasons:

- Linked with visual perception, alpha activity is larger in the occipital region as it is and
- There are fewer artifacts, such as ocular muscle activity, in this region compared to the frontal scalp regions.

The number of subject was 25 and each of the subjects had 16 trials. The age of the subjects were in between 15-30. Each trial had started with close eye condition then the subject asked to close his eye and it ends with open eye again. Active electrodes are used with a conductive gel to have direct electric contact to the scull. The voltage difference was measured with reference at ear and ground on forehead. Hear gel or other cosmetics were prevented for the subjects because the cosmetics can make insulation between skin and electrodes. The artifacts from other sources were regarded as white noise and removed by filtering. The sampling rate of the signals was 256Hz and acquired in a windows operating based computer of quad core processor. The signals were segmented in 68 samples in each segment. Then the segments were passed through a band pass filter. The alpha response of the signal was extract by the band pass filter with 8 Hz lower cut off and 13 Hz higher cut off frequency.

### B. Feature Extraction

We selected 200 of the trials as training signals. The features are extracted from the signals in Labview Engineering platform. The segments are transformed in frequency domain, and then the PSD was measured by the (1). PSD plot of a trial is shown in Fig. 2. Second order plots of the same segments were derived for measuring the CTM. The second order plot and the CTM are showed in Fig. 3 and Fig 4. PSD and CTM value of each segment were selected as a single sample for corresponding feature vector.

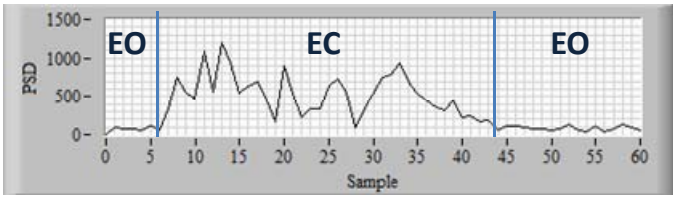



Figure 2: PSD plot of EEG signals.

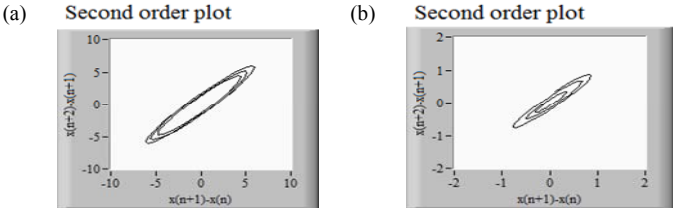



Figure 3: Second Order plot of EEG signals at (a) CE n, (b) OE condition

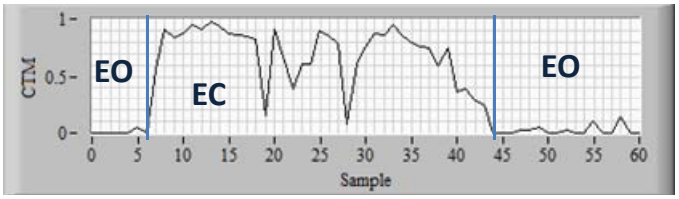



Figure 4: CTM plot of EEG signals.

### C. *Multivariate Gaussian Distribution Analysis*

Multivariate analysis of training samples showed that the discrimination among the EEG signals in close eye and open eye conditions. From the Fig. 5 it is clearly visible that the discrimination among the classes in the PSD and CTM feature space is good enough for classification.

## IV. RESULTS

The parameters of LDA and QDA classifiers are found form the training trials. The remaining 200 trials were selected for testing. The rate of true positive of the classification is showing in table 1. The sensitivities of LDA, QDA and FLDA are found 94.27%, 95.09% and 95.96%. Performance of QDA is slightly higher than LDA because of the nonlinear discrimination boundary. The larger discrimination in one dimensional projection space gives the best performance in FLDA classification. We implied the FLDA to the switching program. The primary result of switching program for real-time trial is found to be satisfactory (Fig. 6). We sated the switching time was 5 sec. The logic of push button SPST switch was applied.

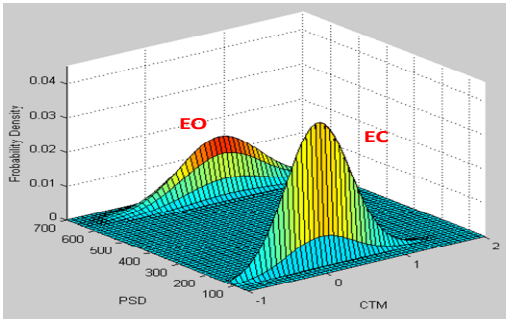



Figure 5: Multivariate Gaussian distribution plot of EEG signals in PSD and CTM feature space.

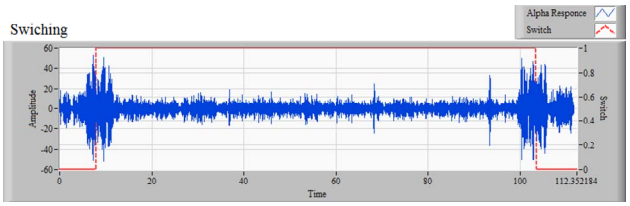



Figure 6: Switching mechanism for ECS using EEG.

TABLE I. PERFORMANCE OF CLASSIFIERS

| Classifier | Close | | Open | | Total | | Sensitivity |
|---|---|---|---|---|---|---|---|
| | TP | FN | TP | FN | TP | FN | |
| LDA | 1189 | 178 | 6299 | 277 | 7488 | 455 | 94.27% |
| QDA | 1131 | 236 | 6422 | 154 | 7553 | 390 | 95.09% |
| FLDA | 1255 | 112 | 6367 | 209 | 7622 | 321 | 95.96% |

## V. DISCUSSION AND CONCLUSION

The multivariate analysis has shown that the there is a linear discrimination among the sample from close and open eye conditions. The result of classification is almost near for quadratic and linear boundary. Even though the performance is slight low in LDA, this can be considered for computational complexity of QDA. Now days, advanced machine learning algorithms can define the classes better. From the study of classifiers on EEG Support Vector Machine found to be more robust and stable than other [14]. Support Vector Machine is a linear classifier and faster than other complex learning algorithms. Hidden Markov Model can is a good model for state prediction of switching. The state of Hidden Markov Model are depended on last state and current observation so that the switching state of ECS. The environment control system for disable should have consistence and as perfect as possible. Though the sample classification rate is good correct, a miss on switching can make a big trouble for disable people. To get a 100% perfect switching system our suggestion is to use a classifier like SVM for sample classification and a prediction model like HMM or Kalman filter for switching.

## REFERENCES


[1] Craig, A., P.M. YvonneTran, and P. Boord, "The efficacy and benefits of environmental control," *Med Sci Monit*, vol. 11 no. 1, pp. 39, 2004.

[2] Jackson, M.M. and G. Birch, "A general framework for characterizing studies of brain interface technology," *Annals of biomedical engineering*, vol. 33, no. 11, pp. 1653-1670, 2005.

[3] Postelnicu, C.C., F. Girbacia, and D. Talaba, "EOG-based visual navigation interface development," *Expert Systems with Applications*, 2012.

[4] Deng, L.Y., et al., "EOG-based Human–Computer Interface system development," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3337-3343, 2010.

[5] Zander, T.O., et al., "Combining eye gaze input with a brain–computer interface for touchless human–computer interaction," *Intl. Journal of Human–Computer Interaction*, vol. 27, no. 1, pp. 38-51, 2010.

[6] Kirkup, L., et al., "EEG-based system for rapid on-off switching without prior learning," *Medical & Biological Engineering & Computing*, vol. 35, no. 5 pp.504-509, 1997.

[7] Blankertz, B., et al., *The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single w.* IEEE Trans. Biomed. Eng., vol. 51, no. 6, pp. 1044-51, 2004.

[8] Bazanova, O., "Comments for Current Interpretation EEG Alpha Activity: A Review and Analysis," *Journal of Behavioral and Brain Science*, vol. 2 no. 2, pp. 239-248, 2012.

[9] Craig, A., et al., "Improving correct switching rates in a'hands-free'environmental control system," *Journal of Neural Engineering*, vol. 2, no. 4, pp. L9, 2005.

[10] Thuraisingham, R.A., et al., "Analysis of eyes open, eye closed EEG signals using second-order difference plot," *Med Biol Eng Comput*, vol. 45, no. 12, pp. 1243-9, 2007.

[11] Malmivuo, J. and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, USA, 1995.

[12] Jung, T.-P., et al., "Estimating alertness from the EEG power spectrum," *IEEE Transactions Biomedical Engineering,*vol.44, no. 1, pp. 60-69, 1997.

[13] Mika, S., et al. "Fisher discriminant analysis with kernels," in *Neural Networks for Signal Processing IX, 1999. Proceedings of the 1999 IEEE Signal Processing Society Workshop.*Aug. 1999, pp. 41-48.

[14] Lotte, F., et al., "A review of classification algorithms for EEG-based brain–computer interfaces," *Journal of Neural Engineering,* vol 4, no. 2, pp. R1-R13, 2007.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Human Ear Recognition Using Geometric Features

Mostafizur Rahman, Muhammad Sheikh Sadi, Md. Rezwanul Islam
Department of Computer Science and Engineering, Khulna University of Engineering and Technology,
Khulna-9203, Bangladesh
[1] E-mail: mony4k@live.com

***Abstract*—Biometrics includes the study of automatic methods for distinguishing human beings based on physical or behavioral traits. Finding good biometric methods has been researched extensively in recent years. Among several biometric features, ear is quite stable because it does not vary with age and emotion. The ear recognition works based on the height of the ear, reference line cut point and corresponding angles. The study is performed on the ear in random orientation and shows a greater accuracy than existing dominant approach. The recognition accuracy is increased by removing the noise in captured ear images and developing new methodologies to work with online images.**

***Keywords—Ear biometrics, curve fitting, ear orientation, ear height line, reference lines, angles.***

## I. INTRODUCTION

Human identification is a great challenge in many commercial system and private system for authorized access [1]. Traditional identification system includes Personal Identification Number (PIN), password, displaying ID, badges and so on. But traditional identification systems have many disadvantages like PIN, password could be forgotten or hacked by masqueraders and ID badges could be lost [2]. Traditional identification process is an active identification process where user has to take part in the identification process; this is another major drawback of this system.

There are numerous biometric identification methods based on image analysis like face recognition, iris recognition, finger print, ear biometrics and so on. Human ear does not change with time and age. Ear satisfies all the properties that should be possessed by a biometric, i.e. uniqueness, universality, performance and collectability [3]. It has complex structure with uniqueness which is used to identify individuals. Even, ear of twins has different structure. In many criminal investigation ears are used as strong evidence [4]. In case of accidents, ears remain unchanged and could be used for investigation. Ear has been used for many years in forensic sciences for recognition. It has not required any direct input to the authentication process.

Ear recognition and identifications has been done in several ways before. Almost all approaches are still facing the problem to reach an acceptable accuracy with online images. There are huge gaps in the pre-processing stage of ear images as well. Here this paper proposes a methodology to recognize online ear with a better accuracy. This paper contains related work, methodology of the ear recognition system, experimental analysis, conclusions in the next section.

## II. REALATED WORK

Ear detection and identification can be done by several ways such as shape model based [5], Helix Shape Model [6], Histograms of Categorized Shapes [7], connectivity graph [8], Cascaded adaboost [9], Edge detection and curvature estimation [10], Ray transform [11], Ovoid model [12], Edge detection and line tracing [13]. In shape model based ear detection ears' image regions are selected by a large local curvature with a technique which is called step edge magnitude. Cascaded adaboost uses weak classifiers based on Haar-wavelets in connection with AdaBoost for ear localization. The ray transform approach proposed in is designed to detect the ear in different poses and to ignore straight edges in the image, which can be introduced by glasses or hair. Ray transform uses a light ray analogy to scan the image for tubular and curved structures like the outer helix. For precise ear segmentation, the contour fitting method based on modified active shape models is introduced. So, ear edge can be detected in many ways but if the edge contain noise or inappropriate for extracting features then detection system will be fall. Curve smoothing and make it continuous it is important to fit the edge curve and find the proper orientation.

Chen and Bhanu [5] proposed three different approaches for ear detection. In the approach from Chen and Bhanu [5] train a classier, this recognizes a specific distribution of shape indices, which are characteristic for the ear's surface. However this approach only works on profile images and is sensitive to any kind of rotation, scale and pose variation.

In their later ear detection approaches Chen and Bhanu detected image regions with a large local curvature [5] with a technique they called step edge magnitude [6]. Then a template, which contains the typical shape of the outer helix and the anti-helix, is fitted to clusters of lines. The main limitation of Chen and Bhanu's proposed method [6] is the fixed data set. In random orientation their method will fail to detect the correct person.

In 2005 Michał Choraś [14] suggested new methodology of ear detection. He mentioned that geometrical features representing shapes of ear contours are more suitable for ear images than texture, color or global features. Method is divided into image normalization, contour extraction (edge detection), calculation of the centroid, coordinates normalization and two steps of geometrical feature extraction. He treats the centroid as the specific point in this method, even though it is not a specific point within the ear topology. The main limitation of this methodology is the contour image centroid determination and continuity of the edges.

Abaza et al. [9] and Islam et al. [16] use weak classifiers based on Haar-wavelets in connection with AdaBoost [9] for ear localization. According to Islam et al., the training of the classifier takes several days, however once the classifier is set up, ear detection is fast and effective. Abaza et al. [9] use a modified version of AdaBoost and report a significantly shorter training phase. The effectiveness of their approach is proved in evaluations on five different databases. They also include some examples of successful detections on images from the internet. As long as the subject's pose does not change, weak classi ears are suitable for images which contain more than one subject. Depending on the test set Abaza et al. [9] achieved a detection rate between 84% and 98.7% on the Sheffield Face database. On average, their approach successfully detected 95% of all ears

In 2012 Zahid et al. introduced new geometric method [1] to recognize the ear in identification purpose. Their method was to find the ear edge using canny edge detection [17] from the static ear image and extract geometric features from the edge. They use three vectors as features, first ear height line, reference lines and reference line cut point and corresponding angles. If the edge of the ear image is broken or vague then the whole methodology will be fall and the ear orientation is another limitation of this method. If the ear height line can be in appropriate due to noisy edge and broken edge. So the angle measured will be also incorrect.

## III. METHODOLOGY OF EAR RECOGNITION

Ear recognition is done by processing the ear image properly and extracting the geometric features from processed image. Processing of ear image includes finding outer ear edge, finding ear height, reference line cut point and corresponding angle. Modified ear recognition process flow cart from Zahid et al.[1] is shown in Fig. 1.

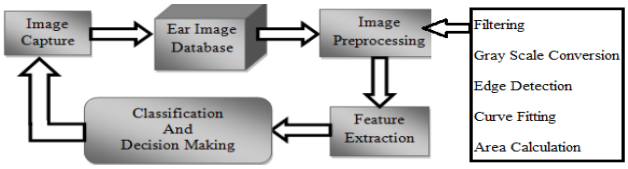



Figure. 1 Flow chart of ear recognition system

This paper proposes a methodology to recognize online ear images using geometric features. The ear recognition system takes place based on ear height line, reference lines and angles [1]. Ear height line is the height of the ear, reference lines are the lines which are parallel to the width of the ear image which divide the image cell into (n+1) parts, where *n* is a positive integer. and angle is outer edge of the ear and one portions of the outer ear edge that creates angles in the midpoint of the ear height line is measured. Fig. 2 shows the reference lines and ear height line sequentially.

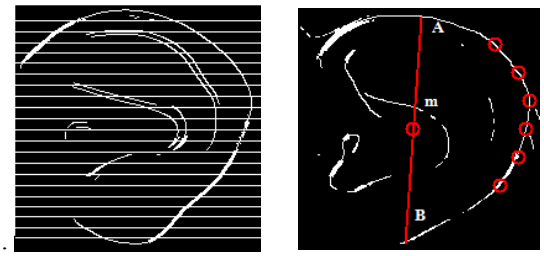



Fig. 2 Reference line and ear height line

Co-ordinate of $b$ is $(b_x,b_y)$ in Fig. 3, which is the top point of ear edge and starting point of ear height line. Similarly co-ordinate of $a$ is $(a_x,a_y)$ which is the midpoint of ear height line. Here point $c$ $(c_x,c_y)$ is the point produced by reference line cutting the outer ear edge. Hence, three points *a, b, c* is known but corresponding angle is unknown. Calculate *ab, bc* and *ca* using Euclidian distance by-

$$ab = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2}$$

Then the cosine of the angle θ between *ba* and *ac* is taken from Fig. 3 is given by-

$$\cos\theta = \frac{\{(ab)^2 + (ac)^2 - (ca)^2\}}{2 * ab * ac}$$

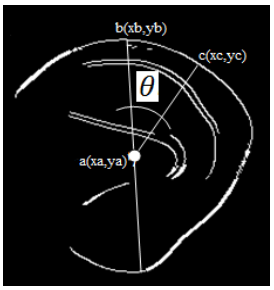



Figure. 3 Angle created by one portion of outer ear edge

Outer ear edge is very important to be detected clearly. So, if there is some noise or broken, incomplete or vague outer edge then the whole methodology will fail [1]. If the orientation of the image is changed, the methodology will not work properly. To overcome the limitations of the method images are processed more precisely. To fix broken or incomplete outer ear edge, 2nd degree curve fitting namely parabolic curve fitting is introduced. To detect ear from random positions, it is needed to identify the proper orientation of the ear. It can be done it from calculating the area occupied by outer ear edge curve.

### *A. Noise Minimization by Curve Fitting*

Ear edges found from the image are not always perfect. So, the noise in the ear image should be removed. Edges are simple 2D curve and some portion may be broken. In Fig. 4 broken edges are indicated clearly.

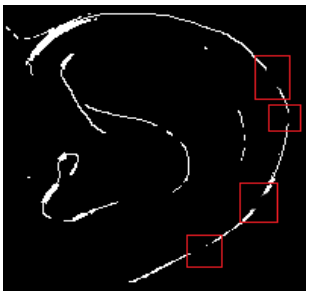


Figure. 4 Ear image with broken edge

Co-ordinates of each pixel are found in 2D co-ordinate system. Then binomial equations are solved to find the value of the co-efficient and fit the curve properly. The outer edge of the ear is parabolic shape. Now, the equation may be found from the outer ear edge as follow:

The curve of outer edge of ear is represented by the parabolic equation:

$$y = a + bx + cx^2 \quad \ldots \quad (1)$$

Here, *a, b* and *c* are three co-efficient in which outer ear edge equation depends on. Equation (1) satisfies given data

$$(x_1, y_1), (x_2, y_2), (x_3, y_3) \ldots \ldots \ldots (x_n, y_n)$$

This is the visible outer ear edge points. Let, $y_\lambda$ be the theoretical value for $x_1$ then,

$$\text{Error, } e_1 = y_1 - y_\lambda$$

$$\Rightarrow e_1 = y_1 - (\text{a} + \text{b}x_1 + \text{c}x^2{}_1)^1$$

$$\Rightarrow e^2{}_1 = (y_1 - \text{a} - \text{b}x_1 - \text{c}x^2{}_1)^2$$

Let,

$$\text{Sum, S} = \sum_{\text{i=1}}^{\text{n}} e_i{}^2$$

$$\Rightarrow \text{S} = \sum_{\text{i=1}}^{\text{n}} (y_i - \text{a} - \text{b}x_i - \text{c}x^2{}_i)^2$$

By the principle of least square, value of S is minimum

$$\frac{\delta S}{\delta a} = 0, \ \frac{\delta S}{\delta b} = 0, \ \frac{\delta S}{\delta c} = 0 \quad \ldots \quad (2)$$

Solving the equation (2) and dropping the suffix we get

$$\sum y = \text{na} + \text{b}\sum x + \text{c}\sum x^2 \quad \ldots \quad (3)$$

$$\sum xy = a\sum x + \text{b}\sum x^2 + \text{c}\sum x^3 \quad \ldots \quad (4)$$

$$\sum x^2 y = a\sum x^2 + \text{b}\sum x^3 + \text{c}\sum x^4 \ldots \quad (5)$$

The equation (3), (4) and (5) are known as normal equation and solving the equation (3), (4) and (5) getting the value of a, b and c. Putting the value of a, b and c in the equation (1), outer ear edge can be obtained and then curve fitting can be introduced in the broken edge curve. It is possible to extract geometric feature from the image of the outer ear at any orientation and this was the big limitation of Zahid and et al. [1] method of ear recognition.

### *B. Ear Orientation and Random Position*

Two ear images shown in Fig. 5 are examined, one is front ear image and another one is back ear, both rear and front ear properties are stored in database. Front ear occupies larger geometric area than the rear one. Comparing their area with a fixed threshold value front and rear ear can be distinguished easily. The threshold value is taken empirically to distinguish rear and front ear. Fig. 3 also shows the ear height line and reference line cut point for both rear and front end.

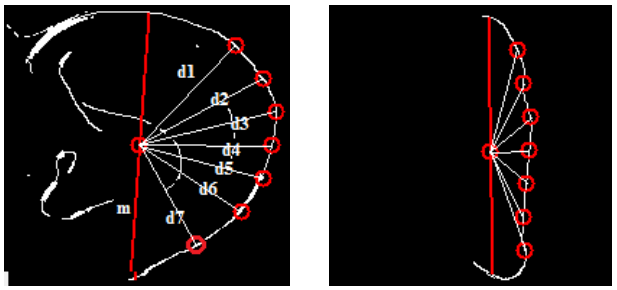



Figure. 5 Front ear and rear ear image

Let, the equation of the ear height line

$$y = b + mx \quad \ldots \quad (6)$$

Where, m is slope of the height line.

Let the equation of the outer ear edge curve

$$y = a + dx + cx^2 \quad \ldots \quad (7)$$

Solving (6) and (7) the two values of x are found. After the limit is being found, the equation (7) is integrated between the limit of x and the outcome is the area between ear height line and outer edge of the ear.

So, area is calculated as the following equation (8)

$$\text{A} = \int_{\min of\ x}^{\max of\ x} y dx \quad \ldots \quad (8)$$

Where, *y* is taken from equation (1).

$$y = a + dx + cx^2$$

Observing the value of the calculated area, orientation of the ear is found.

## IV. EXPERIMENTAL ANLYSIS

The method is implemented in Microsoft Visual Studio 2012. Canny Edge detection algorithm [17] is used here for finding the ear edge. For online identification ear images are captured by Microsoft webcam and particular driver software is used for this purpose. Online ear images are processed according to methodology and then raw images with edges are found as output of the system and curve fitting are applied to the system. Finally, all three feature vectors are extracted and compared with the database of saved ear images and traced out the exact match. The database of ear images contains 50 images and no standard data set is used.

Zahid et al. [1] worked with static images which are captured as front view. It was in appropriate for random orientation, online detection and recognition. Introducing curve fitting and changing the ear orientation Zahid et al. [1] reduces its accuracy at a huge rate. Applying random

positioning of the ear image and without curve fitting it becomes 84.04% accurate. This proposed method has increased accuracy to 96.8% shown in Fig. 4.

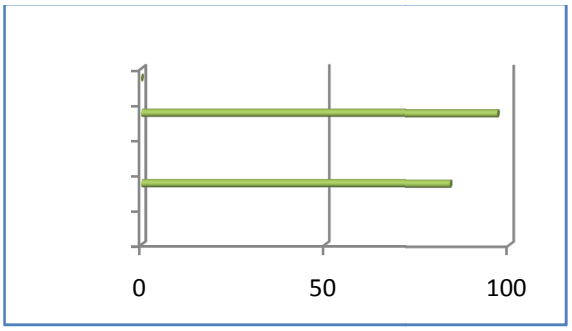



Figure. 6 Comparison between proposed method and other methods

Selecting the value for number of reference lines *n* is a crucial aspect in this approach. A greater value of *n* increases the accuracy but also increases the space requirement and the time taken for classification as more number of features is to be compared. Therefore an optimal value should be selected which satisfies all the three requirements of space, time and accuracy. Error Rate comparison with Zahid et al. [1] is given in Fig. 6.

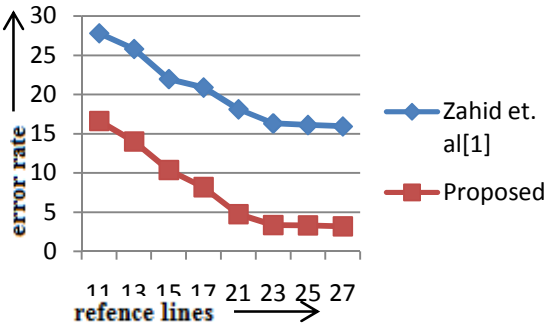



Figure. 7 Graphical representation of our experimental result

The graph in Fig. 6 shows, after taking 21, 23 and 27 reference lines the error rate reduces linearly. But the main fact is that, when curve fitting is introduced in this proposed method experimentally accuracy increased and when random positioning was introduced in earlier methods, its accuracy falls drastically which was overcome in this proposed methodology. Thus proposed methodology proves its best result experimentally.

## V. CONCLUSIONS

This process includes curve fitting of the outer ear edge and finding the proper orientation of the ear. Although proposed methodology works weakly in low hysteresis threshold, it recognizes the ear more accurately and accustomed with the real environment. Moving person's ear can be captured and processed through this methodology and can be traced out with a better accuracy than existing domain approach. This paper proposes a new methodology to recognize ear with a online random orientation. This work can be extended for inner curve fitting of the ear and extracting features of the inner portion of the ear edge.

## REFERENCES


[1] Md. ZahidHasanPolin, A. N. M EnamulKabir, Muhammand Sheikh Sadi, "2D Human-ear Recognition Using Geometric Features," in *Proc. IEEE Int. Conf. on Electrical and Computer Science,* Dec. 2012

[2] M. Choras, "Ear biometrics based on geometrical feature extraction", Electronic Letters on Computer Vision and Image Analysis, 5(3), 84-95

[3] Choras M., "Perspective Methods of Human Identification: Ear Biometrics," Opto-Electronics Review. 2008, pp. 16:85-96

[4] A. Iannarelli, "Ear Identification, Forensic Identification Series", Paramount Publishing Company, 1989

[5] Chen H, Bhanu B, "Shape Model-Based 3D Ear Detection from Side Face Range Images," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition.*, Jun. 2005, pp. 122.

[6] Chen H, Bhanu B., "Contour Matching for 3D Ear Recognition," in *Proc. Of seventh IEEE Workshop on Application of Computer Vision (WACV/MOTION);*Jun. 2005.

[7] M. Abdel-Mottaleb, J. Zhou, "Human Ear Recognition From Face Profile Images," Advances in Biometrics, Vol. 3823, pp. 786-790, 2005.

[8] Prakash S, Gupta P.,"An Efficient Technique for Ear Detection in 3D: Invariant to Rotation and Scale," In: The 5th IAPR International Conference on Biometrics (ICB); 2012.

[9] Abaza A,Hebert C, Harrison MAF, "Fast Learning e Ear Detection for Real-time surveillance," in *Proc. IEEE Int. Conf. Biometrics: Theory Applications and Systems(BTAS-2012)* ., Apr.2010, pp. 1-6.

[10] Ansari S, Gupta P., "Localization of Ear Using Outer Helix Curve of the Ear," In: International Conference on Computing: Theory and Applications; 2007. pp. 688-692.

[11] Alastair H, Cummings AH, Nixon MS, Carter JN.,"A Novel Ray Analogy for Enrolment of Ear Biometrics," In: Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS); 2010.

[12] Alvarez L, Gonzalez E, MazorraL. ,"Fitting ear contour using an ovoid model," In: 39th Annual 2005 International Carnahan Conference on Security Technology (CCST '05); 2005.

[13] Attarchi S, Faez K, Raei A., "A New Segmentation Approach for Ear Recognition," In: Blanc-Talon J, Bourennane S, Philips W, Popescu D, cheunders P, editors. Advanced Concepts for Intelligent Vision Systems. vol. 5259 of Lecture Notes in Computer Science. Springer Berlin Heidelberg; 2008. pp. 1030-1037

[14] MichałChoraś, "Contour Matching for 3D Ear Recognition," in *Proc. IEEE Int. Conf. Electronic Letters on Computer Vision and Image Analysis.*, Jan. 2005, pp. 87.

[15] Abaza A,Hebert C, Harrison MAF, "Fast Learning e Ear Detection for Real-time surveillance," in *Proc. IEEE Int. Conf. Biometrics: Theory Applications and Systems(BTAS-2012)* ., Apr.2010, pp. 1-6.

[16] Islam S, Davies R, Bennamoun M, Mian A, "Efficient Detection and Recognition of 3D Ear," *IEEE Trans. Computer Vision, pp.; 95:52*-73. Sep. 2011.

[17] J. Canny, "A Computational Approach to Edgfe Detection," IEEE Trans. on Pattern Analysis and Machine Intelligence, 8(6), pp. 679-698, 1986.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# An Empirical Robotic Framework for Interacting with Multiple Humans

Mohammed Moshiul Hoque[1,*], Quazi Delwar Hossian[1], Dipankar Das[2], Yoshinori Kobayashi[2], Yoshinori Kuno[2], and Kaushik Deb[1]
[1]Chittagong University of Engineering and Technology, Bangladesh
[2]Saitama University, Japan
[*] E-mail: moshiul_240@cuet.ac.bd

*Abstract*—**currently work in robotics is expanding from industrial robots to robots that are employed in the living environment. For robots to be accepted into the real world, they must be capable to behave in such a way that humans do with other humans. This paper focuses on designing a robotic framework to interact with multiple humans in a natural and social ways. To evaluate the robotic framework, we conducted an experiment to perform an important social function in any conversation such as initiating interaction with the target human in multiparty setting. Results show that the proposed robotic system is functioning to initiate an interaction process in four viewing situations.**

***Keywords—Human-robot interaction, initiating interaction, viewing situations, face-to-face, evaluation.***

## I. INTRODUCTION

Human-robot interaction (HRI) is an interdisciplinary research field aimed at improving the interaction between human beings and robots and to develop robots that are capable of functioning effectively in real-world domains, working and collaborating with humans in their daily activities. To behave in a socially acceptable manner, the robots are capable of performing a lot of social and cognitive functions. Although a number of significant challenges remained unsolved related to the social capabilities of robots, the robot that can initiate an interaction process proactively with a particular human's is also an important research issue in the realm of natural HRI. In this paper we focus on such a social function such as, *initiating interaction*.

Initiating interaction with the intended human when s/he is involving her/his task is one of the fundamental skills in human social interaction and cognition. Initiating interaction plays a critical role in a wide range of social behaviors: it sets the stage for learning, facilitates communication, and supports conveying own intention to the others. Many robots were equipped with the capability to encourage people to initiate interaction [1]-[3]. These robots always wait for people to approach them first, which is one strategy for robots to an initiate interaction process. Alternatively, in order to initiate an interaction process a social robot needs to attract the attention of a target human from his/her current focus to what is sought by the robot. This way of providing services is more proactive than waiting, since it enables robots to help people who have potential needs.

It is apparent that the robot can attract the attention of the people by voice, but using voice is certain to attract other people's attention. Thus, voice should be used as a final resort. The robot should start with a weak action to avoid attracting other people than the target person and use stronger actions if it cannot attract his/her attention. This is the basic design concept of our robot.

Most of the previous works assumed that the human faces to the robot when their interaction begins [4], [5]. However, this assumption may not be practical in natural human-robot communication. Therefore, our major concern is-how the robot can capture the target humans' attention when s/he is not facing toward the robot and engaging his/her current task. In this paper, we simulate the four positional relationships (hereafter we called 'viewing situation') between the robot and the human in terms of the robot's position within the human's field of view. The viewing situations are central field of view (CFOV), near peripheral field of view (NPFOV), far peripheral field of view (FPFOV), and out of field of view (OFOV).

## II. RELATED WORK

The capabilities of social robots that can initiate an interaction process with the target human are still at a rudimentary stage. Moreover, there has not been significant work conducted yet about how humans attract others' attention to initiate an interaction beyond the primary facts that they stop at a certain distance [6], start the interaction with a greeting [7], and arrange themselves in a spatial formation [8]. Some robots were initiate interaction capabilities by offering cues such as approach direction [9], approach path [10], and standing position [11]. These studies assumed that the target person faces the robot and intends to talk to it; however, in actual practice this assumption may not always hold. Robots may wait for a person to initiate an interaction. Although such a passive attitude can work in some situations, many situations require a robot to employ a more active approach [12–14]. Moreover, most of the previous HRI studies were used some robotic heads such as Kismet [15], Andriod [16], and Robovie [17]. These heads are mechanically very complex and as such expensive to design, construct and maintain. Some robots used a flat screen monitor as the robot's head [18]. However, a flat screen

monitor is unnatural as a face. In our previous work [19], we used a retro-projected CG images as the robot's eyes which is very cheap. However, it can only interact with a single participant. In this paper, we extend our previous robotic framework to interact with the multiple participants.

## III. System Overview

The purpose of our research is to develop a module for robots that can initiate an interaction process with a target human while s/he is attending his/her task at different viewing situations. Thus, for serving our purpose, we have developed a robotic head. In the following sections, we discuss the architecture of our robotic systems and its behaviors.

### A. Hardware Components

We have developed a robotic head for human-robot interaction experiments. Fig. 1 shows an overview of our robotic head. The head consists of a spherical 3D mask, an LED projector (3M pocket projector, MPro150), Laser range sensor (URG-04LX by Hokuyo Electric Machinery), three USB cameras (Logicool Inc., Qcam) and a pan-tilt unit (Directed Perception Inc., PTU-D46).

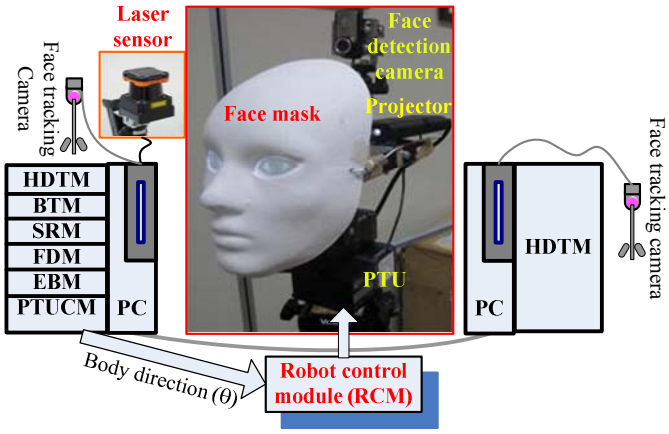



Figure 1. System consists of six modules: HDTM (head detection and tracking module), SRM (situation recognition module), BTM (body tracking module), EBM (eye blinks module), RCM (robot control module) and PTUCM (pan-tilt unit control unit module).

The 3D mask and projector are mounted on the pan-tilt unit. In the current implementation, One USB camera is wired on the top of the mask to detect frontal face of human (shown in Fig. 1). The other two cameras and the laser sensor are put on the tripods placed at appropriate positions for observing participants' heads as well as their bodies to recognize their viewing situations.

To provide a communication channel between the hardware components of the system, there is a standard RS-232 serial port connection between the general purpose PC's (Windows XP) and the pan-tilt unit. To ensure smooth operation, the pan and tilt speeds of the robot's head has adjusted actively when the participant is moving. The LED projector projects CG generated eyes on the mask. Thus, the head can show non-verbal behaviors by its head and eye movements including blinking. The system utilizes the two general purpose computers (Windows XP), each connected with an USB camera.

### B. Softwares Modules

The proposed system has six software modules: the body tracking module (BTM), the head detection and tracking module (HDTM), the situation recognition module (SRM), the eye-contact module (ECM), the robot control module (RCM), and the pan-tilt unit control module (PTUCM).

*a) Body tracking module (BTM):* After detecting the target person and his/her current attentional focus by using the information of HDTM, the robot should turn its head toward him/her for making eye contact. To turn its head toward the target person, the robot should know the location information of his/her gaze, or head or body. Since in our scenario, the gaze and head directions of the target person is changing abruptly due to the watching paintings in several viewing directions. However, he/she is watching paintings from a fixed sitting position. Thus, we can assume that his/her body direction may not varied much due to the different viewing conditions. Our robotic agent continuously tracks the body direction of the target participant in real time using a laser range sensor and computes his/her body positions (x, y), directions (θ), and distance (D). Therefore, the robotic head turn its head toward the target person by using the information from the BTM.

A human body can be modeled as an ellipse (Fig. 2 (a)). We assume the coordinate system is represented with their X and Y axes aligned on the ground plane. Then, the human body model is consequently represented with center coordinates of ellipse [α, β] and rotation of ellipse (ϕ). These parameters are estimated in each frame by the particle filter framework [20]. We assume that the laser range sensor is placed on the participant's shoulder level so that the contour of his/her shoulder can be observed. When the distance data which captured by the laser range sensor is mapped on the 2D image plane, the contour of the participant's shoulder is partially observed.

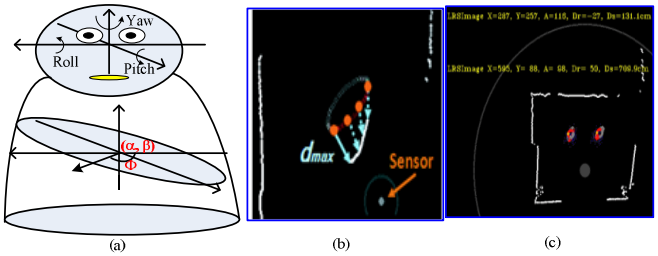



Figure 2. (a) Human body model (b) Likelihood evaluation based on maximum distance (c) Tracking results of BTM

The likelihood of each sample is evaluated the maximum distance between evaluation points and the nearest distance data using the Eq. 1.

$$\theta = \exp\left(\frac{-D_{\max}^2}{\partial_D}\right) \quad (1)$$

where θ is the likelihood score based on the laser image, $D_{max}$ is the maximum distance between evaluation points and the

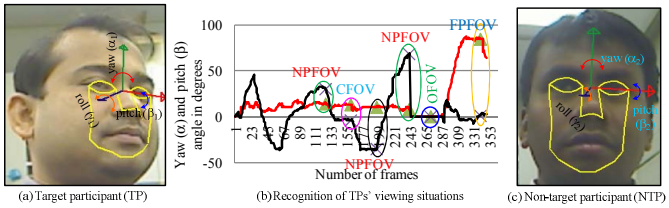
nearest distance data. At each time instance, once the distance image is generated from the laser image, each distance $D_n$ is easily obtained. $\sigma_D$ is the variance derived from $D_n$. Evaluation procedures are repeated for each sample. Conceptual images of evaluation process are shown in Fig. 2 (b). We employ several points on the observable contour as the evaluation points to evaluate hypotheses in the particle filter framework. These points are changes depend on the relational position from the laser range sensor and the orientation of the model. Selection of evaluation points can be performed by calculating the inner product of normal vectors on the contour and its position vector from laser range sensor. An example of the results of the BTM is shown in Fig. 2 (c).

*b) Head detection and tracking module (HDTM):* To detect, track and computes the direction of human head in real time (30 frame/sec), we use FaceAPI [21] by Seeing Machines Inc. It can measure 3D head position (x, y, z) and direction (yaw (α), pitch (β), and roll (γ)) within $3^0$ errors. Coordinates in head frame are usually measured in meters and orientation is radians. Fig shows an example of the tracking result using HDTM. One USB camera is placed in front of the human to track his/her face up to $\pm 90^0$. Figs. 3 (a) and (c) shows an example of the tracking results using HDTM. The head coordinate frame is a right-handed 3D reference frame and the origin is the midpoint of the line that joins the center of the eyes sclera spheres. When viewed from in front of the face, the x-axis points horizontally to the left toward the person's right eye, the y-axis points vertically upward, and the z-axis points away from the viewer, towards the back of the head. The results of the HDTM send to the ECM and PTUCM to determine the current attentional direction toward the robot.

*c) Situation recognition module (SRM):* To perform an effective action, the robot should identify the social situation first. Therefore, to recognize the existing situation (where the human is currently looking), we observe the head as well as body direction estimated by HDTM and BTM respectively. By extrapolating from the person's head and body information, the SRM determines which situation (CFOV, NPFOV, FPFOV, or OFOV) is exists. The HDTM tracks within $\pm 90^0$ (right/left) only, therefore, while the human attend to OFOV situation, the system losses head information, in that case, the robot recognize the current situation based on the body information (laser sensor can tracks up to $270^0$). From the results of HDTM and BTM, the system recognizes the following four situations in terms of yaw (α), pitch (β), and body (θ) movements using a set of *if-then-else* rules.

- *Central field of view (CFOV)*: recognized if the current head direction within $-10^0 \le \alpha \le +10^0$ and $-10^0 \le \beta \le +10^0$ and remains 30 frames in the same direction.
- *Near peripheral field of view (NPFOV)*: recognized if the current head direction within $-10^0 > \alpha \ge +70^0$ or $+10^0 \le \alpha \le +70^0$ and $-10^0 \le \beta \le +10^0$ and remains 30 frames in the same direction.
- *Far peripheral field of view (FPFOV)*: recognized if current head direction within $-70^0 > \alpha \ge +90^0$ or $+70^0 \le \alpha \le +90^0$ and $-10^0 \le \beta + 10^0$ and remains 30 frames in the same direction.
- *Out of field of view (OFOV)*: recognized if the human looking to the opposite direction with respect to robot's direction. That means, the robot cannot capture the target human's face or head. In that case, the current head direction occurs within $\alpha = \beta = 0^0$ or body direction within $90^0 < \theta \le +270^0$ or $-90^0 > \theta \ge -270^0$ and remains 30 frames in the same direction.

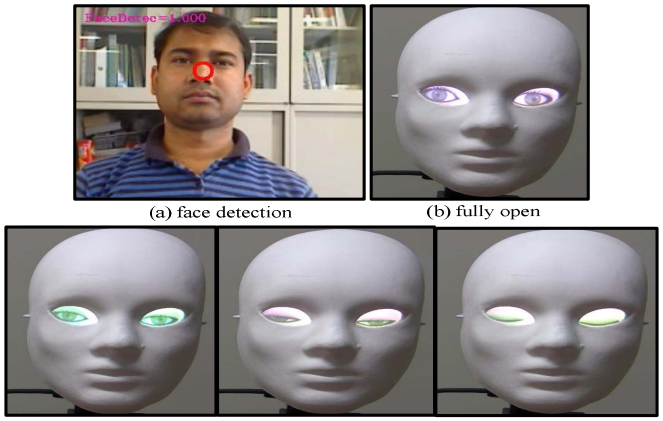
Fig. 3 (b) represents the results of SRM to recognize three situations (i.e., CFOV, NPFOV, FPFOV, and OFOV respectively) in terms of head information.



(a) Target participant (TP) (b) Recognition of TPs' viewing situations (c) Non-target participant (NTP)

Figure 3. Results of HDTM and SRM.

*d) Eye-contact module (ECM):* The ECM mainly consists of two sub modules: the face detection module (FDM), and the eye blinking module (EBM). The robot continuously checks the target person's whether his/her face directed to the robot or not. The robot considers that the human has responded against the robot actions if s/he looks at the robot within expected times. In that case, FDM uses the image of the forehead camera to detect his/her frontal face [22] (Fig. 4. (a)). After face detection, FDM sends the results to EBM for exhibiting eye blinks. Since the eyes are CG images, the robot can easily blink the eyes in response to the human's gazing at it. Figs. 4 (b)-(e)) illustrates some screenshots of eye behaviors of the robot.


(a) face detection (b) fully open

(c) partially closed (d) partially open (e) fully closed

Figure 4. Result of ECM: detect of face (a) and exhibits several actions (b-e) of eye contact behavior.

*e) Robot control module (RCM):* Integrates all sensor processing results of both computers and sends proper control signals to the PTUCM based on the results of other modules.

*f) Pan-tilt unit control (PTUCM): ):* In order to capture the human attention toward the robot, it needs to perform several actions (such as, head turning, head shaking, uttering

reference terms, etc.). The PTUCM provides all kinds o physical actions based on the output of the RCM. Fig. 5 shows the PTUD-46 device used for our current design.

Figure 5. Pan-tile unit (PTUD-46).

### C. *Behavior of the Robot*

Turning the head toward the target person is the most fundamental action of the robot. Simple head turning or eye movements may be enough when the robot exists in the human's central view but not effective in all cases. Therefore, the robot may need to use stronger signal in some situations than others. To initiate an interaction process the robot performs the following steps consecutively. *Step 1*: the robot detects and tracks position of the participants based on the information of HDTM and BTM. *Step 2*: recognize the current viewing situation of the participants. *Step 3*: If the situation is CFOV, or NPFOV, the robot turns its head to the participant. Go to step 4 for checking his/her response. *Step 3.1*: If the situation is FPFOV, the robot turns its head to the participant and shakes its head to capture his/her attention. Go to step 4 for checking the participants' response. *Step 3.2*: If the situation is OFOV, the robot uttering reference terms ('excuse me'). Go to step 4 for checking the participant's response. *Step 4*: the robot waits (in particular up to 2 seconds) for the participants' looking response. If the participant is attracted (i.e. looked at the robot) within this period, s/he turns his/her face toward the robot. Now, the participant and the robot are in face-to-face, and go to step 5. Otherwise, the robot gives up initiating interaction process. *Step 5*: the robot detects the participant's face. The robot considers that the participant has responded to its actions if s/he looks at the robot within expected times, which is recognized by detecting the participant's frontal face in the camera image. *Step 5.1*: the robot performs blinking action [up to three times] to show the gaze awareness.

## IV. EXPERIMNETS

In order to evaluate the system, we performed an HRI experiment. We recruited 36 participants (Average age = 26.8, SD = 3.72). All of them are graduate students of Saitama University, Japan. Two participants interacted at a time.

To prompt participants to look in various directions, we hung seven paintings (P1-P7) on the wall at the same height (just above the eye level of the participants). These paintings were placed in such a way that, when observed from a participant's sitting position, they covered their whole field of view (close to $180^0$). To produce the stimuli, we prepared two robotic heads with the same appearance. Only the eyes of the robotic head are retro-projected and the rest of the face is opaque. The mere existence of such robots in an environment may prompt participants to be attracted to them because of their human-face-like appearance, even if they do not perform any actions. Thus, two robots were placed in the participant's left and right monocular fields of view. One robot is called the static robot (SR) and it was stationary at all times. The other is called moving robot (MR) and it performed the initiating interaction process. The placement of robots was exchanged randomly.

In order to initiate the interaction process, the robot tried to attract the target participant's attention while he/she was looking at different paintings so that it could obtain data for four types of viewing situations: the central field of view (e.g., looking at paintings P1), the near peripheral field of view (e.g., looking at paintings P2, or P3), the far peripheral field of view (e.g., looking at paintings P4 or P5) and the out of field of view (e.g., looking at paintings P6 or P7). The moving robots initiated the attention attraction process after recognizing the viewing situation. In this experimental scenario, if the target participants did not gaze at the robot or the object within the expected time-frame following the robot's actions, then the robot considered the case to be a failure and completed the experiment. Fig. 6 depicts an image from the experiment. Two video cameras were placed in appropriate positions to capture all interactions during the experiment.

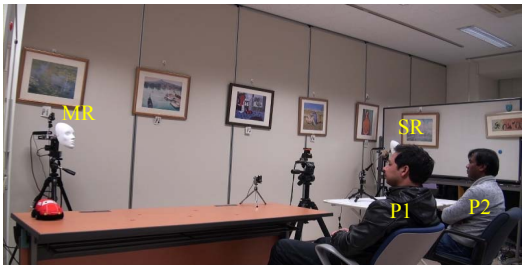



Figure 6: an experimental scene where the participant 1 (P1) was attracted by the robot's action.

## V. RESULTS

The experiment had a within-participant design, and the order of all experimental trials was counterbalanced. In order to evaluate the system, we used the following two measures: quantitative (two items: *Performance evaluation, and success ratio*) and subjective (one item: *impression of robot).*

### A. *Performance Evaluation*

We have identified the several characteristics of the robotic head by performing several experimental trials. The details performance analysis of the body tracking module has been found in [21]. Table 1 summarizes the properties of the robotic head.

TABLE I. PROPERTIES OF THE ROBOT HEAD

| **Characteristics** | **Capabilities** |
|---|---|
| Head turn | (horizontal) |
| from $-159^0$ (left) to $+159^0$ | (right) |
| Head turn | (vertical) |
| from $-47^0$ (down) to $+31^0$ | (top) |
| Eye turn from | $-90^0$ (left) to $+90^0$ (right) |
| Eye blinks | 2/seconds |
| Speed | $300^0$/second |
| Head tracking | from $-90^0$ (left) to $+90^0$ (right) |

| | |
|---|---|
| Body tracking (error $<3^0$) | upto $270^0$ |
| Distance for body detection (error $<6^0$) | upto 3meters |
| Behavioral capabilities | attention attraction, eye contact |
| No. of people tracking | 02 |

### B. Success ratio

Each participant interacted 3 times in each situation. Thus, from the interaction videos (total =3 (times) x 36 (participants) x 4 (situations) = 432), we measured the four kinds of *success ratios.* (i) *Success ratio for attracting attention($SR_{AA}$):* This was measured by the ratio of the number of target participants looking at the robot ($N_R$) and the number of times that the robot attempted to attract their attention toward it ($N_A$), (ii) *Success ratio for situation recognition module* ($SR_{SRM}$): refers to the ratio between the numbers of cases where the robot accurately recognize the situation ($N_{VS}$) and total number of cases where the participant looked at the paintings ($N_L$), (iii) *Success ratio for face detection module ($SR_{FDM}$)*: refers to ratio between the number of times that detect the face ($N_F$) and total number of cases where the participant looked at the robot ($N_{MR}$), and (iv) *Success ratio for eye blinks module ($SR_{EBM}$)*: refers to the ratio between the number of times blinks ($N_B$) and the total number of cases where the participant looked at the robot ($N_{MR}$). Table II summarizes the success ratios for several phases. For each module, success ratio is higher in CFOV and NPFOV situations than the others.

TABLE II RESULTS OF QUANTITATIVE ANALYSIS

| Modules | Viewing Situations | | | | Success Ratio |
|---|---|---|---|---|---|
| | ***CFOV*** | ***NPFOV*** | ***FPFOV*** | ***OFOV*** | |
| $SR_{AA}$ ($N_R/N_A$) | 108/108 | 106/108 | 103/108 | 102/108 | 97% |
| $SR_{SRM}$ ($N_{VS}/N_L$) | 107/108 | 105/108 | 100/108 | 104/108 | 96% |
| $SR_{FDM}$ ($N_F/N_{MR}$) | 108/108 | 104/108 | 96/108 | 93/108 | 92% |
| $SR_{EBM}$($N_B/N_{MR}$) | 108/108 | 103/108 | 95/108 | 91/108 | 91% |

### C. Impression of robots

We administered a questionnaire to obtain the participants impressions in each viewing situations (after three interactions). The measurement was a simple rating on a Likert scale of 1– to-7. The questionnaires had the 2 items (shown in Fig. 7).

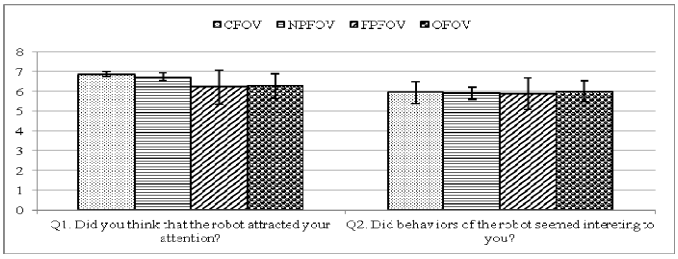



Figure 7: subjective evaluation results

Concerning Q1, repeated measures of analysis of variance (ANOVA) shows the significant differences between conditions ($F(3, 105)=6.4$, $p=0.0004$, $\eta^2=0.08$). However, there are no significant main effects were found between conditions in concerning Q2 ($F(3, 105)=0.16$, $p=0.91$, $\eta^2=0.004$). Results reveal that the robotic system is functioning to attract the target human's attention toward it for initiating an interaction episode between the robot and the target human.

## VI. CONCLUSION

To behave in a socially acceptable manner, the robots are capable of performing a lot of social and cognitive functions. In this paper, we showed that the robot is capable to perform an important social function such as initiating interaction with humans in different situations. This system can be extended for performing other social functions such as eye contact, attention control, and shared attention. Facial expression module, embodiment, other social cues (such as, waving, or touch) can be include for further improvements.

## REFERENCES


[1] K. Hayashi, D. Sakamoto, T. Kanda, M. Shiomi, S. Koizumi H. Ishiguro, T. Ogasawara, and N. Hagita, "Humanoid robots as a passive-social medium-a field experiment at a train station", in *Proc. ACM/IEEE Annual Conf. HRI*, 2007, pp. 137-144.

[2] N. Bergström, et al., "Modeling of natural human-robot encounters", in *Proc*. IEEE/RSJ *Int. Conf. IROS,* 2008, pp. 2623-2629.

[3] M. P. Michalowski, S. Sabanovic, and R. Simmons, "A spatial model of engagement for a social robot", in *Proc. IEEE Int. Works. AMC*, 2006, pp. 762-767.

[4] B. Mutlu, J. Forlizzi, J. Hodgins, "A storytelling robot: modeling and evaluation of human-like gaze behavior", in *Proc. IEEE-RAS Int. Con. Hum. Rob.*, 2006, pp. 518-523.

[5] Z. Yucel, A. A. Salah, C. Mericli, "Joint visual attention modeling for naturally interacting robotic agents", in *Proc. Int. Sym. Com. & Info. Sci.*, 2009, pp. 242-247.

[6] E. T. Hall, *The Hidden Dimension: Man's Use of Space in Public and Private*, The Bodley Head Ltd., London, 1966.

[7] E. Goffman, *Behavior in Public Place*, The Free Press, New York, 1963.

[8] A. Kendon, *Spatial Organization in Social Encounters: The F-Formation System,* Cambridge University Press, Massachusetts, 1990.

[9] K. Hayashi, M. Shiomi, T. Kanda, and N. Hagita, "Friendly patrolling: A model of natural encounters" in *Proc. Int. Conf. RSS*, 2011.

[10] C. Shi, M. Shimada, T. Kanda, H. Ishiguro, "Spatial formation model for initiating conversation, in Proc. Int. Con. RSS, 2011.

[11] F. Yamaoka, T. Kanda, H. Ishiguro, N. Hagita, "A model of proximity control for information presenting robot", *IEEE Trans. Robot.*, vol. 26, no. 1, pp. 187-195, 2010.

[12] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, N. Hagita, "How to approach humans? Strategies for social robots to initiate interaction", in *Proc. ACM/IEEE Int. Conf. HRI*, 2009, pp. 109-116.

[13] M. Buss et. al., "Towards proactive human-robot interaction in human environments", in *Proc. Int. Con. CogInfocom.*, 2011, pp. 1-6.

[14] M. M. Hoque, T. Onoki, Y. Kobayashi, and Y. Kuno, "Effect of robot's gaze behaviors for attracting and controlling human attention, Adv. Robo., vol. 27, no. 11, pp. 813-829, 2013.

[15] C. Breazeal, *Designing Sociable Robots*, Cambridge, MA: MIT Press, 2002.

[16] D. Mastsui, T. Minato, K. Macdorman, and H. Ishiguro, "Generating natural motion in an android by mapping human motion", in *Proc. Int. Conf. IROS,* 2005, pp. 3301-3308.

[17] T. Kanda et al., "Development and evaluation of an interactive robot Robovie", in *Proc. Int. Conf. ICRA*, 2002, pp. 1848-1855.

[18] D. Miyauchi, A. Nakamura, and Y. Kuno, "Bidirectional eye contact for human-robot communication", *IEICE Trans. of Info. & Syst.*, vol. 88-D, no. 11, pp. 2509-2516, 2005.

[19] M. M. Hoque et al., "Design an intelligent robotic head to interacting with humans", in Proc. Int. Conf. ICCIT, 2012, pp. 539-545.

[20] Y. Kobayashi, and Y. Kuno, "People tracking using integrated sensors for human robot interaction", in Proc. IEEE Int. Conf. IT., 2010, pp. 1597-1602.

[21] (2010) FaceAPI: Face Tracking for OEM Product Development. Seeing Machines Int., Available: http://www.faceapi.com.

[22] G. Bradsky, A. Kaehler, and V. Pisarevskky, Learning based computer vision with Intel's open computer vision library, *J. Intel Tech.*, vol. 9, no. 1, pp. 119-130, 2005.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Eye Gaze Behavior of Virtual Agent in Gaming Environment by Using Artificial Intelligence

Avijeet Das[1] and Md Mahmudul Hasan[2,*]
[1] Department of Engineering and Information Technology, University of Technology Sydney, Australia
[2] Department of Computer Science and Engineering, Daffodil International University, Bangladesh
[*] E-mail: mhasan@daffodilvarsity.edu.bd

*Abstract*—**In the gaming environment, building natural looking avatar is a challenging task. An avatar represents a human character which includes speech, action, and eye movement in a virtual world. The movement and the expression of eye modeling is cognitive function. In order to assess virtual agent's eye gaze behavior on the basis of cognitive and emotional characteristics, this paper presents an experimental simulation to identify perfection of the eye gaze behavior. For this reason, this paper splits eye behavior into few levels by using CADIA Populus. But the goal of this research work is to focus on gaze behavior. In addition, it briefly describes the cognitive activities based on different cases. Furthermore this paper introduces an integrated framework for modeling the interaction of agent with the immersive environment. Finally, the gaze strategy and some simulation results from previous studies have been discussed to provide a concrete idea about gaze behavioral science in gaming environment using artificial intelligence.**

***Keywords—Gaming environment, Eye gaze behaviour, Artificial intelligence, Virtual Agent, Non-player character***

## I. INTRODUCTION

The Virtual Agent community is closely associated to Virtual Assistants industry. It is a computer engendered, animated, artificial intelligent virtual character (virtual agent 2011). The agent takes an intelligent communication with users, performs suitable non-verbal behavior. In the gaming environment virtual agent have some capacity to improvise its actions. In gaming environment, the autonomous characters are called as non-player character *(NPC)*. The next sub sections will discuss about the different aspects of virtual agents. Furthermore, additional concentration will be put into the activity of eye movement: steering and gaze behavior. The research work briefly discuss about the virtual agent and different features of the virtual agent. Furthermore it introduces an integrated framework for modeling the interaction of agent with the immersive environment. The framework also discuss on the decision making behavior and autonomous motion control of the avatar.

Finally, this paper also express various gaze strategies based on the framework. Furthermore, analysis of the simulation from previous studies also addressed.

### A. Virtual Agent

Virtual agent or avatar in computer game is the graphical representation of human character. In the world of computer gaming it may represent as a three- dimensional form. There are a number of researches from the past for the improvement of avatar behavior and their physical appearance. According to Barbara Hays (1996), in the gaming environment we should concentrate three kinds of advanced skills exhibit by the virtual agent. First of all, agent should "exhibit life-like qualities". Secondly, agent are able to "follow the directions" which are taken from the external sources (Barbara 1996). Finally, agents are able to "improvise" many different aspects of behavior. Virtual agents are the economical and strongest link between human and service. According to Liu et al. (2010), a virtual agent is a computer generated 3D digital representation of human user. It sometimes refers as a computer program [1].

The term "virtual agent" consists of two parts: virtual and agent. The word virtual dates back to late 16th century as "inspired by physical capabilities". The Latin word virtus refer to excellence and efficiency. The word virtual first recorded in 1959 as: "capable of producing certain effect" (Virtual Agent 2011). The agent is a Latin word founded on back to 15th century which means "on who acts".

### B. Different aspects of virtual agent behavior

It is expected that the virtual agents exhibit more realistic features. The features include communicative and expressive characteristics which is similar to the natural human. The features consist of speech, facial expression, gestures, eye movement (e.g. eye gaze, steering). For controlling the virtual human behavior the scientists are focused on autonomous behavior control ability. For designing the different behaviors of virtual agent physiological and physical behaviors are combined. This is very hard to put all the behavior and perform a decision by the agent. For the simplicity, this work specially focused on eye movement behavior of the controversial human agent.

### C. Eye movement of virtual agent

In the realistic game environments, eye movement is a significant feature that virtual agent perform in a natural way. The eye movement is a cognitive function of an avatar [2]. Eye movement comprises steering behavior, gaze behavior,

idle gaze, and locomotion. The quality of an avatar performance reflects on its eye. For an example, in a situation where an agent is perform fight with another agent. The user expects some realistic eye control behavior on the agent which makes the game more efficient.

### *D. Eye movements in natural behavior*

In the natural behavior, eye movement is a cognitive function [3]. An experiment from "Yarbus" shows some positive findings in this field. They came up with three points about the understanding of the cognitive function of eye. The first is the demonstration of the pervasive role of the task in guiding where and when to fixate [4]. The second has been the recognition of the role of internal reward in guiding eye and body movements, revealed especially in neurophysiologic studies. The third important advance has been the theoretical developments in the fields of reinforcement learning and graphic simulation (Hayhoe & Ballard 2005, p. 188).

## II. BEHAVIOURAL MODEL

### *A. Gaze Behavior*

Gaze behavior is important in communication and social demonstration of a virtual agent. The action of an avatar represents by the movements of its eye. The structured action of gaze represents more realistic avatar. On the other hand, random or uncontrolled gaze behavior is both ambiguous and confusing (Badler, Chi & Chopra 1999, p.4) [4]. The agent has a large number of applications; such as in the interactive environment face to face communication is a vital role for agents. For such kind of situation gaze behavior plays an important role. Gaze model consists of many functions such as signaling, attention, adaptable turn-talking (Kipp & Gebhard 2008, p. 191). An avatar's personality and the current mode also reflect on its gaze behavior. A numerous number of scientists and researchers are working on the development of gaze behavior. They are trying to put the agent into a 3D social environment and investigate the social gaze of the agent [5].

### *B. Idle Gaze behavior*

Idle gaze is a very important feature of human behavior, which shows the human interpersonal skills. In the virtual environment, it is expectable that the agent behaves naturally. Idle gaze is similar to gaze but the difference is that the gaze is in idle position. For example, if an agent is walking down the street or waiting for the bus the gaze behavior should be in idle place. In the later part of this report will show some statistical result related with idle gaze behavior of virtual agent [6, 7].

### *C. Steering Behavior*

Steering behaviors consists of: seek, flee, pursuit, evasion, offset pursuit, arrival, obstacle avoidance, wander, path following, wall following, containment, flow field following, unaligned collision avoidance, separation, cohesion, alignment, flocking, and leader following (Reynolds 1999). In the steering behavior a vector is incorporated which is named as "steering vector". Consider an example agent is running behind a subject avoiding so many obstacles and the subject is fixed. So the agent is trying to run behind the subject and keep its eye on the subject. From figure 1; the agent flee as desired velocity for the target and the agent use seek steering. The gray line denoted as the steering vector. The length of desired velocity could be max_speed to the target which depends on the specific situation. From Reynolds (1999) the equation for steering from figure 1 is;

```
desired_velocity = normalize (position -
           target) * max_speed
  steering = desired_velocity- velocity
```

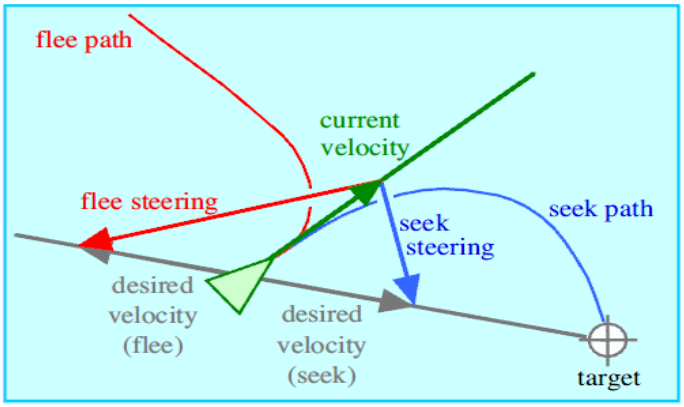



Figure 1. Steering behavior: seek and flee (Reynolds 1999)

Seek can be described as to steer the agent towards a specific location on the map. This behavior modifies the character ability which changes its velocity in the direction of the subject. This is not related to attractive force such as gravity which provides an orbital path around the subject (Reynolds 1999, p. 8).

Flee is opposite of seek which can be described as a steering character. The speed is "radially aligned" away from the target and the desired velocity works in the opposite direction [8].

Pursuit and seek is nearly similar but the main difference is: in seek behavior the subject is constant but in the pursuit behavior the subject is a moving subject. In the pursuit behavior an estimation of the next position is needed. The methodology is: first select a reference point and reexamine each point after a certain period of time. Another problem with the pursuit is if the agent and the subject move to each other from the opposite direction. It appears in a constant heading on the global space for the subjects.

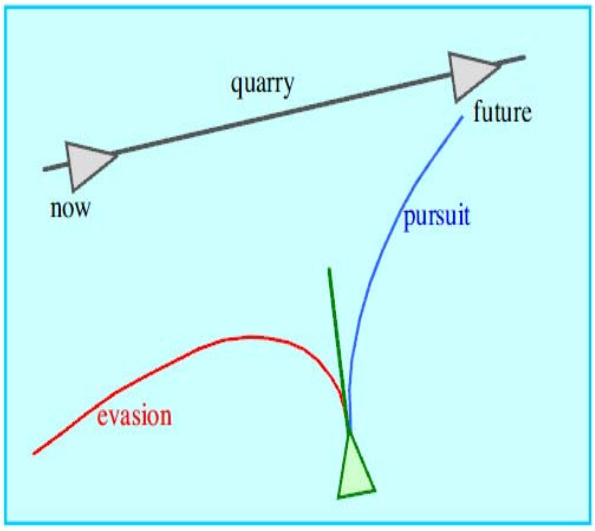



Figure 2. Steering Behavior: pursuit and evasion (Reynolds 1999)

Evasion is similar to pursuit, but the main difference is in fleeing. In the evasion the flee is used to steer away from the expected location of the subject. For controlling the movement optimal technique is applied on pursuit and evasion. But the natural system, the movement is unpredictable and non-optimal [9, 10].

In the arrival behavior, the position of the agent is far from the position of the target. But instead of the moving towards the subject at the full speed, the behavior enforces the character to slow down as it approaches the target which is shown in figure 3. Finally, slowing to a stop equivalent with the subject (Reynolds 1999, p.10).

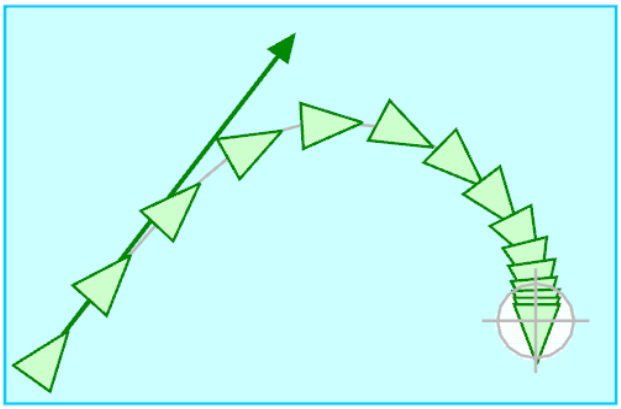


Figure 3. Steering behavior: arrival (Reynolds 1999)

The equation for the arrival behavior is:

```
target_offset = target - position
distance = length (target_offset)
ramped_speed = max_speed * (distance /
            slowing_distance)
clipped_speed = minimum (ramped_speed,
                max_speed)
desired_velocity = (clipped_speed /
      distance) * target_offset
steering = desired_velocity - velocity
```

Obstacle avoidance behavior is a complex behavior, which is comprises seek and flee. This behavior provides the ability to movement in a tangled map by dodging around obstacles (Reynolds 1999, p.11).

Path following behavior allows an agent to follow a preset path like terrain, roadway or path on the space. This is widely used in those virtual environments where paths are predefined. For an example, in the car race the paths are preset and the cars are moving along the track which is shown in figure 4.

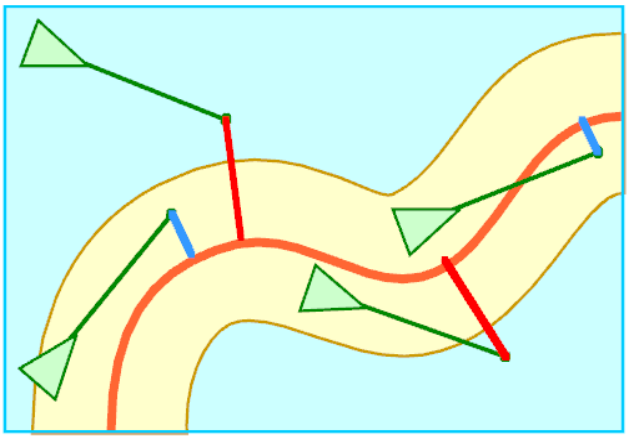


Figure 4. Steering behavior: Path following (Reynolds 1999)

Path following sometimes characterized by flow following and wall following. The goal of the path following is to move a character alongside the path while staying within the specific radius and spine (Reynolds 1999, p .13).

*D. Locomotion*

Locomotion is a kind of activity which collects the information from steering behavior and converts into motion of the character's body [11, 12]. This motion is subject to constraints imposed by the body's physically-based model, such as the interaction of momentum and limitation of forces applied by the body (Reynolds 1999). There are different types of locomotion: an avatar could have a locomotion characterized by physically-based dynamically stabled movement which gives us both genuine animation and behavioral locomotion. On the other hand an agent could have a simple locomotion model which is attached by pre-animated model. An adaptive locomotion is the avatar shall have the ability to learn from the environment. The best approach is to represent hybrid model approach which is to use simple model and adaptive model both. In the locomotion behavior the steering behavior remains same but the movement of the body is shifted. For example, someone wants to ride on the bicycle. So he walks to his cycle, ride onto it and start cycling. In the whole process the agent might steer on its cycle, but the movement to the cycle and riding on it in a physical balanced manner. At the end locomotion can be controlled to the motion with the mixer of some predefined movement like walking, running and so on [13].

## III. FRAMEWORK FOR MODELING EYE BEHAVIOR OF AVATAR

Liu et al. (2010) proposed a framework for virtual agent behavior modeling [3]. The model helps the user to interact with the virtual environment through the agent. The framework divided into different layers. There are basically three layers, which are: i) Information perception, ii) Behavior decision making, iii) Autonomous motion control

### *A. Information Perception*

This part of the framework handles all the cognitive behaviors which are shown in figure 5. Perception and cognition comprises four phases: synthetic vision, audition simulation, intention and attention, and memory. Synthetic vision helps an agent to collect the information in its eye sight. In the complex scenario, the visibility can be accomplished by testing against the virtual human's view frustum (Liu et al. 2010, p. 2) [3]. Audition simulation can be achieved by the cross-examination of system messages. A vocal message passes to the agent and agent check in the system. After perform the examination on system message agent check the scalar distance between the position from the origin of the message to its own position. After acquiring vision information and audition information, the agent performs action to limit its focus to confine the object. Using perceptual attention, the scale is balanced between the perception processing output and decision making behavior. Intention helps the agent to pick the object it focused on. Memory model helps the agent to remember the perceived objects through sensitivity [14].

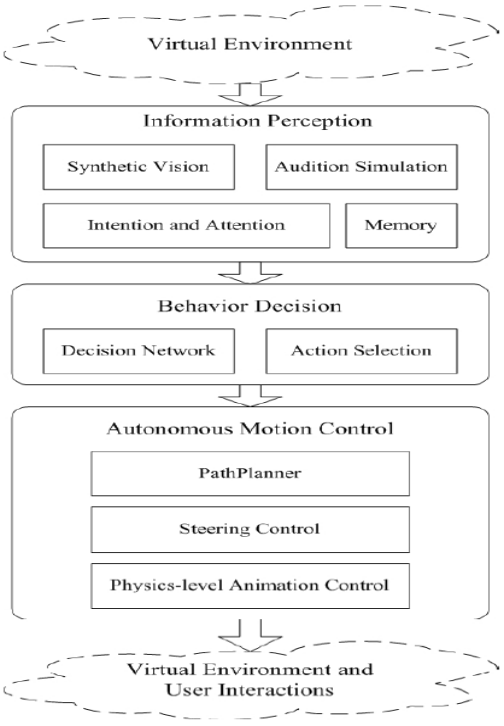



Figure 5. Framework for modeling avatar eye behavior

### *B. Behavior decision making*

The decision making behavior consists of two modules, which are: i) decision network and ii) action selection. Decision network is used to perform agent decision making process. It solves the problem of complexity and uncertainty (Liu et al. 2010, p. 2) [2]. Decision making networks works in a hierarchical manner: the top level response network, the acquaintance behavior network, the attack-response behavior network and energy-restore behavior network. The top level behavior network applied to choose the remaining three networks. After that the chosen network is used to handle the analogous behavior network. The decision network decides to choose the corresponding action for that particular scene.

### *C. Autonomous motion control*

Movement animation control has three phases: path planner, steering control and physics level control. Path planner is responsible for defining where the agent planned to move and engendered a target. For path finding the A* search algorithm is used. The algorithm provides to find optimal path from the source to endpoint. But the A* algorithm cannot deal with the dynamic object in gaming environment. So that, steering control behavior is implemented. Steering control behavior helps the agent to deal with the dynamic objects. A "sidestep repulsion vector" is used to determine the steering path. The steering control is fastest solution to avoid obstacle but the physical action is performed by the physical control behavior. Here in the physical level controlling collision detection mechanism is used. This method help to avoid stuck on any obstacle [15].

### *D. Different Gaze Strategies*

According to Kipp & Gebhard (2008), the three gaze strategies are used for the framework discussed earlier. Three gaze strategies are: i) The Mona Lisa Strategy, ii) Dominant Strategy (Dom), and iii) Submissive Strategy (Sub). The diagrams for three strategies were modeled using "timed finite state automata".

**The Mona Lisa Strategy(ML+, ML-)** comprises of continuous following of user's position at all times which makes 3D effect on agent eyes. When the 3D effect switched on that means ML+, agent looks at the position of virtual camera. When the effect is switched off (ML-) that means avatar looks at the virtual gaming environment which is basically not the camera (Kipp & Gebhard 2008, p. 193).The Mona Lisa is widely used for request for attention, stare, look straight into other agent eyes, shows the anger of the avatar even cold anger.

**Dominant Strategy (Dom):** Dominant strategy involves the eye contact when avatar communicates with another avatar. The strategy consists of speaking and listening with another avatar. It was observed that, the advanced status agent gazes spend less time than the lower status agent. The activities are related to the amount of time for listening and speaking and calculated by Visual Dominance ratio (VDR). According to Kipp & Gebhard (2008, p.138), dominant strategy comprises eye contact while conversation between two agent and arbitrarily changes of gaze behavior while listening. The agent usually holds eye contact while speaking and listening.

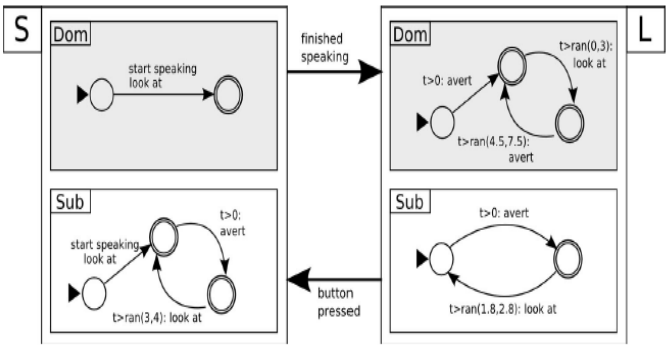



Figure 6. A simulation used to model the gaze behaviors of dominant (upper states) and submissive (lower). S/L refer to speaking/listening modes (Kipp & Gebhard 2008, p. 194)

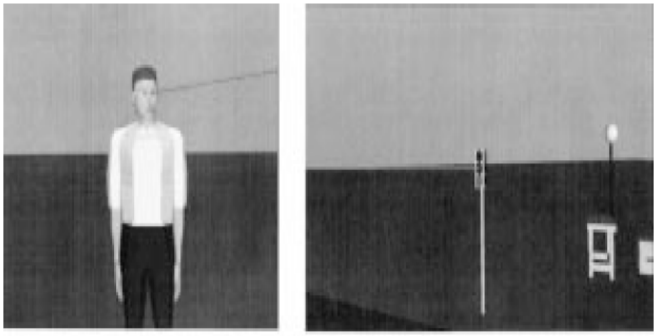
**Submissive Strategy (Sub):** This strategy is related to dominant strategy. If the agent breaks the eye contact and looks back, that agent is lower status. This strategy helps the agent look briefly every from time to time and changes the gaze. The agent maintains while talking but avoids its gaze instantaneously (Kipp & Gebhard 2008, p. 194). The gaze remains avert for 3-4 sec. After that the agent starts eye contact again and looks back instantaneously. In terms of listening, the only difference is time which is 1.8-2.8 sec.

## IV. Simulation of Eye Gaze Behavior

The following simulation is based on gaze behavior, steering and locomotion. Two researchers named S. Chopra and N.I. Badler from "University of Pennsylvania" accomplished the simulation [4]. The result was very impressive than the previous simulation on eye gaze based on cognitive modeling. The problem was to generate a visual attending behavior in controversial virtual humanoid. The behavior comprises some cognitive behavior such as eye control and head motion with the other actions like gaze, locomotion. Their goal was, when an agent moves toward the target, or looks for someone in the social situation; the behavior should appear as actual eye behavior (Chopra & Badler 2001, p.2) [4].

For the simulation the taken scenario was, an agent asked to walk to one point from another. So that to achieve the goal: he has to cross the road, look for oncoming traffic and wait for the suitable traffic signal. A social scenario was animated to investigate the simulation.

A virtual human model was taken to investigate all the tasks. A walking eye behavior produced adds appropriate sites to IntentionList. The IntentionList consists of the destination and the ground in front of agent's feet. The corresponding eye behavior remains active while walking along the road. The simulation showed, when the agent walks along the street more items added into IntentionList. The agent uses its monitoring eye gaze behavior. The line in the second part of the figure 7 indicating agent's LoS (Line of sight).



(a)

Figure 7. Agent look for the traffic using monitoring eye gaze (Chopra & Badler 2001) [4]

A monitoring eye behavior also produced to check oncoming traffic on the road. This behavior remains until the agent crosses the road.

## V. Analyzing the Model of Idle Gaze Behavior

Idle gaze is an important aspect of human behavior. In the gaming environment it is necessary to look avatar more realistic. The simulation was made on CADIA Populus. For analyzing the model, two situations were chosen: i) waiting for the bus and ii) walking down the street. The simulation was done on social situation. It is essential to know how an agent reacts in social gaming environment. The avatar was modeled using CADIA Populus.

Three researchers Cafaro1, Gaito1&Vilhjalmsson (2009) took the video for analyzing the model. The video studies were performed in different situations [8]. The two cases are:

Case 1: The waiting behavior was observed and the camera was placed 30m away from the bus station. The people were waiting for the bus and the eye behavior of that time was taken.

Case 2: The walking behavior was observed, and the camera was placed in a shopping center. While the people were walking they change their gaze depending on the gender.

After analyzing these two cases, some results can be produced. From the case 1 we found:

i. The observed person fabricates a glance and maintains his glance for a short period of time. The shorter glances were used to look at the other person near him.
ii. The longer gazes used to look at the target away from him. Such as oncoming bus.

And from the case 2, we found:

i) The person looks at the ground frequently while walking.
ii) The person closes his eyelids just before moving his head.
iii) Most of the time the person never looks up, up-left or up-right.

Now while observing the case 1 we can produces a table. The table shows the time relationship gaze targets and proxemics where the subject is waiting for a bus. All the time durations are in seconds.

TABLE I. THE RELATIONSHIP BETWEEN GAZE TARGETS AND PROXEMICS FROM CASE 1

| | Objects | | | Persons | | |
|---|---|---|---|---|---|---|
| Proxemics Area | Time on Target | % | Avg. Dur. | Time on Target | % | Avg. Dur. |
| Intimate | 84.11 out of 84.11 | 100% | 6 | 0 out of 6.43 | 0% | - |
| Personal | 10.26 out of 10.26 | 100% | 5.13 | 7.3 out of 61.16 | 12% | 2.43 |
| Social | 14.36 out of 22.56 | 64% | 4.80 | 45.17 out of 96.68 | 47% | 22.60 |
| Public | 1.33 out of 1.33 | 100% | 1.33 | 6.90 out of 29.46 | 23% | 2.30 |
| Extra | 5.90 out of 8.66 | 68% | 5.90 | 9.00 out of 35.56 | 25% | 3.00 |

The following table shows the observation where person look while walking down the shopping street. How their gaze reacts with the same gander and the opposite gender.

After putting all together in the CADIA Populus, there was an improvement for the avatar was observed. The idle gaze required lower update in the case of waiting for the bus, but the gaze required higher update in the case of walking down the street. The goal was to produce autonomous agent for the social environment. The social situation needs steering behavior of the agent to make it more realistic. In the public space the blind cone for the agent is 90 to 150 degree.

TABLE II. OBSERVATION OF CASE 2 WHERE PEOPLE LOOK WHILE WALKING DOWN IN THE SHOPPING STREET

| Targets | | | Durations | | |
|---|---|---|---|---|---|
| Target Type | % | Time on Target | Avg. | Min. | Max. |
| Same Gender | 3% | 7.75 | 0.64 | 0.24 | 1.24 |
| Opposite Gender | 5% | 12.92 | 0.72 | 0.29 | 1.15 |
| Shops | 13% | 33.58 | 0.74 | 0.14 | 1.34 |
| Cars | 9% | 23.25 | 0.93 | 0.1 | 1.8 |
| Ground | 25% | 64.58 | 1.5 | 0.5 | 2.5 |
| Camera | 7% | 18.08 | - | - | - |
| Other Side | 16% | 41.33 | - | - | - |
| Unknown | 22% | 56.83 | - | - | - |
| Total | 100% | 258.32 | | | |

Visual perception is divided into central view and peripheral view. The main focus is to make the gaze behavior more dynamic. To make it more dynamic the avatar has the intention to change the gaze for the surrounding area. So the gaze shifts during to choose a target.

## VI. CONCLUSION

This paper provides an overview of a virtual agent in semi-immersive environment. The simulating model shows the implementation of the gaze strategies using timed automata. Gaze is a dominant interaction modality with many functions like signaling attention, regulating turn-taking or deictic reference (Kipp & Gebhard 2008, p. 191).

From the above studies, it proves that steering behavior and gaze is important aspects of avatar. Improvement of cognitive behavior makes avatar more efficient and realistic. Moreover, the simulation from different cases shows the results to make the avatar more dynamic. To construct realistic reactive systems, we have to develop tools and systems to model where the human agent can freely reactive with the environment in future.

## ACKNOWLEDGMENT

Authors would like to give thanks Prof. Dr Syed Akhter Hossain and Center for Innovation and Technologies (CIT), DIU for their support to conduct this research work.

## REFERENCES

[1] Edward, L., Lourdeaux, D. & Barthes, J. 2009, 'Cognitive Modeling of Virtual Autonomous Intelligent Agents Integrating Human Factors', *Web Intelligence and Intelligent Agent Technologies, 2009. WI-IAT '09. IEEE/WIC/ACM International Joint Conferences on*, vol. 3, pp. 353-6.

[2] Del Bimbo, A. & Vicario, E. 1995, 'Specification by example of virtual agents behavior', *Visualization and Computer Graphics, IEEE Transactions on,* vol. 1, no. 4, pp. 350-60.

[3] Litao Han, Qiaoli Kong, Bing Liu &Zhiqiang Li 2011, 'Building cognitive model of intelligent agent in Virtual Geographical Environment', *Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference on*, vol. 3, pp. 660-4.

[4] Badler, N.I., Chi, D.M. & Chopra, S. 1999, 'Virtual human animation based on movement observation and cognitive behavior models', *Computer Animation, 1999. Proceedings*, pp. 128-37.

[5] Vinayagamoorthy, V., Garau, M., Steed, A. & Slater, M. 2004, 'An Eye Gaze Model for Dyadic Interaction in an Immersive Virtual Environment: Practice and Experience', *Computer Graphics Forum,* vol. 23, no. 1, pp. 1-11.

[6] Lun, E. V. 2011, *Virtual agent,* viewed 4 November 2011, <http://www.chatbots.org/virtual_agent/>

[7] Harber, M., Bosch, K. V. D. 2009, *A Study into Preferred Explanations of Virtual Agent Behavior,* viewed on 4 November 4 2011, <http://people.cs.uu.nl/maaike/Iva_2009.pdf>

[8] Cafaro, A., Gaito, R. &Vilhjálmsson, H., H., 2009, 'Animating Idle Gaze in Public Places',*IVA '09 Proceedings of the 9th International Conference on Intelligent Virtual Agents*, Springer-Verlag Berlin, Heidelberg, p. 250-256.

[9] Garau, M., Slater, M., Vinayagamoorthy, V.,Brogni, A., Steed, A., & Sasse, M., A., 2003,'The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment',*In Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, New York, NY, USA, p. 529-536.

[10] Hayhoe, M. & Ballard, D. 2005, 'Eye movements in natural behavior', *Trends in cognitive sciences,* vol. 9, no. 4, pp. 188-94.

[11] Kipp, M., &Gebhard, P., 2008, '*Intelligent Virtual Agents'*, 1stedn, Springer, Berlin.

[12] Md. Mahmudul Hasan & Most. Tajmary Mahfuz (2013), "Teletraffic Model for K-fold Multicast Network: Performance Evaluation with Buffer using M/M/n/n+q/N Kendal's Notation", *The SIJ Transactions on Computer Networks & Communication Engineering (CNCE)*, The Standard International Journals (The SIJ), Vol. 1, No. 3, Pp. 67-71.

[13] Pelachaud, C., Bilvi, M. 2003, *Modelling Gaze Behavior for ConversationalAgents,* viewed on November 4 2011, <http://perso.telecom-paristech.fr/~pelachau/allpapers/iva03.pdf>.

[14] Grillon, H. &Thalmann, D. 2008, 'Eye contact as trigger for modification of virtual character behavior', *Virtual Rehabilitation, 2008*, pp. 205-11 .

[15] Steptoe, W., Oyekoya, O., Murgia, A., Wolff, R., Rae, J., Guimaraes, E., Roberts, D. & Steed, A. 2009, 'Eye Tracking for Avatar Eye Gaze Control During Object-Focused Multiparty Interaction in Immersive Collaborative Virtual Environments', *Virtual Reality Conference, 2009. VR 2009. IEEE,* pp. 83-90 .

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Highly Constrained University Course Scheduling using Modified Hybrid Particle Swarm Optimization

Tania Ferdoushi, Prodip Kumer Das, and M. A. H. Akhand*
Khulna University of Engineering & Technology (KUET), Bangladesh
E-mail: akhand@cse.kuet.ac.bd

*Abstract*—**At universities, course scheduling problem is an NP-hard dilemma concerned with instructor assignments and class scheduling under various constraints and restricted resources. A number of novel meta-heuristic algorithms based on the principles of particle swarm optimization, genetic algorithm, simulated annealing, tabu search, etc. were proposed and demonstrated to solve University Course Scheduling Problem (UCSP). Among several Particle Swarm Optimization (PSO) based methods, Hybrid PSO (HPSO) is the most recent one and is shown to achieve optimal solution. In those exiting algorithms including HPSO, some important features regarding co-class concept, format of classes, lab scheduling, etc. are either ignored or adopted partially. In this study the modified HPSO (MHPSO) is investigated for highly constrained environment introducing some important features with HPSO. MHPSO takes into account some significant hard and soft constraints that make it unique to solve complex UCSP with higher coverage. The proposed MHPSO has been tested on a real-world highly constrained environment and found to produce feasible solution.**

***Keywords—UCSP, complex preferences, local search, repair process, constraints*.**

## I. INTRODUCTION

Scheduling problems evolve optimally assigning limited resources to tasks over time. Educational scheduling focuses on course scheduling and exam scheduling of which course scheduling is much more complex to solve because achieving optimality of it is a great challenge. Almost every university has to solve the general problem of University Course Scheduling Problem (UCSP) [1].

Course Scheduling problems are to assign days and hours to courses, so that the constraints imposed by the students as well as instructors are satisfied. UCSP is a real life optimization problem with that type of constraints satisfaction [2], [3]. For UCSP, the constraints are different for different institutions. Whatsoever, it is always a difficult task due to limited resources as well as different class pattern such as single theory course conduction by two instructors; necessity of several conjugative time slots for sessionals or lab classes. In UCSP, each instructor should teach one course section at a time and 'each course section should be taught by an instructor'- and other likelihood constraints that can't be violated are termed as hard constraints. Soft constraints evolve instructors' and students' preferences about days and time periods and leisure periods, etc; these are less important than hard constraints and can be disobeyed. Tasks of UCSP also include lectures allocation into suitable rooms considering room facility and enrolled students for the lectures [4], [5].

Numerous approaches have been investigated to solve UCSP from early 90's to till date. Among the methods, Ant Colony Optimization (ACO) [5], Genetic Algorithm (GA) [6], Tabu Search (TS) [7] and Simulated Annealing (SA) [8] are well known with their numerous variations. Recently, Particle Swarm Optimization (PSO), a population based meta-heuristic method, is explored to solve UCSP that is in which particles spread in space and a particle's position indicates a solution. Each particle moves to a new position heading for global optimum based on the global experience of the population and individual experience of its own. Among several PSO based methods, Hybrid PSO (HPSO) [9] is the most recent one and is shown to achieve optimal solution [10]-[13]. The HPSO algorithm, having several important features, seems to construct more feasible timetables compared to other techniques. However, the environment in which HPSO tested seems to have fewer constraints and might not perform well when resources (e.g., number of teacher) are limited and/or the environment is highly constrained.

The objective of the paper is to modify and/or add features to HPSO to solve highly constrained UCSP. As a highly constrained realistic environment, Computer Science and Engineering department of Khulna University of Engineering & Technology (KUET) is considered where course scheduling is always a difficult task due to limited resources. In KUET, a theory course is conducted by two instructors; sessionals/labs are required several conjugative time slots- these and some other features are considered in this paper. Moreover, this study also simulates the minimum resource requirement through resource scheduling and required additional resources to increase student intake.

Rest of the paper is organized as follows: Section II describes the HPSO with proposed modification to solve complex UCSP. Section III discusses experimental setup and results with the performance of the proposed approach. Section IV presents concluding remarks with possible future studies directed from this study.

## II. PROPOSED METHODOLOGY

UCSP is a task of assigning the events of universities (lectures, activities, etc) to the various resources such as instructors, classrooms and time slots. Several approaches (e.g., GA, SA, TS, ACO, PSO, HPSO) have been adopted to solve UCSP; among them HPSO is proved more worthy. But for highly constraints environment HPSO is to be modified to

cope up more realistic atmosphere like KUET. Considering some crucial hard and soft constraints with additional features Modified HPSO (MHPSO) are investigated in this study. Features like lab and room scheduling are included in MHPSO.

A particle in MHPSO changes its velocity ( $V_{id}^{new}$ ), therefore position($X_{id}^{new}$), looking its own best position ($L_{id}$) and swarm's best position ($G_d$) in every iteration. In Eq. (1), $w$ is the inertia and $c_1$ & $c_2$ are the learning factors.

$$V_{id}^{new} = wV_{id} + c_1 r_1 (L_{id} - X_{id}) + c_2 r_2 (G_d - X_{id}) \;\; r_1, r_2 \epsilon\, [0,1] \quad (1)$$

$$X_{id}^{new} = X_{id} + V_{id}^{new} \quad (2)$$

The position of a particle represents a complete scheduling of instructors, class rooms and laboratories when MHPSO works on UCSP. A solution for instructors is the class assignment in the time-slots represented by a matrix of Table I where a time slot in a week for a particular instructor is uniquely identified. According to the table, each day contains nine teaching periods; five of such working days in a week provide continuous 45 time-slots for each instructor; time-slots 1–45 for first instructor, 46–90 for second instructor, and so on. MHPSO allows instructors to give their preferences i.e., their preferred days and time periods to conduct classes, maximum number of teaching-free time periods and lecturing format (consecutive time periods or separated into different time periods). Table II presents a typical arrangement of instructors' preferences where a larger positive value (e.g., 5) indicates an instructor's high preference to take class in the corresponding slot. On the other hand, a negative value represents disfavor of a slot.

TABLE I. REPRESENTATION OF WEEKLY TIME SLOTS FOR COURSE SCHEDULING OF M INSTRUCTORS.

| | | sun | mon | tue | wed | thu | . | sun | mon | tue | wed | thu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PERIOD | 1 | 1 | 10 | 19 | 28 | 37 | . | 45(m-1)+1 | | … | … | … |
| | 2 | 2 | 11 | 20 | 29 | 38 | . | 45(m-1)+2 | … | … | … | … |
| | 3 | 3 | 12 | 21 | 30 | 39 | . | 45(m-1)+3 | … | … | … | … |
| | 4 | 4 | 13 | 22 | 31 | 40 | . | 45(m-1)+4 | … | … | … | … |
| | 5 | 5 | 14 | 23 | 32 | 41 | . | … | … | … | … | … |
| | 6 | 6 | 15 | 24 | 33 | 42 | . | … | … | … | … | … |
| | 7 | 7 | 16 | 25 | 34 | 43 | . | … | … | … | … | 45m -1 |
| | 8 | 8 | 17 | 26 | 35 | 44 | . | … | … | … | … | 45m |
| | 9 | 9 | 18 | 27 | 36 | 45 | | | | | | |
| < Instructor 1 > | | | | | | | | < Instructor m > | | | | |

MHPSO tries to assign courses (populated in Table I) maximizing instructor's preferences that satisfy all hard constrains with total soft constrains violations. The fitness of a solution (particle) is measured in MHPSO considering a total priority value (TPV) of a solution with considering soft constraints satisfaction matter. Since soft constraint satisfaction is an element of a quality solution of a UCSP, this study considered it in the fitness calculation and total soft-constraints violated (*TSCV*) is discarded from TPV to get fitness of solution.

TABLE II. PRIORITY VALUES OF PREFERENCES FOR INSTRUCTORS.

| | | sun | mon | tue | wed | thu | . | sun | mon | tue | wed | thu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PERIOD | 1 | 0 | 0 | 0 | -1 | 5 | . | 0 | 5 | 4 | -1 | 5 |
| | 2 | 1 | 3 | 4 | -1 | 5 | . | 2 | 5 | 4 | 2 | 5 |
| | 3 | 1 | 3 | 4 | -1 | 5 | . | 2 | 5 | 4 | 2 | 5 |
| | 4 | 1 | 3 | 4 | -1 | 5 | . | 2 | 5 | 4 | 2 | 5 |
| | 5 | 3 | 5 | -1 | 3 | 4 | . | 1 | 3 | -1 | 1 | -1 |
| | 6 | 3 | 5 | -1 | 3 | 4 | . | 1 | 3 | -1 | 1 | -3 |
| | 7 | 3 | 5 | -1 | 3 | -1 | . | 1 | 2 | -1 | 1 | -5 |
| | 8 | 0 | 0 | -1 | 3 | -1 | . | 0 | 2 | -1 | 1 | -1 |
| | 9 | 0 | 0 | -1 | 3 | -1 | . | 0 | 2 | -1 | 1 | -1 |
| < Instructor 1 > | | | | | | | | < Instructor m > | | | | |

$$fitness[i] = TPV - \gamma \sum_{t=1}^{T} TSCV(t) \gamma\, \epsilon [0,1]$$

$$= \alpha \left[ \sum_{t \epsilon T} \sum_{l \epsilon L_t} \mathrm{PV_{lt}} \right] + \beta \left[ \sum_{k \epsilon C} \sum_{h \epsilon H_k} \mathrm{PV_{hk}} \right] - \gamma \sum_{t=1}^{T} TSCV(t) \quad (3)$$

$PV_{lt}$ = sum of priority values for those of time periods in teaching course section $l$ for Instructor $t$ ($t \epsilon T$)
$PV_{hk}$= sum of priority values for those of time periods in offering course section $h$ for class k ($k \epsilon C$)
$\alpha$, $\beta$ = weighting factors to specify relative priorities to be given a degree of satisfaction for instructors and classes; specified as $\alpha + \beta = 1$.

In MHPSO velocity is updated according to equation (1) as well as particle moves to a new position from its current position according to equation(2). Since values in particle are timeslot numbers, so we always uses the integer position values for timeslot numbers. The value of each component in V can be clamped to the range [$-V_{min}$, $+V_{max}$] to control excessive roaming of particles outside the search area. we set the $V_{max}$ values of 3, which limit to move forward or backward at most 3 time-slots for each course section.

A local search mechanism is incorporated into MHPSO for modifying the particles position to get better solution by the following equation

$$x_{id}^{new} = x_{id} + v_{id}^{new} + Rand(-L_{max}, +L_{max}) \;\ldots\ldots\ldots\ldots (4)$$

where *Rand* denotes a random integer number between $-L_{max}$ and + $L_{max}$. After a new position is obtained according to Eq. (2), the process then checks the neighborhoods of the new position whether there is a better position (generated by executing $Rand(-L_{max}, +L_{max})$ with a higher PV than the new position. If there is such position existed, then the new position is updated based on Eq. (4), otherwise, the new position is kept unchanged.

**Repair process:**

If there is a conflict between two instructors in the class or lab assignment, then a random value is generated using Eq. (4). And, the new position is then used to check the conflict between instructors. If the conflict exists, the random instructor is selected for a random value which minimizes the conflict rate.

**Efficient room scheduling:**

Efficient room scheduling can be done by MHPSO. If no. of rooms is reduced, existing class scheduling can run smoothly. In CSE Department of KUET Maximum 4 class rooms are needed to do efficient class scheduling. After solving Class Scheduling using MHPSO, 3 rooms are enough to do that efficient class scheduling.

**Lab Scheduling:**

Labs are scheduled based on the room and lab preferences of the rooms and instructors. The Labs are fixed for every subject.

**Constraints that are considered in MHPSO are:**

***Hard Constraints:***

1. Each instructor can only teach one course section at a time.
2. Each course section can only be taught by an instructor.
3. Students can only attend one course section at a time.
4. Each classroom can't be used more than one course section at a time period.
5. Certain time periods can't be scheduled for non-academic activities, such as lunch time and sport.
6. Some course sections have to be scheduled in a particular room, such as computer lab.
7. The subjects must match the predefined lab subjects.
8. Every sessional class must continue for 3 class periods.
9. The class of one subject with 3 credits must have a common class time conflict.
10. Two instructors can't take class at the same time.
11. Lab class must be taken in $1^{st}$ period, $4^{th}$ period and $7^{th}$ period in any schedule date.

***Soft Constraints:***

1. Instructors and students (classes) can indicate their preferences along with their preferred days and time periods in 45 time-slots across five days and nine periods.
2. Instructors can choose to maximize the number of teaching free time periods; i.e., instructors can specify time periods when they prefer not to lecture.
3. A course section within three teaching hours has to be scheduled at the consecutive time periods and cannot be separated into two days of different time periods if required; i.e., instructors can specify their preferred lecturing format for a specific course section (for example, a course section requires three teaching periods, the formats of 3-0, 2-1 and 1-1-1 represent three consecutive periods, two consecutive periods with a further single period and three single periods, respectively).
4. Minimizing students' movement between rooms.
5. Two or one instructors for one subject.
6. The no. of students of every class must be permitted considering the lab capacity

Figure 1 describes the working procedure of the proposed algorithm. The velocity is updated using a new equation to fit every highly constrained environment in our study.

The teachers can provide their class preferences and based on those preferences, the classes are assigned to each teacher and respective class and lab scheduling is generated. The preferences for the instructors are shown in the Table II. Both the class room and lab room preferences are taken in the same format as instructors.

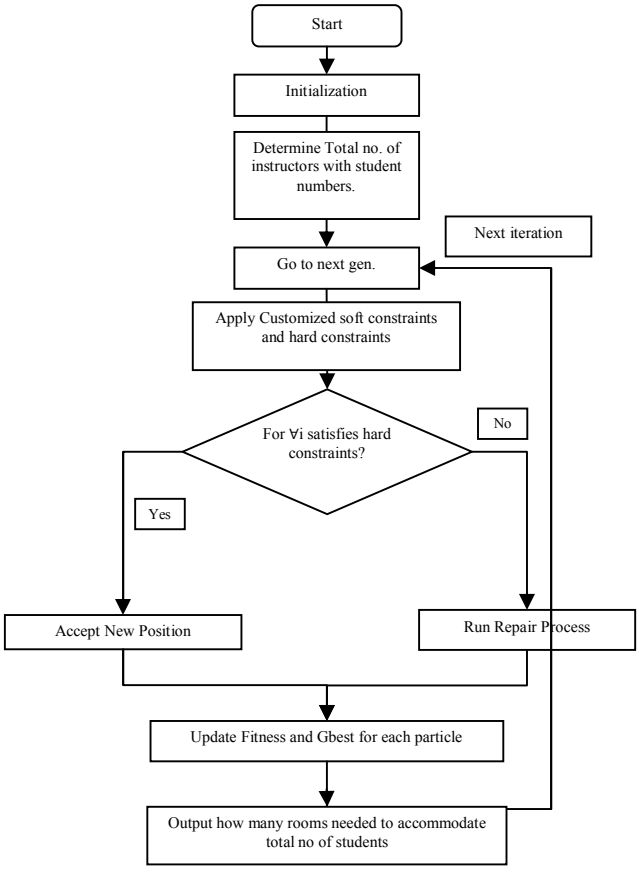



Figure 1. Flow chart of MHPSO.

## III. EXPERIMENTAL STUDIES

This section includes the experimental setup and experimental result analysis.

### *A. Experimental Setup*

The Experimental setup includes the teachers, students, room, lab capacity of Computer Science and Engineering

department of KUET. At present, there are 12 teachers, 60 students in each year, 4 classrooms, one seminar room and 3 labs. Our main objective is to study on these parameters to increase the utilization of resources from all respect of views.

### B. *Experimental Result*

Figure 2 explains the experimental result of the global best, average and generation best of 200 iterations. Size of population was 100 for the experiment. As the figure explains, the generation best tries to populate the global best value.

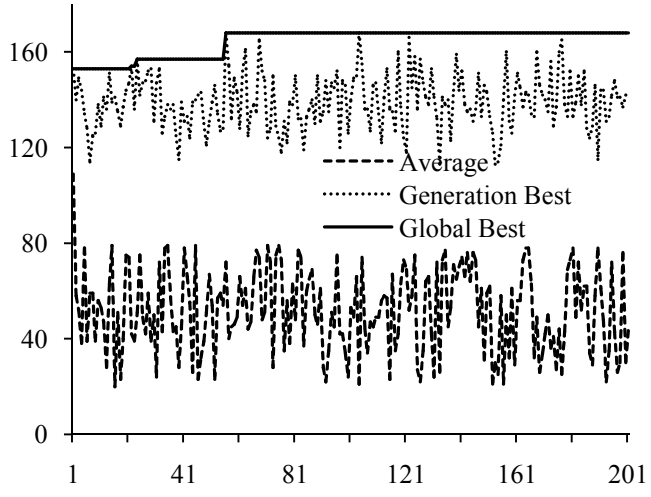



Figure 2. Experimental result analysis of 200 iterations.

### C. *Comparison with GA*

Figure 4 explains the result with MHPSO for 200 iterations for 100 populations. Here we can see that MHPSO grabs the best global best value in a few iteration than Genetic Algorithm (GA).

Parameters that are considered in the comparison are the instructors, room numbers, numbers of students and numbers of classes per day. Same input values are utilized in both MHPSO and the GA.

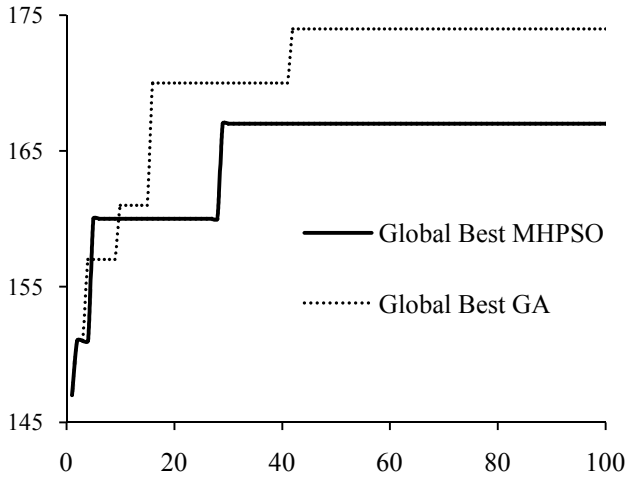



Figure 4. Comparison with GA with 100 populations.

| Instructors | Sunday | | | | | | | | | Monday | | | | | | | | | Tuesday | | | | | | | | | Wednesday | | | | | | | | | Thursday | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| $I_1$ | | | CSE 1107 | | | CSE 3107 | CSE 1101 | | | CSE 4123 | | CSE 3107 | | | CSE 3119 | | | | CSE 1107 | | | CSE 2104 | | | | CSE 2103 | | | CSE 4109 | | | | | | | | | | CSE 1101 | | CSE 2103 | | | |
| $I_2$ | CSE 2103 | | CSE 2107 | | | | | | | | | | CSE 1107 | | | | | | | | | CSE 3101 | CSE 3109 | | | | | CSE 2107 | | | | CSE 3101 | | CSE 4120 | | | | | | | | | | |
| $I_3$ | | CSE 1101 | | | | | CSE 2102 | | | | | | | | | | | | | | | | | | CSE 3107 | | | | | | | | | | | | | CSE 2104 | | | | | | CSE 2101 | |
| $I_4$ | CSE 3104 | | | CSE 3103 | | | | | | CSE 2102 | | | CSE 2107 | CSE 3103 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $I_5$ | | | | CSE 4106 | | | | | | | | | | | | CSE 3104 | | | | | | | CSE 4105 | | | | | | | | | CSE 3119 | | | | | CSE 4105 | | CSE 2132 | | | | | |
| $I_6$ | | | | | CSE 3119 | | | | | | | | | | | | | | CSE 2132 | | | CSE 4107 | | | | | | CSE 3120 | | | | | | | | | | | | | | | | |
| $I_7$ | | | | | | | | | | | | CSE 4109 | | | | | | | | | | CSE 1104 | | | | CSE 4109 | | | | | | | | | | | | | | | | | | | |
| $I_8$ | | | | CSE 1102 | | | | | | CSE 3109 | | | | | | | | | | | | | | | | | | | | CSE 1102 | | | | | | | | | | CSE 3109 | | | CSE 1104 | | |
| $I_9$ | | CSE 4123 | | | | | CSE 3120 | | | | | | | | | CSE 4124 | | | | | | | | CSE 4123 | | | | | | | | | CSE 3108 | | | | | CSE 3101 | | | | | | |
| $I_{10}$ | CSE 4107 | | | | | | | | | CSE 1109 | CSE 4107 | | CSE 4120 | | | | | | | CSE 1109 | | | | | | | | CSE 4113 | | | | | | | | | | | | | | CSE 3103 | | |
| $I_{11}$ | | CSE 2101 | | CSE 2100 | | | | CSE 1109 | | | | | | | | | | | | | | | | | CSE 2101 | | | | | | | | | | | | | | CSE 4113 | | | CSE 4124 | | |
| $I_{12}$ | | | CSE 4105 | | | | | | | | | | CSE 3108 | | | CSE 2100 | | | CSE 4106 | | | | | | | | | | | | | | | | | | | | | | | | | | |

Figure 3. Class schedule for instructors.

Figure 3 shows the result as the final outcome of our proposed algorithm. The classes of the instructors in different days of a week are represented in the figure. The left most column indicates anonymous instructors. In the experimental scenario we have used a data set for nine period of a day and twelve instructors with three classrooms and two lab rooms. This result is optimized in the sense of existing resource utilization of the University. Corresponding schedules for labs and class rooms are also found from the proposed method.

### D. *Experimental Result Analysis*

Our main objective was to satisfy the highly constrained university course scheduling problem using MHPSO. We have experimented results that the class is scheduled preserving the hard and soft constraints. We have also experimented on the resource utilization of the university. The students can be easily migrated to 120 for one year with changing minor modification of the current infrastructure. The accommodation facility can also be utilized in the same manner. The result is satisfactory in respect of the class scheduling, room

scheduling, lab scheduling of the department. And, it can be prescribed that this experiment can result in the improvements of socio-economic condition of the people of Bangladesh because of the nature of our experiment.

The proposed method is superior to other methods because it has several important features that are able to construct more efficient feasible timetables. Traditionally, some important features regarding co-class concept, format of classes, lab scheduling, etc. are either ignored or adopted partially so far which are introduced in the proposed approach that also beget worthy fitness calculation. In room scheduling, minimizing no. of rooms doesn't affect the existing effective class scheduling, and if no. of intakes are maximized, existing class scheduling can also run smoothly.

## IV. CONCLUSION

To meet optimally higher satisfaction at different context, several approaches have been performed to solve UCSP accordingly over the years. But, in question of accomplishment, they differ. Solving UCSP in a highly constraint and realistic environment, this paper proposed MHPSO algorithm which is satisfactorily result-oriented. Most of the universities in Bangladesh have to face such situation where course scheduling is really a difficult task due to crucial constraints and limited resources. In this paper, the practical scenario of Computer Science and Engineering department of KUET, Bangladesh is considered for some remarkable exceptional constraints, different class patterns, and some special features. The outcome scheduled of MHPSO is found feasible for the department and match with the real schedule of the department. In future, it is expected to schedule whole KUET including all the departments. Moreover, it might be interesting to compare performance of the proposed MHPSO with some other methods which remain as future study.

## REFERENCES


[1] M. Chiarandini, M. Birattari, K. Socha and O. R. Doria, "An effective hybrid algorithm for university course timetabling," *Journal of Scheduling*, vol. 9, pp. 403-432, 2006.

[2] G. G .Yen and B. Ivers, "Job shop scheduling optimization through multiple independent particle swarms," *International Journal of Intelligent Computing and Cybernetics*, vol. 2, pp. 5-33, 2009.

[3] I. X. Tassopoulos and G. N. Beligiannis, "Solving effectively the school timetabling problem using particle swarm optimization," *Expert Systems with Applications*, vol. 39, pp. 6029–6040, 2012.

[4] X. Song, "Hybrid particle swarm algorithm for job shop scheduling problems," *Future Manufacturing Systems*, INTECH, pp. 235-268, 2010.

[5] Z. N. Azimi, "Hybrid heuristics for Examination Timetabling problem," *Applied Mathematics and Computation*, vol. 163, no. 2, pp.705-733, 2005.

[6] Y. Z. Wang, "An application of genetic algorithm methods for teacher assignment Problems," *Expert Systems with Applications*, vol. 22, no. 4, pp. 295–302, 2002.

[7] D. Costa, "A tabu search algorithm for computing an operational timetable," *European Journal of Operational Research*, vol. 76, pp. 98-110, 1994.

[8] R. A. Valdes, E. Crespo and J. M. Tamarit, "Design and implementation of a course scheduling system using tabu search," *European Journal of Operational Research*, vol. 137, no. 3, pp. 512–523, 2002.

[9] D. F. Shiau, "A hybrid particle swarm optimization for a university course scheduling problem with flexible preferences," *Expert Systems with Applications*, vol. 38, pp. 235-248, 2011.

[10] X. Luo, D. Wang, J. Tang and Y. Tu, "An improved PSO algorithm for resource Constrained project scheduling problem," in *proc. the 6th world congress on intelligent control and automation (WCICA 2006)*, vol. 1, pp. 3514–3518, 2006.

[11] I. H. Kuo, S. J. Horng, T. W. Kao, T. L. Lin,C. L. Lee and T. Terano, "An efficient flow-shop scheduling algorithm based on a hybrid particle swarm optimization model", *Expert Systems with Applications*, vol. 36, no. 3, pp. 7027–7032, 2009.

[12] S. T. Lo, R. M. Chen, D. F. Shiau and C. L. Wu, "Using particle swarm optimization to solve resource-constrained scheduling problems," in *proc. IEEE conference on soft computing in industrial applications*, pp. 38–43, 2008.

[13] A. Gunawan, J. M. Ng and K. L. Poh, "Solving the teacher assignment-course scheduling problem by a hybrid algorithm", *International Journal of Computer Information, and System Science and Engineering*, vol. 1, no. 2, pp. 136–141, 2007.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# A Neural Network Model for Estimating Global Solar Radiation on Horizontal Surface

Muztoba Ahmad Khan[1,*], Saiful Huque[2] and Azim Mohammad[1]
[1]Department of IPE, Bangladesh University of Engineering & Technology,
[2]Institute of Renewable Energy, University of Dhaka, Bangladesh
[*]E-mail: muztobaahmad@yahoo.com

*Abstract*- **This research focuses on the development of artificial neural network (ANN) model for estimation of daily global solar radiation on horizontal surface in Dhaka. In this analysis back-propagation algorithm is applied. Day of the year, daily mean air temperature, relative humidity and sunshine duration were used as input data, while the daily global solar radiation was the only output of the ANN. The database consists of 1827 daily measured data, between 2008 and 2012, in term of daily mean air temperature, relative humidity and sunshine duration and global solar radiation. The data has been collected from Bangladesh Meteorological Department. The 1461 daily measured data between 2008 and 2011 are used to train the neural networks while the data of 366 (leap year) days from 2012 are used to test the neural network. MATLAB neural network toolbox is used to train and test the network. Both estimated and measured values of daily global solar radiation on horizontal surface were compared during testing phase statistically using two methods: Root Mean Square Error (RMSE) and Regression R Value (R), giving a value of 113.6 Wh/m² and 0.9744, respectively. The results of this study have shown a better accuracy than other conventional prediction models that have been used up to now in Bangladesh. This ANN model may be suitable for predicting solar radiation at any location in Bangladesh, provided that samples of the sunshine duration data from the locations are available.**

***Keywords— Artificial Neural Network, Global Radiation, Solar Radiation, Prediction, Sunshine Duration.***

## I. INTRODUCTION

Energy certainly plays a vital role in development and welfare of human being. There exists a direct correlation between the development of a country and its consumption of energy. To meet the future energy needs the world is looking for a non-exhaustible energy source. If used in a cost effective manner, solar energy can be the best choice among the all non-conventional energies. The amount of solar energy the earth receives each year is ten times more than the energy that can be produced from all fossil reserves available on earth [1]. Generation of electricity from solar energy is gaining popularity as a solution to the growing energy demands. The most important parameter in renewable energy applications is solar radiation. It is expected that the present worldwide research and development program on solar energy will help to solve the future energy crisis of the world.

The applications of solar energy require information of the availability of solar energy for its optimum use. Due to the utilization of solar energy potential in many areas, there is an increasing need for more precise modeling and prediction of solar radiance. Solar radiation data are required by solar engineers, architects and agriculturists for many applications such as solar heating, cooking, drying and interior illumination of buildings [2–5]. Since solar radiation is not uniform over all places on the earth, any solar energy conversion installation at a certain place requires knowledge of the amount of solar radiation at that place which again varies from time to time. But unfortunately, for any cases, solar radiation measurements are not easily available due to the cost and maintenance and calibration requirements of the measuring equipment.

Various solar models have been used to predict daily global solar radiation all over the world. Solar prediction models can be categorized in two distinct groups. First group in this classification are empirical models. Empirical models were used by many researchers to estimate global solar radiation [6–8]. These models usually consist of a few measurable meteorological parameters. Empirical methods to estimate global solar radiation requires the development of a set of equation that relate it to other meteorological parameters. Artificial neural network (ANN) models are the second type of solar prediction models. ANN provides a computationally efficient way of determining nonlinear relationship between a number of inputs and one or more outputs. In recent years, ANN models were used by many researchers to estimate global solar radiation [9–12]. Almost all the literatures concluded that ANN model is superior to other empirical regression models.

Bangladesh is endowed with abundant sunshine for at least 8 months of the year. The prospect of utilization of solar energy is thus very bright. But solar radiation data are not available in many locations of Bangladesh due to absence or malfunction of measuring instruments. However, the climatological data such as sunshine hour, temperature, humidity etc. are available at meteorological department for most districts of Bangladesh. These data can be used in ANN models to estimate the global solar radiation at any location. In Bangladesh, ANN has not been used yet to estimate daily global radiation. There are few papers on solar radiation estimation using empirical models only [13-15]. In this paper, ANN will be used to estimate the daily global solar radiation in Dhaka which lies in the tropics between latitudes 23°43′N and longitudes 90°25′E.

## II. ARTIFICIAL NEURAL NETWORK

Artificial neural network models employ artificial intelligence techniques and are data driven; they learn and memorize a data structure and subsequently simulate the structure. They are able to learn key information patterns within a multidimensional information domain [16]. In a way, artificial neural network mimic the learning process of a human brain and therefore do not need characteristic information about the system; instead, they learn the relationship between input parameters and the output variables by studying previously recorded data. This makes artificial neural network ideal for modeling non-linear, dynamic, noisy data and complex systems [17]. Further, artificial neural networks are good for tasks involving incomplete data sets [18]. Fig. 1 shows a typical neural network, which consists of an input layer, a hidden layer and an output layer. An input $x_j$ is transmitted through a connection, which multiplies its strength by a weight $w_{ij}$ to give a product $x_j w_{ij}$. This product is an argument to a transfer function *f*, which yields an output $y_i$ represented as:

$$y_i = f\left(\sum_{j=1}^{n} x_j w_{ij}\right)$$

Where *i* is an index of neurons in the hidden layer and *j* is an index of an input to the neural network.

There are three steps in solving an ANN problem which are 1) training, 2) generalization and 3) implementation. Training is a process that network learns to recognize present pattern from input data set. We present the network with training examples, which consist of a pattern of activities for the input units together with the desired pattern of activities for the output units. For this reason each ANN uses a set of training rules that define training method. Generalization or testing evaluates network ability in order to extract a feasible solution when the inputs are unknown to network and are not trained to network. We determine how closely the actual output of the network matches the desired output in new situations. In the learning process the values of interconnection weights are adjusted so that the network produces a better approximation of the desired output. ANNs learn by example. They cannot be programmed to perform a specific task. The examples must be selected carefully otherwise useful time is wasted or even worse the network might be functioning incorrectly. The disadvantage is that because the network finds out how to solve the problem by itself and its operation can be unpredictable. In this paper the effort is made to identify the best fitted network for the desired model according to the characteristics of the problem and ANN features.

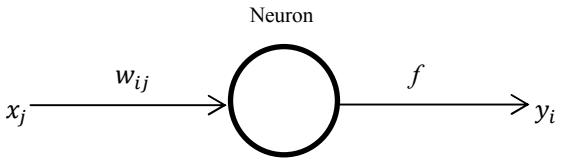



Figure 1. Typical neuron in a neural network system.

## III. METHODOLOGY

Day of the year, daily mean air temperature (°C), relative humidity (%), daily global solar radiation (Wh/m²) and sunshine duration (hour) daily data between year 2008 and 2012 are collected from Bangladesh Meteorological Department. The database is consists of 1827 data sets.

The data set was split into two sets. The training data set: the group of data by which the network adjusts, in order to reach the best fitting of the nonlinear function representing the phenomenon and it consisted of 1461 data sets. The testing data set: a set of new data used to evaluate the developed artificial neural network model generalization and it consisted of 366 (leap year) data sets. The training data set was between the year 2008 and 2011. The testing data set was from the year 2012.

The ANN model was developed using neural network tool of MATLAB version R2011b. For the training process of the ANN, a Bayesian regulation back propagation algorithm was used. This algorithm is a supervised iterative training method that updates the weights and bias values according to Levenberge Marquardt optimization [19]. It minimizes a linear combination of squared errors and weights, and then uses Bayesian regularization to determine the correct combination that results in a network that generalizes satisfactorily. The number of hidden neurons in an ANN is a function of the problem's complexity, the number of input and output parameters, and the number of training cases available. A trial and error process was used to determine the number of hidden neurons. After trying a number of different configurations, and repeating each training process ten times to avoid random errors, it was found that 22 neurons in the hidden layer yielded the best results with a reasonable computational effort. Two transfer functions were investigated, including the tangent sigmoid and log sigmoid functions. Linear transfer function was used for both input layer and output layer. Tangent sigmoid function was used for the hidden layer.

MATLAB representation of the final neural network model is presented in Fig. 2. The characteristics of the developed artificial neural network model used in the present study are presented in Table I.

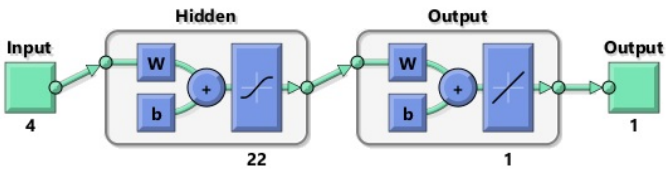



Figure 2. Final neural network architecture.

TABLE I. CHARACTERISTICS OF THE ANN MODEL

| Item | Value |
|---|---|
| Number of Layers | 3 |
| Input Layer Nodes | 4 |
| Transfer Function | Linear |
| Hidden Layer Nodes | 22 |
| Transfer Function | Sigmoid |
| Output Layer Nodes | 1 |
| Transfer Function | Linear |

## IV. RESULTS AND DISCUSSION

Although the tested period seems short but it has appeared that the developed artificial neural network model with one hidden layers based on the standard back propagation algorithm, using tangent sigmoid transfer function in hidden layer and linear transfer function in output and input layers, resulted as a very efficient model to estimate daily global solar radiation on horizontal surface at Dhaka city (Bangladesh). The comparison between estimated and measured values during testing is depicted in Fig. 3 for daily global solar radiation on horizontal surface.

The performance of the neural network model is measured using RMSE (Root Mean Squared Error) and Regression R Value. Regression R Values measure the correlation between outputs and targets. An R value of 1 means a close relationship and 0 means a random relationship. Mean Squared Error is the average squared difference between outputs and targets. Lower values are better. Zero means no error.

The value of R was 0.9777 and 0.9744 in training and testing phases, respectively. The proposed artificial neural network model, which accepts 4 input variables, predicts with a RMSE of 113.6 Wh/m². These lower values of errors demonstrate that, the proposed artificial neural network model can estimate daily global solar radiation on horizontal surface for the testing data set with reasonable accuracy.

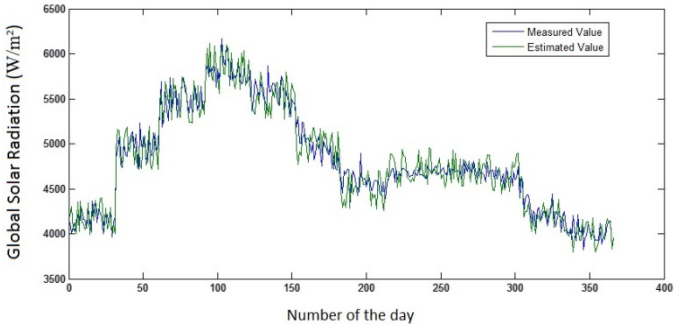



Figure 3. Comparison between estimated and measured values.

When compared with an empirical model [13], it is found that the results calculated by ANN model are better than that calculated by empirical models. Table II summarizes the correlations and error analyses which result from the comparison between estimated values (ANN model and empirical) and measured values.

TABLE II. RESULT OF PERFORMANCE ANALYSES

| Model | R | RMSE (Wh/m²) |
|---|---|---|
| Empirical Model | 0.9304 | 326.8 |
| ANN Model | 0.9744 | 113.6 |

## V. CONCLUSION

The use of ANN technique in modeling daily solar radiation on horizontal surface at Dhaka has been reported. The results of validation and comparative study indicate that the ANN based estimation technique for solar radiation is more suitable to predict the solar radiation than the empirical regression models proposed by other researchers. This study confirms the ability of the ANN to predict solar radiation values more precisely. Therefore, this ANN model may be suitable for predicting solar radiation at any location in Bangladesh, provided that the necessary data from the locations are available.

## REFERENCES


[1] P. Wolfgang, *Solar Electricity: An Economic Approach to Solar Energy*, United Kingdom: Butterworth, 1977.

[2] L. T. Wong and W. K. Chow, "Solar radiation model," *Applied Energy,* vol. 69, no. 3, pp. 191-224, 2001.

[3] D. H. W. Li and J. C. Lam, "Solar heat gain factors and the implications for building designs in subtropical regions," *Energy and Buildings*, vol. 32, no. 1, pp. 47-55, 2000.

[4] Z. Lu, R. H. Piedrahita and C. D. S. Neto, "Generation of daily and hourly solar radiation values for modeling water quality in aquaculture ponds," *Transactions of the ASAE*, vol. 41, no. 6, pp. 1853-1859, 1998.

[5] R. Kumar and L. Umanand, "Estimation of global radiation using clearness index model for sizing photovoltaic system," *Renewable Energy*, vol. 30, no. 15, pp. 2221-2233, 2005.

[6] S. A. Khalil and A. M. Fathy, "An empirical method for estimating global solar radiation over egypt," *Acta Polytechnica*, vol. 48, no. 5, pp. 48-53, 2008.

[7] S. K. Srivastava, O. P. Sinoh, and G. N. Pandy, "Estimation of global solar radiation in uttar pradesh (india) and comparison of some existing correlations," *Solar Energy*, vol. 51, no. 1, pp. 27-29, 1993.

[8] J. Davies, M. Abdel-Wahab, and D. Mekay, "Estimating solar irradiance on horizontal surface," *Int. J. Sol. Energy*, vol. 2, pp. 405, 1984.

[9] J. Mubiru, "Using artificial neural networks to predict direct solar irradiation," *Advances in Artificial Neural Systems*, vol. 2011, pp. 1-6, 2011.

[10] T. Khatib, A. Mohamed, M. Mahmoud, and K. Sopian, "Estimating global solar energy using multilayer perception artificial neural network," *International journal of energy*, vol. 6, no. 1, pp. 82-87, 2012.

[11] M. A. Abdulazeez, "Artificial neural network estimation of global solar radiation using meteorological parameters in gusau, nigeria," *Archives of Applied Science Research*, vol. 3, no 2, pp. 586-595, 2011.
[12] M. Benghanem, A. Mellit, and S. N. Alamri, "ANN-based modelling and estimation of daily global solar radiation data: A case study," *Energy Conversion and Management*, vol. 50, pp. 1644-1655, 2009.
[13] H.R.Ghosh, L.Mariam, M.Sadia, S.K.Khadem, N.C.Bhowmik and M.Hussain, "Estimation of monthly averaged daily and hourly global & diffuse radiation for Bangladesh," *The Dhaka University Journal of Science*, vol. 54, no. 1, pp. 109-113, 2006.
[14] M. Arif and S. Bhuiyan, "Estimation of solar radiation: an empirical model for Bangladesh", *The IIUM Engineering Journal*, vol. 14, no. 1, pp. 103-117, 2013.
[15] M. A. Hena and M. S. Ali, "A Simple Statistical Model to Estimate Incident Solar Radiation at the Surface from NOAA AVHRR Satellite Data," *I.J. Information Technology and Computer Science*, vol. 5, no. 2, pp. 36-41, 2013.
[16] S. A. Kalogirou, "Artificial neural networks in renewable energy systems applications: a review," *Renewable and Sustainable Energy Reviews,* vol. 5, no. 4, pp. 373-401, 2001.
[17] J. Mubiru, "Predicting total solar irradiation values using artificial neural networks," *Renewable Energy*, vol. 33, no. 10, pp. 2329-2332, 2008.
[18] S. A. Kalogirou, 2000. "Applications of artificial neural-networks for energy systems," *Appl. Energy*, vol. 67, no. 10, pp. 17-35, 2000.
[19] F. D. Foresee and M. T. Hagan, "Gauss-Newton approximation to Bayesian learning," *International Joint Conference on Neural Networks*, vol. 3, pp. 1930-1935, 1997.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Complex-valued Neural Network Using Magnitude Encoding Technique For Real-valued Classification Problems & Time Series Prediction

Shahriar Morshed*, Nizam Uddin Ahmed, and Md. Shahjahan
Department of Electrical and Electronic Engineering,
Khulna University of Engineering & Technology,
Khulna – 9203, Bangladesh.
* E-mail: smm2k8@mail.com

*Abstract*— **In this paper a new conversion technique is proposed for complex-valued neuron (CVN) to convert real value into complex value in order to solve real-valued classification problems & Time series analysis. Previously phase encoding system was used to solve these types of problems. In this proposed encoding system, each real-valued input is converted into complex value according to the input real value with a fixed phase. In this model the input magnitude ranges from the lowest and highest value of the given input. The converted value is then multiplied by complex weight and then they sums up to feed into an activation function. The activation function converts the complex value into real value within a certain range. We used this encoding system in solving different Boolean problems. Some real world benchmark problems are also tested by this process. Different time series analysis is performed to test the prediction ability of this encoding system. The result shows that this magnitude encoding provides better accuracy in different benchmark problems. Especially in case of predicting time series data, this encoding system provides better result than the phase encoding system.**

***Keywords—Magnitude encoding, Time series analysis, Complex-valued neural network(CVNN), Classification***

## I. INTRODUCTION

The importance of expressing a real value into a complex value is inevitable for real world problems such as signal processing, where both the phase and amplitude are needed to be treated properly. Without expressing real values in complex domain it is impossible to deal with phase. So for artificial neural networks the complex value is employed through different types of encoding system. Previously complex-valued neural network (CVNN) and its back propagation algorithms were developed to use complex value in artificial neural networks [1], [2], [4]. The complex-valued neural network (CVNN) has been used for real valued classification problem by many other researchers [3], [5], [6]. But in this paper we also discuss about time series analysis using CVNN.

We are conscious of two style of conversion system for the application of CVNN for solving time series prediction and real world benchmark problems. In Ref [3] each input real value is encoded through phase encoding in the range of 0 to π of unity magnitude to employ the complex value into complex valued neurons. In this paper we propose a new encoding system which is magnitude encoding an also can be said line encoding. The gradient descent learning rule was used to learn the network. According to this learning rule both the real and imaginary part of activation function is needed to be differentiable. By considering this, two activation functions were used which is proposed earlier in the paper of Murase and Amin [3]. We also considered a single layered CVNN consists of n CVNs to solve n class problems and for this aim we create the learning and classification format.

In the remaining section of paper we will discuss about the various encoding system, the gradient descent learning rule and update equation, and the ability of CVNN for benchmark problems and time series analysis for different encoding system. Finally we will make the completion by discussion and conclusion.

## II. CVN MODEL & LEARNING RULE

### *A. Proposed encoding for CVN input:*

The real data is necessary to converted into complex value through a conversion process though CVN performs operation over complex data. In Fig. 1 the whole model of CVN architecture is shown. The real inputs are converted into complex value which is multiplied with complex weights. Then these values sum up including the complex bias. This complex sum is then fed into the activation function to provide real output value into a desired range.

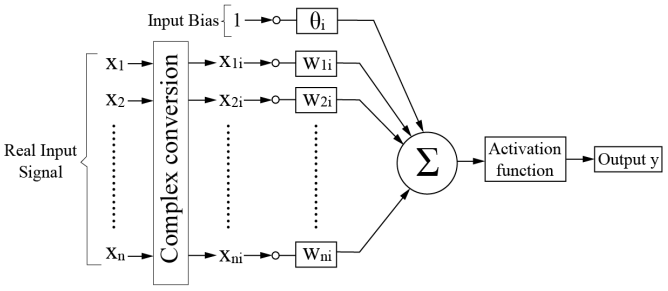



Figure 1. Architectural model of complex neuron

For performing operation in complex domain we need a process which maps the input real data into complex data. To map a real value into complex value we have to maintain following equation.

$$z = \alpha e^{i\varphi} \quad (1)$$

where, $\alpha$ = magnitude and $\varphi$ = phase.

Here two variables are present. One is phase and other is magnitude. If we vary phase and make the magnitude constant then it is called phase encoding. In phase encoding technique, the input value is distributed over the upper half cycle of a unit circle as shown in Fig. 2. Phase encoding is done through following equations:

Let $x \in [a,b]$ where $a,b \in R$

then, $$\varphi = \frac{\pi(x-a)}{(b-a)} \quad (2)$$

and $$z = e^{i\varphi} = \cos\varphi + i\sin\varphi \quad (3)$$

But now we propose a new encoding system which is magnitude encoding. Here the magnitude is varied by making the phase constant. So the input real value is distributed over a line of fixed phase as shown in Fig. 3. So it is also called line encoding. This is done through following equation:

Let, $x \in [a,b]$ where, $a,b \in R$ then,

$$z = x(\cos\varphi + i\sin\varphi) \quad (4)$$

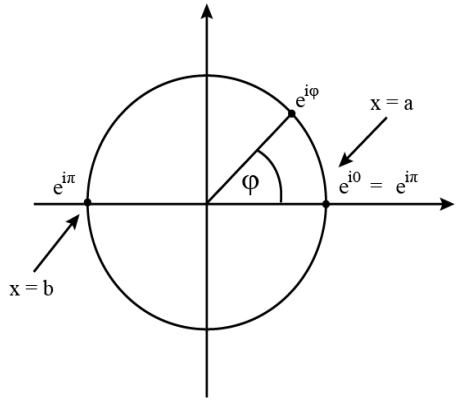



Figure 2. Complex conversion by phase encoding system. If the input value ranges within the interval $[a,b]$, then the corresponding complex value will range within the interval $[e^{i0}, e^{i\pi}]$.

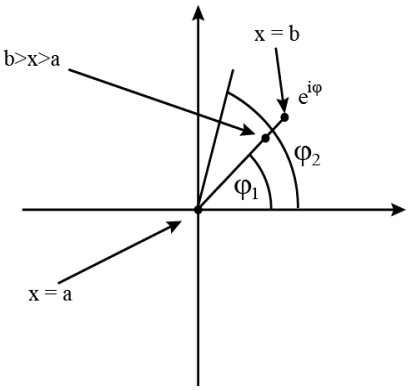



Figure 3. Complex conversion by magnitude encoding. If the input value ranges within the interval $[a,b]$, then the equivalent complex value will range within the interval $[ae^{i\varphi}, be^{i\varphi}]$, where $\varphi$ remains fixed for a particular case analysis.

*B. Learning process*

Here we have developed a gradient descent learning rule for a CVNN which is the combinatin of multiple CVNs without having any hidden layer [3]. Consider a p-q CVNN, where p is the number of input nodes and q is the number of output nodes. Let $X = [x_1 .....x_p]^T$ be a p dimensional complex valued input vector, $W_g = [w_{g1} .....w_{gp}]^T$ the complex valued weight vector of neuron g, and $\theta_g$ the complex valued bias for the $g_{th}$ neuron. To express the real and imaginary parts,

let $x_m = x_m^R + ix_m^I$, $w_{gm}^R + iw_{gm}^I$ and $\theta_g = \theta_g^R + i\theta_g^I$ .Then the net input $z_g$ and the output $y_g$ of the $g_{th}$ neuron are given by,

$$z_g = \sum_{m=1}^{p} w_{gm}x_m + \theta_g \quad (5)$$

$$z_g = (\sum_{m=1}^{p} (w_{gm}^R x_m^R - w_{gm}^I x_m^I) + \theta_g^R)$$

$$+ i(\sum_{m=1}^{p} (w_{gm}^I x_m^R + w_{gm}^R x_m^I) + \theta_g^I) \quad (6)$$

$$z_g = z_g^R + iz_g^I \quad (7)$$

And,

$y_g = f_{C\to R}(z_g)$ , where $f_{C\to R}(z_g)$ is an activation function

Here two activation functions are used to limit the out value within a finite value in the range of [0, 1]. If $z = u + iv$ then the activation functions are,

$$f_{C\to R}(z) = \sqrt{(f_R(u))^2 + (f_R(v))^2} \quad (8)$$

$$f_{C\to R}(z) = (f_R(u) - f_R(v))^2 \quad (9)$$

Where $f_R(x) = 1/(1+\exp(-x))$

If the desired output of the $g_{th}$ neuron is $d_g$, then the error function is to be minimized during the training is given by,

$$E = \frac{1}{2}\sum_{g=1}^{q}(d_g - y_g)^2 = \frac{1}{2}\sum_{g=1}^{q} e_g^2 \quad (10)$$

During the training, the biases and the weights are updated according to the following equations:

$$\Delta\theta_g = -\eta\frac{\partial E}{\partial \theta_g^R} - i\eta\frac{\partial E}{\partial \theta_g^I} = \Delta\theta_g^R + i\Delta\theta_g^I \quad (11)$$

$$\Delta w_{gm} = -\eta\frac{\partial E}{\partial w_{gm}^R} - i\eta\frac{\partial E}{\partial w_{gm}^I} = \Delta w_g^R + i\Delta w_g^I = \overline{x_m}\Delta\theta_g \quad (12)$$

Where the learning rate is $\eta$ and $\overline{x_m}$ is the complex conjugate of the complex number $x_m$.

If the activation function is given by (7), then,

$$\Delta\theta_g^R = \eta e_g \frac{f_R(z_g^R)}{y_g} f_R^{'}(z_g^R) \quad (13)$$

$$\Delta\theta_g^I = \eta e_g \frac{f_R(z_g^I)}{y_g} f_R^{'}(z_g^I) \quad (14)$$

Where $f_R(u) = 1/(1+\exp(-u))$ and $f_R^{'}(z_g^R) = f_R(u)(1-f_R(u))$, $u \in R$. If the activation function is given by (8), then,

$$\Delta\theta_g^R = 2\eta e_g (f_R(z_g^R) - f_R(z_g^I)) f_R^{'}(z_g^R) \quad (15)$$

$$\Delta\theta_g^I = 2\eta e_g (f_R(z_g^I) - f_R(z_g^R)) f_R^{'}(z_g^I) \quad (16)$$

Using this equation the weights and biases can be updated

## III. ANALYZED PROBLEMS

As Boolean problem we used 7 bit input pattern to determine whether the bit pattern is symmetric or not, with respect to the center value of the pattern. To analyze real world classification problems, we used Cancer and Diabetes datasets from PROBEN1 [7]. Table II displays the variables used in both datasets. In case of diabetes testing, our aim was to detect whether it is diabetes positive or not by learning form the data of the input attributes, which are based on pregnancy, age, glucose tolerance level, mass of the body & blood pressure. In Cancer testing, we decided whether the tumor is benign or malignant according to the input data where the attributes are based on cell size and cell shape, bare nuclei frequency and the amount of marginal adhesion.

TABLE I. CANCER & DIABETES DATASETS

| *Variable name* | **Number of variables** | |
|---|---|---|
| | ***in Cancer dataset*** | ***in Diabetes dataset*** |
| Input | 9 | 8 |
| Output | 2 | 2 |
| Examples | 699 | 768 |
| Learning examples | 349 | 384 |
| Testing Examples | 349 | 384 |

In case of time series analysis we used AMEX (American Stock Exchange) stock price close data as input dataset. Among 2541 days of data we used first 2541 data for learning 2541 for testing. We also use the dataset of DOW Chemical Company for time series analysis. Here we have chosen 594 data for learning and 594 for testing.

## IV. EXPERIMENTAL SETUP & RESULT

### *A. Experimental set up:*

The data set of Cancer and Diabetes problems as given in PROBEN1 [7] and time series given in AMEX & DOW analysis contain real value which is converted into complex value through magnitude encoding which we discussed in section 3. Winner takes all method is used to determine the class, which states that for input pattern the highest output from the output neuron, will designate the class.

The complex valued back propagation algorithm was used to train the CVNN for both encoding system. Throughout the training process we kept the learning rate at 0.1 for both case. Weights and biases are randomly initialized from a uniform distribution in the range of [-0.5, 0.5] in case of CVNN real and imaginary parts.

The equation which is used for calculating the mean squared error is as follows:

$$E = \frac{1}{2K}\sum_{k=1}^{K}\sum_{i=1}^{q}(d_{ki} - y_{ki})^2 \quad (17)$$

Where k is the pattern number, K is the number of training patterns, and q is the number of output neurons.

*B. Result:*

Here we have made the comparison between phase encoding & magnitude encoding. Different real world benchmark datasets and time series datasets are used to determine which encoding is better. In every case 10 independent runs are made to achieve the best result. In case of benchmark problem analysis we used diabetes & cancer datasets. For cancer data testing, Fig. 4 depicts that, the mean squared error is higher for magnitude encoding than from phase encoding.

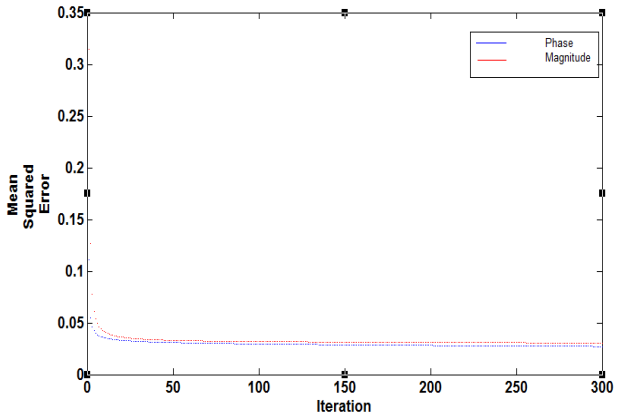



Figure 4. Learning process curve of cancer dataset for both phase and magnitude encoding system

Table II & III shows the results of accuracy for both Diabetes and Cancer testing. In both case accuracies are almost the same. But from Table III it is observed that the accuracy rate is higher for our proposed magnitude encoding for every independent test.

TABLE II. CLASSIFICATION RESULTS FOR DIABETES

| **Encoding process** | ***Classification ability*** | **Activation Function** | |
|---|---|---|---|
| | | ***Activation function 1*** | ***Activation function 2*** |
| Phase Encoding | Classified | 293 | 294 |
| | Misclassified | 91 | 90 |
| | Accuracy | 76.30% | 76.56% |
| Magnitude Encoding | Classified | 293 | 297 |
| | Misclassified | 91 | 87 |
| | Accuracy | 76.30% | 77.34% |

TABLE III. CLASSIFICATION RESULTS FOR CANCER

| **Encoding process** | ***Classification ability*** | **Activation Function** | |
|---|---|---|---|
| | | ***Activation function 1*** | ***Activation function 2*** |
| Phase Encoding | Classified | 337 | 339 |
| | Misclassified | 12 | 10 |
| | Accuracy | 96.56% | 97.13% |
| Magnitude Encoding | Classified | 341 | 341 |
| | Misclassified | 8 | 8 |
| | Accuracy | 97.70% | 97.70% |

In time series analysis, the aim is to predict the distribution of data as accurate as possible. In case of prediction of AMEX stock price close data, we used data of 2541 days. From Fig. 5 it is apparent that for phase encoding, there exists higher deflection between the desired & predicted curve. But from Fig. 6 for magnitude encoding, the predicted curve almost overlaps with the desired curve i.e. the magnitude encoded curve traces the desired curve with high accuracy.

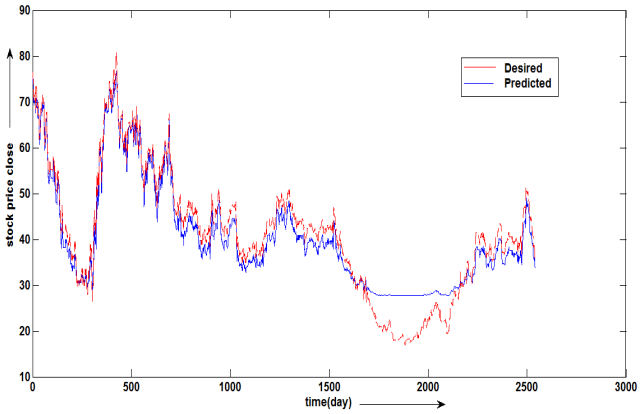



Figure 5. Prediction of AMEX stock price close data using Phase encoding

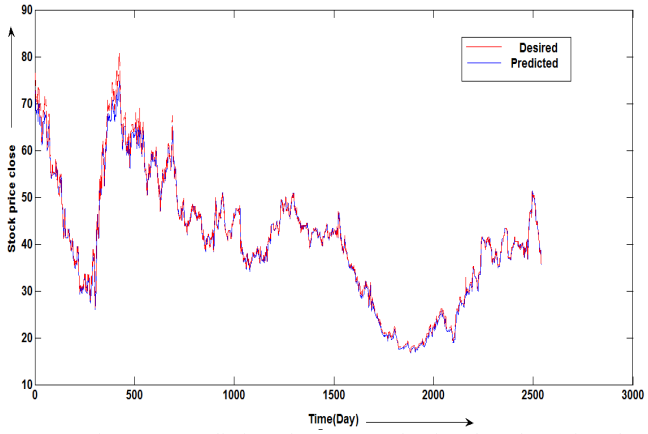



Figure 6. Prediction of AMEX stock price close data using Phase encoding

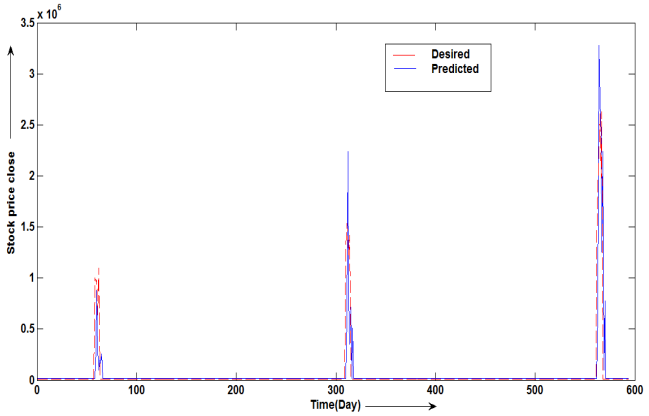



Figure 7. Prediction of DOW stock price close data using Phase encoding

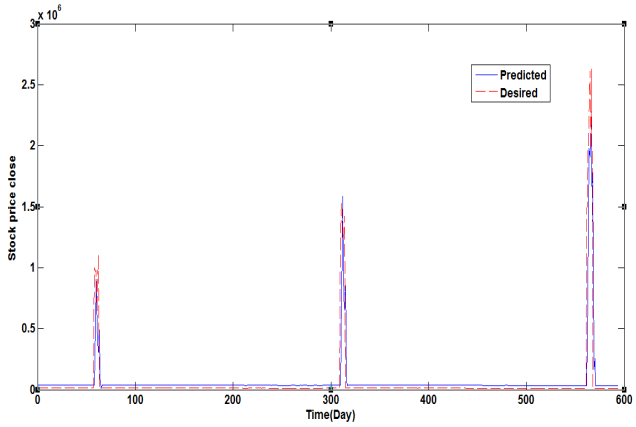



Figure 8. Prediction of DOW stock price close data using Magnitude encoding

In case of DOW close data prediction analysis we also found the similar better performance for magnitude encoding technique.

## V. CONCLUSION

In this paper our aim was to show the better effectiveness of magnitude encoding rather than phase encoding in case of solving different problems. Though in case of real world benchmark problems, the mean square error was higher for magnitude encoding than the phase encoding, but our proposed encoding system shows better accuracy than the previous one to classify dataset in some particular benchmark problems like cancer data testing. In case of other benchmark problem this encoding system shows almost the same accuracy as the phase encoding system. In time series analysis our research shows that, stock exchange data are predicted with less deflection by magnitude encoding than phase encoding system. We tried to find a new way of solving different real world problems which may enhance the ability of complex-valued neural network. From the better result of different problem analysis especially for time series analysis, makes us think that, the conversion technique for CVNN requires further research.

## REFERENCES

[1] A. Hirose, *Complex-Valued Neural Networks*, Springer, Berlin, 2006.

[2] H.E. Michel, A.A.S. Awwal, Enhanced artificial neural networks using complex-numbers, in *Proc.IJCNN'99 International Joint Conference on Neural Network*, vol. 1, 1999, pp. 456 – 461.

[3] Md. Faijul Amin, Kazuyuki Murase , "Single-layered complex-valued neural network for real-valued classification problems", *Neurocomputing* 72 (2009) 945– 955.

[4] T. Nitta, An extension of the backpropagation algorithm to complex numbers, *Neural Networks* 10 (8) (1997) 1391–1415.

[5] I. Aizenberg, C. Moraga, Multilayer feedforward neural network based on multi-valued neurons and a backpropagation learning algorithm, *Soft Comput.* 11 (2) (2007) 169–183.

[6] G. Georgiou, C. Koutsougeras, Complex backpropagation, *IEEE Trans. Circuits Syst.* II 39 (5) (1992) 330–334.

[7] L. Prechelt, "PROBEN1—A Set of Benchmarks and Benchmarking Rules for Neural Network Training Algorithms", University of Karlsruhe, Germany.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Bangla Handwritten Character Recognition Using Deep Belief Network

Md. Musfiqur Rahman Sazal[1,*], Sujan Kumar Biswas[1], Md. Faijul Amin[1], and Kazuyuki Murase[2]
[1]Khulna University of Engineering & Technology, Bangladesh
[2]University of Fukui, Japan
E-mail: musfiqsazal@msdn.net.bd[*], sujan.kuet@msdn.net.bd, mdfaijulamin@yahoo.com, and murase@u-fukui.ac.jp

*Abstract*— **Recognition of Bangla handwritten characters is a difficult but important task for various emerging applications. For better recognition performance, good feature representation of the character images is a primary requirement. In this study, we investigate a recently proposed machine learning approach called deep learning [1] for Bangla hand written character recognition, with a focus on automatic learning of good representations. This approach differs from the traditional methods of preprocessing the characters for constructing the handcrafted features such as loops and strokes. Among different deep learning structures, we employ the deep belief network (DBN) that takes the raw character images as input and learning proceeds in two steps – an unsupervised feature learning followed by a supervised fine tuning of the network parameters. Unlike traditional neural networks, the DBN is a probabilistic generative model, i.e., we can generate samples from the model and it can fit both the semi-supervised and supervised learning settings. We demonstrate the advantages of unsupervised feature learning through the experimental studies carried on the Bangla basic characters and numerals dataset collected from the Indian Statistical Institute.**

***Keywords—Deep belief network, unsupervised feature learning, Bangla handwritten character recognition, supervised learning, backpropagation.***

## I. INTRODUCTION

Handwritten character recognition has received a lot of attention because of its various applications such as postal mail sorting according to zip code [2], signature verification, and bank-check processing. In the Bangla language, however, most of the works have been done for the recognition of printed characters [3]. While each category of the printed characters has relatively less variation in the character images and easier to be dealt with, it is far more difficult to recognize a handwritten character because of various factors like inconsistency in the writer's handwriting, different ways of writing and noise in the data. Moreover, the characters in Bangla language have high variations such as different shapes, sizes, loops and strokes. Therefore, the recognition of Bangla handwritten characters is a challenging task.

In order to achieve better recognition performance, research on character or object recognition from images has been heavily relied on good feature representation, which involves hand-crafted feature engineering and goes through complicated preprocessing steps [4]. Such feature engineering approach is application dependent and requires human prudence or ingenuity. Consequently, when building an application of pattern recognition, most of the human effort is spent for constructing discriminating features and less is given to the classifier design [5].

Recently, hierarchical feature learning by deep architectures in an unsupervised manner has revolutionized the machine learning researches [1]. The unsupervised feature learning is a new perspective whose goal is to learn good representation of input data automatically without considering the labels [6]. In order to achieve this goal layers of neurons are arranged hierarchically to form a deep architecture. Each layer learns a new representation from its previous layer with a goal of modeling different explanatory factors of variation behind the data. It is hypothesized that high level complex features can be learned from low level simpler features [7]. For example, in vision problems, bottom layer might learn edges from the scene. The next layer could discover contours and the subsequent layers could learn even more complex features. Once good features are learnt, they can be exploited in various tasks.

In this paper, we study deep belief network (DBN) [1] for Bangla handwritten character recognition problem. The DBN is a particular type of deep architecture that can gradually learn complex structures of the data by learning its probability distribution. It takes raw image data as input with the hope that subsequent layers would learn good feature representation. Learning in the network takes two steps – an unsupervised feature learning followed by a supervised learning of discriminating function. The unsupervised learning is performed by using contrastive divergence (CD) algorithm which is an approximation of maximum likelihood estimation [8], while in the later stage the network parameters are fine-tuned with the gradient based backpropagation (BP) algorithm [9]. We conducted our experimental studies on a dataset of Bangla handwritten characters and numerals from the Indian Statistical Institute. It is demonstrated here that on the one hand, the unsupervised feature learning relieves human effort for constructing application dependent features, and on the other hand, the neural network classifier achieves much higher recognition performance than the classifier trained only by BP algorithm.

The organization of paper is as follows. Section II provides an overview of the related works. Section III describes deep

belief network. Section IV presents details of our experiment and a performance comparison. Finally, Section V provides conclusion.

## II. Related Works

There has been extensive research works on handwritten character recognition for different languages, such as Roman scripts, English, several European languages, and some Asian languages like Chinese, Korean and Japanese. However, relatively few works have been done for the Bangla language. An important research contributions relating to the handwritten Bangla characters involve a multistage method developed by Rahman et al. [10]. Another interesting work is an MLP classifier developed by Bhowmik et al. [11]. The prime features for the multistage approach include upper part of the character, vertical line, and double vertical line. For the MLP classifier [11], alphabetic features were designed by the human expert.

The DBN, an alternative approach that do not require feature engineering by the human expert has been recently developed and employed for recognizing English handwritten numerals [12]. Here the features are extracted from the pixel data of character image in an unsupervised manner, i.e., without considering the labels. Another technique that extracts local features is the convolutional neural network [13], [14]. Note that these methods have a common goal of learning distributed representation [9] which resulted in the resurgence of research in the neural networks. However, it was realized that gradient based BP algorithm is not suitable for several layered architectures. The DBN and convolutional neural network avoids such difficulties, but retain the original goal of learning distributed representation. We noted that such approaches have not been applied yet on the handwritten Bangla character recognition problem. Therefore, we study here one of aforementioned methods, the DBN, to analyze its recognition performance on the Bangla handwritten character recognition task.

## III. Deep Belief Network (DBN)

The DBN is a multilayer neural network that can work as a probabilistic generative model [1]. The DBN as shown in Figure 1 is composed of several layers of stochastic hidden variables and one layer of visible units. The basic module of a

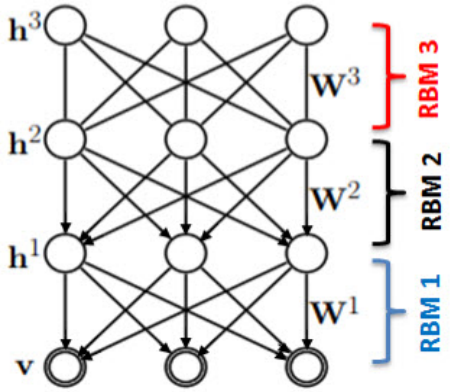



Figure 1. DBN composed with three stacking RBMs

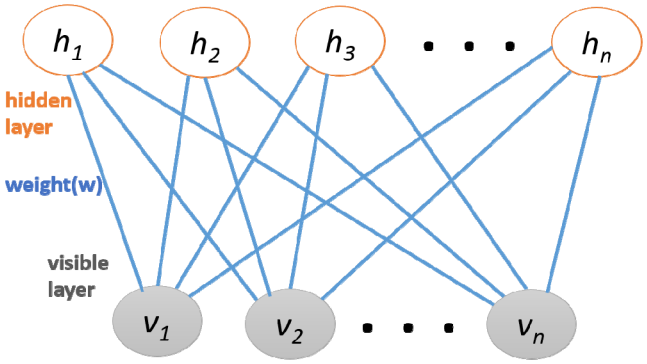



Figure 2. An RBM with no hidden-to-hidden or visible-to-visible connections.

DBN is called restricted Boltzmann machine which we explain next.

### A. Restricted Boltzmann Machine (RBM)

The RBM is a stochastic neural network consisting of one layer of visible units and one layer of hidden units. It can be considered as a bipartite graph where all visible units are connected to all hidden units, but there are no visible-visible or hidden-hidden connections as shown in Figure 2.

The weights on the connecting edges and the bias units define a probability distribution over a binary vector of visible units, $\mathbf{v} = [v_1, v_2, \ldots, v_V]$ and a binary vector of hidden units, $\mathbf{h} = [h_1, h_2, \ldots, h_H]$. The energy function for the joint configuration is given by

$$E(\mathbf{v}, \mathbf{h}; \theta) = -\sum_{i=1}^{V}\sum_{j=1}^{H} w_{ij} v_i h_j - \sum_{i=1}^{V} b_i v_i - \sum_{j=1}^{H} a_j h_j \quad (1)$$

where $\theta = (\mathbf{W}, \mathbf{b}, \mathbf{a})$ and $w_{ij}$ represents the weights between visible units $i$ and hidden units $j$ and $b_i$ and $a_j$ are their biases. $V$ and $H$ are the number of visible and hidden units. The probability of the visible vector $\mathbf{v}$ is defined by

$$p(\mathbf{v}; \theta) = \frac{\sum_{h} e^{-E(v,h)}}{\sum_{n}\sum_{h} e^{-E(u,h)}} \quad (2)$$

As there is no connection among the units of same layer, conditional distributions (see [15] for details) are given by

$$p(h_j = 1 | \mathbf{v}; \theta) = \sigma\left(\sum_{i=1}^{V} w_{ij} v_i + a_j\right) \quad (3)$$

$$p(v_i = 1 | \mathbf{h}; \theta) = \sigma\left(\sum_{j=1}^{H} w_{ij} h_j + b_i\right) \quad (4)$$

where $\sigma(x) = (1 + e^{-x})^{-1}$ is an activation function. Thus an RBM models a joint distribution of visible and hidden units.

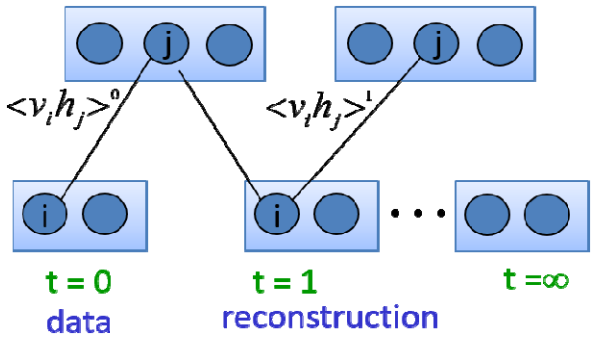



Figure 3. Training a single layer RBM with CD

For training an RBM to model joint distribution of data and class labels [16], an additional visible vector is added with a binary vector of class labels, $\mathbf{l} = [l_1, l_2, \ldots, l_L]$. The energy function with class labels becomes:

$$E(\mathbf{v}, \mathbf{l}, \mathbf{h}; \theta) = -\sum_{i=1}^{V}\sum_{j=1}^{H} w_{ij} h_j v_i - \sum_{y=1}^{L}\sum_{j=1}^{H} w_{yj} h_j l_y - \sum_{j=1}^{H} a_j h_j - \sum_{y=1}^{L} c_y l_y - \sum_{i=1}^{V} b_i v_i \quad (5)$$

and here,

$$p(l_y = 1|\mathbf{h}; \theta) = \text{softmax}\left(\sum_{j=1}^{H} w_{yj} h_j + c_y\right) \quad (6)$$

Then $p(\mathbf{l} \mid \mathbf{v})$ is calculated using

$$p(\mathbf{l} \mid \mathbf{v}) = \frac{\sum_{h} e^{-E(v,l,n)}}{\sum_{l}\sum_{h} e^{-E(u,l,h)}} \quad (7)$$

### B. Training RBM

An RBM is trained to learn probability distribution of the data with the help of hidden stochastic variables. Since maximizing the likelihood requires a Markov chain Monte Carlo simulation taking a long time, an approximation called CD is generally employed [17]. The training process of a single layer RBM using CD is shown in Figure 3. In the CD procedure, the weights of the RBMs are updated to learn the features from the sample data. The weight update rule is defined as:

$$w_{ij}^{new} = w_{ij}^{old} + \Delta w_{ij} \quad (8)$$

$$\text{Where } \Delta w_{ij} = \varepsilon\left(\left\langle v_i h_j \right\rangle_{data} - \left\langle v_i h_j \right\rangle_{model}\right) \quad (9)$$

Here, $\left\langle v_i h_j \right\rangle_{data}$ is the expectation over training data and $\left\langle v_i h_j \right\rangle_{model}$ is the expectation over reconstructed data and $\varepsilon$ is the learning rate parameter.

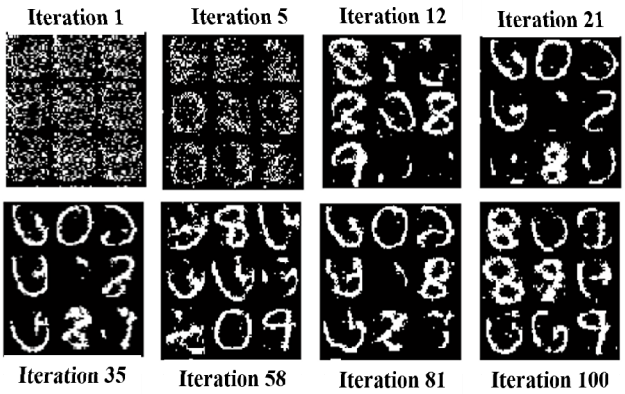



Figure 4. Reconstructed digit images with the corresponding iteration number

### C. Training DBN

Basically, learning in the DBN involves a greedy layer-wise unsupervised learning algorithm. When the weights of an RBM have been learned, the vector of hidden feature activations can be used as visible data for training the next RBM that learns a higher layer of features. In this way, a greedy layer-wise unsupervised training is done in the DBN with RBMs as the building blocks for each layer. This is also called pertaining [16], after which a softmax layer (Eq. (6)) is added for classification problem and the entire weights are optimized using BP algorithm to minimize the error rate.

## IV. Experiment and Performance Analysis

In order to evaluate the effectiveness of unsupervised feature learning approach using DBN, we investigated the performance on the dataset of Bangla numerals and other basic characters collected from Indian Statistical Institute [18]. There are 10 numerals and 50 basic characters in the data set.

We used 27900 images for training and 8600 images for testing. All the images were rescaled to 28×28 pixels and binarized. The architecture of DBN was 784-500-1000-2000-60, i.e., it consisted of three RBMs. The DBN was trained with the algorithms described in Section III. In order to see the usefulness of unsupervised pre-training (feature learning), we also the trained another DBN without unsupervised training part, but only conjugate gradient BP was used. The overall classification accuracy for DBN with and without unsupervised pre-training was 90.27% and 75.30%, respectively, as can be seen from TABLE I. It can be clearly observed that unsupervised feature learning significantly improved the recognition performance.

To show a glimpse of recognition performance, a confusion matrix for only numerals is given in TABLE II. The largest confusion was observed between one (“ ১ ”) and nine (“ ৯ ”). Indeed, the digits look almost similar in the handwritten form. The average accuracy over the numerals was 91.30% which is slightly higher than the overall accuracy.

One of the prime features of the DBN approach is the ability of image reconstruction as it can work as a generative model. The better the reconstruction, the better the learning,

TABLE I: PERFORMANCE COMPARISONS OF DIFFERENT APPROACHES

| Method | Average classification accuracy |
|---|---|
| Only Supervised Learning (Conjugate gradient Backpropagation) | 75.30% |
| Unsupervised Learning + Supervised Fine Tuning | 90.27% |

and it helps attaining better classification accuracy. To illustrate, a few samples of reconstructed images, during the unsupervised feature learning stage, are shown in Figure 4. We can see that the model gradually learns a probability distribution of the visible data such that the sampled (reconstructed) images become meaningful (recognizable) and clearer. The reconstruction error rate curve in the training phase is shown in Figure 5.

We now compare the recognition performance of DBN with other methods, hierarchical learning architecture (HLA) schemes proposed by Bhowmik et al. [19] who reported their results on the same dataset we used in our experiments. The comparison is given in TABLE III. The comparison result reveals that the DBN approach can achieve higher recognition rate even though no handcrafted feature was used. Useful features in the DBN are rather learned during the training process.

TABLE III: PERFORMANCE COMPARISONS BETWEEN DBNS AND HLA FOR BANGLA CHARACTERS

| Scheme | Average test accuracy (%) |
|---|---|
| Using DBNs | 90.27 |
| HLA with disjoint groups (HLA_DG) | 84.78 ± 2.02 |
| HLA with overlapped groups (HLA_OG) | 88.02 ± 1.55 |
| HLA with groups using neural gas (NG) | 83.58 ± 1.98 |

TABLE II: CONFUSION MATRIX FOR BANGLA NUMERALS

| Target Digits \ Predicted Digits | ০ | ১ | ২ | ৩ | ৪ | ৫ | ৬ | ৭ | ৮ | ৯ |
|---|---|---|---|---|---|---|---|---|---|---|
| ০ | 95.7 | 0 | 0 | 1.25 | 0 | 1.30 | 0.05 | 1.65 | 0 | 0.05 |
| ১ | 0.20 | 94.3 | 1.36 | 0.5 | 0 | 0.05 | 0 | 0.25 | 0 | 3.34 |
| ২ | 0 | 1.72 | 89.65 | 1.63 | 0 | 0 | 0 | 0.15 | 0 | 0.20 |
| ৩ | 1.2 | 1.25 | 0 | 90.48 | 0 | 0.02 | 6.85 | 0 | 0 | 0.2 |
| ৪ | 0 | 0 | 0 | 0 | 93.75 | 0 | 0 | 4.2 | 1.52 | 0.53 |
| ৫ | 6.53 | 0 | 0.2 | 0 | 0 | 89.01 | 3.37 | 0.05 | 0.75 | 0 |
| ৬ | 0 | 0 | 0 | 7.5 | 0 | 1.75 | 89.99 | 0 | 0.5 | 0.26 |
| ৭ | 3.62 | 0 | 0 | 0 | 0 | 0.78 | 0 | 95.25 | 0 | 0.35 |
| ৮ | 0 | 0 | 0.05 | 0 | 6.5 | 0 | 2.2 | 0.05 | 90.35 | 0.85 |
| ৯ | 0 | 8.2 | 3.1 | 0 | 0.95 | 0 | 0.12 | 0 | 0.11 | 87.52 |

## V. CONCLUSION

In this paper, we investigated a feature learning based approach by the DBN for handwritten Bangla character recognition problem. We also used traditional supervised learning approach to show the effectiveness of deep learning. The main focus of this paper is to demonstrate the power of unsupervised feature extraction and learning. In this approach, there is no need to create handcrafted features like loops, stroke, and curves, yet the recognition result is satisfactory. The training and test dataset we used is not sufficient with respect to other researches in English language (for example, the MNIST dataset [20]). If it was possible to use larger dataset for training then performance would be better. As a future work, this work can be applied on Bangla compound characters and Bangla speech processing.

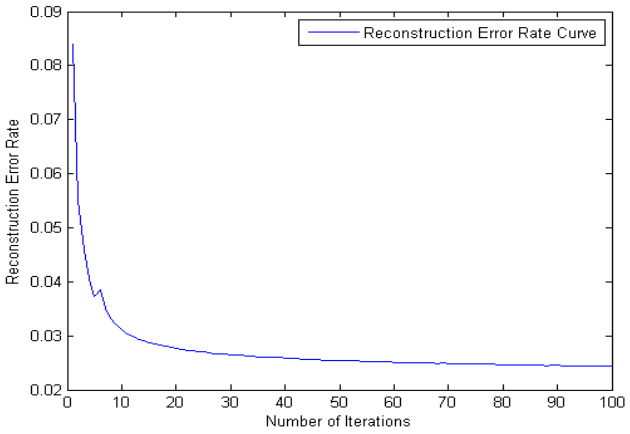


Figure 5. Reconstruction error rate curve in the training phase

## ACKNOWLEDGMENT

We would like to thank Ujjwal Bhattacharya from Indian Statistical Institute for providing us the *ISI handwritten character dataset.*

## REFERENCES

[1] Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh, "A fast learning algorithm for deep belief nets," *Neural computation*, vol. 18, no. 7, pp. 1527-1554, Jul. 2006.

[2] Ofer Matan, H.S. Baird, J. Bromley, Christopher J.C. Burges, J.S. Denker, L.D. Jackel, Y. Le Cun, E.P.D. Stenard, W.D. Pednault, C.E. Satterfield,T.J. Thompson, "Reading handwritten digits: A ZIP code recognition system," *Computer,* vol. 25, no.7, pp. 59-63, Jul. 1992.

[3] S. H.B Shaikh, Nabendu Chaki, Marek Tabedzki, and Khalid Saeed,"A view-based approach for recognition of Bengali printed characters," in Proc. *World Congress on Nature & Biologically Inspired Computing,* Dec. 2009, pp. 1005-1009.

[4] Fahim Irfan Alam and Bithi Banik, "Offline isolated bangla handwritten character recognition using spatial relationships," in Proc. *IEEE International Conference on Informatics, Electronics & Vision,* May 2013, pp. 1-6.

[5] G. G. Rajput, and H. B. Anita,"Handwritten Script Recognition Using DCT, Gabor Filter, and Wavelet Features at Word Level." in *Proc. International Conference on VLSI, Communication, Advanced Devices, Signals & Systems and Networking,* 2013, pp. 363-372.

[6] Rajat Raina, Alexis Battle, Honglak Lee, Benjamin Packer, and Andrew Y. Ng, "Self-taught learning: transfer learning from unlabeled data," in Proc. *ACM 24th international conference on Machine learning*, Jun. 2007, pp. 759-766.

[7] Adam Coates, Blake Carpenter, Carl Case, Sanjeev Satheesh, Bipin Suresh, Tao Wang, David J. Wu, and Andrew Y. Ng, "Text detection and character recognition in scene images with unsupervised feature learning," in Proc. *IEEE International Conference on Document Analysis and Recognition*, Sep. 2011, pp. 440-450.

[8] Tijmen Tieleman and Geoffrey Hinton,"Using fast weights to improve persistent contrastive divergence," in Proc. *ACM 26th Annual International Conference on Machine Learning*, Jun. 2009, pp. 1033-1040.

[9] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams, "Learning representations by back-propagating errors," *Nature,* vol.323, no. 6088, pp. 533-536, 1986.

[10] A.F.R. Rahman,R. Rahman,M.C. Fairhurst, "Recognition of Handwritten Bengali Characters: a Novel Multistage Approach," *Pattern Recognition*, vol. 35, no. 5, pp. 997-1006, May 2002.

[11] T. K. Bhowmik, U.Bhattacharya and Swapan K. Parui, "Recognition of Bangla Handwritten Characters Using an MLP Classifier Based on Stroke Features," *Neural Information Processing*, vol. 3316, p.p. 814-819, 2004.

[12] Geoffrey E. Hinton, and Ruslan R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science,* vol. 313, no. 5786, pp. 504-507, Jul. 2006.

[13] Aiquan Yuan, Gang Bai, Po Yang, Yanni Guo, and Xinting Zhao, "Handwritten English Word Recognition based on Convolutional Neural Networks," in Proc. *IEEE International Conference on Frontiers in Handwriting Recognition,* Sep. 2012, pp. 207-212.

[14] Yann LeCun,, Koray Kavukcuoglu, and Clément Farabet, "Convolutional networks and applications in vision," in Proc. *IEEE International Symposium on Circuits and Systems,* Jun. 2010, pp. 253-256.

[15] M. Welling, M. Rosen-Zvi, and G. E. Hinton, "Exponential family harmoniums with an application to information retrieval," in *Proc. Neural Information Processing Systems*, Dec. 2005, pp. 1481-1488.

[16] A-R.Mohamed, Tara N. Sainath, George Dahl, Bhuvana Ramabhadran, Geoffrey E. Hinton, and Michael A. Picheny, "Deep belief networks using discriminative features for phone recognition," in Proc. *IEEE International Conference on. Acoustics, Speech and Signal Processing*, May 2011, pp. 5060-5063.

[17] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural Computation*, vol. 14, no. 8, pp. 1771–1800, Aug. 2002.

[18] URL: http://www.isical.ac.in/~ujjwal/download/database.html

[19] Tapan Kumar Bhowmik, Pradip Ghanty, Anandarup Roy, and Swapan Kumar Parui, "SVM-based hierarchical architectures for handwritten Bangla character recognition." *International Journal on Document Analysis and Recognition (IJDAR),* vol. 12, no. 2, pp. 97-108, Jul. 2009.

[20] URL: http://yann.lecun.com/exdb/mnist/

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# BER Analysis of Optical Wireless Communication System Employing Neuro-Fuzzy based Spot-Diffusing Techniques

Shamim Al Mamun[1*], M. Shamim Kaiser[2], Muhammad R. Ahmed[3], and Md. Imdadul Islam[4]
[1,2]Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh
[3]Faculty of Information. Sciences and Engineering, University of Canberra, Australia
[4]Dept of Computer Science and Engg, Jahangirnagar University, Dhaka, Bangladesh
*Email:shamim@juniv.edu

***Abstract*—This paper presents financial feasibility analysis of a micro-controller based solar powered tricycle or rickshaw for Bangladesh. The main components of the Rickshaw are: rickshaw structure, a photovoltaic array, a battery bank, a battery charge controller, PIC micro-controllers and a DC motor. Here PIC micro-controllers are used to control the duty cycle and speed of the DC motor. The financial feasibility analysis have been done using standard methods of engineering econometrics such as Net Present Worth (NPW), Benefit-to-Cost Ratio (BCR) and Payback Periods (PP) methods. From the numerical analysis, it has been found that the proposed micro-controller based solar power rickshaw is financially feasible and Life cycle unit cost of the system is lowest compared to grid powered battery driven auto rickshaw.**

***Keywords*—Solar energy, battery operated rickshaw, economic analysis, microcontroller.***

## I. INTRODUCTION

With the increasing development of ultra broadband wireless application the radio frequency (RF) spectrum became congested and scare resources. Thus optical wireless communication (OWC) has drawn considerable attention to the researchers. OWC for indoor application was first proposed by Gfeller and Bapst [1]. In number of applications where higher data throughputs is more of requirement than the mobility, transmission link based on optical wireless would be one of the best options as outlined in [2-5]. The performance of OW systems depends on the propagation and type of system used. The basic system types fall into diffuse or line of sight (LOS) systems. In LOS systems, high data rates in the order of Gbit/s can be achieved [6], but the system is vulnerable to blockage/shadowing because of its directionality. In a diffuse OW system, several paths from source to receiver exist, which makes the system robust to blockage/shadwoing. However, the path losses are high and multipaths create inter-symbol interference (ISI) which limits the achievable data rate [7, 8]. There are several advantages of OWC over traditional RF systems, these are: an abundant free spectrum, extremely high communication speed is possible by all network, does not interfere with the over congested RF spectrum. But limitations are: a beam is short ranged, may be harmful for eye. The first limitation can be overcome by wavelength reuse technique, where as eye safety can be ensured by maximum transmit power. It can be classified as line of sight (LOS) and diffuse system. Many researchers have considered diffuse systems for indoor applications it offers robust link and overcomes the problem of shadowing [7], does not require transmitter-receiver alignment and uses the wall or ceiling for multi path reflection [8]. The multipath reflections increases delay spread or inter-symbol interference. Ambient light such as florescent, incandescent light and Compact Florescent Lamp (CFLs) produces channel noise which reduces signal-to-noise plus interference ratio (SNIR). In order to improve the system performance several spot diffusion configuration using multi beam transmitter have been proposed [9]. Multi beam transmitter is place in center of the room and pointed upward. A multi-spot pattern have been generated, illuminated multiple small areas in the ceiling and then reflected multiple spot have been received by receivers. In [10], to improve the bandwidth, reduce the effect of inter-symbol interference, and increase the signal to noise ratio (SNR) when the transmitter operates at a higher data rate under the impact of multipath dispersion, background noise, and mobility in conjunction with an imaging receiver. It proposed different line streaming multi-beam spot diffusion (LSMS) model to gain about 32.3 dB SNR at worst communication path. But the multi path dispersion reduces the performances due to transmitter power and can be improve using power adaptive system by [9]. User mobility is very important aspect of wireless communication especially with today's hand hold devices. As the user device can mobilize with the room then power adaptation will be a great solution to get higher SNR. In this aspect [12] propose a genetic algorithm for multi spot diffuse system in indoor wireless communication. But it is noted from different research that if the diffuse system has a predefined spot for a room and use an adaptive power allocation for beam using calculation of delay spread then it can improve the performance of the OWC. Neural network and Adaptive Linear Equalizers can be a solution in this case for adaptive power distribution. [12] Presents a comparative study of two equalizers, the adaptive linear and the neural equalizer for indoor optical wireless (OW) links using OOK modulation technique to reduce ISI effect.

This paper introduce adaptive neuro-fuzzy interference system (ANFIS) for selecting spot beam matrix for a room and also distribute power allocation by calculating delay spread in considering Doppler shift effect of the mobile devices.

The paper is organized as follows: the system model is presented in section II; power allocation algorithm is explained in section III; Section IV presents discussion and results. The

concluding remarks and future work is included in section V.

## II. Proposed System Model

Consider an empty room with floor dimensions of $8 \times 4\ m^2$ and ceiling height of $3m$ as shown in Figure 1. The reflection coefficient of the ceiling is considered to be $0.8$. Their are eight spot lights on the ceiling. In the Figure, $\delta$ is the elevation angle, $\alpha$ is the azimuth angle, $d = 8$, $w = 4$ and $h = 3$, $x_0$ and $x$ are the position of the imaging receiver and $v$ is the velocity. Neuro-Fuzzy (NF) adaptive multibeam transmitter is located at the center of the room whereas a imaging receiver is placed at $x_0 = (1, 1, 0.5)$. The transmitter generates multi spot beam matrix on the ceiling where beam power and beam angle $(\alpha, \delta)$ are adapted and the reflected beams are received by the imaging receiver. The transmitter learns receiver position, mobility through the low rate diffuse channel. At low data rate, the beam maintains the fixed power.

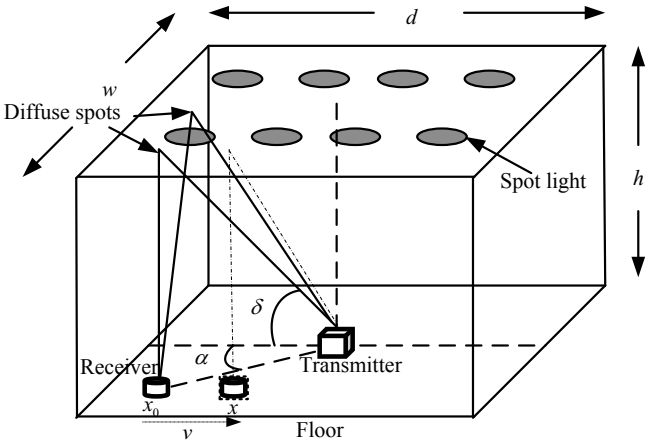



Fig. 1. System Model for OWC based on spot-diffusing technique. Here $\delta$ is the elevation angle, $\alpha$ is the azimuth angle, $d = 8$, $w = 4$ and $h = 3$.

### A. Signal to Noise Plus Interference Ratio

In indoor optical-wireless communication, the ambient light affects signal-to-noise-plus interference (SNIR) at the receiver. Many researchers have considered intensity modulation with direct detection (IM/DD) as most viable approximation. The received signal, denoted by $y(t)$, can be expressed as

$$y(t) = \sum Rx(t) * h(t, \alpha, \delta) + \sum n(t, \alpha, \delta) + \sum I(t, \alpha, \delta) \tag{1}$$

where $R$ is the receiver responsivity , $x(t)$ is the instantaneous optical transmitted power, $h(t, \alpha, \delta)$ is the impulse response of the OW channel, $n(t, \alpha, \delta)$ is the ambient light noise, $I(t, \alpha, \delta)$ is the instantaneous interference power.

The SNIR, denoted by $\gamma$, of the received signal can be calculated by [9]

$$\gamma = \frac{R^2(P_{s1} - P_{s0})h^2}{(\sigma_{s1} - \sigma_{s0})^2} \tag{2}$$

where $P_{s1}$ and $P_{s0}$ are the optical power associated with the binary 1 and binary 0 respectively, $\sigma_{s1}$ are $\sigma_{s0}$ are the shot noise variation component with $P_{s1}$ and $P_{s0}$ respectively.

### B. Adaptive Power Allocation

The achievable data transmission rate, denoted by $b$, of the OWC system is given by

$$b = \frac{1}{M} \sum_{i=1}^{M} \log_2 \left(1 + \frac{R^2 \times (P_{s_{1i}} - P_{s_{0i}})h_i^2}{(\sigma_{1i} - \sigma_{0i})^2}\right) \tag{3}$$

The optimization problem and constraint of the power allocation can be written as

$$\max \quad b \tag{4}$$

$$s.t. \quad \sum_{j=1}^{J} P_j \leq \bar{P} \tag{5}$$

where $\bar{P}$ is the average power. We can use the Lagrange multiplier method to analyze the above optimization problem and the Lagrangian function is defined as

$$L = b + \mu_j(\sum_{j=1}^{J} P_j - \bar{P}) \tag{6}$$

where $\mu_j$ is the Lagrange multiplier. After solving the Eqn. (6),we can write

$$P_j = \left[\frac{\left(P + \sum_{j=1}^{J} \frac{1}{h_i}\right)}{C} - \frac{1}{h_i}\right] \tag{7}$$

$$= \max\left[\lambda(C) - \frac{1}{h_i}, 0\right] \tag{8}$$

### C. Delay Spread

The delay spread of an impulse is expressed as rms value by,

$$D = \sqrt{\frac{\sum (t_i - \mu)^2 P_r^2}{P_r^2}} \tag{9}$$

where $\mu = \frac{t_i P_r^2}{P_r^2}$ and $t_i$ is the delay time and $P_r$ is the received power

### D. Doppler Shift

Light waves require no medium and being able to travel even through vacuum. Let $\bar{v}$ is the relative velocity between transmitter and receiver, the proper frequency of the transmitted information signal from the optical transmitter is $f_0$ . Let $f$ is the frequency of the received signal accepted by the moving receiver with a velocity $\bar{v}$ , then

$$f = f_0 \times \sqrt{\frac{1 \pm \beta}{1 \mp \beta}} \tag{10}$$

where $\beta = \bar{v}/c$, $c$ is the speed of light. For low speed, i.e., $\beta << 1$, and in this case the above eqn. (10) is reduced to

$$f = f_0(1 - \beta)^{-\frac{1}{2}}$$

$$= f_0(1 + \beta + \frac{1}{2}\beta^2) \tag{11}$$

### E. ANFIS Model

Neuro-fuzzy inference system is consider if learning capabilities are required. In this paper, we consider the adaptive neuro-fuzzy inference system (ANFIS) for the implementation of the spot beam matrix selection as shown in Fig. 2. Based on the signal to noise ratio , i.e., $\gamma$, and link delay, i.e., $\Delta\tau$, ANFIS decides a spot is eligible for selection or not. The ANFIS is trained iteratively to achieve the desired output for the input parameters and their membership functions. This can be done by back propagation gradient descendent which evaluates the error signals recursively from the output layer backward to the input nodes. In this way, ANFIS learns the behavior of the system. Sugeno ANFIS model contains if and then rules, e.g., If $x$ is $A_i$ and $y$ is $B_i$ then $f_i = p_i x + q_i y + r_i$, where $\{p_i, q_i, r_i\}$ is the consequent parameter set. Fig. 1 shows ANFIS model for spot beam matrix selection. It consists of five layer: input layer, output layer and three hidden layers.

Each adaptive node in the input layer generates membership grades. If bell shape membership functions are considered, output of this node, denoted by $O_i^1$, can be written as

$$O_i^1 = \mu_{\alpha_i}(x) = \frac{1}{\left[1 + \left|\frac{x-c_i}{a_i}\right|^{2b_i}\right]}$$

where $\alpha_i \in \{A_i, B_i\}$ is the input vector, $\{a_i, b_i, c_i\}$ are the premise parameters.

Nodes in the first hidden layer calculates the firing strength of a rule via multiplication. The output of the each node,denoted by $O_i^2$, can be written as

$$O_i^2 = w_i = \mu_{A_i}.\mu_{B_i}(x)$$

where $i = 1, 2, 3$. ANFIS performs AND operation in this layer.

Nodes in the second hidden layer computes the normalized value of the firing strength. The output of the each node,denoted by $O_i^3$, can be written as

$$O_i^3 = \bar{w_i} = w_i / \Sigma w_i$$

Nodes in the third hidden layer computes the contribution of $i$-th rule towards the overall output. . The output of the each node,denoted by $O_i^3$, can be written as

$$O_i^4 = \bar{w_i} f_i = \bar{w_i}(p_i x + q_i y + r_i)$$

A signal node in the output layer computes the overall output, denoted by $O_i^5$ as follows:

$$O_i^{5`} = \Sigma_i \bar{w_i} f_i$$

.

## III. Adaptive Spot-Beam Selection Algorithm

Fig. 3 shows the block diagram of the adaptive spot-beam selection algorithm. In the first step the beam hologram or matrix generates $40 \times 20$ equal powered spot-beams in the ceiling. The SNIR and delay spread for each beam have been calculated by the image receiver. The receiver periodically evaluates the SNIR after 1 second interval whereas the delay spread for each beam is same if the receive is not moving. In the second step, the receiver sends the spot-beam information which contains SNIR and delay spread to the transmitter. Based on the minimum SNIR and maximum delay spread, transmitter select the spot-beam matrix by NF based algorithm in the third step. The transmitter allocates the power for each selected beam adaptively using eqn. (8) in the forth step. Finally based on the velocity of movement of the receiver, transmitter moves spot-beam matrix for the receiver.

The algorithm is summarized as follows:

The following algorithm will find the spot beam with an equal power allocation over $40 \times 20$ beam hologram or matrix, $H$.

- Step 1 A spot beam scans the ceiling, SNIR, $\gamma$ and delay spread, $\Delta\tau$ for each beam have been calculated by the image receiver using Eqns (2) and (9).
- Step 2 Based on the required minimum SNIR, i.e., $\gamma_{min}$, and maximum delay spread, i.e., $\Delta\tau_{max}$, transmitter selects the spot-beam matrix ($H$) by NF controller.
- Step 3 The transmitter allocates the power for each selected beam adaptively using Eqn (7)
- Step 4 Based on Doppler shift, the transmitter adapts the beam angles $\alpha$ and $\delta$.
- Step 5 Multi-spot optical transmitter further reduce the $\Delta\tau$ by scheduling.
- Step 6 Finally, Multi-spot optical transmitter transmits the spot beam matrix to receiver via ceiling.
- Step 7 Go to Step 1 if transmitter gets receiver's position update.

## IV. Numerical Analysis

In this section, Neuro-Fuzzy based multibeam system (NFMS) is investigated with diversity receiver configuration. It is compared with other spot-beam diffusion method. The ANFIS model, adaptive power allocation and multi-spot diffuse pattern formation are implemented in MATLAB/SIMULINK. ANFIS consider two input such as SNR and delay.

Simulation parameters considered for the analysis are: length, width and height of the room are $8m$, $4m$ and 3 m; the reflection coefficient of the ceiling is $\rho = 0.8$; there is one transmitter which is located at $(2, 4, 1)$ location; there is also one receiver; the area, acceptance semi-angle of the each photo-diode are $2cm^2$ and $65^0$ respectively. The number of pixel at the receiver is 200 (with area of $0.01cm^2$ ) Pedestrians move typically at the speed of 1 m/s. If the SNIR is computed after 10 $\mu s$; there are 8 spot lamp in the room which are located at $(1, 1, 1)$, $(1, 3, 1)$, $(1, 5, 1)$,$(1, 7, 1)$,$(3, 1, 1)$, $(3, 3, 1)$, $(3, 5, 1)$, and $(3, 7, 1)$; and the wavelength of the light is $850nm$

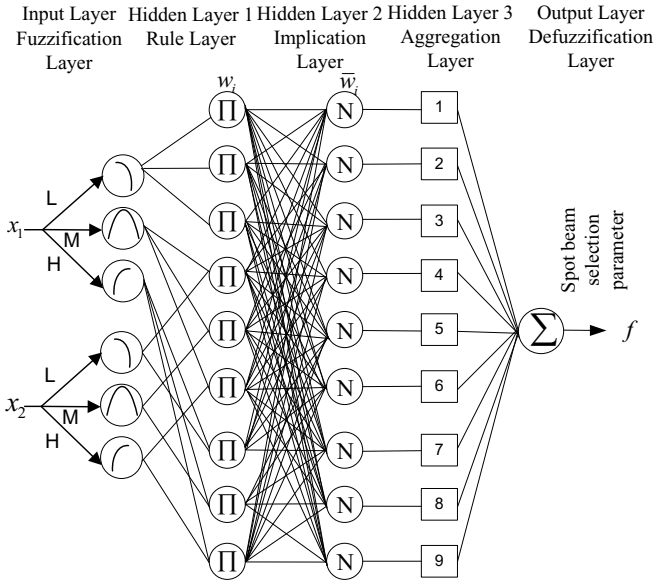



Fig. 2. ANFIS model for Spot beam selection

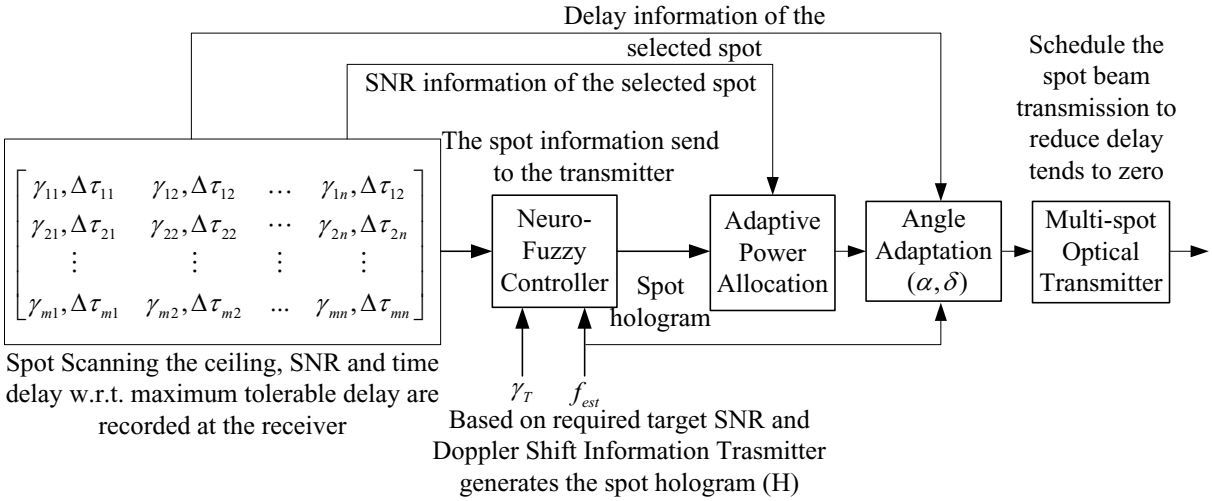



Fig. 3. Spot beam selection algorithm in the presence of Doppler shift due to receiver movement

The 80 ms adaptation time will give overhead of $8\%$. Adaptation time depends on environment. Receiver computes the SNIR and delay spread and sends these information via a low rate channel to the transmitter.

ANFIS consider two inputs. Iterative training of the ANFIS has been done to achieve the desired output. After a predefined simulation time to obtain the simulation result and use them to train. Based on the training data set, ANFIS

Fig. 4 shows the effect of receiver position on the SNIR for proposed model, line strip multi-spot diffuse system (LSMS) and conventional diffuse system. The SNR calculations were performed for the receiver is moving towards the transmitter (i.e., the values of $x$ is increasing) while neglecting the movement along $y-$axis. Significant SNIR improvement of almost 3dB is observed as the NFC moves the spot beam, selects the best positioned spot only, and allocate the power adaptively based on the channel condition of the selected slots. It is also found that the SNIR performances have been degraded as the receiver is moving. This degradation in SNIR increases as the velocity of the receiver increases.

Fig 5 shows the SNIR for proposed model has been improved further (almost 1dB), if we change the slot beam angle of the selected slot.

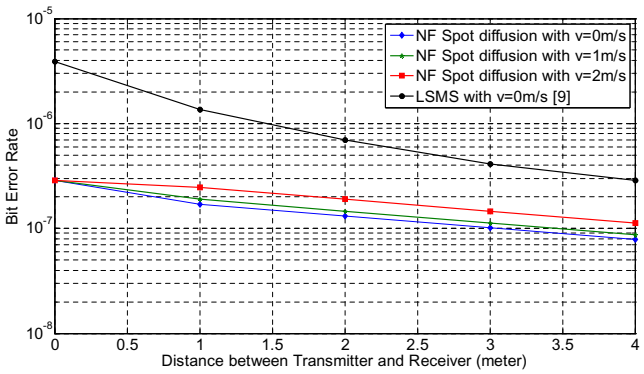



Fig. 4. Effect of receiver position on BER for proposed model and LSMS

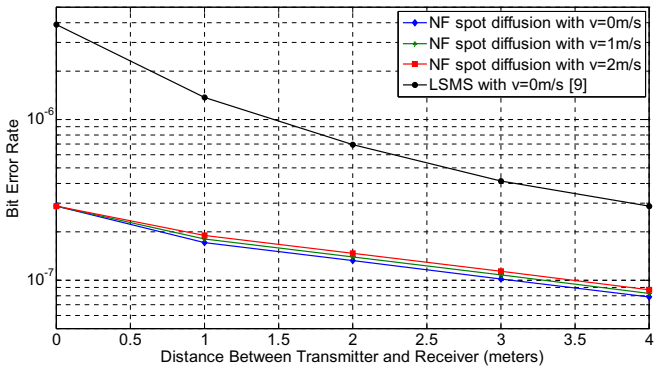



Fig. 5. Effect of receiver position on BER for proposed model and LSMS. Here the beam pattern is shifted with Doppler shift

## V. Conclusion

In this paper, we have proposed a new method of real-time beam and angle adaptation technique for optical wireless communication system using ANFIS. This NF controller has five layers and is trained with back-propagation gradient decent algorithm. The controller is trained with data obtained by simulations. Simulation results show that the proposed NF based OW spot-diffusing communication system outperforms other spot-beam diffusion method in terms of SNIR and delay spread.

## References

[1] M. S. Kaiser, K. M. Ahmed and R. A. Shah, *Power allocation in OFDM-based cognitive relay networks*, IEEE International Conference on Wireless Communications, Networking and Information Security (WCNIS), 2010, Beijing, China, pp. 202-206, June 2010.

[2] K. Son, B. C. Jung and D. K. Sung, "Opportunistic Underlay Transmission in Multi-carrier Cognitive Radio Systems," IEEE WCNC 2009, 2009.

[3] F.R. Gfeller and U. H. Bapst, "Wireless in-house data communication via diffuse infrared radiation," Proc. IEEE, vol. 67, no. 11, pp. 1474-1486. Nov. 1979.

[4] T. Komine, M. Nakagawa, "A Study of Shadowing on Indoor Visible-Light Wireless Communication Utilizing Plural White LED Lightings," Int. Sympo. On Wireless Commun., pp. 36-40, 2004.

[5] M. Z. Afgani, H. Haas, H. Elgala, D. Knipp, "Visible light communication using OFDM," Proc. IEEE Symp. on Wireless Pervasive Computing, TRIDENTCOM 2006.

[6] T. Komine, M. Nakagawa, "Fundamental Analysis for Visible- Light Communication System using LED Lights," IEEE Trans. on Consumer Electron., vol. 50, no., pp. 100-107, 2004.

[7] M. Akbulut, C. Chen, M. Hargis, A.Weiner, M. Melloch, and J.Woodall, "Digital Communications Above 1 Gb/s Using 890-nm Surface-Emitting Light-Emitting Diodes," IEEE Photonics Technology Letters, vol. 13, no. 1, pp. 85-87, Jan. 2001.

[8] Fuad E. Alsaadi, Mohammad Nikkar, and Jaafar M. H. Elmirghani, "Adaptive Mobile Optical Wireless Systems Employing a Beam Clustering Method,Diversity Detection, and Relay Nodes," IEEE Transactions On Communications, Vol. 58, No. 3, March 2010.

[9] Mohammad T. alresheedi and Jaafar M. H. Elmirghani,"10 Gb/s Indoor Optical wireless Systems Employing Beam Delay, Power, and Angle Adaptation Methods with Imaging Dectection," Journal of Lightwave Technology, vol. 30, n0. 12, june 15, 2012.

[10] M. Nikker Esfahani and Jaafar M. H. Elmirghani, "A genetic algorithm Method for Multi spot diffuse Infrared Wireless Communication," London Communication Symposium, 2009.

[11] Z. Ghassemlooy and Sujan Rajbhandari, "Performance of Diffused Indoor Optical Wireless Links Employing Neural and Adaptive Linear Equalizers," ICICS 2007.

[12] D. K. Borah1, A. C. Boucouvalas, C. C. Davis, S. Hranilovic and K. Yiannopoulos, "A review of communication-oriented optical wireless systems," EURASIP Journal on Wireless Communications and Networking 2012, 2012 : 91 doi: $10.1186/1687-1499-2012-91$.

[13] A. G. Al-Ghamdi and J. M. H. Elmirgrani, "Analysis of Diffuse Optical Wireless Channels Employing Spot-Diffusing Techniques, Diversity Receivers and Combining Schemes," IEEE Transaction on Communications, Vol. 52 (10), pp: 1622–1631, 2004.

# An Enhanced Model of Vertical Handoff Decision Based on Fuzzy Control Theory & User preference

Snigdha Khanum [1*], and Mohammad Mahfuzul Islam [2]
[1]Stamford University Bangladesh
[2]Department of CSE, BUET, Dhaka-1000, Bangladesh
[*] snigdha.khanum@yahoo.com

*Abstract*—**With the development of wireless communication technology, various wireless networks will exist with different features in same premises. Heterogeneous networks will be dominant in the next generation wireless networks. In such networks choose the most suitable network for mobile user is one of the key issues. Vertical handoff decision making is one of the most important topics in wireless heterogeneous networks architecture. Here the most significant parameters are considered in vertical handoff decision. The proposed method considered Received signal strength (RSS), Monetary Cost(C), Bandwidth (BW), Battery consumption (BC), Security (S) and Reliability (R). Handoff decision making is divided in two sections. First section calculates system obtained value (SOV) considering RSS, C, BW and BC. SOV is calculated using fuzzy logic theory. Today's mobile user are very intelligent in deciding there desired type of services. User preferred network is choose from user priority list is called User obtained value (UOV). Then handoff decisions are made based on SOV & UOV to select the most appropriate network for the mobile nodes (MNs). Simulation results show that fuzzy control theory & user preference based vertical handoff decision algorithm (VHDA) is able to make accurate handoff decisions, reduce unnecessary handoffs decrease handoff calculation time and decrease the probability of call blocking and dropping.**

***Keywords— Vertical handoff, Heterogeneous networks, Fuzzy control theory, Received signal strength, Cost, Bandwidth, Security and user preference.***

## I. INTRODUCTION

Wireless communication demand is increasing day by day due to globalization. Due to improvement of this communication technology wireless network becomes heterogeneous and making available a wide range of new applications. That is the reason why the development of mobile terminal designs continues to grow and changes our living standards [1]. Wireless network technologies are different from each other usually in terms of bandwidths, frequencies, latencies, Signal strength, Cost and etc. The service of Wireless communication networks is upgrading extremely fast with the development of wireless communication technology. Presently, there are many kinds of wireless networks available to satisfy different needs and requirements of mobile users. But no single technology simultaneously provides a low latency, high bandwidth, and cost-effective services to all mobile users. In these wireless communications system, multiple wireless networks will coexist to offer multiple services anywhere at any time [2] [3]. The Coexist wireless communication networks, which consist of an overlaid structure of different sizes and features, can be termed as Wireless Overlay Networks [4]. When users are roaming among various wireless networks, such as Wireless Local Area Network (WLAN) and Universal Mobile Telecommunications System (UMTS) or stay in an overlaid network, the interconnection of these different networks has become a problem. While a mobile terminal (MT) crosses the coverage boundary of two different systems, its ongoing connection must be seamlessly switched to a new network with a guaranteed quality-of-service (QoS). Such a cross-system transfer of an ongoing connection is usually referred to as inter-system, or vertical handover [5].

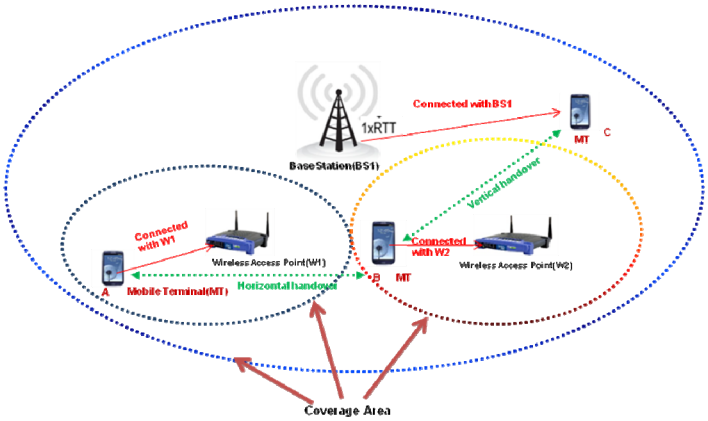



Figure 1: Vertical handoff in heterogeneous network.

The vertical handoff decision technique presented in this paper is based on fuzzy control theory and user preference. Vertical handoff decision technique divided into two parts. One part is SOV and another part is UOV. It takes into account four parameters and use fuzzy control theory to calculate SOV. With the advancement of wireless communication technology users become more familiar with different features of communication systems. Today's users know all about security, signal strength, cost and other common features of wireless communication. So user can easily decide which parameters to give priority to choose a wireless network. According to user priority list UOV is calculated. Finally from SOV and UOV value final decision value (FDV) for each network is calculated. The vertical

handoff decision technique presented in this thesis is very simple and can be easily implemented.

## II. Related Previous Works

The heterogeneous networks become a main focus in the development toward the next generation wireless networks. Vertical handoff is very significant in heterogeneous networks for seamless communication. In heterogeneous environments, it is critical to design an efficient handoff algorithm for supporting seamless communication service. Many wireless communication researchers working in different way for efficient vertical handoff technology. Many researchers apply fuzzy logic into the handoff decision because of the inherent strength of fuzzy logic to process all available context information which has different types of values. In [6] authors propose a handoff decision based on fuzzy logic principle to decrease unnecessary handoff in the homogeneous cellular system. Fuzzy sets and enumerative fuzzy handoff rule base are established. But it is only considered the horizontal handoffs not applying for the vertical ones. A novel fuzzy logic-based handoff decision algorithm for the mobile subsystem of tactical communications systems is introduced in [7]. In [8], an adaptive fuzzy based vertical handoff algorithm is presented to decide the RSS hysteresis values using user speed and traffic load. It only made handoff decision by RSS, while this is insufficient for vertical handoffs decision. In [9], a fuzzy based vertical handoff decision algorithm is aimed between Global System of Mobile (GSM), General packet radio service (GPRS) and WLAN networks with bandwidth, coverage area, power consumption and sojourn time parameters. Another fuzzy based vertical handoff decision algorithm with Elman neural network is proposed between WLAN and UMTS [10]. Authors in [11] provide a multi-criteria decision-making algorithm based on fuzzy theory for multiple access network selection. But it only gives the theoretic descriptions for mobility management without the performance evaluation for this strategy. [12] Proposes a vertical handoff algorithm that considers some network parameters including signal strength, cost and unused bandwidth. The handoff decision is based on the weight vector for each of the input parameter and membership function of each parameter.

## III. System Model & Methodology

Here two different types of wireless network considered for vertical handoff simulation one is UMTS and another is wireless wide area network (WWAN). The UMTS is a third generation mobile cellular system for networks based on the GSM communication standard which is developed and maintained by the 3rd Generation Partnership Project (3GPP). UMTS deliver broadband information at speeds up to 2 Mbits/sec. Besides voice and data, UMTS will deliver audio and video to wireless devices anywhere in the world through fixed, wireless and satellite systems. The larger size of a wide area network compared to a local area network requires differences in technology. Wireless networks of all sizes deliver data in the form of telephone calls, web pages, and streaming video. WWAN connectivity allows a user with a laptop and a WWAN card to surf the web, check email, or connect to a virtual private network (VPN) from anywhere within the regional boundaries of cellular service. Since radio communications systems do not provide a physically secure connection path, WWANs typically incorporate encryption and authentication methods to make them more secure. Wireless WAN allows locations to connect to the corporate infrastructure across AT&T Mobility's Broadband Connect network by using a cellular-enabled router and security enhanced dedicated or Internet connection to data centre or headquarters location.

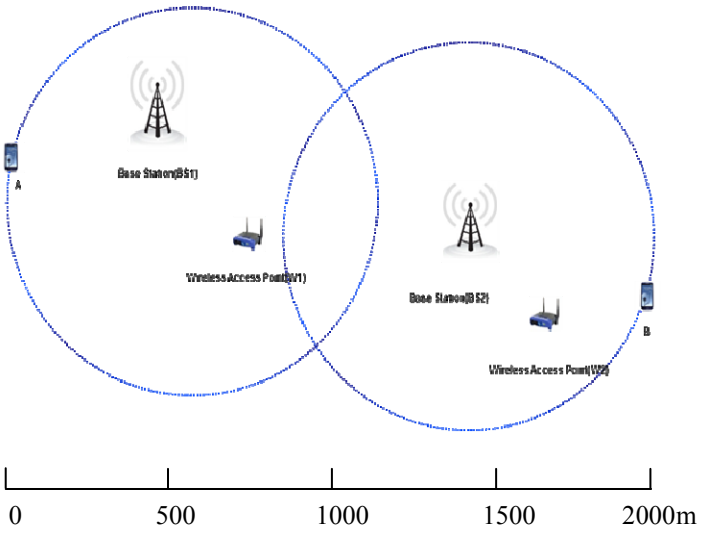



Figure 2: Simulation scenario showing heterogeneous network & MT.

The base station (BS1), base station (BS2) represent UMTS network and circle shows its coverage area. Two WWAN wireless access point (AP) (W1&W2) covers some portions of the UMTS service area. A MT moves from position A to position B in a straight line and the distance from point A to B is 2000 meter. The distance is plot in X axis and the changes of others parameter is presented in Y axis considering based on depicted simulation scenario. The vertical handoff decision will trigger when the MN detects a new wireless link.

Here system input parameters RSS, C, BW, BC. Then membership functions of the input parameters are determined. After establishing the membership functions, the membership degrees of the input parameters for each candidate BS or AP can be calculated. To reduce the handoff decision time and handoff failure dynamic weight vector is used in proposed system. From membership degree weight vector SOV of each BS or AP is achieved. User priority list is also used. In user priority list parameters are S, C, R & BC. From list user can give priority any of the parameter. Normalized value of priority parameter and priority vector calculates UOV. From SOV & UOV for each BS or AP Final decision Value (FDV) is achieved. The FDV Value is used to evaluate a certain BS or AP.

## IV. Membership Functions of Input Parameters for SOV & Dynamic Weight Vector

There are many parameters in wireless communication technology. Conventionally RSS from candidate BSs are take into account for vertical handoff decision. In this paper considered RSS and also the factors of C, BW and BC to decrease unnecessary handoff. Different significant parameters can realize the interconnection of different networks.

### A. Membership function of RSS

RSS the actual received signal strength that is received from the candidate AP and BS, and define $RSS_{TH}$ to be the RSS threshold. $RSS_{max}$ is the maximum value of RSS that can be received from a candidate AP and BS. By using the normalization factor of RSS and $RSS_{max}$, we can get a piecewise linear membership function of RSS denoted by $\mu_{RSS}$ where, $x$ denotes the location of users (The variable $x$ in the following discussion has the same meaning).

$$\mu_{RSS} = \begin{cases} 0 & 0 \le RSS(x) \le RSS_{TH} \\ \frac{RSS(x)-RSS_{TH}}{RSS_{max}-RSS_{TH}} & RSS(x) > RSS_{TH} \end{cases} \quad (1)$$

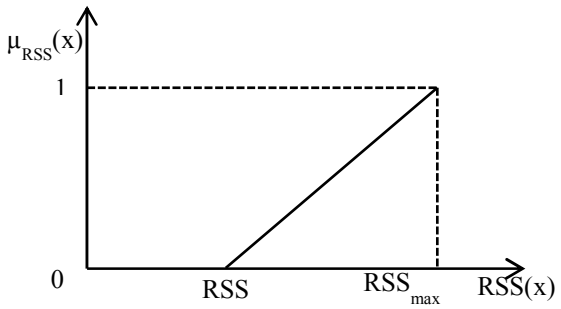



Figure 3: Membership Function for RSS.

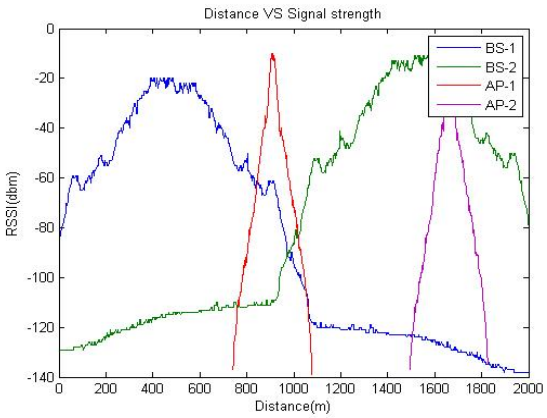



Figure 4: RSS Membership function changes with distance.

The figure represents membership functions of proposed system's network parameter RSS.

### B. Membership function of C

C(x) is the actual cost of the operation network which a candidate AP and BS belongs to. And $C_{TH}$ is the threshold cost. That is, when $C(x) \ge C_{TH}$, the user will consider that the operation network is too expensive to accept. The membership function of C which is denoted as $\mu_C$ is shown below:

$$\mu_C = \begin{cases} 1 - \frac{C(x)}{C_{TH}} & 0 \le C(x) \le C_{TH} \\ 0 & C(x) > C_{TH} \end{cases} \quad (2)$$

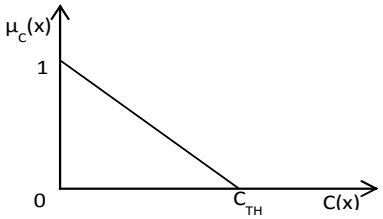



Figure 5: Membership Function for C.

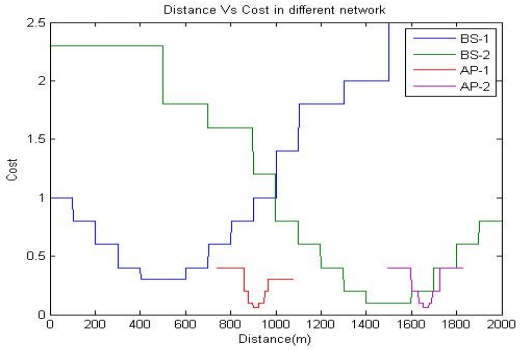



Figure 6: C Membership function changes with distance.

### C. Membership function of BW

Available bandwidth is also a criterion for the handoff. In bandwidth membership function BW(x) is defined as the amount of bandwidth available under a candidate BS or in an AP and $BW_{max}$ to be the maximum amount of bandwidth that a candidate BS or an AP can provide. The membership function of BW which is denoted as $\mu_{BW}$ is shown below:

$$\mu_{BW} = \begin{cases} \frac{BW(x)}{BW_{max}} & 0 \le BW(x) \le BW_{max} \\ 0 & BW(x) > BW_{max} \end{cases} \quad (3)$$

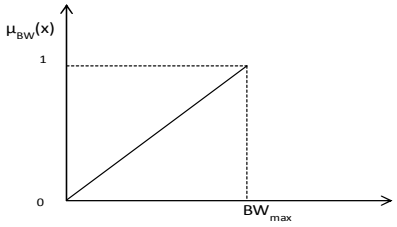



Figure 7: Membership Function for BW.

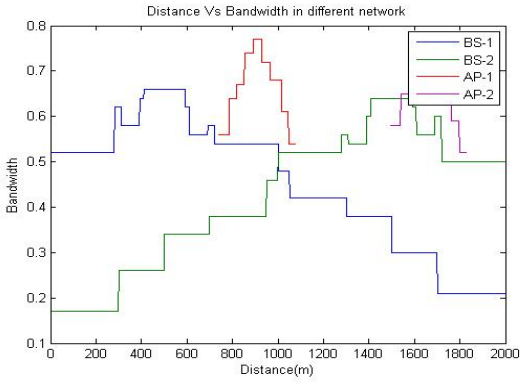



Figure 8: BW Membership function changes with distance.

### D. *Membership function of BC*

When mobile terminal's battery is low the power consumption becomes critical issue. Considering the significance of battery consumption it's also included in handoff decision. In battery consumption membership function BC(x) is defined as battery consumption in candidate AP and BS and $BC_{TH}$ is threshold value of battery consumption. The membership function of BC which is denoted as $\mu_{BC}$ is shown below:

$$\mu_{BC} = \begin{cases} 1 - \frac{BC(x)}{BC_{TH}} & 0 \le BC(x) \le BC_{TH} \\ 0 & BC(x) > BC_{TH} \end{cases} \quad (4)$$

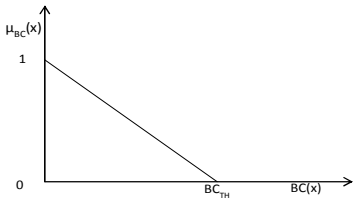



Figure 9: Membership Function for BW.

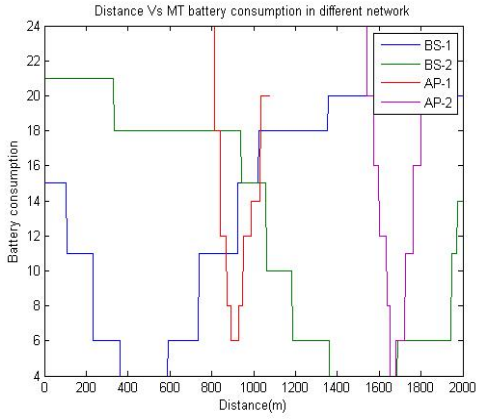



Figure 10: BC Membership function changes with distance.

### E. *Membership Degree*

Using the membership functions in (1), (2), (3) and (4), we can determine the membership degrees of RSS, C, BW and BC for each BS. The results are shown in Table 1. For the membership degree μ in the table, the first subscript denotes the input parameter, and the second subscript denotes the BS or AP.

TABLE I. MEMBERSHIP DEGREES FOR EACH BS OR AP

| | $BS_1$ | $AP_1$ | … | $BS_n$ |
|---|---|---|---|---|
| RSS | $\mu_{1,1}(x)$ | $\mu_{1,2}(x)$ | … | $\mu_{1,n}(x)$ |
| C | $\mu_{2,1}(x)$ | $\mu_{2,2}(x)$ | … | $\mu_{2,n}(x)$ |
| BW | $\mu_{3,1}(x)$ | $\mu_{3,2}(x)$ | … | $\mu_{3,n}(x)$ |
| BC | $\mu_{4,1}(x)$ | $\mu_{4,2}(x)$ | … | $\mu_{4,n}(x)$ |
| F | $F_1(x)$ | $F_2(x)$ | … | $F_n(x)$ |

Using the membership functions the membership degrees of Bandwidth, Received Signal Strength, Cost and Battery Consumption for each BS & AP is determined.

### F. *Dynamic weight vector*

Under various occasions, the values of Bandwidth, Received Signal Strength, Cost and Battery Consumption and their membership degrees are continuously changing in an unpredictable way. In order to achieve the optimized vertical handoff under different occasions, the Dynamic Weight Vector DW has the features that it dynamic, well-adjusted to the unpredictable occasions, it can always properly reflect the relationships and weights of the continuously changing membership degrees; it can 'magnify' the dominant-difference of certain membership degree(s) among candidate access points. For example, assume that the dominant difference among candidate access point AP1, AP2 is Battery Consumption, while their cost, C and Bandwidth, BW are very close or even the same. Then, the weight of $\mu_{1,k}(x)$ (for k = 1, 2, . . , n) , $DW_1$ should be the largest. So, the Dynamic Weight Vector DW can be defined as follows:

$$DW_i = \frac{\sigma_i}{\sum_{i=1}^{4} \sigma_i}$$

or,

$$DW = \left( \frac{\sigma_1}{\sum_{i=1}^{4} \sigma_i}, \frac{\sigma_2}{\sum_{i=1}^{4} \sigma_i}, \frac{\sigma_3}{\sum_{i=1}^{4} \sigma_i}, \frac{\sigma_4}{\sum_{i=1}^{4} \sigma} \right)$$

$$= (DW_1, DW_2, DW_3, DW_4) \quad (5)$$

Where $\sigma_i$ is the standard deviation of $\mu_{i,1}(x)$, $\mu_{i,}(x)$, $\mu_{i,3}(x)$,. . . . $\mu_{i,n}(x)$ (for *i*=1, 2, 3), is shown below.

$$\sigma_i = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} \left[ \mu_{i,k}(x) - \frac{1}{n} \sum_{k=1}^{n} \mu_{i,k}(x) \right]^2} \quad (6)$$

So far, the Dynamic Weight Vector *DW* with desired features has been determined.

## V. FINALIZATION OF SOV

After establishing the membership functions presented above, the membership degrees of the 4 input parameters for each candidate BS& AP can be calculated. In addition, we have to determine the dynamic weights for the 4 input parameters. For the $k_{th}$ ($1 \le K \le n$) base station $BS_k$, its membership degree vector $u_k$ is shown below

$$u_k = \begin{pmatrix} \mu_{1,k}(x) \\ \mu_{2,k}(x) \\ \mu_{3,k}(x) \\ \mu_{4,k}(x \end{pmatrix} \quad (7)$$

At the same time, we define the Weight Vector DW for 4 membership degrees.

$$DW = (DW_1, DW_2, DW_3, DW_4) \quad (8)$$

With (7) and (8), we can get the Value for $BS_k$, as shown in (9).

$$SOV_k(x) = DW * u_k \quad (9)$$

Or equivalently

$$SOV_k(x) = w_1\mu_{1,k}(x) + w_2\mu_{2,k}(x) + w_3\mu_{3,k}(x) + w_4\mu_{4,k}(x)$$

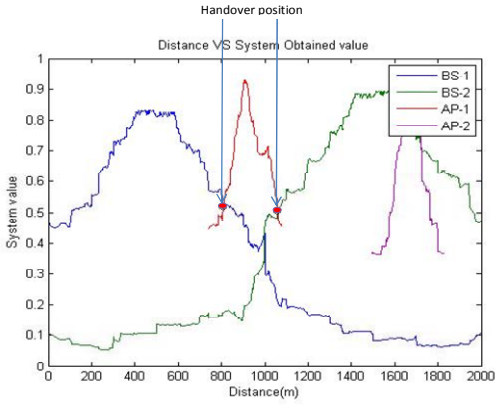



Figure 11: SOV of proposed system without user priority.

## VI. USER PREFERENCE

User consideration is also a very interesting characteristic in vertical handoff decision. It can include user intervention (user preferences), user interaction (with automation or not) or user satisfaction. When handoff happens, the users have more options for heterogeneous networks according to their preferences and network performance parameters. The user preferences could be preferred networks, user application requirements (real time, non-real time), service types (Voice, data, video), QoS (It is a set of technologies for managing network traffic in a cost effective manner to enhance user experiences for wireless environments) etc. Here four parameters S, C, R & BC are kept in user priority list. User can arrange the parameters from more priority to less priority sequence. User can choose any number of parameters from the list or user can keep deactivate user preference as user wish.

### *A. Normalised priority parameters*

In order to compare the attributes of different values and different units of measurement it is necessary to use the process of normalization. To evaluate the UOV from user preference normalized value of each priority parameter is taken. Normalization is needed to ensure that the values in different units are meaningful.

$$N(x) = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Where x=Security, Reliability, Cost & Battery status.
For the normalized n in the matrix, the first subscript denotes the priority input parameter, and the second subscript denotes the BS or AP.
For the $k_{th}$ ($1 \le K \le n$) $BS_k$ or $AP_k$, its membership degree vector $n_k$ is shown below:

$$n_k = \begin{pmatrix} n_{1,k}(x) \\ n_{2,k}(x) \\ n_{3,k}(x) \\ n_{4,k}(x \end{pmatrix} \quad (10)$$

### *B. Priority Weight Vector*

In user priority list most common and important options are kept for user. User can give priority from these options according to their preference. User can choose one more options from the list. Suppose user can choose cost as only one preference or can make a priority list from most prefer to less prefer. For user priority list weight vector are defined as follows:

$$PW_i = \frac{2P - 2i + 1}{\sum_{i=1}^{i=P} W_i}$$

$$W_i^{i=P}_{i=1} = (2P - 2i + 1)$$

Where
P = Number of priority parameter
$PW_i$ = weight for the $i^{th}$ priority parameter

$$PW = \{PW_1, PW_2, PW_3 PW_4\} \quad (11)$$

So as per user preference number priority weight vector is determined. From normalized value of priority parameters in (10) and from priority weight vector in (11) can get the UOV value for each BS and AP, as shown in (12).

$$UOV_k = n_k * PW \quad (12)$$

or equivalently,

$$UOV_k = n_{1,k} * PW_1 + n_{2,k} * PW_2 + n_{3,k} * PW_3 + n_{4,k} * PW_4$$

## VII. FDV FOR VERTICAL HANDOFF

The proposed technique has two sections, first section calculate SOV and second section determine UOV. According to the described procedure above the SOV and UOV can be

obtained. Finally adding SOV and UOV vector final vertical handoff decision vector FDV can be obtained.

$$FDV_k = SOV_k + UOV_k$$

$$FDV = \{FDV_1(x), FDV_2(x), FDV_3(x), \ldots\ldots\ldots, FDV_n(x)\}$$

Vertical handoff decision is executed when the FDV satisfies two conditions. First condition is

$$FDV_k = \max\{FDV_1(x), FDV_2(x), FDV_3(x), \ldots\ldots, FDV_n(x)\}$$

Second condition is

$$FDV_k \geq FDV_{TH}.$$

Where $FDV_{TH}$ is the threshold FDV Value. It helps to decrease unnecessary vertical handoffs. For the $k_{th}$ ($1 \leq k \leq n$) $BS_k$ or $AP_k$ , if it satisfies two conditions then, $k_{th}$ BS or AP is the final base station chosen for vertical handoff. The figures represent when the vertical handoff occurs and their corresponding FDV with distance.

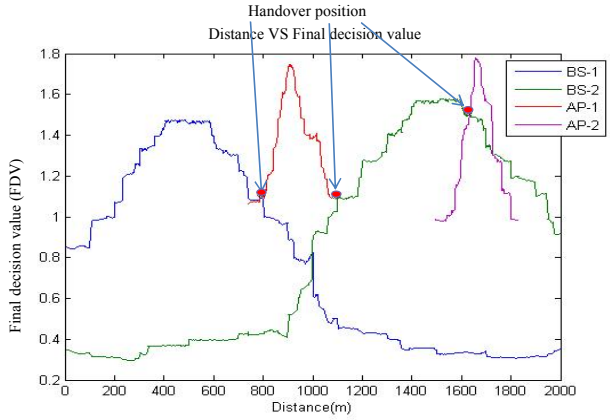



Figure 12: FDV of proposed system when user's 1st priority is security.

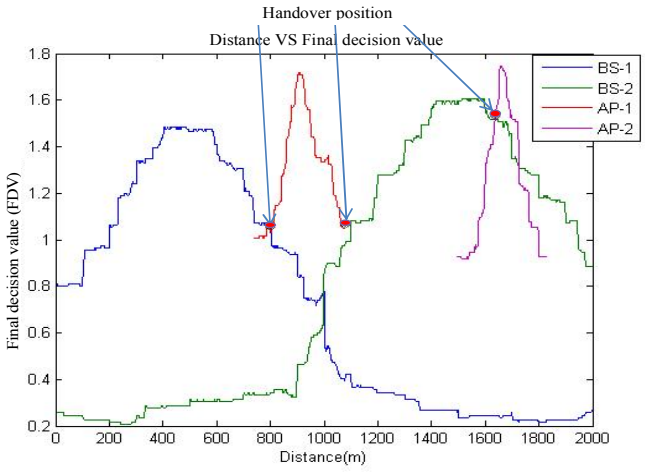



Figure 13: FDV of proposed system when user's 1st priority is cost.

## VIII. CONCLUSION

In this paper considered most Significant parameters which increases successful desired handoff. Dynamic decision vector takes very little time to take handoff decision due to less computation time. User priority choice ensures better satisfaction and handoff performance. Keeping all others parameter constant, just change in user priority list can occur vertical handoff. Combination of system and user part makes the method more reliable and efficient. From the simulation result it is seen that vertical handoff decision triggered at exact position to choose the best network and the proposed method is very effective for multi-criteria vertical handover decision. The method and algorithm used are very simple to carry out through software. Here MATLAB program is used to simulate the vertical handoff decision model.

## REFERENCES


[1] N. Nasser, S. Guizani, and E. Al-Masri, "Middleware vertical Handoff manager: A neural network-based solution," in *Proc. IEEE Int. Conf. on Commun., Glasgow, Scotland, UK.*, Jun. 2007, pp. 5671–5676.

[2] Y. H. Wang, C. P. Hsu, K. F. Huang, and W. C. Huang, "Handoff decision scheme with guaranteed QoS in heterogeneous network," in *Proc. First IEEE Int. Conf. on Ubi-Media Computing*, Jul. 2008, pp. 138 -143.

[3] M. Reza, H. Nezhad, Z. A. Zukamain, N. I. Udzir, and M. Othman, "A connection selection method for vertical handoff in hybrid wireless environment," in *Proc. First Int. Symposium on Applied Sciences on Biomedical and Communication Technologies*, Oct. 2008, pp. I -5.

[4] R. H. Katz and E. A. Brewer, "The case for wireless overlay networks," in *Proc. of SPIE Multimedia and Networking Conference, San Jose, CA*, Jan. 1996, pp. 77-88.

[5] Y. Yuan, D. Huang, and D. Liu, "An Integer Wavelet Based Multiple Logo-watermarking Scheme," in *Proc. of the First Int. Multi-Symposiums on Computer and Computational Sciences,* Jun. 2006, pp. 175-179.

[6] B. Honman and W. Benjapolakul, "A handover decision procedure for mobile telephone systems using fuzzy logic", in *Proc. of IEEE Asia-Pacific Conference on Circuits and Systems*, Nov. 1998, pp. 503-506.

[7] T. Onel, C. Ersoy, E. Cayirci, and G. Par, "A multicriteria Handoff decision schema for the next generation tactical communications systems," t*he Int. Journal of Computer and Telecommunications Networking*, vol. 46, no. 5, pp. 695–708, 2004.

[8] A. Majlesi, and B. H. Khalaj, "An adaptive fuzzy logic based handoff algorithm for interworking between WLANs and mobile networks," in *Proc. of the 13th Int. Symposium IEEE ISPIMRC*, Sep. 2002, vol. 5, pp. 2446 – 2451.

[9] Y. Ling, B. Yi, and Q. Zhu, "An improved vertical Handoff decision algorithm for heterogeneous wireless networks," in *Wireless communications, networking and mobile computing, WiCOM*, 2008, pp. 1–3.

[10] Q. Guo, J. Zhu, and X. Xu, "An adaptive multi-criteria vertical Handoff decision algorithm for radio heterogeneous network," in *Proc. of the IEEE Int. Conf. on Commun.*, May 2005, vol. 4, pp. 2769–2773.

[11] B. Hongyan, H. Chen, and J. Lingge, "Intelligent signal processing of mobility management for heterogeneous networks," in *Proc. of the IEEE Int. Conf. on Neural Networks & Signal Processing, China*, Dec. 2003, pp. 1578-1581.

[12] M. Sharma, and R. K. Khola, "Intelligent Approach for Seamless Mobility in Multi Network Environment," the *Int. Journal of Communication*, vol. 2, no. 2, pp. 43-47, Jul. 2011.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Optical Properties of ZnO Thin Films Prepared by Automatic Sol-gel Method

Shuva paul, Md. Faruk Hossain*, Md. Hasibul Islam, Md. Abu Raihan, Sajal Chakladar
Department of Electrical and Electronic Engineering
Rajshahi University of Engineering and Technology
Rajshahi-6204, Bangladesh
*E-mail: faruk94_ruet@yahoo.com

*Abstract*—**In this work, ZnO thin films were deposited on glass substrate using automatic sol-gel process, rather than conventional sol-gel process, and the optical properties of the prepared films were also measured. Using zinc acetate, ethanol, diethanolamine and de-ionized water, ZnO films were prepared by varying the number of coatings (2-5 dipping iterations) for zinc acetate concentrations of 1.0 mol/L. The films were decomposed at 300$^0$C for 1 h and, after giving final coating, were annealed at 500$^0$C for 2 h. Optical transmittance measurements of the deposited films have been carried out with UV/Visible spectrometer at 200-900 wavelengths. The prepared ZnO films have been found to be highly transparent in the visible range. The average transmittance of ZnO films for 2 coatings has been found to be of 92% in the visible range with high absorbance in the UV-region and the direct and indirect allowed band gap has been found to be of 3.21 eV and 3.08eV respectively.**

***Keywords—Automatic sol-gel, optical properties, band gap, ZnO.***

## I. INTRODUCTION

Among various functional materials, metal oxide films are very attractive for many applications due to their good electrical and optical properties [1]. Now a days, zinc oxide (ZnO) has drawn much attention of the researchers as one of the most promising semiconducting materials for its versatile applications in blue and ultraviolet light emitter, varistors, chemical and gas sensors, surface acoustic devices, spintronic devices, optoelectronic devices and piezoelectric transducers [2]-[5]. Many research groups are trying to achieve p-type conductivity and as well as to control unintentional n-type conductivity of ZnO [6]. It is well known that, ZnO crystallizes in the wurtzite hexagonal structure with c-axis orientation that enhances the anisotropic structures formation [7]. But the availability of ZnO as a large bulk single crystal is the big advantage of ZnO over GaN that also crystallizes in the wurtzite structure [6]. Moreover, band gap of ZnO is 3.37 eV at room temperature [8], while that of GaN is 3.44 eV [6]. Again free-exciton binding energy in ZnO is 60 meV [8] and that of GaN is 25 mev [6]. Lower band gap and higher exciton binding energy makes ZnO a most promising material for application in optoelectronics and excitonic effects based optical devices over GaN. In addition, the electron mobility of ZnO (115-155 $cm^2 V^{-1} s^{-1}$) at room temperature is higher than that of $TiO_2$ ($<10^{-5} cm^2 V^{-1} s^{-1}$) [9]. As a result, with similar band gap of $TiO_2$, ZnO nanoparticles are being widely used in the Grätzel-type solar cells [7]. Last but not least, ZnO or In/Al-doped ZnO are recently being used as transparent conductive oxide (TCO), due to high transparency in visible range, as an alternative of ITO ($F:SnO_2$) and FTO ($F:SnO_2$) because they are very costly.

Electrical and optical properties of ZnO thin films are greatly dependent on the deposition technique applied. So far, various deposition techniques have been employed to prepare ZnO thin films such as RF magnetron sputtering [10], molecular beam epitaxy [11], spray pyrolysis [12], hydrothermal growth [13], pulse laser deposition [14], chemical vapor deposition [15] and sol-gel method [8]-[9], [16]-[17]. But the sol-gel technique offers excellent homogeneity, choice to use both organic and inorganic precursors, simplicity, low cost, high quality films on large area substrate and above all, environment friendly [18]-[19].

This work is the continuation of our previous work [20], where ZnO thin films were successfully deposited on glass substrate using a homemade length controlled automatic sol-gel process. In previous work, it was only shown that the designed automatic system can deposit thin films successfully but the prepared films were not characterized with any properties. The present work reports on the investigation of optical properties of automatic sol-gel derived ZnO thin films that have been deposited on glass substrate with various number of coatings.

## II. OPERATING PRINCIPLE OF AUTOMATIC SOL-GEL METHOD

Fig. 1 shows the homemade automatic sol-gel system for deposition of thin films [20]. The whole system is arranged on a wooden basement. The main features of the system are motor, plastic ribbon, damper weight, clip, beaker base and a keyboard. Clip is attached to a plastic ribbon, one end of which is coiled on the motor shaft. A damper weight has been attached with the ribbon to keep it straight and to eliminate the swing and vibration because the ribbon is too thin to have continuous swing in the natural wind and exposed to vibration while the motor rotates. There are six keys on the keyboard, corresponding to six different lengths (2cm, 3cm, 4cm, 5cm, 6cm and 7cm), and a power switch .

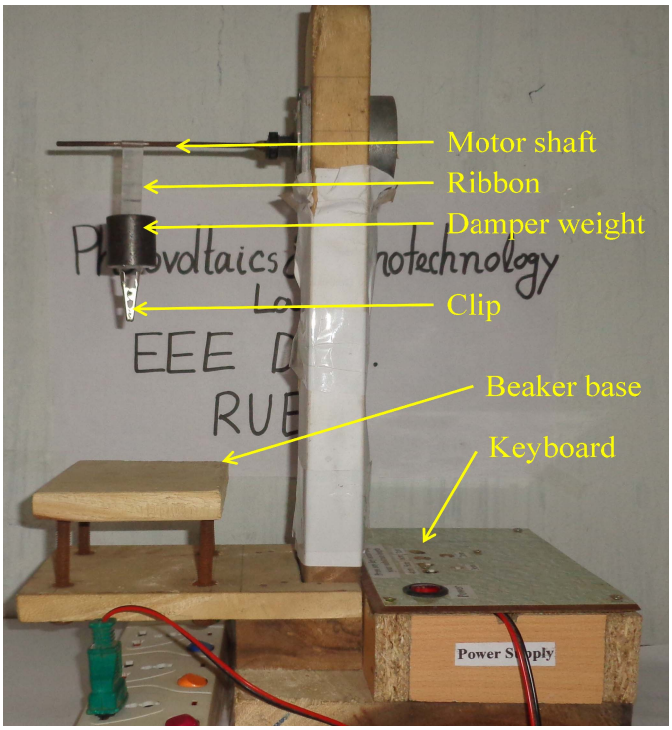



Figure 1. Automatic sol-gel system

After preparing the precursor solution, it is kept on the beaker base and the substrate is hanged on the clip. The length of ribbon from motor shaft is so adjusted that the lower part of the substrate just touches the precursor solution. Then the power switch is turned on and one of the six keys is pressed according to the required length of the substrate on which thin films have to be deposited. This causes the motor to start rotation, at the same time the ribbon is uncoiled from the motor shaft and the substrate starts to dip into the precursor solution at a constant rate (1cm/min).When the required length of substrate is fully immersed into the solution, the motor stops rotation for 10 seconds. After that the motor again starts to rotate but in opposite direction so that ribbon is coiled on the motor shaft automatically that causes the substrate to be withdrawn from the solution at the same rate of dipping (i.e. 1cm/min). When the substrate has been fully withdrawn from the solution it is then unclipped from substrate holder and is moved away for heat treatment.

## III. Experimental Section

### A. Chemical Reagents

ZnO precursor solution was prepared using absolute ethanol ($M$=46.07, $CH_3CH_2OH$, EtOH) as solvent, zinc acetate di-hydrate ($M$=219.50, $Zn(CH_3COO)_2.2H_2O$, ZAD) as precursor, diethanolamine ($M$=105.14, $[CH_2(OH)CH_2]_2.NH$, DEA) as chelating agent and de-ionized water.

### B. Preparation of ZnO Films

In this work, ZnO films were deposited for zinc acetate concentration of 1.0 mol/L with various number of coatings (2-5 dipping iterations). ZnO precursor solution was prepared in the following way. At first, 5.488g of zinc acetate was dissolved in 25 ml of absolute ethanol to yield 1.0 mol/L concentration of precursor solution and then was magnetically stirred for 30 min. A milky solution was obtained where with a molar ratio of DEA/ZAD as 1:1, 4.823 ml of DEA and with a molar ratio of de-ionized water/ZAD as 2:1, 1.8 ml de-ionized water was added into the emulsion. The solution was again stirred and a clear solution was obtained after a certain period. Meanwhile, a homogeneous mixer of absolute ethanol and de-ionized water was prepared with a volumetric ratio of 8:2. Finally, the precursor solution was hydrolyzed with drop by drop addition of prepared mixer of ethanol and de-ionized water keeping the solution under continuous stirring.

Glass micro-slides were used as substrate for deposition of ZnO films. The substrates were cleaned ultrasonically, first in ethanol and subsequently in ion exchanged distilled water for 15 min each at $50^0$C. After that solvents from the surface of substrate were swept away by using an air compressor and kept in an oven at $80^0$C for 10 min. Then 2cm length of the clean substrate was coated with the prepared sol using automatic sol-gel process. The dipping and withdrawing rate was 1cm/min and interval between consequent dipping and withdrawing was 10s. The coated films were dried at room temperature for 24 h and then heated at $300^0$C in air for 1 h to decompose the precursor film to ZnO. After that another deposition cycle was carried out. After completing the required deposition cycles, the films were annealed at $500^0$C at a heating rate of $10^0$C/min and left at $500^0$C for 2 h.

The optical properties of the deposited films were measured by UV/Visible spectrometer at 200-900 nm wavelengths using reference substrate. From transmittance spectra absorbance was determined using Beer-Lambert Law and optical band gap was found using Tauc relationship.

## IV. Results and Discussion

Fig. 2 shows the images of automatic sol-gel derived ZnO films for various numbers of coatings with zinc acetate concentration of 1.0 mol/L. It can be observed that the all films are visually transparent and the transparency decreases with the increase of number of coatings. This implies that, ZnO films with 5 coatings show higher absorbance than the films with decreased number of coatings. This is because increasing number of coatings causes to increase the thickness of the film and the increase in absorbance of ZnO films with the increase of film thickness has been reported by [8], in case of conventional sol-gel derived ZnO films.

Fig. 3 shows the optical transmittance of ZnO films for 2 coatings in the wavelength range 300-900 nm. It is clearly seen that, in the visible range of electromagnetic spectrum the films are highly transparent with an average transmittance value of 92% beyond the wavelength of 400 nm and presents sharp ultraviolet cutoff at approximately 360 nm. This transmittance value is higher than the previously reported transmittance value of ZnO films [21]. Transmittance reaches to its maximum value at higher wavelength.

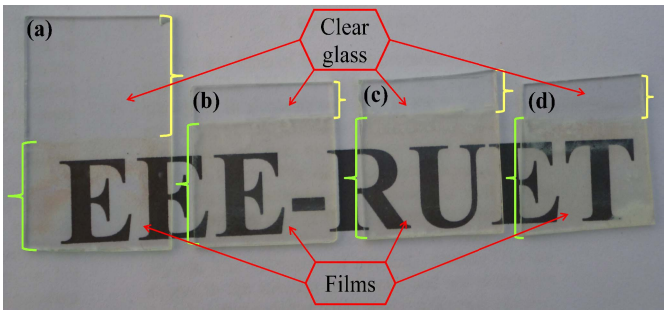



Figure 2. Images of ZnO films for zinc acetate concentration of 1.0 mol/L with (a) 2, (b) 3, (c) 4, and (d) 5 coatings.

Fig. 4 shows the absorption spectrum of automatic sol-gel derived ZnO for 2 coatings. It is clearly observed that the films have a very low absorption at transparent region and high absorption at ultraviolet region. The absorption value of the prepared ZnO films at ultraviolet region is lower than previously reported data [8]. This is because of the smaller thickness of ZnO films as only 2 coatings have been given.

The direct and indirect allowed optical transitions between valance and conduction bands can be evaluated by fitting a straight line in strong absorption spectral region using the Tauc relationship. According to Tauc law dependence of absorption co-efficient (α) on photon energy (hν) can be given by [22]

$$(\alpha h\nu) = A(h\nu - Eg)^{r} \quad (1)$$

Where α is the absorption co-efficient, A is the edge width parameter, and hν is the photon energy, and r is a constant, for direct allowed transition r equals ½ and for indirect allowed transition equals 2.

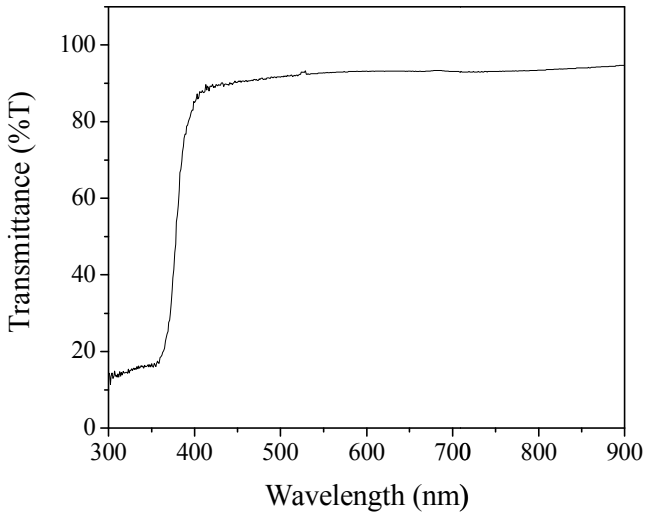



Figure 3. Optical transmittance ZnO thin films

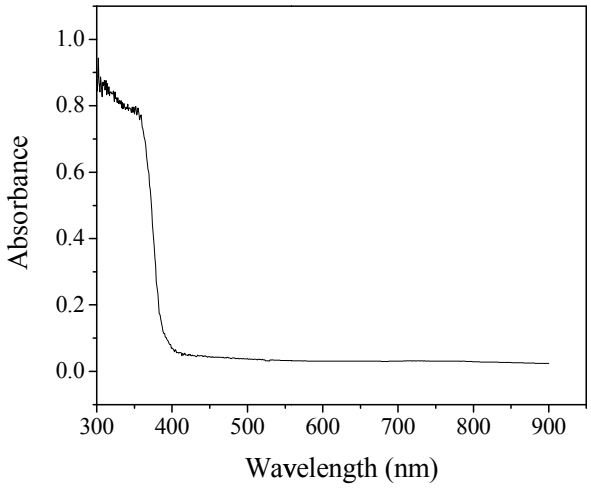



Figure 4. Optical absorbance of ZnO thin films

Fig. 5 shows the plot of $(\alpha h\nu)^2$ vs. hν of prepared ZnO films for 2 coatings. The optical band gap for direct allowed transition of the films have been determined from the extrapolation of the linear portion of $(\alpha h\nu)^2$ vs. hν at α= 0. Direct transition band gap is found to be of 3.21eV. Although this value is smaller than the bulk value of 3.37 eV but still it is in good agreement with previously reported data of ZnO thin films [23].

Fig. 6 shows the plot of $(\alpha h\nu)^{1/2}$ vs. hν of ZnO films for 2 coatings and band gap for indirect allowed transition has been determined by extrapolating the straight line portion of the spectrum at α= 0. The indirect allowed band of prepared ZnO thin films with 1.0 mol/L and 2 coatings has been found to be of 3.08 eV which is greater than the previously reported indirect allowed band gap of 3.0 eV for ZnO films [22].

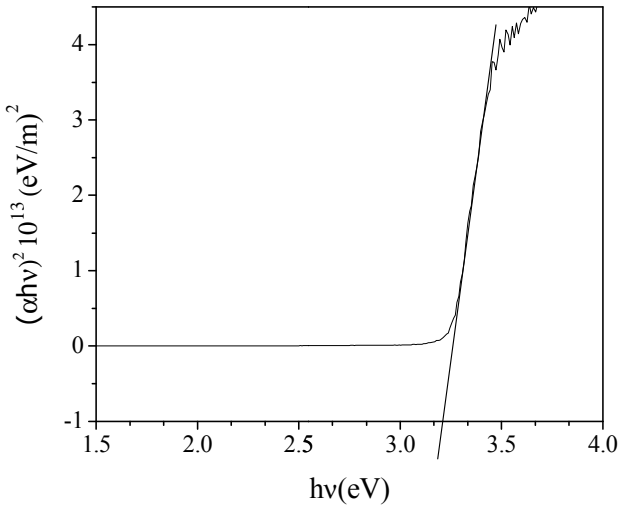



Figure 5. Plot of $(\alpha h\nu)^2$ vs. hν of ZnO thin films.

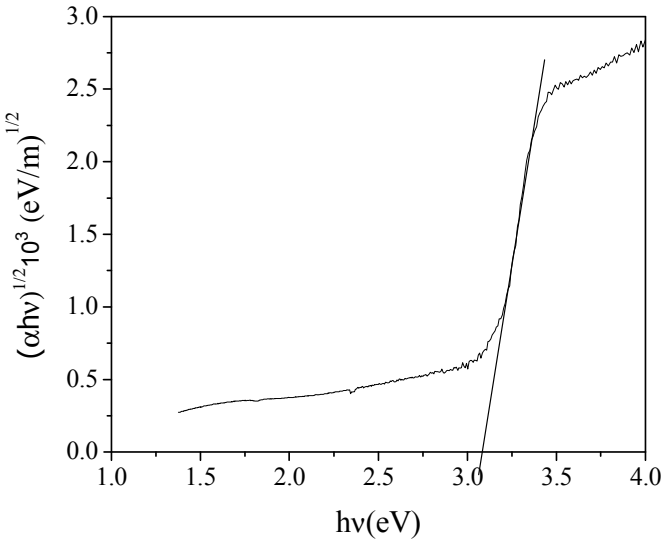



Figure 6. Plot of $(\alpha h\nu)^{1/2}$ vs. hν of ZnO thin films.

## V. CONCLUSION

ZnO thin films were successfully prepared on glass substrate by using effective and user friendly automatic sol-gel technique with zinc acetate as precursor. ZnO thin films were prepared with various numbers of coatings with zinc acetate concentration of 1.0 mol/L. In order to obtain crystal structure of ZnO, as-deposited films were annealed at $550^0$C for 2h. The results have revealed that, ZnO films prepared by automatic sol-gel method are highly transparent at visible region (92%). Direct and indirect allowed band gap has also to be found with satisfactory value. Higher transmittance in the visible region and higher absorption in the UV-region has made the films potentially able for optical window applications and also as transparent conductive oxide (TCO).

## ACKNOWLEDGMENT

The authors would like to thank the Photovoltaic & Nanotechnology Laboratory of Rajshahi University of Engineering & Technology for giving chance to use different equipments.

## REFERENCES

[1] K. C. Yung, H. Liem, and H. S. Cho, "Enhanced redshift of the optical band gap in Sn-doped ZnO free standing films using the sol–gel method", *J. Phys. D: Appl. Phys.,* vol. 42, no. 18, pp. 185002(5pp), Sept. 2009.

[2] Z. Liu, Z. Jin, W. Li, J. Qiu, "Preparation of ZnO porous thin films by sol–gel method using PEG template", *Materials Letters*, vol. 59, no. 28, pp. 3620-3625, Dec. 2005.

[3] S. S. Shariffudin, M. Salina, S. H. Herman, and M. Rusop, "Effect of film thickness on structural, electrical and optical properties of sol-gel deposited layer-by-layer ZnO nanoparticles," *Transactions on Electrical and Electronic Materials*, vol. 13, pp. 102-105, April 2012.

[4] B. Wacogne, M. P. Roe, A. T. Pattinson, and C. N. Pannell, "Effective piezoelectric activity of zinc oxide films grown by radio-frequency planar magnetron sputtering," *Applied Physics Letters,* vol. 67, no.12 , pp. 1674-1676, July 1995.

[5] F. K. Shan, B. I. Kim, G. X. Liu, Z. F. Liu, J. Y. Sohn, W. J. Lee, B. C. Shin, and Y. S. Yu, "Blueshift of near band edge emission in Mg doped Zno thin films and aging," *Journal of Applied Physics*, Vol. 95, No. 9, pp. 4772-4776**,** Feb**.** 2004.

[6] A. Janotti and C. G. V. d. Walle, "Fundamentals of zinc oxide as a semiconductor", *Rep. Prog. Phys.*, vol. 72, no. 12, pp. 126501(29pp), July 2009.

[7] M. F. Hossain, T. Takahashi, S. Biswas, "Nanorods and nanolipsticks structured ZnO photoelectrode for dye-sensitized solar cells", *Electro. Commu.*, vol. 11, pp. 1756-1759, July 2009.

[8] M. F. Hossain, S. Biswas, M. Shahjahan, and T. Takahashi, "Study of sol-gel derived porous ZnO photoelectrode for the application of dye – sensitized solar cells," *J. Vac. Sci. Technol. A*, vol.27, no.4, pp.1047-1051, Jul/Aug. 2009.

[9] M .F. Hossain, Z. H. Zhang and T. Takahashi, "Novel micro-ring structured ZnO photoelectrode for dye-sensitized solar cell," *Nano-Micro Lett*., vol.2, no.1, pp. 53-55, Apr. 2010.

[10] G. Sberveglieri, S. Groppelli, and G. Coccoli, "Radio frequency magnetron sputtering growth and characterization of indium-tin oxide (ITO) thin films for $NO_2$ gas sensors", *Sensors and Actuators,* vol. 15, no. 3, pp. 235–242, Nov. 1988.

[11] A. Ohtomo, M. Kawasaki, T. Koida, K. Masubuchi, H. Koinuma, Y. Sakurai, Y. Yoshida, T. Yasuda, and Y. Segawa, "$Mg_xZn_{1-x}O$ as a II-VI widegap semiconductor alloy," *Applied Physics Letters*, vol.72, no.19, pp. 2466-2468, March 1998.

[12] S. Ilican, Y. Caglar, M. Caglar and B. Demerci, "Poly-crystalline Indium-doped ZnO thin films: preparation and characterization," *Journal of Optoelectronics and Advanced Materials*, vol. 10, no. 10, pp. 2592- 2598, 2008.

[13] H. Endo, M. Sugibuchi, K. Takahashi, S. Goto, S. Sugimura, K. Hane, and Y. Kashiwaba, "Schottky ultraviolet photodiode using a ZnO hydrothermally grown single crystal substrate," *Appl. Phys. Lett*., vol. 90, no. 12, pp. 121906(3pp), March 2007

[14] N. Naghavi, A. Rougier, C. Marcel, C. Gueary, J. B. Leriche and J.M. Tarascon, "Characterization of indium zinc oxide thin films prepared by pulsed laser deposition using a $Zn_3In_2O_6$ target," *Thin Solid Films*, vol. 360, no. 1-2, pp. 233-240, Feb. 2000.

[15] J. H. Lee and B. O. Park, "Transparent conducting ZnO:Al, In and Sn thin films deposited by the sol-gel method," *Thin Solid Films*, vol. 426, no. 1-2, pp. 94-99, Feb. 2003.

[16] C. Y. Tsay, M. C. Wang, and S. C. Chiang, "Effects of Mg additions on microstructure and optical properties of sol-gel derived ZnO thin films", *Materials. Transactions*., vol. 49, no .5, pp. 1186-1191, April 2008.

[17] L. Xu, X. Lu, Y. Chen, and F. Xu, "Structural and optical properties of ZnO thin films prepared by sol-gel method with different thickness," *Applied Surface Science,* vol. 257, pp. 4031-4037, Feb. 2011

[18] A. Muthuvinayagam, N. Melikechi, P. D. Christy and P. Sagayaraj, "Investigation on mild condition preparation and quantum confinement effects in semiconductor nanocrystals of $SnO_2$", *Phys. B,* vol. 405, no. 4, pp. 1067-1070, 2010.

[19] F. E. Ghodsi and H. Absalan, "Comparative study of ZnO thin films prepared", *Acta Physica Polonica A,* vol. 118, no. 4, pp. 659-664, 2010.

[20] S. Paul, D. M. F. Hossain, M. H. Islam, M. F. Ali, and U. Shaha, "Synthesization of ZnO thin films by length controlled automatic sol-gel process", *Int. Conf. Inf. Elec. Vission,* vol. 1, pp. 85, May 2013.

[21] Z. Liu, J. Li, J. Ya, Y. Xin, and Z. Jin, "Mechanism and characteristics of porous ZnO films by sol-gel method with PEG template". *Materials Letters*, vol.62, no. 8-9, pp. 1190-1193, March 2008.

[22] N. Shakti, P. S. Gupta, "Structural and optical properties of sol-gel prepared ZnO thin film", *Applied Physics Research*, vol. 2, no. 1, pp. 19-28, 2010.

[23] W. R. Saleh, N. M. Saeed, W. A. Twej, and M. Alwan, "Synthesis sol-gel derived highly transparent ZnO thin films for optoelectronic applications", *Advances in Materials Physics and Chemistry*, vol. 2, no. 01, pp. 11-16, March 2012.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# An Analytical Approach to Study Energy Band Structure in Strained Graphene

Md.Shamim Sarker [*], Muhammad Mainul Islam[+], Md. Rafiqul Islam[#]
Department of Electrical and Electronic Engineering,
Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh
E-mail: [*]sarkershamimbd@gmail.com, +mainul0803004@gmail.com and [#]islambit@yahoo.com

*Abstract*— **We have proposed a simple analytical approach to study energy band structure in strained graphene. In our study, the strain-induced deformation of primitive unit cell is included in tight binding model. The unit cell deformation is determined in terms of the angle between two primitive unit vectors. The proposed method is applied to evaluate the band gap under the application of biaxial symmetrical and uniaxial strains in graphene lattice structure. We found zero band gap for biaxial strained graphene and this result agrees with that of estimated by first principle method. The band gaps are also estimated in uniaxial strained graphene. The results in the present study are compared with that of determined by first principle method and found to be in good agreement.**

***Keywords—Graphene, planner strain, tight binding model, analytical approach, band structure.***

## I. INTRODUCTION

Graphene, a strictly two-dimensional material having unusual and interesting properties [1] is a rapidly rising star on the horizon of material science and condensed matter physics. It is a material of interest in semiconductor industry because of its exceptionally high crystal and electronic quality, excellent transport properties (i.e. high electron mobility [2] and high thermal conductivity), and as it is planner, it is capable of extreme device scaling comparing with silicon technology. However, these excellent properties are hindered with a major drawback; graphene is a zero bandgap semiconductor or semimetal [3]-[4]. For large scale manufacturing, the absence of bandgap is the most difficult engineering issue to solve. The zero bandgap revels that it is impossible to switch graphene based device from the conductive to the nonconductive state. This implies that it cannot be used in logic circuits.

As the zero bandgap property of graphene limits its application in practical fields, scientists are working to find out the methods to open the bandgap in graphene. Several methods have been proposed, such as creating quantum confinement effect in transverse direction of graphene nanoribbon (GNR) [5]-[8], introducing symmetry breaking between two carbon layers via an external electric field in bilayer graphene [9],[10], by the process of doping [11]-[13], and by the process of external strain [14],[15]. To investigate the band gap, several techniques have been applied such as first principal calculation, tight binding modeling, k.p method etc. Most of them are performed using the software simulation or numerical techniques, which requires a huge computational complexity and time, and need high capacity super computer. Some analytical works have also been done on strained graphene and GNR using TB model which mainly focuses on the strength of the strain or the spin polarization interaction [16].

In our study we have proposed a simple analytical approach to investigate the energy dispersion of graphene under different planner strain conditions. The strain induced unit cell deformation is included in the framework of tight binding model. Using the proposed method the band gap is determined under the application of biaxial symmetrical strain and uniaxial strain parallel and perpendicular to the carbon-carbon (C-C) bonds in graphene structure. The results obtained from the proposed method is compared with the results published by the first principle method [14] and found to be in good agreement.

## II. Proposed model

Graphene is a honeycomb lattice of regular hexagonal structure. But it loses its regular hexagonal structural symmetry under uniaxial/shear strain. The position of carbon atoms shifts relative to each other under the application of planar stress to grapheme. As a result the vector position of lattice point changes. To explain this phenomenon, the angle between the unit vectors $\mathbf{a_1}$ and $\mathbf{a_2}$ is represented by $\theta$ instead of assuming $60^o$ which is true for ideal or relaxed graphene. The strain-induced effect is included in the tight-binding Hamiltonian.

Fig. 1 (a) and (b) shows the ideal and strained lattice structures of graphene, respectively. We have used the nearest neighbor tight binding model. Here the type-A atom in the graphene lattice is σ bonded with three of its equivalent nearest neighbors of type-B atoms. The primitive unit vectors

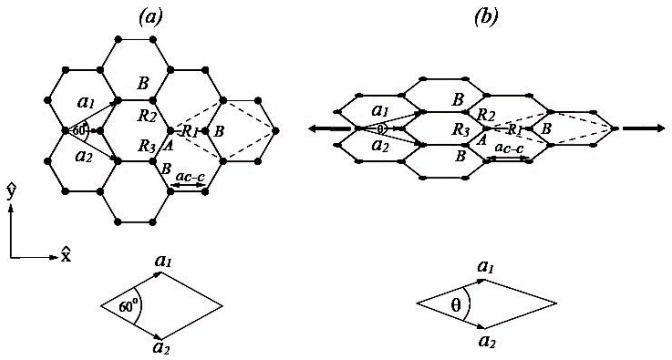



Figure 1. The graphene lattice structures a) ideal structure with angle between two primitive vectors $60^0$ and b) strained structure with angle $\theta$. The corresponding unit cells representation are shown below the lattice structure.

in terms of $\theta$ can be represented by

$$\mathrm{a}_1 = [a\cos(\theta/2), a\sin(\theta/2)]; \quad (1)$$

$$\mathrm{a}_2 = [a\cos(\theta/2), -a\sin(\theta/2)]; \quad (2)$$

where $|\mathrm{a}_1| = |\mathrm{a}_2| = a$. The separation between the type-A atom and its three nearest neighbors type-B atoms represented by the following three vectors $\mathbf{R_1}$, $\mathbf{R_2}$, and $\mathbf{R_3}$

$$\mathrm{R}_1 = [\frac{a}{2\cos(\theta/2)}, 0] \quad (3)$$

$$\mathrm{R}_2 = -a_2 + R_1 = [\frac{a}{2}\cos(\theta/2) - a\cos(\theta/2), a\sin(\theta/2)] \quad (4)$$

$$\mathrm{R}_3 = -a_1 + R_1 = [\frac{a}{2}\cos(\theta/2) - a\cos(\theta/2), -a\sin(\theta/2)] \quad (5)$$

The nearest neighbor tight binding energy dispersion formula can be given by [17]

$$E(k) = \pm\sqrt{H_{AB}(k)H^*_{AB}(k)} \quad (6)$$

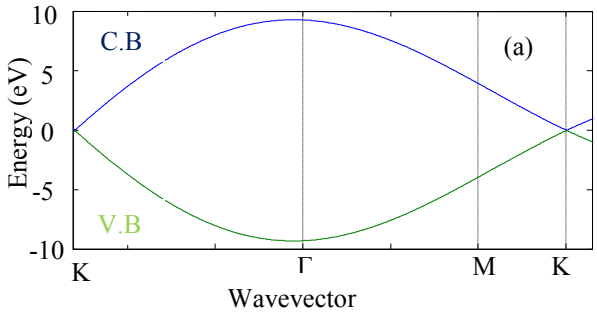
The Hamiltonian matrix element can be calculated by

$$H_{AB}(k) = \gamma(e^{-ik.\mathrm{R}_1} + e^{-ik.\mathrm{R}_2} + e^{-ik.\mathrm{R}_3}) \quad (7)$$

where $\gamma$ is a fitting parameter which is often called the nearest neighbor overlap energy or the C-C interaction energy. Finally, using "(6)" and "(7)" the generalized energy dispersion relation becomes

$$E(k) = \pm\gamma\sqrt{1 + 4\cos[a\cos(\theta/2)k_x]\cos[a\sin(\theta/2)k_x] + 4\cos^2[a\sin(\theta/2)k_x]} \quad (8)$$

The value of $\gamma$ ranges from 2.7eV to 3.3 eV [17].



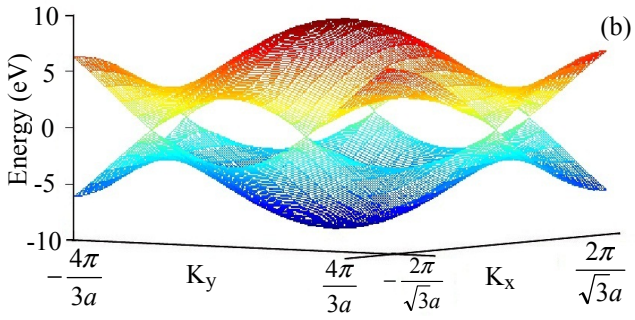




Figure 2. The nearest neighbor tight-binding band structure of relaxed graphene (a) $\Gamma - M$ and $\Gamma - K$ representation of Brillouin zone (b) hexagonal representation of Brillouin zone.

For the unstrained or relaxed condition, the value of the angle, between $\theta = 60^o$

The states of primitive unit cells under the application of biaxial symmetrical strain and uniaxial strain (perpendicular to and parallel to C-C bonds) conditions are shown in Figs. 3(a), (b) and (c), respectively, with corresponding primitive cells, reciprocal lattices and Brillouin zones below the deformed lattices. The hexagonal symmetry of unit cell does not change under the application of biaxial symmetrical strain as seen in Fig. 3(a).

Under the application of uniaxial strain perpendicular and parallel to C-C bonds, the regular hexagonal structure is deformed as seen in Figs. 3(b) and (c). It also causes the deformation in the primitive unit cell due to symmetry breaking. The primitive unit cell deformation causes the change in angle $\theta$ between the primitive unit vectors. When strain is applied gradually up to 12.2 % parallel to C-C bonds the angle between the primitive unit vectors $\theta_{pl}$ can be given by the following equation with respect to $L_x$ and $L_y$

$$\theta_{pl} = -1.1618\,e^{-8}L_y^4 + 2.8783\,L_y^3 - 5816.2155\,L_y^2 + 2436.573\,L_y - L_x - D_{pl} \quad (9)$$

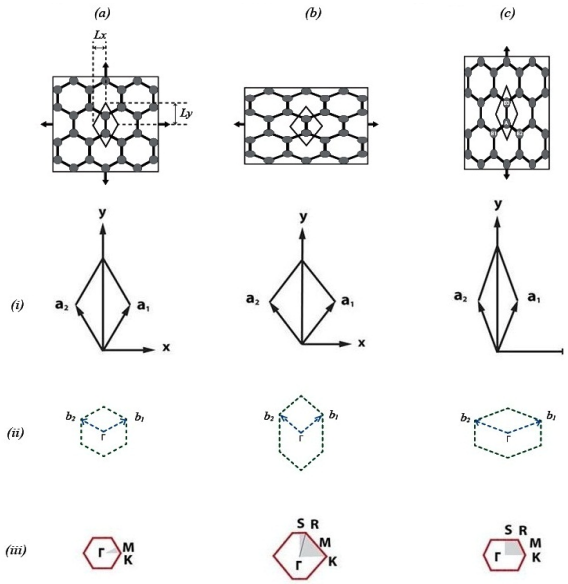



Figure 3. Graphene lattice structures (a) biaxial symmetrical strain (b) uniaxial strain perpendicular to C-C bonds, and (c) uniaxial strain parallel to C-C bonds. Corresponding primitive cells, reciprocal lattices, and Brillouin zones are shown below the deformed lattices. Γ, K, M, R and S are the high symmetrical points. $L_x$ and $L_y$ are the half of the diagonal lengths of the primitive cells in parallel and perpendicular directions of the C-C bonds.

In case of strain applied gradually up to 6.5% perpendicular to C-C bond, the strain-dependent change in angle between the primitive unit vectors $\theta_{pr}$ in terms of $L_x$ and $L_y$ can be given by the following equation

$$\theta_{pr} = 9030.6822L_x^2 - 2090.482L_x + L_y + \mathrm{D_{pr}} \qquad (10)$$

where $D_{pl}$ = 195.028 and $D_{pr}$ = 180.188 are fitting parameters and $L_x$ and $L_y$ are the half of the diagonal lengths of the primitive cells in parallel and perpendicular directions of the C-C bonds in nanometers.

## II. RESULTS AND DISCUSSION

We have studied the electronic band structure of biaxial symmetrical strained graphene as well as uniaxial strained graphene using the analytical approach proposed here. The energy dispersion profiles are determined at different critical points of Brillouin zone using "(8)" for biaxial symmetrical strain condition, that is for $\theta = 60^0$. The results are plotted in 2-dimensional space and hexagonal space in Fig. 2 (a) and (b), respectively, where zero band gap is found at $K$ point and maximum at $\Gamma$ point of the Brillouin zone. The hexagonal symmetry of unit cell does not change under the application of biaxial symmetrical strain as seen in Fig. 3(a), which results the same band property as is found for the relaxed graphene [17].

We also studied the electronic band structures of strained graphene under the application of uniaxial strain perpendicular to and parallel to C-C bonds. Uniaxial strain distribution results in the opening of the bandgap between the maximum of the valance band and the minimum of the conduction band in graphene [14]. When uniaxial strain is applied parallel to C-C bonds, the value of $L_y$ increases and becomes constant for a deformed primitive cell. The value of $L_x$ decreases until the system reaches to its lowest total energy during structural relaxation. Due to the change of $L_x$ and $L_y$, the angle $\theta$ between the primitive unit vectors decreases and causes the symmetry breaking. This angular change is taken as the parameter of deformed primitive cell to calculate the electronic band structure of graphene. For the variation of $L_y$ 0.2196, 0.2236, and 0.2396 nm the corresponding optimized values of $L_x$ are found 0.1228, 0.1224, and 0.1217 nm [14]. Using "(9)" the corresponding angle $\theta$ between the primitive unit vectors are estimated 59.47°, 58.91°, and 54.79°. Substituting the values of $\theta$ into "(8)" the electronic band structure or band diagrams are calculated and plotted in Fig. 4 (a), (b), and (c). To observe the band gap opening clearly the figures are extended near the K- point of the Brillouin zone. The band gap opening in strained graphene is evaluated 0.051, 0.106, and 0.489 eV, respectively, under the application of uniaxial strain parallel to C-C bonds.

We also applied the proposed method to study the strain-induced band gap opening for the graphene system when it is strained perpendicular to C-C bonds. In this case, for the variation of $L_x$ = 0.1268, 0.1292, and 0.1313 nm the corresponding optimized $L_y$ are reported 0.2126, 0.2120, and 0.2115 nm [14]. Using "(10)" the corresponding deformed angles $\theta$ are determined $60.52^0$, $61.05^0$, and $61.605^0$. Using "(8)", the band gap openings corresponding to the deformed angles are estimated 0.053, 0.104, and 0.162 eV. The results are shown in Figs. 5(a), (b), and (c). To observe the results clearly, the dispersion profiles are extended near the K-point of the Brillouin zone.

The results obtained in the present study revel that the zero bandgap or semi-metallic behavior of graphene sheet gets modified or a bandgap is opened under the application of uniaxial strain. Now the question is what is the reason behind this? We know that planner graphene consists of strong $\sigma$ bonds and delocalized $p_z$ electrons. Here $\pi$ orbitals are formed by overlapping the $p_z$ orbitals of the carbon atoms in the hexagonal lattice. These $\pi$ and $\pi^*$ bands touches each other at the K point causing zero bandgap. When the strain is applied the C-C bonds of graphene gets elongated. Due to this elongation of C-C bonds, the $\pi$ electron clouds get localized on the corresponding carbon atoms. Therefore a restriction is imposed on movement of the electrons which causes the opening the bandgap at the K point. In this way, under strained condition, graphene loses its semimetal characteristics and turns into a direct bandgap semiconductor.

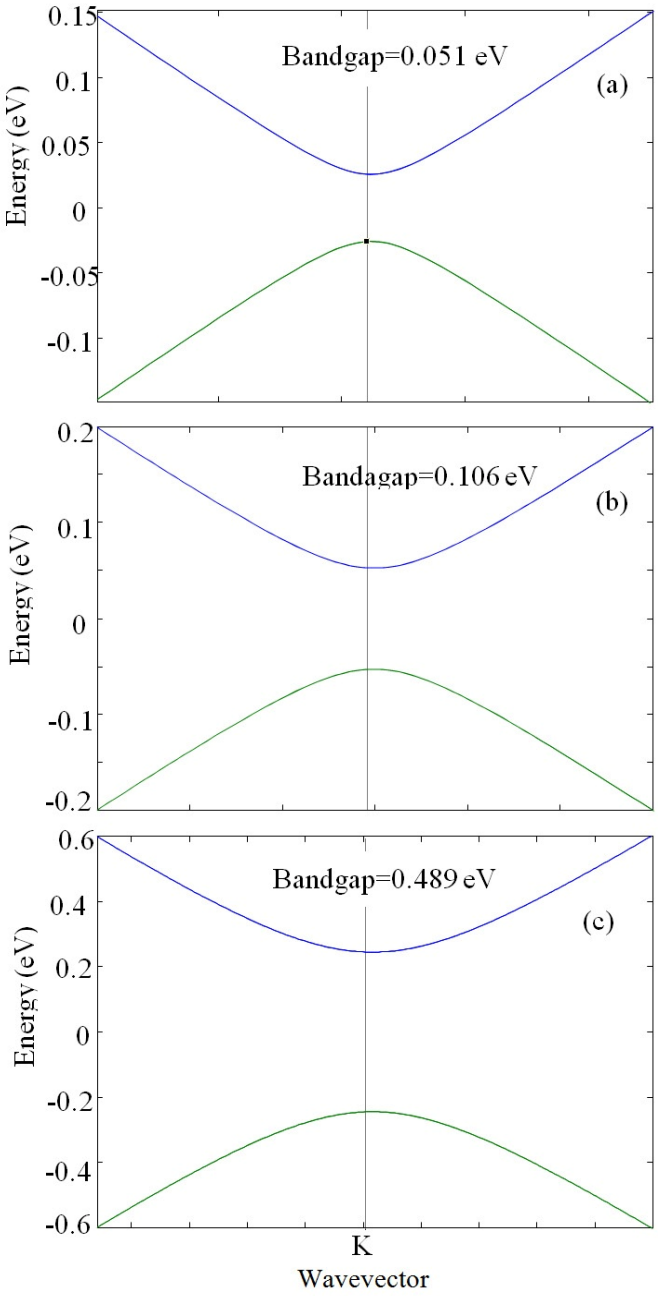



Figure 4. Band gap opening under the application of uniaxial strain parallel to C-C bonds near the K-point of Brillouin zone (a) 2.81 % strained grapheme ($L_y$=0.2196 nm and $L_x$=0.1228 nm) (b) 4.68 % strained grapheme ($L_y$=0.2236 nm and $L_x$=0.1224 nm), and (c) 12.2 % strained grapheme ($L_y$=0.2396 nm and $L_x$=0.1217 nm).

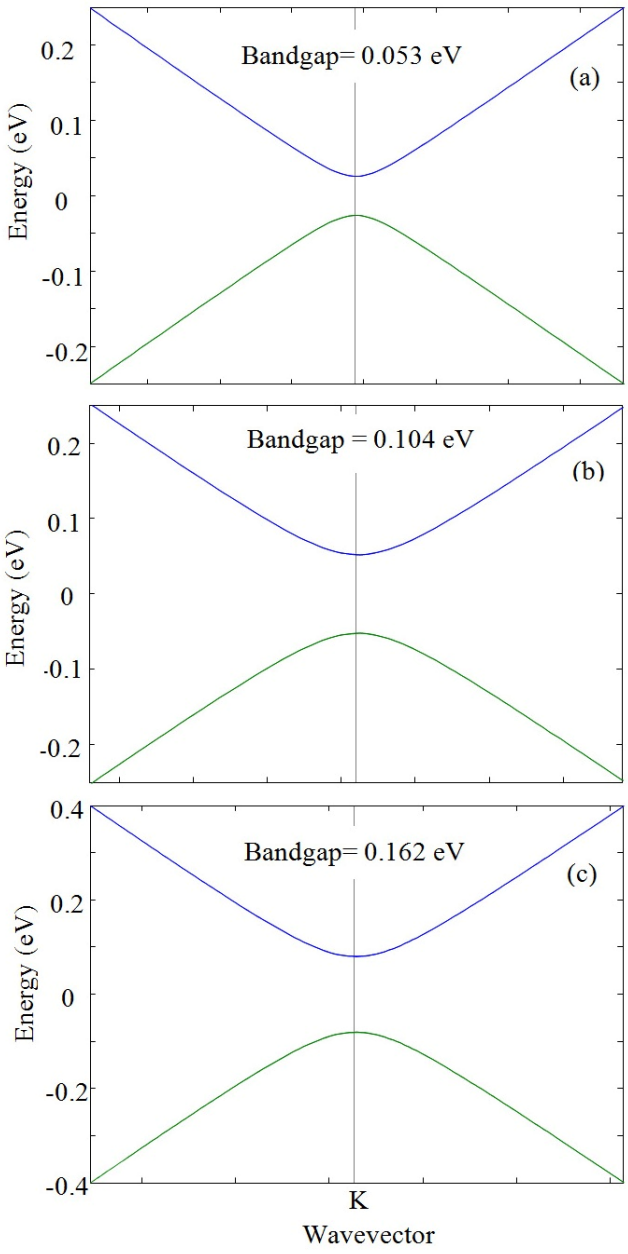



Figure 5. Band gap opening under the application of uniaxial strain perpendicular to C-C bonds (a) 2.84 % strained ($L_x$=0.1268 nm and $L_y$= 0.2126 nm) (b) 4.79% strained ($L_x$= 0.1292 nm and $L_y$=0.2120 nm) (c) 6.5 % strained ($L_y$=0.1313 nm and $L_x$= 0.2115 nm).

## III. CONFIRMATION OF THE PROPOSED METHOD

In order to confirm the proposed method the results estimated in the present study are compared with the published results [14]. A comparison of the results is shown in Table I and Table II. It is found that when strain is applied in parallel or perpendicular to C-C bond the bandgap obtained in both the cases agrees with the results evaluated by first principle method This implies that the proposed method can be applied to estimate the band gap in strained grapheme structure.

Table I

A comparison of the results evaluated by the proposed and by the First principle method [14] under the application of strain parallel to C-C bond.

| $L_y$ (nm) | $L_x$ (nm) | Bandgap (eV) | | Error (%) |
|---|---|---|---|---|
| | | First principle method | Proposed method | |
| 0.2196 | 0.12277 | 0.052 | 0.051 | 1.92 |
| 0.2236 | 0.12245 | 0.107 | 0.106 | 0.94 |
| 0.2296 | 0.12107 | 0.219 | 0.220 | 0.46 |
| 0.2336 | 0.1219 | 0.314 | 0.314 | 0.0 |
| 0.2377 | 0.12172 | 0.442 | 0.431 | 2.49 |
| 0.2396 | 0.12166 | 0.486 | 0.489 | 0.62 |

Table II

A comparison of the results evaluated by the proposed and by the First principle methods [14] under the application of strain perpendicular to C-C bond.

| $L_x$ (nm) | $L_y$ (nm) | Bandgap (eV) | | Error (%) |
|---|---|---|---|---|
| | | First principle method | Proposed method | |
| 0.1252 | 0.2130 | 0.026 | 0.027 | 3.85 |
| 0.1268 | 0.2126 | 0.054 | 0.054 | 0 |
| 0.1278 | 0.2124 | 0.075 | 0.075 | 1.33 |
| 0.1288 | 0.2121 | 0.095 | 0.098 | 4.21 |
| 0.1292 | 0.2120 | 0.103 | 0.107 | 3.88 |
| 0.1303 | 0.2118 | 0.129 | 0.135 | 4.65 |
| 0.1313 | 0.2115 | 0.155 | 0.162 | 4.52 |

## IV. CONCLUSION

In this study an analytical model is proposed to evaluate the energy band gap of strained graphene under the application of planner strain in graphene lattice structure. The proposed model is applied to evaluate the energy gap in symmetrically strained as well as asymmetrically strained graphene structure. Using the model the band gap of symmetrically strained graphene is calculated and it is found to be zero at K-point of the Brillouin zone. Further, the model is applied to determine the band gaps under different planner strains parallel and perpendicular to C-C bonds of graphene lattice structure. The results obtained in both the cases are found to be in good agreement with that of obtained from first principle method .

## REFERENCES


[1] A.K. Geim and K.S. Novoselov, "The rise of graphene ," *Nat. Mater,* vol.6, pp.183-191, 2007.

[2] Ryutaro Sako, Hideaki Tsuchiyaand Matsuto Ogawa, "Influence of bandgap opening on ballistic electron transport in bilayer graphene and graphene nanoribbon FETs," *IEEE Trans. Electronic Devices.*, vol. 58, no. 10, pp. 3300–3306, Oct. 2011.

[3] K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva and A.A. Firsov, *Science* 306, 666 (2004).

[4] K.S. Novoselov, D. Jiang, F.Schedin, T.J. Bhoot, V.V. Khot-kevich, S.V. Morozov and A.K Geim, *Proc.Natl.Acad. Sci*. U.S.A. 102,10451 (2005).

[5] M.Y. Han, B. Ozylmaz, Y. Zhang, and P. Kim, "Energy band gap engineering of grapheme nanoribbons," *Phys. Rev. Lett*. ,vol. 98, no. 20, P. 206805, May 2007.

[6] G. Liang, N. Neophytou, D.E. Nikonov, and M.S. Lundstrom, "Performance projections for ballistic graphene nanoribbon field-effect transistors," *IEEE Trans*. *Electron Devices*, vol. 54, no. 4, pp. 677–682, Apr. 2007.

[7] Y. W. Son, M. L. Cohen, and S. G. Louie, "Energy gaps in graphene nanoribbons," Phys. Rev. Lett., vol. 97, no. 21, p. 216803, Nov. 2006.

[8] X. Li, X. Wang, L. Zhang, S. Lee, and H. Dai, "Chemically derived, ultrasmooth graphene nanoribbon semiconductors," *Science*, vol. 319, no. 5867, pp. 1229–1232, Feb. 2008.

[9] T. Ohta, A. Bostwick, T. Seyller, K. Horn, and E. Rotenberg, "Controlling the electronic structure of bilayer graphene," *Science*, vol. 313, no. 5789, pp. 951–954, Aug. 2006.

[10] Y. Zhang, T.-T. Tang, C. Girit, Z. Hao, M. C. Martin, A. Zettl, M. F. Crommie, Y. R. Shen, and F. Wang, "Direct observation of a widely tunable bandgap in bilayer graphene," *Nature*, vol. 459, no. 7248, pp. 820–823, Jun. 2009.

[11] Jun Ito, Jun Nakamura, and Akiko Natori, "Semiconducting nature of the oxygen-adsorbed graphene sheet," *Journal of applied phys*. 103,113712 (2008).

[12] Paolo Marconcini, Gianluca Fiori, Alessandro Ferretti, Giuseppe Iannaccone, and Massimo Macucci, "Numerical analysis of transport properties ofboron-doped graphene FETs".

[13] Zhipping Xu and Kun Xue, "Engeneering graphene by oxidation: a first principle study," *Nanotechnology* 21 (2010) 045704

[14] Gui Gui, Jin Li, and Jianxin Zhong 'Band structure engineering of graphene by strain: First-principles calculations', *PHYSICAL REVIEW* B 78, 075435 2008.

[15] V.J. Surya, K. Iyakutti, H. Mizuseki, and Y. Kawazoe, "Tuning electronic structure of Graphene: A first-principle calculation," IEEE Trans. Nanotechnology, vol.11, No.3,pp.534-541, May 2012.

[16] Yang Li, Xiaowei Jiang, Zhongfan Liu, Zhirong Liu, "Strain Effects in Graphene and Graphene Nanoribbon: The underlying mechanism," *Nano Res (2010) 3: 545–556*

[17] H.-S Philip Wong, Deji Akinwande, Carbon Nanotube and Graphene Device Physics, Cambridge University Press 2011.

# DC Characteristics of Dual Gated Large Area Graphene MOSFET

Md. Tawabur Rahman*, Ashish Kumar Roy, Hossain Md. Abu Reza Bhuiyan, Md. Tajul Islam and Ashraful G. Bhuiyan
Khulna University of Engineering & Technology, Bangladesh
* E-mail: tawabur_eee06@yahoo.com

***Abstract*— This work reports the DC characteristics of dual gated large area graphene metal oxide semiconductor field effect transistor (MOSFET). The sheet charge density dependent quantum capacitance is obtained self-consistently with considering the impurities concentration of the gate oxide layer. The potential profile as well as sheet charge density of graphene channel is calculated. The *C-V* and *I-V* characteristics are illustrated here. Finally, the velocity- field relation is shown.**

***Keywords*— Graphene MOSFET, GFET, Large area graphene, dual gate effects, self-consistent quantum capacitance.***

## I. INTRODUCTION

Graphene is a flat monolayer of $sp^2$ carbon atoms tightly packed into a two-dimensional (2D) honeycomb lattice including a linear energy dispersion relation [1]. Graphene offers many of the advantages such as high carrier mobilities up to $2,00000\ cm^2\ V^{-1}\ s^{-1}$ in substrate supported devices and high saturation velocity [2]−[6]. The novel electronic properties of graphene lead to intense research into possible applications of this material in nano scale devices such as dual gate graphene MOSFETs. In case of field effect devices, the charge control over conducting channel can be known from *C-V* characteristics clearly. But, graphene is different as a channel material compared to conventional MOSFETs in several respects. It is known from general MOS theory that the gate capacitance per unit area $C_G$ of a MOS structure with a semiconductor having a finite density of states (DOS) cannot be described properly by the oxide capacitance $C_{ox}$ alone. Instead, the gate capacitance has to be considered as a series connection of the oxide capacitance $C_{ox}$ and the quantum capacitance $C_q$ accounting for the limited DOS [7]. Sheet charge density dependent the quantum capacitance was calculated without considering the impurities concentration of the gate oxide layer [7]. Here, we have calculated the sheet charge density dependent quantum capacitance self-consistently considering the impurities concentration of the gate oxide layer [8]-[12]. Also the channel potential as well as sheet charge density variation with channel length are studied here. We have introduced the effects of dual gate on dc characteristics of graphene MOSFET. The *I-V* characteristics of graphene MOSFET show kink-like effect

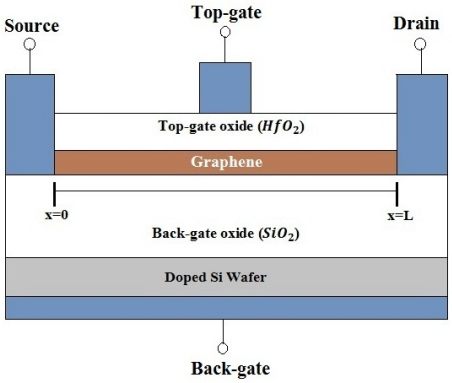



Figure 1. Cross section of the modeled graphene MOSFET.

of ambipolar nature of the conducting graphene channel. The parasitic series resistances are taken into account. The low field electron mobility is considered constant throughout the simulation [13], [14]. The velocity-field characteristics are also illustrated here.

## II. SELF-CONSISTENT QUANTUM CAPACITANCE

A dual gated graphene MOSFET is considered for our work shown in Fig.1. Graphene grown on metal and transferred to a $SiO_2$ covered Si wafer is used as the channel of the MOSFET. The length and width of the graphene channel are 5 $\mu m$ and 1 $\mu m$ respectively. Here, $HfO_2$ ($k$=16) is used as top-gate oxide and $SiO_2$ ($k$=3.9) is used as back-gate oxide [5], [7]. The quantum capacitance of graphene channel depends on the sheet charge density of the channel. The overall net mobile sheet charge density $Q_{sh}$ is simply the difference between the hole and electron sheet densities multiplied by the elementary charge [7].

$$Q_{sh} = q\times(p-n) \quad (1)$$

$$Q_{sh} = q\frac{2}{\pi(\hbar v_F)^2}\int_0^\infty E\left(\frac{1}{\exp\left(\frac{E+E_F}{k_B T}\right)+1} - \frac{1}{\exp\left(\frac{E-E_F}{k_B T}\right)+1}\right) dE \quad (2)$$

The quantum capacitance is defined as the derivative of the net channel charge density $Q_{sh}$ with respect $V_{ch}$,

$$C_q = -\frac{dQ_{sh}}{dV_{ch}} \quad (3)$$

The expression for quantum capacitance was derived based on a two dimensional electron gas (2DEG) model [8]

$$C_q = \frac{2q^2 k_B T}{\pi(\hbar v_F)^2} \ln\left[2\left(1 + cosh\left(\frac{qV_{ch}}{k_B T}\right)\right)\right] \quad (4)$$

Where, $v_F \cong c/300 = \frac{3\times10^8}{300} = 1\times10^6$ ms$^{-1}$ $= 1\times10^8$ cms$^{-1}$ is the Fermi velocity of the Dirac electron and $V_{ch} = E_F/q$ is the potential of the graphene. Here, $k_B$ is the Boltzmann constant, ħ is the reduced Plank's constant and $T$ is the Kelvin temperature. When $qV_{ch} \gg k_B T$, then (4) reduces to as follows [7], [8].

$$C_q \cong \frac{2q^2}{\pi}\frac{q|V_{ch}|}{(\hbar v_F)2} = \frac{2q^2}{\sqrt{\pi}\hbar v_F}\sqrt{n} \quad (5)$$

Where, $n = \frac{1}{\pi}(\frac{qV_{ch}}{\hbar v_F})^2$ is the carrier concentration of the graphene channel. The quantum capacitance has a minimum value at the Dirac point ($C_{q,min}$) [8] which is close to zero. If the value of $V_{ch}$ is zero then from (4) $\cosh(0) \cong 1$. So the minimum value of $C_q$ is as follows:

$$C_{q,min} = \frac{2q^2 k_B T}{\pi(\hbar v_F)2}\, ln\,(4) \quad (6)$$

At room temperature $T$ = 298K the value of $C_{q,min}$ = 0.8358 μFcm$^{-2}$ and at $T$ = 300K then the value of $C_{q,min}$ = 0.8414 μFcm$^{-2}$ *i.e.* approximately $C_{q,min} \cong 0.8$ μFcm$^{-2}$. The theoretical model is based on the assumption of pure and perfect graphene. In reality, various impurities and defects exist in the oxide substrate. The recent theoretical and experimental results have shown that charged impurities have a key role in the transport properties of graphene near the Dirac point. It has been reported that charged impurities in substrates cause local potential fluctuations and electron or hole puddles in graphene. Additional carrier density $n^*$ is induced by the charged impurities. This additional carrier density should take into account by expressing the total carrier concentration as [8]:

$$n = |n_g| + |n^*| \quad (7)$$

Where, $n_g$ and $n^*$are the carrier concentrations caused by the gate potential and the charged impurities respectively. Combining (5) and (7) we can calculate the quantum capacitance of graphene as a function of the graphene potential and the different impurity concentration of the oxide substrate which is acts as the insulator of a GFET.

$$C_q = \frac{2q^2}{\pi(\hbar v_F)2}\frac{qV_{ch}}{} = \frac{2q^2}{\sqrt{\pi}\,\hbar v_F}\sqrt{|n_g| + |n^*|} \quad (8)$$

Where, $n_g = (\frac{qV_{ch}}{\sqrt{\pi}\hbar v_F})^2$ is the carrier concentration due to gate voltage. The net sheet charge density is calculated by using (3) as [7]

$$Q_{sh} = -\int C_q d\mathrm{V_{ch}} = -\frac{1}{2}C_q V_{ch} \quad (9)$$

The top gate oxide capacitance $C_{ox\text{-}top}$ and back gate oxide capacitance $C_{ox\text{-}back}$ are added with the quantum capacitance $C_q$ as shown in equivalent circuit in Fig. 2.

Here, $V(x)$ is the voltage drop in the channel. It is zero at $x = 0$ and equal to the drain-source voltage $V_{ds}$ at $x = L$. The equation of the potential across the quantum capacitance $V_{ch}$ is derived in [7].

$$V_{ch} = (V_{gs-top} - V(x) - V_{gs-top,0})\frac{C_{ox-top}}{C_{ox-top}+C_{ox-back}+\frac{1}{2}C_q} + (V_{gs-back} - V(x) - V_{gs-back,0})\frac{C_{ox-back}}{C_{ox-top}+C_{ox-back}+\frac{1}{2}C_q} \quad (10)$$

Since $C_q$ is not constant but depends on $V_{ch}$ and on the other hand $V_{ch}$ depends on $C_q$. In this case (8) and (10) have to be solved self-consistently.

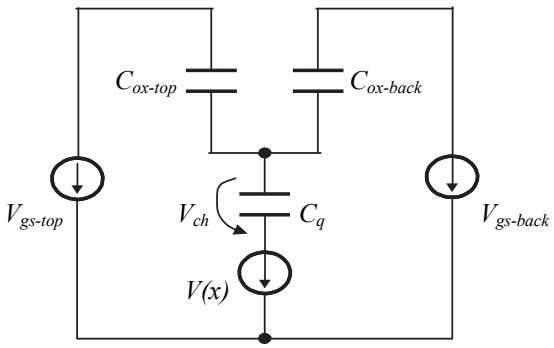



Figure 2. Equivalent capacitive circuit of a graphene MOSFET with both top-gate and back-gate. Here, $C_{ox\text{-}top}$, $C_{ox\text{-}back}$ are the top-gate and back-gate oxide capacitance respectively.

## III. CHANNEL POTENTIAL CALCULATION

The channel potential of dual gated, large area graphene MOSFETs can be calculated in different techniques. We have followed the model [15]. The main feature of this model is to feed a drain current into the device and applying the top-gate and back-gate voltages to calculate the resulting channel potential as well as drain source voltage [15]. Here, only the drift current is considered and the effect of the self-consistent quantum capacitance is taken into account using (8) and (10). First, the channel length is divided into several equidistant segments as shown in Fig. 3. The length of the small segment is $\Delta x = L/N$, where $L$ is the channel length and $N$ is the total number of small segments. Initially a drain current $I_d$ is feed into this model and top-gate voltage $V_{gs\text{-}top}$, back-gate voltage $V_{gs\text{-}back}$ is applied and using the parameters in Table I the channel potential is calculated. Since the current is negative the channel acts as a p-type MOSFET. To get the channel potential $V(x)$ which is dependent on the position of the channel length $x$ from *0* to $L$, the parasitic source-drain resistances (*i.e.* $R_s$ and $R_d$) are not considered. At position $x=0$ the initial values $V(x)$, $\rho(x)$, $v(x)$, $E(x)$ are calculated. At $x=0$, the channel potential is $V(x=0)=V(i=0)=0$. It is mentioned that $x = i\times\Delta x$, where $i$ = *0* to $N$ *i.e.* the number of the values of $i$ is $N+1$. Using $V(x)=0$ the self-consistent quantum capacitance $C_q(i\Delta x=0)$ is calculated as mentioned in section II. Then potential across the quantum capacitance $V_{ch}(i\Delta x=0)$ is calculated by using (10). The corresponding sheet charge density $\rho_{sh}(i\Delta x=0)$, saturation velocity $v_{sat}(i\Delta x=0)$ are calculated by

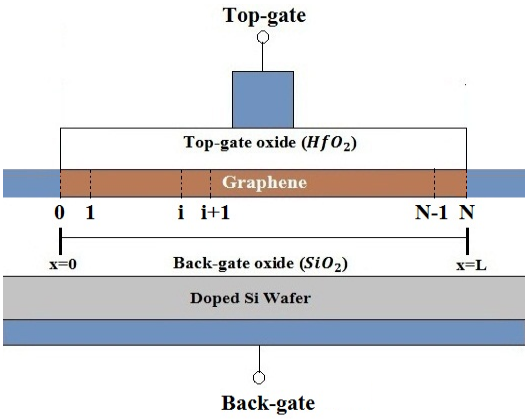



Figure 3. Cross section of the discretized Graphene channel.

using (11), (13) respectively. Due to the current continuity the drift velocity $v(i\Delta x=0)$, electric field $E(i\Delta x=0)$ are calculated as the following equations [15].

$$\rho_{sh}(i\Delta x=0) = \left|-\frac{1}{2q}C_q(i\Delta x=0)\,V_{ch}(i\Delta x=0)\right| + \rho_{sh,0} \quad (11)$$

$$v(i\Delta x=0) = \frac{I_d}{qW\rho_{sh}(i\Delta x=0)} \quad (12)$$

$$v_{sat}(i\Delta x=0) = \frac{2\hbar\Omega}{\pi^2\hbar^2 v_F\rho_{sh}(i\Delta x=0)}\sqrt{\pi(\hbar v_F)^2\rho_{sh}(i\Delta x=0) - \left(\frac{\hbar\Omega}{2}\right)^2} \quad (13)$$

$$E(i\Delta x=0) = \frac{|v(i\Delta x=0)|\times v_{sat}(i\Delta x=0)}{\mu\sqrt{v_{sat}(i\Delta x=0)^2 - v(i\Delta x=0)^2}} \quad (14)$$

The channel potential in the next segment *i.e.* $i=i+1$ is calculated by:

$$V[(i+1)\Delta x = \Delta x] = V(i\Delta x=0) - E(i\Delta x=0)\times \Delta x \quad (15)$$

This algorithm is repeated by calculating the sheet charge density at $x=(i+1)\times\Delta x$. Thus applying this procedure until $i=N$ and the total channel potential $V(N\Delta x)$ is calculated. If the drift velocity $v(i\Delta x)$ exceeds the saturation velocity ($2v_F/\pi$) then the sheet charge density $\rho(i\Delta x)$ will be too small to guarantee current continuity. In this condition the drift velocity is set to maximum saturation velocity and sheet charge density is set to as follows [15].

$$v(i\Delta x) = v_{sat} = 2v_F/\pi \quad (16)$$

$$\rho(i\Delta x) = \frac{I_d}{qWv_{sat}} = \frac{\pi I_d}{2qWv_F} \quad (17)$$

When $v(x) = v_{sat}(x)$ the electric field $E_{sat}$ would be infinitely large and $E_{sat}$ is calculated by

$$E(i\Delta x) = a\frac{v(i\Delta x)}{\mu} = a\frac{2v_F}{\pi\mu} \quad (18)$$

Where, $a$ is an empirical factor such as *4* or *5* used only in the region of constant electric field at $v(x) = v_{sat}(x)$.

## IV. CURRENT VOLTAGE CALCULATION

The current-voltage characteristics are found by following both the section II and section III approximation. Here, $I_d - V_{ds}$ characteristics of graphene MOSFETs are found by setting channel potential at source end with $I_dR_s$ at $x = 0$ [15].

$$V(x=0) = V(i=0) = I_dR_s \quad (19)$$

By applying the procedure of Section III until $i = N$ and adding potential $I_d R_d$ to get the drain-source voltage $V_{ds}$ as

$$V_{ds} = V(N\Delta x) + I_d R_d \quad (20)$$

Now the drain current and gate voltage (*i.e.* top gate voltage or back-gate voltage) characteristics are calculated for different values of drain-source voltages. Here, the given drain source voltage $V_{ds\text{-}const}$ is mentioned with the fixed parameters of Table I and a condition is applied such that if drain source voltage by (20) is same as the $V_{ds\text{-}const}$ then this current $I_d$ is the desired current for the gate voltage. Such the $I_d$-$V_{gs\text{-}top}$ or $I_d$-$V_{gs\text{-}back}$ characteristics can be found.

TABLE I. PARAMETERS OF GRAPHENE CAPACITANCE MOSFET

| Parameters with Units | Parameters Values |
|---|---|
| L(μm) | 5 |
| W (μm) | 1 |
| $t_{ox\text{-}top}$ (nm) | 15 |
| $t_{ox\text{-}back}$ (nm) | 285 |
| $V_{gs-top,0}$(V) | 1.45 |
| $V_{gs-back,0}$(V) | 2.70 |
| $\rho_{sh,0}$ ($cm^{-2}$) | $1.5\times10^{12}$ |
| $R_s = R_d$ (Ω) | 900 |
| $n^*$($cm^{-2}$) | $8\times10^{11}$ |
| $\mu_p = \mu_n$ ($cm^{-2}V^{-1}s^{-1}$) | 1500 |
| ħΩ (meV) | 55 |

## V. RESULTS AND DISCUSSION

### A. C-V Characteristics:

The effect of the quantum capacitance on the overall gate capacitance is investigated. First, the capacitances of a graphene MOSFET with a 3-nm $SiO_2$ top- gate dielectric, $4\times10^{11}$ $cm^{-2}$ impurities concentration $n^*$, zero $V_{gs-top,0}$ and zero applied $V_{ds}$ *i.e.* without feeding current into the model is considered. The quantum capacitance and top-gate oxide capacitance (1.15 μF$cm^{-2}$) affects the gate capacitance $C_g$ as a series combination of $C_q$ and $C_{ox-top}$. The minimum value of the quantum capacitance was found as 1.7325 μF$cm^{-2}$ at zero top-gate voltage and the maximum value was found as 23.5009 μF$cm^{-2}$ for both top-gate voltage −1V and +1V also at Fig. 4. The shape of the quantum capacitance is fully symmetrical on the both side of the Dirac point *i.e.* at the point where the minimum value of $C_q$ is found. The quantum capacitances are shown for different values of the impurities concentrations $n^*$ from $1\times10^{11}$ $cm^{-2}$ to $10\times10^{11}$ $cm^{-2}$ in Fig. 5 [10], [12]. With the increasing value of $n^*$ the minimum value of the quantum capacitance is also increasing. But the shape of $C_q$-$V_{gs\text{-}top}$ curve is unchanged

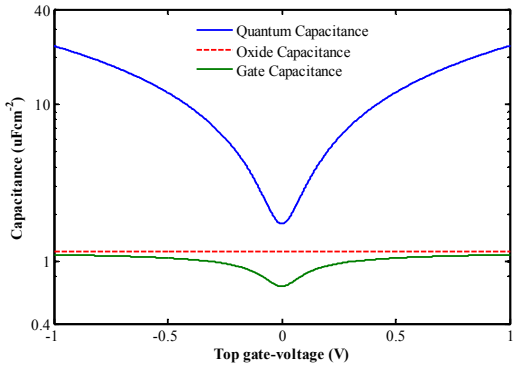



Figure 4. Quantum capacitance $C_q$, top-gate oxide capacitance $C_{ox\text{-}top}$ and gate capacitance $C_g$ of a graphene MOSFET structure as a function of top-gate voltage.

*i. e.* symmetrical with the top-gate voltage. By taking the different values of $n^*$ the $C_q$ are found such as shown in Table II.

TABLE II. QUANTUM CAPACITANCE WITH IMPURITIES CONCENTRATIONS

| **Impurities concentrations, $n^*$ ($10^{11}$ cm$^{-2}$)** | **$C_{q,min}$ (μFcm$^{-2}$) ($V_{gs\text{-}top} = 0V$)** | **$C_{q,max}$ (μFcm$^{-2}$) ($V_{gs\text{-}top} = \pm 1\ V$)** |
|---|---|---|
| 1 | 0.866 | 23.452 |
| 2 | 1.225 | 23.468 |
| 4 | 1.732 | 23.500 |
| 8 | 2.450 | 23.564 |

Thus the slopes of the curves are about same shape with respect to both sides of the top-gate voltage. The gate capacitance of a dual gated graphene MOSFET is the series combination of $C_q$, $C_{ox\text{-}top}$, and $C_{ox\text{-}back}$

$$\frac{1}{C_g} = \frac{1}{C_{ox-top}} + \frac{1}{C_{ox-back}} + \frac{1}{C_q} \tag{21}$$

The total sheet charge density $\rho_{sh}$ has the effect of residual sheet carrier concentration $\rho_{sh,0}$ [7]

$$\rho_{sh} = |-\frac{1}{2\,q} C_q V_{ch}| + \rho_{sh,0} \tag{22}$$

We considered $\rho_{sh}$ by (22) throughout the model simulation.

The gate oxide capacitance $C_{ox\text{-}top}$ or $C_{ox\text{-}back}$ is defined as:

$$C_{ox} = \frac{\epsilon_0 \epsilon_{ox}}{t_{ox}} \tag{23}$$

Where, $t_{ox}$ is the gate oxide thickness and $\epsilon_{ox}$ is the dielectric constant of the gate oxide material.

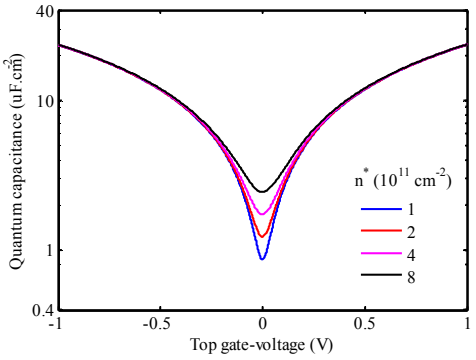
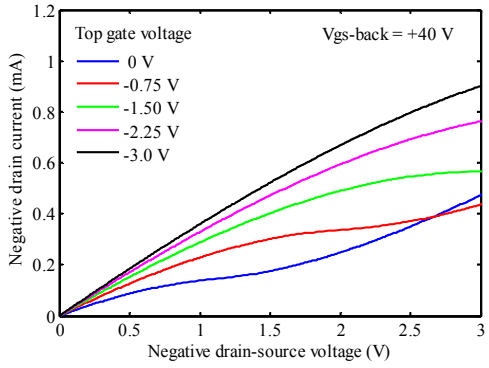



Figure 5. Quantum capacitance $C_q$ with different values of the impurities concentrations $n^*$ in the oxide substrate of a graphene MOSFET structure as a function of top-gate voltage.

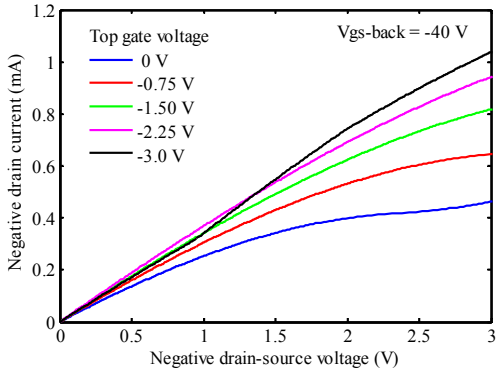
## *B. I-V Characteristics:*

Fig. 6(a) and 6(b) shows the *I-V* characteristics of the graphene MOSFET for different $V_{gs\text{-}top}$ at two back gate bias conditions, $V_{gs\text{-}back}$ of +40V and −40V respectively.

Fig. 6(a) shows that the drain current $I_d$ is increasing with increasing drain to source voltage $V_{ds}$ for different top gate voltages at constant back gate bias $V_{gs\text{-}back}$=$+40V$. The similar curves are displayed on Fig. 6(b) for $V_{gs\text{-}back}$= $-40V$.



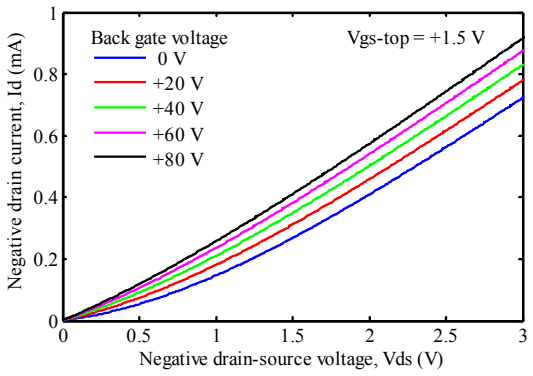

(a)



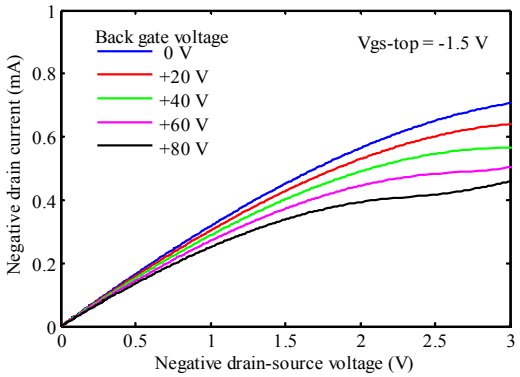

(b)

Figure 6. (a) The drain current $I_d$ as a function of drain source voltage $V_{ds}$ for four different values of $V_{gs-top}$ at $V_{gs\text{-}back}$ = +40 V (b) The drain current $I_d$ as a function of drain source voltage $V_{ds}$ for four different values of $V_{gs-top}$ at $V_{gs\text{-}back}$ = -40 V.

The variation of gate voltage directly affects the channel potential, $V_x$. The quantum capacitance, $C_q$ and the graphene potential, $V_{ch}$ are varied self-consistently with $V_x$ which influences the sheet charge density, $\rho_{sh}$. Thus the drain current, $I_d$ which is a function of gate voltage ($V_{g\text{-}top}$ or $V_{g\text{-}back}$) and sheet charge density, $\rho_{sh}$ is also changed. The $I_d$ curve from Fig. 6(a) for $V_{gs\text{-}top}=0.0$ $V$, which shows a pronounced 'kink' in the characteristics signify the presence of an ambipolar channel. For our device in Fig. 1, with channel length $L$, $V(L)=V_{ds}$ so that for $V_{ds} \leq V_{ds\text{-}kink} \approx V_{gs\text{-}top}-V_0$, current is carried by holes throughout the length of the channel [5]. For $V_{ds} = V_{ds\text{-}kink}$, the vanishing carrier density produces a 'pinch-off ' region at the drain that renders the current in the channel relatively insensitive to $V_{ds}$ and results in the pronounced kink seen in the $I$–$V$ characteristics. The $V_{ds}>V_{ds\text{-}kink}$, gives rise to an accumulation of electrons in the drain side and a corresponding increase in the carrier density leading to a further increase in the drain current. Thus there produces an ambipolar regime in the graphene channel. The pinch-off point becomes a place of recombination for holes flowing from the source and electrons flowing from the drain. Because there is no band gap, no energy is released in this recombination [5].



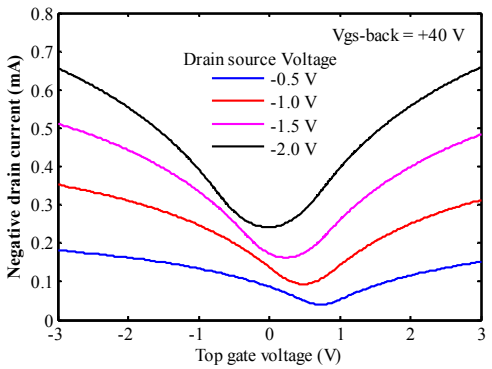

(a)



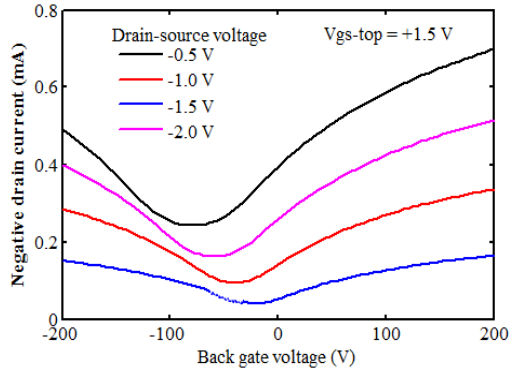

(b)

Figure 7. (a) The drain current $I_d$ as a function of drain source voltage $V_{ds}$ for four different values of $V_{gs-back}$ at $V_{gs\text{-}top}$ = +1.5 V. (b) The drain current $I_d$ as a function of drain source voltage $V_{ds}$ for four different values of $V_{gs-back}$ at $V_{gs\text{-}top}$= -1.5 V.

Fig. 7(a) and 7(b) shows the $I$-$V$ characteristics of the graphene MOSFET for different back-gate voltages at two top gate bias conditions, of +1.5 V and −1.5 V respectively.




(a)




(b)

Figure 8. (a) The drain current $I_d$ as a function of top-gate voltage $V_{gs\text{-}top}$ for four different values of $V_{ds}$ at $V_{gs-back}$ = +40V. (b) The drain current $I_d$ as a function of back-gate voltage for four different values of $V_{ds}$ at $V_{gs-top}$ = +1.5V.

Fig. 8(a) and 8(b) shows the transfer characteristics. In both figures, it is seen that the Dirac point shifts when the drain-source voltage is varied.

Fig. 9 shows the simulation of potential on graphene channel with fixed $V_{ds} = -2.5V$ , $V_{gs-top} = -1.5V$ and $V_{gs-back} = +40\,V$ . At the source end *i.e.* at x=0, channel potential is zero. As the channel length increases the corresponding channel potential increases and becomes equal to drain voltage at the drain end voltage equal to −2.5 V.

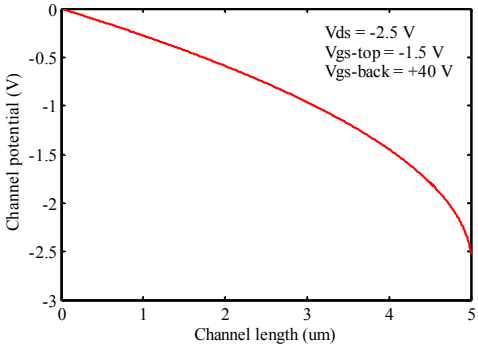



Figure 9. Potential profile of graphene Channel

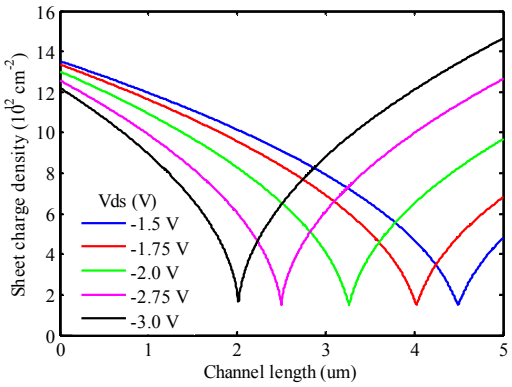



Figure 10. Variation of sheet charge density with channel length at $V_{gs\text{-}top} = -1.5\ V$, $V_{gs\text{-}back} = +40\ V$ with different $V_{ds}$.

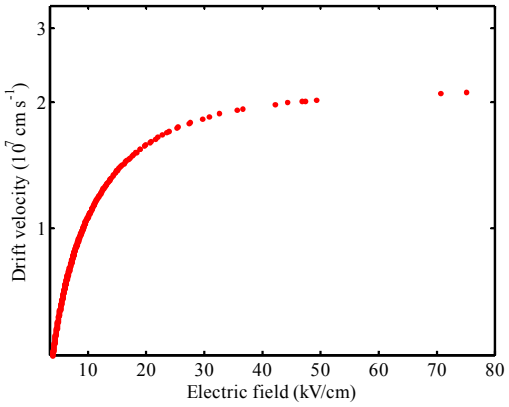



Figure 11. Velocity-field relation for graphene MOSFET at $V_{gs\text{-}top} = -1.5\ V$, $V_{gs\text{-}back} = +40\ V$ when $V_{ds} = -3.0\ V$.

Fig. 10 shows the variation of sheet charge density with channel length at different $V_{ds}$. As $V_{ds}$ is decreased negatively the corresponding sheet charge density decreases up to Dirac point. After Dirac point we obtain a positive gate to channel voltage at the drain end which gives rise to an accumulation of electrons and corresponding increase in sheet charge density. Now the majority charge carriers are holes at the source end and electrons at the drain end *i.e.* ambipolar conduction occurs.

Fig. 11 shows the velocity-field relation for graphene MOSFET. It can be seen that the drift velocity increases linearly with the applied electric field and becomes saturated ($2.12\times10^7$ cms$^{-1}$) at 75.06 kV/cm similar to [16], [17].

## VI. CONCLUSION

The simulation of dc characteristics of dual gated graphene MOSFET has been presented here. We have calculated sheet charge density dependent quantum capacitance self consistently considering charged impurities in the gate oxide. With increasing charged impurities concentration the quantum capacitance increases near the Dirac point. A maximum value of quantum capacitance 23.564 μFcm$^{-2}$ is obtained at $V_{gs\text{-}top} = \pm 1\ V$ with the charged impurities concentration of $8\times10^{11}$ cm$^{-2}$. The effects of dual gate voltage on the *I-V* characteristics of graphene MOSFET is illustrated here. The channel potential as well as the sheet charge density on graphene channel is simulated. Sheet charge density of $13.5 \times 10^{12}$ cm$^{-2}$ has been found at $V_{ds} = -1.5\ V$, $V_{gs\text{-}top} = -1.5\ V$, $V_{gs\text{-}back} = +40\ V$. We also try to point out the ambipolar characteristics of graphene channel. Finally, the velocity-field relation is shown.

## REFERENCES


[1] A. K. Geim, K. S. Novoselov, "The rise of graphene", *Nature Materials*, vol. 6, pp. 183 -191, 2007.

[2] F. Schwierz, "Graphene transistors", *Nature Nanotechnology,* vol. 5, pp. 487–496, 2010.

[3] K. S. Novoselov, A. K. Geim, "Two-dimensional gas of mass less Dirac fermions in graphene", *Nature* vol. 438, pp. 197-200, 2005.

[4] F. Schedin, A. K. Geim, "Detection of individual gas molecules adsorbed on graphene",*Nature Materials,* vol. 6, pp. 652–655, 2007

[5] I. Meric, M. Y. Han, A. F. Young, "Current saturation in zero-bandgap, top-gated graphene field-effect transistors", *Nature Nanotechnology*, vol. 3, pp. 654–659, 2008.

[6] K. I. Bolotina, K. J. Sikesb, Z. Jianga, M. Klimac, G. Fudenberga, J. Honec, P. Kima, H. L. Stormer, "Ultrahigh electron mobility in suspended graphene", *Solid State Communications,* vol. 146 , pp. 351–355, 2008.

[7] S. A. Thiele, J. A. Schaefer, F. Schwierz, "Modeling of graphene metal-oxide-semiconductor field-effect transistors with gapless large-area graphene channels", *J. Appl. Phys.,* vol. 107, pp. 094505, 2010.

[8] J. Xia, F. Chen, J. Li, N. Tao, "Measurement of the quantum capacitance of graphene", *Nature Nanotechnology*, vol. 4, pp. 505–509, 2009.

[9] D. L. John, L. C. Castro, D. L. Pulfrey, "Quantum capacitance in nanoscale device modeling," *J. Appl. Phys.*, vol. 96, pp. 5180, 2004.

[10] E. Pince, C. Kocabas, "Investigation of high frequency performance limit of graphene field effect transistors", *Appl. Phys. Lett.,* 97, 173106, 2010.

[11] J. H. Chen, C. Jang, S. Xiao, E. D. Williams, M. S. Fuhrer, "Diffusive Charge Transport in Graphene on $SiO_2$," *Solid State Communications,* vol.149, pp. 1080-1086, 2009.

[12] J. H. Chen, C. Jang, S. Adam, M. S. Fuhrer, E. D. Williams, M. Ishigami, "Charged-impurity scattering in graphene", *Nature Physics*, vol. 4, pp. 377–381, 2008.

[13] V. E. Dorgan, M. H. Bae, E. Pop, "Mobility and saturation velocity in graphene on $SiO_2$", *Appl. Phys. Lett.,*vol. 97, pp. 082112, 2010.

[14] A. Venugopal, J. Chan, X. Li, C. W. Magnuson, W. P. Kirk, L. Colombo, R. S. Ruoff, E. M. Vogel, "Effective mobility of single-layer graphene transistors as a function of channel dimensions", *J. Appl. Phys.*, vol. 109, pp. 104511, 2011.

[15] S. Thiele and F. Schwierz, "Modeling of the steady state characteristics of large-area graphene field-effect transistors", *J. Appl. Phys.*, vol. 110, pp. 034506 -034506-7, 2011.

[16] R S Shishir, D K Ferry, "Velocity saturation in intrinsic graphene", *Journal of Physics: Condensed Matter*, vol. 21, no. 34, 2009 .

[17] B. W. Scott, J. Leburton, "High-field carrier velocity and current saturation in graphene field-effect transistors", *10th IEEE Conference on Nanotechnology (IEEE-NANO),* pp. 655-658, 2010.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Theoretical Investigation of Quantum Capacitance in Armchair-edge Graphene Nanoribbons

Md. Faruque Hossain*, Asif Hassan, Md. Sohel Rana
Department of Electronics & Communication Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh
* E-mail: fhossain97@yahoo.com

*Abstract*—**Graphene nanoribbons (GNRs) are considered as a prospective material for the next generation of nanoelectroic devices. One of the important properties of GNRs in determining the performance of such devices is capacitance; in particular, the quantum capacitance when the device size approaches in the scale of nanometer. This work presents a comprehensive investigation of the bandgap structure and the classical and quantum capacitance in armchair-edge GNRs (A-GNRs) using semi-analytical method. The method is simple and more realistic considering edge effects of A-GNRs. The results show that the edge effects have significant influence in defining the bandgap which is a necessary input in the accurate analyses of capacitance. The classical capacitance is completely determined by the device geometry and a dielectric constant of the medium. The quantum capacitance is obtained considering edge effects and discussed for both degenerate (high gate voltage) and nondegenerate (low gate voltage) regime. It is demonstrated that the total capacitance is equivalent to the classical capacitance in nondegenerate regime, whereas in degenerate regime, quantum capacitance dominates over the classical capacitance. Such detail analysis of GNRs considering a realistic model would be useful for the optimized design of GNR based nanoelectronic devices.**

***Keywords—Graphene nanoribbon, bandgap, classical capacitance, quantum capacitance.***

## I. INTRODUCTION

Graphene, a flat monolayer of carbon atoms packed into a two-dimensional honeycomb lattice, is recognized as a promising candidate for the nanoelectronic devices [1], [2]. It has attracted enormous interest for possessing remarkable electronic properties. In particular, the high mobility, which promises near-ballistic transport and high speed switching, has made it viable alternate to silicon (Si) for the channel of field-effect transistors [3]. Although carbon nanotubes (CNTs) offer better transport properties, graphene are believed to be more controllable from a fabrication point of view. This is due to the planar nature of graphene, which can be patterned using high-resolution lithography.

Generally graphene is metallic or semi-metallic with zero band gap. For semiconductor applications, a band gap can be achieved by using a narrow strip of graphene in the size of nano-scale called graphene nanoribbon (GNR) [4]. In accordance with the shape of edges, GNRs are classified as zigzag-edge (Z-GNR) and armchair-edge (A-GNR) [5]. The semiconducting properties of A-GNRs can be predicted from

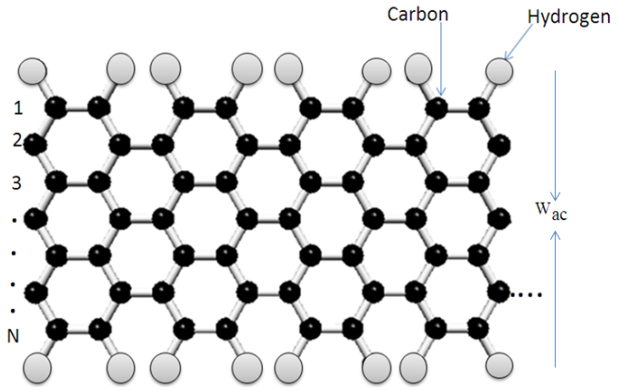



Figure 1. Schematic of a A-GNR. The gray circles denote hydrogen atoms terminating the edge carbon atoms. The number of dimmer line and the ribbon width are represented by $N$ and $W_{ac}$, respectively.

their width which is dependent on its number of dimmer lines ($N$), as shown in Fig. 1. It will be semiconducting for $N = 3p$ or $N = 3p+1$, or $N = 3p+2$, where p is an integer [6]–[9]. In comparison, Z-GNRs are always metallic and independent of $N$ [10].

One of the main properties of controlling the semiconducting transistor is the capacitance formed between the channel and the gate. The capacitance is important for understanding fundamental electronic properties of the material such as the density of states (DOS) as well as the device performance including the $I$-$V$ characteristics and the device operation speed [3]. The capacitance in classical approach is completely determined by the device geometry and a dielectric constant of the medium. However, quantum effects will give prominent impact to the device performance once the transistor approaches the size of nanoscale device and hence the quantum capacitance must be taken into consideration [4], [11]–[13].

A number of works have been found in literature investigating the band gap and quantum capacitance in GNR field-effect transistors [10], [14]–[17]. Many of them consider classical approaches, for example a simple tight binding (TB) model, which have been previously used to simulate Si metal oxide semiconductor field-effect transistors (MOSFETs), CNTFETs, and nanowire FETs [14]–[16]. Such model did not consider edge effects. However, unlike CNTs which do not have edges, the presence of edges in the A-GNRs has a

significant influence on the band gap and thus on the quantum capacitance of GNRFETs. Several approaches have been reported to consider the edge effects, likewise, modified TB model [8], NEGF formalism with self-consistent method [17] and semi-analytical model [4]. This work presents a comprehensive investigation of band gaps and quantum capacitance in A-GNRs considering edge effects. The semi-analytical model which is computationally simple and can produce accurate results is used for the analysis.

An accurate calculation of the band structure is a necessary input for the model used in this paper. We first examine the band gap and band structure in A-GNRs without considering and with considering edge effects. The results show significant influence of edge effects in band gap. Then the quantum capacitance is thoroughly investigated considering edge effects in both degenerate (high gate voltage) and non-degenerate (low gate voltage) regime. The dependency of degeneracy on the gate voltage, which influences the quantum capacitance, is also discussed. Since the graphene technology is at its very early stage, such theoretical analysis can be very useful to get physical insight of GNRFETs and to optimize the design with attainable performance.

## II. THEORY

An armchair-edge graphene nanoribbon can be recognized by the armchair-like pattern that the outermost carbon atoms form along the edges of a narrow strip of graphene. Recent studies have shown that A-GNRs can be semiconducting for $N = 3p$ or $3p+1$ or $3p+2$ depending upon their edges [15]. The width of an armchair ribbon can be defined in terms of the number of dimer lines (N) as [16]

$$W_{ac} = (N-1)\frac{\sqrt{3}}{2}a \tag{1}$$

where $a = 1.42\,\text{A}^0$ is the nearest neighbor distance. Now the relations between band gap energy $E_G$ and width of GNRs can be derived as

$$E_G = -\frac{2\pi(\gamma_1 - 2\gamma_3)}{\sqrt{3}(N+1)} + \frac{6(\gamma_3 + \Delta\gamma_1)}{N+1} \tag{2}$$

where $\gamma_1 = -3.2$ eV is the first nearest neighbor hopping parameter, $\gamma_3 = -0.3$ eV is the third nearest neighbor hopping parameter, $\Delta_{\gamma_1} = -0.2$ eV is the correction of $\gamma_1$ for the bonds due to the edge bond relaxation effect [14].

The energy dispersion relation can be expressed as [4]

$$E = \sqrt{\left(\left(\frac{\Delta}{2}\right)^2 + (\hbar\gamma_s k)^2\right)} \tag{3}$$

where $\hbar$ is the reduced plank constant,

$$\frac{\Delta}{2} = \gamma_1\left(2s\cos\frac{p\pi}{N+1} + 1\right) + \gamma_3\left(2s\cos\frac{2p\pi}{N+1} + 1\right) + 4\frac{(\gamma_3 + \Delta\gamma_1)}{N+1}\sin^2\frac{p\pi}{N+1}$$

and

$$(\hbar\gamma_s)^2 = (3d)^2\left\{-\frac{1}{2}\gamma_1 s\cos\frac{p\pi}{N+1}\times\left[\gamma_1 + \gamma_3\left(2\cos\frac{2p\pi}{N+1} + 1\right) + 4\frac{(\gamma_3 + \Delta\gamma_1)}{N+1}\sin^2\frac{p\pi}{N+1}\right] - \gamma_3\left(\gamma_1 + 2\gamma_3\cos\frac{p\pi}{N+1} + 4\frac{(\gamma_3 + \Delta\gamma_1)}{N+1}\sin^2\frac{p\pi}{N+1}\right)\right\}.$$

Capacitance is one of the important characteristics in semiconducting device. The gate voltage $V_g$ which is applied to metallic gate induces extra carrier with a density *n,* where another extra carrier is produced because of the opposite polarity into gate itself. Thus, the gate voltage $V_g$ is the summation of onsite potential $V_s$ and the voltage drop over the gate oxide $V_{ox}$ ,

$$V_g = V_s + V_{ox} \tag{4}$$

If the electrostatic potential on the GNR is uniform, then the voltage drop over the gate oxide is also uniform, and the following expression can be obtained from (4)

$$\frac{dV_g}{dQ} = \frac{dV_s}{dQ} + \frac{dV_{ox}}{dQ} \tag{5}$$

If one defines the quantum capacitance $C_{QND} = dQ/dV_s$ and the classical capacitance as $C_{ins} = dQ/dV_{ox}$ , equation (5) becomes

$$\frac{1}{C_G} = \frac{1}{C_{QND}} + \frac{1}{C_{ins}} \tag{6}$$

This indicates that the gate capacitance is the serial combination of the quantum capacitance and the classical capacitance.

An electrostatic potential $V_{ox}$ shows classical behavior when the density of states is infinite and thus $V_s = 0$. This definition describes the classical capacitance which can be denoted as [8]

$$C_{ins} = N_G\varepsilon_o k\left(\frac{W_{ac}}{t_{ox}} + \alpha\right) \tag{7}$$

where $N_G$ is the number of gates, $k$ is the relative dielectric

constant of the gate insulator, $\varepsilon_0$ is the permittivity of air, $t_{ox}$ represents the gate insulator (oxide) thickness and $\alpha = 1$ is a dimensionless fitting parameter [8].

The quantum capacitance is simply denoted as a differential change of charge density with respect to electrostatic potential $V_s$ as

$$C_{QND} = \frac{dQ}{dV_s} = e\frac{\partial n}{\partial V_s} = e^2 D_o \tag{8}$$

where $D_o$ is the weighted average of the density of states (DOS) and $e$ is the charge of the carrier.

$$D_o = \int_{-\infty}^{\infty} dE.D(E + eV_s)[-\frac{df(E)}{dE}] \tag{9}$$

where $D(E + eV_s)$ is the density of states (DOS) of GNR which is shifted by electrostatic potential $V_s$ . The Fermi probability function is given by

$$f(E) = \frac{1}{\exp\left(\frac{E - E_F}{k_B T}\right) + 1} \tag{10}$$

It is stated that under thermal equilibrium the Fermi level is unique and exactly at the Dirac point $E_F = 0$.

So, the general expression of quantum capacitance can be obtained from (8) as [17]

$$C_{QND} = \frac{e^2}{3ta\pi}\left(\frac{x + \frac{E_G}{2k_B T}}{\sqrt{x + x\frac{E_G}{k_B T}}}\right) f(E) \tag{11}$$

where $k_B$ = $1.38\times10^{-23} J/K$ (Boltzmann constant), $e = 1.6\times10^{-19}$ Columb, $T$ = Room temperature, $t$ = 2.7 eV, hopping integral of perfect GNR [10], and $x = E - E_G/2$ .

If local electrostatic potential in a GNR is tuned by a gate voltage, degeneracy can be determined when the Fermi level ( $E_F = eV_s$ ) will located within the conduction band. In degenerate regime, $E - E_F < 3k_B T$ and therefore, the Fermi function can be approximated as $f(E) = 1$. On the other hand, for a nondegenerate condition when the Fermi level is located inside the GNR bandgap, the condition is given by $E - E_F > 3k_B T$. Therefore, the Fermi function becomes $f(E) = \exp[(E_F - E)/k_B T]$. Since the Fermi level is not accessible experimentally, the quantum capacitance in this regime can be calculated considering reduced Fermi function $f(E) = \exp(\eta)$ in (11), where the reduced Fermi function is related to gate voltage, $V_g$ by [16]

$$\eta = \frac{eV_g - eV_T}{k_B T} \tag{12}$$

Here, $V_T$ is the threshold voltage. In the following section, the quantum capacitance will be investigated as a function of the gate voltage as shown in (12).

## III. Results and Discussion

### A. Width of GNR

The electronic properties of GNRs can be controlled by varying the width of the ribbon. The width of the GNRs is derived by their number of dimer lines which is obtained from equation (1) and shown in Fig. 2. The width of GNR increases linearly with the increases of dimer line.

### B. Bandgap Energy

Fig. 3 plots the bandgap energy ($E_G$) as a function of the width ($W_{ac}$) of A-GNR for $N$ = 3$p$. The bandgap energy is calculated using (2) for two cases: without considering edge effect (model 1) and with considering edge effect (model 2). In model 1, modeling parameters are $\gamma_1$ = -3.2 eV, $\gamma_3$ = 0 eV, $\Delta_{\gamma_1}$ = 0 eV. The corresponding energy dispersion curves ($E$-$k$ diagram) for N = 21 are shown in Fig. 4.

As shown in Fig. 3, the A-GNRs are semiconductor with energy gaps which is inversely proportional to the channel width. In order to go to the conduction band from valence band the electron must have to achieve an amount of energy which is equal to the bandgap energy. However, the energy gaps obtained without considering edge effects (model 1) are quite different from those with considering edge effects. Consideration of edge effects produces lower value of bandgap for a given width. This difference indicates the importance of considering edge effects in the analysis of GNR based devices.

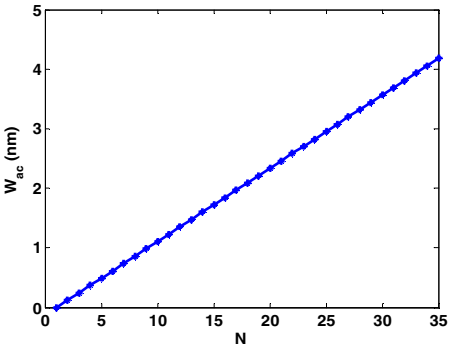



Figure 2. Width of armchair nanoribbon ($W_{ac}$) in nm vs. number of dimer line ($N$).

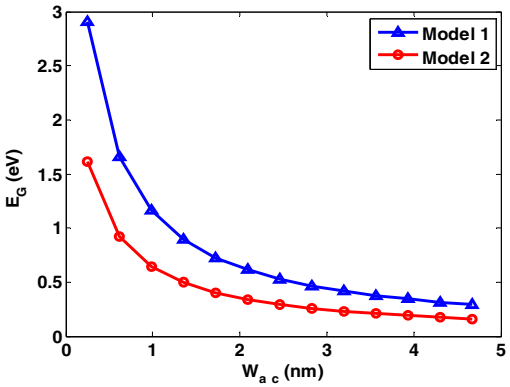



Figure 3. Bandgap Energy ($E_G$) vs. width ($W_{ac}$) of A-GNR with $N=3p$ for model 1 (without considering edge effects) and model 2 (with edge effects).

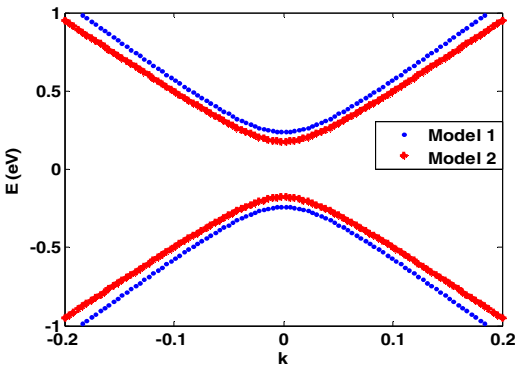



Figure 4. First sub-band structure of A-GNRs using $N=21$ for model 1 (without considering edge effects) and model 2 (considering edge effects).

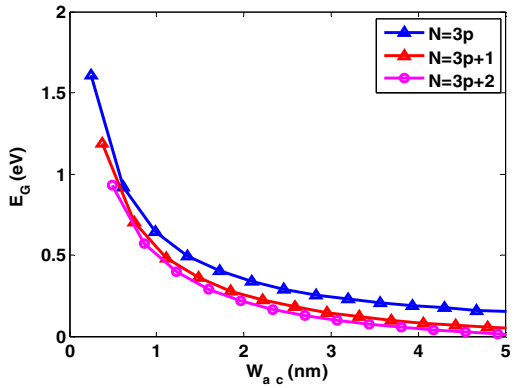



Figure 5. Bandgap Energy ($E_G$) as a function of width ($W_{ac}$) of A-GNR in case of *N=3p, N=3p+1 and N=3p+2* considering edge effects (model 2).

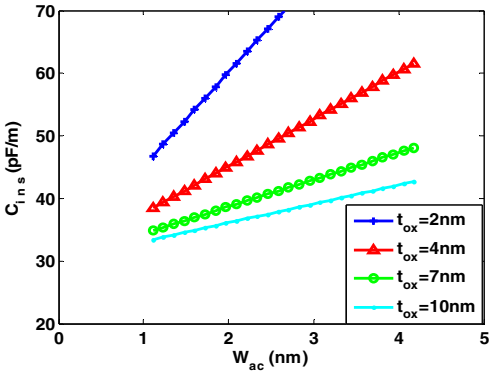



Figure 6. Classical capacitance ($C_{ins}$) vs. A-GNR width ($W_{ac}$) at different oxide thickness ($t_{ox}$) of 2 to 10 nm for a single gate GNRFET.

Fig. 5 shows the bandgap energy ($E_G$) as a function of the width ($W_{ac}$) for three different conditions of dimer line as *N=3p, N = 3p+1 and N = 3p+2*. The variations in energy gap exhibit three distinct family behaviors. According to the simple TB model (without considering edge effects), the A-GNR with $N = 3p+2$ is metallic [10]. However, considering edge effects it is found to be semiconductor and among three conditions of dimmer lines, this gives the lowest bandgap for a given width. This result is consistent with the previous works found in literature [14].

From Fig. 5, it is observed that $N = 3p+2$ shows metallic behavior ($E_G = 0$) if the width increases above few nm (~ 4 nm). So it is fruitful to take *N=3p, N=3p+1* structure based device to work as a transistor since these give higher bandgap energy with designable width.

### C. Classical capacitance

The effects of classical capacitance (i.e., gate insulator capacitance) give knowledge about the storage of charge with respect to electrostatic potential $V_{ox}$ . It is proportional to device GNR width ($W_{ac}$), insulator thickness ($t_{ox}$) and dielectric constant. Fig. 6 shows the gate insulator capacitances ($C_{ins}$) achieved from (7) for a single gate GNRFET, $k = 3.8$, and the oxide (gate insulator) thickness $t_{ox}$ of 2 to 10 nm.

The classical capacitance ($C_{ins}$) increases linearly as the GNR width increases as shown in Fig 6. This is due to the proportional increase in area of A-GNR with the increasing width. The variation of $C_{ins}$ with the oxide thickness ($t_{ox}$) is shown in Fig. 7 for two different widths of 2.09 nm and 4.18 nm. It shows that for a given width of GNR the $C_{ins}$ vary nonlinearly with the $t_{ox}$. The capacitance decreases with the increase in oxide thickness and this influence is more significant for lower value of oxide thickness.

### D. Quantum capacitance

As the device structures lessening quantum capacitance becomes dominant and therefore must be taken into account. As discussed in section II, the Fermi level defines the quantum capacitance in two states: degenerate regime and nondegenerate regime. These two types give clear understanding of a nanoscale device from device turn-off to turn on. The quantum capacitance ($C_{QND}$) in response to gate voltage ($V_g$) is shown in Fig. 8. Here, the band gap energy is considered as $E_G = 0.34$ eV corresponding to $W_{ac} = 2.34$ nm with $N = 3p$ and $E = 1$eV. The value of $C_{QND}$ for general case is calculated directly using (11).

As shown in Fig. 8, the capacitance increases with the increasing gate voltage. In general case, the applied gate voltage is increasing means the Fermi level is shifting from in

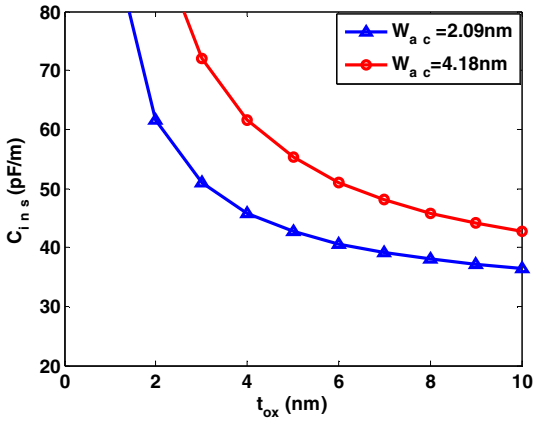



Figure 7. Classical capacitance ($C_{ins}$) vs. insulator thickness ($t_{ox}$) at different A-GNR width ($W_{ac}$) of 2.09 nm and 4.18 nm for a single gate GNRFET.

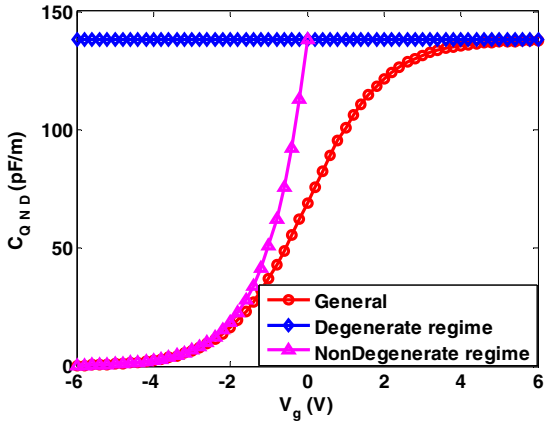



Figure 8. Quantum capacitance ($C_{QND}$) vs. gate voltage ($V_g$) in general case, nondegenerate regime and degenerate regime for $W_{ac}$ = 2.34 nm ($N = 3p$).

between position of bandgap to the conduction band. In degenerate regime, the concentration of electrons in the conduction band exceeds the DOS and the Fermi level located inside conduction band. This produces a maximum and constant value (138 pF/m) of quantum capacitance in degenerate regime (high gate voltage). Otherwise, in nondegenerate regime, the Fermi level lies in between the bandgap. So, increasing the negative channel voltage drops down the fermi level signifying the decrease of free electrons and thus the drecrease of quantum capacitance.

Fig. 9 shows the comparison of gate insulator (classical) capacitance and quantum capacitance as a function of gate voltage. It is observed that the classical capacitance dominates the total capacitance at lower value of $V_g$, whereas at higher value of $V_g$ the quantum capacitance dominates the total capacitance.

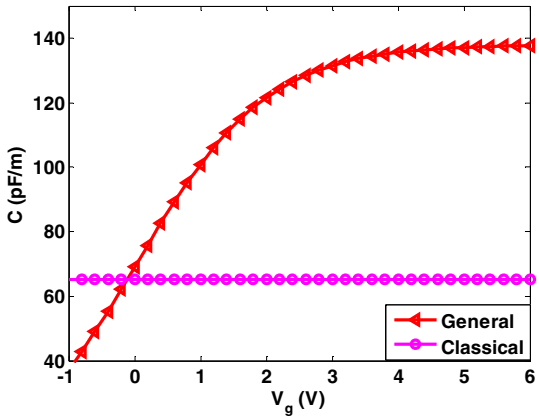



Figure 9. Capacitance (Quantum capacitance in general case and Classical) vs. gate voltage ($V_g$) for ribbon width of $W_{ac}$ = 2.34 nm.

## IV. Conclusion

Firstly, the band structure of A-GNRs is investigated as a function of ribbon width. By comparing the computed results with and without considering edge effects (edge bond relaxation and 3NN interaction), it is shown that edge physics in nanometer sized ribbons plays an important role in determining the bandgap and thus a careful consideration of edge effects is important to evaluate a device performance. Then we have investigated in detail the classical (gate insulator) capacitance as well as the quantum capacitance in A-GNRs with considering edge effects. The classical capacitance decreases with the increase in oxide thickness and the rate of change is more significant near the lower value of oxide thickness. The quantum capacitance is discussed in both nondegenerate and degenerate regime. The results show that the classical capacitance dominates the total capacitance in the nondegenrate regime (low gate voltage). On the other hand, quantum capacitance is dominant compared to the classical capacitance in the degenerate regime. All the results are discussed in terms of two controlling parameters: width (from the fabrication point of view) and gate voltage (from the operating point of view). Since the results are obtained using realistic model (considering edge effects), it is expected that the analysis of this work will be helpful for the design of GNR based nano-electronic devices.

## References


[1] Y. M. Lin, C. Dimitrakopoulos, K. A. Jenkins, D. B. Farmer, H. Y. Chiu, A. Grill, and Ph. Avouris, "100-GHz Transistors from Wafer-Scale Epitaxial Graphene," *Science*, vol. 327, no. 5966, pp. 662, 2010.

[2] X. Liang, Y. Jung, S. Wu, A. Ismach, D. Olynick, S. Cabrini, and J. Bokor, "Formation of Bandgap and Subbands in Graphene Nanomeshes with Sub-10 nm Ribbon Width Fabricated via Nanoimprint Lithography," *Nano Lett.* ,vol. 10, pp. 2454–2460, 2010.

[3] A. A. Shylau, J. W. Klos and I. V. Zozoulenko, "Capacitance of Graphene Nanoribbons," *Phys. Rev. B*, vol. 80, pp. 205402, 2009.

[4] P. Zhao, M. Choudhury, K. Mohanram, J. Guo, "Computational Model of Edge Effects in Graphene Nanoribbon Transistors," *Nano Research*, vol. 1, no. 5, pp. 395-402, 2008.

[5] K. Nakada, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, "Edge state in graphene ribbons: Nanometer size effect and edge shape dependence," *Phys. Rev. B*, vol. 54, no. 24, pp. 17954–17961, 1996.

[6] M. Choudhury, Y. Yoon, J. Guo and K. Mohanram, "Technology Exploration for Graphene Nanoribbon FETs", *Design Automation Conference(DAC) 45th ACM/IEEE* 2008, pp. 272-277.

[7] Z. F. Wang, Q. Li, Q. W. Shi, X. Wang, J. Yang, J. G. Hao and J. Chen, "Chiral Selective Tunneling Induced Negative Differential Resistance in Zigzag Graphene Nanoribbon: A theoretical study," *Apl. Phy. Lett.*, vol. 92, no. 13, pp. 133114, 2008.

[8] J. Guo, Y. Yoon and Y. Ouyang, "Gate Electrostatics and Quantum Capacitance of Graphene Nanoribbons," *Nano Lett.*, vol. 7, no. 7, pp. 1935-1940, 2007.

[9] D. Unluer, F. Tseng, W. Ghosh, and M. R. Stan, "Atomistic Deconstruction of Clear Performance Advantages of a Monolithically Patterned Wide-Narrow-Wide All-Graphene FET," *Device Research Conference* 2009, pp. 75-76.

[10] Y.-W. Son, M. L. Cohen, and S. G. Louie, "Energy Gaps in Graphene Nanoribbons," *Phys. Rev. Lett.,* vol. 97, pp. 216803, 2006.

[11] Z. Chen and J. Appenzeller, "Mobility Extraction and Quantum Capacitance Impact in High Performance Graphene Field-effect Transistor Devices," *Electron Devices Meeting, IEDM IEEE Int'l*, 2008, pp. 1-4.

[12] Serge Luryi, "Quantum Capacitance Devices," *Appl. Phys. Lett*., vol. 52, no. 6, pp. 501, 1988.

[13] D. L. John, L. C. Castro and D. L. Pulfrey, "Quantum Capacitance in Nanoscale Device Modeling," *J. of Appl. Phys.*, vol. 96, no. 9, pp. 5180, 2004.

[14] D. Gunlycke and C. T. White, "Tight-binding Energy Dispersions of Armchair-edge Graphene Nanostrips," *Phys. Rev. B*, vol. 77, no. 11, pp. 115116, 2008.

[15] T. Fang, A. Konar, H. Xing, and D. Jena, "Carrier Statistics and quantum capacitance of graphene sheets and ribbons," *Apl. Phy. Lett.*, vol. 91, no. 9, pp. 092109, 2007.

[16] Z. Johari, N. A. Amin, M. T. Ahmadi, D. C. Y. Chek, S. M. Mousavi and R. Ismail, "Modeling of Quantum Capacitance in Graphene Nanoribbon," *Int'l Conf. Enabling Science And Nanotechnology,* vol. 1341, 2010, pp. 384-387.

[17] G. Fiori, and G. Iannaccone, "Simulation of Graphene Nanoribbon Field-Effect Transistors," *Electron Device Lett. IEEE*, vol. 28, no. 8, pp. 760-762, 2007.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Modeling of Crosstalk Induced Effects in Nanoscale Copper Interconnects

Manodipan Sahoo and Hafizur Rahaman
Department of Information Technology, Bengal Engineering & Science University, Shibpur, Howrah-711103, India
Email: manodipansahoo@gmail.com, rahaman_h@it.becs.ac.in

***Abstract*—Due to continued scaling of feature sizes, signal integrity and performance of today's copper based nanoscale interconnects are severely impacted. In this work, an ABCD parameter based model is presented for fast and accurate estimation of crosstalk delay and noise for identically coupled copper based nano-interconnect systems. Using the proposed analytical model, the crosstalk delay and noise are estimated in copper based nano-interconnects for intermediate and global interconnects at the future Integrated circuit technology nodes of 21 nm and 15 nm respectively. The proposed model has been compared with SPICE and it is found that both the crosstalk delay and noise model are almost 100% accurate as SPICE and in an average ∼22 times and ∼40 times faster than SPICE respectively.**

***Keywords*—signal integrity , nanoscale , ABCD parameter , crosstalk , delay , noise , nano-interconnects , SPICE.**

## I. Introduction

With the continued scaling of feature sizes, the clock frequency increases, separation between adjacent interconnects decreases, signal edge rates become faster and interconnects become longer. Aggressive interconnect scaling also leads to significant coupling capacitance among adjacent interconnects. Moreover, because of the higher operational frequencies and lower resistivity copper interconnects, inductive impedance of the on-chip wires become comparable to or larger than the resistive impedance. Due to these reasons, on chip capacitive and inductive effects have become a major concern for signal integrity and performance of today's copper (Cu) based nano-interconnects [1]. So, conventional approaches of lumped or distributed RC model of interconnects are not adequate for delay and crosstalk prediction especially in intermediate and global wires in the nanometer regime [2].

In [3], the investigation on a single distributed RLC line has been done without any consideration of coupling capacitances. Delay and crosstalk have been modeled in [4] for loosely coupled interconnects. Time domain expressions for output of capacitively coupled interconnects have been developed using the transmission line theory in [5], [6]. However, the delay and crosstalk noise expressions ignore the effect of capacitive loading at the receiver end. In [7], the lumped parameter approximation was considered for crosstalk delay and noise estimation, which did not correctly model the intermediate and global interconnects of the nanometer regime. In [8]-[10], two coupled interconnects are considered for the analytical estimation of crosstalk delay and noise using classical ABCD parameter based approach. ABCD matrix based method is also reported to be advantagious to SPICE simulations [11] and FDTD (Finite Difference Time Domain) method [12]. In this work, we have used the concept of ABCD parameter matrix to accurately model the crosstalk delay and noise of three identically coupled copper interconnects for intermediate and global interconnects at the integrated circuit technology nodes of 21 nm and 15 nm [3], [8]-[13].

The rest of the paper is organised as follows. Section II briefly describes the equivalent RLC parameters of copper based nano-interconnects. The proposed delay and crosstalk analysis model is explained in section III. Section IV describes the crosstalk analysis results using our proposed model for various technology nodes and levels of interconnects. The work concludes in section V.

This work is partially supported by the DIT, Government of West Bengal, India under VLSI Design Project.

## II. Equivalent electrical parameters of copper interconnects

A typical copper interconnect system is shown in Fig. 1. In Fig. 1, $L$ is interconnect length, $r$ is per unit length (p.u.l) resistance, $l$ is p.u.l self inductance, $C_g$ is p.u.l electrostatic ground capacitance and $C_c$ is p.u.l electrostatic coupling capacitance. The driver is implemented using inverting buffers and load is capacitive, denoted by $C_L$. The buffers can be modeled as an equivalent RC circuit with a high degree of accuracy [3]. The parameters of the driving buffer are $R_s$ and $C_{out}$. $R_s$ and $C_{out}$ are the equivalent switching resistance and the equivalent diffusion capacitance of a minimum sized inverter buffer. RC parameters of the driver and load inverting buffer are estimated using the device parameters from ITRS-2011 roadmap [14]. The estimated parameters are shown in Table I. The technology parameters for various levels of copper interconnects are obtained from ITRS-2011 roadmap [14] and shown in Table II. The analytical equations from [15] are used to calculate the distributed RLC parameters (i.e., $r$, $l$, $C_g$ and $C_c$) of copper interconnects.

TABLE I. RC PARAMETERS OF THE MINIMUM SIZED BUFFER FOR VARIOUS TECHNOLOGY NODES

| Parameters \ Technology node | 21 nm | 15 nm |
|---|---|---|
| $R_s$ (in KΩ) | 34.45 | 47.4 |
| $C_{out}$ (in fF) | 0.049 | 0.03 |

## III. Proposed Model for analysis of crosstalk effects

In this work, a computationally efficient model for crosstalk analysis in identically coupled interconnect systems is presented. The model is applied for analyzing crosstalk effects in copper based nano-interconnect systems and compared with

TABLE II. INTERCONNECT PARAMETERS (ITRS-2011)

| Year | 2015 | 2018 |
|---|---|---|
| DRAM 1/2 pitch (nm) (contacted) | 25.26 | 17.86 |
| MPU/ASIC Metal1 1/2 pitch (nm) (contacted) | 21.24 | 15.02 |
| MPU Physical gate length (nm) | 16.80 | 12.78 |
| No. of Metal levels | 13 | 14 |
| Wiring pitch (nm) | 42 | 30 |
| | 42 | 30 |
| | 84 | 60 |
| A/R (Cu) | 1.9 | 2 |
| | 1.9 | 2 |
| | 2.34 | 2.34 |
| A/R (Via) | 1.7 | 1.8 |
| | 1.7 | 1.8 |
| | 1.5 | 1.5 |
| Effective Resistivity ($\mu$ $\Omega$-cm) | 6.61 | 8.09 |
| | 6.61 | 8.09 |
| | 4.82 | 6.08 |
| VDD (Volt) | 0.8 | 0.73 |
| $T_{ox}$ (nm) | 0.67 | <0.55 |
| Effective dielectric constant ($k_{eff}$) | 2.55-3.00 | 2.10-2.78 |
| Average dielectric constant ($k_{avg}$) | 2.78 | 2.44 |

SPICE. In our proposed methodology first, we develop an analytical expression for the crosstalk delay and noise in laplace domain. Then we adopt the inverse laplace transform to find out the time-domain response of the crosstalk delay and noise. The proposed methodology is briefly illustrated as follows.

*A. ABCD parameter based approach:*

Infinitesimal section of an identically coupled three conductor interconnect system is shown in Fig.2. Here, net1 and net3 are the aggressor nets and net2 is the victim net. The Kirchhoff equation for an infinitesimally small segment of these three coupled interconnects is given by the following equation,

$$\phi_i = P\phi_o \tag{1}$$

where,

$$\phi_i = [V_{i1}, V_{i2}, V_{i3}, I_{i1}, I_{i2}, I_{i3}]^T \tag{2}$$
$$\phi_o = [V_{o1}, V_{o2}, V_{o3}, I_{o1}, I_{o2}, I_{o3}]^T \tag{3}$$

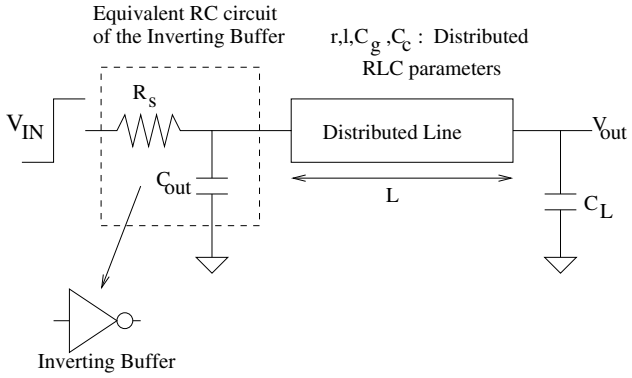



Fig. 1. Electrical equivalent model of a typical copper interconnect system.

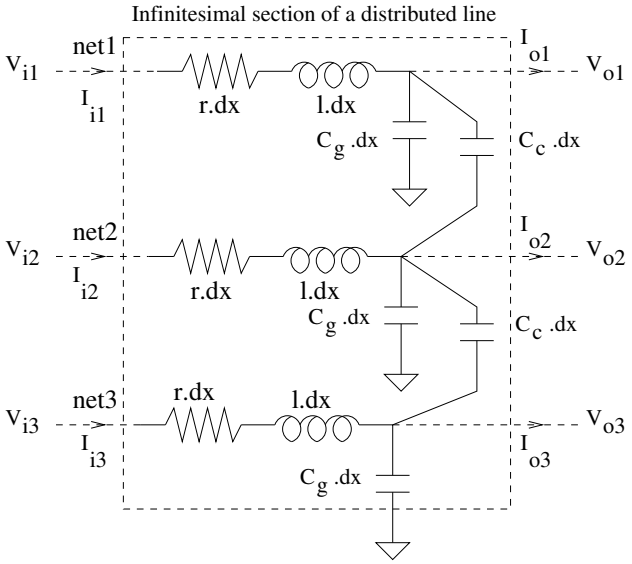



Fig. 2. Schematic diagram of an infinitesimal section of an identically coupled copper interconnect system.

and, matrix P is shown in (23). The matrix P can be diagonalized as $P = VWV^{-1}$, where, V and W are the eigen vector and eigen value matrices of P respectively. Using,

$$P^n = (VWV^{-1})^n = (VW^nV^{-1}) \tag{4}$$

and the identity,

$$\lim_{n\to+\infty}(1+\frac{x}{n})^n \to e^x \tag{5}$$

$P^n$ can be written as shown in (25), where, n is the number of infinitesimal lumped sections. Other parameters are defined as follows.

$$ta = \cosh(L\theta_1) \tag{6}$$
$$tb = \cosh(L\theta_2) \tag{7}$$
$$tc = \cosh(L\theta_3) \tag{8}$$
$$td = \sinh(L\theta_1) \tag{9}$$
$$te = \sinh(L\theta_2) \tag{10}$$
$$tf = \sinh(L\theta_3) \tag{11}$$

Here, $\theta_1$, $\theta_2$ and $\theta_3$ are the propagation constants and, $Z_1$, $Z_2$ and $Z_3$ are the characteristic impedances of the decoupled interconnects [10]. The definitions are given below.

$$dx = L/n \tag{12}$$
$$\theta_1 = \sqrt{sC_g(r+sl)} \tag{13}$$
$$\theta_2 = \sqrt{s(C_g+3C_c)(r+sl)} \tag{14}$$
$$\theta_3 = \sqrt{s(C_g+C_c)(r+sl)} \tag{15}$$
$$Z_1 = \sqrt{\frac{(r+sl)}{sC_g}} \tag{16}$$
$$Z_2 = \sqrt{\frac{(r+sl)}{s(C_g+3C_c)}} \tag{17}$$
$$Z_3 = \sqrt{\frac{(r+sl)}{s(C_g+C_c)}} \tag{18}$$

TABLE III. REPRESENTATION OF TRANSFER FUNCTION H(S)

| Evaluation Condition | Representation of H(s) |
|---|---|
| Crosstalk delay evaluation in the victim net($V_{i2}$) when both the aggressors($V_{i1}$ and $V_{i3}$) switch in the same direction | $H_1(s)$ |
| Crosstalk delay evaluation in the victim net($V_{i2}$) when both the aggressors($V_{i1}$ and $V_{i3}$) switch in the opposite direction | $(\frac{4}{3}.H_2(s) - \frac{1}{3}.H_1(s))$ |
| Crosstalk noise evaluation in the victim net($V_{i2}$) when both the aggressors($V_{i1}$ and $V_{i3}$) switch in the same direction | $\frac{2}{3}(H_1(s) - H_2(s))$ |

The ABCD matrices of the driver and load ends of the three identically coupled interconnects can be written as,

$$P_{dr} = \begin{bmatrix} 1 & 0 & R_s & 0 \\ 0 & 1 & 0 & R_s \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{19}$$

$$P_{dc} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ sC_{out} & 0 & 1 & 0 \\ 0 & sC_{out} & 0 & 1 \end{bmatrix} \tag{20}$$

$$P_{load} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ sC_L & 0 & 1 & 0 \\ 0 & sC_L & 0 & 1 \end{bmatrix} \tag{21}$$

The final Kirchoff equation for the coupled interconnects can be written as,

$$\Phi_i = P_{dr}P_{dc}P^nP_{load}\Phi_o \tag{22}$$

Here, $P_{dr}$ and $P_{dc}$ are the ABCD matrices of the resistive ($R_s$) and capacitive ($C_{out}$) component of the driving buffer. $P_{load}$ is the ABCD matrix representation of the capacitive load. By solving (22), the final output voltages $V_{o1}$(s), $V_{o2}$(s) and $V_{o3}$(s) can be represented in terms of the inputs $V_{i1}$(s), $V_{i2}$(s) and $V_{i3}$(s) in matrix form as shown in (26). The exact expressions of $H_1$(s), $H_2$(s) and $H_3$(s) in (26) are shown in (27), (28) and (29) respectively. $H_1$(s), $H_2$(s) and $H_3$(s) are the transfer functions of the decoupled interconnects. The transfer function for the crosstalk delay or noise evaluation in the victim net will be denoted as H(s) throughout the paper. The representation of H(s) for different evaluation conditions (i.e., crosstalk delay and noise) are shown in Table III. We consider a seventh-order Padé expansion for approximating hyperbolic functions in (27), (28) and (29) with high degree of accuracy [3]. So the transfer function H(s) will be of the form given in (30) after the Padé expansion:

$$H(s) = \frac{1}{1 + sb_1 + s^2b_2 + s^3b_3 + s^4b_4 + s^5b_5 + s^5b_5 + s^6b_6 + s^7b_7} \tag{30}$$

where, the coefficients are shown below,

$$b_1 = \frac{L^2rC}{2} + R_s(C_{out} + C_L) + R_sLC + LC_Lr \tag{31}$$

$$\begin{aligned} b_2 = {} & \frac{L^2lC}{2} + \frac{R_s(C_{out} + C_L)L^2rC}{2} + \frac{R_srL^3C^2}{6} \\ & + C_LLl + \frac{C_LL^3r^2C}{6} + LC_LR_sC_{out}r + \frac{L^4r^2C^2}{24} \end{aligned} \tag{32}$$

TABLE IV. REPRESENTATION OF PARAMETER C

| Transfer function | Parameter C |
|---|---|
| $H_1(s)$ | $C_g$ |
| $H_2(s)$ | $(C_g + 3C_c)$ |
| $H_3(s)$ | $(C_g + C_c)$ |

$$\begin{aligned} b_3 = {} & \frac{L^3R_slC^2}{6} + \frac{L^4rlC^2}{12} + \frac{L^6r^3C^3}{720} + \frac{R_s(C_{out} + C_L)L^2lC}{2} \\ & + \frac{R_s(C_{out} + C_L)L^4r^2C^2}{24} + \frac{C_Lr^3L^5C^2}{120} + \frac{R_sC^3r^2L^5}{120} \\ & + \frac{C_LrlL^3C}{3} + LC_LR_sC_{out}l + \frac{L^3C_LR_sC_{out}r^2C}{6} \end{aligned} \tag{33}$$

$$\begin{aligned} b_4 = {} & \frac{L^4l^2C^2}{24} + \frac{L^6r^2lC^3}{240} + \frac{R_s(C_{out} + C_L)rlL^4C^2}{12} + \frac{R_sC^3rlL^5}{60} \\ & + \frac{R_s(C_{out} + C_L)r^3L^6C^3}{720} + \frac{C_Lr^2lL^5C^2}{40} + \frac{C_Ll^2L^3C}{6} \\ & + \frac{C_LR_sC_{out}L^3rlC}{3} + \frac{C_LR_sC_{out}r^3L^5C^2}{120} \end{aligned} \tag{34}$$

$$\begin{aligned} b_5 = {} & \frac{rl^2C_LL^5C^2}{40} + \frac{rl^2L^6C^3}{240} + \frac{R_s(C_{out} + C_L)l^2C^2L^4}{24} \\ & + \frac{R_sl^2C^3L^5}{120} + \frac{R_s(C_{out} + C_L)r^2lL^6C^3}{240} + \frac{C_LR_sC_{out}r^2lL^5C^2}{40} \\ & + \frac{C_LR_sC_{out}l^2L^3C}{6} \end{aligned} \tag{35}$$

$$\begin{aligned} b_6 = {} & \frac{l^3L^6(C_g + 2C_c)^3}{720} + \frac{R_s(C_{out} + C_L)rl^2L^6(C_g + 2C_c)^3}{240} \\ & + \frac{C_Ll^3L^5(C_g + 2C_c)^2}{120} + \frac{C_LR_sC_{out}rl^2L^5(C_g + 2C_c)^2}{40} \end{aligned} \tag{36}$$

$$b_7 = \frac{R_s(C_{out} + C_L)l^3L^6(C_g + 2C_c)^3}{720} + \frac{C_LR_sC_{out}l^3L^5(C_g + 2C_c)^2}{120} \tag{37}$$

Similarly, $H_2$(s) and $H_3$(s) can be represented in a form as given in (30). The parameter C in the coefficients $b_1$ to $b_7$ varies according to the transfer function of the decoupled interconnect and it is shown in Table IV. Here, the other parameters are basically the distributed parameters of the copper interconnect, length of the interconnect, driver and load RC parameters.

TABLE V. CROSSTALK DELAY COMPARISON IN INTERMEDIATE LEVEL INTERCONNECTS IN 21 $nm$ NODE

| Length | $T1_{PLH}$(ns) | | | $T2_{PLH}$(ns) | | |
|---|---|---|---|---|---|---|
| | SPICE | Model | % diff | SPICE | Model | % diff |
| 10 $\mu$m | 0.104 | 0.103 | 0.96 | 0.106 | 0.105 | 0.94 |
| 50 $\mu$m | 0.325 | 0.324 | 0.31 | 0.355 | 0.354 | 0.28 |
| 100 $\mu$m | 0.606 | 0.605 | 0.165 | 0.712 | 0.709 | 0.42 |
| 500 $\mu$m | 3.01 | 3.01 | 0 | 5.4 | 5.34 | 1.11 |
| 1 mm | 6.42 | 6.4 | 0.31 | 15.9 | 15.6 | 1.88 |
| Average simulation time in SPICE is ∼65.6 sec and our model takes ∼4.25 sec. So our model is in average ∼15.5 times faster than SPICE. | | | | | | |

## IV. Results and Discussions

The proposed analytical model is implemented using MATLAB 7.1 under standard desktop environment on Intel Core 2 Duo Processor running at 3.0 GHz with 4.0 GB of physical memory. The simulation is performed in various technology nodes namely 21 nm and 15 nm and various interconnect levels viz. intermediate and global. We have chosen the buffer size to be 100 times the minimum sized buffer for global level and 50 times the minimum sized buffer for intermediate level interconnects. The capacitive load is considered to be 100 fF for all the simulations.

$$P = \begin{bmatrix} 1 & 0 & 0 & (r+sl)dx & 0 & 0 \\ 0 & 1 & 0 & 0 & (r+sl)dx & 0 \\ 0 & 0 & 1 & 0 & 0 & (r+sl)dx \\ s(C_g+C_c)dx & -sC_c.dx & 0 & 1 & 0 & 0 \\ -sC_cdx & s(C_g+2C_c)dx & -sC_cdx & 0 & 1 & 0 \\ 0 & -sC_cdx & s(C_g+C_c)dx & 0 & 0 & 1 \end{bmatrix} \tag{23}$$

$$W = \begin{bmatrix} (1+\theta_1 dx) & 0 & 0 & 0 & 0 & 0 \\ 0 & (1-\theta_1 dx) & 0 & 0 & 0 & 0 \\ 0 & 0 & (1+\theta_2 dx) & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-\theta_2 dx) & 0 & 0 \\ 0 & 0 & 0 & 0 & (1+\theta_3 dx) & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-\theta_3 dx) \end{bmatrix} \tag{24}$$

$$P^n = \begin{bmatrix} (\frac{ta}{3}+\frac{tb}{6}+\frac{tc}{2}) & (\frac{ta}{3}-\frac{tb}{3}) & (\frac{ta}{3}+\frac{tb}{6}-\frac{tc}{2}) & (\frac{td.Z1}{3}+\frac{te.Z2}{6}+\frac{tf.Z3}{2}) & (\frac{td.Z1}{3}-\frac{te.Z2}{3}) & (\frac{td.Z1}{3}+\frac{te.Z2}{6}-\frac{tf.Z3}{2}) \\ (\frac{ta}{3}-\frac{tb}{3}) & (\frac{ta}{3}+\frac{2tb}{3}) & (\frac{ta}{3}-\frac{tb}{3}) & (\frac{td.Z1}{3}-\frac{te.Z2}{3}) & (\frac{td.Z1}{3}+\frac{2te.Z2}{3}) & (\frac{td.Z1}{3}-\frac{te.Z2}{3}) \\ (\frac{ta}{3}+\frac{tb}{6}-\frac{tc}{2}) & (\frac{ta}{3}-\frac{tb}{3}) & (\frac{ta}{3}+\frac{tb}{6}+\frac{tc}{2}) & (\frac{td.Z1}{3}+\frac{te.Z2}{6}-\frac{tf.Z3}{2}) & (\frac{td.Z1}{3}-\frac{te.Z2}{3}) & (\frac{td.Z1}{3}+\frac{te.Z2}{6}+\frac{tf.Z3}{2}) \\ (\frac{td}{3Z1}+\frac{te}{6Z2}+\frac{tf}{2Z3}) & (\frac{td}{3Z1}-\frac{te}{3Z2}) & (\frac{td}{3Z1}+\frac{te}{6Z2}-\frac{tf}{2Z3}) & (\frac{ta}{3}+\frac{tb}{6}+\frac{tc}{2}) & (\frac{ta}{3}-\frac{tb}{3}) & (\frac{ta}{3}+\frac{tb}{6}-\frac{tc}{2}) \\ (\frac{td}{3Z1}-\frac{te}{3Z2}) & (\frac{td}{3Z1}+\frac{2te}{3Z2}) & (\frac{td}{3Z1}-\frac{te}{3Z2}) & (\frac{ta}{3}-\frac{tb}{3}) & (\frac{ta}{3}+\frac{2tb}{3}) & (\frac{ta}{3}-\frac{tb}{3}) \\ (\frac{td}{3Z1}+\frac{te}{6Z2}-\frac{tf}{2Z3}) & (\frac{td}{3Z1}-\frac{te}{3Z2}) & (\frac{td}{3Z1}+\frac{te}{6Z2}+\frac{tf}{2Z3}) & (\frac{ta}{3}+\frac{tb}{6}-\frac{tc}{2}) & (\frac{ta}{3}-\frac{tb}{3}) & (\frac{ta}{3}+\frac{tb}{6}+\frac{tc}{2}) \end{bmatrix} \tag{25}$$

$$\begin{bmatrix} V_{o1}(s) \\ V_{o2}(s) \\ V_{o3}(s) \end{bmatrix} = \begin{bmatrix} (\frac{1}{3}H_1(s)+\frac{1}{6}H_2(s)+\frac{1}{2}H_3(s)) & \frac{1}{3}(H_1(s)-H_2(s)) & (\frac{1}{3}H_1(s)+\frac{1}{6}H_2(s)-\frac{1}{2}H_3(s)) \\ \frac{1}{3}(H_1(s)-H_2(s)) & (\frac{1}{3}H_1(s)+\frac{2}{3}H_2(s)) & \frac{1}{3}(H_1(s)-H_2(s)) \\ (\frac{1}{3}H_1(s)+\frac{1}{6}H_2(s)-\frac{1}{2}H_3(s)) & \frac{1}{3}(H_1(s)-H_2(s)) & (\frac{1}{3}H_1(s)+\frac{1}{6}H_2(s)+\frac{1}{2}H_3(s)) \end{bmatrix} \begin{bmatrix} V_{i1}(s) \\ V_{i2}(s) \\ V_{i3}(s) \end{bmatrix} \tag{26}$$

$$H_1(s) = \frac{1}{(1+sR_s(C_{out}+C_L))\cosh(\theta_1 L)+(\frac{R_s}{Z_1}+sC_L(1+sR_sC_{out})Z_1)\sinh(\theta_1 L)} \tag{27}$$

$$H_2(s) = \frac{1}{(1+sR_s(C_{out}+C_L))\cosh(\theta_2 L)+(\frac{R_s}{Z_2}+sC_L(1+sR_sC_{out})Z_2)\sinh(\theta_2 L)} \tag{28}$$

$$H_3(s) = \frac{1}{(1+sR_s(C_{out}+C_L))\cosh(\theta_3 L)+(\frac{R_s}{Z_3}+sC_L(1+sR_sC_{out})Z_3)\sinh(\theta_3 L)} \tag{29}$$

TABLE VI. CROSSTALK DELAY COMPARISON IN INTERMEDIATE LEVEL INTERCONNECTS IN 15 $nm$ NODE

| Length | $T1_{PLH}$(ns) | | | $T2_{PLH}$(ns) | | |
|---|---|---|---|---|---|---|
| | SPICE | Model | % diff | SPICE | Model | % diff |
| 10 $\mu$m | 0.192 | 0.191 | 0.52 | 0.196 | 0.195 | 0.51 |
| 50 $\mu$m | 0.696 | 0.695 | 0.14 | 0.753 | 0.751 | 0.26 |
| 100 $\mu$m | 1.33 | 1.33 | 0 | 1.54 | 1.54 | 0 |
| 500 $\mu$m | 6.74 | 6.73 | 0.15 | 11.7 | 11.5 | 1.71 |
| 1 mm | 14.3 | 14.2 | 0.7 | 33.9 | 33.4 | 1.47 |
| Average simulation time in SPICE is ∼65 sec and our model takes ∼2.35 sec. So our model is in average ∼ 27.5 times faster than SPICE. | | | | | | |

TABLE VII. CROSSTALK DELAY COMPARISON IN GLOBAL LEVEL INTERCONNECTS IN 21 $nm$ NODE

| Length | $T1_{PLH}$(ns) | | | $T2_{PLH}$(ns) | | |
|---|---|---|---|---|---|---|
| | SPICE | Model | % diff | SPICE | Model | % diff |
| 100 $\mu$m | 0.109 | 0.108 | 0.92 | 0.133 | 0.132 | 0.75 |
| 500 $\mu$m | 0.473 | 0.472 | 0.21 | 0.912 | 0.901 | 1.2 |
| 1 mm | 0.997 | 0.992 | 0.5 | 2.67 | 2.63 | 1.5 |
| 1.5 mm | 1.6 | 1.59 | 0.62 | 5.29 | 5.19 | 1.89 |
| 2 mm | 2.27 | 2.25 | 0.88 | 8.74 | 8.56 | 2.06 |
| Average simulation time in SPICE is ∼67 sec and our model takes ∼3.9 sec. So our model is in average ∼ 17 times faster than SPICE. | | | | | | |

TABLE VIII. CROSSTALK DELAY COMPARISON IN GLOBAL LEVEL INTERCONNECTS IN 15 $nm$ NODE

| Length | $T1_{PLH}$(ns) | | | $T2_{PLH}$(ns) | | |
|---|---|---|---|---|---|---|
| | SPICE | Model | % diff | SPICE | Model | % diff |
| 100 $\mu$m | 0.24 | 0.239 | 0.42 | 0.284 | 0.282 | 0.7 |
| 500 $\mu$m | 1.13 | 1.12 | 0.88 | 2.04 | 2.01 | 1.47 |
| 1 mm | 2.38 | 2.38 | 0 | 5.94 | 5.85 | 1.51 |
| 1.5 mm | 3.81 | 3.79 | 0.52 | 11.7 | 11.5 | 1.71 |
| 2 mm | 5.39 | 5.36 | 0.56 | 19.3 | 19 | 1.55 |
| Average simulation time in SPICE is ∼68 sec and our model takes ∼2.25 sec. So our model is in average ∼ 30 times faster than SPICE. | | | | | | |

TABLE IX. CROSSTALK NOISE COMPARISON IN INTERMEDIATE LEVEL INTERCONNECTS IN 21 $nm$ NODE

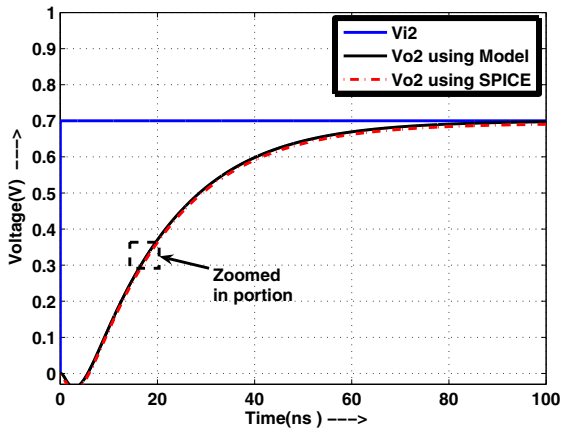
| Length | NP(mV) | | NW(ns) | | NA(V-ps) | | |
|---|---|---|---|---|---|---|---|
| | SPICE | Model | SPICE | Model | SPICE | Model | % diff |
| 10 $\mu$m | 3.12 | 3.1 | 0.817 | 0.822 | 1.27 | 1.27 | 0 |
| 50 $\mu$m | 12.66 | 12.5 | 2.53 | 2.52 | 16 | 15.8 | 1.25 |
| 100 $\mu$m | 23.45 | 23.1 | 4.78 | 4.77 | 56 | 55.1 | 1.6 |
| 500 $\mu$m | 92.67 | 91.4 | 27.1 | 27 | 1260 | 1230 | 2.38 |
| 1 mm | 151.88 | 151 | 66.3 | 65.6 | 5040 | 4950 | 1.78 |
| Average simulation time in SPICE is ∼150 sec and our model takes ∼4.25 sec. So our model is in average ∼ 35 times faster than SPICE. | | | | | | | |

### *A. SPICE Simulation procedure*

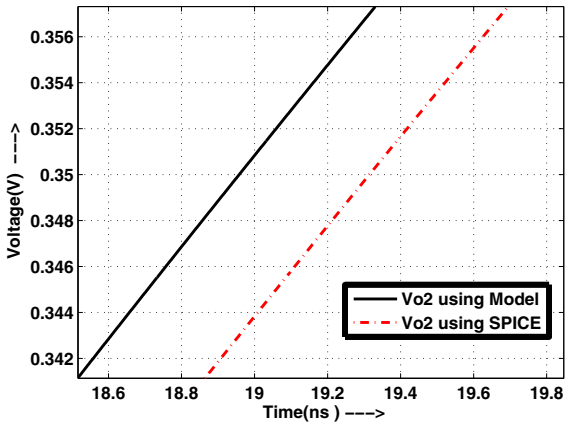
We have chosen the Spectre simulator version IC 6.1.4.500 of Cadence and found out the noise and delay values. Simulations are run in the same environment where MATLAB simulations are run. RC parameters of the driver/load inverting buffer are taken from Table I. Because of the distributed nature of the interconnects, we implement that by cascading 200 infinitesimal lumped RLC sections [16].

### *B. Results of Crosstalk delay analysis*

Table V, VI, VII and VIII show the crosstalk delay values for various technology nodes and level of interconnects in three identically coupled copper interconnect system. Our model is in an average 22 times faster as compared to SPICE with accuracy within 2.06% of SPICE. Average simulation time is computed by averaging the simulation run time for individual runs for each interconnect length considered here. We have




(a)




(b)

Fig. 3. Crosstalk delay waveform for 2 mm long copper interconnect in 15 nm technology node. (a) Crosstalk delay waveform as obtained from SPICE and proposed model when aggressor and victim nets switch in opposite direction. (b) Zoomed crosstalk delay waveform.

TABLE X. CROSSTALK NOISE COMPARISON IN INTERMEDIATE LEVEL INTERCONNECTS IN 15 $nm$ NODE

| Length | NP(mV) | | NW(ns) | | NA(V-ps) | | |
|---|---|---|---|---|---|---|---|
| | SPICE | Model | SPICE | Model | SPICE | Model | % diff |
| 10 $\mu$m | 2.35 | 2.33 | 1.49 | 1.49 | 1.75 | 1.74 | 0.57 |
| 50 $\mu$m | 9.78 | 9.66 | 5.37 | 5.37 | 26.3 | 26 | 1.14 |
| 100 $\mu$m | 18.44 | 18.2 | 10.5 | 10.5 | 96.5 | 95 | 1.55 |
| 500 $\mu$m | 75.53 | 74.6 | 60 | 59.7 | 2260 | 2220 | 1.77 |
| 1 mm | 125.99 | 125 | 145 | 143 | 9130 | 8980 | 1.64 |
| Average simulation time in SPICE is ∼114 sec and our model takes ∼2.35 sec. So our model is in average ∼ 48.5 times faster than SPICE. | | | | | | | |

TABLE XI. CROSSTALK NOISE COMPARISON IN GLOBAL LEVEL INTERCONNECTS IN 21 $nm$ NODE

| Length | NP(mV) | | NW(ns) | | NA(V-ps) | | |
|---|---|---|---|---|---|---|---|
| | SPICE | Model | SPICE | Model | SPICE | Model | % diff |
| 100 $\mu$m | 29.38 | 29.1 | 0.875 | 0.875 | 12.9 | 12.7 | 1.55 |
| 500 $\mu$m | 105.74 | 105 | 4.39 | 4.36 | 232 | 228 | 1.72 |
| 1 mm | 167.54 | 167 | 10.7 | 10.6 | 898 | 883 | 1.67 |
| 1.5 mm | 209.66 | 210 | 19.3 | 19 | 2020 | 1990 | 1.48 |
| 2 mm | 240.34 | 241 | 30.1 | 29.6 | 3620 | 3560 | 1.66 |
| Average simulation time in SPICE is ∼121 sec and our model takes ∼3.9 sec. So our model is in average ∼ 31 times faster than SPICE. | | | | | | | |

calculated the delay values due to the crosstalk for two cases : (a) $T1_{PLH}$ : when aggressor and victim nets switch in the same direction, from low to high, and (b) $T2_{PLH}$ : when aggressor and victim nets switch in the opposite direction, victim net switches from low to high and aggressor nets switch from high to low. Fig. 3(a) shows crosstalk delay waveforms as obtained by SPICE simulations and proposed model when the nets switch in opposite direction. The difference between the waveforms is shown in Fig. 3(b).

TABLE XII. CROSSTALK NOISE COMPARISON IN GLOBAL LEVEL INTERCONNECTS IN 15 $nm$ NODE

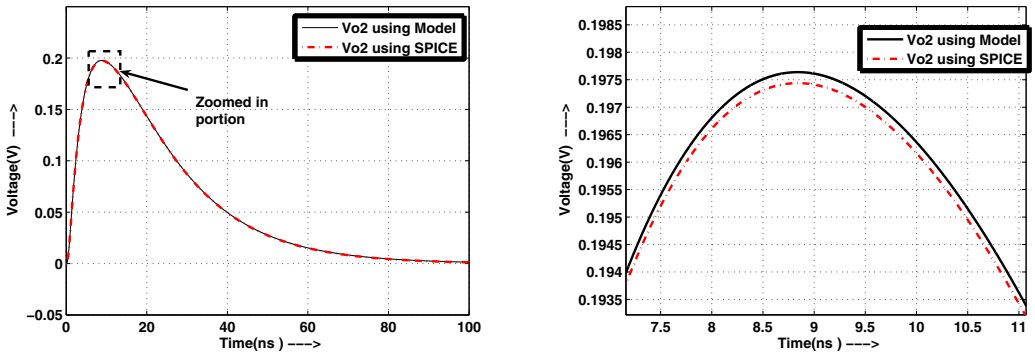
| Length | NP(mV) | | NW(ns) | | NA(V-ps) | | |
|---|---|---|---|---|---|---|---|
| | SPICE | Model | SPICE | Model | SPICE | Model | % diff |
| 100 $\mu$m | 21.56 | 21.3 | 1.9 | 1.9 | 20.5 | 20.2 | 1.46 |
| 500 $\mu$m | 82.41 | 81.4 | 10.2 | 10.1 | 420 | 412 | 1.9 |
| 1 mm | 134.16 | 133 | 24.7 | 24.5 | 1660 | 1630 | 1.81 |
| 1.5 mm | 170.49 | 170 | 44.1 | 43.5 | 3760 | 3700 | 1.6 |
| 2 mm | 197.46 | 198 | 68.3 | 67.2 | 6740 | 6640 | 1.48 |
| Average simulation time in SPICE is ∼101 sec and our model takes ∼2.25 sec. So our model is in average ∼ 45 times faster than SPICE. | | | | | | | |

### C. Results of Crosstalk noise analysis

Table IX, X, XI and XII show the crosstalk noise values for various technology nodes and level of interconnects in three identically coupled copper interconnect system. Our model has accuracy within ∼2.38% of SPICE and in an average 40 times faster than SPICE. We have calculated various noise parameters to characterize the crosstalk induced noise. These parameters include Noise peak amplitude (NP), Noise width (NW) and Noise Area (NA). NA is calculated as, $NA = \frac{1}{2}NP \times NW$. In all these simulations, the victim net remains at low and aggressor nets switch from low to high, thus inducing a noise in the victim net. Fig. 4(a) shows crosstalk noise waveforms as obtained by SPICE simulations and proposed model. It can be observed that the waveforms obtained using SPICE and the proposed model is nearly the same. The difference between the waveforms is shown in the Fig. 4(b).

## V. Conclusions

In this work, an analytical model for crosstalk induced delay and noise has been presented for identically coupled copper based interconnects for various interconnect levels and technology nodes in the nanometer regime. The result shows that the proposed model is almost 100% accurate as SPICE. It is shown that the proposed crosstalk delay model is in an




(a) (b)

Fig. 4. Crosstalk noise waveform for 2 mm long copper interconnect in 15 nm technology node. (a) Crosstalk noise waveform as obtained from SPICE and proposed model when victim net remains quiet at ground and aggressor nets switch. (b) Zoomed crosstalk noise waveform.

average ∼22 times and crosstalk noise model is ∼40 times faster than SPICE. It is also shown that the crosstalk noise peak voltage reduces as we move towards scaled technology nodes but noise width increases with scaling. It is also observed that the noise area increases as we scale more which is a concern for the overall signal integrity in the nanometer technology nodes. All the noise parameters increase with the length of the interconnect. It is also observed that the crosstalk delay increases monotonically with the interconnect length. The crosstalk delay increases as we scale deeper. Hence, we may conclude that the crosstalk noise and delay will be a concern for the performance and integrity of the copper based nano-interconnects in future IC technology nodes.

## REFERENCES


[1] E. E. Davidson, B. D. McCredie, and W. V. Vilkelis, "Long lossy lines ($L^3$) and their impact upon large chip performance," *IEEE Trans. Components, Packaging and Manufacturing Technology*, Part B: Advanced Packaging, vol. 20, pp. 361-375, Nov. 1998.

[2] A. Deutsch et. al., "When are transmission-line effects important for on-chip interconnections?," *IEEE Trans. Microwave Theory Tech.*, vol. 45, pp. 1836-1846, Oct. 1997.

[3] K. Banerjee and A. Mehrotra, "Analysis of On-Chip Inductance Effects for Distributed RLC Interconnects", *IEEE TCAD*, vol. 21, pp. 904-915, Aug. 2002.

[4] K. T. Tang and E. G. Friedman, "Peak Crosstalk Noise Estimation in CMOS VLSI Circuits", *IEEE ICECS*, pp. 1539-1542, September 1999.

[5] J. A. Davis and J. D. Meindl, "Compact distributed RLC models for multilevel interconnect networks," *Symp. VLSI Technol. Dig. Techn. Papers*, 1999, pp. 165-166.

[6] J. A. Davis and J. D. Meindl, "Compact distributed RLC interconnect models-Part II. Coupled line transient expressions and peak crosstalk in multilevel networks," *IEEE Trans. Electron. Devices*, vol. 47, pp. 2078-2087, Nov. 2000.

[7] D. Das and H. Rahaman, "Unified Model for Analyzing Timing Delay and Crosstalk Effects in Carbon Nanotube Interconnects", *IEEE ASQED*, pp. 100-109, 2012.

[8] A. K. Palit, V. Meyer, W. Anheier and J. Chloeffel, "ABCD modeling of crosstalk coupling noise to analyze the signal integrity losses on the victim interconnect in DSM chips", *IEEE 18th International Conference on VLSI Design*, 2005.

[9] A. K. Palit, S. Hasan and W. Anheier, "Decoupled Victim Model for the Analysis of Crosstalk Noise between On-chip Coupled Interconnects," *11th Electronics Packaging Technology Conference*, pp. 697-701, 2009.

[10] J. Zhang and E. G. Friedman, "Decoupling Technique and Crosstalk Analysis for Coupled RLC Interconnects", *IEEE ISCAS*, pp. 521-524,2004.

[11] M. Sahoo, P. Ghosal and H. Rahaman, "An ABCD parameter based Modeling and Analysis of Crosstalk Induced Effects in Single-Walled Carbon Nanotube Bundle Interconnects", *IEEE ASQED*, Aug. 2013.

[12] M. Sahoo, P. Ghosal and H. Rahaman, "An ABCD Parameter Based Modeling and Analysis of Crosstalk Induced Effects in Multiwalled Carbon Nanotube Bundle Interconnects", *IEEE* $27^{th}$ *International Conference on VLSI Design*, 2014 (accepted).

[13] M. Sahoo and H. Rahaman, "Analytical Modeling of Crosstalk Effects in Coupled Copper Interconnects in Deep Sub Micron Technology," $5^{th}$ *International Conference on Computers and Devices for Communication*, 2012.

[14] International Technology Roadmap for Semiconductors (ITRS-2011) Reports, [Online]. Available: http://www.itrs.net/reports.html.

[15] Predictive Technology Model [Online]. Available: http://ptm.asu.edu/.

[16] T. Sakurai, "Approximation of Wiring Delay in MOSFET LSI," *IEEE JSSC*, vol. SC-18, no. 4, Aug. 1983.

# Design and Performance Analysis of a 4×1 MIMO Inverted F- Shaped Patch Antenna Array for X-Band Applications and Link Budget Optimization

Anindya Kumar Kundu, MD. Tofael Hossain Khan*, Md. Ashikur Rahman, Md. Osman Goni, Kazi Abul Barket[2]
Dept. of Electronics & Communication Engineering, Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh.
[2]Physics Discipline, Khulna University, Khulna-9208, Bangladesh.
*tofael18@gmail.com

*Abstract*—**In the modern communication technologies, the development of very small size, low-cost, low-profile, high gain and high directivity antennas is essential. In this paper, the design and performance of a 10 GHz 4×1 MIMO Inverted F-shaped patch antenna (PIFA) array has been analyzed for the X-band applications. The novelty of the proposed PIFA antenna array is that it has a gain of 12.87dB, directivity of 13.24dBi; return loss value of -36.61dB, Voltage Standing Wave Ratio (VSWR) of 1.03, bandwidth of 240MHz, antenna efficiency of 91.90% and 3dB angular beamwidth i.e., the Half Power Beamwidth (HPBW) of 64.4deg. The resonant frequency of the antenna array is 10.16 GHz. This antenna array can be used for X-band applications like satellite communications, radars, medical applications and other wireless systems. This antenna array is designed and simulated in the CST Microwave Studio. Link budget optimization is performed in order to analyze the critical factors in the transmission chain and to optimize the performance characteristic. The link budget determines what size antenna is to use, power requirements and in general, the overall customer satisfaction.**

***Keywords— Array, gain, CST Microwave Studio, satellite communications, X-band.***

## I. INTRODUCTION

With the growth of modern communication technology, the use of microstrip antennas have increased due to their very small size, low-cost, low-profile, high gain and high directivity. Patch antennas and patch antenna array are widely presented in books and papers in the last decade as presented in [1-4]. A 10 GHz phased cylindrical antenna array system incorporating IF phase processing is analyzed in [5]. A 12 GHz Planar Array Antenna for Satellite Communication is studied in [6]. The proposed antenna has a 10dB bandwidth of 240MHz. These wide bandwidths are applicable to X-band applications like satellite communications, radars, medical applications and other wireless systems. The total antenna efficiency $e_0$ is used to take into account losses at the input terminals and within the structure of the antenna. Such losses may be due to the (i) reflections because of the mismatch between the transmission line and the antenna, (ii) $I^2R$ losses (conduction and dielectric) [7]. In general the overall antenna efficiency can be written as-

$$e_0 = e_r e_{cd} = e_{cd}(1 - |\Gamma^2|) \quad (1)$$

where, $e_r$ is the reflection efficiency $=(1 - |\Gamma^2|)$ and $e_{cd}$ is the antenna radiation efficiency, which is used to relate the gain and directivity and the relation is-

$$Gain(dB) = e_{cd}(dB) + Directivity(dB) \quad (2)$$

$$e_{cd}(dB) = 10log_{10}(e_{cd}) \quad (3)$$

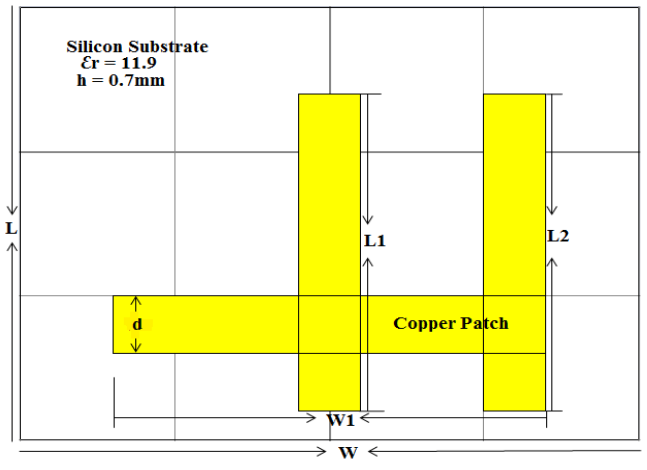
The organization of this paper is as follows- Section II conducts the design of the proposed antenna system. Section III deals with simulation results obtained from CST Microwave Studio. Section IV represents result analysis. Section V conducts link budget optimization. Section VI resembles the cost calculation of the proposed antenna system. Finally, Section VII provides some concluding remarks.

## II. ANTENNA DESIGN

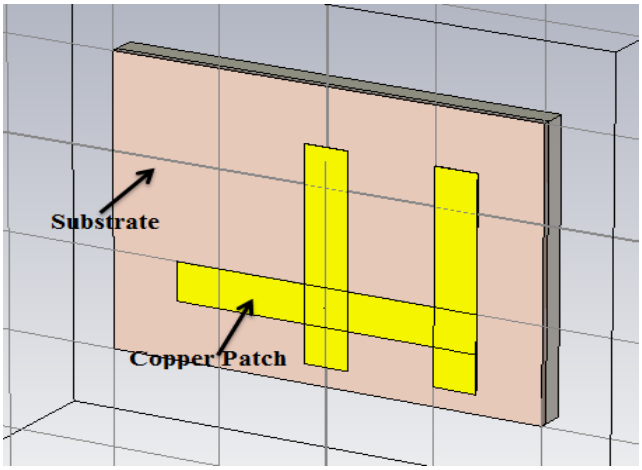
The silicon (lossy) material is used as substrate for the proposed antenna and the substrate area is $40\times30mm^2$. The relative permittivity ($\varepsilon_r$), relative permeability ($\mu_r$) and electrical conductivity of the substrate material are 11.9, 1.0 and $2.5\times10^4$ (S/m) respectively. The copper (annealed) material is used as the radiation patch element. The relative permeability ($\mu_r$), electrical conductivity and relative density (ρ) of this copper (annealed) element are 1.0, $5.8\times10^7$ (S/m) and 8900 ($kg/m^3$) respectively. Fig. 1 shows the structure of the front view and perspective view of the proposed antenna optimized for 10 GHz. The thickness of the patch is 0.07mm. The ground plane is 2.1mm thick and the substrate plate is 0.7mm thick. The size of an antenna can be determined approximately by the following equation-

$$f_0 = \frac{c}{4(a+b)} \quad (4)$$

where $c$ is the velocity of light, $a$ and $b$ are the width and length of the radiating element and $f_0$ is the operating frequency. Table I shows the antenna variables and their values.




(a)




(b)

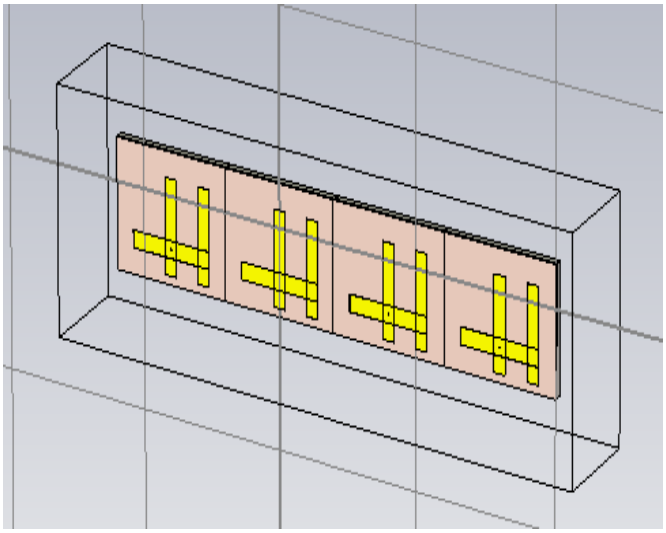
Figure 1. Structure of the proposed antenna system (for single element)

(a) Dimension and (b) Substrate

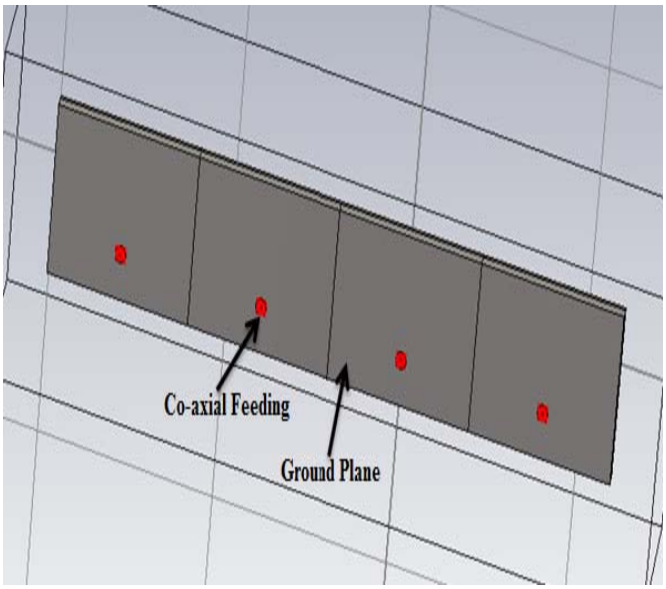
TABLE I. ANTENNA VARIABLES AND THEIR VALUES

| Antenna Variables | W | L | W1 | L1 | L2 | d | h | Mt |
|---|---|---|---|---|---|---|---|---|
| Values (mm) | 40 | 30 | 28 | 22 | 22 | 4 | 0.7 | 0.07 |

The antenna performance changes with the variation of these variables. Fig. 2 shows the structure of proposed 4×1 Inverted F shaped MIMO antenna array. In Fig. 2(b), the back view of the antenna array is shown. The ground plane and the co-axial feeding of the array have been shown. Four waveguide ports are arranged for the 4-element array feeding named port 1, 2, 3 and 4 respectively. The results are obtained by simultaneously exciting all the 4-elements of the array. The phase difference between the excitation signals is 0 deg.



(a)




(b)

Figure 2. 4×1 MIMO PIFA array structure (a) Front View (b) Back View.

## III. Simulation Results

### A. Voltage Standing Wave Ratio (VSWR) and Return Loss

Fig. 3 shows the graph of VSWR vs. Frequency. From this figure it can be seen that the VSWR value is very near to the unity which is mostly expected. The VSWR value of the PIFA antenna is 1.029987 at 10.168 GHz. From Fig. 4 it is seen that the value of return loss is -36.61 dB at 10.168 GHz.

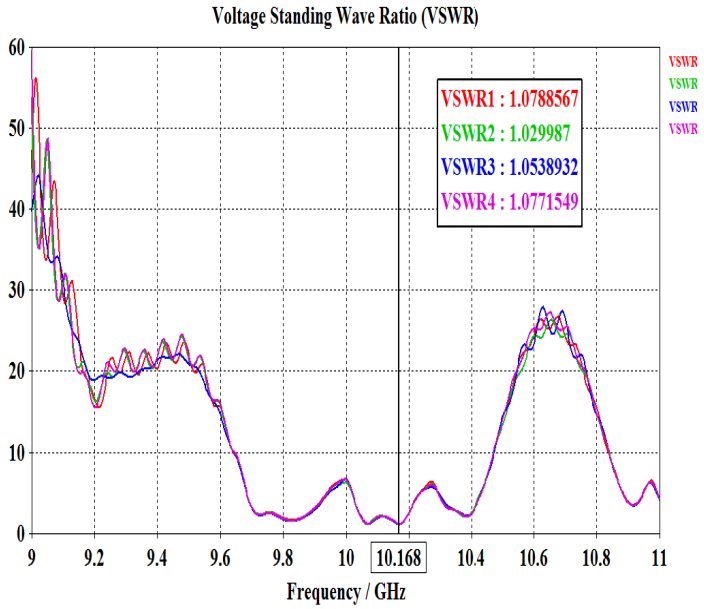



Figure 3. VSWR of the PIFA array.

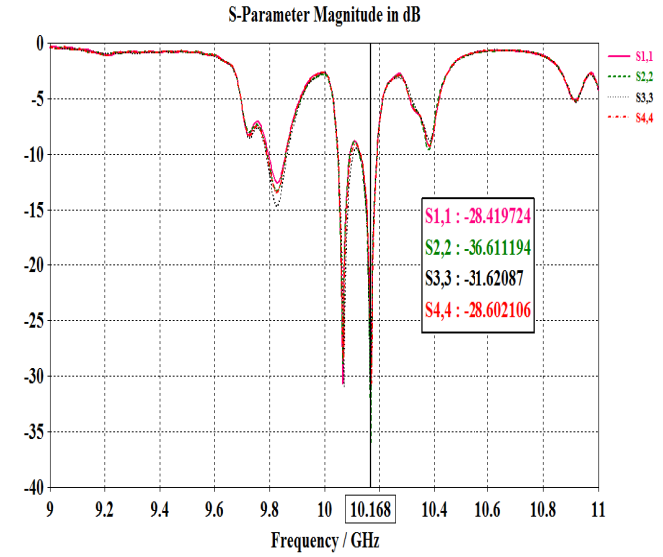



Figure 4. Return loss of the PIFA array.

### B. Farfield Radiation Pattern (3D), Gain Pattern and Antenna Efficiency

Fig. 5 shows the farfield radiation pattern of the PIFA antenna array. From this figure it can be seen that the directivity and gain of the antenna is about 13.24 dBi and 12.87 dB respectively. The radiation efficiency is -0.3655dB. The total antenna efficiency can be calculated as 91.90%.

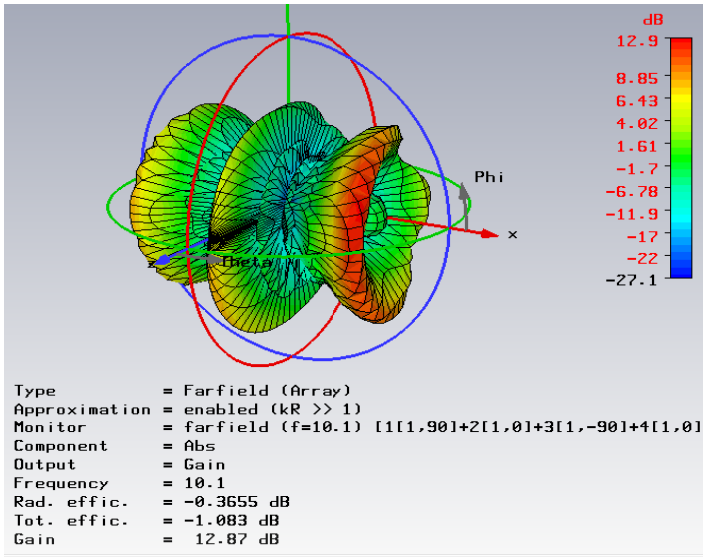



Figure 5. Farfield radiation pattern.

Fig. 6 shows the farfield gain pattern of the proposed PIFA antenna array. The main lobe magnitude is 7.3 dB, main lobe direction is 0.0 deg., 3dB angular beamwidth is 64.4 deg. and side lobe level is -10.9dB.

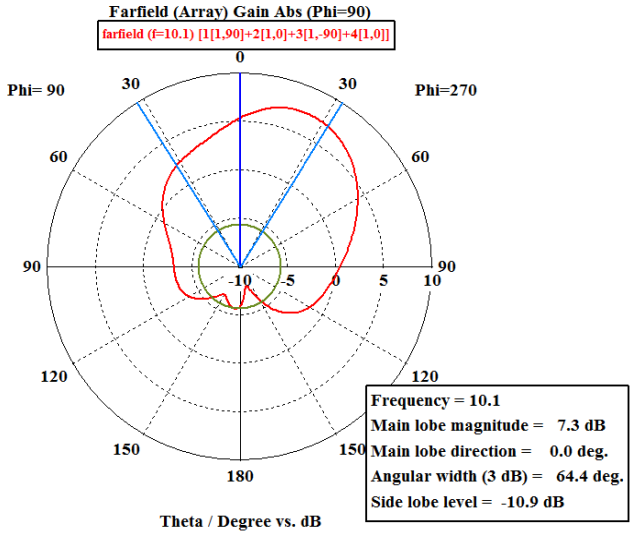



Figure 6. Farfield gain pattern.

## IV. Result Analysis

VSWR, return loss, farfield radiation pattern, antenna gain, directivity and antenna efficiency of the proposed antenna array show reasonable characteristics. The antenna has a VSWR of almost unity (1.029987) and very good return loss value (-36.61 dB), gain (12.87 dB) and a very good value of directivity (13.24 dBi). The antenna efficiency is reasonably good (91.90%). The farfield radiation pattern shows a good 3 dB angular beamwidth (64.4 deg.). The performance of the antenna is quite good. This antenna can be used in the X-band applications for its good performance.

TABLE II. Designed Pifa Antenna Array Parameters and Their Values at A Glance

| **Designed PIFA Antenna Parameters** | **Simulation Results** |
|---|---|
| VSWR | 1.029987 |
| Return Loss (in dB) | -36.61 dB |
| Gain (in dB) | 12.87 dB |
| Directivity (in dBi) | 13.24 dBi |
| 3dB Angular Beamwidth (in deg.) | 64.4 deg. |
| Antenna Efficiency (%) | 91.90% |
| Bandwidth (MHz) | X- band→240 MHz |

The results in the Table II reveal that the designed 4×1 MIMO Inverted F-shaped patch antenna array is a good antenna for X-band applications like satellite communication, radar, medical applications, and other wireless systems.

## V. Link Budget Optimization

The results obtained from the link budget calculator are shown below [8]-

### *A. Uplink Budget*

| | |
|---|---|
| Uplink frequency GHz | 6 |
| Uplink antenna diameter m | 0.08 |
| Uplink antenna aperture efficiency e.g. 0.65 | 0.919 |
| Uplink antenna transmit gain dBi | 13.65854 |
| Uplink antenna, power at the feed W | 850 |
| Uplink EIRP dBW | 42.95273 |
| Range (35778 - 41679) km | 38500.0 |
| Uplink path loss dB | 199.7222 |
| Uplink pfd at satellite dBW/m^2 | -119.756 |
| Bandwidth Hz | 240000000 |
| Satellite uplink G/T dB/K | 40.938 |
| Uplink C/N dB | 28.96638 |

### *B. Downlink Budget*

| | |
|---|---|
| Downlink frequency GHz | 4 |
| Downlink receive antenna diameter m | 0.08 |
| Downlink receive antenna aperture efficiency e.g. 0.65 | 0.919 |
| Downlink system noise temperature (antenna+LNA) K | 120 |
| Downlink receive antenna gain dBi | 10.13671 |
| Downlink receive antenna G/T dB/K | -10.6550 |
| Downlink satellite EIRP dBW | 35.5 |
| Downlink path loss dB | 196.2004 |
| Downlink C/N dB | -26.5576 |

The results obtained from the link budget calculator show that our proposed antenna system is practically realizable by showing similarity between the gain obtained by simulation and the gain obtained by the calculator which is approximately equal to 13 dB.

## VI. COST CALCULATION

Large antennas are expensive to construct and install, with costs exceeding $1M for 30m diameter fully steerable antennas [9]. The cost of large fully steerable antennas has been quoted as [10]-

$$\text{Cost} = \$\ y(D)^{2.7} \quad (5)$$

where, D is the diameter of the antenna aperture in feet.The constant y in equation (5) depends on the currency used and inflation, but might typically be around five U.S. dollars in the early 1980s.

The diameter of the proposed PIFA antenna array is 0.08m i.e.; 0.262467191 feet which in turn gives-

$$\begin{aligned}\text{Cost} &= \$\ y(D)^{2.7} \\ &= \$\ \{5 \times (0.262467191)^{2.7}\} \\ &= \$\ 0.2\end{aligned}$$

## VII. CONCLUSION

The design of a PIFA antenna array and its performance analysis for X-band applications along with the link budget optimization has been illustrated in this paper. The simulation results of the proposed PIFA antenna resemble very good performance. The results obtained from the link budget calculator show that our proposed antenna system is practically realizable by showing similarity between the gain obtained by simulation and the gain obtained by the calculator. Also the cost calculation reveals that our proposed antenna system is cost effective. In this paper, CST Microwave Studio software has been used for all the simulations which provide effective and satisfactory results. The proposed antenna system provides high gain, directivity, half power beamwidth, efficiency and bandwidth. It also shows very low value in case of VSWR which is near about unity that satisfies the antenna specification.

## VIII. FUTURE WORK

We will try to design an Ultra-Wide band satellite antenna system and to develop practical infrastructure of the above designed antenna system.

## REFERENCES


[1] James, J. R., and P. S. Hall, *Handbook of Microstrip Antennas*, Peter Peregronic Ltd., London, 1989.

[2] A. Sabban. (1981) Wideband Microstrip Antenna Arrays, IEEE Antenna and Propagation Symposium MELCOM, Tel-Aviv, 1981. [Online]. Available: http:// www.intechopen.com/

[3] A. Sabban. (1983) A New Wideband Stacked Microstrip Antenna, IEEE Antenna and Propagation Symposium, Houston, Texas, U.S.A, June 1983. [Online]. Available: http:// www.intechopen.com/

[4] A. Sabban, and E. Navon. (1983) A MM-Waves Microstrip Antenna Array, IEEE Symposium, Tel-Aviv, March 1983. [Online]. Available: http:// www.intechopen.com/

[5] N. C. Athanasopoulos, and N. K. Uzunoglu, "Development of a 10 GHz Phased Array Cylindrical Antenna System incorporating IF Phase Processing," *Progress in Electromagnetics Research, PIER* 59, 17–38, 2006.

[6] Adel M. Abdin, "12 GHz Planar Array Antenna for Satellite Communication," *PIERS Online*, Vol. 4, No. 4, 2008.

[7] Constantine A. Balanis, *Antenna Theory: Analysis and Design,* ISBN: 0-471-66782-X.

[8] Satellite Link Budget Calculator. [Online]. Available: http://www.satsig.net/ linkbugt.htm/

[9] Timothy Pratt et al., *Satellite Communications*, ISBN: 9971-513-22-6.

[10] N.E. Feldman, "The link from a communications satellite to a small ground terminal," *Microwave journal*, 7, 39-44, 1964.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Priority Based Adaptive Guard Channel for Multi-class Traffic in Wireless Networks

Rahul Bhattacharjee, Tahsin Ahmed Chowdhury, and Mostafa Zaman Chowdhury
Department of Electrical and Electronic Engineering
Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh
E-mail: rahul0803036@gmail.com, tawrinkle@gmail.com, mzceee@gmail.com

*Abstract*— **An increasing demand of multi-class traffic has become a prime concern in wireless communication system. Quality of service is seen to be degrading to fulfill these demands. To handle multi-class traffic, already different schemes exist such as non-priority scheme, fixed guard channel reservation scheme, queuing scheme, etc. These schemes do not meet the requirements of minimizing the new call blocking probability and increasing the channel utilization at the same time. Here, we are proposing a new adaptive guard channel scheme for multi-class traffic system to minimize both the new call blocking probability of higher priority calls and keeping the channel utilization in a suitable range so that, both the performance metrics are optimized and thus provides good quality of service. In our adaptive guard channel scheme, guard channels are reserved adaptively on the basis of the arrival rates of the traffic classes and the number of channels occupied. In this paper at first we discuss different schemes for multi-class traffic system, and then discuss different problems in these schemes. Then, we introduce our proposed scheme, mathematical modeling and queuing analysis of it and at last we show the improvement in our scheme through performance analysis.**

**Keywords—Multi-class, wireless network, new call blocking probability, channel utilization, adaptive guard channel.**

## I. INTRODUCTION

Whenever to consider the multi-class traffic system, both the real time traffic e.g. *voice call*, *video call, etc.* and non-real time traffic e.g. *data, message, etc.* come under the consideration [1]. The priority of real time calls is much of important compared to the non-real time calls; which may requires channel allocation for higher priority calls, resulting a degradation of channel utilization [1], [2].

From the origin of the wireless system it is becoming complex from being single-class to multi-class. Once it was only limited to the handling of only single-class traffic such as data or voice. In modern age as the number of traffic is increasing, the system is becoming complex. At the present time, the wireless network has to deal with voice call, video call, data and so many types with limited resources [3], [4].

Multi-class traffic system was introduced with the beginning of 2G network where SMS service was first introduced with voice call service. The 2G network afterwards has been superseded with 2.5G, 2.75G, 3G and today with 4G networks with MMS service, video calls, video conference, internet service, etc. Generally, the services in a multi-class traffic system can be categorized into voice, video, and data calls [5]. Among these services real-time voice and video calls should be prioritized in wireless networks rather than the non-real time data traffic.

For the establishment of successful wireless networks for multi-class traffic system, we need to work with limited quantity of resources. To handle multi-class traffic system, the efficient use of these limited resources is an important concern. To deal with the multi-class traffic different schemes such as- non priority scheme and fixed guard channel scheme have been developed [6]. A dynamic channel allocation scheme [1] and an adaptive bandwidth allocation scheme [2] were also proposed for multi-class traffic system. They are also developed on different basis. While in our proposed scheme the number of guard channel is made adaptive on two factors but in [1] only one factor is considered and in [2] the bandwidth is made adaptive, not the guard channel.

In non-priority scheme there is given no priority to any types of traffic classes [6]. So, blocking probabilities of all the traffic classes are same. As there is given no priority, the blocking probability of higher traffic classes is higher in this scheme.

In fixed guard channel scheme, the real time calls have been given priority by reserving some guard channels for them [6]. In this scheme the higher priority calls such as voice and video calls have been given more priority than data. So by sacrificing some resources of the lower priority classes the blocking probability of the higher priority classes can be reduced. In this case the channel utilization falls. Because in this scheme if the number of higher priority traffic arrival rates are low; then some channels remain empty and these channels cannot be used by the lower priority traffic classes. It causes the reduced utilization of channels.

In our adaptive guard channel scheme, guard channels are reserved adaptively on the basis of the arrival rate of the traffic and the number of channels occupied. In this scheme the blocking probability of the higher priority traffic call reduces than the other schemes along with a better utilization of channels.

The paper is organized as follows. Section II presents the mathematical modeling and queuing analysis of the proposed scheme. Performance analysis of different schemes is presented in Section III. Finally, in Section IV, some concluding remarks are drawn.

## II. ADAPTIVE GUARD CHANNEL SCHEME

We propose an adaptive channel reservation scheme in which we only deal with three traffic classes- class 1, class 2, and class 3. Class 1 has the highest priority and class 3 has the lowest priority. No fixed guard channel is reserved for any classes. Here we introduce some factors on the basis of which the reservation of channels for the different classes is determined.

Let, the total number of channels be $N_1$, fixed number of minimum common channels for class 1 and class 2 be $M_2$ and fixed number of minimum common channels for all classes be $M_3$ for fixed guard channel scheme where, $M_2 > M_3$. These ranges can be varied from $M_2$ to $N_1$ and $M_3$ to $N_1$, respectively in our proposed idea. If average call life time is $1/\mu_c$ and average dwell time is $1/\eta$ then channel holding time can be expressed as [7],

$$\frac{1}{\mu} = \frac{1}{\mu_c + \eta} \tag{1}$$

Let, $\lambda_1$, $\lambda_2$ and $\lambda_3$ are the call arrival rate and $P_{B1}$, $P_{B2}$, $P_{B3}$ are the blocking probability for traffic class 1, class 2 and class 3, respectively.

Consider, factors $f_1$, $f_2$ which calculate the number of channels occupied by class 1 and class 2 traffic, while factors $\alpha_1$, $\alpha_2$ calculate the traffic arrival rates of traffic class 1 and class 2, respectively.

Factors for priority 1 traffic class is,

$$f_1 = \frac{\{(1-P_{B1})\lambda_1\}}{\mu N_1} \tag{2}$$

$$\alpha_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \tag{3}$$

Factors for priority 2 traffic class is,

$$f_2 = \frac{\{(1-P_{B2})\lambda_2\}}{\mu N_1} \tag{4}$$

$$\alpha_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \tag{5}$$

By calculating these factors, the number of guard channels is made adaptive.

Number of channels reserved only for class 1 traffic is,

$$X_1 = f_1 \alpha_1 (N_1 - M_2) \tag{6}$$

Number of channels reserved only for class 1 and class 2 traffic is,

$$X_2 = f_2 \alpha_2 (N_1 - M_3) \tag{7}$$

The number of channels available for the class 2 traffic,

$$N_2 = N_1 - X_1 \tag{8}$$

The number of channels available for the class 3 traffic is,

$$N_3 = N_1 - (X_1 + X_2) \tag{9}$$

Here, $N_2 > N_3$ is always maintained.

The proposed scheme can be modeled as an M/M/S/S queuing system. The Markov Chain for the proposed scheme is shown in Fig. 1

The steady-state probability $P_i$ is found as follows [1]

$$P_i = \begin{cases} \dfrac{(\lambda_1+\lambda_2+\lambda_3)^i}{i!\mu^i} P_0; & 0 \le i \le N_3 \\ \dfrac{(\lambda_1+\lambda_2+\lambda_3)^{N_3}(\lambda_1+\lambda_2)^{i-N_3}}{i!\mu^i} P_0; & N_3 < i \le N_2 \\ \dfrac{(\lambda_1+\lambda_2+\lambda_3)^{N_3}(\lambda_1+\lambda_2)^{N_2-N_3}\lambda_1^{i-N_2}}{i!\mu^i} P_0; & N_2 < i \le N_1 \end{cases} \tag{10}$$

$$P_0 = \left[ \sum_{i=0}^{N_3} \frac{(\lambda_1+\lambda_2+\lambda_3)^i}{i!\mu^i} + \sum_{i=N_3+1}^{N_2} \frac{(\lambda_1+\lambda_2+\lambda_3)^{N_3}(\lambda_1+\lambda_2)^{i-N_3}}{i!\mu^i} + \sum_{i=N_2+1}^{N_1} \frac{(\lambda_1+\lambda_2+\lambda_3)^{N_3}(\lambda_1+\lambda_2)^{N_2-N_3}\lambda_1^{i-N_2}}{i!\mu^i} \right]^{-1} \tag{11}$$

The blocking probability $P_{B1}$ for traffic class 1 is given by,

$$P_{B1} = \frac{(\lambda_1+\lambda_2+\lambda_3)^{N_3}(\lambda_1+\lambda_2)^{N_2-N_3}\lambda_1^{N_1-N_2}}{N_1!\mu^{N_1}} P_0 \tag{12}$$

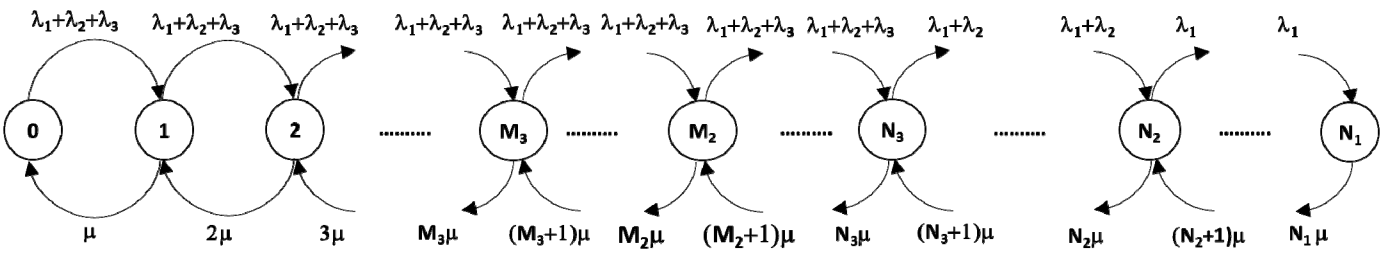



**Fig. 1:** State transition diagram for proposed adaptive guard channel scheme

The blocking probability $P_{B2}$ for traffic class 2 is given by,

$$P_{B2} = \sum_{i=N_2}^{N_1} \frac{(\lambda_1+\lambda_2+\lambda_3)^{N_3}(\lambda_1+\lambda_2)^{N_2-N_3}\lambda_1^{i-N_2}}{i!\mu^i} P_0 \quad (13)$$

The blocking probability $P_{B3}$ for traffic class 3 is given by,

$$P_{B3} = \sum_{i=N_3}^{N_1} P_i$$
$$= P_{B2} + \sum_{i=N_3}^{N_2-1} \frac{(\lambda_1+\lambda_2+\lambda_3)^{N_3}(\lambda_1+\lambda_2)^{i-N_3}}{i!\mu^i} P_0 \quad (14)$$

The % of channel utilization is calculated as,

$$U = \frac{\left[\{(1-P_{B1})\lambda_1\}+\{(1-P_{B2})\lambda_2\}+\{(1-P_{B3})\lambda_3\}\right]}{\mu N_1} \times 100 \quad (15)$$

The guard channel reservation strategy of the proposed scheme is shown in Fig. 2. For any traffic arrived in the cell, at first it searches the minimum number of fixed channels for all the classes whether empty or not and if it finds free channels in the minimum fixed channels range $M_3$, the call will be accepted. If no channel is empty in this range, then by calculating the factors $\alpha_1$, $\alpha_2$, $f_1$, and $f_2$ on the basis of the channels occupied and the arrival rate of the traffic classes, the range $N_3$ and $N_2$ are determined which is the variable channels

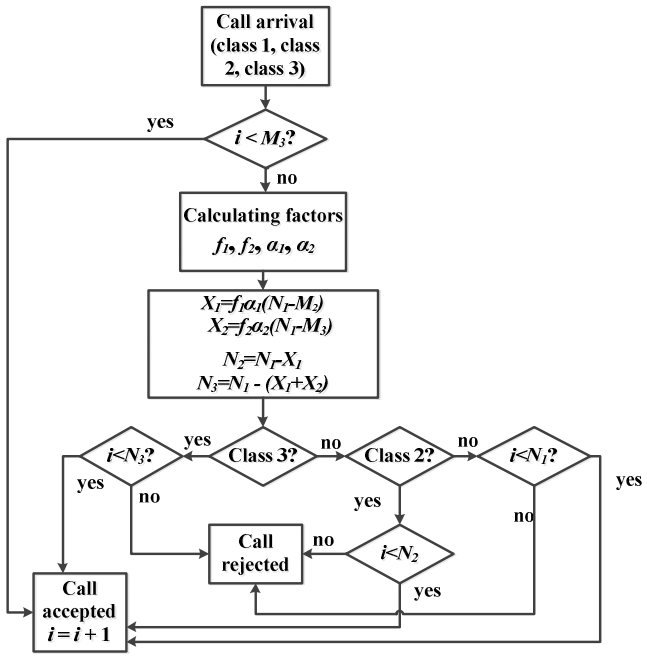



**Fig. 2**: Guard channel reservation strategy of the proposed adaptive scheme.

range available for the priority 3 and priority 2 traffic classes. If the call arrival is of traffic class 3 and no channel is empty in the range $N_3$ then the call will be rejected. When class 2 call arrives then if in the range $N_2$ any channel is empty, the call will be assigned and if no channel is empty then the call is rejected. Now when the call is of class 1 then in the range $N_1$ if no channel is empty, the call will be rejected but if any channel is empty then the call will be assigned.

## III. PERFORMANCE ANALYSIS

In this section, we verified the performance of our proposed scheme. We consider voice, video and data as the priority 1, priority 2 and priority 3 classes respectively for our analysis. $M_3$ and $M_2$ are considered as 70% and 80%, respectively of $N_1$.

Fig. 3 presents comparison of new call blocking probability. Here our proposed scheme shows a better performance than the non-priority and fixed guard channel scheme. For non-priority scheme, the blocking probabilities of all three classes are same. In case of fixed guard channel scheme, the blocking probability of the class 1 and class 2 are lower than the non-priority scheme but for class 3 it is very high. The reason is class 1 and class 2 have been given more priority than the class 3 traffic and the blocking probability of the class 1 and class 2 are improved by sacrificing for class 3. In our proposed scheme, the blocking probability for class 2 and class 3 have been more improved than the fixed guard channel scheme. The blocking probability of the class 1 traffic is very low and almost same as fixed guard channel scheme. So our proposed scheme shows an overall better performance than the non-priority and the fixed guard channel scheme.

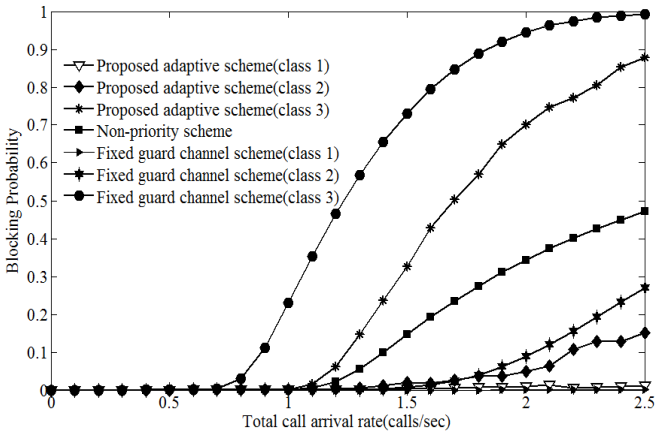



**Fig. 3:** Comparison of new call blocking probability

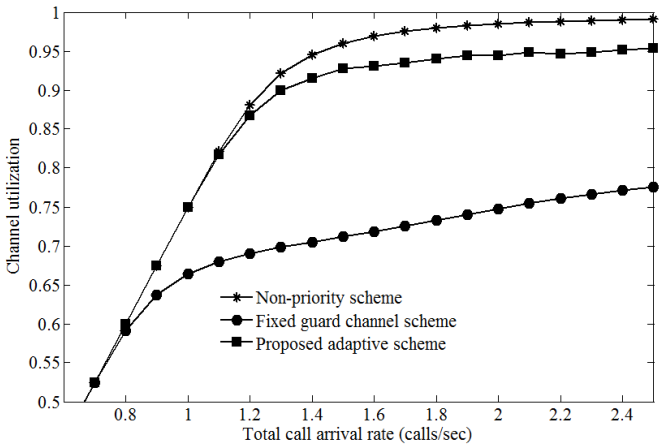



**Fig. 4:** Comparison of channel utilization.

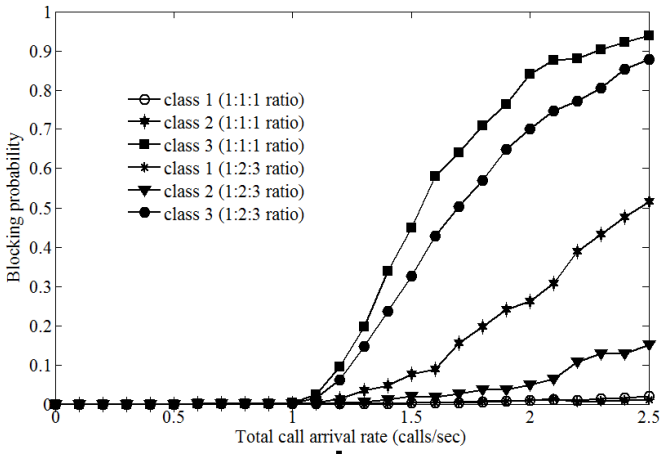



**Fig. 5:** Comparison of blocking probability of the proposed scheme for different ratio of call arrival rates.

Fig. 4 shows the comparison of the channel utilization of the proposed scheme with the non-priority and the fixed guard channel schemes. From this figure we can see that our proposed scheme has a very good utilization than the fixed guard channel scheme and slightly less than the non-priority scheme. The utilization is about 95.5% for the proposed adaptive scheme because some guard channels are always reserved for higher priority traffic classes.

The performance analysis shows different responses at different ratio of the call arrival rate. In Fig. 5, comparison of the blocking probabilities of the proposed scheme at ratio of 1:2:3 and 1:1:1 has been shown. Fig. 6 shows the variation of channel utilization with different call arrival rates in using the proposed scheme. It seems that due to the variation of call arrival rates the channel utilization does not fall significantly.

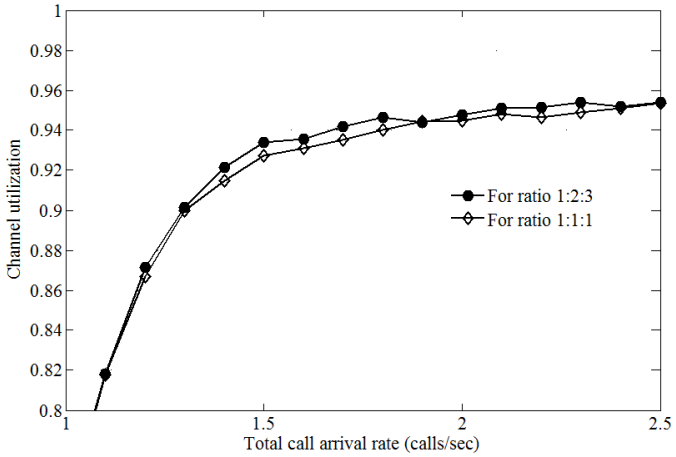



**Fig. 6:** Comparison of the channel utilization for two different ratio of call arrival rates for the proposed scheme.

## IV. CONCLUSION

In comparison with the other schemes we found that our proposed scheme provides a better utilization and minimization of blocking probability of the higher priority classes. In non-priority scheme, utilization of channels is high but as there is given no priority to any of the classes, the blocking is higher. In fixed guard channel scheme priority has been given thus reduces the blocking of higher priority traffics but because of fixed guard channel the utilization reduces to a very low value. In our proposed scheme the guard channels have been made adaptive so that the channel utilization is higher than the fixed guard channel scheme and slightly less than the non-priority scheme along with a reduction of blocking probability.

## REFERENCES


[1] M. Z. Chowdhury, M. S. Uddin, and Y. M Jang, "Dynamic Channel Allocation for QoS Provisioning in Visible Light Communication," *IEEE International Conference on Consumer Electronics(ICCE)*, Las Vegas, U.S.A. , pp. 13-14, January 2011.

[2] M. Z. Chowdhury, Y. M. Jang and Z. J. Haas, "Call Admission Control Based on Adaptive Bandwidth Allocation for Wireless Networks," *Journal of Communications and Networks(JCN)*, vol. 15, no. 1, pp. 15-24, February 2013.

[3] W. C. Y. Lee, *Wireless and Cellular Communications*, McGraw-Hill, 2005.

[4] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd edition, Prentice Hall, 2002.

[5] F. A. Cruz-Pérez and L. Ortigoza-Guerrero, "Flexible Resource Allocation Strategies for Class-Based QoS Provisioning in Mobile Networks," *IEEE Transactions on Vehicular Technology,* vol. 53, No. 3, pp. 805-819, May 2004.

[6] I. Stojmenovic, *Handbook of Wireless Networks and Mobile Computing*, John Wiley & Sons, Inc, 2002.

[7] M. Schwartz, *Mobile Wireless Communications, Cambridge University Press*, Cambridge, 2005.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Comparative Analysis of GNSS Reliability: GPS, GALILEO and Combined GPS-GALILEO

Md Hossam-E-Haider[1], Asma Tabassum[2,*], Rafiul Hossain Shihab[3], and Mahbub Hasan[4]
[1]Electrical, Electronics and Communication Engineering Department
[2, 3, 4] Aeronautical Engineering Department
Military Institute of Science & Technology
Mirpur Cantt, Dhaka
[*] tabassum_1708@yahoo.com

*Abstract*—**Global navigation satellite systems (GNSS) is the collective term for those navigation systems that provide the user with a three-dimensional positioning solution by passive ranging using radio signals transmitted by satellites. The most well known is the Navigation by Satellite Timing and Ranging (NAVSTAR) Global Positioning System (GPS). It is the most advanced but still not free from erroneous positioning and needs better accuracy. Now European GALILEO is under development, which is interoperable with GPS. In this paper performance of GPS and GALILEO alone and combined GPS-GALILEO system has been analyzed. Simulation has also done using dual and triple frequency. This paper has also presented which system and combination gives better precision and reliability. Also a time analysis of getting highest availability is studied. Results show that the improvement of the combined system is as high as 50% as present GPS alone system.**

***Keywords— GPS, GALILEO, Interoperability, Reliabilty, MDB, MDE, BNR, Dual frequency and Triple frequency.***

## I. INTRODUCTION

Satellite navigation, positioning, and timing have already found widespread applications in a large variety of fields. Global positioning system (GPS) is the most advanced among all the satellite navigation systems, which was developed by the U.S. government mainly for military purpose [1]. Now it has found its place in the field of ranging, surveying, navigation, transportation and geodesy. At present application of dual frequency GPS has made a better precision. But still there are situations, where GPS itself sometimes cannot provide required precision and integrity.

Recognizing the strategic importance, the European Union (EU) and European Space Agency (ESA) took initiative to launch their own global navigation satellite system, GALILEO providing a highly accurate, guaranteed global positioning service under civilian control [1]. It is interoperable with GPS. Interoperability means combining information (e.g., pseudo range measurements, navigation data) from both the GPS and GALILEO systems at the user receiver to achieve better performance than employing either system separately [2].

A number of studies have been done on the performance of GNSS system i.e. GPS, GLONASS (Global navigation satellite system which was developed by Russia) and GALILEO in respect of availability of satellite, biases, outliers and least square ambiguities [3-7]. In this paper, performance of GPS, GALILEO and combined GPS-GALILEO has analyzed for some specific case and the percentage of improvement has shown. The paper has also presented the comparison of time required to gain highest availability. Design parameters, which don't need any actual observations [7], are considered in this purpose. Parameters considered are internal reliability, represented as minimal detectable bias (MDB) and external reliability represented as minimal detectable effect (MDE) and bias-to-noise ratio (BNR). Also a comparison of dual and triple frequency ranging system has been carried out. The results show that a triple frequency combined system gives a better performance. A graphical user interface (GUI), VISUAL is used to simulate and analyze the results [8].

## II. STATUS OF GPS AND GALILEO

The U.S. government baseline configuration for the nominal constellation consists of 24 satellites divided unevenly over six orbital planes. The orbital radius (i.e., nominal distance from the center of mass of the Earth to the satellite) is approximately 26,600 km. The orbits are nearly circular and equally spaced around the equator at a 60° separation with a nominal inclination relative to the equatorial plane of 55° [2]. In each plane there are four operational satellites. There are 10 different GPS navigation signals, broadcast across three bands, known as link 1 (L1), link 2 (L2), and link 5 (L5). The carrier frequencies are 1575.42 MHz for L1, 1227.60 MHz for L2 and 1176.45 MHz for L5, while the declared double-sided signal bandwidth is 30.69 MHz in each band. Although at present 28 satellites are operational, we will consider the nominal configuration of 24 satellites.

GALILEO is in its development phase and its initial operational capability is planned for 2010–2012 with full operational capability by 2014 [1]. The fully deployed Galileo system consists of 30 satellites (27 operational + 3 active spares), positioned in three circular Medium Earth Orbit (MEO) planes at 23,222 km altitude above the Earth, and at an inclination of the orbital planes of 56 degrees to the equator. Spare satellites will be kept in each orbital plane, but will not be brought into operation until a satellite needs replacing [9]. So far four operational satellites launched - the basic minimum for satellite navigation in principle. Galileo broadcasts 10 different navigation signals across three frequency bands: E5, E6, and E1-L1-E2. The E5 band is 92.07 MHz (90 × 1.023

MHz) wide and centered at 1,191.795 MHz [10]. It is partitioned into E5a and E5b sub bands, with carrier frequencies of 1,176.45 and 1,207.14 MHz, respectively. The E6 and E1-L1-E2 bands are both 40.92 MHz wide and centered at 1278.75 and 1575.42 MHz, respectively. In our paper nominal constellation of 27 satellites without three active spares has been considered.

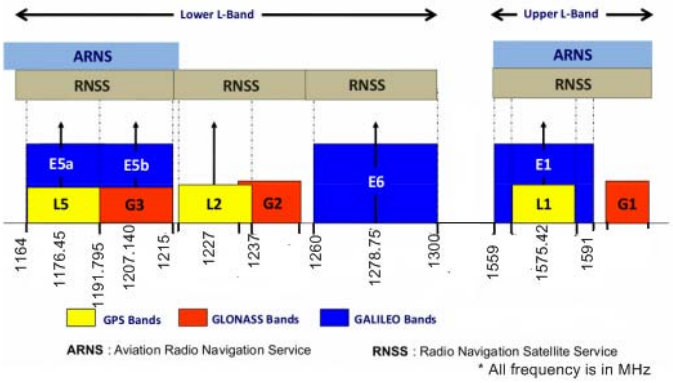



Fig 1: Frequency span of GNSS system

## III. Satellite Navigation Observation Models

The observations of satellite navigation are pseudo range code p and carrier phase Φ. Observation can be of single difference (observation of single receiver) and double difference (observations collected by two receivers). If we consider double difference (DD) observations, equations for satellite r and s can be given by [11]:

$$p_i^{rs}(t) = \rho^{rs}(t) + \frac{f_{L1}^2}{f_i^2} I^{rs}(t) + e_i^{rs}(t) \quad (1)$$

$$\Phi_i^{rs}(t) = \rho^{rs}(t) - \frac{f_{L1}^2}{f_i^2} I^{rs}(t) + \lambda_i N_i^{rs} + \varepsilon_i^{rs}(t) \quad (2)$$

These equations hold for both GPS and GALILEO. Unknown parameters are the DD satellite-receiver range, ρ, the DD ionospheric effect, I, the DD integer carrier ambiguities, N, the measurement noise, e and ε, of code and phase respectively. The carrier wavelengths are denoted by, λ. And i is the frequency any L-band/ E-band frequency. Geometry free (GF) baseline model is considered, where observation equations remain parameterized in terms of the unknown DD receiver-satellite ranges.

Any least square adjustment deals with two equally important components, the mathematical model and the stochastic model.

### A. *Mathematical Model:*

For a single epoch the complete DD mathematical measurement model for two receivers observing m satellites, on f frequencies is:

$$y_k = (e_{2f} \otimes I_{m-1})\rho_k + (C_2 \otimes I_{m-1})a + n_k \quad (3)$$

where,

$C_2 = c_2 \otimes I_f$;

$y_k = (p_i(k)^T, \Phi_i(k)^T)^T$; $\rho_k = \rho(k)$; $a = (\lambda_i N_i)^T$;

$n_k = (e_i(k)^T, \varepsilon_i(k)^T)^T$

⊗ Kronecker product,

$$M_{p\times q} \otimes N = \begin{pmatrix} m_{11}N & \cdots & m_{1q}N \\ \vdots & \ddots & \vdots \\ m_{p1}N & \cdots & m_{pq}N \end{pmatrix}$$

The baseline model for k epochs of data can be written as:

$$y = [I_k \otimes M \;\; e_k \otimes N] + n \quad (4)$$

$$M = e_{2f} \otimes I_{m-1}; N = C_2 \otimes I_{m-1} \quad . \quad (5)$$

The observations are collected by type, first the code observations and then the phase observations.

The combined observation model for one epoch is written as:

$$\begin{pmatrix} y_{GPS} \\ y_{GAL} \end{pmatrix} = \begin{pmatrix} M_{GPS} & N_{GPS} & \\ M_{GAL} & & N_{GAL} \end{pmatrix} \begin{pmatrix} b \\ a_{GPS} \\ a_{GAL} \end{pmatrix} + n \quad (6)$$

where subscript GPS refers to GPS observations and GAL refers to Galileo observations. For geometry-free model the unknowns, b is simply the ranges ρ. Double differences are formed with respect to a GPS satellite and a Galileo satellite respectively.

### B. *Stochastic Model:*

The variance-covariance (vc) matrix of DD satellite observations is:

$$Q_y = I_k \otimes C \otimes E \quad (7)$$

where,

$$E = D^T D$$

The vc-matrix that corresponds to the combined GPS and GALILEO is given by [7]:

$$Q_y = I_k \otimes \begin{pmatrix} C_{GPS} \otimes E_{GPS} & \\ & C_{GAL} \otimes E_{GAL} \end{pmatrix} \quad (8)$$

Note that,

$$D^T = \begin{pmatrix} D_{GPS}^T & \\ & D_{GAL}^T \end{pmatrix} \quad (9)$$

## IV. Reliability of Network

Reliability of a network refers to how consistent that network is. Reliability combines two aspects: Internal reliability and external reliability. Internal reliability referred to the ability of a system to detect outlier and the measures of the influence of outlier below the detectable bias threshold on parameter estimation, referred to as external reliability.

### A. *Internal Reliability:*

The internal reliability is determined by the minimum detectable bias (MDB), which is the magnitude of smallest error that can be detected for a specific level of confidence. In another words, MDB represents a value of the least observation error possible to be detected using statistical test [12]. High MDB indicates a small possibility of observation or

coordinate control; hence the higher MDB, the lower is the network reliability. Consider the following null hypothesis and alternative hypothesis:

$$H_0: E\{y\} = Ax;\ Q_y$$
$$H_a: E\{y\} = Ax + c\nabla;\ Q_y$$

with $E\{.\}$ the expectation operator, y the m-vector of normally distributed observables, A the design matrix, x the n-vector of unknown parameters, and c the unknown bias vector. The bias vector is assumed to describe the model error. Hence it is absent under the null-hypothesis, but present under the alternative hypothesis. Corresponding size of the bias can be obtained as:

$$|\nabla| = \sqrt{\frac{\lambda_0}{c^T Q_y^{-1}(I_m - P_A)c}} \qquad (10)$$

$P_A$ is the orthogonal projector on the range space of A.

Using the properties of the kronecker product and orthogonal projectors, the MDB for a code outlier is given by:

$$|\nabla_p| = \sqrt{\frac{\lambda_0}{d^T Q_y^{-1}\left(I_{m-1} - \left(1 - \frac{1}{k}\right)P_M - \frac{1}{k}\left(P_N + P_{P_N^{\perp}M}\right)\right)d}} \qquad (11)$$

*B. External Reliability:*

External reliability, applied to determine systematic error of observation on adjusted coordinate. A dimensionless parameter BNR (bias to noise ratio) is used to express external reliability. External reliability can also be determined by MDE (minimal detectable effect). MDE is a vector which describes the impact of an MDB sized bias in the observations, on each of the unknown parameters to be estimated. It follows from equation (10) as:

$$\nabla\hat{x} = (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} c\nabla = Q_{\hat{x}} A^T Q_y^{-1} c\nabla \qquad (12)$$

In this GPS model, if we consider only the impact on the unknown baseline coordinates, then the corresponding BNR can be computed by subtracting the contribution of the ambiguities. The resulting squared BNRs, written in a similar form as equation (11) and equation (12) are:

$$\lambda_{\hat{b}} = v.\nabla^2\left(d^T Q_y^{-1} d - \frac{v}{k} d^T Q_y^{-1} P_N d\right) - \lambda_0 \qquad (13)$$

## V. Simulation & Analysis

*A. Simulation:*

A Matlab® software tool, VISUAL, is used to analyze the result. The input parameters chosen for the simulations are:

- System: GPS, GALILEO and combined GPS-GALILEO;
- Almanac: a Yuma almanac file for current GPS [13] and GALILEO ;
- Time and date: 15-08-2013 0:00h (spatial variation) or 0:00-0:24h (temporal variation);
- Number of epochs: 1-300s;
- Frequencies : 2 or 3;
- Cutoff elevation: 15° ;
- Ionospheric model: weighted, s=0.02m ;
- Tropospheric model: fixed;
- Receiver type: Stationary;
- Baseline model: Geometry Free;
- Location: World (spatial variations), Dhaka 23.7000° N, 90.3833° E (temporal variations).

MDB, MDE and BNR for code outlier on L1 or E1 are the output parameters.

*B. Result Analysis:*

We analyze the performance for two types of variations; one is spatial variations where simulation has done for entire earth and other is temporal variation where one specific location has been chosen for the same day.

*a) Spatial variation:* The spatial variation of the MDBs , MDEs and BNRs have been studied for the world. The results showing the MDBs of 3 configuartions for dual frequency are shown in fig-2.

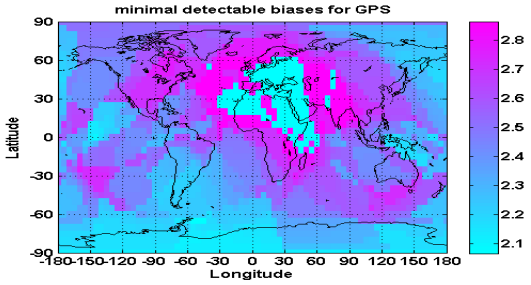



Fig-2(a) Minimal detectable bias of GPS

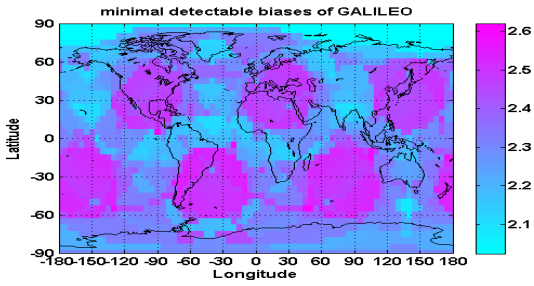



Fig-2(b): Minimal detectable bias of GALILEO

Analyzing fig-2(a) and fig-2(b), it has found that for GPS in most places MDB lies between 2.1 ~ 2.4 and highest value is 2.8, for GALILEO the value lies between 2.2 ~ 2.4 in most places and highest 2.6 in some places.

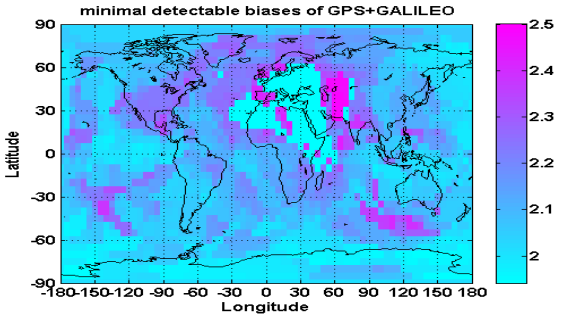


Fig-2(c): Minimal detectable bias of combined system

In case of combined system highest MDB reduced to 2.5. However, for GALILEO and combined GPS-GALILEO system this improvement is marginal. The maximums and average values of MDB obtained with GALILEO are lower than with GPS. This is because of the even distribution of GALILEO satellites over three plans, where the GPS satellites are divided over six planes which are not evenly distributed.

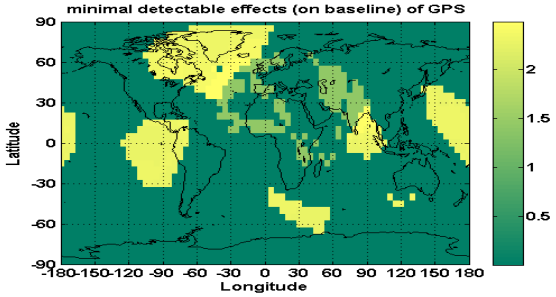


Fig-3(a): Minimal detectable effects of GPS

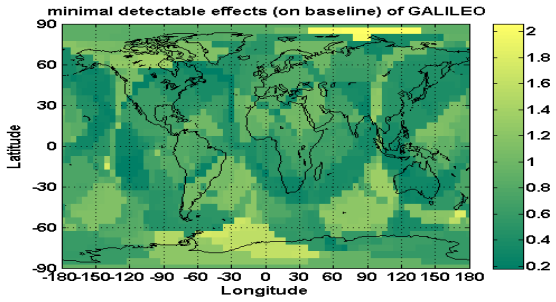


Fig-3(b): Minimal detectable effects of GALILEO

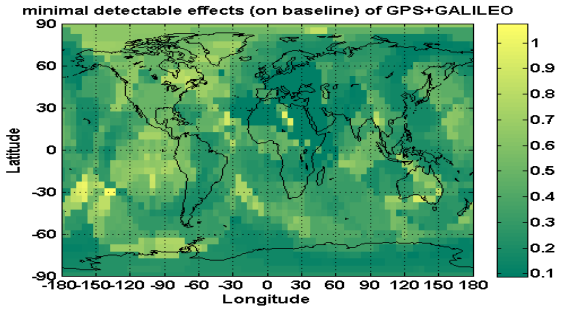


Fig-3(c): Minimal detectable effects of combined system

A similar observation is done for the MDEs for either code outlier in L1 GPS observation, or a code outlier in an E1 GALILEO observation. Fig-3 shows the MDE for different systems. MDE which indicates the impact of MDB on the observation also reveals the improved reliability of combined system. The highest value of combined sytem is reduced to 1, where GPS and GALILEO alone system gives a highest value of 2. The minimum value of combined system is as low as 0.1. GALILEO gives better reliability than GPS.

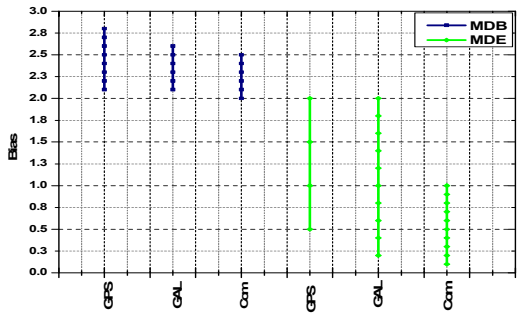


Fig-4: Comaparison of MDB and MDE of three system

In comparson with the GPS alone solution, the MDE of the combined GPS-GALILEO system is improved almost 50%. Fig-4 summerize the overall scenary of all three combinations where the range of errors has shown.

*b) Temporal Variation:* The temporal variation is done at the same day for a specific location, Dhaka. Fig-4 and fig-5 display the variation of MDB and BNR with satellite availability. Simulation for triple frequency combination has also carried out.

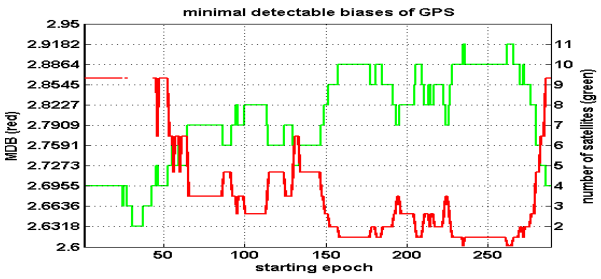


Fig-5(a): Minimal detectable bias of GPS

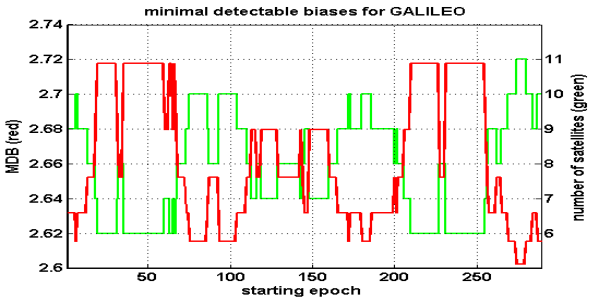


Fig-5(b): Minimal detectable bias of GALILEO

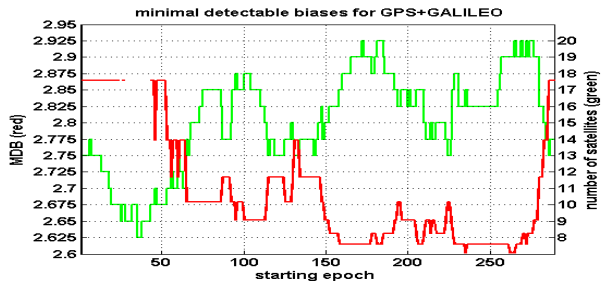



Fig-5(c): Minimal detectable bias of combined system

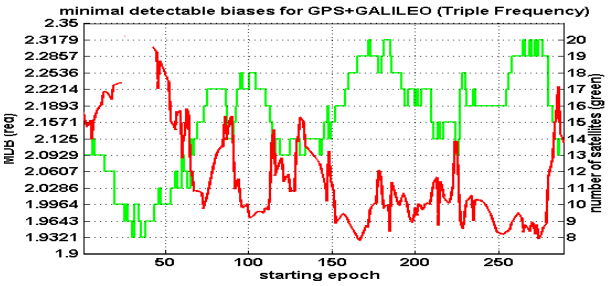



Fig-5(d): Minimal detectable bias of combined system with a third frequency added

It is clear from the results that combined system offers more availability then single system as the combined constellation consist more satellites. Satellite available in combined system is more than double than GPS alone system. Gaps in the output are found when the no. of satellite is less than 8 indicating worse reliability. Reliability increases when more satellites are available. Analyzing results it has obtained that the improvement of MDB of the combined system with dual frequency is marginal. When we add a third frequency in combined system huge improvements is apparent. The lowest value of MDB is 1.9 when maximum no. of satellites is available. Adding a third frequency in the combined system improved over 50% in comparison to single system.

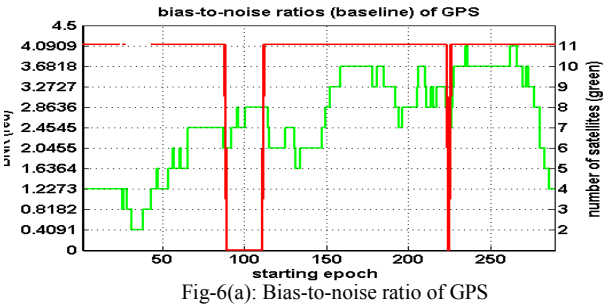



Fig-6(a): Bias-to-noise ratio of GPS

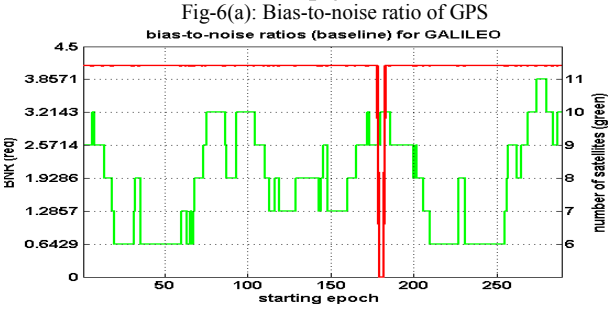



Fig-6(b): Bias-to-noise ratio of GALILEO

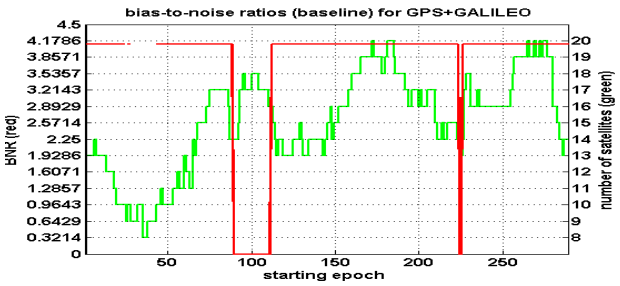



Fig-6(c): Bias-to-noise ratio of combined system

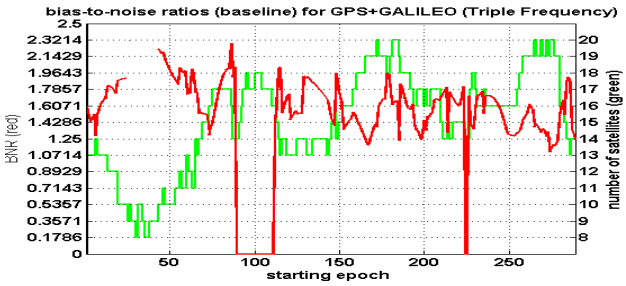



Fig-6(d): Bias-to-noise ratio of combined system with a third frequency added

In case of BNR combined system with dual frequency improvement is marginal where with triple combination system improved over 44%.

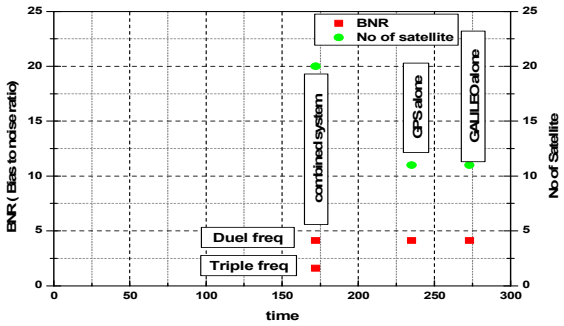



Fig-7: Comparison of minimum time to get highest no. satellite and BNR at that moment

Fig-7 shows a comparison between three combinations in respect of minimum time to get highest availability. It reveals that a combined system takes much less time than GPS and GALILEO alone system and in a triple frequency combination BNR reduces to 1.597 from 4.12 which is the highest value of GPS alone system.

## VI. CONCLUSION

A comparative analysis of internal and external reliability of various systems is presented and compared to detect the possible best solution. The minimum time needs to get highest availability for a specific case has shown. The combined system takes minimum time to achieve highest availability.

Simulation also depicts that a combined system gives better performance regards to reliability than GPS alone solutions, where as the improvements are marginal compared to

GALILEO. If we add a third frequency the combined system works far better than current GPS alone solution and the improvements is as high as 50%.

## ACKNOWLEDGMENT

The authors would like to thank Dr. Sandra Verhagen for providing the Simulation software and the constellation information of GALILEO.

## REFERENCES

[1] *Principles of GNSS, Inertial and Multisensor Integrated system.* Paul D. Groves, Artech House Inc, 2008.

[2] *Understanding GPS, Principles and Applications.* Elliott D. Kaplan, Christopher J. Hagerty, Artech House, Inc, 2006.

[3] Hawkins D.M. *Identifications of outliers,* Chapman & Hall, London/ New York, 1980.

[4] Gao Y. *Reliability assurance for GPS integrity test, ION GPS 1992,* Salt Lake City, Utah, September 22-24. 1992, 567-574.

[5] Teunissen, P.J.G. *Minimal detectable biases of GPS data. Journal of Geodesy*, vol. 72, 236-244, 1998.

[6] Salgado, G., S. Abbondanza, R Blondel, and S. Lannelongue, *Constellation availability concepts for Galileo. Proc. Of ION NTM,* Long Beach CA, January 22-24. 2001, 778-786.

[7] Verhagen, S. *Performance Analysis of GPS, Galileo and Integrated GPS-Galileo, ION GPS 2002*, Portland, Oregon, September 24-27. 2002, 2208-2215.

[8] Verhagen, S. *A New Software tool: Studying the Performance of Global Navigation Satellite Systems. GPS World June 2002.*

[9] Dinwiddy, S. E., E. Breeuwer, and J. H. Hahn, *''The Galileo System,'' Proc. ENC-GNSS 2004,* Rotterdum, Netherlands, 2004.

[10] J Ruiz, L., R. Crescinberi, and E. Breeuwer, ''*Galileo Services Definition and Navigation Performance,'' Proc. ENC-GNSS 2004,* Rotterdam, the Netherlands, 2004.

[11] *GPS Satellite Surveying,* Alfred Leick, John Wiley & Sons, Inc

[12] Baarda, W. *A testing procedure for use in geodetic networks. Netherlands Geodetic Commission, Publications on Geodesy*, vol.2, no.5, Jan 1968.

[13] The U.S. Coast Guard Navigation Centre website. Current almanac file. [Online]. Available: http://navcen.uscg.gov/?pageName=gpsAlmanacs

# Design of a Circular Polarization Switchable Microstrip Array Antenna using Magic-T Bias Circuit

Md. Azad Hossain[1], Piyas Chowdhury[1,*], Quazi Delwar Hossain[1],
Eisuke Nishiyama[2] and Ichihiko Toyoda[2]
[1]Chittagong University of Engineering and Technology, CUET, Chittagong-4349, Bangladesh
[2]Saga University, 1 Honjo-machi, Saga-shi, Saga 840-8502, Japan
[*] E-mail: {azad, piyas, headete}@.cuet.ac.bd, nisiyama@ceng.ec. saga-u.ac.jp, toyoda@cc.saga-u.ac.jp

*Abstract*—**In this paper, a 4-element circular polarization switchable patch array antenna is proposed. The array antenna is realized by the successful employment of the both-sided MIC technology, where the odd and even mode of the slot line and microstrip lines are used for the realization of the orthogonal feed and magic-T switching circuit respectively. The advantage of the magic-T circuit is the excellent isolation between the RF signal and the switching signal. In order to realize the polarization switching, two switching diodes are integrated with each patch elements. Unequal feed line is introduced to realize 90º phase difference between the patch elements and this way, circular polarization switchable array antenna can be realized. The array antenna unit is realized in very simple and compact structure as all the antenna elements, feeding circuit and switching circuit are arranged on both sides of a Teflon glass fiber substrate. The structure and the basic behavior of the circular polarization switchable array antenna are demonstrated in this paper. In addition, the simulated S-parameters and the phase difference of the feed circuit are explained.**

***Keywords—Both Sided MIC technology, Polarization switching, Planar array antenna.***

## I. INTRODUCTION

Microstrip antennas pose some attractive features for modern communication system, such as low profile, light weight, ease of manufacturing and conformability [1]-[3]. Among the microstrip antennas, circular polarization microstrip antennas are more suitable because of their inherent capabilities of reduced multi-path fading, improved coverage and fixed polarization [4]-[5]. The distinct advantage of the circular polarization antenna over the linear polarization antenna is the orientation flexibility [6]. Some basic studies of the circular polarization antenna have been reported [1]-[4],[7]-[9]. In [1]-[3], the circular polarization and the polarization diversity are achieved by using PIN diode. There are also some reports for the circular polarization switching [10]-[11], In these reports, the circular polarization switching is realized by the integration of the SPDT switch circuit and the 90º hybrid circuit with the antenna elements. In this proposed research, the authors have proposed a novel circular polarization array antenna using the linear polarization patches. In 2012, the authors have reported a 4-element linear polarization switchable microstrip array antenna using Magic-T bias circuit [12]. In this report, the linear polarization switching is achieved by the ON/OFF condition of the PIN diodes. The authors have modified the diode orientation and the feed circuit in order to realize the circular polarization switchable array antenna. Microwave momentum of the Advanced Design System (ADS) is used for the simulation of the feed circuit and the bias circuit of the array antenna elements. The simulation of the antenna elements integrated with two switching diodes is simulated by the EMPro software. In section 2, the structure of the array antenna, Magic-T and the basic behavior of the Magic-T is explained. In section 3, the basic behavior of the array antenna is explained. Brief discussion is made regarding the simulation results.

## II. STRUCTURE

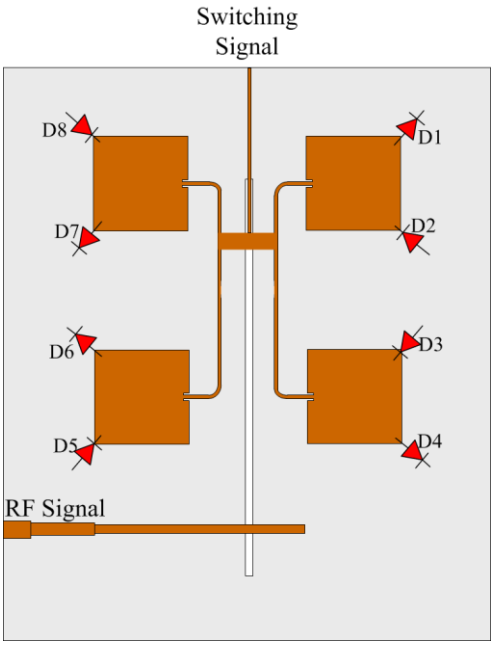



Figure 1. The proposed array antenna.

Fig. 1 shows the schematic structure of the proposed array antenna. The proposed array antenna consists of four patch elements and 2 switching diodes are loaded on each patch element totaling in 8 diodes. The feed and bias circuits are realized by arranging the magic-T and air bridge as seen in the dotted square of Fig. 1. The antenna elements, diodes and microstrip lines are arranged on the obverse side and the slot line is arranged on the reverse side of a Teflon glass fiber substrate. The position of the Magic-T is arranged in such way that there is a quarter wave length physical length difference between the patch #1,#4 pair and the patch #2,#3 pair of the array antenna. The diodes D1, D4, D6 and D7 are arranged outward and the diodes D2, D3, D5 and D8 are arranged inward to the antenna.

The feed circuit is realized by the employment of the micristrip-slot parallel branch circuit and the slot-microstrip series branch circuit. This type of array antenna needs no impedance matching circuits and has a very simple circuit configuration mainly due to the excellent combination of both the microstrip-slot parallel branch circuit and the slot-microstrip series branch circuit [13]-[15]. However, the design of this feed circuit for impedance matching is clearly explained in [15]. Therefore, any number array antenna such as 16×16, 64×64 and so on can be realized using the proposed array antenna as a unit. In this case, the gain of the array antenna will be increased according to the number of antenna elements. Feed network also can be realized by meaner line. However, in this case the meander line impedance might very high, which will need additional impedance matching circuits.

Two PIN diodes are loaded on two corners of each antenna element in order to realize polarization switching. The other ends of the diodes are connected to the ground plane by via hole. The characteristic impedance of the microstrip lines connected to the antenna elements is 110Ω and the width of the microstrip line is 0.55mm. The input impedance of the antenna element is adjusted to the microstrip lines by properly inserting a pair of notches at the patch [16]-[17]. The width of the notch is 0.2 mm and length is 0.4 mm. The 110Ω microstrip lines connect with another microstrip line which is used to realize the magic-T. The impedance of this microstrip line is 55Ω and the width is 2.4 mm. The width of the switching signal line is maintained at 0.2 mm in order to achieve a high impedance of 154Ω. A slot line of 0.2 mm width is arranged exactly below the switching signal line on the ground plane of the array antenna. The switching signal line is connected with the input port through air-bridge. A 110Ω RF microstrip line is arranged on the obverse plane upon the slot line as a microstrip-slot parallel branch circuit. This RF microstrip line is connected with the 50Ω RF port using a quarter wavelength ($\lambda_g/4$= 5.55 mm) impedance transformer whose impedance is 71Ω and width is 1.36 mm. The size of the Teflon glass fiber substrate and the ground plane is 48 × 54 $mm^2$. The thickness of the substrate is 0.8 mm with the relative dielectric constant of 2.15.

### *A. The Magic-T*

Planar magic-Ts are used in microwave integrated circuits to split or combine in-phase or anti-phase signals. The magic-T is very useful for balanced-mixers, discriminators, and beam-forming networks. The advantages of the magic-T are low insertion loss, high isolation, compact size and fabrication simplicity [18]. There are some reports for realizing the magic-T using the co-planar waveguide or microstrip to slot line mode conversion techniques [19]-[22].

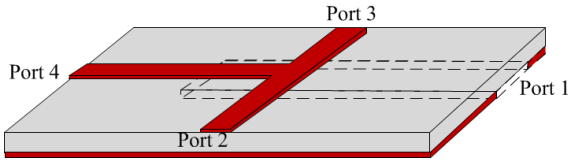



Figure 2. Magic-T structure.

Fig. 2(a) shows the structure of the magic-T. The RF signal can be propagated through the slot line according to the slot line mode which can be referred as the orthogonal feed. When the RF signal is input to port 1, the RF signal is split to the microstrip line in anti-phase and propagates to port 2 and port 3 which can be referred as the odd mode or slot line mode. The switching signal from port 4 can propagate to port 2 and port 3 by the means of even mode. As the even mode and odd mode are orthogonal to each other, a good isolation between the RF signal and the bias signal can be realized. The S11 for the structure is matched at the design frequency of 10 GHz. S41, which is the isolation, also achieved better than -28 dB. In addition, the S21 and S31 are achieved around -4 dB for the simulation.

### *B. The unequal feed line for 90° phase difference*

The schematic structure of the unequal feed line is shown in Fig. 3. The main two conditions for the realization of the circular polarization are

1. Two orthogonal signals
2. 90° phase difference between the orthogonal signals.

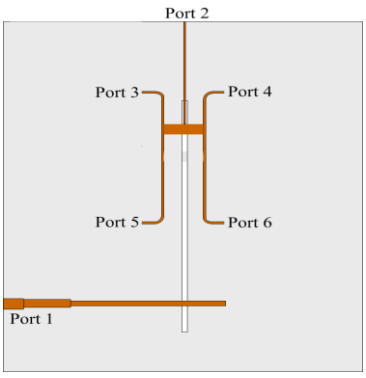



Figure 3. The unequal feed circuit.

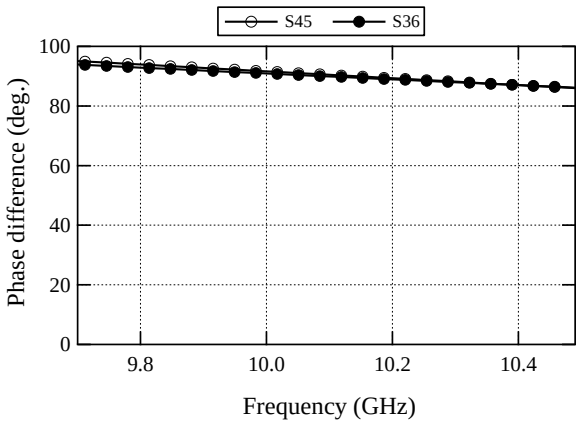



Figure 4. Phase differnce of the feed circuit.

The length of the feed lines is adjusted such a way that there is a 90° phase difference between port 4 and port 5, and port 3 and port 6 is achieved at the design frequency of 10 GHz which is shown in Fig. 4.

### C. *Basic Behavior of the array antenna*

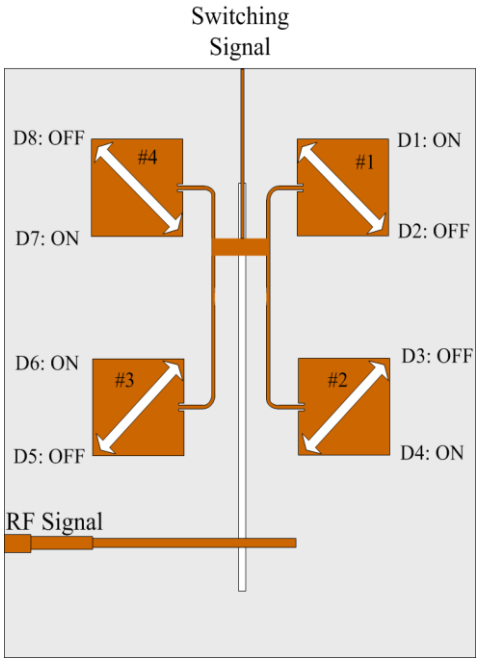



Figure 5. Basic behaviour of the array antenna.

The basic behavior of the circular polarization array antenna will be explained using Fig. 5. When positive bias voltage is applied, the diodes D1, D4, D6 and D7 will be ON and the Patch #1 and Patch #4 will be -45° linear polarized and patch #2 and patch #3 will be +45° deg linear polarized. However, there is a 90° phase difference between the signals of the patch #1,#4 and the patch #2,#3 which fulfills the above two conditions. In this condition, the array antenna will be right hand circularly polarized. When negative bias voltage is applied, the diodes D2, D3, D5 and D8 will be ON and the Patch #1 and Patch #4 will be +45° linear polarized and patch #2 and patch #3 will be -45° linear polarized. In this condition, the array antenna will be left hand circularly polarized. This antenna will also exhibit circular polarization it the phases are changed sequentially instead of diode ON/OFF. However, phase shifter will be needed in this case.

### D. *Performance of the single antenna element*

Performance of the single antenna element is investigated by the ADS simulation software. The ON state diode is replaced by short and the OFF state diode is replaced by open. Fig. 6 shows the S11 characteristics for the single antenna where the antenna impedance is matched for the design frequency of 10 GHz. However, the S11 performance of the feed circuit of the full array antenna is also investigated by the simulation. In this case also, the input impedance of the array antenna is matched for the design frequency of 10 GHz.

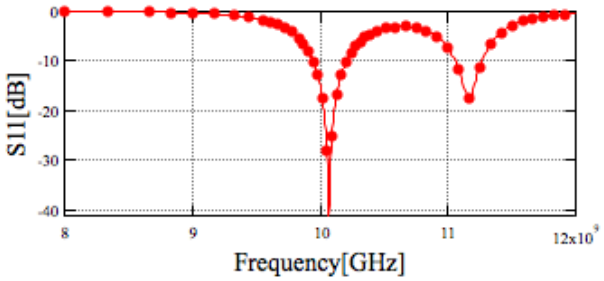



Figure 6. S11 characteristics of the single antenna.

## III. CONCLUSION

In this paper, a 4-element circular polarization switchable patch array antenna is proposed. The array antenna is realized by the successful employment of the both-sided MIC technology, where the odd and even mode of the slot line and microstrip lines are used for the realization of the orthogonal feed and magic-T switching circuit respectively. In order to realize the circular polarization switching, two switching diodes are integrated with each patch elements. Unequal feed line is introduced to realize 90 deg, phase difference between the patch elements and this way, circular polarization switchable array antenna can be realized. The circular polarization array antenna will be fabricated and will be investigated by the experiment.

## REFERENCES

[1] F. Yang and Y. Rahmat-sami, "Patch Antennas with Switchable Slots(PASS) in wireless communications: concepts, Design and

Application," *IEEE Trans. on Antennas and propagation*, Vol. 47, No. 2, pp. 13-29, 2005.

[2] M.K. Fries, R.Vahldieck, "Uniplaner circularly polarized slot-ring antenna architectures", Radio Sci., vol. 38, no. 2, 2002 November.

[3] J. Y. Park, C. Caloz, Y. Qian and T. Itoh, "A Compact Circularly Polarized Subdivided Microstrip Patch Antenna," *IEEE Microwave and Wireless Components Letters*, Vol. 12, No. 1, pp. 18-19, 2002.

[4] X. M. Qing, Y. W. M. Chia, "Circularly polarized circular ring-slot antenna fed by stripline hybrid coupler", *Electronic Letter*, vol, 35, no. 25, pp. 2154- 2155, 1999 December.

[5] J. S. Rpw, C. Y. D. Sim, K. W. Lin, "Broadband printed ring-slot array with circular polarization", Elect. Lett., vol. 41, no. 3,

[6] M. A. Hossain, Y. Ushijima, E. Nishiyama, I. Toyoda, and M. Aikawa, "Orthogonal circular polarization detection patch array antenna using double-balanced RF multiplier," *Progress In Electromagnetics Research C*, Vol. 30, 65-80, 2012.

[7] M. Boit, L. Dusspot, and J.-M. Laheurte, "Circularly olarized antenna with switchable polarization sense", *Electron. Lett*.,vol. 36, no. 18, pp. 1518- 1519, 2005 August.

[8] J. S. Rpw, C. Y. D. Sim, K. W. Lin, "Broadband printed ring-slot array with circular polarization", *Elect. Lett.*, vol. 41, no. 3, pp. 110-112, 2005 February.

[9] X. M. Qing, Y. W. M. Chia, "Circularly polarized circular ring slot antenna fed by stripline hybrid coupler", *Elect. Lett*., vol, 35, no. 25, pp. 2154- 2155, 1999 December.

[10] Yu Ushijima, E. Nishiyama and M. Aikawa, "Circular Polarization Switchable Microstrip Antenna with SPDT Switch Circuit", *IEEE Antennas and Propagation Society International Symposium (APSURSI)*, 2010.

[11] Yu Ushijima, E. Nishiyama, I. Toyoda and M. Aikawa, "Circular Polarization Switchable Single Layer Microstrip Antenna", *IEEE Antennas and Propagation Society International Symposium (APSURSI)*, 2012.

[12] T. Onishi, M. A. Hossain, E. Nishiyama and I. Toyoda, "Linear Polarization Switchable Microstrip Array Antenna using Magic-T Circuit ", *2012 International Symposium on Antennas and Propagation (ISAP)*,2012 .

[13] M. Aikawa and H. Ogawa, "Double-sided MIC's and their applications," *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-37, No. 2, pp. 406-413, 1998.

[14] K. Kodama and E. Nishiyama and M. Aikawa, "Slot array antenna using Both-sided MIC technology," *IEEE International symposium on Antennas and Propagation* , Vol. 3, pp. 2715-2718,

[15] K. Egashira and E. Nishiyama and M. Aikawa, "Microstrip array antenna using Both-sided MIC feed circuits," *Asia-pacific Microwave conference*, APMC, 2002.

[16] R. E. Munson, "Conformed microstrip antennas microstrip patched arrays," *IEEE Trans. Antennas Propagat*., Vol.AP-22, pp. 7478, 1974.

[17] P. S. Hall and C. M. Hall, "Coplanar corporate feed effects in microstrip patch array design," *Proc. Inst. Elect. Eng., pt. H*, Vol. 135, pp. 180186, 1988.

[18] K. U-yen, E. J. Wollack, S. H. Moseley, J. Papapolymerou and J. Laskar, J, "A Compact Low-loss Magic-T using Microstrip-Slotline Transitions," *IEEE/MTT-S International Microwave Symposium*, pp. 37-40, 2007.

[19] C. H. Ho, L. Fan and K. Chang, "New uniplanar Coplanar Waveguide Hybrid-Ring Couplers and Magic-T's," *IEEE Trans. Microwave Theory and Technique*, Vol. 42, No. 12, pp. 2440-2448, 1994.

[20] J. P. Kim and W. S. Park, "Novel Configurations of Planar Multilayer Magic-T using Microstrip-Slotline Transitions," *IEEE Trans. Microwave Theory and Technique*, Vol. 50, No. 7, pp. 1683-1688, 2002.

[21] M. Aikawa and H. Ogawa, "A New MIC Magic-T Using Coupled Slot Lines," *IEEE Trans. Microwave Theory and Technique*, Vol. 28, No. 6, pp. 523-528, 1980.

[22] T. Hiraoka, T. Tokumitsu and M. Aikawa, "Very small wide-band MMIC magic T's using microstrip lines on a thin dielectric film," *IEEE Trans. Microwave Theory and Technique*,Vol. 37, No. 10, pp. 1569-1575, 1989.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Sinusoidal Appearance of Nonuniform Dumbbell Shape EBGSs in Microstrip Transmission Line

S. M. Shakil Hassan[1,*], S. M. Anayetullah[1], and Mohammad Nurunnabi Mollah[2]
[1]Dept. of Electrical and Electronic Engineering, Eastern University, Bangladesh
[2]Khulna University of Engineering and Technology, Bangladesh
[*] E-mail: shakil.bq@gmail.com

*Abstract*—**A novel technique of designing nonuniform dumbbell shape electromagnetic bandgap structure (EBGS) assisted microwave transmission line has been proposed that shows distinguished performance. The proposed design has sinusoidal appearance of binomially distributed nonuniform EBG structures. Tapered dumbbell shape EBGS assisted designs using the proposed method have been investigated and the result is compared with the performance of corresponding uniform design. Distinguished low-pass filtering (LPF) performance has been observed from the proposed design that not only shows sharp attenuation; but also significantly low ripples. The basis of the performance has been demonstrated and verified. The proposed technique, moreover, has significantly overcome the limitation of getting increased number of effective tapered EBGSs from the conventional method (i.e. binomial distribution) of tapering.**

***Keywords—Electromagnetic bandgap structure, EBGS, defected ground structure, low-pass filter, sinusoidal appearance of nonuniform EBGS.***

## I. INTRODUCTION

Periodic electromagnetic bandgap structures (EBGS), basically, introduce a periodic discontinuity at the beneath of transmission line (T-line) where electromagnetic waves are impeded, hence form slow wave; therefore, effective wavelength gets increased i.e. compact performance achieved [1]. In photonic engineering such kind of patterns are termed as photonic bandgap (PBG) elements. Electromagnetic bandgap structures (EBGSs), however, are periodic structures that perform wide band-pass and band rejection property at microwave and millimeter wave frequencies; since electromagnetic waves behave in photonic substrates as electrons behave in semiconductors. Due to their unique behavior in microwave engineering they have found potential applications in designing various types of microwave filters, antenna, waveguide, amplifiers, and resonators [2]-[8]. Various planer etched EBG patterns have been introduces in literatures; some of them are in general geometric shapes such as circular and square and some of them are modified and complicated like dumbbell shape [9], [10].

Dumbbell shape EBGS is a kind of electromagnetic bandgap (EBG) pattern that has found tremendous interest of researchers in the field of electromagnetic and photonic engineering due to their distinguished compact performance. This particular type of design has lots of factors that impact individually to the performance of the microstrip T-line. Depending on the sizes and shapes it sometimes performs as a stopband or as a low-pass filters, even sometimes as dual-band or multi-band filters [10]. Dumbbell shape EBG structure is also termed as defected ground structure (DGS) [7]. In contradiction, any kind of pattern etched in the ground plane at the beneath of microstrip T-line, principally, makes the plane defected and disturbs the propagation of electromagnetic waves. The advantage of dumbbell pattern is that only few EBG elements show the property of conventional periodic structure, hence the size becomes relatively smaller; but one shortcoming is that there is no proper formulation of designing.

A design with tapered dumbbell shape EBG assisted microstrip transmission line has been proposed in this paper. Such kinds of tapered design are sometimes called nonuniform EBG assisted design and has been reported in too many literatures [11]-[14]. In this paper the tapering has been done using binomial distribution; chebyshev distribution can also be applied to make EBGSs nonuniform [13]. We have proposed a new novel design with sinusoidal appearance of tapered EBG elements that has greatly modified the performance of conventional binomially distributed design. Repeating the tapered EBGS lattice we can obtain sinusoidal appearance of EBG elements. Proper demonstration has been provided regarding the performance on the basis of potential effect of number of electromagnetic bandgap structures at the beneath of microstrip transmission line.

## II. NONUNIFORM DUMBBELL SHAPE EBGS

### A. Binomial Distribution

Binomial distribution is a popular distribution that is being studied in probability and statistics theory to determine the probability distribution for the discrete number of successes in independent sequence of experiments. The coefficients of the polynomial are determined by the expression (1). By using positive coefficients of the binomial series expansion for different values of $n$ a triangular array can be formed that is called Pascal's triangle [15] shown in Fig. 1.

$$(1+x)^{n-1} = 1 + (n-1)x + \frac{(n-1)(n-2)}{2!}x^2 + \frac{(n-1)(n-2)(n-3)}{3!}x^3 + \cdots \quad (1)$$

n=1 1
n=2 1 1
n=3 1 2 1
n=4 1 3 3 1
n=5 1 4 6 4 1
n=6 1 5 10 10 5 1
n=7 1 6 15 20 15 6 1
n=8 1 7 21 35 35 21 7 1
n=9 1 8 28 56 70 56 28 8 1

Figure 1. Binomial coefficients in Pascal's triangle.

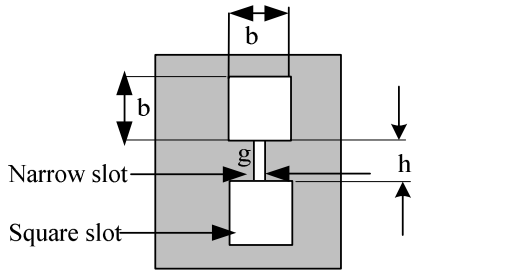
Binomial coefficients have potential use in EBG engineering in desingig nonuniform EBG elements. From the Fig. 1, if $n$ represents the number of EBG elements at the ground plan then the ampletudes of the different EBG elements are calculated from the coeffiiceants. In this work the relative ampletude of the dumbbell shape EBG elements are varied proportionally using the coefficents of the Pascal's triangle. We have taken five EBG elements in our desigs, hence the proportional ampletude coefficents ($P$) are as fllows:

$P_3$=0.17 $P_2$=0.67 $P_1$=1 $P_2$=0.67 $P_3$= 0.17

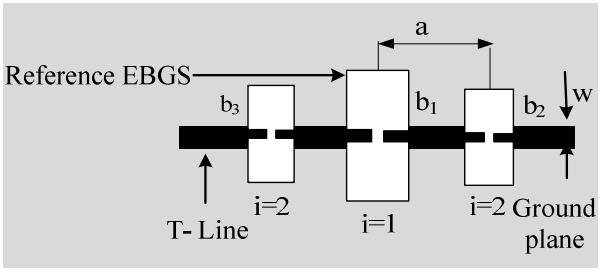
These values are calculted by deviding each of the binomial coefficents (i.e. 1, 4, 6, 4, 1) with the largest value (i.e. 6) that are shown in the Pascal's triangle for $n$ = 5.

### *B. Tapered Dumbbell Shape EBGS*

Fig. 2(a) shows the geometry of dumbbell shape EBGS. From the figure it is seen that two bigger square slots are etched at the both ends of a narrow slots; it has been reported that the etched square slots and narrow slot are equivalent to the added effective inductance ($L$) and capacitance ($C$) respectively with characteristic impedance $Z_0$ (= 50 Ω) that yield the parallel $L$-$C$ equivalent circuit [7]. Fig. 2(b) shows the design of periodic tapered dumbbell shape EBG assisted microstrip transmission line. Arm length of the square slots ($b_i$) of nonuniform dumbbell shape EBG elements, however, are calculated by following expression:




(a)




(b)

Figure 2. (a) Unit dumbbell shape EBGS, and (b) tapered dumbbell shape EBG elements at the beneath of T-line.

$$b_i = P_i \times b \quad (2)$$

Where

$b$ = arm of square slot of the biggest EBG element or $b_1$

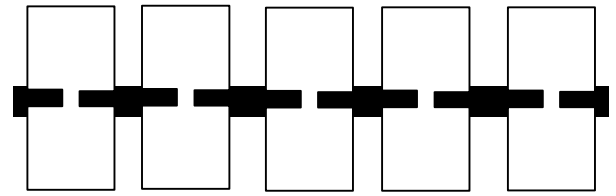
$i$ = position of EBG elemnts

$P_i$ = proportional amplitude coefficient for the $i^{th}$ EBG

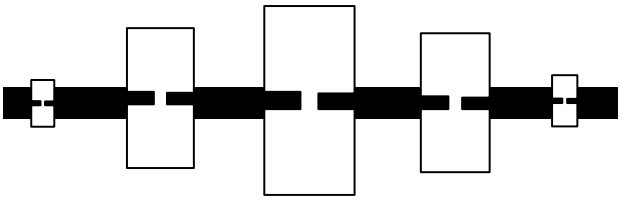
$b_i$ = arm of square slot of the $i^{th}$ EBG

## III. DUMBBELL SHAPE EBGS ASSISTED MICROSTRIP LINE

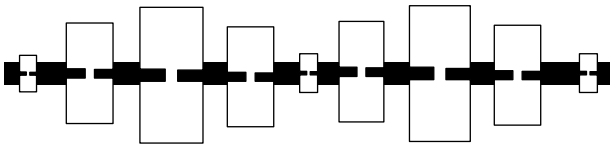
In this section we have proposed a dumbbell shape EBGS assisted microstip transmission line. Its performance has been investigated into different forms – (a) as uniform patterns, (b) as binomially distributed nonuniform EBGS, and (c) nonuniform EBGS with sinusoidal appearance. In [16] and in [17] sinusoidal variation of characteristic impedance of microstip T-line and continuous sinusoidal patters in the



(a)



(b)



(c)

Figure 3. (a) Uniform dumbbell shape EBGS, (b) tapered dumbbell shape EBGS, and (c) sinosoidal appearance of nonuniform dumbbell shape EBGS.

ground plane of a microstrip T-line are reported respectively. In this work, however, we have proposed nonuniform EBGS that has variation in the size of EBGS in a way that looks like a sinusoid and it results distinguished low-pass performance.

For designing we have chosen the substrat that has dielectric constant ($\varepsilon_0$) and thickness are 2.45 and 31 mil (0.7874 mm) respectively. The width of microstrip transmission line is 2.2642 mm to keep characteristic impedance ($Z_0$) 50 Ω. The designs are as follows:

### *A. Uniform Dumbbell Shape EBGS*

In this design five dumbbell shape EBGSs are placed with a periodicity ($a$) of 10.43 mm shown in Fig. 3(a). The arm ($b$) of the larger square slot is 7 mm and the height ($h$) and width ($g$) of the narrow slot are 0.5 mm and 0.5 mm respectively.

### *B. Tapered Dumbbell Shape EBGS*

With the same periodicity and size of the narrow slot we have designed a microstrip transmission line having nonuniform EBGSs, where the tapering has been done by (2). Thus the sizes of the bigger slots become 7 mm, 4.69 mm, and 1.19 mm from bigger to smaller EBGS shown in Fig. 3(b).

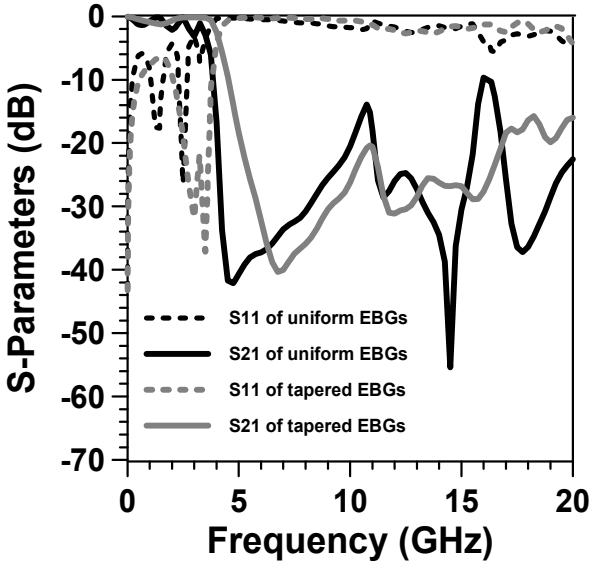
### *C. Nonuniform Dumbbell Shape EBGS with Sinusoidal Appearance*

Fig. 3 (c) shows a novel design that has tapered dumbbell shape EBGSs that are used in the design shown in Fig. 3 (b). We have repeated the lattice of the tapered EBGSs to make two cycles that can be termed as sinusoidal appearance of tapered EBGSs; therefore, it increases the number of EBGSs and makes significant improvement in performance.

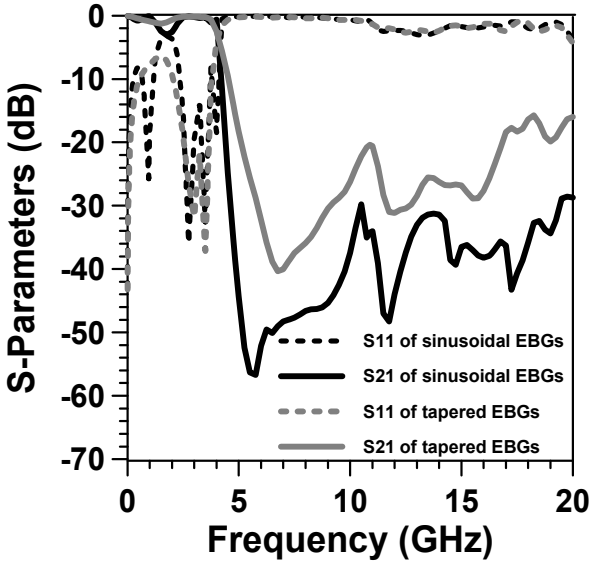
## IV. PERFORMANCES OF THE DESIGNS

In this section scattering parameter (S-parameter) performances of the designs have been depicted in terms of insertion loss (IL) and return loss (RL) properties. Insertion loss and return loss curves are, basically, $S_{21}$ and $S_{11}$ of the scattering parameters. We observed stopband at -20 dB of $S_{21}$ curve and passband at -10 dB of $S_{11}$ curve; also -3 dB cutoff frequency and maximum ripple of $S_{21}$ in passband to justify the performances. However, the simulations have been done with a renowned method of momentum (MOM) based software named Zeland ie3d that has found much use in this field in literature [13]-[14].

Fig. 4 shows s-parameter performances of uniform, nonuniform and sinusoidally appeared nonuniform dumbbell shape EBGSs assisted designs. The results have been accumulated in the Table 1. From the Fig. 4(a), however, it is seen that tapered EBGSs have reduced the ripples and have increased stopband width. Proposed sinusoidal appearance of nouniform EBGSs have increased the insertion loss hence resulted low-pass performance; but the performance pattern is likely to the performance of the binomially tapered design with the exception of deeper $S_{21}$ curve. This is because of having almost double EBG elements than the simple binomially tapered design. Basically increased numbers of EBG elements increases the insertion loss, since it introduces more impedance in the microstrip transmission line [18]. The proposed technique, therefore, offers more effective tapered EBGS at the beneath of T-line unlike conventional chebyshev distribution [13], binomial distribution [14], and some other tapering techniques [19]-[20]. It has overcome the limitation of much smaller EBGSs at the ends that have no potential impact on the performance of the microstrip line while using conventional methods of tapering.




(a)




(b)

Figure 4. (a) S-parameter performance of uniform EBGSs and binomially tapered EBGSs assisted T-lines, and (b) S-parameter performance of binomially tapered EBGSs and nonuniform EBGSs with sinusoidal appearance assisted T-lines.

TABLE I. S-PARAMETER PERFORMANCES OF ALL DESIGNS

| Design Title | Different Values of S-parameter | | | |
|---|---|---|---|---|
| | *Passband at -10 dB (GHz)* | *-3 dB cutoff frequency (GHz)* | *Stopband at -20 dB* | *Max.RH (dB)* |
| Uniform EBGSs | 2.69 | 3.505 | 6.01 | -3.11 |
| Nonuniform EBGSs | 3.8 | 4.04 | 11.73 | -1.2 |
| Sinusoidal Appearance | 4.12 | 4.054 | LPF started at 4.41 | -3.05 |

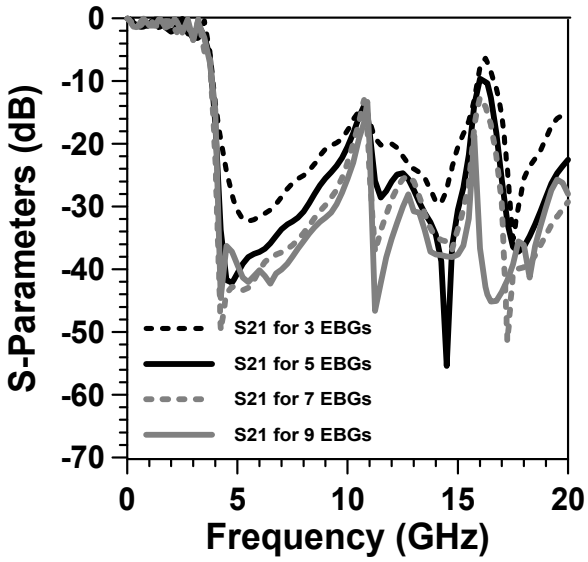
## V. REASON & VALIDATION OF THE RESULTS: EFFECT OF VARIOUS NUMBER OF EBGS

Using the size of proposed dumbbell shape EBGS that has arm of bigger and narrower square slots are 7 mm and 0.5 mm respectively, we have investigated the effect of different number of uniform EBGSs at the beneath of microstrip transmission line. The results of 3, 5, 7, and 9 EBGSs assisted microstrip T-line has been shown in the Fig. 5. From the figure it is seen that $S_{11}$ curves of all the above mentioned designs are almost similar; nearly overlapped with each other. $S_{21}$ curves, on the other hand, have little dissimilarity that a design that has more number of EBGSs has more insertion loss i.e. deeper $S_{21}$ curve in terms of negative dB. However, for all the cases passbands are observed of 2.69 GHz, -3 dB cutoff frequencies are observed at 3.5 GHz, and stopband widths are around 6 GHz. Maximum ripple heights, in fact, get increased as number of EBGSs increased. The performance, actually, have satisfied the concept of [18]. It is mentionable that proposed sinusoidally appeared EBGSs assisted T-line has performed worth compared to the design having nine uniform EBGSs; since proposed design has shown LPF performance in Fig. 4(b), whereas uniform EBGSs have shown just a stopband performance in Fig. 5.

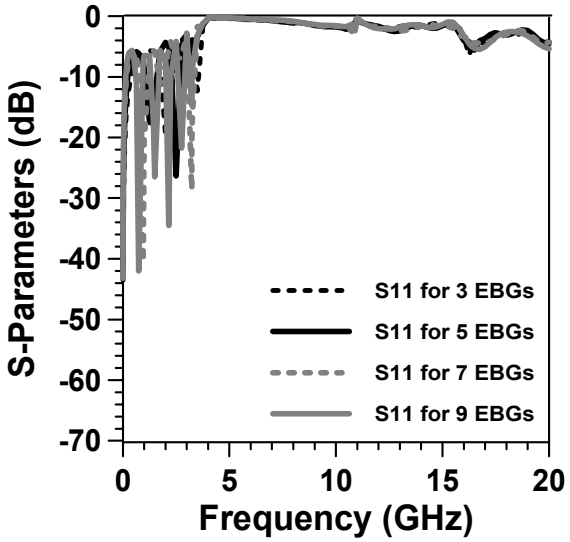
Results that have been reported in this work are observed by simulating the designs using Zeland ie3d. Experts have found simulation results using this software are in well agreement with the measured performances [5], [9], [14]. In [10] and [20] the congruency of measured performance and simulation results has been illustrated. However, earlier we have verified the concept of effect of number of EBGSs that has been reported in [18]; therefore, we may consider that our simulation results will find good agreement with the measured results as well.

## VI. CONCLUSSION

Sinusoidal appearance of the nonuniform EBGSs can be achieved by placing binomially tapered EBG elements consequently to make the second cycle. Normal tapered design realized by binomial distribution, therefore, can also be termed as sinusoidally appeared EBGSs assisted design having one cycle only. Proposed sinusoidally appeared nonuniform design of two cycles has importance over simple tapered nonuniform designs i.e nonuniform EBGSs of single cycle. However, this is actually a potential technique to modify the performance of binomially tapered nonuniform designs. In this paper we have




(a)




(b)

Figure 5. (a) $S_{21}$ curves of different number of dumbbell shape EBGSs assisted T-lines, and (b) $S_{11}$ curves of different number of uniform dumbbell shape EBGSs assisted T-lines.

reported a very eminent low-pass performance; whereas the uniform design of its base EBG elements is very worst – unwanted peaks of insertion loss ($S_{21}$) curves are observed around at 10 GHz and 17 GHz shown in Fig. 5. These peaks, indeed, have degraded the performance and have prevented to form an attractive low-pass performance. Binomially tapered design, in contrast, has fined the performance; but finer result has been observed from the proposed sinusoidally appeared nonuniform designs of two cycles. In essence, using tapered design though reduces the ripples and fines the results; but the the sinusoidally appeared nonuniform designs offers better option if it is needed to increase the number EBG elements and to abate annoying peaks from $S_{21}$ performance.

## REFERENCES


[1] F.-R. Yang, K.-P. Ma, Y. Qian and T. Itoh, "A uniplanar compact photonic-bandgape (UC-PBG) structure and its applications for microwave circuits," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, no. 8, pp. 1509-1514, Aug. 1999.

[2] R. Gonzalo, P. D. Maagt, and M. Sorolla, "Enhanced patch-antenna performance by suppressing surface waves using photonic-bandgap substrates," IEEE Trans. Microw. Theory Techn., vol. 47, no. 11, pp.2131–2138, Nov. 1999.

[3] V. Radisic, Y. Qian, R. Coccioli, and T. Itoh, "Novel 2-D photonic bandgap structures for microstrip lines," IEEE Microwave and guided wave lett., vol. 8 no. 2, pp. 69-71, Feb. 1998.

[4] M. N. Mollah, and N. C. Karmakar, "Pure harmonic suppression of a bandpass filter using binomially distributed photonic bandgap structures," Microw. Opt. Technol. Lett., vol.44, no. 2, pp.194–196, Jan. 2005.

[5] N. C. Karmaker, and M. N. Mollah, "Potential applications of PBG engineered structures in microwave engineering - part one," Microwave Journal, vol. 47, no. 7, pp. 22-44, July 2004.

[6] V. Radisic, Y. Qian, and T. Itoh, "Broad-band power amplifier using dielectric photonic bandgap structure," IEEE Microw. Guided Wave Lett., vol. 8, no. 1, pp. 13–14, Jan. 1998.

[7] J. S. Lim, C. S. Kim, D. Ahn, Y. C. Jeong, and S. Nam, "Design of lowpass filters using defected ground structure," IEEE Trans. On Microwave Theory and Techniques, vol. 53, no. 8, pp. 2539-2545, Aug. 2005.

[8] T. Y. Yun, and K. Chang, "Uniplanar one-dimensional photonic-bandgap structures and resonators," IEEE Trans. On Microwave Theory and Techniques, vol. 49, no. 3, pp. 549-553, Mar. 2001.

[9] M. N. Mollah, and N. C. Karmakar, "Compact hybrid defected ground plane," Microw. Opt. Technol. Lett., vol. 44, no. 3, pp.266–270, Feb. 2005.

[10] S. M. S. Hassan, M. N. Mollah, M. A. Rashid, N. H. Ramly, and M. Othman, "Dumbbell shape EBGS structure – worth to EBG assisted microwave filter designing," *2012 IEEE Asia-Pacific Conf. on Applied Electromagnetics*, Dec. 2012, Melaka, Malaysia, pp. 1-5.

[11] N. C. Karmakar, "Theoretical investigations into binomial distributions of photonic bandgaps in microstripline structures," Microw. Opt. Technol. Lett., vol.33, no. 3, pp.191–196, Jan. 2002.

[12] S. Y. Huang, and Y. H. Lee, "Tapered dual-plane compact electromagnetic bandgap microstrip filter structures," IEEE Trans. Microw. Theory Techn., vol. 53, no. 9, pp.2656–2664, Sep. 2005.

[13] N. C. Karmakar, "Improved performance of photonic band-gap microstripline structures with the use of chebyshev distribution," Microw. Opt. Technol. Lett., vol. 33, no. 1, pp.1–5, Apr. 2002.

[14] N. C. Karmakar, and M. N. Mollah, "Investigations into nonuniform photonic-bandgap microstripline low-pass filters," IEEE Trans. Microwave Theory Tech., Vol. 51, no. 2, pp. 564-572, Feb. 2003.

[15] C. A. Balanis, Antenna Theory Analysis and Design, 2 ed. New York: Wiley, 1997.

[16] T. Lopetegi, M. A. G. Laso, M. J. Erro, D. Benito, M. J. Grade, F. Falcone, and M. Sorolla, "Novel photonic bandgap microstrip structures using network topology," Microw. Opt. Technol. Lett., vol. 25, no. 1, pp.33–36, Apr. 2000.

[17] D Nesic, and A. Nesic, "Bandstop microstrip PBG filter with sinusoidal variation of the characteristic impedance and without etching in the ground plane," Microw. Opt. Technol. Lett., vol. 25, no. 1, pp.33–36, Apr. 2000.

[18] M. N. Mollah, and N. C. Karmakar, "Effect of number of uniform photonic bandgap (U-PBG) elements on the performance of microstrip transmission lines," PACRIM-2003, Victoria, Canada.

[19] M. A. G. Laso, T. Lopetegi, M. J. Erro, D. Benito, M. J. Grade, and M. Sorolla, "Novel wideband photonic bandgap microstrip structures," Microw. Opt. Technol. Lett., vol. 24, no. 5, pp.357–360, Mar. 2000.

[20] S. M. S. Hassan, S. M. Anayetullah, M. N. Mollah, K. R. Hossain, M. M. Ahmed, M. A. A. Shufian, and M. M. Hassan, "Realization of low-pass filters from arbitrarily designed nonuniform EBG structures," in *Proc. Int. Conf. on Electrical Information and Communication Technology*, Dec. 2013, Khulna, Bangladesh.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Re-Evaluating Chain-Code as Features for Bangla Script

Minhaj N. Alam* and M. A. Naser
Islamic University of Technology, Bangladesh
*E-mail: pulak0017@gmail.com

*Abstract*— **Most of the characters in Bangla script have similar shapes. Since chain-code is one kind of shape descriptor, we argue that it would not be able to distinguish between two similar characters and hence not be able to provide good recognition rate. Even though, the chain-code has widely been used as feature for Bangla script, none of the literature has considered the fact that chain-code may not be compatible with the script itself. We assume that chain-code cannot provide the variation necessary to describe the contours of similar characters, especially which are almost identical in the first place. We validated our proposal through a statistical test called one-way Analysis of Variance (ANOVA), to verify whether chain-codes of two similar looking characters are truly different. The results substantiate our assumption, suggesting that chain-code based features may not be the best features for Bangla character recognition.**

***Keywords— Chain-code, Optical Character Recognition (OCR), Bangla script, Analysis of Variance (ANOVA), F-distribution.***

## I. INTRODUCTION

Feature extraction plays an important role in any Optical Character Recognition (OCR) system. Selection of a proper extraction method is probably the single most important factor in achieving high recognition performance [1]. Some commonly used feature extraction methods for English alphabet are described in [1] while a comprehensive study of OCR work on Bangla scripts can be found in [2]. One such method employed for Bangla script is chain-code and we shall confine our discussion to this method.

In image processing, chain-code is used as a descriptor for recognition purposes. Nixon and Aguado asserted that to be useful for recognition, descriptors should have four important properties such as completeness, congruence, invariance and compactness [3]. Completeness is achieved when no two objects have the same descriptors if they have different shapes. Congruence property assures that similar objects can be recognized when they have similar descriptors. Invariance enables descriptors to recognize objects irrespective of their orientation, position or size. In addition to these three properties, the descriptors should be a compact set that retains the essence of an object i.e. the uniqueness or difference from the other objects. Even though chain-code has these properties and can provide good recognition performance, one has to understand that chain-code cannot characterize general objects. In other words, the description by chain-code is strongly related to each particular application with a particular type of object. According to Devijver and Kittler [4], features are "the information which is most relevant for classification purposes, in the sense of minimizing the with-in class pattern variability while enhancing the between–class pattern variability."

Chain-code is basically used to represent a contour where the direction of border pixels is translated or coded into numbers [5], [6]. The successive direction from one pixel to the next pixel becomes an element in the code and all these elements are concatenated to form a chain. Hence, the name chain-code. This process is repeated for each point until the starting point is reached. The contour of a character, along with an 8 directional chain-code is shown in Fig. 1. Chain-code was originally used for English character recognition [1] and provided good accuracy. However, unlike English characters, Bangla characters have a lot of inter-class similarity. Four pairs of characters are almost identical, except a dot (under the character) that differentiates between characters in a pair. These characters are grouped in Table 1.

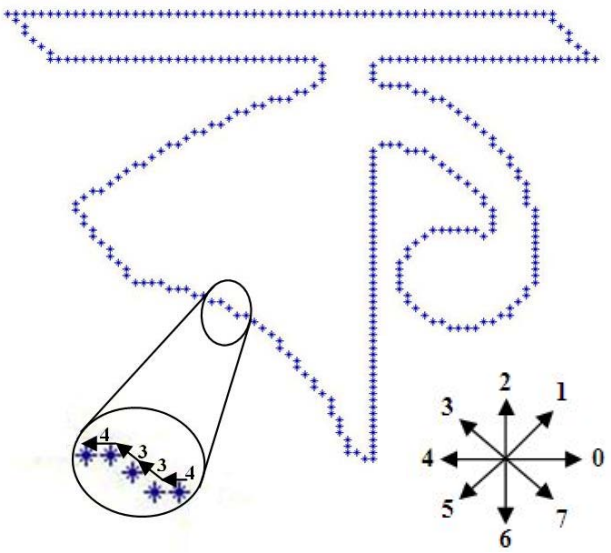



Figure 1. Contour of a character with 8 directional chain-codes.

Most of the rest have partial similarity among them and we have grouped all such characters in Table 2. Considering Nixon's statement, we argue that as a descriptor chain-code

fails to provide completeness due to such similarity in shapes. Thus two different characters with similar shapes have similar chain-code that degrades between-class variability mentioned by Devijver. In other words, the congruence property makes it difficult to classify characters from similar chain-codes. We, therefore, propose to re-evaluate chain-code as features for Bangla characters.

Table 1. Identical pairs of characters (character number)

| |
|---|
| ড (23) & ড় (44) |
| ঢ (24) & ঢ় (45) |
| য (36) & য় (46) |
| ব (33) & র (37) |

Table 2. Characters with similar shapes (character number)

| | | | |
|---|---|---|---|
| অ(1) আ (39) | উ(4) ঊ(5) | ই(2) ঈ(3) | এ(7) ঐ(8) |
| ও(9) ঔ(10) | ঋ(6) ঝ(19) | ঘ(14) ম(35) | ক(11) ফ(32) |
| ঘ(14) ম(36) ষ(41) য়(46) | | | |

In this paper, we shall present a statistical way to validate our hypothesis. The method is called Analysis of Variance (ANOVA) and it is described in section 4. Section 2 presents some related works, section 3 describes chain-code histogram features, and results are presented in section 5. Finally section 6 concludes the work.

## II. RELATED WORKS

Many researchers [7, 8, 9, 10, 11] employed chain code as feature extraction method in Bangla OCR. However, we will limit our discussion only to the works where characters are grouped based on their similarity. Bhattacharya et al. [7] employed local chain-code histograms as features in Multilayer Perceptrons (MLP) classifier. They reported 94.65% recognition accuracy (best situation). However, this result comes after multiple stages of classification in which the initial classification does not give satisfactory recognition. The approach is contingent to the fact that a character must be classified as one of the similar shaped characters in the first stage. If it is classified as one of the characters from another group, it does not have any chance to be recognized correctly. Moreover, multistage classification requires higher computational time.

Sikdar et al. proposes a step-by-step grouping and recognition based on pixel density and chain-code features [8]. In the first step of this hierarchical approach, horizontal and vertical lines are identified using pixel density and based on that the characters are grouped accordingly. Then they search for a distinguishing feature (i.e. dot under the character) by matching the chain-code of the whole character against a stored code of that particular feature. The problem with such approach is that for certain characters it may employ too many stages of recognition, which consequently increases the computational time. The most important drawback of such approach is that it is not rotation invariant. If character is rotated 90˚, horizontal and vertical lines interchange their position leading to erroneous classification. Nevertheless, it is very laborious to carry on an exhaustive study of each and every feature and write algorithms for that.

Reza and Khan used chain-code for feature extraction in [9] and reported almost 90% accuracy. But the main goal of this work was to group similar characters. The classifier was trained only to classify characters among pre-defined groups and determine inter-group and between-group misclassification rate.

In [10], not only chain-code but a combination of three other feature extraction methods was used (intersection, shadow, straight line fitting and chain-code histogram) for handwritten Devnagari datasets. Accuracy was around 92.8%. However, it would decrease if chain-code feature was used only instead of the combination.

Pal et al. presented a lexicon driven segmentation-recognition scheme where chain-code based features were used [11]. Primitive components were obtained through pre-segmentation and used to match against lexicon. Dynamic programming was employed to maximize average character likelihood. This method provided 87.21% accuracy.

From the above discussion it is clear that most of the chain-code based approaches use multiple techniques to enhance the recognition. Major techniques include multistage classification [7], grouping of similar characters [8], [9] and combination of other feature extraction methods [10]. Although the reported accuracy is high, with increasing number of test characters, we assume it would degrade. For instance, the highest accuracy rate of 94.65% would give 5350 misclassified characters for a book of 100,000 characters.

We argue that the chain-code based features do not provide enough distinguishing properties due to the inherent similarity of Bangla characters and thus accuracy gets saturated at some point. It may not be possible to increase the recognition beyond the reported values. In this paper, we tested whether chain-code histogram features of two similar characters are truly different in nature and capable to provide distinguishing properties. First the test statistics were obtained by a statistical test called Analysis of Variance (ANOVA). The result of ANOVA was then used to make a hypothesis test which finally decides whether the features are truly different.

Since our work is primarily based on measuring the efficacy of chain-code features through statistical parameters and does not concern the performance of classifiers directly, further post processing methods and the range of classifiers used in literature are not discussed here.

## III. FEATURE EXTRACTION

We represent each character by a feature vector and 50 such feature vectors of 50 Bangla characters are used to run statistical test on the characters. We employed chain-code based feature extraction technique which is described in detail in [7], [9]. First, we took a bounding box (the smallest rectangular frame enclosing the character skeleton or its contour) from the character image. Then contour of the character is taken. The bounding box image is then divided into 7×7 blocks. Thus, the image is divided into 49 rectangular

blocks. In each block, local histogram of the chain-codes is calculated. Each histogram is quantized into one of 4 possible values, i.e. 0 or 4, 1 or 5, 2 or 6 and 3 or 7. Thus, the feature vector has 4×7×7 = 196 components. We down sampled the 7×7 feature vector to a 4×4 feature vector using Gaussian pyramid for better performance [7]. The pyramid generation equation can be found at [12]. After down sampling, the feature vector has 4×4×4 = 64 components now.

Instead of comparing chain-codes of the same characters from different fonts, we conducted our work only on one font because we are interested in the variation a character possesses compared to the other in the same font type. If the chain-code of these two similar looking characters give high similarity statistics, then we would say chain-codes doesn't give enough information to differentiate between them. And this applies to any two similar looking characters of any specific font.

## IV. ANALYSIS OF VARIANCE (ANOVA)

In statistics, the ANOVA tests the null hypothesis that samples in *k* groups are drawn from populations with the same mean values;

$H_0$: $\mu_1 = \cdots = \mu_k$.

$H_a$: *at least one of the* $\mu_k$*'s is different* (1)

The alternative must include everything else, which can be expressed as at least one of the *k* population means differs from all of the others[13].To do this, two estimates are made of the population variance. These estimators will then be used to calculate a test statistics called *F* value. This *F* value will help determine whether to reject or fail to reject the null hypothesis. For our experiment, we assume *k*=50 for 50 characters.

The feature distribution is normal as chain-code is itself an independent process having its values Gaussian distributed; each time a chain-code is assigned to the contour, the code can have any value from 0 to 7 and it does not depend on the previous value. This normality makes the feature fit for ANOVA testing.

After extracting features from characters, we will test them by ANOVA. Now, if two characters have similar shape, we predict that null hypothesis is failed to be rejected. In other words, features of similar characters have equal mean and do not possess variation that is crucial for classification. Likewise, we predict that the null hypothesis will be rejected for characters with dissimilar shapes. This is because the features of dissimilar characters inherently possess variation and hence their means are not equal.

### A. Test Estimators

The two estimates are variance within groups and variance between groups, which are basically mean squares (*MS*) and will be denoted by *MSW* and *MSB* respectively. Both *MSW* and *MSB* are good estimates of variance, $\sigma^2$. Mean squares are calculated from sum squares (SS), which is the sum of all squared deviations.

$$SS_i = \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_i)^2 \quad (2)$$

The deviation for subject *j* of group *i* is mathematically equal to $Y_{ij} - \overline{Y}_i$, where $Y_{ij}$ is the observed value for subject *j* of group *i* and is the sample mean for group *i*. It can be shown that

$$MSW = \frac{SSW}{dfW} \quad (3)$$

Where, $SSW = \sum_{i=1}^{k} SS_i$ and
$dfW = \sum_{i=1}^{k} df_i = \sum_{i=1}^{k} (n_i - 1) = N - k$; *dfW* is the degree of freedom within the group and *N*= Total number of samples of all groups.

*SSB* is the sum of the N squared between-group deviations, where the deviation is the same for all subjects in the same group. The formula is:

$$SSB = \sum_{i=1}^{k} n_i \, (\overline{Y}_i - \overline{\overline{Y}})^2 \quad (4)$$

where $\overline{\overline{Y}}$ is the grand mean. Because the *k* unique deviations add up to zero, we are free to choose only *(k−1)* of them, and then the last one is fully determined by the others, which is why degree of freedom between group is, *dfB = k − 1*.

$$MSB = \frac{SSB}{dfB} \quad (5)$$

### B. Test Statistic

The next step is to select a statistic for which we can compute the null sampling distribution and that tends to fall in a different region for the alternative than the null hypothesis. The ANOVA produces an *F*-statistic, the ratio of the variance calculated among the means to the variance within the samples.

$$F_{obs} = \frac{MSB}{MSW} \quad (6)$$

If the group means are drawn from populations with the same mean values, the variance between the group means should be lower than the variance of the samples. In other words, if $F_{obs}$ is small, then variability in features between characters is negligible compared to variation in features of a character. On the other hand, if $F_{obs}$ is large, then variability between characters is large compared to variation within character.

### C. Hypothesis test

Restating Eq. 1, the rule for the ANOVA test statistic:
Reject $H_0$, if $F_{obs} \geq F_{\alpha;\, k-1,N-k}$
Fail to reject $H_0$, if $F_{obs} < F_{\alpha\,;\, k-1,N-k}$

Here, $F_{\alpha,k-1,N-k}$ is the critical value on the $F_{k-1,N-k}$ distribution that is used as a cutoff and gives an area in the upper tail= $\alpha$. The $\alpha$ is the significance level. The distribution is shown in Figure 2.

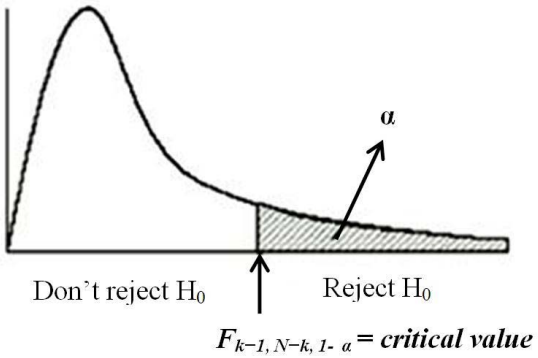



Figure 2. F-distribution [15]

To establish chain-codes as reliable feature extraction method, we should gather enough evidence to reject the null hypothesis, which states that chain-codes of different characters have equal means. In other words, we need evidence that characters are actually different, based on the measured characteristics of chain-codes. If, however, we fail to reject null hypothesis, we have to conclude that chain-code does not provide enough variability between characters. The results of ANOVA are presented in section 5.

## V. RESULT AND ANALYSIS

We run ANOVA for all 50 characters. Our special interest was the characters listed in Table 1 and Table 2. We wanted to see whether chain-codes of these characters could provide high between-class variability. For the groups with two characters in the tables, we give features of these two characters as input to ANOVA. The $\alpha=0.05$, $k$=2, $N$=64×2=128, $dfB$=2-1=1, $dfW$=128-2=126. Hence the critical value of $F_{0.05,1,126}$= 3.92 which can be obtained from [14]. Similarly, for the groups with four characters, $\alpha=0.05$, $k$=4, $N$=64×4=256, $dfB$=4-1=3, $dfW$=256-4=252 and $F_{0.05,3,252}$=2.64. For example, we run ANOVA test for character 9 and 10 in Matlab and the result is shown in Table 3. It is obvious that 0.03<3.92 and hence we fail to reject null hypothesis concluding that there is no strong evidence that the mean value of characters 9 and 10 differ. Matlab gives a probability value for the test called p-value. Here, p-value is greater than the significance level of 0.05 indicating that the observed result would be highly likely under null hypothesis. This result can also be shown using box-plot in Fig. 4. Location of the red lines indicates equal mean value for both feature sets.

Table 3. ANOVA result for character 9 and 10.

| Character | Source | SS | df | MS | $F_{obs}$ | P | $H_0$ |
|---|---|---|---|---|---|---|---|
| ও(9), ঔ (10) | W | 0.00 | 1 | 0.395 | 0.03 | 0.86 | Fail to reject |
| | B | 169.48 | 126 | 1.345 | | | |

To give an idea how much characters differ from each other, we have quantified the p-values and gave denotations as listed on Table 4. However, this is to give a general idea of how anyone can quantify the range of p-values. The ANOVA results for all other similar character groups are listed in Table 5. Not all characters of Table 1 and 2 give low F-value like the previous example. But, it is still less than the critical value, resulting failure to reject the null hypothesis. The groups that show high variation include (7, 8), (24, 45), (36, 46), (36, 41). But the variation is not high enough to provide evidence that their mean value differ.

Following the opposite argument, dissimilar characters should give greater F-value than corresponding critical value. Some examples are listed in Table 6. From there, it is obvious that the features from the character with dissimilar shapes truly differ from each other. There is strong evidence to reject the null hypothesis. Also the p-value is considerably smaller than the significance level which implies that the observed result is highly unlikely under null hypothesis.

Table 4. Ranges of p-value

| Ranges of p-value | Denotations |
|---|---|
| 0.00-0.06 | Very High |
| 0.061-0.10 | High |
| 0.11-0.40 | Moderate |
| 0.41-1.00 | Low |

Table 5. F-statistic for similar characters

| Characters | | F | P | Variation | $H_0$ |
|---|---|---|---|---|---|
| অ (1) | আ (39) | 0.96 | 0.33 | Moderate | Fail to reject |
| উ(4) | ঊ(5) | 0.29 | 0.592 | Low | Fail to reject |
| ই(2) | ঈ (3) | 0.05 | 0.82 | Low | Fail to reject |
| এ (7) | ঐ (8) | 1.7 | 0.187 | Moderate | Fail to reject |
| ও(9) | ঔ (10) | 0.03 | 0.864 | Low | Fail to reject |
| ঋ (6) | ঝ (19) | 0.43 | 0.513 | Low | Fail to reject |
| ড (23) | ড় (44) | 0.65 | 0.42 | Low | Fail to reject |
| ঢ (24) | ঢ় (45) | 1.25 | 0.267 | Moderate | Fail to reject |
| য (36) | য় (46) | 3.1 | 0.08 | High | Fail to reject |
| ব (41) | য় (46) | 0.015 | 0.89 | Low | Fail to reject |
| য (36) | ব (41) | 3.4 | 0.07 | High | Fail to reject |
| ক (11) | ব (33) | 0.28 | 0.595 | Low | Fail to reject |
| ক (11) | র(37) | 0.18 | 0.672 | Low | Fail to reject |
| ঘ(14) | য (36) | 0.00 | 0.99 | Low | Fail to reject |
| ঘ(14) | ম (35) | 0.004 | 0.95 | Low | Fail to reject |
| ব (33) | র (37) | 0.01 | 0.91 | Low | Fail to reject |
| ক (11) | ক্ষ(32) | 0.06 | 0.80 | Low | Fail to reject |

Table 6. F-statistic for dissimilar characters

| Characters | | F | P | Variation | $H_0$ |
|---|---|---|---|---|---|
| আ (39) | ম (35) | 10.2 | 0.00 | Very high | Reject |
| ঘ (14) | ই(2) | 9.08 | 0.003 | Very high | Reject |
| ঘ (14) | ঊ(5) | 17.5 | 0.00 | Very high | Reject |
| ই(2) | ম (35) | 8.8 | 0.003 | Very high | Reject |
| ও(9) | ট(21) | 8.7 | 0.004 | Very high | Reject |
| ড (23) | ঘ (14) | 11.3 | 0.00 | Very high | Reject |
| ঊ(5) | ঐ (8) | 6.5 | 0.01 | Very high | Reject |
| ঔ (10) | জ(18) | 4.57 | 0.03 | Very high | Reject |
| জ (18) | ঋ(12) | 5.02 | 0.02 | Very high | Reject |

## VI. CONCLUSION

In this paper, efficacy of chain-code as features for Bangla script is measured. We proposed that chain-code do not provide necessary variation in describing boundaries of characters due to high similarity in shapes. We validated our proposal through one-way ANOVA, a statistical test to verify whether chain-codes of different characters are truly different. From this test we observed that for similar characters null hypothesis is failed to be rejected which concludes that the characters have similar features. The test gives sufficient evidence that chain-code is not robust to similarity factor and hence we can conclude that chain-code based local histogram features are not strong features for Bangla script.

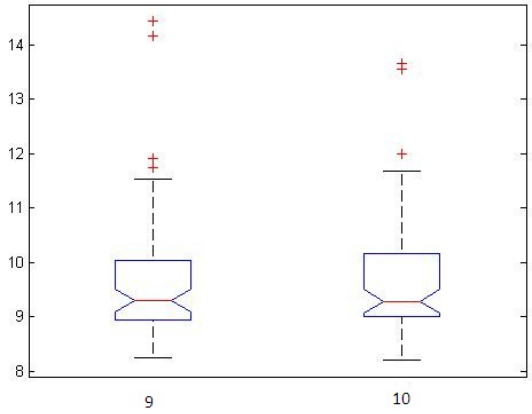



Figure 4. Box-plot indicating equal mean of character ও (9) and ঔ (10)

## References


[1] Trier, D., Jain, A. K. and Taxt, T. 1996 Feature extraction methods for character recognition - a survey, Pattern Recognition, Vol. 29 (4), pp. 641-662

[2] Bag, S., Harit, G. 2013 A survey on optical character recognition for Bangla and Devanagari scripts, Sa dhana , Vol. 38 (1), pp. 133–168.

[3] Nixon, M. and Aguado, A. S. 2008 Feature Extraction and Image Processing, 2nd ed., Academic Press, pp. 247-251.

[4] Devijver, P. A. and Kittler, J. 1982 Pattern Recognition: A Statistical Approach. London: Prentice-Hall

[5] Freeman, H. 1961 On the Encoding of Arbitrary Geometric Configurations, IRE Trans., EC- 10(2), pp. 260–268.

[6] Freeman, H. 1974 Computer Processing of Line Drawing Images, Computing Surveys, 6(1), pp. 57–95.

[7] Bhattacharya U., Shridhar M. and Parui S.K., 2006. On Recognition of Handwritten Bangla Characters, 5th Indian Conference, ICVGIP 2006, Madurai, India, December 13-16, 2006. Proceedings. DOI: 10.1007/11949619_73

[8] Sikdar A., Roy P., Mukherjee S., Das M., and Banerjee S., 2012A Feature based chain code method for identifying printed

[9] Reza, K.N. and Khan, M., 2012. Grouping of Handwritten Bangla Basic Characters, Numerals and Vowel Modifiers for Multilayer Classification. 10.1109/ICFHR.2012.206

[10] Arora S., Bhattacharjee D., Nasipuri M., Basu D.K., and Kundu M., 2010 Combining Multiple Feature Extraction Techniques for Handwritten Devnagari Character Recognition. arXiv:1005.4032v1

[11] Pal U., Roy K., and Kimura F., 2010. A Lexicon Driven Method for Unconstrained Bangla Handwritten Word Recognition. citeulike:9243964

[12]http://sepwww.stanford.edu/data/media/public/docs/sep124/ssen1/paper_html/node3.html

[13] Howell, D. 2002 Statistical Methods for Psychology, Duxbury, pp. 324–325

[14] http://www.danielsoper.com/statcalc3/calc.aspx?id=4

[15] http://www-hsc.usc.edu/~eckel/biostat2/notes/notes7.pdf

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Motion Region Detection and Tracking Based on Temporal Differencing and Adaptive Background Subtraction

Kaushik Deb*, Sayem Mohammad Imtiaz, Priyam Biswas, Md. Moshiul Hoque
Chittagong University of Engineering & Technology (CUET), Chittagong-4349, Bangladesh
E-mail: debkaushik99@gmail.com*

***Abstract*—Now-a-days video surveillance is indispensable in security sensitive areas. Hence, a significant amount of work has been done in this field. This paper proposes a hybrid algorithm for motion region detection and an appearance based real time motion tracking system. Initially, foreground map is extracted through subtraction process from a background model applying temporal differencing method. Then shadow elimination and morphological operations are used to remove several noises. Finally models are initiated for each detected motion region by extracting features like center of mass, color correlogram etc. which are then used for tracking purpose. As similarity distance within a certain radius is measured, probability of confusing object is reduced considerably and therefore performance is optimized significantly. The proposed algorithm also uses a robust technique to label people within group. This system has the capability to work in indoor and semi-outdoor and even in outdoor environment that generates penumbra shadow as well as it handles occlusion, group effectively. The algorithm takes good care of false foreground pixels due to penumbra shadow. Hence this system will play a pivotal role in providing security in highly confidential areas**.

**Keywords—Background subtraction, temporal differencing, shadow elimination, Morphological operation, Color correlogram.**

## I. INTRODUCTION

In today's modern world, video surveillance has numerous applications including smart video data mining, public and commercial security, in military and law enforcement agencies etc. Developing algorithm for core parts like object detection and tracking does not only increase the odds in favor of video surveillance, rather other application domains like video compression, augmented reality; virtual reality, human machine interface etc also get directly benefitted from it. Hence video surveillance system has caught attention of researchers.

The target of video surveillance system is to detect motion region along with analysis of suspicious move detection as well as tracking them over the time considering information such as color, shape, size, center of mass etc.

Segmentation and tracking is the most vital part of the video surveillance system. Because later phase like event recognition largely depends on output of previous stages. A lot of challenges are faced during real time surveillance like light change, shadow that alters the segmentation process, dynamic background etc. Among them shadow is important that need to be dealt with carefully to get correct segmentation. Labeling entity separately in group is another great challenge to overcome.

In this paper, hybrid model has been proposed which combines adaptive background subtraction with temporal differencing for foreground extraction. Then shadow region are detected and finally eliminated with the computation of brightness and chromaticity distortion. In addition, model for each entity with color correlogram, center of mass is initiated which are used for tracking purpose. We have confined our study only in video sequences with static background.

The rest of this paper is organized as follows. Section II gives a brief description of related research. Section III introduces our proposed moving object detection and tracking scheme. Section IV compares the performance of our proposed scheme with some other approach. Finally, section V concludes this paper.

## II. RELATED RESEARCH

Segmentation and tracking is two paramount steps of the automatic video surveillance system. Over the past years a significant amount of work has been done in this area.

There are two kinds of techniques for segmentation: optical flow computation [2] and background subtraction. The optical flow performs better, but computationally expensive method and unsuitable for real time system. Background subtraction is most commonly used technique for segmentation in [6]. We have combined adaptive background subtraction method that is compatible with gradula light change and temporal differencing to compensate missing pixels.

Shadow is a big obstruction for acquiring correct segmentation. Reference [5] uses chromaticity and brightness distortion to detect shadow and its most commonly used technique in literature.

Tracking algorithm establishes the correspondence between the objects in current frame with the objects in previous frame. This phase is important because it provides cohesive temporal data about moving objects which are used in both lower level processing like motion segmentation and higher level processing like behavior recognition. In order to track the

people, the object model is needed to be initiated. In this work, our primary target object is human. So we consider human model that might include features, such as the color, aspect ratio, edge, velocity etc. Correspondence matching and motion estimation are two methods commonly used for tracking in [4]. We use feature such as center of mass and color correlogram to track the human.

In tracking object, full and partial occlusion is a major challenge to overcome with. In order to solve the problem of occlusion appearance based tracking method [1] were proposed. In our system color correlogram is used for occlusion handling.

## III. PROPOSED METHOD

Our proposed method consists of steps shown in below system block diagram (Figure 1). A brief overview on these steps is given below.

### *A. Foreground detection*

For foreground detection we used a hybrid background subtraction algorithm. One is adaptive background subtraction algorithm where a pixel x is detected as foreground if following condition holds:

$$|\mathrm{I_n(x)} - \mathrm{B_n\,(x)}| > T \qquad (1)$$

Here $\mathrm{I_n}$ is current frame and $\mathrm{B_n}$ is background model at time n and T is the threshold value. Background model is updated with:

$$\mathrm{B_{n+1}(x)} = \alpha \mathrm{B_n(x)} + (1-\alpha)\mathrm{I_n(x)} \ \text{ if x} \in \mathrm{BG} \quad (2)$$

$$\mathrm{B_{n+1}(x)} = \mathrm{B_n(x)} \text{ if x} \in \mathrm{FG} \qquad (3)$$

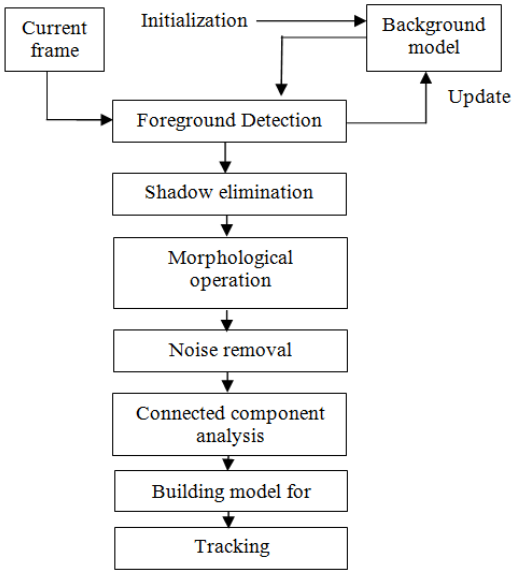



Figure 1: The proposed system block diagram

Here, BG=Background pixel and FG=Foreground pixel. $\alpha$ And T is empirically calculated constant.

Another used method is temporal differencing where a pixel x is detected as foreground if following condition holds.

$$|\mathrm{I_n(x)} - \mathrm{I_{n-1}\,(x)}| > T \qquad (4)$$

Here $\mathrm{I_{n-1}}$ is the frame at time n-1. This method is highly adaptive to dynamic scene changes. So, combining both method yields better segmentation.

### *B. Shadow elimination*

Brightness distortion and chromatic distortion is calculated to find the shadow of moving objects. $I(x) = [I_R(x), I_G(x), I_B(x)]$ Are the *RGB* components of the pixel *x* from the current image and $B(x) = [B_R(x), B_G(x), B_B(x)]$ those of its corresponding pixel in the background image. Then we define the brightness distortion $\delta Br(x)$ of $I(x)$ with regard to $B(x)$ as the difference between $B(x)$ and $I`(x)$ the projection of $I(x)$ on B(x) (Figure 2), formally,

$$\delta\mathrm{Br(x)} = \|\mathrm{B(x)}\| - \frac{\mathrm{I(x).B(x)}}{\|\mathrm{B(x)}\|} \qquad (5)$$

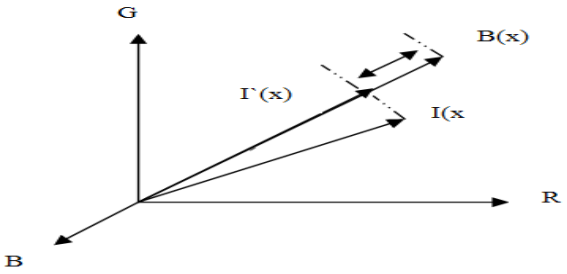



Figure 2: Brightness Distortion

Three cases arise: $\delta Br(x)$ take a null value if no brightness change is observed in the scene, has a positive value if *x* is a shadowed pixel in this case we rather speak about a darkening level, and finally a negative distortion denotes an intensity gain.

Chromatic distortion $\delta Cr(x)$ of the vector $I(x)$ with regard to the background vector $B(x)$ as the angle □ *(I(x), B(x))*:

$$\delta\mathrm{Cr(x)} = \arccos\left[\frac{\mathrm{I(x).B(x)}}{\|\mathrm{I(x)}\|.\|\mathrm{B(x)}\|}\right] \qquad (6)$$

So, a pixel is classified as shadow if its darkening level is reasonable with a weak chromatic distortion:

$$\delta Br(x) < \tau$$

$$\text{And } \delta Cr(x) < \theta$$

Here, θ and τ is empirically selected threshold value.

Again a pixel is classified as highlighted if it gains intensity also with a weak chromatic distortion. Those pixels are resorted as background pixel.

*C. Pre-processing before tracking*

Several kinds of noises might arise like highlighted pixels, insignificant objects and inaccurate segmentation due to camera noise, light reflectance etc. For better result in later phases, these noises should be reduced.

Morphological operations are applied to foreground objects to remove noise and discontinuities remaining after background subtraction process. We have applied two dilation followed by three erosion, again, followed by one dilation with various structuring element which acquired through experiment. Due to these operations, holes in image get eroded and discontinuities get filled.

We, then, use result of shadow detection phase and use it to detect foreground false pixels due to highlight. A graph algorithm is then used to label different objects and insignificant objects are removed. Through that graph traversal important features of detected object such as bounding box rectangle, center of mass etc. are calculated.

*D. Tracking*

For tracking we have used two features such as center of mass (CM) and color correlogram. Center of mass is used to reduce the search space to match current objects with previous objects which means we considered only few objects or possibly none of previous frames which are located within a certain radius of current object and for multiple candidates we have used comparatively more effective feature color correlogram than histogram to get the correct match. .

CM is $c_m = (x_{c_m}, y_{c_m})$ calculated from:

$$x_{c_m} = \frac{\sum_i^n x_i}{n} \tag{7}$$

$$y_{c_m} = \frac{\sum_i^n y_i}{n} \tag{8}$$

Two objects are said to be close if Euclidean distance between current object and previous objects center of mass is below a threshold.

Correlogram is defined with:

$$\gamma_I(c_i, c_j, k) = \frac{r(c_i,c_j,k)}{8kH_I(c_i)} \tag{9}$$

$r(c_i, c_j, k)$ Is the count of total occurrences of two pixels one with color $c_i$ and another with $c_j$ at a distance k. And 8k is a factor representing *how* many neighboring pixels in distance k. $H_I(c_i)$ Is the count of pixels of color $c_i$ in the image I.

Distance between two image I and Í is measured with following equation:

$$D_\gamma(I, Í) = \frac{\sum_{\forall i,j,k} |\gamma_I(c_i,c_j,k) - \gamma_Í(c_i,c_j,k)|}{1 + \sum_{\forall i,j,k} \gamma_I(c_i,c_j,k) + \sum_{\forall i,j,k} \gamma_Í(c_i,c_j,k)} \tag{10}$$

Similarity measure can be derived by taking its complement to 1:

$$S_\gamma(I, I) = 1 - D_\gamma(I, Í) \tag{11}$$

Histogram and correlogram is updated with:

$$h_I(c_i, t) = \alpha h_I(c_i, t-1) + (1-\alpha) h_I^{new}(c_i, t) \tag{12}$$

$$\gamma_I(c_i, c_j, k, t) = \alpha \gamma_I(c_i, c_j, k, t-1) + (1-\alpha)\gamma_I^{new}(c_i, c_j, k, t) \tag{13}$$

Correlogram is a new feature which is much more effective than histogram. But we only checked those previous objects which are reasonable within reach of current object which is a great performance optimizing tonic we are using in our proposed method. This slightly different approach brings faster execution and little confusion in object matching.

*E. Occlusion handling*

At first, event of occlusion needs to be detected. If number of object increases and any object in current frame overlaps with more than one previous object then we say occlusion has occurred. Then we try following approach to label object correctly in occlusion.

Let p be a pixel located in object G, and let G(p) denote its color. Let $\prod_p(G|M_m)$ be the likelihood of pixel p belonging to model $M_m$. Then p will be labeled as belonging to model $M_m$ if and only if

$$m = \text{argmax}_i \prod_p (G|M_i) \tag{14}$$

Where

$$\prod_p(G|M_i) = \beta \pi_{G_{(p)},h}(G|M_i) + (1-\beta)\pi_{p,\gamma}(G|M_i) \tag{15}$$

With

$$\pi_{G_{(p)},h}(G|M_i) = \min\left\{\frac{H_{M_i}(G(p))}{H_G(G(p))}, 1\right\} \tag{16}$$

$$\pi_{p,\gamma}(G|M_i) = 1 - D_\gamma(M_i, \{p\}) \tag{17}$$

Here $D_\gamma(M_i, \{p\})$ is the distance between the local correlogram calculated at pixel p and the part of correlogram for $M_i$ that corresponds to color G(p) and $H_{M_i}(G(p))$ is the count corresponding to $h_{M_i}(G(p))$.

We detect spilt in order to track them after occlusion breaks. If any previous object overlaps with more than one current object and number of object decrease then we say split has occurred. Then they are tracked normally.

## IV. EXPERIMENTAL RESULT

Proposed system has been tested with different video sequences in different environment and in varying lighting condition. A machine with 32-bit operating system, 1 gigabyte (GB) RAM, Intel® Pentium® Dual Core 2.20GHz has been used for testing purpose.

Now, here we include some of our experimental results. For some same sample video sequences, we have generated standard CAMSHIFT tracked result and result of the work proposed in [3] and our result. Below is the data set, we have generated.

TABLE I. COMPARISON DATA

| Ground Truth Data (x, y co-ordinates) | CAMSHIFT Result | | Amin et. al Result | | Our Result | |
|---|---|---|---|---|---|---|
| | *(x, y)* | *DE* | *(x, y)* | *DE* | *(x, y)* | *DE* |
| 598, 376 | 596, 368 | 8.25 | 600, 375 | 2.24 | 599 375.5 | 1.12 |
| 598, 375 | 572, 361 | 29.53 | 596, 375 | 2.00 | 599 375.5 | 1.12 |
| 593, 376 | 572, 363 | 24.70 | 596, 375 | 3.16 | 598.5 375 | 5.59 |
| 593, 376 | 568, 363 | 28.18 | 593, 375 | 1.00 | 592.5 375.5 | 0.71 |
| 588, 375 | 567, 364 | 23.71 | 593, 375 | 5.00 | 589.5 375.5 | 1.58 |
| 581, 375 | 561, 368 | 21.19 | 587, 375 | 6.00 | 583.5 375.5 | 2.55 |
| 581, 375 | 554, 367 | 28.16 | 581, 375 | 0.00 | 580.5 375.5 | 0.71 |
| 576, 375 | 552, 363 | 26.83 | 581, 375 | 5.00 | 578.5 375 | 2.50 |
| 572, 375 | 545, 364 | 29.15 | 575, 375 | 3.00 | 573 375 | 1.00 |
| 572, 375 | 541, 365 | 32.57 | 570, 375 | 2.00 | 569.5 375.5 | 2.50 |
| 567, 375 | 540, 365 | 28.79 | 570, 375 | 3.00 | 568.5 375 | 1.50 |
| 566, 375 | 535, 367 | 32.02 | 564, 375 | 2.00 | 563.5 375 | 2.50 |
| 560, 375 | 535, 367 | 26.25 | 564, 375 | 4.00 | 562.5 375 | 2.50 |
| 560, 375 | 530, 369 | 30.59 | 558, 375 | 2.00 | 557.5 375 | 2.50 |
| 554, 375 | 529, 370 | 25.50 | 558, 375 | 4.00 | 557.5 374.5 | 3.50 |
| 547, 374 | 526, 371 | 21.21 | 551, 375 | 4.12 | 550 374.5 | 3.00 |
| 547, 375 | 523, 369 | 24.74 | 545, 374 | 2.24 | 545 374.5 | 2.00 |
| 541, 375 | 524, 368 | 18.38 | 545, 374 | 4.12 | 543.5 374 | 2.23 |
| 539, 374 | 523, 370 | 16.49 | 541, 374 | 2.00 | 539.5 374 | 0.50 |
| 538, 374 | 520, 370 | 18.44 | 537, 374 | 1.00 | 537 374 | 1.00 |
| 536, 373 | 520, 369 | 16.49 | 537, 374 | 1.41 | 536.5 374 | 1.12 |
| 536, 373 | 518, 368 | 18.68 | 535, 373 | 1.00 | 534.5 374 | 1.12 |
| 534, 374 | 518, 368 | 17.09 | 535, 373 | 1.41 | 534 373.5 | 0.50 |
| 534, 373 | 516, 367 | 18.97 | 533, 373 | 1.00 | 32.5 373.5 | 0.50 |
| 532, 373 | 516, 367 | 17.09 | 533, 373 | 1.00 | 532 373.5 | 0.71 |

For a same sequence of 25 frames, we have compared our system's output with standard CAMSHIFT tracked output and a system proposed in [4] which employs an appearance based tracking system with color correlogram feature. Here ground truth data represents center of rectangular area covered by original objects i.e. here human. Displacement error (DE) is the distance between ground truth data and tracked center of location of Region of Interest (ROI) in image frame. Comparison shows average DE of CAMSHIFT result is 30 and average DE of approach described in [3] is 2.62 and our system produces 1.74 averages DE. That is, our system is detecting and tracking ROI in image frame more accurately than of these two methods.

Detecting and tracking people in case of occlusion is a challenging task.


Figure 3: Left column shows original image and right column shows detected and tracked image. Sample frames show performance during occlusion. (a) Before occlusion, (b) During occlusion, (c) After split

Above indoor video sequences shows performance before and after occlusion. As we see, system is tracking people right during occlusion and after split.

Another set of data on similar condition from different video sequence is shown in Figure 4. There also we see, system is handling well challenge posed by occlusion. Frame no. 2-5 is showing performance of the system during partial occlusion and frame no. 6 shows system is good with full occlusion too. Frame no. 7 shows performance after split.

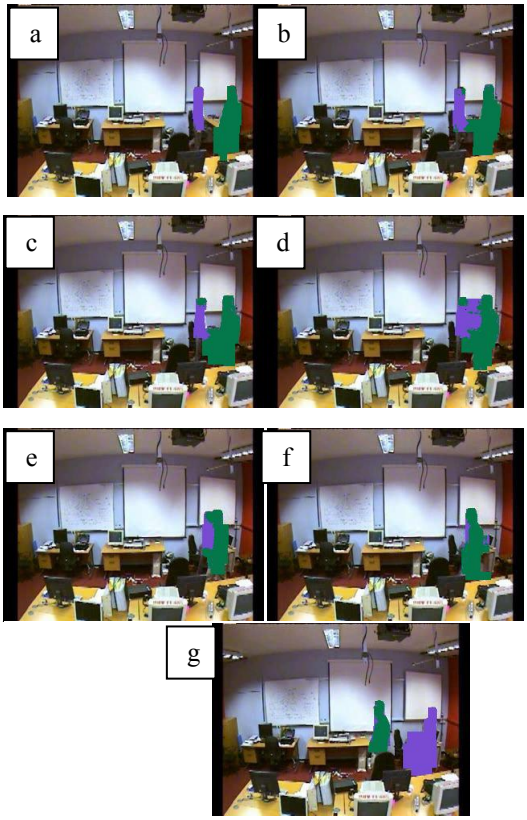



Figure 4: Sequences showing performance before occlusion and after split. (a) Before occlusion, (b), (c), (d), (e) During partial Occlusion, (f) Full occlusion and (g) After split

Another obstacle faced by video surveillance system is shadow. System not dealing with shadow suffers from inaccurate segmentation problem hence affecting result of later phases. Our system deals greatly with shadow.

Below figure 5 shows performance of system with shadow:

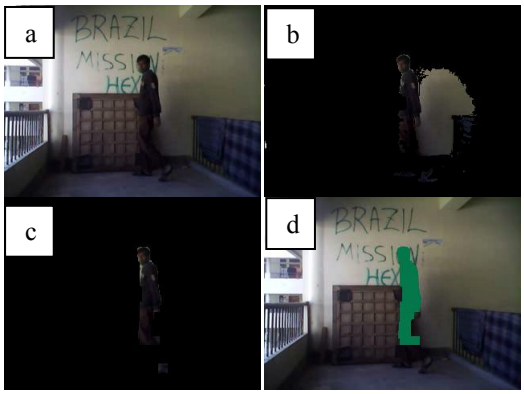



Figure 5: (a) Original image, (b) inaccurate segmentation due to shadow, (c) after elimination of shadow, (d) finally tracked object

Sample two set of frames shown in figure 6, where in first one, illumination is uniform but with shadow effect which system handles well and second set which got non-uniform illumination. Hence resulting in falsely detected foreground pixels and figures show our system copes well with this situation as well.

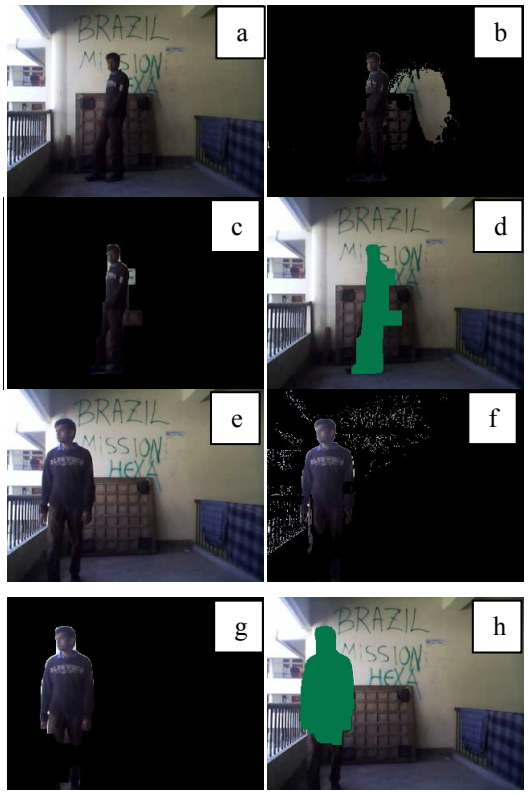



Figure 6: Samples showing system's performance with highlighted pixels. Here (a) Original image where illumination is uniform, (b) Segmentation with false foreground pixels due to shadow, (c) after elimination of noise, shadow, (d) finally tracked object. Again (e) Original image where illumination is not uniform, (f) Segmented foreground along with false detected pixels due to highlight, (g) after removal of effect of highlight and other noise and (h) finally tracked object

All the experimental result shown here indicates that our proposed system is compatible with various obstructions faced during real time video surveillance. It deals effectively with occlusion and split and consistent with reflection and rotation of objects etc. Comparison with CAMSHIFT tracked result and another approach proposed in reference [3] shows that our system detects and tracks object more accurately than those two methods. Also combining center of mass with color correlogram feature significantly reduces total processing time as it ensures that only some previous objects need to be checked with current object and hence reducing probability of conflict in matching.

## V. CONCLUSION

The study presents a robust and fast approach for real time video surveillance. It combines two different algorithms for foreground detection which is very adaptive with both global illumination change and local illumination change like

highlight and shadow. We also used a robust method to detect shadow that has no problem with coping with lighting change over the frame. In tracking, we brought significant changes from traditional appearance based model which merely works with color histogram or correlogram. We used center of mass of each object to reduce the search space. That is number of similarity distance measure required is reduced and if there is more than one candidate then we used comparatively more effective feature color correlogram than histogram to get the correct match. Hence system performs faster. Finally we compared our result with standard CAMSHIFT tracked output and output of a system proposed in [3] which reveals our system is more accurately detecting and tracking human. Also, our system is good with penumbra shadow, highlight, rotation, disturbing background etc. These results indicate clearly that our proposed system is totally fit to be played in real time. In future, we will extend our system to work with non-penumbra shadow and behavior recognition.

## REFERENCES


[1] M. Balcells-Capellades, D. DeMenthon and D. Doermann. "An Appearance-based Approach for Consistent Labeling of Humans and Objects in Video." Pattern Analysis and Applications, pp. 1433-7541, November 2004.

[2] S. S. Beauchemin and J. L. Barron, "The Computation of Optical flow" ACM Computing Surveys, Vol.27, 1995, pp 433 – 466.

[3] Aı Amin Hosain, Anik Saha, "An appearance based human motion tracking", in proceeding of ICIEV'2013.

[4] A. A*mer*, "Voting-based simultaneous tracking of multiple video objects", In proceeding of the SPIE International Symposium on Electronic Imaging, pp. 500–511, Santa Clara, USA, January 2003.

[5] S.J. McKenna, S. Jabri, Z. Duric, and H. Wechsler, "Tracking interacting people", In proceeding of the International Conference on Automatic Face and Gesture Recognition, pp.348–353, 2000.

[6] A. M. McIvor, "Background subtraction techniques", In proceeding of the Image and Vision Computing, auckland, New Zealand, 2000.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Layout Analysis of Technical Documents in a Universal Reader

AHM Mahfuzur Rahman[1,*], Albert Astals Cid [2]
[1] Bangladesh University of Engineering and Technology, Bangladesh
[2] The KDE Community, www.kde.org
*Email: 65mahfuz90@gmail.com

*Abstract*— **A significant amount of work has been done on layout analysis of PDF documents, but there is no system that can work on versatile types of electronic documents. In this paper we have presented a system that works on varieties of document types (PDF, Djvu, Epub etc.) and does layout analysis to find text reading order. The only assumption that we have made is that every document format should give us the bounding blocks of each of the basic text elements: letters, words or sub-words; for the case study we have chosen a popular open source document reader: Okular. We first collect texts generated from Okular, execute some pre-processing on those texts and form words and lines; after that, with the help of statistical information elicited from basic text elements, words and lines, we formulate a modified recursive XY Cut based algorithm. Finally after some post-processing we return the output back to Okular. We do not use any high level information from document types, and our system is fault tolerant to ill-formed technical documents where there may be limited overlapping between column boundaries (in multi-column documents).**

***Keywords*— Document Analysis, Universal Reader, Okular, Reading Order, XY Cut.**

## I. Introduction

Document Layout Analysis is a key element in document image retrieval and understanding. A document (technical) is typically formatted in different regions: title, affiliations, abstract, sections etc. The target of a document layout analysis system is to understand different regions in a document image and find out relationship among them in order to present information provided in a document automatically, and in correct order to any human being or to another system.

Document layout analysis is typically called document image analysis as most of the works are done on scanned images of documents in high resolution. A lot of works [1], [2] has been done for general document image layout understanding. In this paper, however, our concentration is specifically on technical documents, mostly conference and journal papers. A main advantage of technical documents is that they consist of non-overlapping rectangular blocks, and therefore easier to detect through an automated system. From our experiment, we have found that Recursive XY Cut algorithm proposed by Nagy and Seth [3] and its later improvements [4]–[6] are time-efficient and well-suited techniques for layout analysis of technical documents.

Apart from document image layout analysis, there has been some works in literature those consider electronic documents; most of these works [7]–[12] regard PDF as the primary electronic document to perform the algorithms. There are document readers which work with more than one type of documents; Okular [13] is one of them, and probably one of the most versatile document readers to date. If we want to integrate document layout analysis and reading order detection facilities in this kind of readers, we have to make a general framework through which it is possible to incorporate support for several types of electronic documents.

### A. Overview of the System

Okular is a universal document viewer which supports varieties of electronic file formats including PDF, Djvu, Chm, Epub, PS, Dvi etc. Elements in all kind of documents can be divided in two parts: text and image; here we keep ourselves concerned with the text part only. To provide text information irrespective of any document format supported by Okular, it has a general text generation module called the generator which provides all of the texts in a document, not necessarily in order. With our proposed system, it is possible to turn the unordered jumble of words into an ordered, human readable, combination of texts.

We first take the ungraded texts from Okular text generator for any document format where the available text can be as characters, words or sub words; we then remove unnecessary spaces from the text, construct words from characters (if there is already a word, it is reserved) and create lines from words; after that, we generate some statistical information considering word spacing, line spacing and column spacing which are further used in the main part of the algorithm: modified XY Cut algorithm. In the algorithm, we divide any text page to different regions and then find out text order within all regions distinctly and later merge the regions together to provide the final reading order of the document. To

This research was greatly assisted monetarily by Google, and technically by The KDE Community during Google Summer of Code(GSOC), 2011 (https://www.google-melange.com/gsoc/homepage/google/gsoc2011). Technical details about the implementation was documented in my blog (http://nightcrawlerinshadow.wordpress.com/2011/08/20/advanced-text-selection-in-okular/) previously.

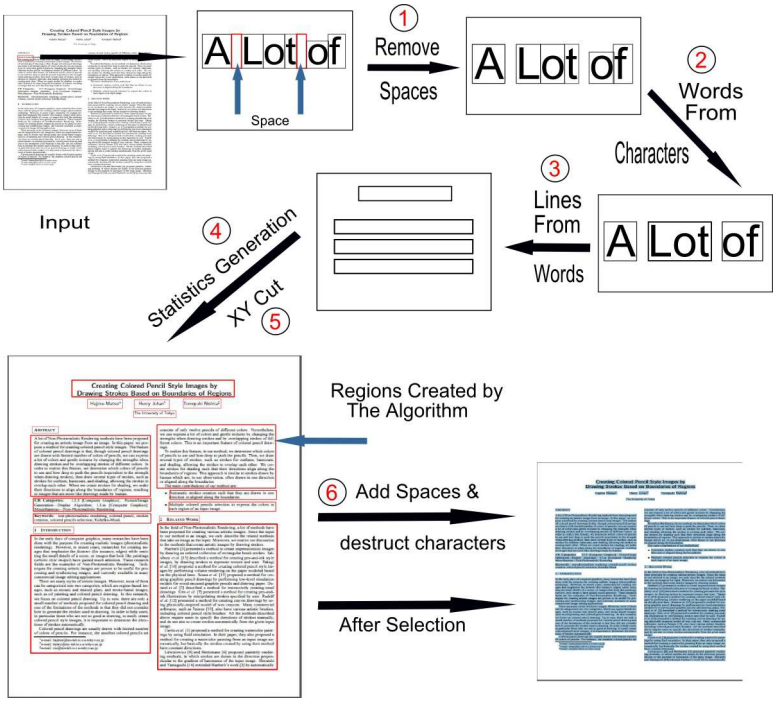



Figure 1: Process Overview (output after no. 6, Add Spaces & destruct characters is exactly opposite of steps 1 and 2, combined; output of step 4 is some numerical values).

keep congruence with the form in which the generator gives us texts for processing, before returning the output to generator, we add the spaces within the words in the output, and also extract characters from words, if necessary. An overview of the whole process is provided in Fig. 1.

The rest of the paper is organized as follows. In section II, we give a detailed description of the overview we provided in this section and devise our algorithm. Section III provides the results of our algorithm and compares the results with the outputs before the implementation of the algorithm; finally we conclude in section IV.

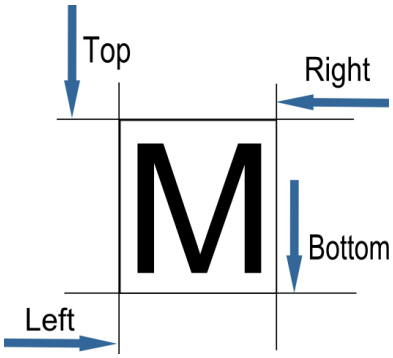



Figure 2: Bounding rectangle of a character.

## II. Document Layout Analysis

Before we are ready to employ XY cut algorithm, we have to perform some pre-processing. As we have already discussed in section I-A, a universal document reader like Okular has to support a lot of different document formats; it therefore needs to provide texts in a document page in some pre-defined format; Okular gives texts as characters (in PDF), words (in Djvu) or sub words (in DVI). For the development of our algorithm, we presume that the generator or any generalized text provider will give us text in any of the three formats discussed above; we call it the basic building block of text. We also assume that, any basic building block will have a bounding rectangle; for example, a character has a defined boundary concerning PDF file formats as shown in Fig. 2.

Basing on these assumptions, we divide our algorithm in several parts, which we are going to describe in the following several subsections.

### A. Remove Spaces

From the Okular generator, we take as input a list of words (may not be full words) within a document page. Then as a pre-processing, we remove the spaces from the texts. In case of DVI generator, we get text as sub words, and there are unnecessary and invalid spaces within sub words, that are removed through this process. Document formats like djvu is untarnished by this process, as djvu directly provides us with a full word or words without any spaces in between them (djvu generator in Okular does not preserve spaces between words).

### B. Construct word from characters

As we have already removed the unnecessary spaces, now we can construct words from characters or sub words. We should notice that it is quite unlikely for every character in a word to have similar boundaries, that is

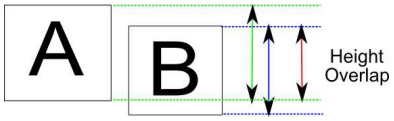



Figure 3: Overlap of Heights between two bounding blocks; Green: Height of Block 1(left), Blue: Height of Block 2 (Right) and Red: Overlap of Heights.

similar vertical start point (top) and end point (bottom), when we consider horizontal lines. Keeping this in mind, we make our word construction module flexible; from our experiment on documents we have allowed characters or sub words to merge into a word only if their mutual overlap of heights (according to Y coordinates) is more than 60% (according to Fig. 3). A significant advantage of this step is that we can now run all of our computation on words in spite of characters, which is more time efficient. This step of the algorithm depends a lot on the output of the generator, because if there are unnecessary spaces within words (as is the case in DVI generator for Okular), then we get only sub words as output. Also in this step, we keep a corresponding character mapping for every constructed word so that we can revert to the basic character blocks at the last step of our algorithm.

### C. Create Lines from words

In this step, we will create lines from words which is very crucial as every line sorts words within a line and provides statistical information about line spacing between consecutive lines in a text. Given the words and bounding rectangles of the words, line drawing procedure is described in algorithm 1.

### D. Generate Statistical Information

In this step, we generate statistical information regarding word spacing, line spacing and column spacing. Word spacing is necessary to provide information about the average spacing within words in a document and is very conducive to determine column spacing, which provides us essential information to segment a multi-column document in different columns; line spacing is necessary to divide text regions which are separated considerably by more than usual line spacing. This method is used repetitively by Modified XY Cut algorithm to generate global (within a document page) and local (within regions of a whole page) statistical information; to know how to generate statistical information see algorithm 2.

### E. Modified XY cut algorithm

We have used a modified version of XY cut algorithm; we first take the whole page as our root of the XY Cut tree (though we call it a tree according to the literature, here we see it more like a list) and try to divide this tree in different regions until it is impossible to make further division. To make a decision of partition, we first calculate projection profiles along the X axis and Y axis.

**Algorithm 1** $LineGrowing(Words)$

Sort all of the $Words$ according to their Y coordinate in ascending order (according to top values of bounding rectangles).
**for** each $Word$ in $Words$ **do**
  $newLine$ = true.
  **for** each $Line$ in generated $Lines$ (Initially there will be none) **do**
    **if** Height of bounding rectangles (according to Y-coordinates) of $Line$ and $Word$ overlaps by more than 70% **then**
      Append $Word$ to $Line$.
      Redefine the boundary of $Line$.
      Set $newLine$ = false.
    **end if**
    **if** $newLine$ == false **then**
      **BREAK**
    **end if**
  **end for**
  **if** $newLine$ == true **then**
    Make a new Line and insert into $Lines$
  **end if**
**end for**
**for** each $Line$ in $Lines$ **do**
  Sort the $Words$ in a $Line$ according to their X coordinate in ascending order (of left value of bounding rectangle)
**end for**
**return** $Lines$

**Algorithm 2** $Statistics(Words)$

$Lines$ = $LineGrowing(Words)$
Find line spacing between lines from $Lines$ and take a weighted average to generate $LSpacing$.
**for** each $Line$ from $Lines$ **do**
  Find the maximum spacing within the line and keep it in $COLSpaces$ ($COL$ refers to column).
  Keep other spaces in $WORDSpaces$.
**end for**
Take weighted average from $COLSpaces$ to generate $CSpacing$.
Take weighted average from $WORDSpaces$ to generate $WSpacing$.
**return** $LSpacing$, $WSpacing$, $CSpacing$ (for Line spacing, Word spacing and Column spacing respectively)

We calculate the statistical information of word spacing, line spacing and column spacing from algorithm 2; using these information and two region threshold parameters, we take the decision of dividing in further regions; but prior to that, we do some pre-processing like boundary white space removal, which is basically shrinking the text region where there is no text and to make the system fault-tolerant to some extent, we make provision to avert stray text overlapping (noise) between columns. Algorithm 3 describes the process in detail.

**Algorithm 3** $XYCut(Words, PageBoundary)$

1: Make the first element of a list $XYTree$.
2: **while** We have not reached the end of $XYTree$ **do**
3: $LSpacing$, $WSpacing$, $CSpacing$ = $Statistics(Words)$.
4: $ThreshX$ = 2 * $WSpacing$, $ThreshY$ = 2 * $Lspacing$.
5: Generate $HorProj$ and $VertProj$ from projection of text elements on the Y axis and X axis of $PageBoundary$.
6: For Noise (Trivial text overlapping between columns), find average of $VertProj$ and subtract this average from all of the elements in $VertProj$.
7: Cleanup boundary white spaces (where there is no text).
8: Let maximum gap from $HorProj$ be $GapX$ and from $VertProj$ be $GapY$.
9: Segment the current region into two regions only if $GapX$ or $GapY$ exceeds corresponding thresholds, $ThreshX$ and $ThreshY$; In cases where both of the gaps meet the requirement, segment according to the greater gap between $GapX$ and $GapY$. Remove the current region and insert two newly created regions.
10: **end while**
11: Traverse all the regions in list in order and get the ordered Lines from there to insert in $TextLines$.
12: **return** $TextLines$

### *F. Add Spaces and Destruct characters*

In this part, we have all of the words of the input page in sorted lines $TextLines$, but we do not have any space in between words, which seems unnatural; for this reason, we will insert space characters in places where there should be spaces. We will find two consecutive words and whether there is a gap in between their boundary rectangles; if there is empty space between those rectangles, we put a space character there instantly. This step will insert spaces in djvu files, as generator for this file format only gives words and no spaces within them are reserved.

After the addition of spaces, we destruct characters from the words, using our word to character correspondence, in order to return to the original characters that we got from the generator.


(a) (b) (c) (d)

Figure 4: djvu selection comparison 4a: before implementation, 4b: after implementation; PDF selection comparison 4c: before implementation, 4d: after implementation

## III. Results and Performance Analysis

As there is no present system for multiple document layout analysis, we cannot give a direct comparison with previous works. Also the works done on PDF document layout analysis have not provided any presentable methodology to compare with previous works. For this reason, we have presented comparison of our work in respect of the previous outputs of Okular.

A graphical comparison of the results before and after the implementation of our algorithm in Okular for Djvu and PDF file formats are shown in Fig. 4. For the selection in djvu files, the improvement is twofold:

1) There has been added space between words after the implementation of algorithm as we can see from the selection, which is the contribution of the last sub step: addition of spaces within words.
2) Previously every word had a different bounding rectangle, but now we can see a continuous rectangle in selection; this has been possible for the redefinition of bounding rectangles for words and lines in the steps words from characters and lines from words.

There are some cases where our algorithm does not perform well, which are given below:

- **Poor Performance of Generator:** For our algorithms, we have assumed that the generator gives us texts as characters, words or sub words; but in cases of chm files, the current generator gives text as multiple lines. As we depend much on line

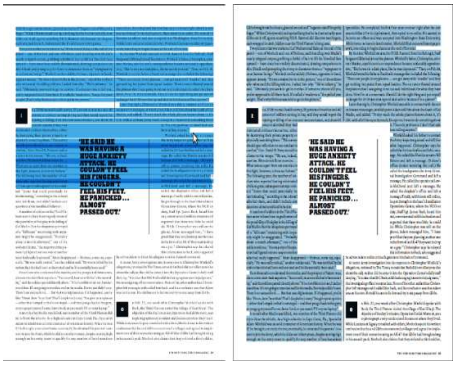


Figure 5: Error of the algorithm: column overlaying; left: our output, right: intended output (From Adobe Acrobat Reader).

segmentation for our algorithm and do not handle input like multiple lines, it results as a failure.

- **Column Overlapping and Overlaying:** As our algorithm uses a variety of XY Cut algorithm, it suffers from the same problems as the original XY Cut algorithm does [6]. Though our algorithm can support a little column overlapping detecting it as noisy condition (Algorithm 3), massive overlapping is not supported. It is not much problematic as for now, as our target in this work is mainly to give support to technical documents. An example of failure is explained through Fig. 5.

## IV. Conclusion

We have proposed a document layout system for the Open Source document reader Okular, which supports versatile types of electronic documents. To support all of these document formats, we have devised a modified XY Cut algorithm; also we have added some pre-processing and post-processing steps to contain all the document formats. Our system takes input from the output of Okular generator as generated texts in an undefined manner; we then pre-process the texts, generate words and lines from there and extract some statistical information to use in the XY Cut algorithm. After running the algorithm, we do some post-processing and return Okular back the output in the format of the original input.

Our work shows promising results for technical documents but with some additional changes in the algorithm it is possible to support more complex documents like consumer magazines, which are now available in different formats like (PDF, Epub, Mobi) etc. Though the work [14] supports consumer magazines for PDF file format, none is available for other formats; exploiting Okular's versatility in document reading, it is possible to devise new algorithms which will support complex documents in different file formats.

## References


[1] T. M. Breuel, "High performance document layout analysis," in *Proc. Symp. Document Image Understanding Technology*, 2003.

[2] K. Kise, A. Sato, and M. Iwata, "Segmentation of page images using the area voronoi diagram," *Computer Vision and Image Understanding*, vol. 70, no. 3, pp. 370–382, 1998.

[3] G. Nagy, S. Seth, and M. Viswanathan, "A prototype document image analysis system for technical journals," *Computer*, vol. 25, no. 7, pp. 10–22, 1992.

[4] J. Ha, R. Haralick, and I. Phillips, "Recursive x-y cut using bounding boxes of connected components," in *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, vol. 2. IEEE, 1995, pp. 952–955.

[5] J.-L. Meunier, "Optimized xy-cut for determining a page reading order," in *Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on*. IEEE, 2005, pp. 347–351.

[6] P. Sutheebanjard and W. Premchaiswadi, "A modified recursive xy cut algorithm for solving block ordering problems," in *Computer Engineering and Technology (ICCET), 2010 2nd International Conference on*, vol. 3. IEEE, 2010, pp. V3–307.

[7] W. S. Lovegrove and D. F. Brailsford, "Document analysis of pdf files: methods, results and implications," *Electronic Publishing–Origination, Dissemination and Design*, vol. 8, no. 3, pp. 207–220, 1995.

[8] A. Anjewierden, "Aidas: Incremental logical structure discovery in pdf documents," in *Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on*. IEEE, 2001, pp. 374–378.

[9] F. Rahman and H. Alam, "Conversion of pdf documents into html: A case study of document image analysis," in *Signals, Systems and Computers, 2003. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 1. IEEE, 2003, pp. 87–91.

[10] K. Hadjar, M. Rigamonti, D. Lalanne, and R. Ingold, "Xed: a new tool for extracting hidden structures from electronic documents," in *Document Image Analysis for Libraries, 2004. Proceedings. First International Workshop on*. IEEE, 2004, pp. 212–224.

[11] H. Chao and J. Fan, "Layout and content extraction for pdf documents," in *Document Analysis Systems VI*. Springer, 2004, pp. 213–224.

[12] H. Déjean and J.-L. Meunier, "A system for converting pdf documents into structured xml format," in *Document Analysis Systems VII*. Springer, 2006, pp. 129–140.

[13] T. O. developers, "Okular - more than a reader. [online]. available: http://okular.kde.org/," 2012.

[14] J. Fan, "Text segmentation of consumer magazines in pdf format," in *Document Analysis and Recognition (ICDAR), 2011 International Conference on*. IEEE, 2011, pp. 794–798.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# ANFIS based Opportunistic power control for cognitive radio in spectrum sharing

Joyraj chakraborty* , J. V. K. C Varma and Maria Erman
Blekinge Institute of Technology, Sweden
* E-mail: engr.joyraj@outlook.com

*Abstract*— **Cognitive radio is an intelligent technology that helps in resolving the issue of spectrum scarcity. In a spectrum sharing network, where secondary user can communicate simultaneously along with the primary user in the same frequency band, one of the challenges is to obtain balance between two conflicting goals that are to minimize the interference to the primary users and to improve the performance of the secondary user. In our paper we have considered a primary link and a secondary link (cognitive link) in a fading channel. To improve the performance of the secondary user by maintaining the Quality of Service (Qos) to the primary user, we considered varying the transmit power of the cognitive user. For this we proposed ANFIS based opportunistic power control strategy with primary user's SNR and primary user's interference channel gain as inputs. By using fuzzy inference system, Qos of primary user is adhered and there is no need of complex feedback channel from primary receiver. The simulation results of the proposed strategy shows better performance than the one without power control.**

**Keywords**— ***ANFIS, Fuzzy inference system, spectrum sharing***

## INTRODUCTION

COGNITIVE RADIO is a technology which helps in the efficient utilization of the spectrum. Spectrum is a limited resource, hence the use of spectrum is regulated by government agencies like the Post and Telecom service ( PTS) in Sweden, Federal Communications Commission (FCC) in United states, Telecom Regulatory Authority of India(TRAI) in India, Bangladesh telecommunications regulatory commission (BTRC) in Bangladesh. It has been found that most of the time the spectrum is underutilized even in highly populated urban areas. There are two types of users who use the spectrum. They are the Primary users (licensed) and the secondary users (Unlicensed). The Primary users (licensed) are those that have a license to use the spectrum. The Secondary users (Unlicensed) are those who don't have a license to use the spectrum. When the assigned spectrum is not completely or if only partially utilized by the primary user, the unutilized spectrum is referred to as a spectrum hole or a white space. Cognitive radio identifies these spectrum holes and assigns them to the secondary users without causing any interference to the primary users. The concept of Cognitive radio was first introduced by Joseph Mitola III and Gerald Q. Maguire Jr. from Royal Institute of Technology, Sweden in 1999 [1],[2].

Cognitive radio comprises of spectrum sensing, spectrum management and spectrum mobility. The process of sensing the spectrum, identifying the primary users and spectrum holes without causing interference to primary users is called spectrum sensing. The process of selecting spectrum bands which is appropriate to perform communication by the cognitive user is called spectrum management. It depends on the cognitive user's requirement and the quality of service to the primary users. The process of exchanging the operating frequencies by the cognitive users for better communication is called spectrum mobility.

For Cognitive radio users, optimal power control in spectrum sharing is one of the most important research issues. Opportunistic spectrum access and spectrum sharing are two types of mechanisms in cognitive radio networks. In opportunistic spectrum access either the primary user (PU) or the secondary user (SU) can use the spectrum. When the PU is idle, the SU will be able to use the spectrum. When the PU needs to use the spectrum again, SU has to vacate the spectrum. In Spectrum sharing both PU and SU can access the spectrum simultaneously as long as there is no interference to PU's Quality of service (QoS). The operation of the SU depends on the peak transmit power constraint and an average interference constraint at the primary receiver [4]. It is important to balance the interference to the PU and improve the performance of the SU. Power control is one of the constraints to improve the performance of the secondary users.

In this paper we proposed that Adaptive neuro fuzzy inference system (ANFIS) can be efficiently used for optimal power control in cognitive radio network by using the ratio of primary user's signal to noise ratio (SNR) to the threshold value and the ratio of primary user's interference channel gain to its maximum value as two inputs. There by improving the performance of the secondary user without disturbing the Quality of service (Qos) of the primary user.

## I. SURVEY OF RELATED WORKS

In [5], the authors have considered a spectrum sharing network with a pair of primary users and secondary users in a fading channel. Using three input parameters: The PU's SNR, the PU's interference channel gain and the relative distance between PU's link and CR's link, a fuzzy based opportunistic power control strategy using mamdani fuzzy control has been proposed. For mamdani based FIS, input/output membership functions and fuzzy rules has to be formulated by using human expert knowledge and also a feedback is needed from the primary receiver.

Optimal power control under four different power constraints are derived via convex optimization by considering interference and transmission power constraints in [6]. The peak power constraint at the secondary transmitter is considered to characterize the power adaption strategies that

maximize the SNR and capacity of the secondary user in [4]. In [9], a cognitive user can adapt transmit power opportunistically to achieve maximum transmission rate without affecting the outage probability of the primary user. ANFIS is used for copper grade prediction in Sarcheshmeh porphyry copper system in [7] and for predicting data rate of a particular radio configuration in [10]. A FIS and ANFIS based downlink power control schemes are proposed for fixed broadband wireless access system in [8].

## II. Overview of Cognitive radio

"Necessity is the mother of invention." Scarcity of Spectrum has led to the invention of new technology called Cognitive radio, which can be used for effective utilization of spectrum. Cognitive radio is a technology used by the unlicensed users to use underutilized spectrum without causing any interference to the licensed users.

According to Simon Haykin [3] , the definition of Cognitive radio is "Cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment (i.e., outside world), and uses the methodology of understanding-by-building to learn from the environment and adapt its internal states to statistical variations in the incoming RF stimuli by making corresponding changes in certain operating parameters (e.g., transmit-power, carrier-frequency and modulation strategy) in real-time, with two primary objectives in mind:

- Highly reliable communications whenever and wherever needed;
- Efficient utilizations of the radio spectrum."

Most important features of cognitive radio is awareness, intelligence, learning, adaptability, reliability, reconfigurability. Different technologies like Software defined radio, digital signal processing, networking, machine learning and computer software & hardware helped to bring these features together in the form of Cognitive radio.

There are three fundamental tasks in the operation of Cognitive radio as shown in Fig. 1. They are

1. Radio scene analysis.
2. Channel state estimation and predictive modeling.
3. Transmit power control and spectrum management.

The first and second tasks are performed at receiver's end however the third task is performed at the transmitter end.

In Radio scene analysis there are two important actions. One is to estimate the interference temperature. Second is to identify spectrum holes. These actions are performed at the receiving end and the information is transmitted back to the transmitter through a feedback channel. This information is necessary for the transmitter in order to perform the third task. Channel state estimation and predictive modeling comprises of estimating the channel state information and to calculate the channel capacity of the cognitive link which is used by the transmitter. Transmit power control and spectrum management is performed at the transmitter end. It comprises of selecting the best transmit power levels of the unlicensed users which maximizes the data transmission rates without exceeding interference temperature. Spectrum management is necessary for effective utilization of the spectrum.

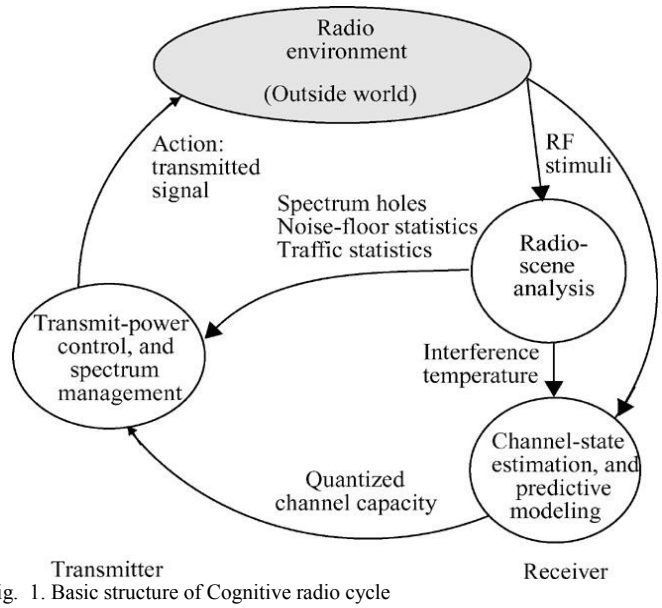


Fig. 1. Basic structure of Cognitive radio cycle

## III. Overview of ANFIS

Adaptive neuro fuzzy inference system(ANFIS) is a type of fuzzy inference system (FIS) which formulates the mapping of inputs to output. It uses both fuzzy logic(FL) and artificial neural networks(ANN) in the process of mapping the inputs to output. In FIS the most difficult part is to obtain membership functions, distribution of membership functions and setting fuzzy rules. These parameters are obtained by using trial and error method. ANFIS uses neural networks to adjust these parameters. The ANN part in ANFIS helps in reducing the error and optimizing the parameters. FL deals with uncertainty very well and known for structured knowledge representation. ANN has the learning capability. ANFIS has the advantages of both FL and ANN. Hence it has become an important step of research in the fields of automatic control, data classification, decision analysis, expert systems and computer vision where FIS has been successfully used. The main objective of ANFIS is to identify the near optimal membership functions and other parameters of the equivalent fuzzy inference system by applying a hybrid learning algorithm using input-output data sets and then to achieve a desired input-output mapping.

ANFIS has a layered architecture just like neural networks. ANFIS consists of five layers. The nodes in these layers can be adaptive or fixed. For identification, the adaptive nodes are represented by squares and fixed nodes are represented by circles in the architecture. The general architecture of ANFIS is shown in Fig. 2. For description of the architecture, we consider first order sugeno with two inputs x and y. A Takagi-sugeno based ANFIS has rules of the form of

If x is A_i and y is B_i then :

$$f_i = p_i\, x + q_i\, y + r_i \qquad (1)$$
$$where\ \ i = 1,2\ \ldots$$

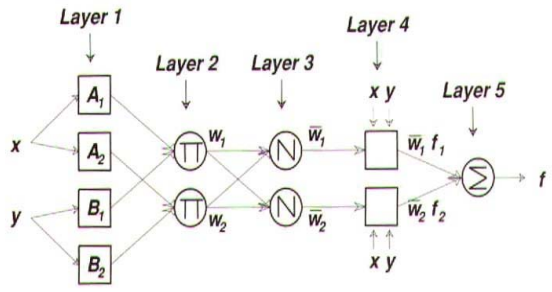



Fig. 2. Generalized structure of ANFIS.

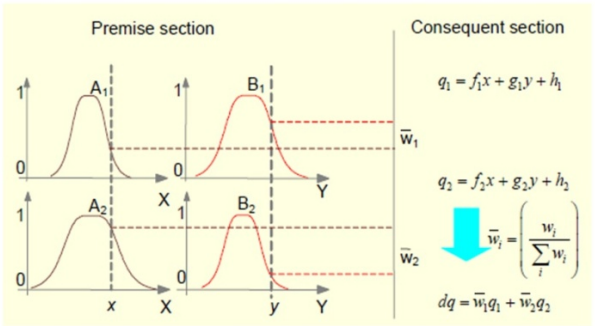



Fig. 3. Calculation of output from inputs

Layer 1 : In this layer the nodes are adaptive. The membership values of the inputs are calculated in this layer. The output of the nodes is a membership grade of the inputs.

$$\text{Output } O_{1,i} \text{ for nodes } i = 1,2 \quad O_{1,i} = \mu_{A_i}(x)$$

$$\text{Output } O_{2,i} \text{ for nodes } i = 1,2 \quad O_{2,i} = \mu_{B_i}(y)$$

x and y are the inputs to the node i. $A_i$ and $B_i$ are the linguistic labels like small, medium, large etc. $\mu_{A_i}(x)$ and $\mu_{B_i}(y)$ can be any membership function like bell shaped, triangular or trapezoidal shaped etc. The number of nodes represents the number of fuzzy sets into which each input is quantified. Usually all the membership functions are in the interval of [0,1]. $A_i$ and $B_i$ are also called as antecedent parameters or premise parameters which are determined and adaptively modified during the training process of ANFIS.

In this paper we used trapezoidal membership functions which are expressed as

$$\mu_{A_i}(x) = \begin{cases} 0 & x < a \text{ or } x > b \\ 1 - \dfrac{x-a}{c-\frac{d}{2}-a} & (2)\ a < x < c - \dfrac{d}{2} \\ 1 & c - \dfrac{d}{2} < x < c + \dfrac{d}{2} \\ \dfrac{b-x}{b-\left(c+\frac{d}{2}\right)} & (3)\ c + \dfrac{d}{2} < x < b \end{cases}$$

Where a and b are left and right bounds. c is the center of the trapezoid and d is the top width of the trapezoid.

Layer 2: In this layer the nodes are fixed. In this layer the firing strength of each rule is calculated and it represents each node output. T-norm operators are used like min, product, fuzzy AND etc.

$$W_I = M_{A_i}(X).M_{B_i}(Y) \quad \text{FOR I} = 1,2 \qquad (4)$$

Layer 3: The nodes are fixed in this layer. In this layer all the rules or firing strengths are normalized. The ratio of a particular rule to that of sum of all the rules. The output of the node is usually called as normalized firing strength and is fed to layer 4. Output of layer 3 is:

$$\overline{w_I} = \frac{w_i}{\sum w_i} \quad \text{where i} = 1,2 \ldots \qquad (5)$$

Layer 4: The nodes are adaptive in this layer. The output or the consequent parameters are determined in this layer. The function of each node is a combination of the output of the layer 3 and a simple linear equation (sugeno rule). In this layer, the contribution of each rule to the overall output is calculated. Output of layer 4 is:

$$\overline{w_I}.f_i = \overline{w_I}(p_i x + q_i y + r_i) \qquad (6)$$

Layer 5: This layer is the output layer. There is only a single node and it is fixed. The function of this layer is the summation of all the inputs or incoming signals (outputs from layer 4). In Fig. 3 it is shown as dq. Output of layer 5 is:

$$\sum_i \overline{w_I}.f_i = \frac{\sum_i w_i . f_i}{\sum_i w_i} \qquad (7)$$

From layer 1 to layer 4 represents the "if -then" rules of takagi sugeno fuzzy model and also determines the number of fuzzy rules. In layer 1 the input membership functions and antecedent parameters are determined and in layer 4, the consequent parameters are determined. If a zero order sugeno model has to be implemented, all the steps will be similar expect the output in layer 4 where p=0 and q=0. The performance of the system is evaluated using root mean squares error (RMSE) and is defined as

$$\text{RMSE} = \frac{1}{n}\sqrt{\sum_{k=1}^{n}(y_k - o_k)^2} \qquad (8)$$

Where $y_k$ is the desired output and $o_k$ is the actual output obtained. n is the number of training samples.

## IV. System Model

In this paper we have considered a single primary link (a transmitter and receiver) and a single secondary or cognitive link (transmitter and receiver) in a fading channel as shown in Fig. 3. $P_p$ and $P_c$ are the transmitted powers of primary and secondary(cognitive) transmitters. $g_{pp}$ and $g_{cc}$ are the channel gains between the primary link and the secondary link. $g_{cp}$ and $g_{pc}$ are the interference channel gains of secondary transmitter to primary receiver and primary transmitter to secondary receiver. All the channel gains are assumed to be random variables acting independently with continuous probability density function. $N_0$ is the variance of the additive white Gaussian Noise that affects both the primary receiver and the secondary receiver independently in the same way. In this spectrum sharing network, primary link has high priority to transmit data then secondary link. As long as the Qos of the primary user is maintained, the secondary user will be able to use the spectrum.

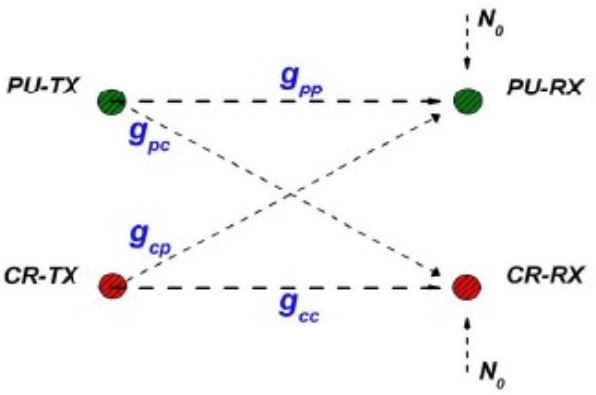



Fig. 4. System model

We have considered the propagation environment without path loss. In this scenario we have considered the ratio of primary user's signal to noise ratio (SNR) to a threshold value as one input and the ratio of $g_{cp}$ to its maximum value as the second input to the ANFIS. The output we considered is the ratio of $P_c$ to its maximum value.

The primary user's Signal to Noise Ratio without the presence of Cognitive user is defined in [9] as

$$\text{SNR} = \alpha = \frac{P_p\, g_{pp}}{N_o} \tag{9}$$

primary user's Signal to interference noise ratio with the presence of cognitive user in the spectrum sharing network is given in [9] as

$$\text{SINR} = \beta(k) = \frac{P_p\, g_{pp}}{k.P_{cmax}\, g_{cp} + N_o} \tag{10}$$

$P_{cmax}$ is the peak power transmitted by the cognitive user and k is the instantaneous power control parameter or power scale ratio. The value of k varies in between 0 to 1 whose value will be determined by the power control strategy [9].

In order to maintain the Qos a threshold level $\beta_t$ is considered which is given in [9] as $2^{R_p} - 1$ . $R_p$ is the data rate of the primary link. Primary user's SINR should be greater than the threshold value during spectrum sharing. The Cognitive user varies its power to keep the primary user's SINR greater than the threshold value.

When the SNR of the primary link falls below the threshold level the primary link is already in outage. In this case the transmission from the cognitive user does not have any negative affect on the primary user, no matter how much transmit power it uses. Hence cognitive user can transmit with its peak power.

In case when the SNR of the primary link is near to the threshold value then the primary user in not in outage and sensitive to interference from the Cognitive user. In this case the cognitive user transmits with fraction of its peak power to maintain Qos to Primary link, so that the primary user can still transmit $R_p$ with $P_p$.

In case when the SNR of the primary link is far above the threshold value, the primary user's SINR will be greater than the threshold value even if the cognitive user transmits with its peak power. As the Qos is assured regardless of cognitive user's transmit power, The Cognitive user transmits with peak power.

When the primary user's interference channel gain ($g_{cp}$) is low, the Primary user's received interference intensity from the Cognitive link is low. In this case the Cognitive user can transmit with peak power. When the primary user's interference channel gain ($g_{cp}$) is high, the primary user's received interference intensity from the Cognitive link is high and the Cognitive user should transmit with low power.

## V. Implementation of ANFIS

The training of ANFIS is carried either by back propagation algorithm or Hybrid algorithm. In this paper we have used hybrid algorithm for training purpose. A hybrid algorithm is an algorithm which uses both back propagation algorithm and least mean squares algorithm to determine the input (antecedent) and the output (consequent) parameters. Usually the consequent parameters are updated first by using least mean square algorithm and the antecedent parameters are updated later by back propagating the errors that still exist using gradient descent algorithm. Since two different types of algorithms are used to reduce the error it is called as hybrid algorithm. The training data is manipulated from the concepts discussed in system model and used for training purpose. We used 45 sets of noisy data and trained for 80 epochs. The error is reduced to 0.082 after 40 epochs as shown in Fig. 5. We performed simulations using MATLAB 2010a. After training, the obtained structure of ANFIS and the membership functions are shown in Fig. 4.

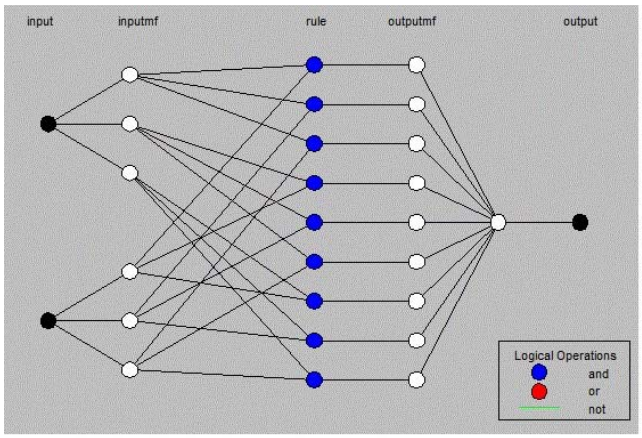
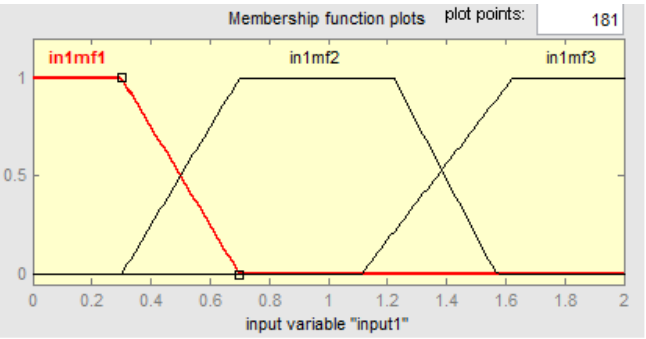
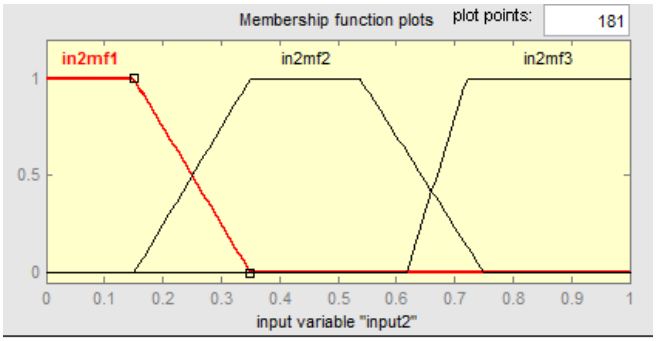



Fig. 5. (a) structure of ANFIS.




(b) Membership function of input 1, $\left(\frac{\alpha}{\beta_t}\right)$.




(c) Memberhsip function of input 2, $\left(\frac{g_{cp}}{g_{cp\,max}}\right)$

The inputs that we considered for ANFIS are the ratio of primary user's SNR to threshold value $\left(\frac{\alpha}{\beta_t}\right)$ and the ratio of primary user's interference channel gain to its maximum value $\left(\frac{g_{cp}}{g_{cp\,max}}\right)$ as inputs 1 & 2 respectively.

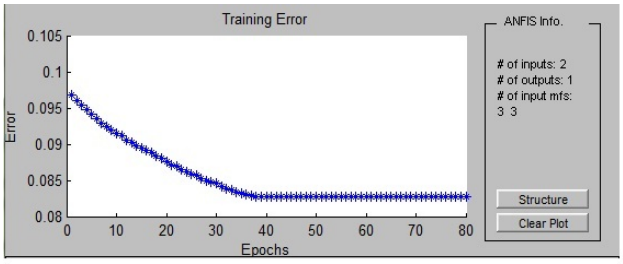



Fig. 6. Error plot

## VI. RESULTS

The surface graph of the two inputs is shown in Fig. 5. From the surface graph we can tell how the inputs affect the output.

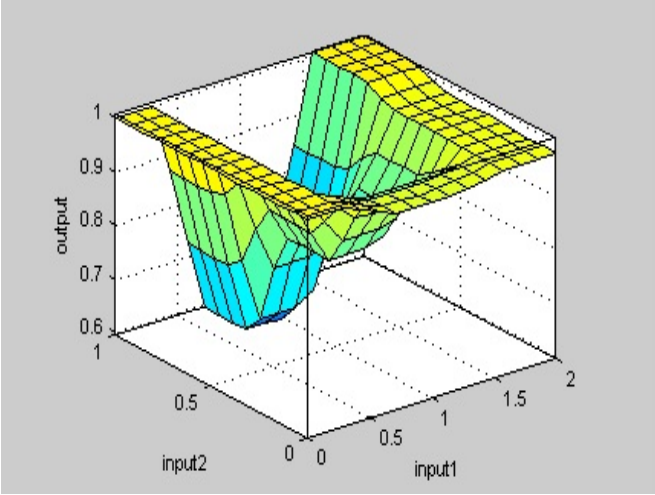



Fig. 7. Surface graph

We have applied FIS based power control strategy for improving the performance of the secondary user. The performance of a system can be observed by plotting the Bit Error rate versus the ratio of energy per bit to Noise power using BPSK modulation scheme. We did the comparison between system without power control and system with ANFIS based power control in the propagation environment without path loss in a fading channel. From the graph we can say that the bit error rate is considerably reduced (20 dB) for the one with ANFIS based power control. From this we can say that the secondary user's performance is improved while maintaining the QoS of the primary link.

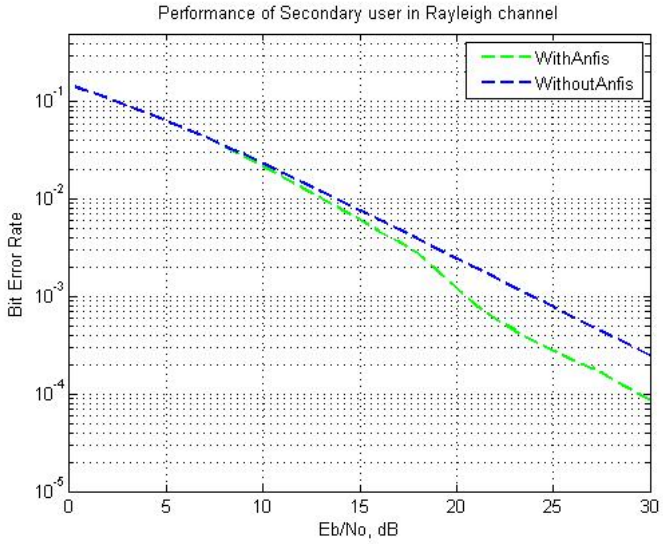



Fig. 8. Performance of secondary user

## VII. CONCLUSION

For the huge demand for mobile & broadband applications, most important demand is that pictures and video are very efficient in describing a situation. Modern cell phones have broadband internet access, and public safety personnel may start to rely on it. Therefore, broadband internet access is a requirement for the next generation public safety communication system, and problem is that spectrum in the public safety bands is scarce. Public safety communication has such a high peak-to-average ratio that it is uneconomical to reserve the needed spectrum in the conventional way. Cognitive radio could be a way out of this problem. However, cognitive radio is not allowed to interfere with the primary user communication in any condition. So in our paper we proposed how power management could be done without interfering primary user and to maintain a QoS.

In a spectrum sharing network, improving the performance of the secondary user and minimizing the interference to the primary link are two conflicting goals. So we proposed ANFIS based power control strategy to improve the performance of the secondary user by maintaining the QoS to the primary link. Simulation results show that bit error rate of the proposed strategy is less than the one without power control. Thus we conclude that ANFIS can be efficiently used for power control in cognitive radio which helps in improving the performance of the secondary user. Using FIS, QoS of the primary user is adhered and there is no need of complex feedback channel. Depending on user's requirement or model design and human expert knowledge, one can use either ANFIS or FIS. However using ANFIS reduces the efforts, and it can learn and adapt from the environment (data sets).

For Future work, the proposed strategy can be extended to the network with multiple primary users and multiple cognitive users so that each cognitive user has a choice to select the best primary link that it can coexist with, in order to improve the efficient utilization of spectrum.

## REFERENCES


[1] J. Mitola III and G. Q. Maguire, "Cognitive radio: making software radios more personal," IEEE J. personal communications, vol. 6, pp. 13-18, Aug. 1999.

[2] J. Mitola III , "Cognitive radio- An integrated agent architecture for software defined radio " Ph.D. dissertation, KTH, Stockholm, 2000.

[3] S. Haykin, "Cognitive radio: brain empowered wireless communications," IEEE J. Selected areas in Commun., vol. 23, pp. 201-220, Feb. 2005.

[4] S. Srinivasa and S. A. Jafer, "Soft sensing and optimal power control for cognitive radio" IEEE J. transactions on wireless commun., vol. 9, pp. 3638-3649, 2010.

[5] A. Dey, S. Biswas and S. Panda, "A new fuzzy rule based power management scheme for spectrum sharing in cognitive radio," in Communication and Industrial Application International Conf., kolkatta, West Bengal, 2011, pp. 1-4.

[6] X. P. Gao, W. J. Lei and X. Z. Xie, "Optimal power control for secondary system based on spectrum sharing," in Communications and Information Technologies International symposium, Hangzhou, 2011, pp. 84-88.

[7] P. Tahmasebi and A. Hezarkhani. (2010). Application of Adaptive Neuro-Fuzzy Inference system for Grade Estimation; Case Study, Sarcheshmeh Porphyry Copper Deposit, Kerman, Iran. Aust. J. Basic & Appl. Sci., [online]. 4(3), pp. 408-420. Available: http://stanford.academia.edu/PejmanTahmasebi/Papers/178689/Application_of_Adaptive_Neuro-Fuzzy_Inference_System_for_Grade_Estimation_Case_Study_Sarcheshmeh_Porphyry_Copper_Deposit_Kerman_Iran.

[8] Z. S. Lee, "ANFIS based Power control for Fixed Broadband Wireless access systems," Ph.D. dissertation, Dept. Comm. Eng., National Central Univ., Taiwan, 2008.

[9] Y. Chen, G. Yu, Z. Zhang, H. Chen and P. Qiu, "On Cognitive radio networks with opportunistic power control strategies in fading channel," IEEE Trans. on Wireless Commun., vol. 7, no. 7, pp. 2752-2761, Jul. 2008.

[10] S. Hiremath and S. K. Patra, "Transmission rate prediction for Cognitive Radio using Adaptive Neural Fuzzy Inference System," in International conf. on Industrial and Information Systems, 2010, pp. 92-97.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Robust Facial Expression Recognition Based on Median Ternary Pattern (MTP)

Farhan Bashar*, Asif Khan, Faisal Ahmed, and Md. Hasanul Kabir
Department of Computer Science and Engineering, Islamic University of Technology,
Gazipur–1704, Bangladesh
*shadhon@iut-dhaka.edu

***Abstract*—Facial expression is a strong medium to express one's feelings and emotions. Accurate detection of facial expression can convey a lot of information about a person's mood. The major barriers to accurate recognition are the presence of noise, illumination variation and occlusion. This paper proposes an effective appearance-based facial feature descriptor constructed with a new local texture pattern, namely the median ternary pattern (MTP) for facial expression recognition. The proposed MTP operator encodes the texture information of a local neighborhood by thresholding against a local median gray-scale value and quantizing the intensity values of the neighborhood around each pixel into three different levels. The MTP codes generated for an image or image patch is then used as the feature representation of the facial expression image. The effectiveness of the proposed method has been evaluated using images from the Cohn-Kanade (CK) Expression Database. Classification was done using Support Vector Machine (SVM). Experimental results show increased accuracy in recognition rates when using the proposed approach in comparison to other popular gray-scale based methods.**

***Index Terms*—Median ternary pattern, facial expression recognition, feature descriptor, support vector machine.**

## I. Introduction

In recent years there has been a growing interest for human behavioural analysis and human-computer interaction. Facial expression is the most powerful nature of human beings to communicate their emotions and intentions [1]. Automatic facial expression recognition is an interesting and challenging task which has attracted much attention over the last two decades due to its potential applications in human computer interactions and data driven animation. Also capturing image using various facial expression has become a common trend in digital video cameras [2]. Although much work has already been done on automatic facial expression recognition system [2], [3], but the recognition accuracy in uncontrolled environment i.e., noise, illumination variation, pose, aging,alignment, and occlusions is still a great challenge [1]. As a result, an efficient face descriptor is desired to increase the robustness of facial expression recognition system against these factors.

At beginning of the research on facial representation, face features are extracted based the geometric relationships i.e., distance, positions, and angles between different facial components (nose, mouth, and eyes). One of the most popular feature-based feature extraction method is Facial action coding system (FACS) [4] which represents facial expression with the help of a set of action units (AU). After that, using the geometric position of different facial components, manually selected fiducial points were used as facial feature representation [5]. Previously, this geometric based approach using fiducial point data gained much popularity for its high recognition rate [6], [7] but as geometric based approach depends on the accurate detection of facial components, in uncontrolled environment the performance of this approach deteriorates.

Currently, appearance-based methods have attracted much research attention and are highly used for feature extraction because it does not rely on the facial components, rather it uses a filter or filter bank on the whole image to extract facial features. Principal component analysis (PCA) [8], [9] is one of the most popular appearance based method which utilizes holistic information of the image to generate efficient features. Independent component analysis (ICA) came out with the theme of generating local feature information of face regions [10], [11]. Some other popular local feature-based methods includes Gabor wavelets [12], Local feature analysis (LFA) [13], and Dynamic link architecture (DLA) [14]. In recent times, Local binary patterns (LBP) [15] and its variants are getting attention for their approach of generating facial features from local neighborhood pixels. LBP generates feature representation by thresholding the gray value of the neighbor pixels with respect to gray value of the center pixel of a local region. This LBP feature provide robustness in monotonic illumination changes. But the performance of LBP feature degrades under the presence of random noise and large illumination variation. To address the issue of random noise Median Binary Pattern (MBP) [16] came with the idea of generating binary codes by thresholding with respect to the median of the neighbor cells around the center pixel. By generating comparatively more stable feature, MBP overcomes the problem of random noise. Local Ternary Pattern (LTP) [17] took further step of robustness in uncontrolled environment problem by using one additional discrimination level than LBP. LTP introduces a 3-level coding scheme, which results from thresholding the neighbors with respect to a certain threshold region ($\pm t$) about the center pixel. Although LTP performs much better in presence of uniform illumination variation but still it generates inconsistent feature if the image is corrupted with large noise and random illumination variation.

This paper introduces a robust appearance-based feature descriptor constructed with the proposed Median Ternary Pattern (MTP) that overcomes the limitation of existing local texture patterns. After finding a successful result on human face recognition [18] MTP face descriptor has also showed robustness in the field of facial expression recognition. The proposed method uses the median value of neighbor cells around each pixel and quantizes the value into three different levels (-1,0,1) in order to generate local texture patterns. The aim of the proposed method is to generate a feature vector which is more stable to random noise and lighting variation and generate a consistent feature for both smooth and high-textured regions. The performance of MTP feature vector is

evaluated by Cohn-Kanade (CK) facial expression database [19] and classification is done by using the well-known classifier, Support vector machine (SVM) [20]. The next section gives a brief summary on existing micro-pattern based face descriptors, followed by a description of the proposed method and using it for generating feature representation of facial expression images. Finally, the recognition accuracy of the proposed method and a comparison between the proposed and existing methods is shown.

## II. Median Binary Pattern (MBP)

### A. Local Binary Pattern (LBP)

Local binary pattern (LBP) is a local texture operator and was originally introduced as a gray-scale invariant method for texture analysis [15]. This method was later successfully applied in facial expression recognition. The LBP operator selects a $3 \times 3$ local neighborhood around each pixel and generates an 8-bit binary code by thresholding the intensity values of the neighbor pixels with respect to the center. The binary code of LBP is generated using (1).

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{p-1} s(i_c - i_p), \;\; s(x) = \begin{cases} 1 & x > 0 \\ 0 & otherwise \end{cases} \tag{1}$$

Here, $i_c$ denotes the gray value of the center pixel ($x_c$, $y_c$) and $i_p$ corresponds to the gray value of the local neighbors. Fig. 1 shows the method of LBP code generation.

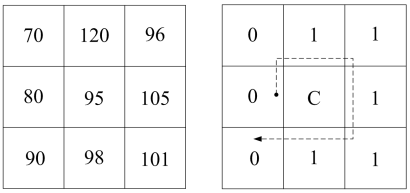



Fig. 1. Illustration of LBP operator (a)Original Image, (b)LBP Code= 00111110

The LBP operator is heavily dependent on the gray-scale value of the center pixel as it thresholds exactly at the center pixel. A change in the value of the center pixel due to noise or illumination variation will generate varying LBP code for similar expression images. This makes LBP very sensitive to noise and illumination variation [17]. To overcome the dependency problem, the Median Binary Pattern (MBP) was proposed [16].

### B. Median Binary Pattern

Median binary pattern (MBP) [16] is an extension of Local Binary Pattern (LBP). To generate the MBP codes, a $3 \times 3$ local neighborhood is defined around each pixel and the median gray-scale value, $M_c$ is computed. The 8-bit binary code of the center pixel is calculated by comparing the intensity of the neighbor cells against the median gray-scale value, $M_c$. The binary code of MBP is generated using (2).

$$MBP(x_c, y_c) = \sum_{p=0}^{p-1} s(i_p) \times 2^p, \;\; s(x) = \begin{cases} 1 & x > M_c \\ 0 & otherwise \end{cases} \tag{2}$$

Here, $M_c$ denotes the median gray value of the local neighbor pixels around the center pixel $(x_c, y_c)$ and $i_p$ corresponds to the gray value of the local neighborhood pixels. The MBP code generation is illustrated in FIg. 2.

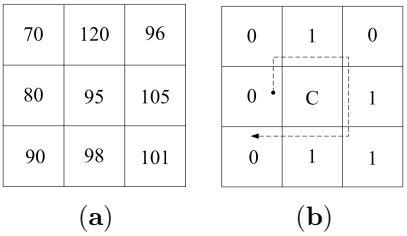



Fig. 2. Illustration of MBP operator (a)Original Image, (b)MBP Code= 00101110

The use of the median gray-scale value for encoding the pixels of the local neighborhood provides more robustness under the presence of random noise but inconsistent code generation problem remains in uniform and near-uniform regions since thresholding is done at exactly the median value. Therefore, a new method is proposed to overcome this problem.

## III. Median Ternary Pattern

A filter which uses local neighborhood median is more robust against the presence of noise in an image [21]. Multi-level gray-scale quantization increases the robustness against illumination variation. Both techniques on their own have proved to be effective, but to some certain extent. However, combing the features of the two methods into a single face descriptor have provided a superior and more accurate facial representation. Our proposed method, median ternary pattern (MTP), combines the advantages of median filter and quantization of gray-scale values into three different levels. As a first step, a $3 \times 3$ local neighborhood is defined around each pixel and the median gray-scale value of the nine pixels in the neighborhood is calculated. Next, those within $\pm t$ about the median M are set to 0, those above $M + t$ are set to $+1$ and those below $M - t$ are set to $-1$. Finally, the MTP code is generated using (3).

$$S_{MTP}(v) = \begin{cases} 1, & v > M_c + t \\ 0, & M_c - t \leq v \leq M_c + t \\ -1, & v < M_c - t \end{cases} \tag{3}$$

Here, $v$ is a neighbor gray level, $M_c$ is the local median, and $t$ is a user specified threshold. The MTP code generation process is shown in Fig. 3(a-b).

The total possible encoded values, or states, that can be generated using the 3-level quantization method is $3^8$. This would require $3^8$ bins when representing the values in a histogram. To reduce the state-space, each MTP code is further split into its corresponding positive and negative parts and are treated as two separate binary patterns called $P_{MTP}$ and $N_{MTP}$. This reduces the feature size from $3^8$ to $2 \times 2^8$. $P_{MTP}$ and $N_{MTP}$ are calculated using (4) and (5), respectively.

$$P_{MTP} = \sum_{p=0}^{7} S_P(S_{MTP}(i_p)) \times 2^p, \;\; S_P(v) = \begin{cases} 1, & v = 1 \\ 0, & otherwise \end{cases} \tag{4}$$

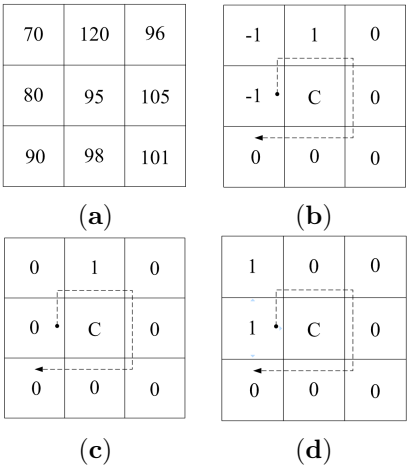



Fig. 3. Illustration of MTP operator (a)Original Image, (b)MTP Code= (-1)(-1)100000 (c)$P_{MTP}$ code= 00100000 (d)$N_{MTP}$ code= 11000000

$$N_{MTP} = \sum_{p=0}^{7} S_N(S_{MTP}(i_p)) \times 2^p, \;\; S_N(v) = \begin{cases} 1, & v = 1 \\ 0, & otherwise \end{cases} \tag{5}$$

Here, $M_c$ is the median of the local neighborhood and $t$ is a margin threshold. $P_{MTP}$ and $N_{MTP}$ are codes of positive and negative binary patterns respectively. The generations of $P_{MTP}$ and $N_{MTP}$ code from the original MTP code is shown in fig. 3(c) and 3(d).

## IV. Face Representation Using MTP

Applying MTP operator on an whole face image generates two encoded images, $P_{MTP}$ and $N_{MTP}$. The histograms generated from these two histograms are then concatenated to create final MTP histogram. This combined histogram is then used as face representation for that face image to perform facial expression recognition. The generation of MTP histogram from an face image is illustrated step by step in Fig. 4.

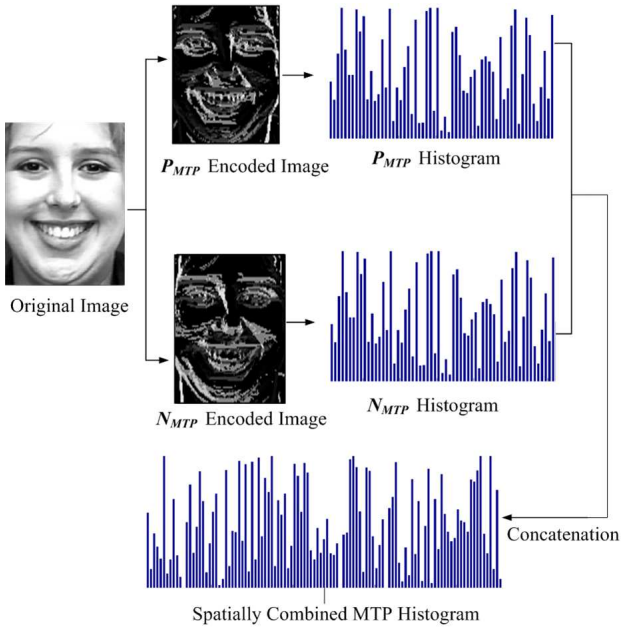



Fig. 4. Illustration of the MTP feature vector generation process.

The histogram generated from the whole encoded image does not contain any location information of the MTP micro-patterns. It only expresses their occurrences. However, presence of location information of a facial image helps to describe the face more accurately and thus the feature generated provides better recognition rate. Therefore, taking some degree of locality in our feature vector, we extend the MTP histogram by concatenating the facial image into a number of regions and individual MTP histograms are generated from those regions. Finally, the extended MTP histogram is generated by concatenating all those local histograms. This extended histogram is then used as final facial feature vector. This process is illustrated in Fig. 5.

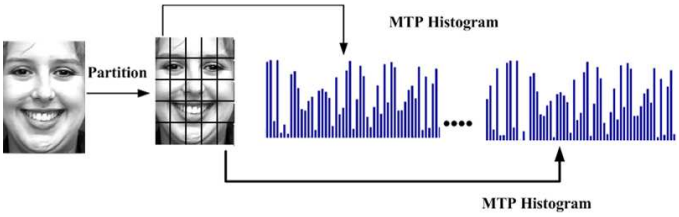



Fig. 5. Localization of an image into several parts and concatenating all histograms together.

## V. Facial Expression Recognition

The facial expression of an face can be found by classify the feature vector of that image to a particular group of expression. We have worked with a de facto database which consists facial images of various expression. At present, there are several classification methods for finding the actual group of an input feature vector i.e, template matching, support vector machine (SVM) etc. We have used both of them to classify the features and found that SVM performs more accurately than the other method. Thus we used SVM with Radial-basis function (RBF) kernel to classify MTP feature vectors representing facial images. SVM makes binary decisions and multi-class classification can be achieved by adopting the one against-rest or several two-class (happiness-anger, sadness-fear, fear-anger, joy-anger, etc.) problems [20]. In our experiment, we trained a binary classifier which uses one-against rest technique to discriminate one expression from others and 10-fold cross-validation approach is performed with SVM using RBF kernels for finding accuracy of the facial expression.

## VI. Experimental Results

There can be different levels of facial expression for a person. As a result, most of the expression recognition systems uses a set of prototypic emotional expression which includes anger, disgust, fear, joy, sadness and surprise [1]. In this work, we have tried to recognize this 6-class expression. Also, the proposed method was carried out after extending the 6-class expression set to 7-class by adding neutral expression to the imageset.

The performance of our proposed method is evaluated by using the well-known image dataset, namely Cohn-Kanade (CK) facial expression database [19]. The CK image database consists of 100 university students within 18 to 30 years of age having different expressions. Separating their images by six expression we get 1224 facial images, we refer as 6-expression dataset and after including additional 408 neutral images, we get 1632 images which we refer as 7-expression

dataset. Sample images of 7-expression dataset is shown in Fig. 6. The facial images were cropped based on the ground truth positions of the two eyes and mouth and normalized to $150 \times 110$ pixels. The classification accuracy was computed for face images divided into $3 \times 3$, $5 \times 5$, $7 \times 6$ sub-regions. The existing methods also used a dimension of $7 \times 6$ for sub-dividing the image, so the comparison is kept upto $7 \times 6$ subdivision, to enable fair comparison between the existing and the proposed method. Feature size of an image depends on the sub-division, for a face without any subdivision, the feature size is 512 and for $7 \times 6$ subdivision, it becomes $512 \times 7 \times 6$ which is 21504. The threshold value $t$ was set to 10 as it gives the best recognition accuracy. The performance of MTP descriptor was also compared to LBP [1], MBP [16], LTP [17], PCA [8] methods.


Fig. 6. Sample images of the seven prototypic expressions.

Table I and II shows the recognition rate of different feature representation methods using 6-expression and 7-expression dataset respectively. All the methods, including the proposed one, were applied on the same imageset. For LTP and MTP, the threshold value is set to 30. In both table, it shows that MTP over-performs other well-known methods. Also for $7 \times 6$ sub-regions, the recognition rate is highest because more number of regions can hold much local information. But large size of feature vector sometimes reduces the recognition time. We choose $7 \times 6$ sub-regions to create the confusion matrix (CM) of 6-class and 7-class recognition shown in Table III and IV respectively. Both the table shows MTP descriptor holds superior performance in recognizing facial expressions.

TABLE I
RECOGNITION RATE (%) USING THE 6-CLASS DATASET

| Operator | No of Region | | |
|---|---|---|---|
| | $3 \times 3$ | $5 \times 5$ | $7 \times 6$ |
| **MTP** | **95.2** | **97.46** | **98.1** |
| MBP | 80.2 | 87.3 | 93.1 |
| LTP | 91.3 | 92.3 | 94.6 |
| LBP | 79.1 | 89.7 | 90.1 |
| PCA | 75.3 | 85.1 | 88.9 |

TABLE II
RECOGNITION RATE (%) USING THE 7-CLASS DATASET

| Operator | No of Region | | |
|---|---|---|---|
| | $3 \times 3$ | $5 \times 5$ | $7 \times 6$ |
| **MTP** | **89.1** | **92.4** | **94.2** |
| MBP | 77.2 | 85.3 | 90.1 |
| LTP | 87.3 | 89.3 | 91.6 |
| LBP | 75.1 | 84.7 | 86.1 |
| PCA | 70.3 | 81.1 | 83.9 |

It can be easily observed from Table I-IV that facial feature representation based on MTP is more robust than other existing local texture patterns. The superiority of MTP is due to the utilization of local medians with the threshold in order to provide consistency in both smooth and high-textured facial images, even under the presence of noise and illumination variation.

TABLE III
CONFUSION MATRIX FOR THE 6-CLASS RECOGNITION USING MTP FEATURE REPRESENTATION

| | Anger (%) | Disgust (%) | Fear (%) | Joy (%) | Sad (%) | Surprise (%) |
|---|---|---|---|---|---|---|
| Anger | **99.4** | 0 | 0.6 | 0 | 0 | 0 |
| Disgust | 0 | **98.2** | 0 | 0 | 1.8 | 0 |
| Fear | 1.5 | 0 | **97.5** | 0 | 0.8 | 0.2 |
| Joy | 0 | 0 | 0 | **98.2** | 0 | 1.8 |
| Sad | 0 | 0 | 0 | 1.5 | **98.3** | 1.2 |
| Surprise | 0.4 | 0 | 1.6 | 0.4 | 0 | **97.6** |

TABLE IV
CONFUSION MATRIX FOR THE 7-CLASS RECOGNITION USING MTP FEATURE REPRESENTATION

| | Anger (%) | Disgust (%) | Fear (%) | Joy (%) | Sad (%) | Surprise (%) | Neutral (%) |
|---|---|---|---|---|---|---|---|
| Anger | **94.5** | 0 | 3.5 | 0 | 0.8 | 1.2 | 0 |
| Disgust | 0 | **93.2** | 0.8 | 0 | 3.3 | 0 | 3.2 |
| Fear | 1.5 | 0 | **95.5** | 0 | 1.8 | 1.2 | 0 |
| Joy | 0 | 0 | 0.5 | **93.2** | 0 | 3.8 | 2.5 |
| Sad | 0 | 2.5 | 0 | 1.5 | **93.3** | 1.2 | 3.0 |
| Surprise | 1.4 | 0 | 1.6 | 0.8 | 0 | **92.6** | 3.6 |
| Neutral | 2.4 | 0 | 1.6 | 1.4 | 1.6 | 1.2 | **91.8** |

## VII. CONCLUSION

In this paper we proposed a new approach for accurate facial expression recognition, using a face descriptor which uses a new local texture filter called Median Ternary Pattern(MTP). The MTP codes are used to generate an effective representation of the expression images, also known as the feature vector of the image. MTP codes are computed by defining a local neighborhood around each pixel of the image, calculating the median gray-scale value of the neighborhood pixels and quantization of the gray-scale value of each pixel based on the computed median and a certain threshold. The classification of the images were done using Support Vector Machine. Experiments carried out on the Cohn-Kanade Expression database show that, the proposed method produces better accuracy in presence of major uncontrolled factors, including illumination variation, noise, alignment and occlusion.

## REFERENCES

[1] C. Shan, S. Gong, and P. W. McOwan,"Facial Expression Recognition based on Local Binary Patterns: A Comprehensive Study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803-816, May 2009.

[2] Y. Tian, L. Brown, A. Hampapur, S. Pankanti, A. Senior, and R. Bolle, "Real World Real-Time Automatic Recognition of Facial Expressions," in Proc. *IEEE Workshop on Performance Evaluation of Tracking and Surveillance*, USA, pp. 9-16, 2003.

[3] Y. Tian, "Evaluation of Face Resolution for Expression Analysis," in Proc. *IEEE Workshop on Face Processing in Video*, 2004.

[4] P. Ekman and W. Friesen, "Facial Action Coding System: A Technique for Measurement of Facial Movement," *Consulting Psychologists Press*, 1978.

[5] Z. Zhang, "Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiment with a Multi-Layer Perceptron," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 13, issue 6, pp. 893911, 1999.

[6] A. Samal and P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey," *Pattern Recognition*, vol. 25, pp. 6577, 1992.

[7] M. Valstar and M. Pantic, "Fully Automatic Facial Action Unit Detection and Temporal Analysis," *IEEE CVPR Workshop*, p. 149, June 2006.

[8] C. Padgett and G. Cottrell, "Representation Face Images for Emotion Classification," *Advances in Neural Information Processing Systems*, vol. 9, 1997.

[9] M. A. Turk and A. P. Pentland, "Face Recognition using Eigenfaces," in Proc. *International Conferencence on Computer Vision and Pattern Recognition*, pp. 586-591, 1991.
[10] M.S. Bartlett, J.R. Movellan, and T.J. Sejnowski, "Face Recognition by Independent Component Analysis," *IEEE Transactoin on Neural Networks*, vol. 13, no. 6, pp. 14501464, 2002.
[11] C.C. Fa and F.Y. Shin, "Recognizing Facial Action Units using Independent Component Analysis and Support Vector Machine," *Pattern Recognition*, vol. 39, no. 9, pp. 17951798, 2006.
[12] Al-Amin Bhuiyan, and Chang Hong Liu, "On Face Recognition using Gabor Filters," *World Academy of Science and Technology*, no. 28, pp. 51-56, 2007.
[13] P. Penav, and J. Attick, "Local feature analysis: A general statistical theory for object representation", *Network: Computation in Neural Systems*, vol. 7, pp. 477500, 1996.
[14] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Wurtz, and W. Konen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture," *IEEE Transactions on Computers*, vol. 42, issue 3, pp. 300-311, 1993.
[15] T. Ojala, and M. Pietikainen, "Multiresoulation Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.
[16] A. Hafiana, G. Seetharaman, and B. Zavidovique, "Median Binary Pattern for Textures Classification," in Proc. *International Conference on Image Analysis and Recognition*, Aug. 2007.
[17] X. Tan and B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions," in Proc. *IEEE International Workshop on Analysis and Modeling of Faces andGestures (AMFG'07)*, 2007, pp. 168-182.
[18] A. Khan, F. Bashar, F. Ahmed, and M. H. Kabir, "Median Ternary Pattern (MTP) for Face Recognition," in Proc. *International Conference on Informatics, Electronics and Vision (ICIEV)*, May 2013.
[19] T. Kanade, J. Cohn, and Y. Tian, "Comprehensive Database for Facial Expression Analysis," in Proc. *IEEE International Conference on Automated Face and Gesture Recognition*, pp. 46-53, Mar. 2000.
[20] C. Hsu and C. Lin, "A Comparison on Methods for Multi-Class Support Vector Machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415-425, 2002.
[21] R. C. Gonzalez, and R. E. Woods, *Digital Image Processing*, 3rd Ed, Pearson Education, 2009.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Basis Expansion Model (BEM) Based MIMO-OFDMA Channel Capacity and Estimation

Mohammad Fazle Rabbi
Department of Electrical and Electronic Engineering
Leading University, Sylhet, Bangladesh
fzrabbi@gmail.com

***Abstract*- MIMO-OFDMA has become an efficient method characterizing diversity to multi-user transmission over doubly selective fading channels with high spectral efficiency. In this paper we develop a Basis Expansion Model (BEM) based MIMO-OFDMA system and we analyze the channel capacity for the system. It is shown that, by using BEM based channel approximation for MIMO-OFDMA system we can achieve diversity without any space-time coding. Specifically, BEM based system inherently exploits the receive diversity and maximizes channel capacity similar to existing diversity technique like Maximal Ratio Combining (MRC).**

***Index Terms*- MIMO, OFDMA, Basis Expansion Model (BEM), Receive diversity.**

## I. INTRODUCTION

Present cellular mobile systems are being designed to provide high-speed multimedia services for high mobility users. Such high data rates give rise to the frequency selectivity in the multipath environment while user's mobility causes time selective propagation. MIMO-OFDMA is one of the most promising technologies for multi-user wideband communication over frequency selective fading channel with high spectral efficiency. Multiple transmit and receive antennas in MIMO system offer channel capacity to achieve high data rate and space-time diversity to mitigate fading effects. On the other hand OFDMA as one of the promising multiple access techniques provide flexibility in resource allocation as well as multi-user diversity embedded into the wireless communication networks. Consequently, OFDMA has become a popular candidate for a number of future broadband wireless systems, including IEEE 802.16e and the Universal Mobile Telecommunications System (UMTS). Considering the doubly selective channel, in this paper we develop an appropriate MIMO-OFDMA uplink system model and then we investigate the channel capacity and spatial diversity of the system.

Basis Expansion Model (BEM) has been given huge attention in the recent studies [1]-[6] for its ability to approximate the time varying feature of the channel precisely. Based on the channel statistics for modeling, several BEM methods are proposed in recent studies. Discrete prolate spheroidal BEM (DKL-BEM) in [1] is proposed for rectangular spectrum and is specifically featured by a set of orthogonal functions. The complex exponential BEM (CE-BEM) in [2] has become popular for its easy algebraic expression. Though it suffers from modeling error a good number of studies [2]-[4] have proposed it for its capability to reform the time varying channel in urban mobile radio environment. Another version of BEM known as polynomial BEM (P-BEM) is also studied in [5], [6] which is basically suitable for lower Doppler scenario. In this paper we will use the principle of complex exponential BEM (CE-BEM) which exploits Fourier bases to give an approximate model of the time varying channel over a period of time. This can be regarded as a frequency domain representation of temporal variation of the channel. In principle, the Fourier transform of the Channel Impulse Response (CIR) over time completely characterizes the nature of the channel variation and linear combination of the Fourier bases sampled in the Doppler frequency domain gives us an adequate accurate approximation of the CIR.

As it is well known that, MIMO system can achieve spatial diversity by using the channel matrices to detect data streams at Base Station (BS). Generally transmit diversity is achieved applying Space-Time Code (STC) and receive diversity can be achieved using Selection Combining (SC) or MRC to maximize the capacity of the MIMO channel and to increase diversity order. But at the same time this causes the transmitter and receiver structure to be more complex. BEM based MIMO system (BEM-MIMO) developed in this paper achieves spatial diversity as well as the receive diversity without any space-time code. In addition, in the BEM-MIMO system, instead of using large channel matrices we use lower dimension BEM coefficients autocorrelation matrix to estimate the BEM coefficients.

The remainder of this paper is organized as follows. In Section II, we will develop a BEM based system model for MIMO-OFDMA uplink transmission. Based on the model, a time domain pilot aided channel estimation scheme will be formulated in Section III. The channel capacity and diversity analysis will be given in Section IV. Section V will present

simulation results and some concluding remarks will be given in Section VI.

## II. SYSTEM MODEL

Consider an OFDMA system with $N$ subcarriers and $U$ users. The signal transmitted from an arbitrary user $u$ at $t^{\text{th}}$ transmitted antenna is given mathematically by

$$x_u^{(t)}(n)=\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}M_u(k)X(k)e^{\frac{j2\pi nk}{N}},\ -N_g\le n\le N-1, \quad (1)$$

where $X(k)$ are complex data modulated by inverse discrete Fourier transform (IDFT) and

$$M_u(k)=\begin{cases}1 & \text{if } k \text{ is used by user } u\\ 0 & \text{otherwise}\end{cases}$$

is the masking variable identifying the subcarriers used by user $u$, and $N_g$ is the length of the Cyclic Prefix (CP) appended at the head of each OFDMA block as a guard interval. Note that $N_g$ should be larger than the delay spread caused by both multipath spread and user time misalignment in the doubly selective fading channel.

As mentioned earlier we will use CE-BEM based channel approximation which exploits the Fourier bases to represent the channel variations over a period of time. Specifically, the Channel Impulse Responses (CIR) of a path $h(t,\tau)$ can be well approximated by a linear combination of finite Fourier bases in a time period $[0\quad T]$. Mathematically,

$$h(t,\tau)=\sum_{q=-Q}^{Q}a_q(\tau)\exp\left(j2\pi tf_q\right),\ 0\le t\le T, \quad (2)$$

where $f_q$ is the $q$-th sampled frequency in the Doppler domain, and $a_q$ is the corresponding basis coefficient. With $T_S$ giving the temporal sampling interval, the discrete time model of (3) is given by

$$h(n,l)=h(t=nT_S,\tau=lT_S)=\sum_{q=-Q}^{Q}a_q(l)\exp\left(j2\pi nT_Sf_q\right),$$
$$0\le n\le N-1=\lfloor T/T_S\rfloor,\ 0\le l\le L-1=\lfloor\tau_{\max}/T_S\rfloor \quad (3)$$

$\tau_{\max}$ is the actual delay spread in the time domain, and $L$ is the delay spread in the discrete time model [7].

If the sampled frequencies in the Doppler domain are chosen to be

$$f_q=\frac{q}{gT},\ -\lceil gf_{\max}T\rceil\le q\le\lceil gf_{\max}T\rceil, \quad (4)$$

then, the discrete time channel model of (3) will become

$$h(n,l)=\delta(n-l)\sum_{q=-Q}^{Q}a_q(l)\exp\left(\frac{j2\pi nq}{gN}\right), \quad (5)$$

where $g$ is a positive integer and determines the frequency resolution in the Doppler domain. The case of $g=1$ gives rise to the conventional BEM model described in [2], [3], while the case of $g>1$ gives an oversampled version of the BEM model as used in [7], [8]. For convenience, $g$ will be termed as the BEM oversampling index.

From (5), we can now formulate the BEM based MIMO-OFDMA system uplink. While it is not really necessary, we will assume, for the sake of convenience, perfect time and frequency synchronization so that OFDMA block received at every $r^{th}$ receiving antenna from user $u$, after removal of the CP, has the form

$$y_u^{(r)}(n)=\sum_{t=1}^{N_t}\sum_{l=0}^{L-1}h_u^{(t,r)}(n,l)x_u^{(t)}(n-l)+z^{(r)}(n),$$
$$=\sum_{t=1}^{N_t}\sum_{q=-Q}^{Q}\exp\left(\frac{j2\pi nq}{gN}\right)\sum_{l=0}^{L-1}a_u^{(t,r)}(l,q)x_u^{(t)}(n-l)+z^{(r)}(n),$$
$$0\le n\le N-1, \quad (6)$$

where $N_t$ is the number of transmit antenna and $z^{(r)}(n)$ is complex AWGN with zero mean and variance $\sigma_n^2$ at the $r$-th receiving antenna. Substituting (1) into (7), we have

$$y_u^{(r)}(n)=\sum_{t=1}^{N_t}\sum_{q=-Q}^{Q}e^{\frac{j2\pi nq}{gN}}\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}\left(\sum_{l=0}^{L-1}a_u^{(t,r)}(l,q)M_u(k)X(k)e^{\frac{j2\pi lk}{N}}\right)e^{\frac{j2\pi nk}{N}}+z^{(r)}(n)$$
$$,\ 0\le n\le N-1 \quad (7)$$

Following (7), each OFDMA block received from user $u$ can be expressed in matrix form as

$$\boldsymbol{y}_u^{(r)}=\sqrt{N}\sum_{t=1}^{N_t}\sum_{q=-Q}^{Q}\boldsymbol{D}_q^{(r)}\boldsymbol{F}\text{diag}\left[\boldsymbol{M}_u\boldsymbol{X}\right]\boldsymbol{F}^H(:,1:L)\boldsymbol{a}_{u,q}^{(t,r)}+\boldsymbol{z}, \quad (8)$$

where

$$\boldsymbol{D}_q^{(r)}=\text{diag}\left[1\quad e^{\frac{j2\pi q}{gN}}\quad\ldots\quad e^{\frac{j2(N-1)\pi q}{gN}}\right],$$

$$[\boldsymbol{F}]_{x,y}=\frac{1}{\sqrt{N}}e^{\frac{j2\pi xy}{N}},$$

$$\boldsymbol{M}_u=\text{diag}\left[M_u(0)\quad M_u(1)\quad\ldots\quad M_u(N-1)\right],$$

$$\boldsymbol{X}=\left[\text{X}(0)\quad\text{X}(1)\quad\ldots\quad\text{X}(\text{N-1})\right]^T,$$

$$\boldsymbol{a}_{u,q}^{(t,r)}=\begin{bmatrix}a_u^{(1)}(0,q) & a_u^{(2)}(0,q) & \ldots & a_u^{(N_tN_r)}(0,q)\\ a_u^{(1)}(1,q) & a_u^{(2)}(1,q) & & a_u^{(N_tN_r)}(1,q)\\ \vdots & & \ddots & \vdots\\ a_u^{(1)}(L-1,q) & \ldots & & a_u^{(N_tN_r)}(L-1,q)\end{bmatrix},$$

$$\boldsymbol{z}=\left[z^{(1)}(n)\quad z^{(2)}(n)\quad\ldots\quad z^{(r)}(n)\right],$$

$\boldsymbol{F}^H(:,1:L)$ is the matrix formed from columns 1 to $L$ of $\boldsymbol{F}^H$. The basis coefficient matrix $\boldsymbol{a}_{u,q}^{(t,r)}$ in (8) represents BEM bases for all paths of each transmit and receive antenna pair.

By adding all the user signals, the OFDMA block received at the BS from all the users can be obtained from (8) as

$$\boldsymbol{y}^{(r)}=\sum_{u=1}^{U}\boldsymbol{y}_u^{(r)}=\sqrt{N}\sum_{u=1}^{U}\sum_{q=-Q}^{Q}\boldsymbol{D}_q^{(r)}\boldsymbol{F}\text{diag}\left[\boldsymbol{M}_u\boldsymbol{X}\right]\boldsymbol{F}^H(:,1:L)\boldsymbol{a}_{u,q}^{(t,r)}+\boldsymbol{z}. \quad (9)$$

With $\boldsymbol{x}_u^{(t)} = \boldsymbol{F}\boldsymbol{M}_u\boldsymbol{X}$ being the time domain symbols transmitted from user $u$, it is easy to show that

$$\boldsymbol{C}^{(t)}(\boldsymbol{x}_u) = \boldsymbol{F}diag[\boldsymbol{M}_u\boldsymbol{X}]\boldsymbol{F}^H(:,1:L) = \begin{bmatrix} x_u^{(t)}(0) & x_u^{(t)}(N-1) & \dots & x_u^{(t)}(N-L+1) \\ x_u^{(t)}(1) & x_u^{(t)}(0) & \dots & x_u^{(t)}(N-L+2) \\ \dots & \dots & \dots & \dots \\ x_u^{(t)}(N-1) & x_u^{(t)}(N-2) & \dots & x_u^{(t)}(N-L) \end{bmatrix} \quad (10)$$

is a matrix with $L$ columns formed by circularly shifting $\boldsymbol{x}_u$. By using augmented matrix we can rewrite (10) for all transmit antenna as given by

$$\boldsymbol{C}_u = \left[\boldsymbol{C}_u^{(1)} \quad \boldsymbol{C}_u^{(2)} \quad \dots \quad \boldsymbol{C}_u^{(N_t)}\right] \quad (11)$$

and (9) can be written in a simpler form in terms of the basis coefficients as

$$\boldsymbol{y}^{(r)} = \sum_{t=1}^{N_t}\sum_{u=1}^{U} \boldsymbol{\Gamma}_u^{(r)}\boldsymbol{a}_u^{(t,r)} + \boldsymbol{z}, \quad (12)$$

where

$$\boldsymbol{\Gamma}_u^{(r)} = \sqrt{N}\left[\boldsymbol{D}_{-Q}^{(r)}\boldsymbol{C}_u \quad \boldsymbol{D}_{-Q+1}^{(r)}\boldsymbol{C}_u \quad \dots \quad \boldsymbol{D}_{Q}^{(r)}\boldsymbol{C}_u\right] \quad (13)$$

$$\boldsymbol{a}_u = \sqrt{N}\left[\boldsymbol{a}_{u,-Q}^{(t,r)\,T} \quad \boldsymbol{a}_{u,Q+1}^{(t,r)\,T} \quad \dots \quad \boldsymbol{a}_{u,Q}^{(t,r)T}\right].$$

Augmenting $\boldsymbol{\Gamma}_u$ for all receive antenna, (13) can be written as

$$\boldsymbol{\Gamma}_u = \left[\boldsymbol{\Gamma}_u^{(1)} \quad \boldsymbol{\Gamma}_u^{(2)} \quad \dots \quad \boldsymbol{\Gamma}_u^{(N_r)}\right]$$

Finally, we can write the augmented matrix $\boldsymbol{\Gamma}$ and $\boldsymbol{a}$ for all users as given by

$$\boldsymbol{\Gamma} = \left[\boldsymbol{\Gamma}_1 \quad \boldsymbol{\Gamma}_2 \quad \cdots \quad \boldsymbol{\Gamma}_U\right],$$
$$\boldsymbol{a} = \left[\boldsymbol{a}_1^T \quad \boldsymbol{a}_2^T \quad \cdots \quad \boldsymbol{a}_U^T\right]$$

The received time domain OFDMA signal can be concisely expressed as

$$\boldsymbol{y} = \boldsymbol{\Gamma}\boldsymbol{a} + \boldsymbol{z}. \quad (14)$$

Note that the matrix $\boldsymbol{\Gamma}$ given in (14) has a dimension of $L_P \times (2Q+1)LU$, where $L_P$ is the length of each pilot block and $(2Q+1)LU$ is the number of BEM coefficients to be estimated.

As derived above, (14) can be regarded as corresponding to the time domain BEM models for the OFDMA uplink. Using these, we will now discuss how the BEM coefficients vector $\boldsymbol{a}$ can be estimated and used to calculate the CIR according to (5).

## III. CHANNEL ESTIMATION

### *A. Time Domain Pilot and Linear MMSE Estimator (LMMSE-T)*

In the time-selective channel estimation pilot symbols are normally needed and the pattern of pilot has significant effects on estimation performance. There are basically two types of pattern used in the current studies [7], [9], [11]. The pilot placed over the Frequency-time grid and pilot placed on the Time axis. In this paper we will use the time domain pilot pattern shown in Fig. 1. This purely time domain pilot pattern first proposed in [9], and has been investigated to assist OFDM channel estimation in [7], [10]. As shown in Fig. 1, all subcarriers in the blocks indexed by $i, i+\mathrm{K}, \dots .. i+(\Lambda-1)\mathrm{K}$ are used for pilots. Since these pilots are basically a series of time domain symbols, they can be added in the time domain interleaved with the IFFT modulated data blocks. As shown in [13], the time domain pilot block received from all users, using (14), can be given by

Fig. 1 Time domain Pilot pattern

$$\boldsymbol{y}^{(p)} = \boldsymbol{\Gamma}^{(p)}\boldsymbol{a} + \boldsymbol{z} \quad (15)$$

We will now investigate the use of the time domain based pilot pattern in conjunction with the BEM model in (15) to estimate the channel response in the time domain. First we will formulate a biased time domain estimator.

In a linear MMSE estimator, the overall MSE is minimized. After solving the standard optimization problem $\min_{\boldsymbol{G}} E\left\{\left\|\boldsymbol{a} - \boldsymbol{G}\boldsymbol{r}^{(p)}\right\|^2\right\}$ for the BEM signal model given by (15), the LMMSE estimator of the BEM coefficients is given by

$$\hat{\boldsymbol{a}}_{LMMSE} = \left(\boldsymbol{\Gamma}^{(p)H}\boldsymbol{\Gamma}^{(p)} + \sigma_n^2\boldsymbol{R}^{-1}\right)^{-1}\boldsymbol{\Gamma}^{(p)H}\boldsymbol{y}^{(p)}, \quad (16)$$

where $\boldsymbol{R} = E\{\boldsymbol{a}\boldsymbol{a}^H\}$ is the autocorrelation matrix of the BEM coefficients and can be pre-calculated as will be shown in next section. According to [13], since the LMMSE-T is solved out by minimizing MSE, the overall MSE will not be greater than those obtained by optimal unbiased estimators. After estimating BEM coefficients as shown above, we can easily calculate the estimated CIR using (5).

### *B. Autocorrelation Function of BEM Coefficients*

In this section we will derive the BEM coefficients autocorrelation matrix $\boldsymbol{R}$ used in (16). BEM coefficients are zero mean and Gaussian random variables [2]. The matrix $\boldsymbol{R}$ consists of nonzero entries associated with same path. According to (12), for the $l$-th path of user $u$, we can collect the CIR samples during the time interval $[0, N_t]$ to form the channel matrix mathematically given by

$$\boldsymbol{h}_u(l) = \boldsymbol{T}\boldsymbol{a}_u, \quad (17)$$

where

$$\boldsymbol{a}_u = \left[\boldsymbol{a}_u^{1}(l) \quad \boldsymbol{a}_u^{2}(l) \quad \dots \quad \boldsymbol{a}_u^{N_t N_r}(l)\right]^T \quad (18)$$

with the BEM coefficient of the $l$-th path of user $u$,

$$\boldsymbol{a}_u(l) = [a_u(-Q;l) \quad a_u(-Q+1;l) \quad \dots \quad a_u(Q;l)]^T \quad (19)$$

and

$$\boldsymbol{T}=\begin{bmatrix} \boldsymbol{T}^{(1,1)} & \boldsymbol{T}^{(1,2)} & \dots & \boldsymbol{T}^{(1,N_r)} \\ \boldsymbol{T}^{(2,1)} & \boldsymbol{T}^{(2,2)} & \dots & \boldsymbol{T}^{(2,N_r)} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{T}^{(N_t,1)} & \boldsymbol{T}^{(N_t,2)} & \dots & \boldsymbol{T}^{(N_t,N_r)} \end{bmatrix}, \tag{20}$$

where

$$\boldsymbol{T}^{(t,r)}=\begin{bmatrix} t^{(t,r)}(0,-Q) & t^{(t,r)}(0,-Q+1) & \dots & t^{(t,r)}(0,Q) \\ t^{(t,r)}(1,-Q) & t^{(t,r)}(1,-Q+1) & \dots & t^{(t,r)}(1,Q) \\ \dots & \dots & \dots & \dots \\ t^{(t,r)}(N_t,-Q) & t^{(t,r)}(N_t,-Q+1) & \dots & t^{(t,r)}(N_t,Q) \end{bmatrix}, \tag{21}$$

is the bases matrix corresponding to interval $[0,N_t]$, with

$$t^{(t,r)}(n,q)=\exp(j2\pi nq/gN)\cdot \tag{22}$$

The relation between CIR autocorrelation matrix $E\{\boldsymbol{h}_u(l)\boldsymbol{h}_u^H(l)\}$ and BEM coefficient autocorrelation matrix $R$ can be mathematically given by

$$E\{\boldsymbol{h}_u(l)\boldsymbol{h}_u^H(l)\}=\boldsymbol{T}\boldsymbol{R}_u(l)\boldsymbol{T}^H. \tag{23}$$

Using well-known Jake's model [11], $E\{\boldsymbol{h}_u(l)\boldsymbol{h}_u^H(l)\}$ can be calculated for SISO channel as given by

$$\phi(\Delta t)=J_0(2\pi f_d\Delta t), \tag{24}$$

where $J_0$ is the zero-order Bessel function of first kind. The CIR autocorrelation matrix for MIMO channel can be written as

$$\boldsymbol{\varphi}=\begin{bmatrix}\phi^{(1)}(\Delta t) & \phi^{(2)}(\Delta t) & \dots & \phi^{(N_tN_r)}(\Delta t)\end{bmatrix}^T \tag{25}$$

By taking $N_t>2Q$, $\boldsymbol{R}_u(l)$ can be calculated as

$$\boldsymbol{R}_u(l)=\boldsymbol{T}^{\dagger}\boldsymbol{\varphi}(\boldsymbol{T}^H)^{\dagger}, \tag{26}$$

where $\boldsymbol{T}^{\dagger}$ is the Moore-Penrose pseudo inverse of $\boldsymbol{T}$. Finally $\boldsymbol{R}_u(l)$ for all paths is calculated using (25) and (26).

## IV. CHANNEL CAPACITY AND DIVERSITY

Considering all received pilot block from all users as given by (15), the signal to noise ratio (SNR) can be calculated as

$$\gamma_{avg}=\frac{E\left[\left|\boldsymbol{\Gamma}^{(p)}\boldsymbol{a}\right|^2\right]}{E\left[\left|\boldsymbol{z}\right|^2\right]} \tag{27}$$

$$=\frac{E\left[\left|\boldsymbol{C}_p\right|^2\right]}{E\left[\left|\boldsymbol{z}\right|^2\right]}E\left[\left|\sum_{i=1}^{N_r}\boldsymbol{D}_i\sum_{j=1}^{N_tN_r}\boldsymbol{a}_j\right|^2\right], \tag{28}$$

$$=\frac{E_s}{N_t\sigma_n^2}E\left[\left|\sum_{i=1}^{N_r}\boldsymbol{D}_i\sum_{j=1}^{N_tN_r}\boldsymbol{a}_j\right|^2\right], \tag{29}$$

where matrix $\boldsymbol{C}_p$ in (28) represents all pilot symbols transmitted by all users and $E_s/N_t$ and $\sigma_n^2$ in (29) are the normalized pilot symbol energy and adaptive white Gaussian noise power respectively. Now let $\left|\sum_{i=1}^{N_r}\boldsymbol{D}_i\sum_{j=1}^{N_tN_r}\boldsymbol{a}_j\right|^2$ in (29) be the instantaneous value of the numerator and the instantaneous output SNR can be written as

$$\gamma_{ins}=\frac{E_s}{N_t\sigma_n^2}\left|\sum_{i=1}^{N_r}\boldsymbol{D}_i\sum_{j=1}^{N_tN_r}\boldsymbol{a}_j\right|^2 \tag{30}$$

Intuitively, to maximize the capacity of the channel we have to maximize the $\gamma_{ins}$ in (30) with respect to $\boldsymbol{D}_i$. As described in [12], according to the Cauchy-Schwarz inequality, we can rewrite (30) as

$$\gamma_{ins}\le\frac{E_s}{N_t\sigma_n^2}\sum_{i=1}^{N_r}\left|\boldsymbol{D}_i\right|^2\sum_{j=1}^{N_tN_r}\left|\boldsymbol{a}_j\right|^2 \tag{31}$$

Since BEM coefficients are zero mean Gaussian random variable, we can also show that

$$E\left[\sum_{j=1}^{N_tN_r}\left|\boldsymbol{a}_j\right|^2\right]=\sum_{j=1}^{N_tN_r}E\left[\left|\boldsymbol{a}_j\right|^2\right]=N_tN_r \tag{32}$$

So, following (31) and (32) the average output SNR can be given by

$$\gamma_{avg}\le\frac{E_s}{\sigma_n^2}N_rE\left[\sum_{j=1}^{N_r}\left|\boldsymbol{D}_j\right|^2\right] \tag{33}$$

From (33) using matrix norm,

$$\gamma_{avg}\le\frac{E_s}{\sigma_n^2}N_rE\left[\sum_{j=1}^{N_r}tr\left(\boldsymbol{D}_j\boldsymbol{D}_j^H\right)\right] \tag{34}$$

$$\gamma_{avg}\le\frac{E_s}{\sigma_n^2}N_rE\left[\sum_{j=1}^{N_r}\alpha_j^2I_N\right] \tag{35}$$

$$\gamma_{avg}\le\frac{E_s}{\sigma_n^2}N_r\sum_{j=1}^{N_r}\alpha_j^2, \tag{36}$$

where $\alpha_j^2$ represents the variances for each $\boldsymbol{D}_j\boldsymbol{D}_j^H$. From the derivation above we can see that, (36) represents the similar expression for MRC [14] which is used for receive diversity. Though we are not using any type of space-time coding or diversity techniques explicitly, BEM based system model inherently exploits the receive diversity and achieve same diversity like MRC. Moreover, the averaged SNR at the receiver is improved by a factor of $N_r\sum_{j=1}^{N_r}\alpha_j^2$ over the single input single output (SISO) link. Thus an array gain of $10\log_{10}\left(N_r\sum_{j=1}^{N_r}\alpha_j^2\right)$ is achieved.

Assuming the equal power allocation for all subcarrier, the capacity of the BEM-MIMO channel can be given by

$$I_{BEM-MIMO}=\sum_{i=1}^{M}\log_2(1+\gamma_{avg}), \tag{37}$$

where $M=\min(N_t,N_r)$. Substituting (36) to (37) we can rewrite the capacity of the BEM-MIMO channel as given by

$$I_{BEM-MIMO}\le\sum_{i=1}^{M}\log_2\left(1+\frac{E_s}{\sigma_n^2}N_r\sum_{j=1}^{N_r}\alpha_j^2\right) \tag{38}$$

Now (38) gives us the upper bound of the channel capacity and the equality holds when perfect knowledge of BEM coefficient matrix $\boldsymbol{a}$ is available at the BS.

## V. SIMULATION RESULT

We will now provide some simulation results to show the capacity of BEM based OFDMA. In the simulation, a 4-user 64-subcarrier interleaved OFDMA system uplink is used in a 3-ray doubly selective Rayleigh fading channel with $g$ = 10 and $f_{\max}T$ =0.01. QPSK modulated symbols are used as data and pilots.

In Fig. 2 we plot the channel capacity for the different number of transmit and receive antenna ( $Nt, Nr$ ) pairs. Since in the practical situation, it is difficult to increase the number of the mobile antenna, we keep the number of transmit antennas to 2 and increase the number receive antennas. From the figure it is clear that in high SNR region, capacity increases by 1 bit/sec/Hz for every 3 dB increase in SNR.

In Fig. 3 we demonstrate the channel estimation using BEM model. Specifically, we estimate the BEM coefficients of (16) prior to reconstruct the channel response. The mean square error (MSE) of channel estimation is plotted against SNR in the figure. As shown in the figure the BEM based channel estimation is also very effective for low range of SNR. Increasing the number of antenna also reduces the estimation MSE.

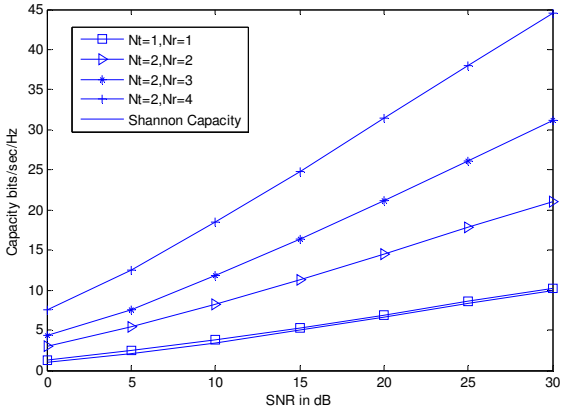



Fig. 2 Capacity comparison for BEM-MIMO Rayleigh fading channels with different number of antennas

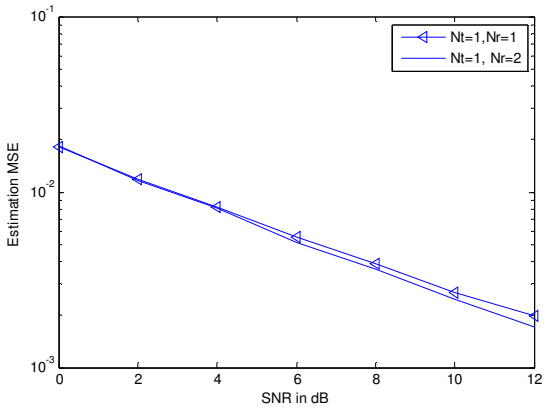



Fig. 3 Channel estimation MSE versus SNR

## VI. CONCLUSION

In this paper we derive formulation of a BEM based MIMO channel capacity for OFDMA uplink system. Specifically, we show that the similar diversity can be achieved as MRC based receive diversity without using any space time coding to the transmit signal. Also, we achieve the receive diversity without generating additional complex weight in the receiver. Moreover, our simulation results show that the channel capacity at the high-SNR region can be maximized by increasing receiving antennas in the Base Station instead of increasing number of transmitting antenna of MS.

## REFERENCE


[1] Thomas Zemen and Christoph F. Mecklenbräuker, "Time-Variant Channel Estimation Using Discrete Prolate Spheroidal Sequences", *IEEE Transactions on Signal Processing,* vol. 33, no. 9, pp. 3597 – 3607, September 2009.

[2] G. B. Giannakis and C. Tepedelenlioğlu, "Basis Expansion Models and Diversity Techniques for Blind Identification and Equalization of Time-Varying Channels", *Proceedings of IEEE*, vol.86, no.10, pp. 1969 – 1986, October 1998.

[3] X. Ma and G. B. Giannakis, "Maximum-Diversity Transmission over Doubly Selective Wireless Channels", *IEEE Transactions on Information Theory,* vol.49, no.7, pp. 1832 – 1840, July 2003.

[4] G.Leus and M.Moonen, "Deterministic subspace based blind channel estimation for doubly-selective channels", in *Proc. IEEE Signal Processing Advances in Wireless Communications (SPAWC),* June 2003, pp. 210-214.

[5] D.K. Borah and B.D. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model", *IEEE Trans. Commun.,* vol. 47, no. 6, pp. 862-873, June 1999

[6] S. Tomasin and A. Grorkhov, H. Yang and J.P. Linnartz, "Iterative interference cancellation and channel estimation for OFDM", *IEEE Trans. Commun.,* vol. 52, no. 1, pp. 238-245, Jan. 2005

[7] T. Cui, C. Tellambura and Y. Wu, "Low-Complexity Pilot-Aided Channel Estimation for OFDM Systems Over Doubly-Selective Channels", in *Proc. IEEE Inter. Conference on Commun.* 2005, pp. 1980 – 1984.

[8] I. Barhumi, G. Leus and M. Moonen, "Equalization for OFDM over doubly selective channels", *IEEE Trans. on Signal Processing*, vol. 54, no.04, pp. 1445 – 1458, April 2006.

[9] H. K. Lau and S. W. Cheung, "A Pilot Symbol-Aided Technique Used for Digital Signals in Multipath Environments", in *Proc. IEEE International Conference on Communications (ICC)*, May 1994., pp. 1126 – 1130,

[10] W. Hou and B. Chen, "ICI Cancellation for OFDM Communication Systems in Time-Varying Multipath Fading Channels", *IEEE Trans. on Wireless Comm.*, vol.4, no.5, pp.2100-2110, Sept. 2005.

[11] Z. Tang, R. C. Cannizzaro, G. Leus and P. Banelli, "Pilot-Assisted Time-Varying Channel Estimation for OFDM Systems", *IEEE Trans. on Signal Process.*, vol.55, no.5, pp.2226-2238, May 2007.

[12] S. Haykin and M. Moher, *Modern Wireless Communication,* Pearson Prentice Hall, 2005.

[13] Mohammad Fazle Rabbi, Sheng-Wei Hou and C. C. Ko, "High Mobility OFDMA Channel Estimation using Basis Expansion Model", *IET Communications*, vol. 4, no.3 pp. 353-367, Feb. 2010.

[14] A. Paulraj, R. Nabar and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, 2003.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Speech Enhancement Using Modified Magnitude and Phase Spectra

Sk. Imran Hossain[1], Md. Fahim Hossain Chowdhury[1], Md. Faijul Amin[1], Kazuyuki Murase[2]
[1]Khulna University of Engineering and Technology, Bangladesh
[2]University of Fukui, Japan
ranacsekuet@gmail.com, fahim9n@yahoo.com, mdfaijulamin@yahoo.com, murase@u-fukui.ac.jp

*Abstract*— **Degrading the quality and intelligibility of the speech signals, background noise is a severe problem in communication and other speech related systems. In order to get rid of this problem, it is important to enhance the noisy speech signal mainly through noise reduction. All most all the speech enhancement methods modify the frequency domain spectrum of the noise-corrupted speech to suppress the noise. Although both magnitude and phase spectra together contain the frequency domain information, the traditional speech enhancement procedures either work with magnitude or phase spectrum. This paper presents a speech enhancement method that exploits both magnitude and phase spectra. Experimental studies of the proposed method exhibits better PESQ (Perceptual Estimation of Speech Quality) score than that of other existing methods. We also found better speech quality with our proposed method in several subjective experiments.**

***Keywords—Speech Enhancement, Magnitude Spectrum, Phase Spectrum, Modified Magnitude and Phase Spectra (MMPS).***

## I. INTRODUCTION

Development and widespread deployment of digital communication systems during the last few decades have brought an increased attention to the role of speech enhancement in various speech processing problems. The main goal of speech enhancement is to reduce noise from corrupted speech signal with a view to improving its intelligibility and quality. Speech enhancement is the term used to describe such algorithms and devices whose purpose is to improve some perceptual aspects of the speech for the human listener or to improve the speech signal so that it may be better exploited by other speech processing algorithms.

A number of different methods for speech enhancement can be found in the literature that can be grouped into magnitude spectrum based methods (e.g., spectral subtraction [1], Wiener Filter [2], and minimum mean squared error (MMSE) estimator [3]) and phase spectrum based methods (e.g., phase spectrum compensation (PSC) [4]). All most all the methods employ a general framework called analysis-modification-synthesis (AMS) framework [5].

The basic AMS framework works as follows. Let us consider an additive noise model

$$x(n) = s(n) + d(n) \tag{1}$$

where $x(n)$, $s(n)$ and $d(n)$ denote discrete-time signals of noisy speech, clean speech and noise, respectively. Speech is analyzed frame-wise in the AMS framework through the short-time Fourier analysis because it can be assumed to be quasi-stationary for short duration. The discrete short-time Fourier transform (DSTFT) of the corrupted speech signal $x(n)$ is given by

$$X(n,k) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)e^{-j2\pi km/N} \tag{2}$$

where $k$ denotes the $k$th discrete-frequency of $N$ uniformly spaced frequencies and $w(n)$ is an analysis window function of short duration. Using DSTFT analysis Eq. (1) can be represented as

$$X(n,k) = S(n,k) + D(n,k) \tag{3}$$

where $X(n,k)$, $S(n,k)$, and $D(n,k)$ are the DSTFTs of noisy speech, clean speech, and noise, respectively. Each of these can be stated in terms of the DSTFT magnitude spectrum and the DSTFT phase spectrum. For example, the DSTFT of the noisy speech signal can be written in the polar form as

$$X(n,k) = |X(n,k)|e^{j\angle X(n,k)} \tag{4}$$

where $|X(n,k)|$ denotes the magnitude spectrum and $\angle X(n,k)$ denotes the phase spectrum.

Most of the existing AMS-based speech enhancement methods modify only magnitude spectrum keeping the noisy phase spectrum unchanged. A major reason for doing so is that the phase spectrum has been found to have less perceptual effect at significantly higher signal to noise ratio (SNR) levels [6]. The other reason is the difficulties to work with the phase spectrum, for example, the non-unique phase unwrapping problem [7]. Only recently, it has been shown that the phase spectrum may be useful in various speech processing applications [8]. Consequently, several speech enhancement methods based on the phase spectrum has been proposed very recently [4], [9]. However, all these research works consider a modification to either magnitude or phase spectrum. To the best of our knowledge, no attempt has been made to exploit the modification of both magnitude and phase spectra.

In this paper, we propose a new approach to speech enhancement, where the modified magnitude spectrum is

recombined with a changed phase spectrum to produce a modified complex spectrum (in the sense of real and imaginary parts). During modification and synthesis, low energy components of the modified complex spectrum cancel out more than the high energy components, resulting in background noise reduction. Thus, our proposed method is appropriate for situations where the noise energy is lower than the speech energy. It is demonstrated that the proposed method results in amended speech quality in terms of an objective speech quality measure PESQ (Perceptual Estimation of Speech Quality) score [12], spectrogram analysis and informal subjective listening test.

The rest of this paper is structured as follows. The proposed method is elaborately presented in Section II. Section III describes the enhancement experiments. Section IV presents results and discussion. Conclusion is drawn in Section V.

## II. PROPOSED METHOD

Our proposed speech enhancement method is established on the generally used speech processing framework, namely, AMS framework which consists of three stages: 1) the analysis stage, where DSTFT analysis is used to process the input speech signal; 2) the modification stage, where some modifications are applied on the noisy complex spectrum; and 3) the synthesis stage, where the output signal is constructed by applying the inverse discrete short-time Fourier transform (IDSTFT) operation followed by the overlap-add (OLA) synthesis process. We call our proposed method as modified magnitude and phase spectra (MMPS) algorithm, which is described in the block diagram of Fig. 1.

Let $\left|\widehat{D}(n,k)\right|$ denotes an estimate of the short-time magnitude spectrum of the noise. Then a pseudo magnitude spectrum is calculated as

$$\left|\hat{X}(n,k)\right| = \begin{cases} |X(n,k)| - \alpha\left|\widehat{D}(n,k)\right| & if\ |X(n,k)| > (\alpha+\beta)\left|\widehat{D}(n,k)\right| \\ \beta\left|\widehat{D}(n,k)\right| & otherwise \end{cases} \quad (5)$$

where $\alpha$, $\beta$ are tuning parameters. This pesudo magnitude spectrum is remcombined with the phase spectrum $\angle X(n,k)$ to create the pseudo complex spectrum (in the sense of real and imginary parts) as follows

$$\tilde{S}(n,k) = \left|\hat{X}(n,k)\right| e^{j\angle X(n,k)} \quad (6)$$

We then obtain the phase spectrum compensation function exploiting the conjugate symmetry property of phase [9], which is given by

$$\Lambda(n,k) = \lambda\Psi(k)\left|\widehat{D}(n,k)\right| \quad (7)$$

where $\lambda$ is a real-valued empirically determined constant, $\Psi(k)$ is the antisymmetry function. The time-invariant antisymmetry function is given by

$$\Psi(k) = \begin{cases} 1, & if\ 0 < \frac{k}{N} < 0.5 \\ -1, & if\ 0.5 < \frac{k}{N} < 1 \\ 0, & otherwise \end{cases} \quad (8)$$

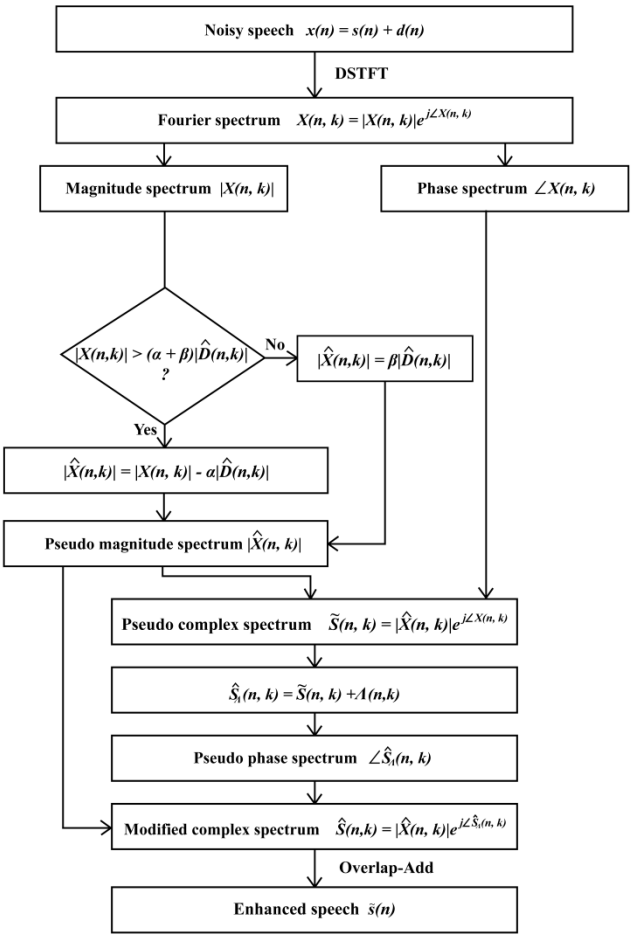



Figure 1. Block Diagram of MMPS.

where zero weighting is given to the values corresponding to nonconjugate part of the DSTFT (i.e. the $k = 0$ value and possible singleton at $k = N/2$ for $N$ is even). Note that $k$ varies form 0 to $N - 1$. Then this phase spectrum compensation function is added with the pseudo complex spectrum using the following equation:

$$\hat{S}_\Lambda(n,k) = \tilde{S}(n,k) + \Lambda(n,k) \quad (9)$$

Then the pseudo phase spectrum $\angle\hat{S}_\Lambda(n,k)$ is calculated. After that, we recombine the pseudo phase spectrum with the pseudo magnitude spectrum to get the modified complex spectrum.

$$\hat{S}(n,k) = \left|\hat{X}(n,k)\right| e^{j\angle\hat{S}_\Lambda(n,k)} \quad (10)$$

In the synthesis stage, the IDSTFT is used to convert the frequency-domain frames, $\hat{S}(n,k)$, to the time-domain representation. Resulting time-domain frames may be complex (real part and very small but nonzero imaginary part) due to numerical accuracy limitation of computer systems and additive offset. The imaginary part is discarded for this reason. Finally, the OLA procedure [10] is employed to produce enhanced time-domain signal, $\hat{s}(n)$.

## III. ENHANCEMENT EXPERIMENTS

### A. Speech Corpus and Noise Types

For evaluations, the NOIZEUS speech corpus is used [11]. The NOIZEUS is composed of 30 phonetically-balanced sentences belonging to six speakers (three males and three females). The corpus is sampled at 8 kHz and filtered to simulate receiving frequency characteristics of telephone handsets. For comparison purpose, as was done in [4], we generated a stimuli set corrupted by additive white Gaussian noise at five SNR levels: 0 dB, 5 dB, 10 dB, 15 dB and 20 dB.

### B. Evaluation Methods

To evaluate and compare with existing methods, we employ an objective speech quality measure, the PESQ score [12]. The score produces robust estimates of speech quality in the presence of a wide range of noise types. The PESQ prediction maps mean opinion score (MOS) estimates to a range between 1.0 and 4.5, where 1.0 corresponds to worst and 4.5 corresponds to distortion less [13]. Mean of PESQ scores over a subset of the NOIZEUS corpus are computed for our evaluation. In addition, we employ spectrogram analysis and informal subjective listening tests.

### C. Experimental Procedure

We evaluate two variants of our proposed methods – (i) using and (ii) ignoring the estimate of the short-time magnitude spectrum of the noise $\left|\widehat{D}(n,k)\right|$ of (7). Our proposed method incorporating the noise estimate is called MMPSWN and the method ignoring the noise estimate is called MMPSWON. We compare our result with the existing phase compensation methods PSCWN (renamed in this paper for consistency) [9] and PSCWON (renamed in this paper for consistency) [4]. The modified AMS procedure shown in Fig. 1 is employed to evaluate our proposed MMPS method. Samples of each of the sentence files are converted to zero-mean and normalized to be between -1.0 and +1.0. The frame duration is set to 32 ms and the frame shift to 4 ms for our experiment. The modified Hanning window [14] is used as the analysis window and the FFT length of 1024 samples is adopted. We use an anti-symmetric Λ function as given in Eq. (7). The value of $\lambda$ is set to 3.74 according to [4]. As it is mentioned before $\alpha$ and $\beta$ are two tunable parameters that govern the degree to which noise is suppressed. The values of α and β are set to empirically observed best values 0.6 and 0.35, respectively, from some preliminary trials. To illustrate, the PESQ values at different SNR are shown in Fig. 2 for two sets of values of α and β.

## IV. RESULTS AND DISCUSSION

Mean PESQ improvement scores for the white noise case investigated in our experiments are shown in Fig. 3. It can be observed that our proposed method performed better than the PSC method.

In order to illustrate better performance of our proposed method pictorially, the results of spectrogram analysis are shown in Fig 4. As can be seen, our proposed methods exhibit better noise cancellation than the PSC methods.

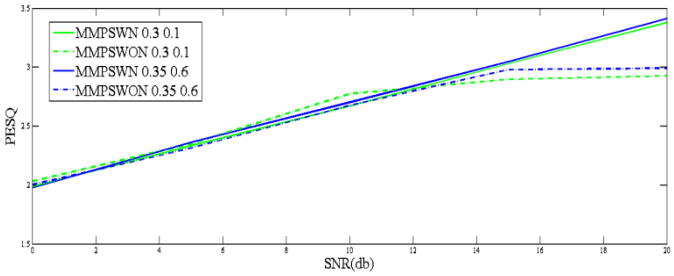



Figure 2. Mean PESQ vs SNR plot of MMPS method for different values of $\alpha$ and $\beta$.

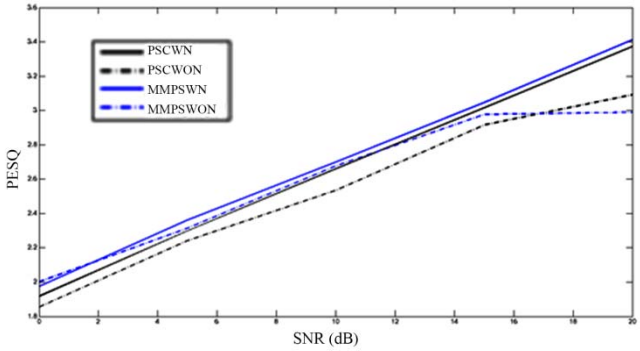



Figure 3. Mean PESQ VS SNR plot of MMPS (ours) and PSC methods [4], [9] for white noise case.

The mean PESQ values are shown in Table I to give an indicative performance of the proposed method against PSC methods [4], [9] that have shown better enhancement than the other existing speech enhancement techniques. The scores clearly show that the proposed method achieves better results than the PSC methods.

TABLE I. MEAN PESQ SCORES FOR THE WHITE NOISE CASE FOR THE MMPSWN, MMPSWON, PSCWN, PSCWON FOR MODIFIED HANNING WINDOW

| INPUT SPEECH SNR (dB) | METHODS | | | |
|---|---|---|---|---|
| | MMPSWN (proposed) | PSCWN[9] | MMPSWON (proposed) | PSCWON[4] |
| 0 | 1.975633 | 1.917909 | 2.001817 | 1.855557 |
| 5 | 2.362274 | 2.298900 | 2.314431 | 2.241660 |
| 10 | 2.700222 | 2.659496 | 2.677725 | 2.534646 |
| 15 | 3.046317 | 3.016357 | 2.977671 | 2.915680 |
| 20 | 3.411373 | 3.374276 | 2.990276 | 3.092118 |

We also conducted informal listening experiments where the clean signal, the noisy signal, and the enhanced signal were provided to the listeners. We found that, noisy speech enhanced by our method was more pleasant than that of PSC method. In other words, the residual noise present in the enhanced speech seemed to be non-distracting with our proposed method.

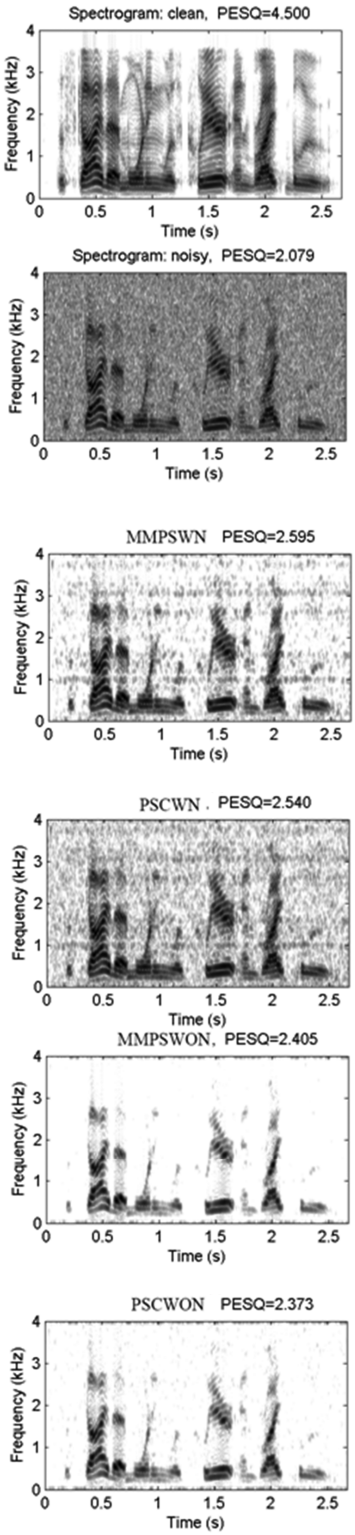



Figure 4. Spectrogram for the sentence "The sky that morning was clear and bright blue" at 10 dB SNR, white noise.

## V. CONCLUSION

In this paper, we have presented a novel approach to the speech enhancement problem. In the proposed method, the noisy short-time magnitude spectrum is recombined with a modified short-time phase spectrum to produce a modified short-time complex spectrum. During modification and synthesis, the low energy components of the modified complex spectrum cancel out more than the high energy components, thus reduces the background noise. Using an objective speech quality measure, informal subjective listening tests, as well as spectrogram analysis, we showed that the proposed method results in improved speech quality. In future, we would like to employ Voice Activity Detector (VAD) for more accurate and robust noise estimation. We also hope to use complex Kalman filter in the complex data of speech spectrogram.

## REFERENCES


[1] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction", *IEEE Trans. Acoust., Speech, Signal Process*., vol. ASSP-27, no. 2, pp. 113–120, Apr. 1979.

[2] N. Wiener, *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series With Engineering Applications*. New York: Wiley, 1949.

[3] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 443–445, Apr. 1985.

[4] Kamil Wójcicki, Mitar Milacic, Anthony Stark, James Lyons, and Kuldip Paliwal, "Exploiting Conjugate Symmetry of the Short-Time Fourier Spectrum for Speech Enhancement", *IEEE Signal Processing Letters*, vol. 15, 2008.

[5] J. Allen and L. Rabiner, "A unified approach to short-time Fourier analysis and synthesis", *Proc. IEEE*, vol. 65, no. 11, pp. 1558–1564, 1977.

[6] D.L. Wang and J.S. Lim, "The unimportance of phase in speech enhancements", *IEEE Trans. Acoust., Speech and Signal Process.,* Vol. 30, pp. 679-681, Aug. 1982.

[7] K. Paliwal, K Wójcicki, B Shannon, "The importance of phase in speech enhancement", *Speech communication,* Vol. 53 Issue 4, Apr. 2011.

[8] K. Paliwal, L Alsteris, "Usefulness of phase in speech processing", *Proc. IPSJ Spoken Language Processing Workshop*, Gifu, Japan, pp. 1-6, 2003.

[9] Stark, A., Wójcicki, K.K., Lyons, J.G. and K. Paliwal, "Noise driven short time phase spectrum compensation procedure for speech enhancement", *Proceedings of the 10th International Conference on Spoken Language Processing (INTERSPEECH-ICSLP),* Brisbane, Australia, pp. 549-552, Sep 2008.

[10] Lawrence R. Rabiner, Bernard Gold, *Theory and application of digital signal processing*. Englewood Cliffs, N.J.: Prentice-Hall. pp. 63–67. ISBN 0-13-914101-4, 1975.

[11] Y. Hu and P. Loizou, "Subjective comparison of speech enhancement algorithms", in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'06), Toulouse, France*, 2006, pp. 153–156.

[12] A. Rix, J. Beerends, M. Hollier, and A. Hekstra, "Perceptual Evaluation of Speech Quality (PESQ)—A new method for speech quality assessment of telephone networks and codecs", in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'01), Salt Lake City, UT*, 2001, vol. 2, pp. 749–752.

[13] URL: http://www.opticom.de/technology/pesq.php.

[14] Griffin, D., Lim, J., "Signal estimation from modified short-time Fourier transform", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-32 (2), 236–243, 1984.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Brightness Preserving Bi-Histogram Equalization Using Edge Pixels Information

Md. Moniruzzaman*, Md. Shafuzzaman and Md. Foisal Hossain
Khulna University of Engineering and Technology, Bangladesh
*mdzamankuet@gmail.com

*Abstract*— **Image enhancement plays an important role to the image or video display systems. It is used to process an input image and the output image is more pleasing than the input one. Many contrast enhancement techniques proposed so far. Among them histogram equalization (HE) is widely used for contrast enhancement technique. In this paper, we propose a new technique of histogram equalization method which preserves the brightness of the image. We tested the proposed method on a variety of images. The proposed method is also compared with some other related methods. We also use some parameters for comparison and the proposed method gives better performance than the related methods.**

***Keywords—Global Histogram Equalization, Brightness preserving Bi-Histogram Equalization, image brightness , canny edge detection.***

## I. INTRODUCTION

Image contrast enhancement is the most appealing areas of digital image processing. The idea behind enhancement techniques is to bring out detail that is obscured. The principle objective of contrast enhancement is to process an image so that the result is more suitable than the original image.

An effective contrast enhancement technique is HE [1] technique which distributes pixel values uniformly such that enhanced image have linear cumulative histogram. The HE is a global operation technique. This technique has one drawback and the drawback is- it does not preserve the image brightness.

To overcome this problem various methods have been proposed. Brightness preserving Bi-Histogram Equalization (BBHE) [2] and Equal area Dualistic Sub-Image Histogram Equalization (DSIHE) [3] split the input image histogram into two sub-histograms according to mean and median, respectively. In Recursive Mean-Spread Histogram Equalization (RMSHE) [4], scalable brightness preservation is achieved by partitioning the histogram recursively more than once. Minimum Mean Brightness Error Dynamic Histogram Equalization (MMBEDHE) [5], an extension of BBHE, provides maximal brightness preservation. Further two Multi Histogram Equalization (MHE) [6] methods, i.e. Minimum Middle Level Squared Error MHE (MMLSEMHE) and Minimum Within-Class Variance MHE (MWCVMHE), have been proposed to produce image with natural looks. Also some adaptive histogram equalization (AHE) [7], [8] techniques have been proposed which is a natural extension of GHE technique. Again the contrast can be enhanced not only for the normal images but also for weather degraded images [9].

In this paper we propose a histogram equalization technique for image contrast enhancement which preserves the brightness of the image. The proposed technique is compared with two other methods namely- HE and BBHE methods. The comparison results give better performance in case of our proposed technique. It preserves the brightness of the image and also gives natural looks of the images.

This rest of paper is organized as follows: section-II explains global histogram equalization (GHE) method and section-III explains brightness preserving bi-histogram equalization (BBHE) method. Section-IV presents the proposed algorithm. Results and comparison of proposed method with some other existing methods is presented in section-V. Finally section-VI concludes the paper followed by references.

## II. GLOBAL HISTOGRAM EQUALIZATION METHOD

Suppose input image I(x, y) composed of discrete gray levels in the dynamic range of [0, L-1]. The transformation function $C(r_k)$ is defined as

$$S_k = C(r_k) = \sum_i (r_i) = \sum_i (n_i / n) \quad (1)$$

Where, k = 0, 1, 2, 3……., L-1 and i = 0, 1, 2 ……..k.

In (1), the number of pixels having gray level $r_i$ is represented by $n_i$, n is the total number of pixels in the input image, and Probability Density Function (PDF) is represented by $P(r_i)$. The Cumulative Density Function (CDF) is calculated from PDF, defined as $C(r_k)$. The mapping in (1) is called Global Histogram Equalization (GHE).

## III. BRIGHTNESS PRESERVING BI-HISTOGRAM EQUALIZATION METHOD

The BBHE decomposes an input image into two sub-images based on the mean of the input image. The method is used to preserve the input image brightness. Among two sub-images, one sub-image is the set of gray-levels less than or equal to the mean. The other sub-image is the set of gray-levels greater than the mean. Then the BBHE equalizes the sub-images independently based on their respective histograms. The gray-levels in the formal set are mapped into the range from the minimum gray level to the input mean. The

gray-levels in the latter set are mapped into the range from the mean to the maximum gray level.

## IV. PROPOSED ALGORITHM

(1). Consider an image I(x, y) of N pixels composed of discrete gray levels in the dynamic range of [0, L-1].

(2). Evaluate the distribution of details in a given image.

(3). Use Canny detection technique to evaluate the distribution of the details. The Canny edge detection algorithm is a robust and accurate edge detection method. Therefore an edge detected image E(x, y) is obtained from the original image I(x, y).

E(x, y) gives the binary result. The result of E(x, y) is zero (0) in areas of constant gray level and is one (1) where there is change in gray levels.

(4). Find out the pixels from the original image I(x, y) which are responsible to give the value of one in binary edge detected image E(x, y). These pixels form D(x, y).

(5). Detect the pixel (M) of maximum occurrence from D(x, y). In other word, find out the pixel of maximum probability from D(x, y), indicated by M.

(6). Now divide the original image into two sub-images ($I_L$ and $I_U$) based on the value of M (pixel of maximum probability in D(x, y)).

$I_L(x, y)$: if $I(x, y) <= M$

$I_U(x, y)$: otherwise

(7). Calculate the cumulative distribution function (CDF) of each sub-image

(8). Equalize the sub-images independently based on their respective histograms using classical HE technique which has been described in section 2 in (1). The range for $I_L(x, y)$ is from the minimum gray level 0 to M and the samples in $I_U(x, y)$ are mapped into the range from M+1 to the maximum gray level (L-1).

## V. EXPERIMENTAL RESULTS AND COMPARISON

The proposed algorithm is compared with Histogram Equalization (HE) method and Brightness Preserving Bi Histogram Equalization (BBHE) method. We use two parameters (1) AMBE and (2) PSNR to demonstrate the performance.

The low value of AMBE and high value of PSNR is required for good enhancement.

We use five images shown later in "Fig. 1", "Fig. 2" and "Fig. 3" . Only in case of Bridge image the experimental results give same results as BBHE. Except Bridge image other four images give better performance than HE and BBHE methods.

TABLE I. AVERAGE MEAN BRIGHTNESS ERROR (AMBE)

| | HE | BBHE | Proposed method |
|---|---|---|---|
| City with river | 24.3350 | 29.7404 | 3.4737 |
| Palace | 42.0428 | 35.4111 | 2.5079 |
| Bridge | 53.9887 | 13.6073 | 13.6073 |
| Sea and stone | 11.386 | 8.5763 | 3.3188 |
| Forest | 17.076 | 5.6584 | 5.1136 |

TABLE II. PEAK SIGNAL TO NOISE RATIO (PSNR)

| | HE | BBHE | Proposed method |
|---|---|---|---|
| City with river | 13.8592 | 13.5702 | 15.1270 |
| Palace | 14.1601 | 14.9148 | 20.4816 |
| Bridge | 9.3525 | 11.6074 | 11.6074 |
| Sea and stone | 13.5354 | 13.6592 | 13.8156 |
| Forest | 18.3803 | 21.8639 | 22.9536 |

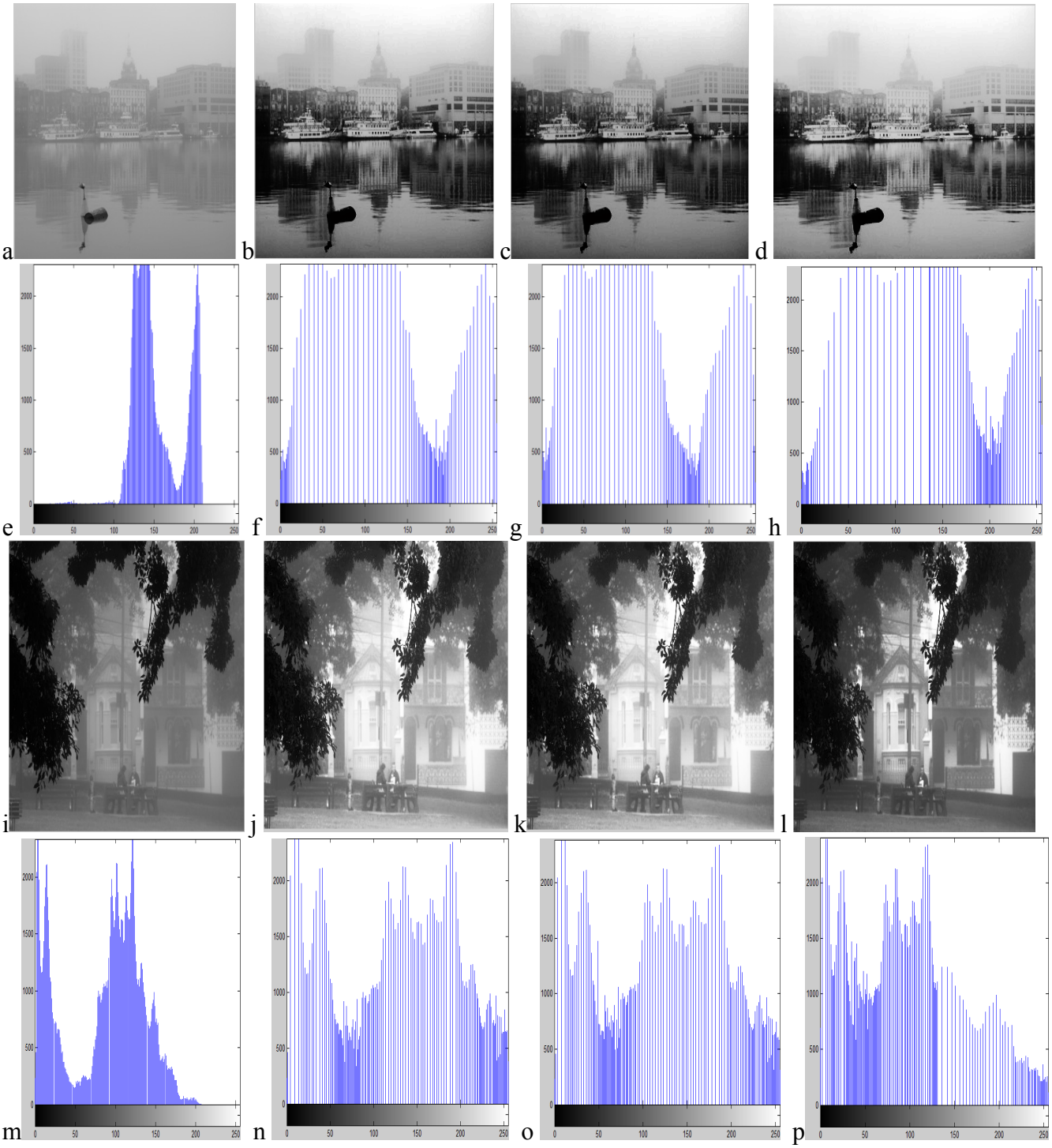



Fig. 1 (a),(i)-original images and (e),(m)-corresponding histograms;(b),(j)-HE results and (f),(n)-corresponding histograms;(c),(k)-BBHE results and (g),(o)-corresponding histograms;(d),(l)-results of proposed method and (h),(p)-corresponding histograms.

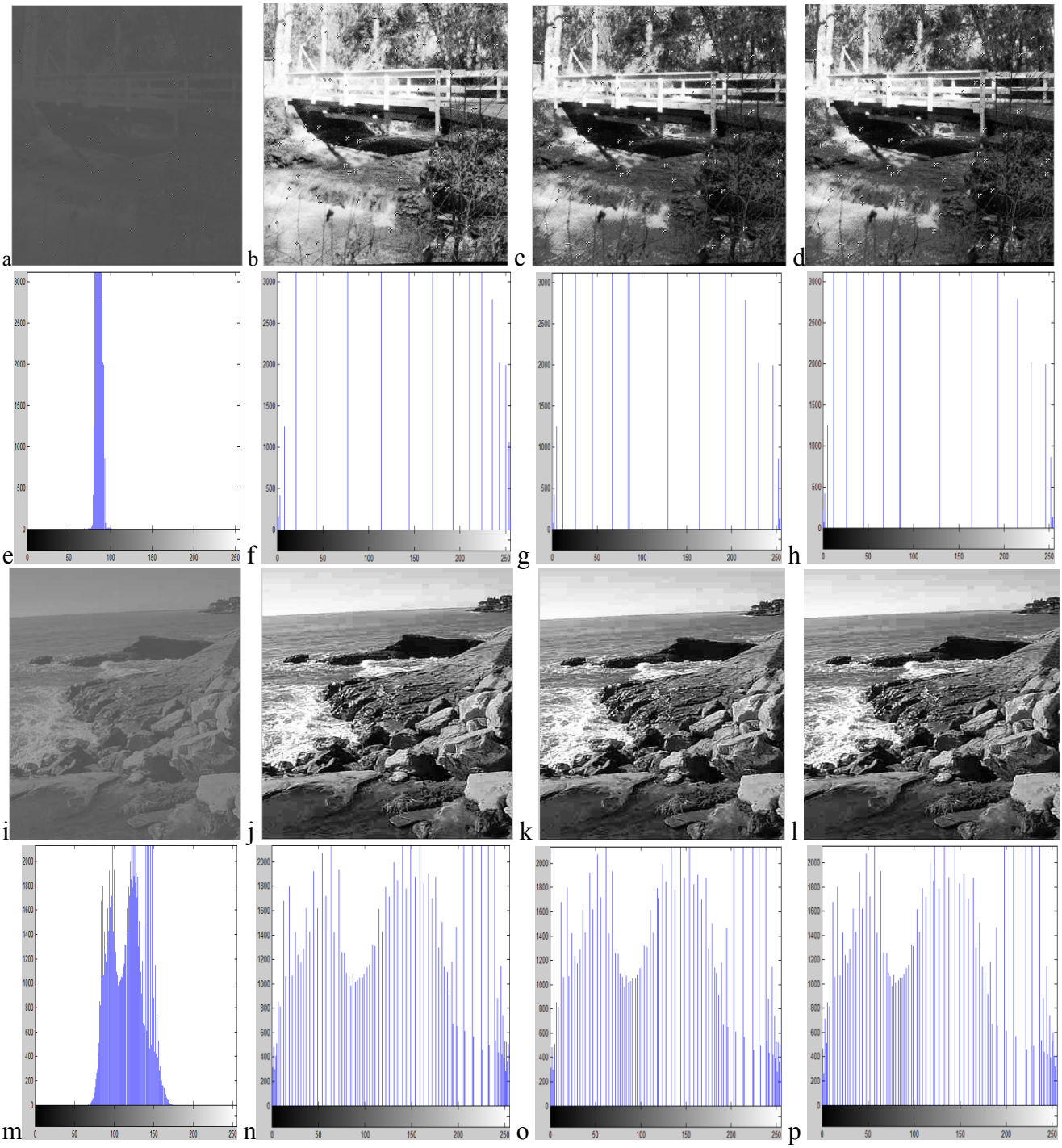


Fig. 2 (a),(i)-original images and (e),(m)-corresponding histograms;(b),(j)-HE results and (f),(n)-corresponding histograms;(c),(k)-BBHE results and (g),(o)-corresponding histograms;(d),(l)-results of proposed method and (h),(p)-corresponding histograms.

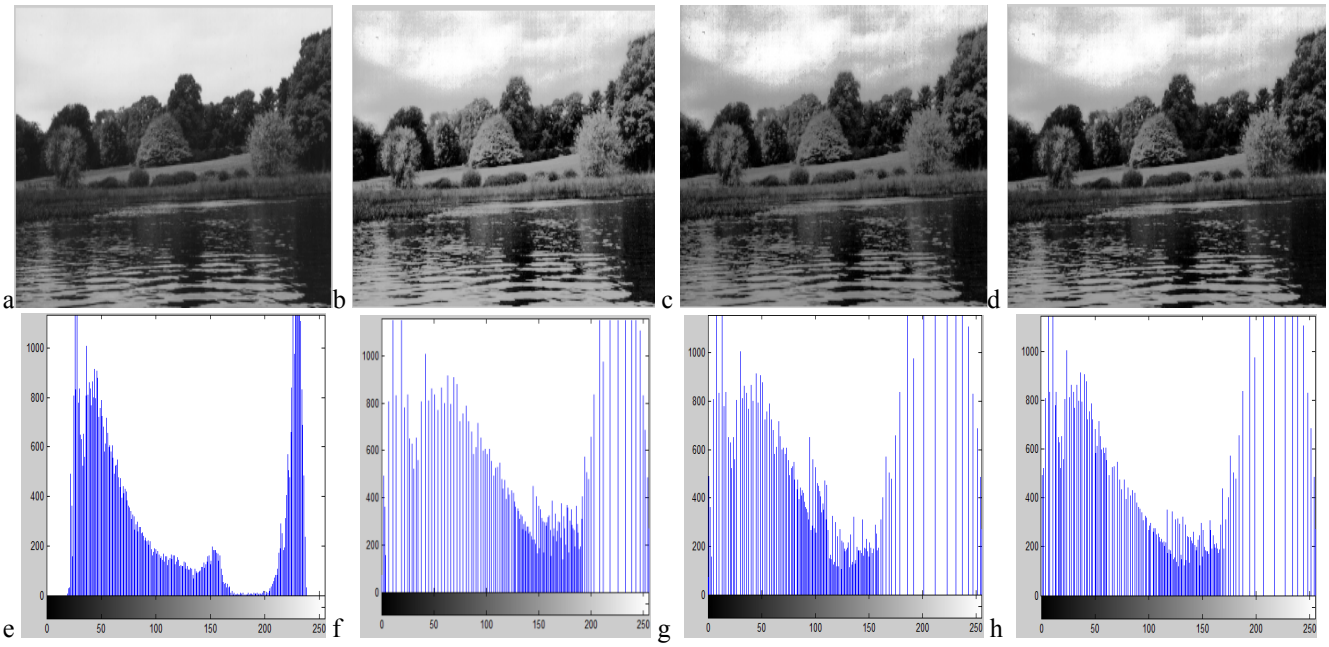



Fig. 3 (a)-original image and (e)-corresponding histogram;(b)-HE result and (f)-corresponding histogram;(c)-BBHE result and (g)-corresponding histogram;(d)-result of proposed method and (h)-corresponding histogram.

## VI. CONCLUSION

In this paper we propose a histogram equalization technique for image contrast enhancement which preserves the brightness of the image. Generally for good contrast enhancement low value of AMBE and high value of PSNR is required. As shown by the results of TABLE-I & TABLE-II, better results are obtained from proposed method. Therefore the low contrast images are increased by the proposed method shown in "Fig. 1", "Fig. 2" and "Fig. 2". The method was coded and tested successfully by using Matlab.

## REFERECES

[1] A. Rafael C. Gonzalez, and Richard E. Woods, "Digital Image Processing," 2nd edition, Prentice Hall, 2002.

[2] [2] Y.-T. Kim, "Contrast enhancement using brightness preserving bi - histogram equalization," IEEE Trans. On Consumer electronics, vol. 43, no. 1, pp. 1-8, Feb. 1997.

[3] [3] Y. Wang, Q. Chen, B. Zhang, Image enhancement based on equal area dualistic sub-image histogram equalization method, IEEE Transactions on Consumer Electronics, vol. 45, no. 1, pp. 68–75, 1999.

[4] Chen, S.-D., Ramli, A.R., 2003. Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation. IEEE Trans. on Consumer Electronics, vol. 49, no. 4, pp. 1301–1309.

[5] Md. Foisal Hossain and Mohammad Reza Alsharif, "Minimum Mean Brightness Error Dynamic Histogram Equalization for Brightness Preserving Image Contrast Enhancement", International Journal of Innovative Computing, Information and Control (IJICIC), Vol. 5, No. 10 (A), pp.3263-3274, October, 2009.

[6] D. Menotti, L. Najman, J. Facon and Arnaldo de A. Araujo, Multi-histogram equalization methods for contrast enhancement and brightness preserving, *IEEE Trans. Consumer Electronics*, vol. 53, no. 3, pp. 1186-1194, 2007.

[7] S.M. Pizer, J.B. Zimmerman, E.V. Staab, Adaptive grey level assignment in CT- scan display, J. Comput. Assist. Tomogr. 8 (1984) 300–305.

[8] Md. Foisal Hossain and Mohammad Reza Alsharif, "Image enhancement based on logarithmic transform coefficient and adaptive histogram equalization," International Conference on Convergence Information Technology(ICCIT 07), Gyeongju , Korea, pp-1439-1444, November 21-23, 2007

[9] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," IEEE Trans. PAMI,vol. 25, pp. 713-724, 2003.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Real-time Numeric Character Recognition System Based on Finger Movements

M. M. Farhad, Md. Sohorab Hossain*, S. M. Nafiul Hossain, Mohiuddin Ahmad
Dept. of Electrical and Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
* E-mail: sohorabhossain2k9@yahoo.com

*Abstract*—**In the field of computer vision, numeric character recognition plays an important role for various applications such as in Human Computer Interaction system. A numeric character recognition system is advantageous in a Human Computer Interaction system as it enables us to interact with a computer without a keyboard or other input devices. In this paper a simple numeric character recognition algorithm has been proposed where colored fingertip has been used for the purpose of tracking. Numeric characters are detected by the finger movements made during the handwriting of the characters. For each frame the tip position is compared to the previous tip position to decide whether the movement is a vertical or horizontal movement. Vertical movements can take place in upward or downward direction and horizontal movement can take place in rightward or leftward direction. This variation of fingertip movements gives rise to the classification criteria for the numeric characters. This method has shown promising result in real time environment.**

***Keywords—color detection, fingertip tracking, blob detection, numeric character recognition.***

## I. INTRODUCTION

Over the last few years with the advancements in the computer vision the necessity of recognizing characters at different sectors has increased. For the simplicity of use these recognizing methods are getting more popularity day by day. Character recognition is a major concern of the researchers in these days. The use of the character recognition systems finds its application in license plate recognition systems [1], [2], hand writing recognition systems [3], [4], [5]. In some cases the characters are identified from gesture from any part of our body [6], [7] or from the natural events [8], [9] to aid the controlling of any digital systems.

From the viewpoint of the recognition the recognition method may be two types. The first type includes the use of any classifier [10] and the second one is without a classifier. Though the use of classifier often shows better result but the complexity in calculation and accuracy constraints limits its use in character recognition. Again some methods use specialized devices for the detection of character using body movements [11]. Among the devices the MEMS are mostly used [12], which is actually a sensor to detect the three dimensional body part movement. The hand writing recognition largely depends on the segmentation process [13]. For this reason many segmentation algorithms such as recognition based segmentation [14], over segmentation [15] fusion based segmentation [16] has been proposed.

In this paper a new method of recognition of numeric characters have been proposed that uses a red fingertip to detect the finger movements. The practical direction of our finger that we use in order to write any numeric character is taken into account for the purpose of detection. The real time video is taken from the camera and the position of red fingertip is detected from the consecutive frames. This position is compared with the previous tip position of the finger tip. These positions are plotted on a two dimensional plane and a number is assigned for the movement in each direction. During the writing of any character the finger movement is fixed and follows a definite sequence. So this sequence is different for different characters. By detecting the finger movements and comparing with the predefined sequence of movements the character is detected. This algorithm can detect the characters without high computational complexity and classifier.

## II. CONCEPT DEVELOPMENT

The numeric character recognition system is dependent on the basis of fingertip movement that if someone moves finger during writing the character. The movements are categorized into 4 types and designated with the numbers 1, 2, 3 and 4. The 4 kinds of movements are horizontal left, horizontal right, vertical up and vertical down. The directions and designated numbers are shown in the Fig. 1 by plotting it on a two dimensional space. The direction of the fingertip is detected during the recognition process.

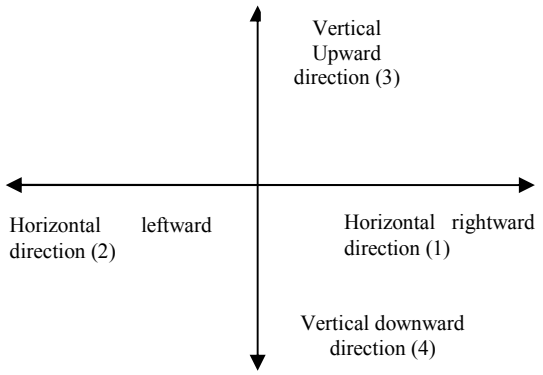



Figure 1. The direction of fingertip movements

During the writing sequence the classification of the characters are necessary. Fig. 2 shows the movements made by the fingers during the handwriting process. Suppose we want to write 2. So we will move our finger firstly from left to right horizontally. Then from up to down, right to left, again up to down and finally left to rightconsecutively as shown in Fig. 2. So the sequence made by the fingertipis 14241.

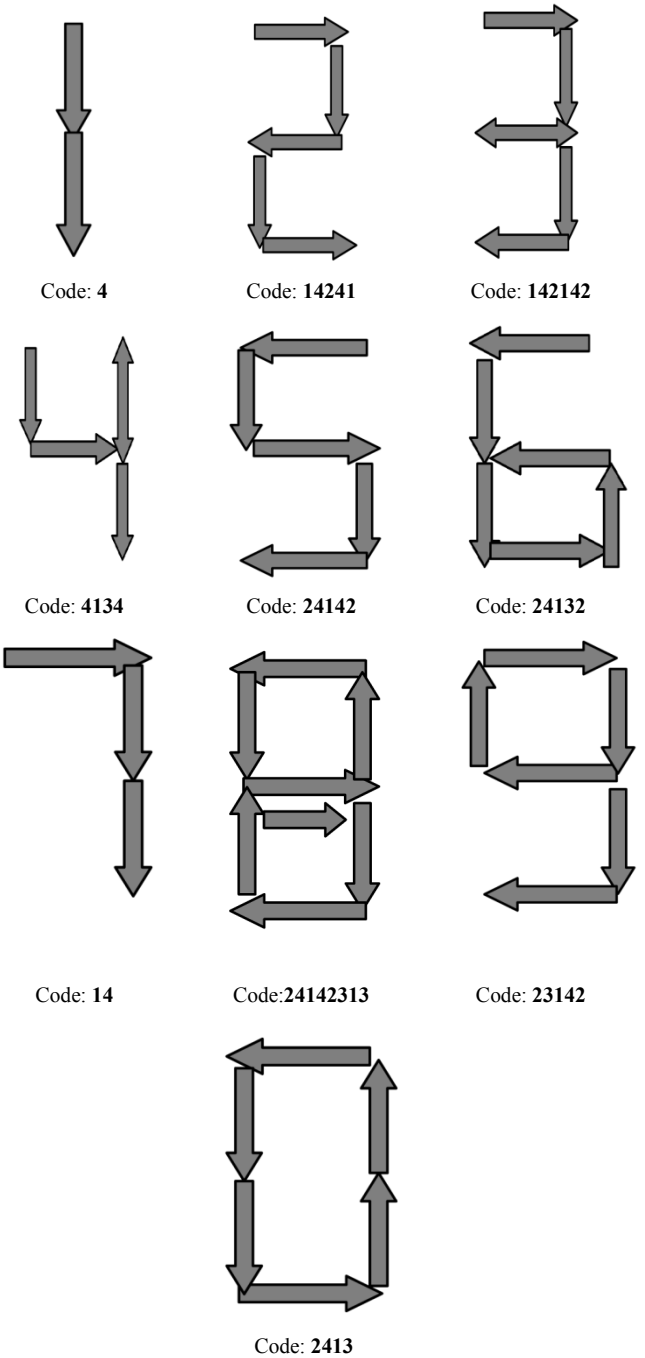



Figure 2. The direction of movements made by finger tips during the handwriting of numeric characters (direction of arrow conforms the direction of finger tip movement)

For the handwriting of 5 the direction of movement of fingertip is shown in the Fig. 2. The consecutive steps are horizontal leftward, vertical downward, horizontalrightward, vertical downward and horizontal leftward. So the code of 5 is 24142. Similarly for 3, the directions are horizontal rightward, vertical downward, horizontal leftward, horizontal rightward, vertical downward, and horizontal leftward. So the code for 3 is 142142. The codes for the rest of the characters are shown in Fig. 2. To make the detection irrelevant with the frame rate two consecutive movements in the similar direction is taken as a single movement in that direction. So the code for 1 is 4 instead of 44.

## III. PROPOSED METHOD

Our proposed method consists of the following steps:

(i). Acquisition of Video sequence

(ii). Fingertip detection (Blob detection)

(iii). Detection of movement

(iv). Compare with the ideal movements and final recognition

Fig. 3 represents the blog diagram of the proposed method.

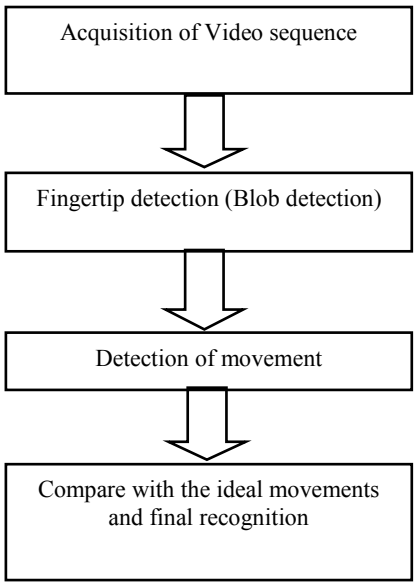



Figure 3. Block diagram of the proposed algorithm

### *A. Acquisition of Image*

For the acquisition of image STH-DCSG-9CM camera was used due to the availability of its plug and play characteristics. The STH-DCSG is a revolutionary stereo head with a global shutter specially designed for machine vision tasks of scenes with motion and most importantly Pre-calibrated facility. It has Global shutter, low-noise, high-sensitivity CMOS imager, Color or monochrome, uncompressed video at VGA resolution (30 fps), includes miniature lenses and SRI Small Vision System Software for real-time stereo analysis shown in Fig. 4. The camera is interfaced with computer with the IEEE 1394 card shown in Fig. 5. IEEE1394 card is a serial transfer standard used for real time data transfer and fast communication.

Figure 4. Pictorial view of a STH-DCSG - 9CM camera

Figure 5. Pictorial view of IEEE1394 card

### B. *Finger tip detection*

The fingertip is covered with a red colored stripe as shown in the Fig. 6(a). To detect the red colored fingertip blob tracking method is used. Blob tracking is a simple method that can be used for many computer vision tasks. Blob detected method detects an area where the value of different properties of a pixel such as Hue, Saturation or Illumination lies within a prescribed range. Firstly the image are converted RGB to HSV format so that the colors are separable. After using the thresholding function the pixels that lie in the range of the desired HSV values are detected. In here the red fingertip is detected by blob tracking and marked by the green circle. Fig. 6(b) shows the detection of the blob as the red fingertip moves frame to frame.

(a)


(b)

Figure 6. (a) Arrangement of writing character, (b) Detected finger tip with blob track

### C. *Detection of Movement*

After detecting the fingertip position of each frame the movement of the position of tips in two consecutive frames is detected. There may be four types of movements: a) horizontal rightwards, b) horizontal leftward, c) vertical upward and d) vertical downward. For the movements there may be angular movements too so that detection of the movement is confusing because it may be either vertical or horizontal movement. For this the threshold value of angle is used as $45^o$. So any upward or downward movement taking place with an angle less than $45^o$ is treated as vertical movement and movements greater than $45^o$ will be treated as the horizontal movement. Similar logic holds for the rightward or leftward movements. Again a movement less than 10 pixels are ignored for the purpose of reduction of noise which occurs at the start and end of writing any character.

### D. *Compare with the ideal movements and final recognition*

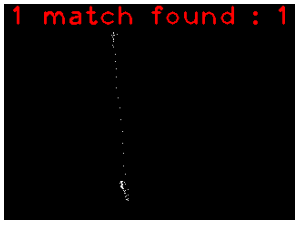
For each character the ideal sequence of movements has already been shown in Fig. 2. After detecting the fingertip movements these are compared with the ideal movements shown in the Fig. 2. The movements that match the ideal movements are recognized as the corresponding characters. For example, if the first detected direction is downward (i.e. direction no is 4) then the system detects that possible characters are 1 and 4. After that if the fingertip is kept rest at a certain position i.e. if no finger movement is recorded for consecutive 30 frames then the system assumes that desired character has been written completely, so for the above situation the detected code is 4 and detected character is 1. Now the system is ready to take the next input. After completing one character to write next one the hand is moved from end position of last character to initial position of next characterin upward direction, so the detected direction code for this action is 3 and there is no character whose start code is 3. This is used as a cue to identify the start of new character.

## IV. EXPERIMENTAL RESULT

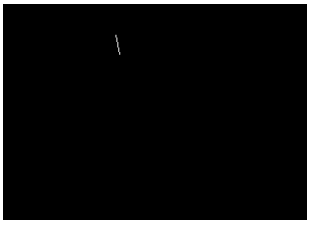
A number of real time experiments were concluded in order to determine the validity and accuracy of the algorithm. The experiments were concluded for all the ten numeric characters. The output was checked for a single character several times. In Fig. 7 the detection of all the numeric characters are shown.

In Fig.7 the detected positions of the finger tips during the process of writing 1(one). 2(two), 3(three), 4(four), 5(five), 6(six), 7(seven), 8(eight), 9(nine) and 0(zero)are shown. In the $2^{nd}$ column of Fig. 7 shows the traced path during these frames. In the first case, we can see that the primary movement was a vertical downward. Later as there was no movement to other directions so it must be 1.Similarly for 3 the code recorded from the movements is 142142 which are identical with the code assigned for 3 shown in Fig. 2. So the character is detected as 3. Similarly for rest of the characters, the result is shown in Fig. 7.

Recognition of 1(one)

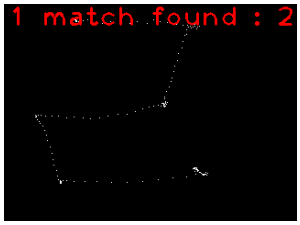



Detected fingertip positions for all frames for 1 (one)

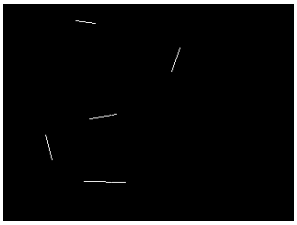


Traced direction of movements

Recognition of 2(two)



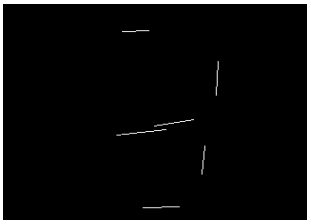

Detected fingertip positions for all frames for 2 (two)



Traced direction of movements

Recognition of 3(three)

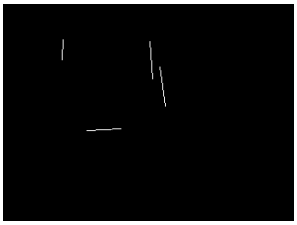

Detected fingertip positions for all frames for 3 (three)



Traced direction of movements

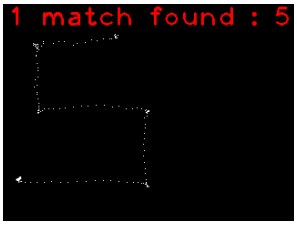
Recognition of 4(four)


Detected fingertip positions for all frames for 4 (four)

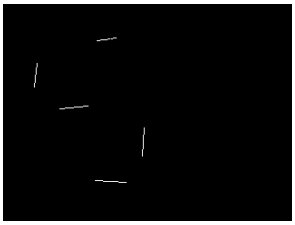


Traced direction of movements

Recognition of 5(five)

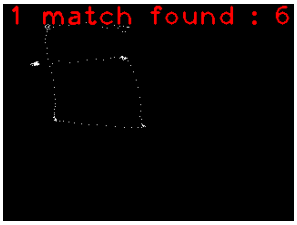



Detected fingertip positions for all frames for 5 (five)

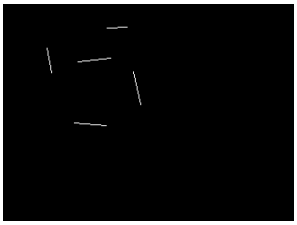


Traced direction of movements

Recognition of 6(six)



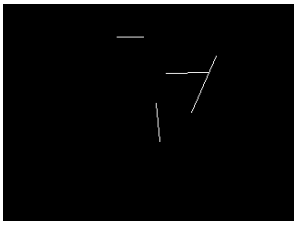

detected fingertip positions for all frames for 6 (six)



Traced direction of movements

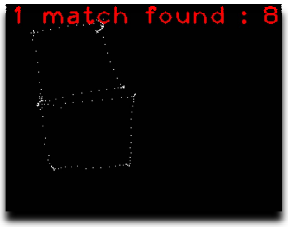
Recognition of 7 (seven)


detected fingertip positions for all frames for 7 (seven)

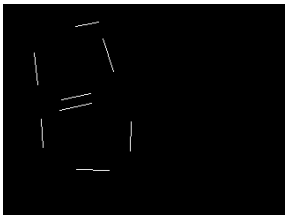


Traced direction of movements

Recognition of 8(eight)

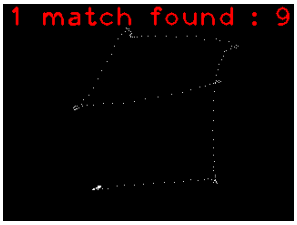



detected fingertip positions for all frames for 8 (eight)

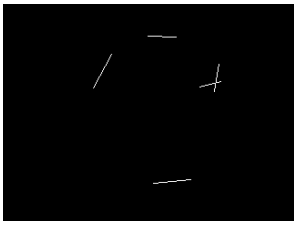


Traced direction of movements

Recognition of 9(nine)



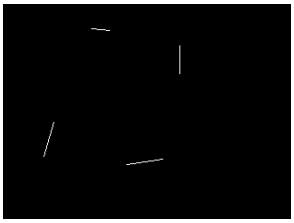

detected fingertip positions for all frames for 9 (nine)



Traced direction of movements

Recognition of 0 (zero)


detected fingertip positions for all frames for 9 (nine)



Traced direction of movements

Figure 7. Detected positions of fingures and traced direction of movements

## V. CONCLUSION

Throughout the paper a new method of the numeric character recognition system was shown. The main feature of this method is it can detect the numeric characters using a simple camera arrangement. Unlike other algorithms this method did not involve unnecessary computation and did not limit the processing time. The movements were recorded to recognize numeric characters. NO false recognition was given by the proposed system unless one missed the sequence of movements. Again as the consecutive similar movement is ignored it is independent of frame rate at which the video was taken. From the experiments, it was shown that the proposed system can process approximately 20 frames per second while the input is taken as 30 frames per second. So the proposed method is suitable for real time applications.

## REFERENCES


[1] Wengang Zhou, Houqiang Li, Yijuan Lu and Qi Tian "Principal Visual Word Discovery for Automatic License Plate Detection," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 4269 – 4279, September 2012.

[2] Rob G. J. Wijnhoven and Peter H. N. de With "Identity Verification using Computer Vision for Automatic Garage Door Opening," *IEEE Transactions on Consumer Electronics*, vol. 57, No. 2, pp. 906 – 914, May 2011.

[3] R. Plamondon and S.N. Srihari, "On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 68-89, Jan. 2000.

[4] A. Vinciarelli, S. Bengio, and H. Bunke, "Offline Recognition of Unconstrained Handwriting Texts Using HMMs and Statistical Models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 709-720, June 2004.

[5] A.L.I. Oliveira, E.R. Silva and C.A.B. Mello "Handwritten Digit Segmentation in Images of Historical Documents with One-Class Classifiers" *In the proc. of 20th IEEE International Conference on Tools with Artificial Intelligence*, vol. 2, pp. 41 – 44, Nov 2008.

[6] Z. Dong, Wejinya, U.C. S. Zhou ; Q. Shan ; Li W.J. ," Real-time written-character recognition using MEMS motion sensors: Calibration and experimental results," *IEEE International Conference on Robotics and Biomimetics*, pp. 687-691 February 2009.

[7] Md. Mehedi Hasan, Arifur Rahaman, Mirza Md. Shahriar Maswood, Md. Mostafizur Rahman"An Improved Method for 7- Segment Numeric Character Recognition*,*"*IEEE International Conference on Informatics, Electronics and Vision*, pp. 1–6, May 2013.

[8] J. Gllavata, R. Ewerth, and B. Freisleben, "Text detection in images based on unsupervised classification of high-frequency wavelet coefficients," *In the Proc of 17th International Conference on Pattern Recognition*, Cambridge,pp. 425–428, U.K. 2004.

[9] C. Mancas-Thillou and B. Gosselin, "Spatial and Color Spaces Combination for Natural Scene Text Extraction," *In the Proc. of IEEE International Conference on ImageProcessing*, pp. 985–988, Oct. 2006.

[10] Simone Marinai, Marco Gori and Giovanni Soda, "Artificial Neural Networks for Document Analysis and Recognition,"*IEEE Transactions on Pattern Analysisand Machine Intelligence*, vol. 27, no. 1, pp. 23-35, January 2005.

[11] S. Zhou, Z. Dong, Li. W.J., C. P. Kwong,"Hand written character recognition using MEMS motion sensing technology," *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 1418-1423, July 2008

[12] Zhou, Q. Shan, F. Fei, Li. W.J., C.P. KwongB. Meng, Chan, "Gesture recognition for interactive controllers using MEMS motion sensors," *4th `IEEE International Conference onNano/Micro Engineered and Molecular Systems,* pp. 935-940, January 2009

[13] A. Elnagar and R. Alhajj, "Segmentation of connected handwritten numeral strings," *Pattern Recognition*, vol. 36, no. 3, pp. 625 – 635, 2003.

[14] R. Nopsuwanchai, A. Biem and W. F. Clocksin, "Maximization of mutual information for offline Thai handwriting recognition," *IEEETransactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1347 – 1351, 2006.

[15] H. Lee,"*Segmentation of cursive handwritten text,"* School of Comp. Sci. Central Queensland University, QLD, Australia, 2007.

[16] Hong Lee and Brijesh Verma, "A Novel Multiple Experts and Fusion Based Segmentation Algorithm for Cursive Handwriting Recognition," *In the proc. of IEEE International Joint Conference on Neural Networks*, pp. 2994 – 2999, June 2008.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Blind Estimation and Compensation of IQ Imbalance in OFDM System

Nilanjon Chakraborty*, Md. Rashidul Kadir and Md. Alamgir Hossain
Khulna University of Engineering & Technology, Bangladesh
*E-mail: ncdhrubokuet@yahoo.com

***Abstract*—The presence of In-phase and Quadrature-phase (IQ) imbalance caused by the local oscillators in the direct-conversion Orthogonal Frequency Division Multiplexing OFDM) transceiver system severely degrade the performance of the communication system by increasing the bit error rate (BER). In this paper, a Blind Least Square Estimation (LSE) based "1 step" scheme is proposed to estimate and compensate for the IQ imbalance caused by the receiver of OFDM systems. Several computer simulations via MATLAB were performed to ensure the performance of the proposed method. Simulations show that the proposed technique outperforms the existing method.**

***Keywords*—IQ imbalance, OFDM, Blind Compensation, Least Square Estimation**.

## I. Introduction

In advanced communication technology such as in wireless communication system, efficient utilization of communication resources (bandwidth and power) is the key challenge to meet the high bit rate demand for future application. Now-a-days broadband wireless systems are needed to afford higher data rates, better redundancy, and higher user mobility while implementing lower cost, lower power consumption, and higher levels of integration. To combine all of these requirements OFDM has been adopted as the standard technology of choice for most broadband wireless standards (such as IEEE 802.16 and IEEE 802.11) [1]. OFDM has developed into a popular scheme for wideband digital communication, whether wireless or over copper wires, used in applications such as digital television and audio broadcasting, DSL broadband internet access, wireless networks, and 4G mobile communications [2]. One of the main reasons to use this technique is to increase the robustness against frequency selective fading channel or narrow-band interference [3]. Such robustness is very attractive, particularly in high-speed communication. In a single carrier system, a single fade can cause the entire link to fail, but in a multi-carrier system, only a small percentage of the sub carriers will be affected. In OFDM system the transmission band width is splitted into many narrow subchannels which are transmitted in parallel. Each subchannel is narrow enough so that it experiences a flat fading although the overall radio propagation environment is frequency selective. Orthogonality between the carriers allows their overlapping while disabling the occurrence of crosstalks [4]. Thus, a significant power saving can be achieved by using an orthogonal multicarrier technique. In a conventional OFDM system, the orthogonality between the subcarriers is achieved by means of the discrete Fourier transform (DFT). As a maximum of a subcarrier corresponds to zeros of other subcarrier, each subcarrier can be demodulated independently of the others. Error correction coding can then be used to correct for the few erroneous sub carriers [5]. There are many impairments which degrade the performance of OFDM systems such as IQ imbalance, Inter carrier interference (ICI), multipath channels, carrier frequency offset (CFO), inter symbol interference (ISI) etc. Nonlinear effects generated by the power amplifier may introduction intercarrier interference and thus destroy the orthogonality. IQ imbalance occurs due to gain and phase mismatches between I and Q branches of local oscillator [6]. In this paper, receiver was simulated according to the parameters established by the standard, to evaluate the performance and different possibilities in the implementation. Also, some considerations about forward error correction coding, synchronization and channel estimation are given oriented to improve the system performance. For accomplishing the proposed method blind estimation technique is used where some reasonable assumptions are made about the statistics of the received signal. In case of pilot-based estimation techniques two major drawbacks are found. Firstly the estimation algorithms often depend on a particular pilot structure that reduces the effective bandwidth. Secondly in case of receiver caused IQ imbalance, the transmitted pilots are affected by several additional RF impairments such as noisy fading channel, spurious signals etc.

In this paper, at first a receiver model is proposed to describe the effect of IQ imbalance in received signal. Then a new algorithm is proposed to estimate and compensate the IQ imbalance parameters and finally computer simulations are performed to ensure the improvement of performance of the proposed algorithm.

## II. System Model

In this section the mathematical modeling of IQ imbalance will be presented.

### A. Transmitter

Fig. 1 shows the transmitter model for OFDM system. The transmitted base band signal is

$$x_\nu(t) = x_i(t) + jx_q(t) \tag{1}$$

The pass band signal $s_\nu(t)$ that is transmitted is

$$s_\nu(t) = x_i(t)\cos(\omega_c t) - x_q(t)\sin(\omega_c t) \tag{2}$$

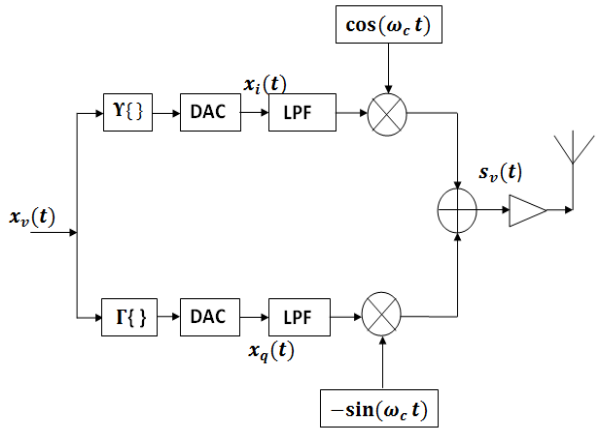



Fig. 1. Transmitter model for OFDM system

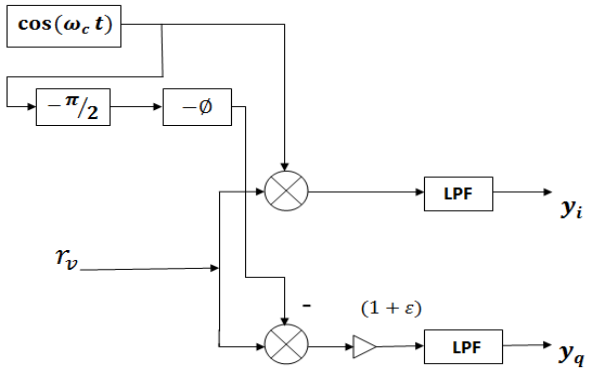



Fig. 2. Asymmetrical model for frequency independent receiver-caused IQ imbalance

### B. Receiver

At the receiver, this signal has passed through a channel with impulse response $h_{RF}(t)$. Amplification by the LNA adds noise to the signal, and thus, the received signal after the LNA is

$$r_\nu(t) = [s_\nu \otimes h_{RF}](t) + n_{RF}(t) \tag{3}$$

where, $[s_\nu \otimes h_{RF}](t)$ denotes the convolution of $s_\nu(t)$ and $h_{RF}(t)$ and $n_{RF}(t)$ is the radio frequency (RF) noise. This signal is mixed with $cos(\omega_c t)$ and low pass filtered in the in-phase branch. After the low pass filter, this signal equals

$$y_i(t) = \cos(\omega_c t) r_\nu(t) \tag{4}$$

The corresponding signal in the quadrature-phase branch equals

$$y_q(t) = -\sin(\omega_c t) r_\nu(t) \tag{5}$$

The DSP can create signal $y_\nu(t) = 2[y_i(t) + j y_q(t)]$, where the factor 2 is added for notational convenience.

IQ imbalance can be modeled into two parameters: amplitude error $\varepsilon$ and phase error $\phi$. The system block of receiver is depicted as shown in Fig. 2. After the low pass filter, the signal in in-phase branch equals

$$y_i(t) = \frac{1}{2}\Upsilon[x_\nu \otimes h](t) \tag{6}$$

corresponding signal in quadrature-phase branch becomes,

$$y_q(t) = \frac{1}{2}(1+\varepsilon)[-\sin\phi\Upsilon(x \otimes h)(t) + \cos\phi\Gamma(x \otimes h)(t)] \tag{7}$$

where, $\Upsilon$ and $\Gamma$ denote the real and imaginary part of the convolution respectively.
Then the received signal can be written as

$$y_\nu(t) = \lambda(t)[x_\nu \otimes h(t)] + \mu[x(t) \otimes h(t)]^* \tag{8}$$

where, $x_\nu(t)$ is the baseband signal, $h(t)$ is the channel impulse response and "$*$" denotes conjugate operation. Here $\lambda$ and $\mu$ are called imbalance parameters and are expressed as follows:

$$\lambda = \frac{1 + (1+\varepsilon)\exp[-j\phi]}{2} \tag{9}$$

and

$$\mu = \frac{1 - (1+\varepsilon)\exp[j\phi]}{2} \tag{10}$$

The received signal in an OFDM system in frequency domain can be written as

$$Y_k = \lambda X_k H_k + \mu X^*_{-k} H^*_{-k} \tag{11}$$

## III. IQ Imbalance Parameters Estimation

### A. Existing method

In existing estimation technique [8] while estimating IQ imbalance parameters only the real part is considered and transmitter is assumed to be noisy.
In case of receiver caused IQ imbalance, the estimated channel on $k$-th subcarrier can be simplified to

$$\hat{H}_k = \lambda_{RX} H_k + \mu_{RX} H^*_{-k} \tag{12}$$

here, dependent estimator is used where the imbalance parameters are related as $\lambda_{RX} = 1 - \mu^*_{RX}$ [9]. To maintain a linear relationship in the components of $\mu_{RX}$.
The received signal can be explicitly split up into real and imaginary parts as shown in equations (13) and (14) where, the subscripts $r$ and $i$ symbolize the real and imaginary part respectively.

The receiver only obtains a probable estimation of $\mu_{RX,r}$ and $\mu_{RX,i}$ [10]. An estimation of such kind can be based on subcarrier $k$ and $-k$ as follows

$$\begin{bmatrix} \hat{\mu}_{RX,r} \\ \hat{\mu}_{RX,i} \end{bmatrix} = \hat{R}^{-1}_{1,k} \left( \begin{bmatrix} Y_{k,r} \\ Y_{k,i} \end{bmatrix} - \hat{R}^{-1}_{2,k} \right) \tag{15}$$

where, $\hat{p}$ denotes the estimate of variable $p$. $\hat{R}_{1,k}$ and $\hat{R}_{2,k}$ depend on the channel coefficients and transmitted symbols. The original estimate of $\mu_{RX}$ can be obtained by considering both real and imaginary parts, and averaging over all subcarriers. The receiver caused imbalance parameter $\hat{\lambda}_{RX}$ can be estimated with the help of previously estimated $\hat{\mu}^*_{RX}$ [11].

$$Y_{k,r} = H_{k,r}X_{k,r} - H_{k,i}X_{k,i} + \mu_{RX,r}(-H_{k,r}X_{k,r} + H_{k,i}X_{k,i} + H_{-k,r}X_{-k,r} - H_{-k,i}X_{-k,i}) + \mu_{RX,i}(-H_{k,r}X_{k,i} - H_{k,i}X_{k,r} + H_{-k,r}X_{-k,i} + H_{-k,i}X_{-k,r}) \quad (13)$$

$$Y_{k,i} = H_{k,r}X_{k,i} + H_{k,i}X_{k,r} - \mu_{RX,r}(H_{k,r}X_{k,i} + H_{k,i}X_{k,r} + H_{-k,r}X_{-k,i} + H_{-k,i}X_{-k,r}) + \mu_{RX,i}(H_{k,r}X_{k,r} - H_{k,i}X_{k,i} + H_{-k,r}X_{-k,r} - H_{-k,i}X_{-ki}) \quad (14)$$

### B. Proposed method

Our proposed method is based on considering both real and imaginary part of imbalance parameters and the transmitter is assumed to be noise free. The new estimation algorithm then becomes:

$$\hat{\mu}'_{RX,r} = \frac{\hat{\mu}'_{RX,k,r} + \hat{\mu}'_{RX,-k,r}}{2} \quad (16)$$

After simplification, the expression for $\hat{\mu}'_{RX,r}$ can be written as

$$\hat{\mu}'_{RX,r} = \mu_{RX,r} - \frac{2\mu^2_{RX,i}}{1 - 2\mu_{RX,r}} \quad (17)$$

The error in $\mu_{RX,r}$ becomes

$$\Xi_{\mu'_{RX,r}} = \hat{\mu}_{RX,r} - \mu_{RX,r} \quad (18)$$

By further simplification, the expression for error becomes

$$\Xi_{\mu'_{RX,r}} = \frac{-2\mu^2_{RX,i}}{1 - 2\mu_{RX,r}} \quad (19)$$

Again the expression for $\hat{\mu}'_{RX,i}$ can be simplified as

$$\hat{\mu}'_{RX,i} = \mu_{RX,i} + \frac{2\mu_{RX,r}\mu_{RX,i}}{1 - 2\mu_{RX,r}} \quad (20)$$

(17) and (20) express the relation between the actual and the estimated value of the required IQ imbalance parameters. The error in this case can be expressed as

$$\Xi_{\mu'_{RX,i}} = \frac{2\mu_{RX,r}\mu_{RX,i}}{1 - 2\mu_{RX,r}} \quad (21)$$

So, the total error for $\mu_{RX,r}$ can be obtained by combining the two error expressions of (19) and (21) above as

$$\Xi_{\mu'_{RX}} = \Xi_{\mu'_{RX,r}} + j\Xi_{\mu'_{RX,i}} \quad (22)$$

After further simplification, the expression for error leads to

$$\Xi_{\mu'_{RX}} = \frac{2j\mu_{RX,i}\mu_{RX}}{1 - 2\mu_{RX,r}} \quad (23)$$

Therefore, the estimated value for $\mu_{RX}$ will have an error as described by (23). In the previous Subsection III-A only the quadratic part of the expression (17) and (20) are taken as error term. Proper expression for the IQ imbalance as a function of the estimated values can be obtained by considering both the quadratic and linear portion of the quadratic equations. In that case, (16) and (19) are combined to establish the following two equations

$$\mu_{RX,r} = \frac{1}{2} - \frac{\mu_{RX,i}}{2\hat{\mu}'_{RX,i}} \quad (24)$$

and

$$\mu^2_{RX,i}(-\frac{1}{2} - 2\hat{\mu}^2_{RX,i}) + \mu_{RX,i}(-\hat{\mu}'_{RX,i}\hat{\mu}'_{RX,r} + \frac{1}{2}\hat{\mu}'_{RX,r}) = 0 \quad (25)$$

Solving (25) results either $\mu_{RX,i} = 0$ or

$$\mu_{RX,i} = -\frac{(2\hat{\mu}'_{RX,r} - 1)}{1 + 4\hat{\mu'}^2_{RX,i}}\hat{\mu}'_{RX,i} \quad (26)$$

Substituting (26) into (24) gives

$$\mu_{RX,r} = \frac{2\hat{\mu'}^2_{RX,i} + \hat{\mu}'_{RX,r}}{1 + 4\hat{\mu'}^2_{RX,i}} \quad (27)$$

This solution provides the compensated value of receiver caused IQ imbalance parameter $\mu_{RX}$ according to proposed method. Using the relation $\lambda_{RX} = 1 - \mu^*_{RX}$, $\lambda_{RX}$ can also be estimated in one step.

## IV. Simulation Result and Discussion

To evaluate the proposed algorithm, a typical OFDM system based on WLAN is considered. Radio frequency signal is shown in Fig. 3. Here, $1/4$ data coding system is used and sampling frequency is taken as 2.5 MHz. Channel is assumed to be known and number of tapping is taken as 1.

Applying the proposed method the performance characteristics of OFDM system is shown in Fig. 4. This simulation is done under the influence of IQ imbalance. It can be observed from the figure that the performance increases when

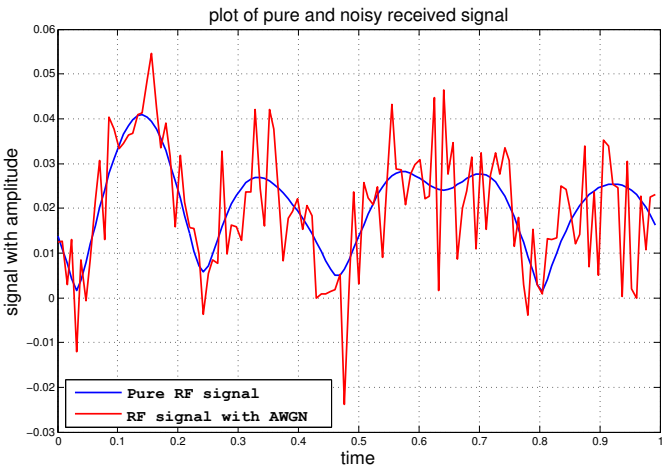



Fig. 3. RF signal with and without AWGN

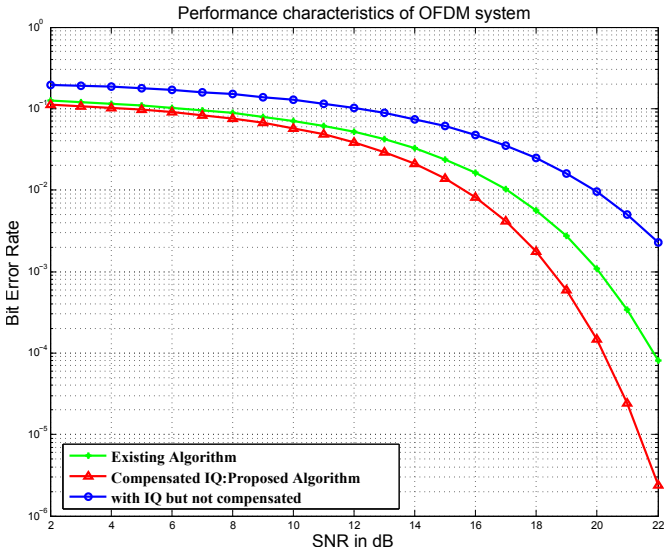



Fig. 4. Performance improvement of OFDM system

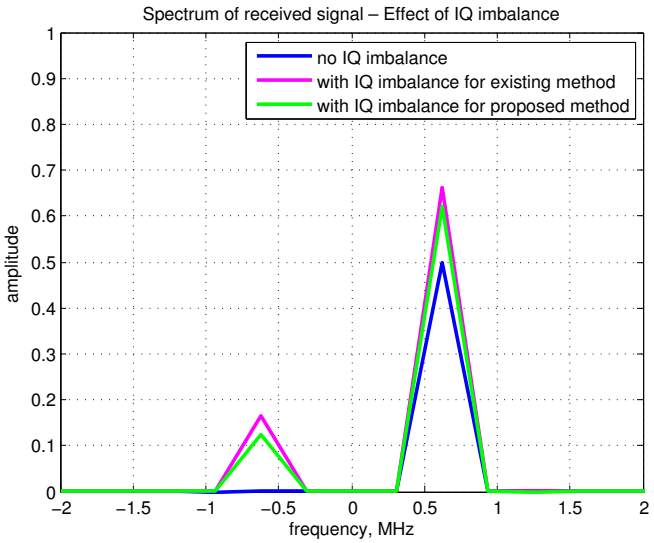



Fig. 5. Effect of IQ imbalance

IQ imbalance is compensated. Then this model is compared with an existing complex algorithm. From Fig. 4 it can be seen that the proposed method outperforms the existing algorithm and the main concern here is that complexity of the existing model is reduced to a greater extend. Here "1 step" approach is applied where no iteration is required. Hence computational complexity is also reduced as the number of steps are less. Here blind compensation technique is used so as not to limit the bandwidth imposed by pilot estimation technique. The improvement of estimating the IQ parameters by using the proposed method are depicted in Table. I and Fig. 5.

Thus the received signal not only contains the desired signal but also the copy of it but mirrored in frequency as shown in Fig. 5. Due to this interference of desired signal IQ mismatch occurs. In this process the received signal gets weaker than the original signal and poor performance of the whole system is observed. From Fig. 5 it can be observed that our proposed method has reduced the effect of IQ imbalance.

## V. Conclusion

Blind signal processing techniques are very suitable for estimating and compensating unwanted RF impairments in direct conversion receiver like IQ imbalance. A new concept is applied here which is mainly the combination of LSE and "1 step" method. From the simulation results it is easily obtainable that our proposed method has updated the BER performance of receiver caused IQ imbalance. In future the task will be to extend the method for the joint estimation of IQ imbalance, CFO and channel response in OFDM systems. Moreover, pilot based estimation techniques are also to be improved. Further research related to this work can be carried out to investigate the influence of IQ imbalance and other non-linear distortion effects on OFDM system. BER performance due to variation of channel length, number of symbol, bit rate can also be investigated.

TABLE I
Numerical Values of IQ Parameters

| Existing | Proposed |
|---|---|
| $\lambda_{RX} = 0.6567$ | $\lambda_{RX} = 0.4969$ |
| $\mu_{RX} = 1.6567$ | $\mu_{RX} = 0.5031$ |

## References


[1] A. Schuchert, R. Hasholzner, and P. Antoine, "A novel IQ imbalance compensation scheme for the receiption of OFDM signals," *IEEE Trans. Consum. Electron.,* vol. 47, no. 3, pp. 313–318, Aug. 2001.

[2] Asad A. Abidi, "Direct-conversion radio transceivers for digital communications," *IEEE J. Solid-State Circuits*, vol. 30, no. 12, pp. 1399–1410, Dec. 1995.

[3] Behzad Razavi, "Design Considerations for Direct-Conversion Receivers," *IEEE Trans. Circuits Syst. II*, vol. 44, no. 6, pp. 428–435, Jun. 1997.

[4] Marcus Windisch, "Estimation and Compensation of I/Q Imbalance in Broadband Communications Receivers," Ph.D. thesis, Technische University, Dresden, Germany, Mar. 2007.

[5] A. Tarighat and A. H. Sayed, "Joint compensation of transmitter and receiver impairments in OFDM systems," *IEEE Trans. Wireless Commun.,* vol. 6, no. 1, pp. 240–247, Jan. 2007.

[6] J. Tubbax, B. Come, L. Van der Perre, S. Donnay, M. Engels, H. D. Man, and M. Moonen, "Compensation of IQ imbalance and phase noise in OFDM systems," *IEEE Trans. Wireless Commun.,* vol. 4, no. 3, pp. 872–877, May 2005.

[7] A. Seyedi and G.J. Saulnier, "General ICI self cancellation for OFDM systems," *IEEE Trans. Veh. Tech.*, vol. 54, no. 1, pp. 198–210, Jan. 2005.

[8] Admar Schoonen, "IQ imbalance in OFDM wireless LAN Systems," M. Eng. thesis, Eindhoven University of Technology, Eindhoven, Netherland, Jan. 2006.

[9] C. L. Liu, "Impacts of I/Q imbalance on QPSK–OFDM–QAM detection," *IEEE Trans. Consum. Electron.,* vol. 44, no. 3, pp. 984–989, Aug. 2004.

[10] S. Alamouti, "A simple transmit technique for wireless communications," *IEEE J. Sel. Areas Commun.,* vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[11] A. Tarighat and A. H. Sayed, "MIMO OFDM receivers for systems with IQ imbalances," *IEEE Trans. Signal Proc.,* vol. 53, no. 9, pp. 3583–3596, Jun. 2004.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Remote Metering of Electricity Consumers Using Frequency Division Multiplexing over XLPE Power Transmission Cable

Syed Mohammad Taukir Imam[1], Abu Mohammad Zafar Sadeque [2], Asif Islam[3], and Mohammad Shariful Islam[1,*]
[1]Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Bangladesh
[2]Electrical and Computer Engineering, Southern Illinois University, USA
[3]Planning and Development Division, Power Grid Company of Bangladesh (PGCB) Ltd., Bangladesh
[*] E-mail: sabuz03@gmail.com

*Abstract*—**In Bangladesh, the electricity consumers have analog energy meters in their home premises as no automated remote metering is deployed yet. These meters need to be inspected on a monthly basis to prepare invoice for electricity consumption. This requires an extra manpower along with the waste of time and this of course paves a way to corruption. Remote Metering of Electricity Consumers is an approach to resolve these menaces by employing a digital meter at the consumer end and developing a technique to monitor the energy consumption of each consumer from the substation using frequency division multiplexing over the existing XLPE power line cables. This paper represents a technique to transmit and measure the energy data which has been consumed by the subscribers using the existing power distribution system. The proposed technique is based on frequency division multiplexing (FDM) and has similarity with power line cable communication (PLCC).**

***Keywords—Remote Metering, Power Line Communication, XLPE Cable, Bandpass Filter, Analog to Digital Conversion, Frequency Division Multiplexing.***

## I. INTRODUCTION

Measurement of consumed electric energy and invoicing for bill collection is an integral part of power management system. The user can be billed for his consumption [1] after a certain period or the power generation and distribution authority can also make a prepaid arrangement. In both cases it's important that consumed electrical energy by a subscriber can be accurately measured and that data can be effectively communicated up to the billing or control center which can be automatic or needs concerned authority's inspection. Though quite a many digital smart meters have been developed for energy measurement in last few years, in Bangladesh it is not yet deployed in large scale. Research is also being carried out all over the world for efficient automatic transmission of measured data up to the billing center which can be wired or wireless.

Campos et al. [10] have evaluated economic impacts of implementation of smart meters and their future prospects at low-voltage subgroup using the available technologies, whereas, Asadi et al. [11] have developed an algorithm using PSO to maximize impacts on real-time smart grid. Wenshu et al. [12] have reported that, Orthogonal Frequency Division Multiple Access (OFDMA), widely adopted in many wired and wireless communication systems, is multicarrier in nature and results in severely high peak-to-average power ratio (PAPR), which is undesirable for low-cost terminal equipment. On the other hand, Single Carrier-Frequency Division Multiple Access (SC-FDMA), a promising alternative to OFDMA, can greatly reduce PAPR while achieving similar performance and complexity. Arif et al. [13] have explored the possibilities of developing GSM based smart meters, and Madueno et al. [14] have assessed the potential of GSM/GPRS/EDGE to operate as a dedicated network for smart meters and estimated the limits of GSM. In contrast, Li et al. [15] have devised efficient smart meter data collection by exploring the secondary spectrum market in cellular networks. Jixuan et al. [16] have outlined some smart meter's aspects and functions of smart meter. They have introduced two basic types of smart meter system's communication technologies: Radio Frequency (RF) and Power Line Carrier (PLC) and recent advances with regard to these two technologies.

Wireless communication is costly in terms of the additional transmitters and protection methods against noise. Power line communication is a better technique for this sort of activity. But in Bangladesh currently no automated remote metering is deployed. But implementation of such a system can improve the total energy distribution process remarkably. In this paper, a tentative technique has been described to measure and transmit consumed energy data of subscribers using the existing power distribution system similar to power line communication [2]-[5]. The discernible contribution of this work is that, it would aid to the widespread rollout of smart meter in Bangladesh, which is very costly till date.

## II. THEORY AND METHODOLOGY

In electrical consumer system where the supply is single phase AC (Alternating Current), the consumed energy at a given time (t) can be expressed as in (1).

$$E = P\,t \tag{1}$$

where P is the power used by the consumer in watt and it can be measured by (2).

$$P = V_{rms}I_{rms}\cos\theta \quad (2)$$

Therefore the total energy consumed [6] at an interval (t) is calculated by (3),

$$E = (V_{rms}I_{rms}\cos\theta)t \quad (3)$$

where, $V_{rms}$ is the terminal RMS voltage at consumer end, $I_{rms}$ is the RMS current value being consumed by customer, $\cos\theta$ is the power factor and t is an interval of time.

To determine the RMS (Root Mean Square) value of an AC voltage, the peak voltage value is detected first for a cycle. Because the governing equation for the RMS value is related to the peak value (4)

$$V_{rms} = V_{peak}/\sqrt{2} \quad (4)$$

The measurement of current flowing through the consumer load can be measured by a CT (Current transformer). But the task has been accomplished in this project using Ohm's law with existing circuitry. A small resistance has been added (small enough compared to the load) in series with the load so that the total current flowing in load doesn't change and the same current flows through this small resistance. As a result a small voltage is dropped across this resistance. This voltage has been measured and later on has been divided by the preset resistance and which is independent of the load as per Ohm's law (5).

$$V = IR \quad (5)$$

And thus $I_{rms}$ can also be measured (6).

$$I_{rms} = I_{peak}/\sqrt{2} \quad (6)$$

The time difference between the $V_{peak}$ and the following $I_{peak}$ (considering Inductive load and thus Current lagging) will provide the value of power factor $\cos\theta$.

In this paper a prototype is described that can measure the consumed energy at the consumer end and then transmit this energy information of two users after proper modulation and multiplexing to the substation, which is 1km away from the consumers. The real life scenario is that the voltage at the home user level is 220 volts 50Hz in Bangladesh. But in the laboratory a model having a supply voltage of ±5 volt has been developed. As microcontroller is used for measurement and control, digital data is required for processing. But the supply voltage from the function generator in the laboratory that is simulated as the generator of real life is purely analog and time varying. A microcontroller is not able to read and process this voltage values directly. For this reason a measure has been adopted to convert these analog values into digital first using ADC (Analog to Digital Converter).

In this paper, an 8086 microcontroller [7] based digital meter to be placed at consumer's end has been devised to measure the electric energy consumption. The first task is converting the analog data at consumer end into digital data via ADC for the competitive immunity of digital signal against noise in case of transmission and processing through microcontroller. As digital signal is discrete in time domain, Sampling is required for analog to digital conversion. As per Nyquist theorem, sampling frequency greater than or equal to twice of the original signal frequency is appropriate to reconstruct it properly. As power signal is 50Hz, therefore a sampling frequency of 100Hz or more will work in this case. In this prototype, a sampling frequency of 667 kHz (ADC arrangement) has been used.

An n-bit ADC is able to determine 2n levels of equidistant amplitudes. An 8-bit ADC defines 28 or 256 discrete amplitude levels denoting them from 00000000 to 11111111, a combination of 8 bits. The highest and lowest level of the input analog signal is determined first. Then the ADC divides the total span (difference between the highest and lowest value) into 256 equidistant levels with an 8-bit representation (7).

$$V_a=(SV_a)/256 \quad (7)$$

where $V_a$ is input analog value, S is span and $V_d$ is corresponding decimal output value.

Any analog value between two distinct levels are not encoded as it is, rather its converted to the next upper or the previous lower level posing a potential error. This error is termed as Quantization Error. It is reduced as the resolution or the digital levels are increased. Another problem is that ADC can't measure negative values. But function generator used supplies an AC sinusoidal pulse from +5V to -5V (Maximum) of a frequency of 50Hz.So the negative part had to be eliminated completely to measure the voltage with ADC. So the voltage shape is shifted with a DC offset of 5V to eliminate the negative part completely. The problem is that, the peak value now reaches 10V. But the ADC used can measure a maximum of 5V.To accomplish the measurement by not touching the source, a voltage divider is arranged and thus cut the voltage to half of its peak and then shifted it with a DC offset. Now the output of this voltage divider needs to be shifted by 2.5V to make it a pure time varying DC by keeping the AC variations intact. A DC offset of 2.5V had to be added to this signal. But to avoid loading phenomena the voltage has taken at the divider end through a buffer and then added with a 2.5V DC. This signal has been sampled and fed to the microcontroller through ADC for determining the $V_{peak}$ by continuous monitoring and comparison by a C based programming code.

As the current flowing through the complete prototype arrangement was too low, it couldn't be measured accurately in this project. Therefore transmitting the voltage information over the transmission line to the substation end has been focused. To accomplish this, the RMS voltage value was converted into a 10-bit number that was enough to represent a huge value practical for consumed energy at a certain period. In the prototype as the voltage reading is a value with a decimal point i.e. 2.86 volt (Table II), the result was simply multiplied by 10 to take the first two significant digits by eliminating the decimal point. After taking the value multiplied by 10, the

value was rounded up to its nearest integer. Suppose the voltage reading indicates 2.86V and after multiplication by 10, it was 28.6. Then this value is rounded to 29 and converted to a 10 bit stream of 1 and 0 through the microcontroller by a method called type casting in programming. This value is transmitted instead of 2.86 V.

When the data is computed and ready to be transmitted [8] over the power line then the most important thing is Modulation as digital signal form has very low energy content. The basic objective of modulation is to transmit this digital information over a powerful carrier or high frequency signal that can keep the information intact till the end of line considering interference, attenuation etc. The power depends on frequency as follows (8):

$$P \propto f^2 \tag{8}$$

where P is power and f is signal frequency. In this prototype, a digital modulation technique ASK (Amplitude Shift Keying) is used where the amplitude of a high frequency carrier signal has been varied according to the digital information. In lab, the job of modulation was accomplished by a simple N-MOSFET. MOSFET acts like a switch according to its gate pulses. The carrier signal of high frequency was simply connected at the Drain of the MOSFET. The Source remained always grounded. And the gate is driven from the microcontroller output which emits the digital signal or information of energy (voltage in this prototype).In this case if the Gate gets 1 the output is zero and if the Gate gets 0 then it outputs the carrier signal which is inverted information of the original. This can easily be detected at the reception end by inverting it once more. Here in the lab, it was simulated for 2 users and as this is a FDM (Frequency Division Multiplexing) based prototype, two different carrier frequencies for 2 users was chosen for communication over the same channel. The chosen frequencies are 600Hz and 1.5 kHz that ensures a good band separation to avoid adjacent channel interference during transmission.

After the two set of modulated information signal [9] of energy reading are ready at the both consumer unit at different frequencies, the next job is to transmit them simultaneously over the same channel. In this case, an Adder circuit was made that adds these two different frequency signals and outputs a combined one for transmission. Bit rate is a crucial parameter [8]-[9] in case of serial communication for reconstructing the signal at the recipient end. In this case a bit rate of 20 bits per second has been used to ensure synchronous communication which is far lower than the modulation frequency. A sampling frequency similar to the bit rate will be used at the reception end to get the transmitted information again. The transmission line is the most important part [6] in this prototype as this is used as the only transmission medium. This approach reduces the establishment of a separate medium. And this idea can be readily implemented without incurring any additional cost for communication in Bangladesh.

In laboratory, a transmission line was simulated with resistance and capacitance that resembles a standard 11KV transmission line. The resistance and inductance of 11KV overhead lines are not available in any technical reference of any cable manufacturing company. After studying a few technical references, the characteristic of the BRB XLPE (insulated) cables was finally selected. According to the technical reference of XLPE cables from BRB Cable Industries Limited the following parameters were chosen as table I:

TABLE I. TYPICAL 11KV XLPE CABLE PARAMETERS/KM

| $V_{base}$ | $I_{base}$ | Resistance per km ($R_{pu}$) | Inductance per km ($L_{pu}$) | Capacitance per km ($C_{pu}$) |
|---|---|---|---|---|
| 11 kV | 157 A | 0.7270 ohm | 0.492 mH | 0.21 μF |

Therefore,

$$Z_{base} = 11\ kV/157A = 70.06$$

The per unit values should be,

$$R_{pu} = 0.7270/70.06 = .01\ ohm$$
$$L_{pu} = 0.492/70.06 = 7.02\ \mu H$$
$$C_{pu} = 0.21/70.06 = 2.99\ pF$$

In this prototype,

$$V_{base} = 5\ V \text{ and } I_{base} = 0\ .1\ A$$

Therefore,

$$Z_{base} = 5/0.1 = 50\ ohm.$$

In this prototype (Fig. 2),

$$R = .01 \times 50 = 5\ ohm$$
$$L = 7.02\ \mu H \times 50 = 0.35\ mH$$
$$C = 2.99\ pf \times 50 = 0.150\ nF$$

As working frequencies in this case are very low of 600Hz and 1.5 kHz, the capacitance of the cable was neglected. Here in the total arrangement, R1 and L4 is simulated subscriber load representation (Fig. 2), and R2 and L2 is Transmission line equivalent representation (Fig. 2)

The information signal from the adder is fed into the transmission line through a capacitance C4 (Fig. 2) .Now the most important task was to ensure that both the data signal carrying 600Hz and 1.5 kHz and the low frequency power signal is flowing through the transmission line in opposite direction without interfering with each other.

The Fourier transform [6] of this signal in the transmission line revealed the existence of all the frequencies (Fig. 3).All the signals travelled through the designed transmission line

successfully. Now another concern was that the high frequency data signals shouldn't enter the substation that can cause damage to the power generator. So an inductance L1 was connected that restricts the high frequency signal to enter the transformer or substation (Fig. 2). Before reaching the generator the high frequency signals were filtered out through a wave trap. Here the capacitance C2 of 600pf works as a wave trap that bypassed the data signals of high frequency (Fig. 2). Fourier transform of the captured signal at generator end depicts that the high frequency data signals wasn't disturbing the substation voltage, which was perfectly a sinusoidal signal of 50 Hz (Fig. 4). And contained no 600Hz or 1.5 kHz signal. With the help of the wave trap, the data signal was separated at the substation end.

At the substation end all the signals of 50Hz, 600Hz and 1.5 kHz were received. The latter two components were required to extract the data information of the consumers. For this purpose filters were designed to separate the desired signal and curtail the others. Since 600Hz and1.5KHzwere chosen for the communication, the band separation was almost 900Hz which was quite a handsome separation that provides the flexibility to work with a band pass filter. Two High Q narrow band pass filter were designed as per requirement. The Q factor is an indicator of the bandwidth in case of a band pass filter.

In typical narrow band pass filter, the bandwidth is governed by (9).

$$B = f_r / Q \tag{9}$$

where, fr is the resonating frequency and Q is the quality factor.

Here resonating frequency is the modulating frequency. Bandwidth is related with the design parameters as in (10).

$$B=0.1591/(RC) \tag{10}$$

From this equation, R can be calculated for any given C. In the circuit (Fig. 1),

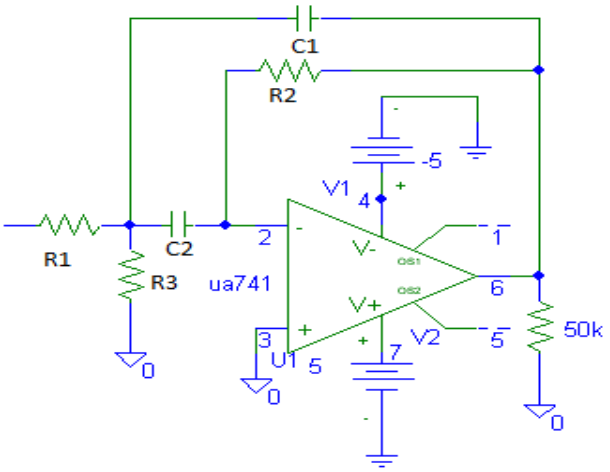



Fig. 1 Schematic diagram of a Bandpass Filter

$$C1 = C2 = C$$

$$R1 = R$$

$$R2 = 2R$$

The most important parameter is R3 which determines the resonating frequency and that can be calculated as in (11) (Fig. 1).

$$R3=R/(2Q^2-1) \tag{11}$$

2 narrow Band-pass Filter were designed with a Q of 3 and 600Hz and 1.5 kHz as resonating frequencies (Table IV and V). After separating the data signals with appropriate filter, an envelope detector was arranged as the demodulator to extract the data from the carrier for each signal. It detects the envelope which is literally the digital information with a time constant of RC that is 4.7ms. The signal amplitude further deteriorates after passing through the detector. For the final data reading through the microcontroller a perfect 1 and 0 representation was needed, which was almost improbable after all these processes encountered. The signal was corrupted by noise, attenuated and was no more in a square pulse shape. To retain the square pulses again a Schmitt trigger can be used. But the result is not that bright due to very low level corrupted signal at the input.

The final job to retrieve the information was to sample the output from the level detector to get the stream of 1s and 0s. This can be done in the substation server. In this prototype, the substation server was simulated with an 8086 based microcontroller as earlier. The two output from the two separate filters followed by two separate envelop detector and level detector were connected to the two pins of the substation side microcontroller. By regular monitoring the two pins at 20 Hz would reveal the 10 bit digital information. The sampling should be done at the midpoint of the bits to ensure not to miss an edge.

## III. Results and Analysis

The equivalent prototype developed in this case can be described by the below circuit (Fig. 2).

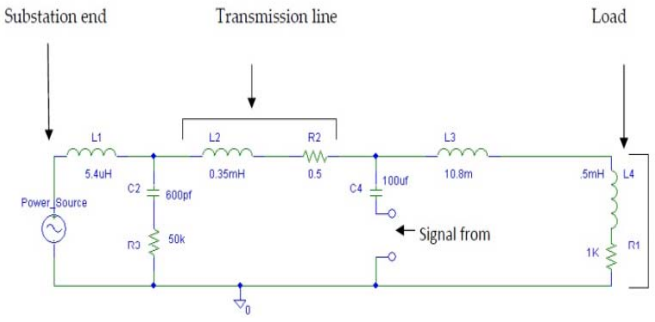



Fig.2 A schematic representation of the prototype developed

The digital meter developed in this case measured terminal voltage and current for two simulated consumers which were compared with the standard meter reading for a comparative

analysis (Table II and III) that shows the authenticity of the digital meter developed. In this case consumer load has been simulated as purely resistive and power can be calculated by (2) considering power factor as (1).

TABLE II. MEASURED DATA FOR USER 1 AT THE USER END ( LOAD = 1K OHM )

| Lab Replica Data | | | Real Meter data | |
|---|---|---|---|---|
| ***Voltage across load (Volt)*** | ***Current (mA)*** | ***Calculated Power (mW)*** | ***Voltage across load (Volt)*** | ***Current (mA)*** |
| 1.1 | 1.0 | 1.1 | 1.12 | 1.09 |
| 1.5 | 1.4 | 2.1 | 1.54 | 1.5 |
| 1.7 | 1.6 | 2.7 | 1.90 | 1.85 |
| 2.6 | 2.4 | 6.2 | 2.65 | 2.58 |
| 3.3 | 2.4 | 7.92 | 3.47 | 3.38 |

TABLE III. MEASURED DATA FOR USER 2 AT THE USER END ( LOAD = 470 OHM )

| Lab Replica Data | | | Real Meter data | |
|---|---|---|---|---|
| ***Voltage across load (Volt)*** | ***Current (mA)*** | ***Calculated Power (mW)*** | ***Voltage across load (Volt)*** | ***Current (mA)*** |
| 1.1 | 2.2 | 2.4 | 1.13 | 2.27 |
| 1.5 | 2.4 | 3.6 | 1.54 | 3.09 |
| 1.7 | 2.1 | 3.6 | 1.85 | 3.72 |

Fourier analysis of the captured signal in the transmission line reveals the existence of the high frequency data signal and the low frequency power signal (Fig. 3). That makes it evident that data signal modulated with high frequency can co-exist with power signal in power line.

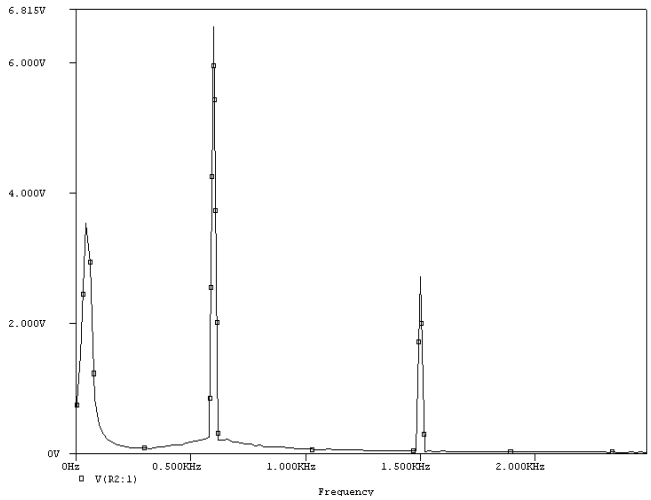



Fig. 3 The Fourier analysis of the signal reveals the existence of 600Hz and1.5kHz signal along with the 50Hz power signal in the transmission line

The capacitance C2 of 600pf works as a wave trap that bypasses the data signals of high frequency and keeps the generator end secured (Fig. 2). Fourier analysis of the signal at the substation generator end shows that it is purely a 50Hz sinusoidal signal that ensures uninterrupted power distribution (Fig. 4).

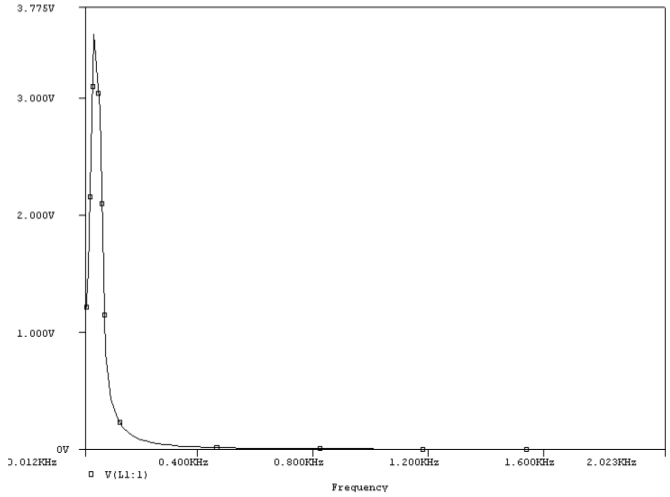



Fig. 4 Fourier analysis depicts that the high frequency data signals are eliminated and only contains 50Hz component of power signal

Two filters have been designed for fetching the data signals separately at substation control point for further processing and detection using equation 9, 10, 11. The two high Q bandpass filter design parameter specification are given below.

### *A. 600Hz Filter*

The design specifications are as follows (Table IV)

TABLE IV. DESIGN SEPCIFICATION PARAMETER OF A NARROW BANDPASS FILTER RESONATING FREQUENCY OF 600HZ

| Parameter | Values |
|---|---|
| $f_r$ | 600Hz |
| Q | 3 |
| B | 200Hz |
| R1 | 8K |
| R2 | 16K |
| R3 | 470 |
| C1, C2 | 0.1µF |

### *B. 1.5KHz Filter*

The design specifications are as follows (Table V)

TABLE V. DESIGN SEPCIFICATION PARAMETER OF A NARROW BANDPASS FILTER RESONATING FREQUENCY OF 1.5KHZ

| Parameter | Values |
|---|---|
| $f_r$ | 600Hz |
| Q | 3 |
| B | 500Hz |
| R1 | 3.2K |
| R2 | 6.4K |
| R3 | 190 |
| C1, C2 | 0.1µF |

The Fourier transform of the received signal at the substation end after the wave trap reveals the existence of both the carrier frequency in Fig. 5. This indicates successful transmission of the high frequency data signals.

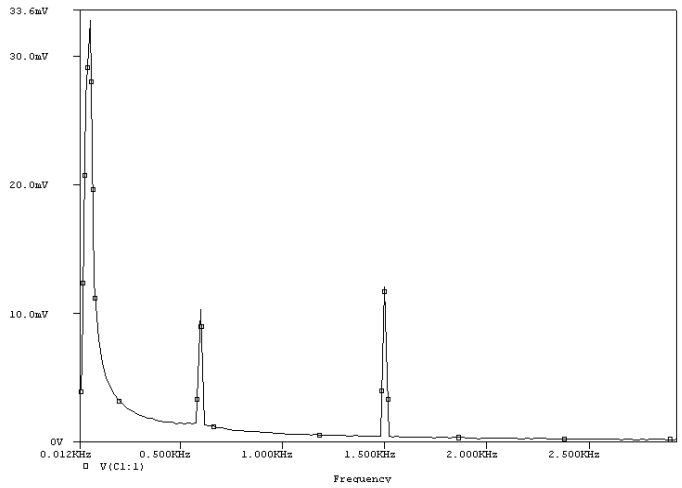



Fig. 5 Fourier analysis after the wave trap at substation end depicts the existence of 50Hz power signal along with 600Hz and1.5KHz data signals

It got almost three signals after the wave trap but at very low amplitude which is certainly a problem couldn't be solved in this prototype. That can be solved by amplifying the signal, it was assumed. So far the targets were achieved in case of transmission. That implied the success of such an arrangement in real life 11 KV overhead conductors.

## IV. CONCLUSION

In this prototype, two different consumer's end voltage data was measured, encoded, modulated and sent over XLPE cable successfully though the reception and decoding wasn't quite successful due to attenuation and interference in this case. But in real life scenario, this challenge can be overcome as all the relevant voltage levels will be higher. In fact, in real cases, for thousands of users, FDM won't bring the best of the results. In that case Code Division multiplexing can be used for effective communication over transmission lines. In this model XLPE transmission cable has been considered as most of the overhead power cables are XLPE in Bangladesh. But the theory is applicable to any conductor. For a proper simulation the actual specification of lines, are needed and in this case, accurate manufacturer data were available only for XLPE and were collected from BRB cables, Bangladesh. This paper is intended to assist in further development of power line communication to cater automatic remote metering of electricity consumers in Bangladesh using the existing infrastructure for a better power distribution system and real time monitoring. The microcontroller based meter deployed at consumer end in this paper can store historical data of a consumer which can be relayed to the substation in case of line fault or protection relay operation. It can also be used for audit purpose later, if required.

## ACKNOWLEDGMENT

The authors would like to acknowledge the contribution of CES (Centre for Energy Studies) of BUET (Bangladesh University of Engineering & Technology) for their logistic and simulation support. They are also grateful to RTS (Research & Testing System) department of PGCB (Power Grid Company of Bangladesh) Ltd. for their data support regarding power and communication cables.

## REFERENCES

[1] M. H. Albadi and E. F. El-Saadany, "A summary of demand response in electricity markets" *Electric Power Systems Research*, vol. 78, issue 11, Nov. 2008, pp. 1989–1996.

[2] F. Yang, W. Ding and J. Song, "Non-intrusive power line quality monitoring based on power line communications," in *Proc. IEEE 17th International Symposium on Power Line Communications and Its Applications*, March 2013, pp. 191-196.

[3] J. Liu, B. Zhao, J. Wang, Y. Zhu and J. Hu, "Application of power line communication in smart power Consumption," in *Proc. IEEE International Symposium on Power Line Communications and Its Applications*, March 2010, pp. 303-307.

[4] A. Islam, S. I. Khan and A. Hoque "Detection of mechanical deformation in old aged power transformer using cross correlation co-efficient analysis method," *Energy and Power Engineering*, vol. 3, no. 4, pp. 585-591, Sep. 2011.

[5] C. Y. Park, K. H. Jung and W. H. Choi, "Coupling circuitry for impedance adaptation in power line communications using VCGIC," in *Proc. IEEE International Symposium on Power Line Communications and Its Applications*, April 2008, pp. 293-298.

[6] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing*, Prentice-Hall, India, 2006.

[7] D. V. Hall, *Microprocessor and Interfacing*, Tata McGraw-Hill, New Delhi: India, 2007.

[8] Y. J. Zhang and K. B. Letaief, "Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems," *IEEE Trans. Wireless Communication*, vol. 3, no. 5, pp. 1566–1575, Sep. 2004.

[9] W. Xu, C. Zhao, P. Zhou, and Y. Yang, "Efficient adaptive resource allocation for multiuser OFDM systems with minimum rate constraints," in *Proc. IEEE Int. Conf. Communication*, Jun. 2007, pp. 5126–5131.

[10] R. A. Campos, D. D. Grego, and D. S. Ramos, "Efficient pricing: An opportunity for cost reduction system expansion to benefit the consumer as bounce deployment of smart metering," in *IEEE PES Conference on Innovative Smart Grid Technologies Latin America (ISGT LA)*, 15-17 April 2013, pp.1-7.

[11] G. Asadi, M. Gitizadeh, and. A. Roosta, "Welfare maximization under real-time pricing in smart grid using PSO algorithm," in *21st Iranian Conference on Electrical Engineering (ICEE)*, 14-16 May 2013, pp.1-7.

[12] W. Zhang and L. Yang, "SC-FDMA for uplink smart meter transmission over low voltage power lines," in *IEEE International Symposium on Power Line Communications and Its Applications (ISPLC)*, 3-6 April 2011, pp.497-502.

[13] A. Arif, M. Al-Hussain, N. Al-Mutairi, E. Al-Ammar, Y. Khan, and N. Malik, "Experimental study and design of smart energy meter for the smart grid," in *International Renewable and Sustainable Energy Conference (IRSEC)*, 7-9 March 2013, pp.515-520.

[14] G. C. Madueno, C. Stefanovic, and P. Popovski, "How many smart meters can be deployed in a GSM cell?" in *IEEE International Conference on Communications Workshops (ICC)*, 2013, pp.1263-1268.

[15] P. Li, S. Guo, and Z. Cheng, "Joint Optimization of Electricity and Communication Cost for Meter Data Collection in Smart Grid," IEEE Transactions on Emerging Topics in Computing, no.99, pp.1-10.

[16] J. Zheng, D. W. Gao, and L. Lin, "Smart Meters in Smart Grid: An Overview," in *IEEE Green Technologies Conference*, 2013, pp.57,64,

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Performance Enhancement of MIMO based Visible Light Communication

Ratan Kumar Mondal*, Nirzhar Saha, and Yeong Min Jang†
Department of Electronics Engineering, Kookmin University, Seoul, Korea
E-mail: *rataneee2k6@yahoo.com; †yjang@kookmin.ac.kr

*Abstract*—**The camera based visible light communication (VLC) is the merger of VLC with vision technology in order to deploy VLC features in hand-held devices e.g., in Smartphone, employing light emitting diode (LED) transmitter to camera communication. However, the most advantageous features of VLC technology have not been achieved due to the low frame handling rate in camera module. In contrast, the spatially light source separation characteristic in camera module explores the scope to deploy multiple-input multiple-output (MIMO) concept for enhancing the overall system capacity and robust signal reception in camera based VLC system. In this paper, the performance of spatial multiplexing in MIMO based VLC system is evaluated.**

***Keywords—Visible light communication, image sensor, LED, MIMO, Spatial Multiplexing.***

## I. INTRODUCTION

Visible light communication (VLC) is a promising technology for indoor wireless communications. By using visible light as the transmission medium, VLC has many advantages, such as harmlessness to human health, high data rate, high security, unlicensed frequency band usage, high *signal-to-noise ratio* (SNR), and etc. The release of IEEE 802.15.7 MAC and PHY specifications represents a significant milestone in promoting deployment of visible light communication in short range wireless personal area networks (WPAN) [1]. Illumination is typically provided by an array of LEDs, with levels of 400-800 lux [2] being specified as the minimum required for sufficient illumination. The modulation by the LEDs light must provide the dimming support for the human eye safety. Achieving high data rates is challenging simultaneously with illuminating consistently due to the low modulation bandwidth of the sources (several MHz) [3], but the levels of illumination specified for occupation ensure that a very high SNR is available.

The availability of a large number of high SNR channels with sufficient bandwidth makes MIMO techniques a promising approach both using as a diversity and a multiplexing [3-6]. In this proposed model, due to the modulation and detection method the MIMO is being used as spatial multiplexing. The potential of this paper is to investigate the performance of MIMO spatial multiplexing in camera based visible light communication. At present, plenty of modulation schemes have been proposed for VLC system. Most of the cases, some advantageous features are considered for modulation and demodulation in VLC system. It is quite difficult to detect OOK signal by using the camera sensor because the frequency of the OOK signal is not suited with the frame rate of camera. Here *frequency shift OOK* (FSOOK) is used for easy detection by using the standard camera frame rate. Each pixel and/or group of pixels will act as multiple receivers whereas each pixel is separated with arrayed by a constant distant. As the spatial separation of pixels provide the spatial multiplexing of the MIMO system. The data rate of the MIMO system is improved by using the multiplexing feature. Spatial multiplexing is not intended to make the transmission more robust; rather it increases the data rate. To do this, data is divided into separate streams; the streams are transmitted independently via separate LED transmitters.

Rest of this paper is organized as follows: the proposed MIMO spatial multiplexing is outlined in section II. The performance evaluation is described in section III. Finally, we conclude the proposed scheme in section IV.

## II. MIMO SPATIAL MULTIPLEXING IN IMAGE SENSOR BASED VLC SYSTEM

### A. Receiver and Transmitter Model

Consider the VLC-MIMO system using camera module for signal reception as shown in Fig. 1. Here, the modulation scheme is guaranteed to provide dimming support by the LED transmitters. The average light intensity must be constant over any given interval not longer than 10ms and not shorter than 1ms during the transmission of data for providing dimming support. As mentioned before, an unmodified smartphone's camera platform to be considered as the receiver, which means the embedded camera will be running at a common commercial video frame rate, e.g. 30 *frames per second* (fps). For modulating the signal, we assume that frequency shift on-off keying (FSOOK) is employed at the LEDs transmitter for the communication over the optical channel. The modulation process is similar to frequency shift keying inasmuch as there are defined mark and space ON-OFF keying frequencies for encoding bits [3].

In FSOOK modulation scheme, the OOK frequency is a sinusoidal at $\alpha_{OOK}$ frequency with a random phase $\delta_{OOK}$. If the sampling frequency is harmonic with the OOK frequency and then the OOK frequency as a harmonic of sampling frequency $\alpha_s$, where $\alpha_{OOK} \geq \alpha_s$, thus the OOK frequency is, $\alpha_{OOK} = n\alpha_s \pm \alpha_\Delta$, where $\alpha_\Delta \leq |\alpha_s / 2|$. In the receiver, after the

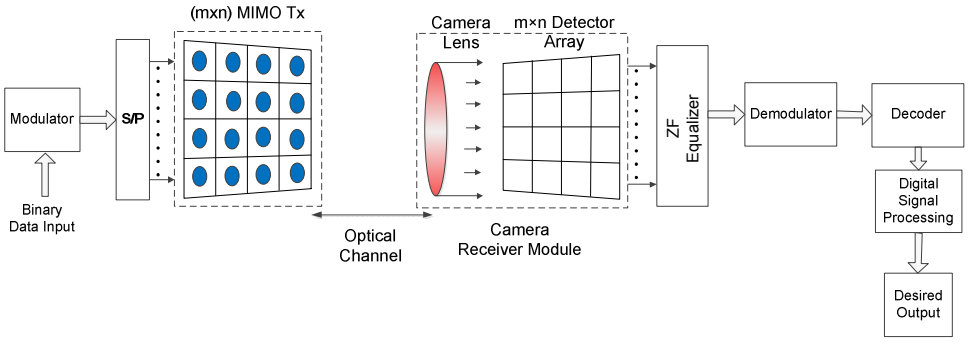



Fig. 1. MIMO spatial multiplexing offered by imaging plane

integration and hold, the comparatively low frequency signal would be expressed as below

$$s(t,\alpha_\Delta)=\mathrm{sgn}\left[\sin\left(\frac{\alpha_s}{2}t+\delta_{OOK}\right)\right] \tag{1}$$

The bit decision will depend on the following frame decoding rules,

$$s(t,\alpha_\Delta)=\begin{cases}\mathrm{sgn}\left[\sin\left(\frac{\alpha_s}{2}t+\delta_{OOK}\right)\right], & logic\ 1\\ \mathrm{sgn}\left[\sin\left(\delta_{OOK}\right)\right], & logic\ 0\end{cases} \tag{2}$$

Consider that the imaging plane can be demodulated the bit as a two dimensional lightwave-to-digital conversion. The pixel photodetector produces a signal proportional to the incoming integrated light intensity which is then held for the scanning ADC (*analog-to-digital conversion*), thus establishing the frame rate of the video camera. In the MIMO system, the reception of one light signal on pixel may be interfered by the adjacent light signal. The signal detection will be effected by this coupling of multiple light signals in one receiver. To mitigate this effect, equalizer has been introduced in the receiver.

### *B. Performance Enhancement by Spatial Multiplexing*

The VLC-MIMO with multiple LED transmitters and camera module as a receiver has been shown in Fig. 1 where an imaging receiver is used in place of the non-imaging devices (PD) [3-4]. Light propagates from the multiple LED transmitters to the receiver, and each transmitted signal is imaged onto an image sensor, where images may strike any pixels or group of pixels on the imaging plane, and be in arbitrary alignment with them. Each pixel or group of pixels on the image sensor will act like receiver, and it will measure the **H** matrix describing the optical connection between receiving pixels and LED transmitters, allows the received signals to be separated as described in following.

At the receiver, each pixel fulfills *direct detection* (DD) to produce an electrical current signal proportional to the received intensity signal. In this VLC-MIMO spatial multiplexing communication system having $N_t$ transmitters and $N_r$ receivers, where $\min\left[N_t, N_r\right]=2$. Thus the received signal by the pixels can be expressed as follows

$$\mathbf{r}=\mathbf{Hs}+\mathbf{n} \tag{3}$$

where **s** is the $N_t\times 1$ transmitted signal vector; **H** is the $N_t\times N_r$ channel matrix; and **n** is the $N_r\times 1$ additive white Gaussian noise vector. The measured channel matrix by the pixel receiver would be as follows

$$\mathbf{H}=\begin{bmatrix}h_{11} & h_{12} & \cdots & h_{1N_t}\\ h_{21} & h_{22} & \cdots & h_{2N_t}\\ \cdots & \cdots & h_{ji} & \cdots\\ h_{N_r1} & h_{N_r2} & \cdots & h_{N_rN_t}\end{bmatrix}$$

Each element of the channel matrix is made up of two components and can be written as $h_{ji}=a_{ji}h_i'$, where $h_i'$ is the normalized power in the image of the *i*-th LED transmitter at the aperture of imaging lens when the receiver is assumed at a particular position, and $a_{ji}$ quantifies how much of this power falls on the *j*-th receiver. Then, $h_i'$ can be written as

$$h_i'=\begin{cases}\sum_{i=1}^{N}\frac{A_{image}}{d_{ji}^2}R_0(\phi_i)\cos(\varphi_{ji}), & 0\le\varphi_{ji}\le\varphi_c\\ 0 & \varphi_{jik}>\varphi_c\end{cases} \tag{4}$$

where $A_{image}$ is the receiving image area of the *i*-th transmitter projected on the *j*-th detecting pixels in the imaging plane, $d_{ji}$ is the distance between the *i*-th LED in transmitter and the center of the receiver collection lens, $\phi_i$ is the angle of irradiance, $\varphi_{ji}$ is the angle of incidence of the light at the receiver, and $\varphi_c$ is the receiver *field-of-view* (FOV).

According to equation (1) and (3), assume that the link consists of $N$ sub-channels with no ICI (*inter-carrier interference*). Although, the noise process is in each signal-dependent, in here we assume the background and circuit noise

dominate over the signal-dependent portion of the noise. Thus the noise variance of $n_{ji}$ at the received pixel is given by

$$\sigma_{ji}^2 = \alpha I_{ji} + \beta \tag{5}$$

where the constant $\alpha$ quantifies the signal dependence of the noise, and $\beta$ represents signal-independent components like circuit noise, dark current noise, and shot noise due to background illumination. SNR is the signal-to-noise of the *i*-th LED to the *j*-th receiver pixels which is expressed in terms of the transmit power $s_j$, the channel dc gain $h_{ji}$ from the equation (4) and AWGN noise from equation (5).

$$SNR_{cam,i} = \frac{\sum_{\forall I_{ji}=1} \left(h_{ji} s_i\right)^2}{\sum_{\forall I_{ji}=1} n_{ji}^2} \tag{6}$$

Here is an indicator function indicating whether a pixel at $(m,n)$ coordinate receives a signal from the LED transmitters or not. Of course the indication function should satisfy the relation of the camera calibration between its focus length and the distance of the object without any blur effect on the imaging plane [5]. For simplification of analysis, we assume white noise and input, i.e., $\mathrm{E}\left[\mathbf{s}\mathbf{s}^H\right] = E_s I_{N_t}$ and $\mathrm{E}\left[\mathbf{n}\mathbf{n}^H\right] = N_0 I_{N_t}$, thus the input SNR $\gamma_s \triangleq E_s / N_0$ and FSOOK is used without loss of generality.

The measured MIMO channels have some mutual coupling effects. The equalizer can be applied to decouple the *N* sub-streams. If ZF equalizer is used then the equalization matrix is

$$\mathbf{H}_{\mathbf{ZF}} = \left(\mathbf{H}^{\mathbf{H}}\mathbf{H}\right)^{-1} \mathbf{H}^{\mathbf{H}} \tag{7}$$

With ZF equalization, the estimate of the transmitted signal $\mathbf{s}$ is given by $\hat{\mathbf{s}} = \mathbf{H}^{\dagger}\mathbf{r} = \mathbf{s} + \mathbf{H}^{\dagger}\mathbf{n}$, where $(\cdot)^{\dagger}$ denotes the matrix Moore-Penrose pseudo-inverse. Thus, at the decision-point SNR of the k-th signal stream, i.e. the signal from the k-th transmit LED, $1 \le i \le N_t$, is obtained as

$$\gamma_{ZF,j} = \gamma_s \left[\left(\mathbf{H}^{\mathbf{H}}\mathbf{H}\right)^{-1}\right]^{-1} \triangleq \gamma_s h_{ZF,j}^2 \tag{8}$$

where $h_{ZF,j}^2 \triangleq \left[\left(\mathbf{H}^{\mathbf{H}}\mathbf{H}\right)^{-1}\right]^{-1}$ denotes the channel optical power gain of the *i*-th stream. The average bit error rate (BER) of the above considered system is obtained by the following equation

$$P_b = \frac{1}{N_t} \sum_{i=1}^{N_t} Q\left(\sqrt{\gamma_s h_{ZF,j}^2}\right) \tag{9}$$

Where $Q(x) = \frac{1}{2\pi}\int_x^{\infty} e^{-y^2/2} dy$ and $h_{ZF,j}$ will depend on the reception imaging plan which has been explained in equation (4).

Generally, the capacity of the MIMO spatial multiplexing of VLC system depends on the channel impulse response, the noise process as well as a constraint on the average optical power, i.e., average amplitude, which can be transmitted. Here, we consider an efficient condition of $h_{ji}$ and compute a upper bound on the channel capacity that provides insight into how the channel capacity scales with average optical power constraint, the threshold distance for multiplexing operation and the number of pixels. The capacity of a MIMO channel can be estimated by the following equation

$$C = \max_{tr(R_{ss}) \le p} \log_2 \left[\det\left(\mathbf{I} + \mathbf{H}\mathbf{R}_{ss}\mathbf{H}^H\right)\right] \tag{10}$$

where $\mathbf{R}_{ss}$ is the covariance matrix of the transmitted vector $\mathbf{s}$, $\mathbf{H}^H$ is the transpose conjugate of the $\mathbf{H}$ matrix and $p$ is the maximum normalized transmit power. Nevertheless, we can perform linear transformations at both the transmitter and receiver converting the MIMO channel to *N* SISO sub-channels (given that the channel is linear) and, hence, reach more insightful results. These transformations can be found in [5] and the capacity will evaluated as follows

$$C = \sum_{i=1}^{N} B \log_2 \left(1 + SNR_{cam,i}\right) \tag{11}$$

where $B$ is the receiver sampling rate (i.e., camera frame rate).

TABLE I. SIMULATION PARAMETERS

| Paramenter | Value |
|---|---|
| Number of LED Transmitter | 1, 2, 4, 9, 16 |
| Number of LEDs per transmitter | 3600 (60x60) |
| Distance between transmitter and receiver | 1 [m] |
| LED transmitter separation | 3 [cm] |
| Transmitted optical power | 20 [mW] |
| Receiver's diagonal FOV | 42 [deg] |
| Receiver optical resoulation | 640x480 [pixel] |
| Focal length | 2.3 [mm] |
| Frame rate at camera | 30 [fps] |
| Filter's transmission coefficient | 1.0 |
| Modulation scheme | FSOOK |

## III. PERFORMANCE EVALUATION

Table I shows the parameters used for evaluating the performance of the spatial multiplexing of MIMO based VLC system. For simplicity, here we assume that the separation between two adjacent LED transmitters will satisfy the perfect imaging detection on the receiving plane. The blur effect on the imaging plane is not considered and the distance between the transmitter and the receiver is fixed at a particular distance (1m) from where the image of the two adjacent LED transmitters could be separated without any advance image

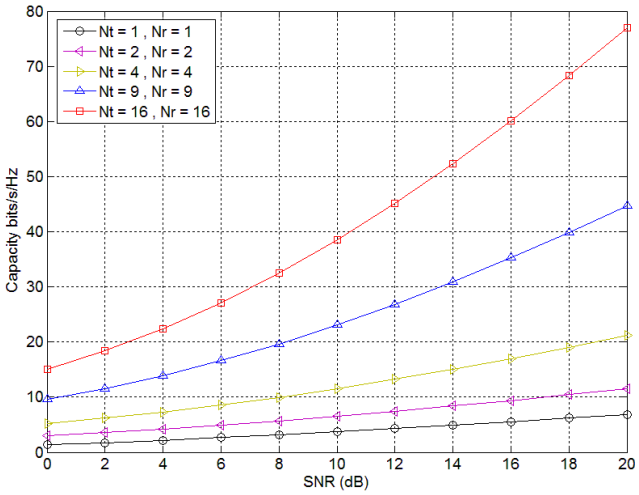



Fig. 2. Capacity [bit/s/Hz] as a function of SNR of VLC system for different transmitters-receivers combinations.

processing. And also consider that, transmitters and receivers are in perfectly aligned in order to satisfy the field-of-view (FOV) criterion of VLC system for detecting the signal.

The main objective of proposed spatial multiplexing of VLC-MIMO operation was to enhance the system capacity as explained in the above section. In Fig.2, the simulated capacity performance has been depicted with different numbers of transmitters and receivers combinations. The system average capacity is rapidly increased with respect to SNR when the number of transmitters and receivers are increased. Spatial light source separation characteristic of camera module has opened the scope to use the same spectral bandwidth on every multiple parallel channels. Even though cross-coupling may occur between the parallel channels due to the multiple channels having the same spectral band but the imaging plane can separate them accurately and spatially for a limited distance between the transmitters and receivers. Normally, the channel gain in the VLC system is very high which implies the high SNR value in every channel in our proposed model. On the other hand, achieving this high SNR value is not inefficient in this VLC model. So, the system capacity has increased when the MIMO multiplexing was introduced without splitting the spectral bandwidth.

Fig.3 depicts the bit error rate (BER) of several transmitters-receivers combinations with respect to the SNR. Here the BER is increasing with respect to the SNR when the numbers of transmitters-receivers are also increasing. BER is increasing because the adjacent pixel couldn't able to separate perfectly when the numbers of transmitters are increasing. For example, at 14 dB SNR the BER of single transmitter system is less than $10^{-4}$ and on the other hand, to get the BER under $10^{-3}$ for $2\times2$ combination, it needs approximately 16dB. From Fig.3, it can be summarized that if the numbers of transmitters-receivers are increased then the MIMO system needs more SNR to achieve a reliable BER. To improve the BER performance, equalizer has also been introduced with the MIMO system. In Fig.3 shows the BER performance of ZF, MMSE and ML equalizer for $16\times16$ combination. In this figure, the BER under $10^{-3}$ is achieved for ZF, MMSE and ML

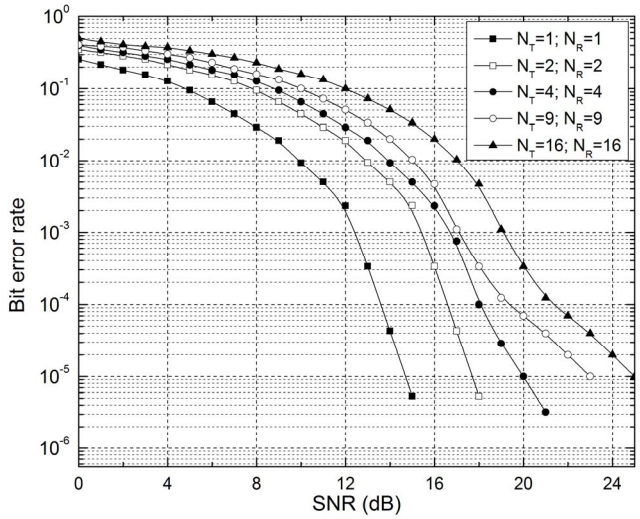



Fig. 3. BER vs. SNR for different transmitters and receivers combinations (without equalization)

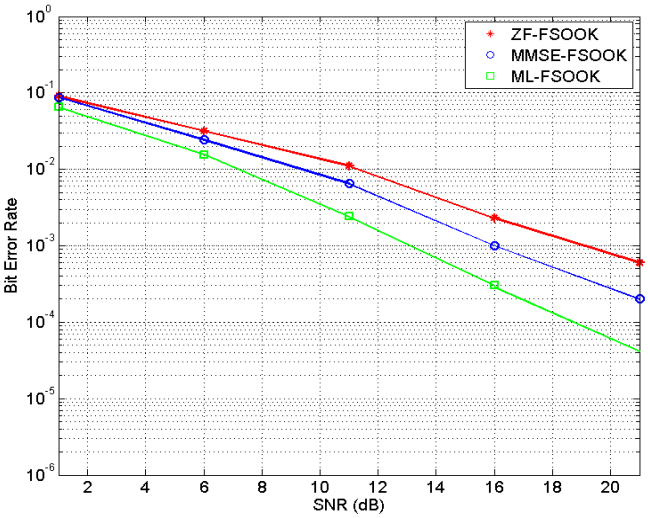



Fig. 4. BER vs. SNR for 16x16 VLC-MIMO system with different equalizations.

equalizer in 19dB, 16dB and 13dB respectively which are much better performance than as shown in Fig.2. If the numbers of transmitters-receivers combinations are increased very largely then the BER performance will approximately same for ZF and MMSE based system. In contrast, for large numbers of transmitters-receivers the BER of ML based system will be low but the system complexity will much increase with respect to the numbers of transmitters-receivers. For considering low complexity with capability to handle a large numbers of transceivers, we consider the ZF equalizer in our proposed spatial multiplexing system.

## IV. CONCLUSION

In this paper, we present a performance analysis of MIMO spatial multiplexing of image sensor based VLC system. In the receiver, the imaging plane i.e., image sensor is considered with lens same as the camera setup. Here, spatial multiplexing

of MIMO operation is deployed for improving the system capacity. In contrast, the probability of BER increases with respect to the increasing numbers of transceivers due to the mutual coupling between close adjacent channels. To mitigate this effect and recognize the multiple data streams separately and more accurately with less complexity, ZF equalizer has been introduced in the MIMO spatial multiplexing operation. Thus, MIMO operation with spatial multiplexing has enhanced the performance and in future image sensor based MIMO operation will contribute in the multiple access technology of VLC system.

## ACKNOWLEDGMENT

This research work is funded by the "MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013".

## REFERENCES

[1] IEEE Std., "Local and metropolitan area networks-part 15.7: Short-range wireless optical communication using visible light," Sept. 2011.

[2] T. Q. Wang, Y. A. Sekercioglu, and J. Armstrong, "Analysis of an optical wireless receiver using hemispherical lens with application in MIMO visible light communications," in *IEEE J. Lightwave Tech.*, vol. 31, no. 11, pp. 1744-1754, Jun. 2013.

[3] R. D. Roberts, "Space-time forword error correction for dimmable undersampled frequency shift ON-OFF keying camera communications (CamCom)," WOCC 2013.

[4] L. Zeng, D. C. O'Brien, H. L. Minh, G.E. Faulkner, K. Lee, D. Jung, Y. J. Oh, and E. T. Won, " High data rate multiple input multiple output (MIMO) optical wireless communications using white LED lighting," in *IEEE J. Sel. Areas Commun.,* vol. 27, no. 9, pp. 1654-1662, Dec. 2009.

[5] A.Ashok, M.Gruteser, N. B. Mandayam, J. Silva, K. Dana, and M.Varga., "Challenge: Mobile optical networks through visual mimo," In MobiCom'10: Proceedings of the sixteenth annual international conference on Mobile computing and networking, pp. 105–112, New York, NY, USA, 2010. ACM.

[6] N. Letzepis, I. Holland, and W. Cowley, "The Gaussian free space optical MIMO channel with Q-ary pulse position modulation," in *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1744-1753, May 2008.

[7] A. H. Azhar, T. A. Tran, and D. C. O'Brien, "Demonstration of high-speed data transmission using MIMO-OFDM visible light communications," in Proc. IEEE Globecom 2010 Workshop on Optical Wireless Commun., pp. 1052-1056, 2010.

[8] Y. A. Alquadh and M. Kavehrad, "MIMO characterization of indoor wireless optical link using a diffuse-transmission configuration," in IEEE Trans. Commun., vol. 51, no. 9, pp. 1554-1560, Sep. 2003.

[9] N. Wang and S. D. Blostein, "Approximate minimum BER power allocation for MIMO spatial multiplexing systems," IEEE Trans. Commun., vol. 55, no. 1, pp. 180-187, Jan. 2007.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Low Complexity SDNLMS Adaptive Channel Estimation for MIMO-OFDM Systems

Tariq Ahmad Dewan*, Sabbir Hasan and Foisal Hossain
Khulna University of Engineering and Technology
Khulna, Bangladesh
* tariq_shoffer@yahoo.com

***Abstract*—Channel Estimation process is an important technical issue for Multiple Input Multiple Output (MIMO)-Orthogonal Frequency Domain Multiplexing (OFDM) wireless communication systems. Adaptive Channel Estimation (ACE) is a widely used estimation technique for MIMO-OFDM systems. Different types of channel estimator like least mean square (LMS), normalized least square (NLMS) are applied to ACE. LMS algorithm has low complexity but it suffers from high minimum mean square error (MMSE) performance. Normalized least mean square algorithm (NLMS) provides low MMSE performance with poor convergence rate. In this paper, an improved sign data normalized least mean square (SDNLMS) channel estimation technique is proposed which improves the convergence rate and computational complexity of NLMS algorithm, while maintaining low MMSE performance for channel estimation (CE). Simulation results show that SDNLMS CE method provides faster convergence rate with low MMSE and computational complexity, which is clearly better than NLMS and LMS method in a MIMO-OFDM system. Therefore SDNLMS CE method can be highly compatible for ACE process.**

***Keywords-MIMO , OFDM , LMS , NLMS, SDNLMS.***

## I. INTRODUCTION

Use of multimedia services and different internet related services are gradually increasing nowadays. This makes people interested to high speed communication. It is difficult to design a transmission system having wide bandwidth and flexibility. To achieve greater bandwidth and channel capacity in wireless communication system different types of method has been proposed recent years [1]. The multiple-input multiple-output (MIMO) combined with orthogonal frequency division multiplexing (OFDM), is one of the effective and popular communication scheme that ensure large system capacity without extra power or bandwidth consumption for which is necessary for high speed wireless communication systems [2].

IN MIMO-OFDM system multiple input multiple output antenna structure facilitates increase channel capacity and reliability, meanwhile orthogonal frequency division multiplexing ensures efficient high speed data transmission and compatibility to frequency selective channels. Hence, the integration of these two technologies has the potential to meet the requirements of the next-generation wireless services such as wireless local area networks (WLANs), worldwide interoperability for microwave access (WiMAX), wireless fidelity (WiFi), cognitive radio, and 3rd generation partnership project (3GPP) long term evolution (LTE) and other ever growing demands of future communication systems [3]. In MIMO-OFDM system inter-carrier interference (ICI) and inter-symbol interference (ISI) introduced in the received signal. Channel estimation techniques can flexibly detect the signals in both cases. Therefore channel estimation is an important issue in MIMO-OFDM systems. The channel can be estimated by basically two ways. One is called pilot tone based or supervised channel estimation, which can be estimated by using the training sequence from transmitted signal to estimate the channel. Other is blind channel estimation (BCE), where approaches can be taken into consideration in order to avoid the payload of using training sequences. Here pilot or tones knowledge of channel and noise statistics are not necessary, only knowledge of received signal is required, which updates channel parameters. [4]. BCE has its advantage of no overhead cost, but it also needs a long data records and complex algorithm [5]. In the pilot based estimation, training symbols in a block are used. This can compromise the performance of systems, but because of simplicity, it is appreciated to use the training symbols in the first part of each OFDM block [6]. In digital communication system, adaptive channel estimation process can be done by periodically transmitting a training sequence which is known to the receiver. It is now a widely used estimation scheme for channel estimation.

Different types channel estimation techniques have been proposed for MIMO-OFDM systems [7-13]. Pilot-based channel estimation schemes have been proposed in [14-15]. But these schemes suffer from high computational complexity, memory requirements and prior knowledge of the second order statistics of channel. Reference [16] describes Adaptive channel singular value decomposition estimation for MIMO-OFDM system. Despite these estimations, there are many adaptive filter algorithms that are widely used for channel estimation. But they suffer from different problems. They either have a low computational complexity with high mean-square error (such as LMS and SLMS algorithms) or low mean-square error with a high computation complexity and slow convergence rate and (such as NLMS algorithm). This implicates that minimum mean-square error (MMSE), convergence rate and computational complexity are three important points that should be considered in selecting of the adaptive algorithms for channel estimation [3], [17].

In this paper sign data normalized least mean square (SDNLMS) channel estimation technique has been proposed. The MMSE of the SDNLMS algorithm is slightly higher than the NLMS algorithm but the convergence speed and

computational complexity significantly outperform the NLMS algorithm. SDNLMS method is compared with some offspring algorithms like leaky least mean square (LELMS), sign error least mean square (SLMS), and basic LMS, NLMS CE algorithm for MIMO-OFDM systems.

In this paper some notations are used. Subscripts are used as time indices for vectors, e.g $u_i$, and parenthesis as time indices for scalars, e.g., $u(i)$ the boldface letter ***d*** refer to the random variable, and the normal letter *d* refer to observations (or realizations) of it. Superscript $x^*$ denotes the Hermitian transpose of the matrices, and $\|x\|^2$ is the squired Euclidean norm of $x$.

The paper is organized as follows. Adaptive channel estimation scheme LMS, LELMS, SLMS and NLMS are presented in section II, and SDNLMS CE algorithm is described in section III, simulation results are shown in section IV, performance of SDNLMS method is analyzed in section V. Finally, some remarks are given in section VI.

## II. ADAPTIVE ALGORITHMS

Adaptive filter is used, to estimate MIMO-OFDM channel. Fig. 1 shows general adaptive channel estimate procedure.

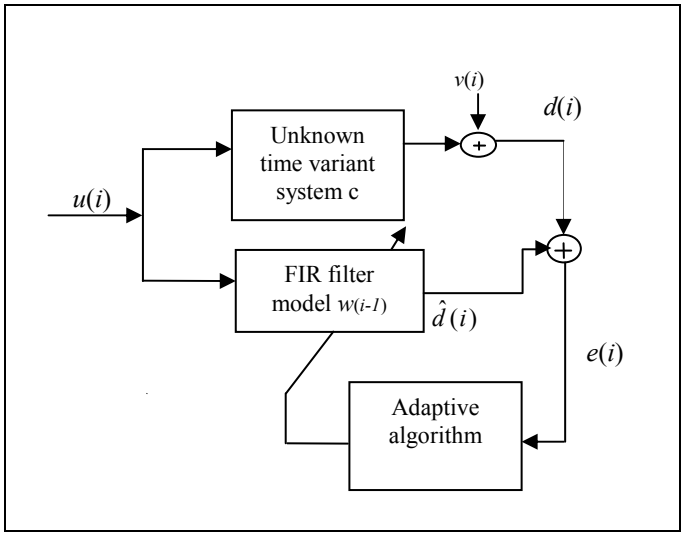



Figure 1. Scheme for adaptive CE [18]

If a zero-mean random sequence signal $\{\boldsymbol{u}(i)\}$ is transmitted via a time-varying channel and received signal is another zero-mean random sequence $\{\boldsymbol{d}(i)\}$ at any time instant *i,* the state of the channel is captured by the regressor

$$\boldsymbol{u_i} = [\ \boldsymbol{u}(i)\ \boldsymbol{u}(i\ -\ 1)\ \ldots\ \boldsymbol{u}(i\ -\ \mathrm{M}\ +\ 1)] \tag{1}$$

And output in the receiver can be described by

$$\boldsymbol{d}(i)\ =\ \boldsymbol{u_i}\mathrm{c} + v(i) \tag{2}$$

where the column vector *c* denotes the channel impulse response sequence, and *v*(*i*) is the additive white Gaussian noise (AWGN) with zero-mean and variance $\sigma^2$, that is uncorrelated with $\boldsymbol{u_i}$. Now consider an FIR filter with adjustable weights in receiver side which is excited by same training sequence. Where FIR filter output can be described by

$$\hat{d}(i) = u_i w_{i-1} \tag{3}$$

Where $w_{i-1}$ is the filter weight vector at (i-1) time instance.

So the error is $e(i) = d(i) - u_i w_{i-1}$ (4)

The error is then used to adjust the filter coefficients from $w_{i-1}$ to $w_i$. The output of the adaptive filter $\hat{d}(i)$ will assume values close to $d(i)$. Consequently, from an input/output perspective, the (adaptive filter) mapping from $u(i)$ to $\hat{d}(i)$ will behave similarly to the channel, which maps $u(i)$ to $d(i)$ Therefore, if we can determine optimum value of $w$ then we recover c.

In LMS algorithm weight is estimated by [18-19]

$$w_i\ = w_{i-1} + \mu u_i^*[d(i) - u_i w(i-1)] \tag{5}$$

$\mu$ is the step size.

In NLMS, LELMS, SLMS method they have different calculation for updating tap weights [19]. Other procedures are same as LMS filter.

NLMS is likely LMS method; except that now update is derived from regularized Newton's recursion method.
Here the update eqn is

$$w_i\ = w_{i-1} + \frac{\mu}{\varepsilon + \|u_i\|^2} u_i^*[d(i) - u_i w(i-1)] \tag{6}$$

Where $\varepsilon$ is a small positive constant to avoid division by zero. Selection of step size in the NLMS is comparatively easier than that of LMS algorithm.

Leaky LMS update is described as

$$w_i\ = (1 - \mu\alpha) w_{i-1} + u_i^*[d(i) - u_i w(i-1)] \tag{7}$$

If $\alpha$ = 0, the leaky LMS algorithm becomes the same as the standard LMS algorithm. A large leaky factor creates a large steady state error. Normally $\alpha$ is in the range of [0, 0.1].

In SLMS the weight update is determined by

$$w_i = w_{i-1} + u_i^* \ \mathrm{csgn}[d(i) - u_i w(i-1)] \quad (8)$$

SLMS method has comparatively low computational complexity. But due to high MMSE it is difficult to analyze.

## III. SDNLMS ADAPTIVE ALGORITHM

Sign data LMS has undesirable MMSE and high convergence performance. NLMS can improve MMSE performance but convergence rate is considerably slow. A unique method is proposed which can improve convergence keeping the MMSE performance. This paper combines two algorithms. Here NLMS, derived from newton recursion from (6) we know the approximation for estimation is,

$$w_i = w_{i-1} + \frac{\mu}{\varepsilon + \|u_i\|^2} u_i^* [d(i) - u_i w(i-1)] \quad (11)$$

Here $\mu$ is the step size, and $\varepsilon$ is the regularization factor. From Sign data algorithm, the weight vector can be approximated iteratively via the recursion:

$$w_i = w_{i-1} + \mu \, \mathrm{csign}(u_i^*)[d(i) - u_i w(i-1)] \quad (12)$$

Where, $\mathrm{csgn}(z) \triangleq \mathrm{sign}(z_r) + j\mathrm{sign}(z_i)$

And, $$\mathrm{sign}(a) \triangleq \begin{cases} +1 & \text{if } a > 0 \\ -1 & \text{if } a > 0 \\ 0 & \text{if } a = 0 \end{cases} \quad (13)$$

The resultant combine algorithm provides high convergence and low MMSE. The update estimation for optimum performance is

$$w_i = w_{i-1} + \frac{\mu}{\varepsilon + \|u_i\|^2} \mathrm{csgn}(u_i^*)[d(i) - u_i w(i-1)] \quad (14)$$

As it is known, Sign LMS can be found in different types. Sign data, Sign error, Sign Sign LMS. Here, SDNLMS algorithm deals with Sign data algorithm where only input signals are performed through sign operation.

## IV. EXPERIMENTAL RESULTS

This section compares MSE Vs. iteration and computational cost for LMS, NLMS, LELMS, SLMS, and SDNLMS algorithms for MIMO systems. In these simulations, 4 by 4, 6 by 6 and 8 by 8 antenna combinations are considered for MIMO OFDM systems. 4 by 8, 8 by 4 antenna combinations (number of transmitting and receiving antenna are not equal) are also considered. In this simulation figures are the simulated output for signal to noise ratio (SNR) 15 dB. Modulation system used in simulation is QPSK. AWGN noise is considered in the simulation. Simulation is done in MATLAB.

### *A. MSE performance observation:*

MMSE performance for described methods is shown in Fig. 2 to Fig. 6. SLMS provides very high MMSE performance. SDNLMS provides almost same MMSE performance as NLMS. In addition, the utilizing of more multiple antennas at the transmitter and/or receiver in the system SDNLMS, NLMS both provides a much higher and steady state MSE performance than other adaptive estimator compared with fewer antennas. Here Fig.2 provides lowest MMSE and Fig.6 provides highest MMSE for a particular method.

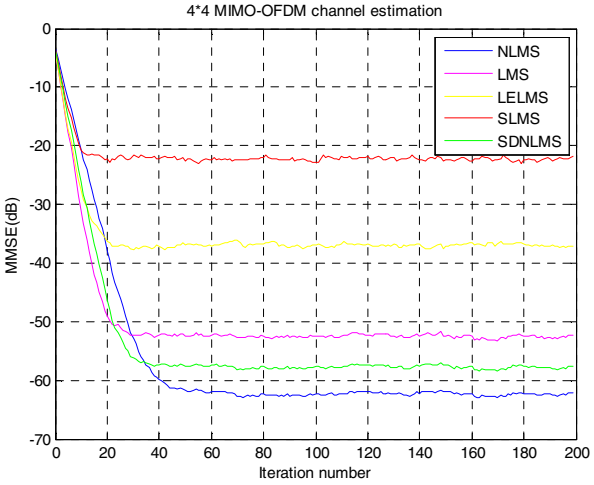



Figure 2. 4 by 4 MIMO-OFDM system

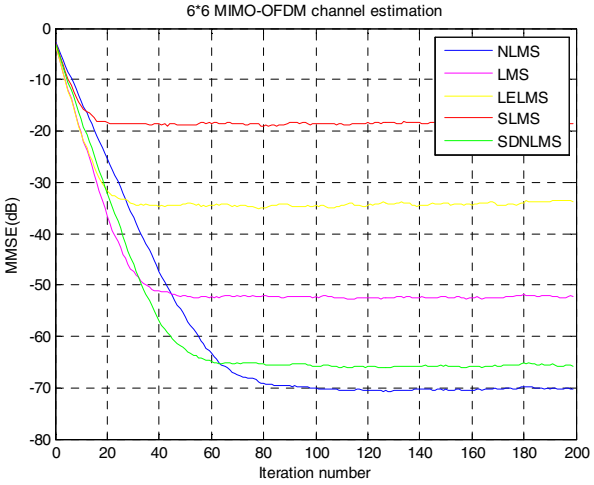



Figure 3. 6 by 6 MIMO-OFDM system

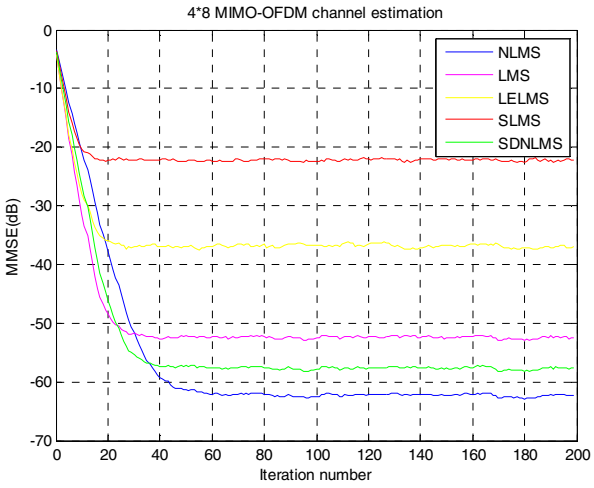



Figure 4. 4 by 8 MIMO-OFDM system

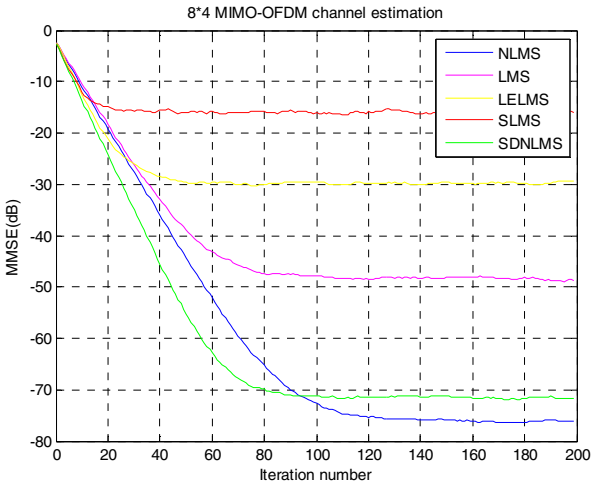



Figure 5. 8 by 4 MIMO-OFDM system

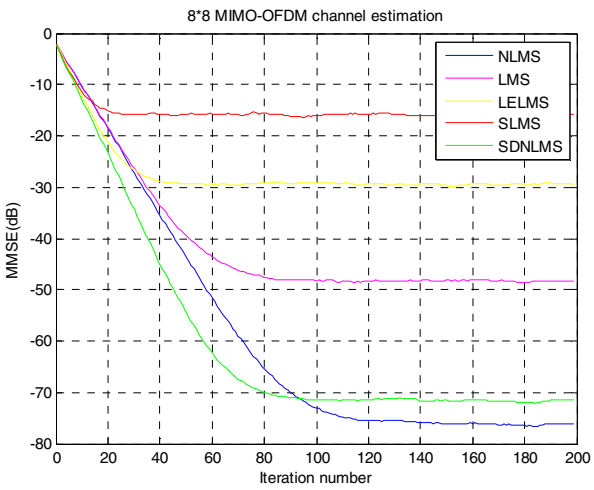



Figure 6. 8 by 8 MIMO-OFDM system

### *B. Convergence:*

Simulation results show that SLMS CE has faster convergence than other CE. LELMS CE also provides faster convergence. LMS CE is slower than SLMS CE. And NLMS CE is the slowest among them. SDNLMS provides better convergence than NLMS CE method. Using more multiple antennas, SDNLMS provides much better convergence than NLMS method with a constant MMSE difference.

### *C. Computational complexity:*

Computational complexity of LMS, NLMS LELMS, SLMS, SDNLMS algorithms has been shown in this section. Table I shows the computational complexity of L length complex-valued data in terms of the number of complex multiplications, complex additions, complex divisions, and comparisons with zero (or sign evaluations), in each iteration [17], [19].

| Method | Addition | Multiplication | Division | Sign |
|---|---|---|---|---|
| LMS | 2L | 2L+1 | - | - |
| SLMS | 2L | 2L | - | 1 |
| NLMS | 3L | 3L+1 | 1 | - |
| LELMS | 2L+1 | 3L+2 | - | - |
| SDNLMS | 3L | 2L+1 | 1 | 1 |

TABLE I: Computational complexity in each iteration for different ACE method.

Calculation shows that LMS algorithm has less computational complexity. LELMS, SLMS algorithm complexities are also low. NLMS algorithm also has a low addition and multiplication complexity but it requires another division. SDNLMS requires less number of complex multiplication operations than NLMS algorithm.

## V. ANALYSIS

MMSE performance, convergence and computational complexity are the key factor for adaptive estimation process. This observation shows that SLMS channel estimation provides faster convergence with a high MMSE, which is difficult to analyze. LMS and LELMS channel estimation also have poor MMSE performance, which is also undesirable. NLMS method provides high MMSE performance with a very slow convergence rate and higher computational complexity. SDNLMS provides significantly high convergence than NLMS. MMSE performance of SDNLMS is slightly less than NLMS, which is considerable and does not vary with number of antenna in the MIMO-OFDM systems. But convergence of SDNLMS gradually improves than other channel estimation methods, when more antennas are used in the system. Therefore, considering all key factors (accuracy, convergence and complexity) SDNLMS CE is a suitable ACE process for MIMO-OFDM systems.

## VI. CONCLUSION

At recent, MIMO-OFDM transmission has become popular, as it offers significant increases in data throughput and link range without additional bandwidth or increased transmit power. It also provides improved reliability and performance for both third and fourth generation wireless network. In this paper, LMS, LELMS, SLMS, NLMS and SDNLMS adaptive

channel estimator are described for MIMO OFDM systems. Experimental result and analysis above state that the proposed SDNLMS CE provides smooth performance than other CEs. So, it can be concluded that SDNLMS channel estimation technique is an efficient scheme MIMO-OFDM systems.

## REFERENCES


[1] Vaishali B. Niranjane, Nivedita A. Pande, and Anjali A. Keluskar, "Adaptive Channel Estimation Technique for MIMO-OFDM," International Journal of Emerging Technology and Advanced Engineering, (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 1, January 2013).

[2] M. N. Seyman and N. Taspinar "MIMO-OFDM Channel Estimation Using ANFIS," ELECTRONICS AND ELECTRICAL ENGINEERING, ISSN 1392 – 1215 2012. No. 4(120).

[3] S. A. Hosseini, and S. A. Hadei, "Low-Complexity Adaptive Channel Estimation for MIMO-OFDM Systems over Rayleigh FadingChannels," NCNIEE.

[4] M. Hossain, S. M. Farhad and Md. T. Riasat "Performance analysis of RLS and VSS-LMS channel estimation techniques for 4G MIMO OFDM systems," Computer and Information Technology (ICCIT), 2012, pp.267-270, 22-24 Dec. 2012.

[5] Pei-Sheng Pan and Bao-Yu Zheng, "An Adaptive Channel Estimation Technique in MIMO -OFDM Systems," JOURNAL OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA, VOL. 6, NO. 3, SEPTEMBER 2008.

[6] M. M. Rana, and M. K. Hosain, "Adaptive Channel Estimation Techniques for MIMO OFDM Systems," International Journal of Advanced Computer Science and Applications, United States, Vol. 1 No. 6, December 2010.

[7] T. Chang, W. Chiang, and Y. P. Hong, "Training sequence design for discriminatory channel estimation in wireless MIMO systems," IEEE Transactions on Signal Processing, vol. 58, no. 12, pp. 6223–6237, 2010.

[8] M. N. Seyman and Taspinar "Channel Estimation Based on Adaptive Neuro–Fuzzy Inference System in OFDM," IEICE Transactions on Commun., 2008. – No. E91–B(7). – P. 2426–2430.

[9] J. Du and Y. Li, "MIMO-OFDM channel estimation based on sub-space tracking," in Proc. IEEEVTC, Jeju, Korea, Apr.2003, vol.2, pp.1084–1088. Korea, Apr.2003, vol.2, pp.1084–1088.

[10] Z. J. Wang, Z. Han, and K. J. R. Liu, "A MIMO-OFDM channel estimation approach using time of arrivals," IEEE Transactions on Wireless Communications , vol. 4, no. 3, pp. 1207–1213, 2005.

[11] C. Shin, W. Heath Jr., and E. J. Powers, "Blind Channel Estimation for MIMO-OFDM Systems," Vehicular Technology, IEEE Transactions on (Volume:56 , Issue: 2 ), March 2007, pp:670 – 685.

[12] J. K. Tugnait, "Blind estimation and equalization of MIMO channels via multi delay whitening," IEEEJ. Sel. Areas Commun., vol. 19, no.8, pp. 1507–1519, Aug.2001.

[13] Y. (G.) Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," IEEE Trans. Wireless Commun., vol. 1, Jan 2002.

[14] Y. Li, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channel," IEEE JSAC vol. 17, no. 3, pp. 461-471, 1999.

[15] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," IEEE Trans. Wireless Comm., vol. 1, no. 1, pp. 67-75, 2002.

[16] H. Z. Jafarian and G. Gulak, "Adaptive Channel SVD Estimation for MIMO-OFDM Systems ".

[17] S. Haykin, "*Adaptive Filter Theory*," Prentice-Hall International Inc, 1996.

[18] S. Salivahanan, A. Vallavaraj and C. Gnanapriya "Digital Signal Processing," .

[19] A. H. Sayed, " *Fundamentals of Adaptive Filtering*," John Wiley & Sons, June 13, 2003.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Power Allocation Grouping Scheme Considering Constraints in Two Separate Stages for OFDM-Based Cognitive Radio System

Elham HOSSEINI and Abolfazl FALAHATI
Iran University of Science and Technology (IUST), Iran
ehosseini@iust.ac.ir

***Abstract*—A power allocation grouping scheme for orthogonal frequency division multiplexing (OFDM)-based Cognitive Radio systems is proposed that can improve capacity while the interference power for primary user receivers stays at constant level. In this scheme, all idle subcarriers are divided into four groups based on the inserted interference power level into the primary user frequency band. Then, the power is allocated to subcarriers in two stages. At the first stage, considering interference constraint, the power is assigned to subcarriers into groups 1, 2 and 3. At the second stage, the remaining power is allocated to other idle subcarriers with water filling algorithm. The numerical results show that the capacity is improved by 20 bps up to 45 bps with the increase in maximum power budget over cognitive radio system. Indeed, the interference power level is unchanged with increasing the transmit power for proposed cognitive radio system.**

***Keywords—Cognitive Radio, Power Allocation Algorithm, OFDM, Water Filling Algorithm.***

## I. INTRODUCTION

Since demands for new services and expansion of existing services are increased dramatically, spectrum scarcity is a most challenge in wireless communication in recent years. According to measurements of ITU, frequency spectrum is underutilized with fixed resource allocation policy. So, cognitive radio (CR) scenario was introduced by Mitola to share licensed frequency spectrum opportunistically with primary user (PU). Since CR system implement dynamic resource allocation with coexisting primary and secondary user (SU), utilization of the limited spectral resource is enhanced. Licensed spectrum is shared between PUs and SUs in three schemes named interweaved, underlay and overlay. In first scheme, SUs continuously sense license spectrum to find spectrum holes which is permissible to transmit their data. However, this scheme has high sensitivity to sensing error. Underlay scheme is developed to transmit simultaneously SUs and PUs in the same frequency so that interference power into primary user which results of secondary transmission is not higher than threshold. The overlay scheme is similar interweave with this difference that SUs cooperate PUs as relay when they is not allowable to transmit their data [1-4].

In order to develop CR network, OFDM platform is a best candidate for this network whose physical layer has to be highly agile to able adapt with radio environment variation fast.

Power allocation algorithm in wireless communication is applied to control interference and maximize capacity. Power allocation problem in cognitive networks is a major issue since the SUs is used PUs spectrum under different scenario. This challenge is more significant in CR system because SUs do not have to be adverse effect on PU' signal. The interference power of PU depends on the SU's transmission power, the interference channel gain and sharing scheme. Hence, power allocation in SU network should been performed using CSI of both the main channel and the interference channel. Therefore, radio resource allocation in cognitive radio network is more complex than primary network because of considering additional interference constraint. However, conventional power allocation algorithm should been modified to apply in cognitive radio system. Interference power level is a major limiting factor for communication performance in many wireless networks. This factor has significant effect in resource allocation of cognitive radio system in two point views. The first perspective is interference from PU into SU which degrade performance of SU gained jamming SU signal. It is assumed that this interference is insignificant and is modeled additive Gaussian noise. The latter is interference from SU into PU which map to second power constraint and is our focus. Iteratively, interference can be categorized in two groups named in-band and out-of-band interference. In-band interference is considered underlay sharing that primary and secondary users are used the same frequency spectrum simultaneously. Out-of-band interference is significant when the PU physical layer is not OFDM and primary and secondary user exist in adjacent channel named overlay/interweave sharing [5]. Our focus in this paper is over out-of-band interference.

In [6], joint power and subcarrier allocation is considered to satisfy individual user quality of service (QoS) requirement for MU-OFDM, in addition to interference constraints. In this reference, the adaptive rate problem is formulated as an integer linear programming problem that is solved by numerical method named branch and bound.

In [7], a linear water-filling power allocation algorithm is proposed to determine rapidly appropriate subcarriers in

OFDM system. The throughput for this algorithm is close to iterative water-filling power allocation algorithm. In [9], the CR system resource allocation problem is solved by Linear Water-Filling (LWF) scheme considering transmit power and interference constraints.

Some researchers [8] have studied the power allocation problem in multiuser case for the downlink OFDM-based CR system, and multilevel water-filling algorithm has proposed according to user weight factor.

In [5], the interference from PU to SU and the interference from SU to PU are considered in the proposed scheme as the transmission power constraint in each subcarrier for the SU under jointly overlay and underlay sharing. The space distance between PU and SU transceivers as well as the frequency distance between PU band and subcarriers are the actual effective parameters in this power allocation scheme.

In all mentioned works, interference constraint is considered as upper boundary which is applied after analytical solution. However, power efficiency and SU's capacity are decreased. In order to improve this scheme, the interference channel coefficients are contributed in modified Water-Filling (WF) algorithm.

The power allocation for single SU is performed in view of interference channel gain and frequency distance between PU's band and the idle subcarriers [9]. So, more power is loaded into a subcarrier that is far away from PU's band and has the lower interference channel gain.

In [10], statistical properties of injected interference power to PU channel are used to determine the container bottom for each subcarrier in WF algorithm. In other words, the container bottom level of each subcarrier depends on injected interference to PU Simulation Results show that the PU injected interference is decreased while the SU capacity does not change significantly.

In [11], the optimization problem is introduced to maximize the received information bit considering PU's traffic model and CSI as well as PU injected interference. The proposed MLWF algorithm raises the power efficiency and reduces the interference power by using radio environment information well. In the proposed algorithm, the best groups of idle subcarriers in proportion to PU activities are selected. Thus, the probability that PUs, by the proposed algorithm, reclaims a channel is lower than previous algorithms.

In previous works, the power is assigned to all subcarriers in one stage simultaneously. Since it is typically very hard to solve the resource allocation with power and interference constraint jointly, these two constraints are carried out separately to lower the complexity in the proposed scheme. In this scheme, primarily subcarriers are divided to four groups. The new metric that depends on the interference channel gain in addition to the frequency distance between PU band and the idle subcarriers, is introduced for subcarriers grouping. Then, the power that is assigned to any subcarrier is proportionate to its group and the main channel gain. So, power efficiency is increased while complexity is lower than other algorithm. Our contribution includes:

Out-of-band interference constraint in interweave sharing is simplified based on frequency distance and the interference channel coefficients.

Power is separately allocated to adjacent subcarriers of PU's spectrum and then the residual power is assigned to other subcarriers by solving Lagrangian method considering only total power constraint.

This paper is organized as follows. In section 2, the power allocation optimization problem is defined considering interference and maximum power constraints usually present in a CR system model. Within section 3, in order to decrease complexity algorithm, power is assigned to subcarriers in two stages. In the first stage, subcarriers are grouped based on frequency distance from PU's frequency band. Then, power allocation is applied to neighboring subcarriers considering interference coefficient. In the second stage, the remaining power is assigned to no interfering subcarriers. Section 4 presents the simulation result. Finally, section 5 concludes the paper.

## II. SYSTEM MODEL

In this paper, a spectrum sharing network with four PUs and one SU is modeled. The system model considers the resource allocation with an OFDM-based CR system. It is also assumed that the SU is allowed to access the spectrum in the frequency band with no PU. This spectrum sharing is an interweave approach. The total frequency band is divided into N subcarriers. The main channel refers to the channel between SU transmitter (SU-Tx) and SU receiver (SU-Rx). And, the channel between SU-Tx and PU-Rx is known as interference channel. The main and interference channels are assumed to be a point-to-point flat Raleigh fading channels that the main channel gain and the interference channel gain as well as variance of AWGN are $g_i^{sp}$, $g_i^{ss}$, $\sigma_{n_0}^2$ respectively.

The channel gain is related to the squared coefficient ( $g_i = |h_i|^2$ ). Hence, distribution function for channel gain is exponential with mean value of unity. CR channel is assumed as an interference channel that its model is shown in Fig. 1.

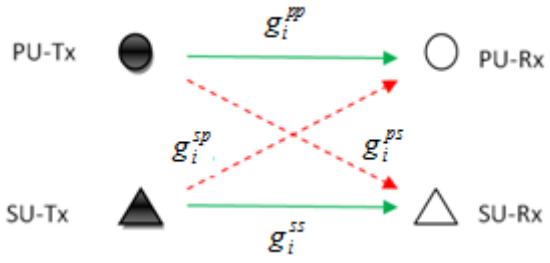



Figure 1. The CR channel model

The optimization variable is secondary transmitter power vector in each subcarrier ( $p=[p_1,\ldots,p_k,\ldots,p_N]$ ). The optimization problem objective is to maximize the secondary capacity while keeping the instantaneous interference injected to the PUs frequency band below a certain threshold, expressed mathematically as,

$$\max_{p=[p_1,\ldots,p_N]} C = \frac{1}{2}\sum_{i=1}^{N}\log_2(1+\frac{g_i^{ss}p_i}{n_0})$$
$$s.t. \quad \sum_{i=1}^{N} I_i^{(l)}(d_{il}, g_i^{sp}, p_i) \le I_{th}^{(l)} \qquad (1)$$
$$\sum_{i=1}^{N} p_i \le p_m \ , \ p_i \ge 0$$

where C, $n_0$, $I_i^{(l)}, I_{th}^{(l)}$ and $d_{il}$ are secondary capacity, AWGN coefficient, the interference power that is inserted into $l^{th}$ PU by the $i^{th}$ subcarrier, interference threshold in PU receiver and frequency distance between $l^{th}$ PU frequency band and the $i^{th}$ subcarrier respectively.

## III. The Proposed Power Allocation Grouping Scheme

Since it is typically very hard to solve jointly the power allocation subject to power and interference constraints, these two constraints are considered separately to reduce the complexity presented here as a second scheme. This second proposed scheme is based on an idle subcarriers grouping approach similar to ladder profile in [9] with a difference that in our proposed scheme the step size is variable and depends on the interference channel gain, pulse shape and frequency distance. So, power efficiency is increased while complexity is decreased in comparison with the existing algorithms.

In the proposed power allocation grouping scheme, all idle subcarriers are divided into four groups based on the inserted interference power level through the PU frequency band. The inserted interference to PU receiver by SU signal depends on SU pulse type, distance frequency, subcarrier power and the interference channel gain. The ith subcarrier power density spectrum assuming an idle Nyquist pulse can be formulated as follows [12]:

$$\phi_i(f) = p_i T_s (\frac{\sin \pi f T_s}{\pi f T})^2 \qquad (2)$$

where $p_i$ and $T_s$ are the ith subcarrier power and the symbol duration respectively. However, the interference power in PU receiver can be written as [9]:

$$I_i^{(l)}(d_{il}, g_i^{sp}, p_i) = g_i^{sp} p_i T_s \int_{d_{il}-\frac{B_l}{2}}^{d_{il}+\frac{B_l}{2}} \phi_i(f)\, df \qquad (3)$$

In order to decrease the algorithm complexity, it is proposed that the interference component in relation (3) is replaced by interference coefficient ($\alpha_i$). This is assigned to each subcarriers based on the interference channel gain and their group index.

$$\alpha_i = \frac{1}{g_i^{sp} c_i} \qquad (4)$$

where $c_i$ and $g_i^{sp}$ are the group index and the interference channel gains of the $i^{th}$ subcarrier respectively. $c_i$ is determined relative to the value of $d_i$ as:

$$c_i = 2^{(d_i-1)}, \qquad d_i = 1,2,3 \qquad (5)$$

where $c_i$ and $g_i^{sp}$ are the group index and the interference channel gains of the $i^{th}$ subcarrier respectively. $c_i$ is determined relative to the value of $d_i$ as:

$$c_i = 2^{(d_i-1)}, \qquad d_i = 1,2,3 \qquad (5)$$

$d_i$ is minimum frequency distance between PU's frequency band and subcarriers where:

$$d_i = \min_l d_{il} \ , \qquad l = 1,\ldots,L \qquad (6)$$

The ith subcarrier power density spectrum is not significant after the fourth zero. However, three neighboring subcarriers are considered as interfering ones. The ith idle subcarrier power density spectrum is not significant after the fourth zero. However, three neighbouring subcarriers are considered as interfering ones. The ladder profile of idle subcarriers based on the effect of interference for the system model in section II is shown in Fig. 2. In order to reduce the algorithm complexity, the subcarrier power is allocated in two stages. In the first stage, the subcarriers are grouped by the frequency distance from PU. Then, the power allocation to PU neighboring subcarriers is performed with subcarriers interference coefficient. This subcarriers interference coefficient is increased by nearing PU frequency band and raising interference channel gain. Finally, the remaining power is allocated to fourth group subcarriers without interference to achieve maximum capacity.

As mentioned previously, power is assigned to group 1 to 3 considering the present interference constraint. This is performed without attention to power constraint and cost function, power is allocated by solving (8) below,

$$\sum_{i=1}^{N} \alpha_i p = I_{th} \quad considering \quad I_{th} = \sum_{l=1}^{L} I_{th}^{l} \qquad (7)$$

Therefore,

$$p = \frac{I_{th}}{\sum_{i=1}^{N} \alpha_i} \qquad (8)$$

In the second stage, water-filling algorithm is deployed to allocate the remaining power to the fourth group subcarriers to

maximize secondary capacity without increasing interference power for PUs.

$$p_i = \max(0, \frac{1}{\lambda} - \frac{n_0}{g_i^{ss}}), \qquad i \in g_4 \tag{9}$$

and $\lambda$ is determined by:

$$\sum_{i \in g_4} \max(0, \frac{1}{\lambda} - \frac{n_0}{g_i^{ss}}) = p_m - p \sum_{i \in g_{1,2,3}} \alpha_i \tag{10}$$

Hence, in general, the power in the ith subcarrier can be written as:

$$p_i = \begin{cases} \alpha_i p, & i \in g_{1,2,3} \\ \max(0, \frac{1}{\lambda} - \frac{n_0}{g_i^{ss}}), & i \in g_4 \end{cases} \tag{11}$$

All the proposed algorithm stages are summarized in Table I.

TABLE I. THE PROPOSED POWER ALLOCATION ALGORITHM STAGES

| step | Process |
|---|---|
| 1. | Define: $g_i^{sp}$, $g_i^{ss}$, $n_0$, $I_{th}$, $L$, $N$, $p_m$. |
| 2. | Calculate $d_i = \min_l d_{il}$, $l = 1,...,L$. |
| 3. | All idle subcarriers is grouped based on $d_i$ in 4 groups and assign $c_i$ to each subcarriers of groups 1,2 and 3. |
| 4. | Calculate $\alpha_i = \frac{1}{g_i^{sp} c_i}$. |
| 5. | Allocate power to group 1,2,3 based on $I_{th}$ constraint and $\alpha_i$. |
| 6. | Allocate power to group 0 based on water filling algorithm. |

The proposed algorithm performance is evaluated with Conventional Water-Filling (CWF) algorithm as well as uniform distribution. In CR scenario, a power allocation scheme should satisfy interference and maximum power constraints. However, the power allocated to idle subcarriers with uniform distribution can be expressed as [7]:

$$p_{uniform} = \min(\frac{p_m}{N}, \frac{I_{th}}{\sum_{i=1}^{N} \sum_{l=1}^{L} K_i^l}) \tag{12}$$

where $K_i^l$ is determined by:

$$I_i^l = p_i \times K_i^l \tag{13}$$

CWF algorithm is performed based on the total power used by the power uniform distribution.

## IV. THE PROPOSED GROUPING SCHEME SIMULATION RESULTS

An OFDM-based CR system with N=128 subcarriers with 4 PUs and single SU is assumed for simulation purposes. Every PU occupies sixteen subcarriers that are distributed between 128 subcarriers. Both main and interference channels are assumed a Rayleigh flat fading model. AWGN variance of the main channel ($\sigma_{n_0}^2$) is assumed to be $10^{-8}$W and the average channel power gains for $g_i^{sp}$, $g_i^{ss}$ are assumed to be 1. The value of $I_{th}$ is $1\times10^{-5}$W. Average transmitted data rates for different algorithms under investigation are obtained by 10,000 independent simulation runs. The proposed algorithm is compared with CWF algorithm and power uniform distribution. The simulation parameters are summarized in Table II.

In Fig. 3a, the achievable transmission data rate for the CR user versus the total power budget for three mentioned algorithms is plotted. By this figure, it is distinguished that the proposed algorithm can achieve transmission rate higher than CWF and uniform distribution algorithms for a given power budget, while injected interference power is unchanged as observed in Fig. 3b. As shown in Fig. 3b, the interference is not increased versus the rising power budget. The reason is that, excess power is assigned to non-interference subcarriers.

Fig. 4 shows power distribution for the proposed algorithm as well as CWF algorithm. This figure also shows that interference is injected into PUs in 33rd and 97th subcarriers by CWF algorithm while the proposed power allocation does not interfere with PUs. Total transmit power in WF algorithm and uniform distribution is limited because of interference constraint since the proposed algorithm operates at maximum power.

Simulation result proves that the proposed algorithm taking into account frequency distance and interference channel gain, is more efficient than conventional water filling or uniform distribution in CR scenarios.

The optimal values of Lagrangian variable are calculated by an iterative search algorithm associated with constraints given in (1). Thus, to solve conventional water-filling, a long processing time is needed. So, the subcarriers grouping approach with two separate stages that can well decrease complexity with grouping subcarriers based on interference power into PU is proposed in this manuscript without iterative algorithm.

TABLE II. SIMULATION PARAMETERS

| ***Simulation Parameter*** | ***value*** |
|---|---|
| Channel type | Rayleigh flat fading |
| $\overline{g^{sp}}$, $\overline{g^{ss}}$ | 1 |
| $\sigma_{n_0}^2$ | $10^{-8}$ watt |
| Number of subcarriers (N) | 128 |
| $p_m$ | $10^{-3}$ to $10^{-2}$ [in Watt] |
| No. SU and PU | 1,4 |
| PU frequency bandwidth | 16 subcarriers |
| Ith of PU receivers | $10^{-5}$ [in mWatt] |

## V. CONCLUSIONS

In this paper, the effects of frequency distance and interference channel gain are simultaneously considered for power allocation problem over CR system. In the proposed algorithm, subcarrier power allocation is deployed in two stages that, primarily power allocation is applied with interference constraints. Then, the remaining power is assigned to non-interfering subcarriers. However, with rising transmit power budget, CR's capacity is increased.

Thus, the proposed algorithm increases power efficiency without increasing the injected interference power for PU receivers. Simulation results prove this claim.

## REFERENCES


[1] Devroye N, Mitran P, Tarokh V, "Limits on communications in a cognitive radio channel," IEEE Commun Mag, vol. 44, pp. 44-49, 2006.

[2] Zhao Q, Sadler B, "Dynamic Spectrum Access: Signal Processing, Networking, and Regulatory Policy," IEEE Signal Proc Mag, vol. 24, pp. 79-89, 2007.

[3] Etkin R, Parekh A, Tse D, "Spectrum Sharing for Unlicensed Bands," IEEE J Sel Area Comm, vol. 25, pp. 517-528, 2007.

[4] Haykin S, "Cognitive radio: brain-empowered wireless communications," IEEE J Sel Area Comm, vol. 23, pp. 201-220, 2005.

[5] Zhao C, Kwak K, "Power/Bit Loading in OFDM-Based Cognitive Networks With Comprehensive Interference Considerations: The Single-SU Case," IEEE T Veh Technol, vol. 59, pp. 1910-1922, 2010.

[6] Y. Rahulamathavan, K. Cumanan, L. Musavian and S. Lambotharan, "optimal subcarrier and bit allocation techniques for cognitive radio networks using integer linear programming," in Proc. IEEE/SP 15th Workshop on Statistical Signal Processing, Sept. 2009, pp. 293-296.

[7] W. Jian , Y. Longxiang , and Liu Xu, "Resource allocation based on linear waterfilling algorithm in CR systems," in Proc. IEEE 7th international conference on wireless communications, networking and mobile computing (WiCOM), Sept. 2011, pp. 1-4.

[8] Z. Tang, G. Wei, Y. Zhu, "Weighted Sum Rate Maximization For OFDM-Based Cognitive Radio Systems," Telecommun. Sys., Vol. 42, pp. 77-84, October 2009.

[9] G. Bansal, J. Hossain , and V.K. Bhargava, "Optimal and suboptimal power allocation schemes for OFDM-based cognitive radio systems," IEEE TRAN. ON WIRELESS COMMU. ,vol. 7, no. 11, pp. 4710–4718, Nov. 2008.

[10] E. Hosseini, A. Falahati, "improving water-filling algorithm to power control cognitive radio system based upon traffic parameters and QoS," Wireless Pers Commun, vol. 70, no. 4, pp. 1747-1759, June 2013.

[11] E. Hosseini and A. Falahati, "Resource Allocation Scheme for Cognitive Radio System Based on COFDM Signaling Considering Secondary User's Channel Uncertainty," accepted in IETE Journal (2013).

[12] T.Weiss, J.Hillenbrand, A.Krohn and F.K.Jondral, "Mutual interference in OFDM-based spectrum pooling systems [C]," in Proc. IEEE Vehicular Technol. Conf., 2004, pp. 1873-1877.

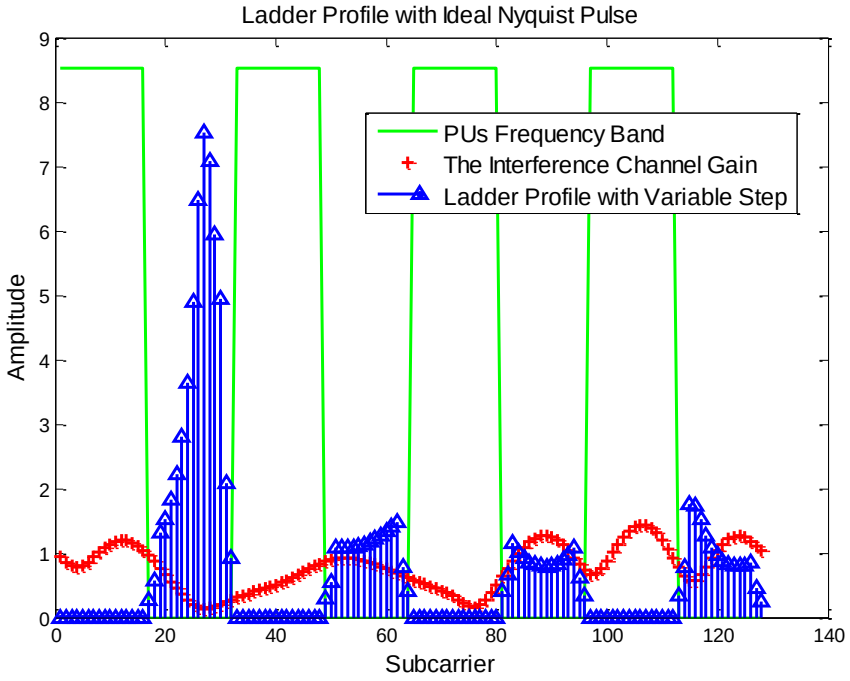
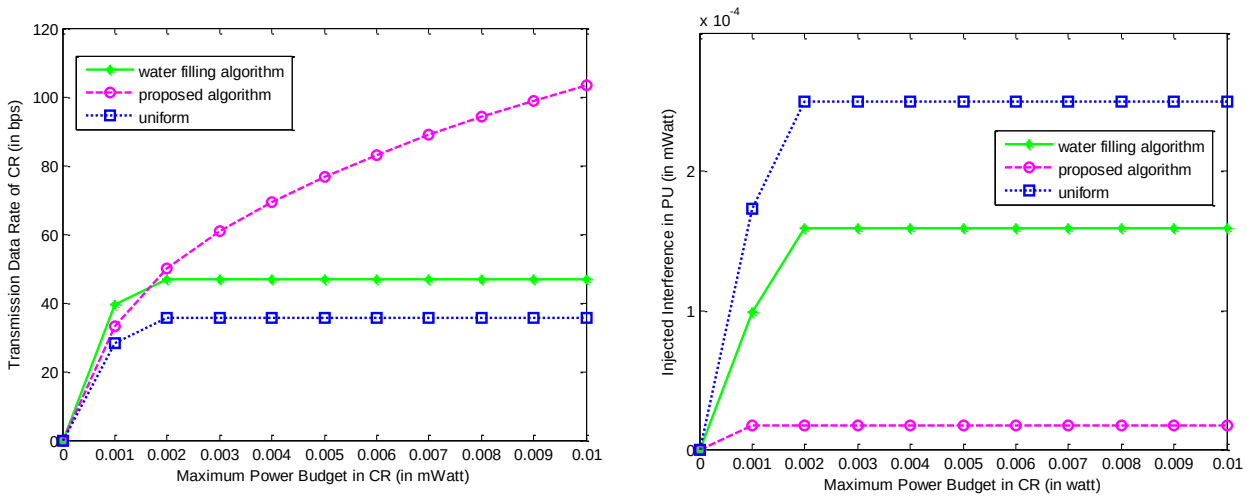




Figure 2. Ladder profile with variable step size for overlay subcarriers




(a) (b)

Figure 3. (a) The comparison of the transmission data rate vs. maximum power for CR with the conventional WF, Uniform and the proposed algorithm (b) The comparison of injected interference in PU vs. maximum power in CR with the conventional WF, Uniform and the proposed algorithm

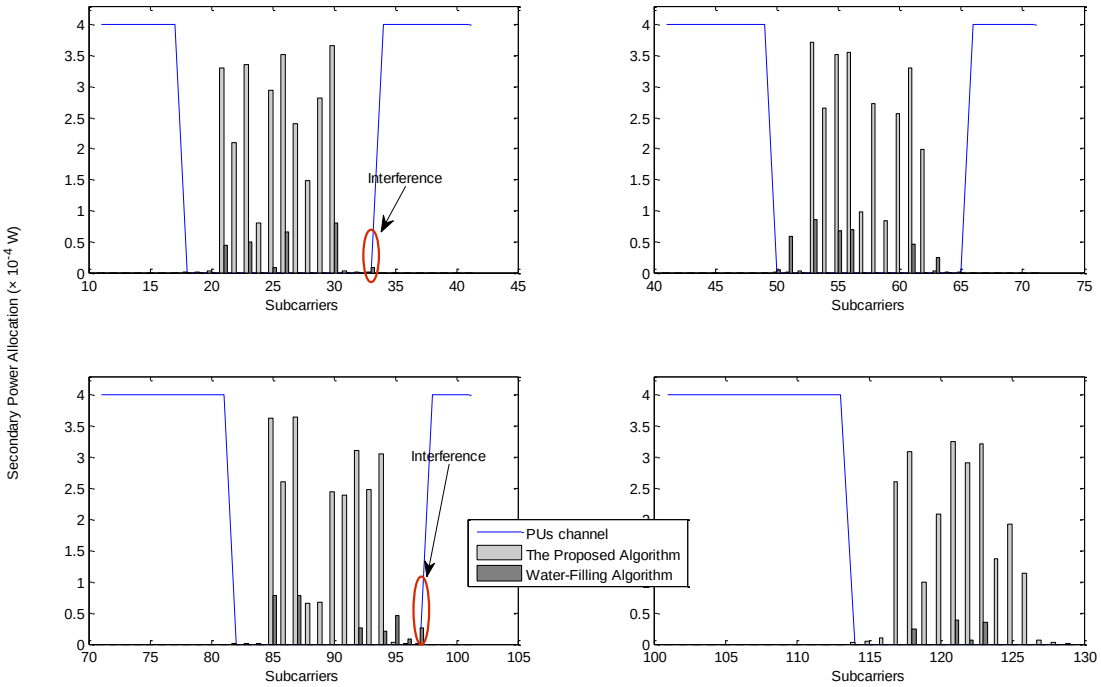



Figure 4. Power allocation scheme for CWF algorithm and the proposed algorithm under investigation

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Proposal for Dispersion Compensating Square-lattice Photonic Crystal Fiber

A. H. Howlader*, Md Asaduzzaman Shobug
Department of Electrical & Electronic Engineering, Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
* E-mail: a_h_howlader@ymail.com

*Abstract*—**A dispersion compensating square-lattice photonic crystal fiber (PCF) for broadband compensation which covers the S, C and L- communication bands i.e. wavelength ranging from 1460 to 1625 nm is proposed in this paper. Theoretically it is shown negative dispersion coefficient of about -595 to -1288 ps/(nm.km) over S to L bands and -975 ps/(nm.km) at the operating wavelength 1550 nm. The relative dispersion slope (RDS) is perfectly matched to that of conventional single mode fiber (SMF) of about 0.0036nm$^{-1}$. Besides the proposed microstructure optical fiber (MOF) shows large non-linear coefficient of 61.88 W$^{-1}$km$^{-1}$ at the operating wavelength 1550 nm. Moreover variation of structural parameters is also studied and discussed here.**

***Keywords—Microstructure optical fiber, dispersion compensating fiber, square-lattice PCF.***

## I. INTRODUCTION

Photonic crystal fiber (PCF) is made of pure silica optical fiber having small air holes arranged in the host silica matrix and goes along the propagation axis. PCFs have gain remarkable attention from the optical scientific society [1]. The most appealing characteristics of PCFs is they can possess dispersion properties which are very different from those of the conventional optical fibers. A good property of PCF is that, the additional design parameters of air hole diameter and hole pitch which gives greater flexibility and independence in the design of dispersion to get the required application for various optical sectors [2]. The chromatic dispersion because of pulse spreading should be compensated in the long distance optical data communication system to minimize the spreading of pulse and this can be compensated by using the dispersion compensating fiber (DCF) having large negative dispersion coefficient [3]. The DCF should be kept as short as possible in length to reduce the insertion loss and the cost. Again the value of negative dispersion coefficient should be as large as possible. The negative dispersion coefficient of DCF should cover a wideband spectrum to successfully and efficiently compensate the dispersion at all the frequencies of dense wavelength division multiplexing (DWDM). Moreover with the dispersion the relative dispersion slope (RDS) should be matched with conventional single mode fiber at the same time [4]. Therefore, it is important to keep special attention on dispersion, dispersion slope, RDS, bandwidth and mode property when designing DCF [5]. Generally the tiny air holes are embedded on the vertex of an equilateral triangle with six air holes in the first ring surrounding the core, this type of PCF is called the hexagonal-lattice PCF (H-PCF) or conventional PCF. Besides the hexagonal structures, other design structures such as square-lattice, octagonal-lattice, decagonal-lattice have been proposed for the MOF. But the demand for a simple dispersion compensating PCF structure still exists in for realizing large negative dispersion coefficient and moderate confinement loss. Previous designs which are proposed all they are based on triangular PCF but proposal for square-lattice PCF with large negative dispersion coefficient over S to L wavelength bands is very few [6].The square-lattice PCF proposed by Nejad et al. [7] has very low negative dispersion coefficient of about -130 to -138 ps/(nm.km) over S to L bands with no RDS match and nonlinearity issues are considered here.

In this paper, we show a square-lattice PCF with six rings of air holes for dispersion compensation over S to L wavelength bands. The attractive property of our proposed PCF is the design simplicity with high negative dispersion coefficient and large nonlinearity which is needed in high-bit-rate transmission and non-linear optics applications. According to simulation work it is shown that, the designed MOF shows high negative dispersion of -975 ps/(nm.km) and high nonlinearity of 61.88 W$^{-1}$km$^{-1}$ at the operating wavelength of 1550 nm.

## II. DESIGN METHODOLOGY

The proposed PCF is made of fused silica and has a square array of air holes running along its length. This is an index

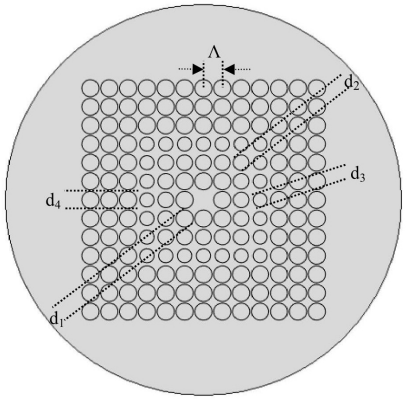



Figure 1. Transverse cross-section of the squre-lattice PCF.

guiding MOF. The transverse cross-section of the PCF is shown in Fig. 1, where $\Lambda$ is the pitch of the lattice, $d_1$ is the air hole diameter of first ring, $d_2$ is the air hole diameter of second ring, $d_3$ is the air hole diameter of third ring and $d_4$ is the air hole diameter of other rings. The air hole diameter of the first ring is relatively large, because it is known that an increase in the air hole diameter of the first ring results in the dispersion coefficient to decrease. The air hole diameters of the second and the third ring are relatively smaller, because a ring of reduced diameter provides a change in the slope of the evolution of the effective refractive index versus wavelength [8]. The total number of air hole rings is chosen to be six in order to simplify as much as possible the structural combination of the PCF.

## III. Numerical Method and Equation

The PCF is simulated by finite element method (FEM) with perfectly matched layer (PML). The FEM directly solves the Maxwell equations to best approximate the value of the effective refractive index. If the modal effective refractive index, $n_{eff}$ is obtained by solving an eigen value problem came from Maxwell equations using the COMSOL MULTIPHYSICS 4.2, chromatic dispersion and effective area of PCF can be easily calculated.

Chromatic dispersion $D$ can be obtained using the following relation [9]

$$D(\lambda) = -(\lambda/c) \times (d^2 Re[n_{eff}]/d\lambda^2) \quad (1)$$

Where *Re [neff]* is the real part of effective refractive index $n_{eff}$ , $\lambda$ is the wavelength, $c$ is the velocity of light in vacuum. The material dispersion given by Sellmeier formula is directly included in the calculation. Therefore, $D$ in (1) corresponds to the chromatic dispersion of the PCF.

The confinement loss $Lc$ is obtained from the imaginary part of $n_{eff}$ as follows [9]

$$Lc = 8.686 \times k_0 \times (Im[n_{eff}]) \quad (2)$$

where $Im[n_{eff}]$ is the imaginary part of the refractive index, $k_0 = 2\pi/\lambda$ is the wave number in the free space.

The effective area $A_{eff}$ is calculated as follows [10]

$$A_{eff} = \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |E|^2 dxdy \right)^2 / \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |E|^4 dxdy \quad (3)$$

Where, $E$ is the electric field derived from the Maxwell equations.

The nonlinear co-efficient $\gamma$ is calculated as follows [10]

$$\gamma = (2\pi/\lambda)\,(n_2/A_{eff}) \times 10^3 \quad (4)$$

Where, $n_2$ in (4) is the nonlinear index coefficient in the nonlinear part of the refractive index and $A_{eff}$ is the effective area.

The relative dispersion slope (RDS) is calculated as follows [11]

$$RDS = S_{SMF}(\lambda)/D_{SMF}(\lambda) = S_{DCF}(\lambda)/D_{DCF}(\lambda) \quad (5)$$

Where, $S_{SMF}(\lambda)$ and $S_{DCF}(\lambda)$ are the dispersion slope of the SMF and DCF respectively. The RDS of SMF is 0.0036 $nm^{-1}$.

## IV. Simulation Results and Discussion

The effective index curve for both x and y-polarized (optimum) modes of the PCF is shown in Fig. 2.

The dispersion characteristics of the proposed PCF for its optimum value is shown in Fig. 3. According to simulation result, the proposed PCF shows negative dispersion coefficient of about -595 to -1288 ps/(nm.km) over S and L-bands for y-polarization. The dispersion value of the proposed PCF at 1550 nm is about -975 ps/(nm.km) which far exceeds the dispersion values of conventional dispersion compensating single mode fiber [12] [typically -100 ps/(nm.km)] and [7].

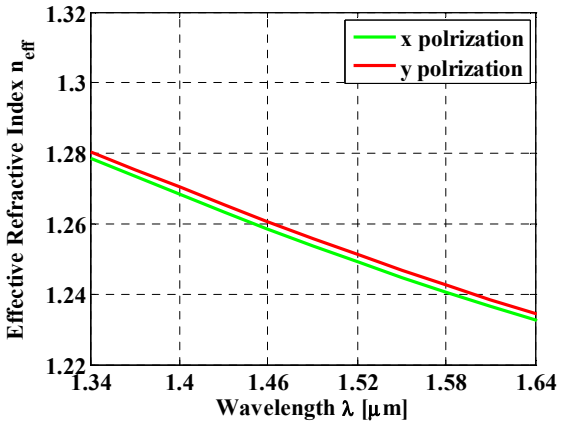



Figure 2. Effective refractive index curve of the PCF as a function of wavelength for $\Lambda$=0.81μm, $d_1/\Lambda$=0.91, $d_2/\Lambda$=0.83, $d_3/\Lambda$=0.75, $d_4/\Lambda$=0.91.

The dispersion characteristics of the proposed PCF for optimum parameters, variation of pitch, variation of first ring radius, variation of second ring radius and third ring radius are shown in Fig. 3, Fig. 4, Fig. 5, Fig. 6 and Fig. 7 respectively for y-polarization while remaining parameters are kept constant.

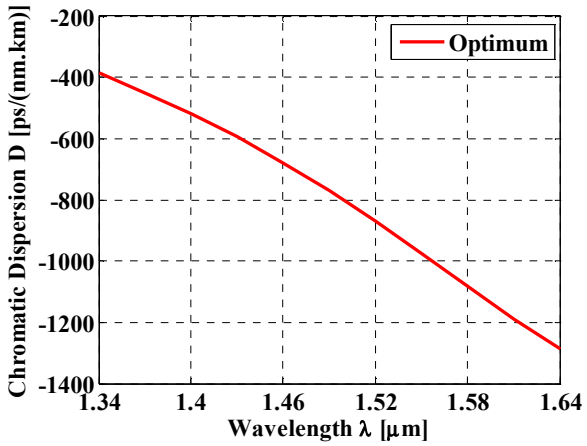



Figure 3. Chromatic dispersion curve of the PCF as a function of wavelength for y-polarization and optimum condition $\Lambda$=0.81μm, $d_1/\Lambda$=0.91, $d_2/\Lambda$=0.83, $d_3/\Lambda$=0.75, $d_4/\Lambda$=0.91.

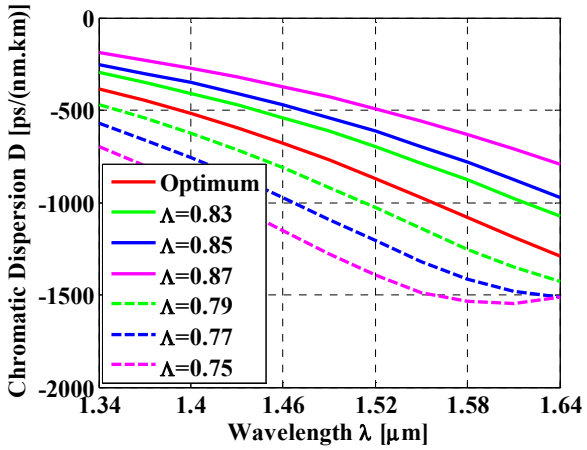



Figure 4. Chromatic dispersion curve of the PCF as a function of wavelength for pitch variation.

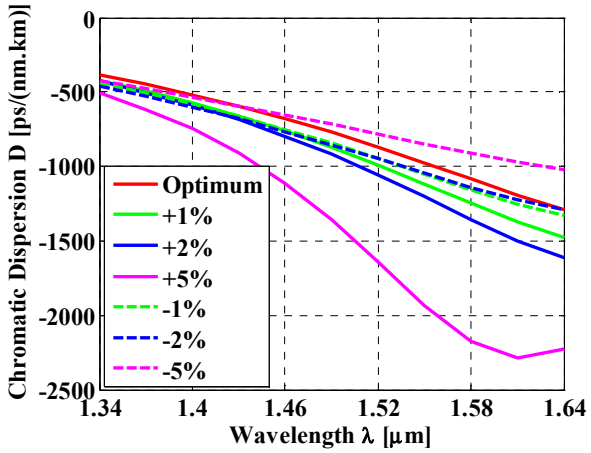



Figure 5. Chromatic dispersion curve of the PCF as a function of wavelength for first ring variation.

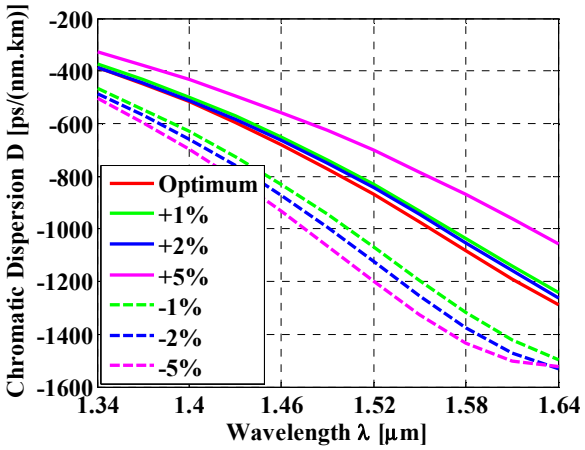



Figure 6. Chromatic dispersion curve of the PCF as a function of wavelength for second ring variation.

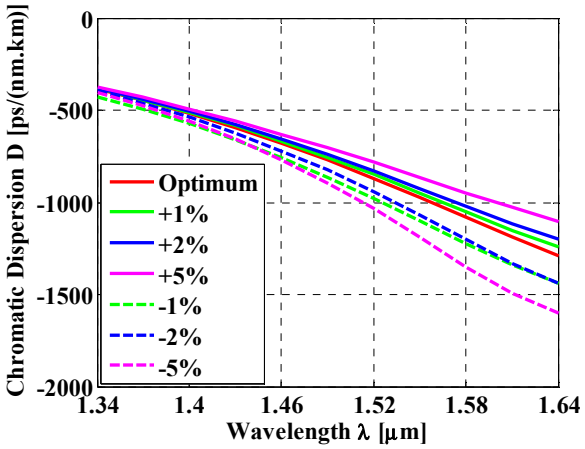



Figure 7. Chromatic dispersion curve of the PCF as a function of wavelength for third ring variation.

The dispersion slope, RDS, residual dispersion of the PCF for optimum condition is shown in Fig. 8, Fig. 9 and Fig. 10 respectively. From Fig. 8 it is shown that the proposed PCF has better dispersion slope of -2.55 to -3.26 for S to L wavelength bands. From Fig.9 it is shown that perfectly matched RDS of 0.003599 $nm^{-1}$ of conventional SMF at the operating wavelength of 1550 nm. Fig. 10 shows the calculated residual dispersion obtained after the dispersion compensation by a 0.71 km long optimized MOF for the dispersion considered in one span (40 km long) of the transmission SMF fiber. It can be showed that the residual dispersion value which ranges from -39.42 to -26.78 ps/nm in the wavelength range of 1430 to 1640 nm enables the proposed PCF to be a better candidate for high-bit-rate transmission systems of S, C and L band.

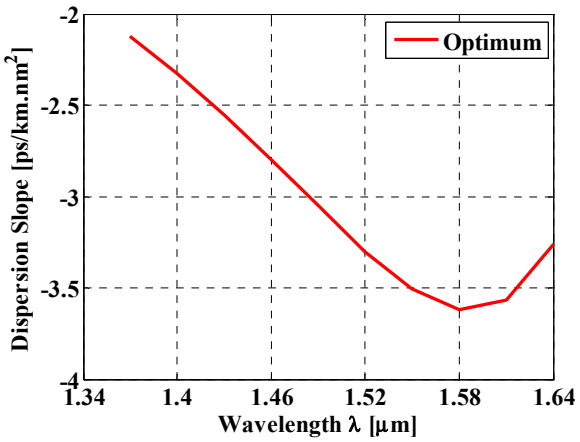



Figure 8. Dipersion slope curve as a function of wavelength for y-polarization and optimum condition $\Lambda$=0.81µm, $d_1/\Lambda$=0.91, $d_2/\Lambda$=0.83, $d_3/\Lambda$=0.75, $d_4/\Lambda$=0.91.

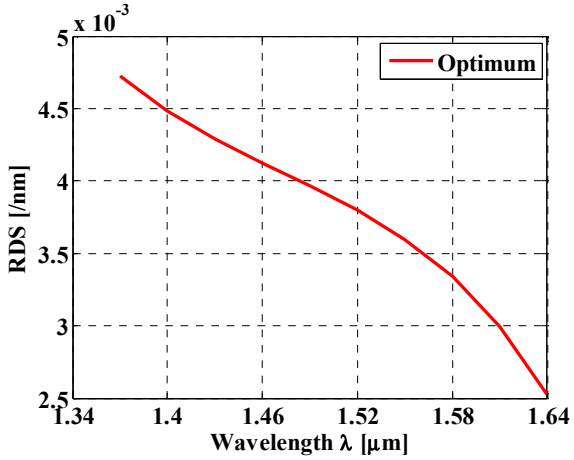



Figure 9. RDS curve as a function of wavelength for y-polarization and optimum condition $\Lambda$=0.81µm, $d_1/\Lambda$=0.91, $d_2/\Lambda$=0.83, $d_3/\Lambda$=0.75, $d_4/\Lambda$=0.91 which shows perfectly matched RDS of with SMF at wavelength 1550 nm.

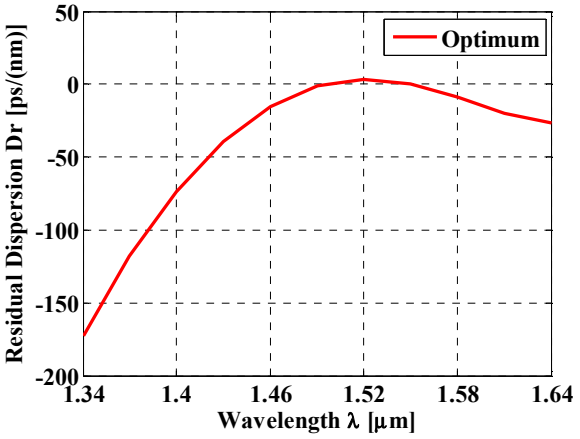



Figure 10. Residual dispersion curve as a function of wavelength for y-polarization and optimum condition $\Lambda$=0.81µm, $d_1/\Lambda$=0.91, $d_2/\Lambda$=0.83, $d_3/\Lambda$=0.75, $d_4/\Lambda$=0.91.

Fig. 11 represents effective area as a function of wavelength and Fig.12 represents wavelength dependence of nonlinear coefficient, $\gamma$ as a function of wavelength for optimum design parameters. From these figures it is found that the proposed PCF is highly non-linear and the non-linear coefficient is found 61.88W$^{-1}$Km$^{-1}$ for y-polarization at operating wavelength 1550 nm, which is higher than [11] and suitable for optical parametric amplification, supercontinuum generation and soliton pulse generation.

Confinement loss of the proposed PCF for y-polarization is shown in Fig.13 which shows moderate confinement loss at the operating wavelength of 1550 nm.

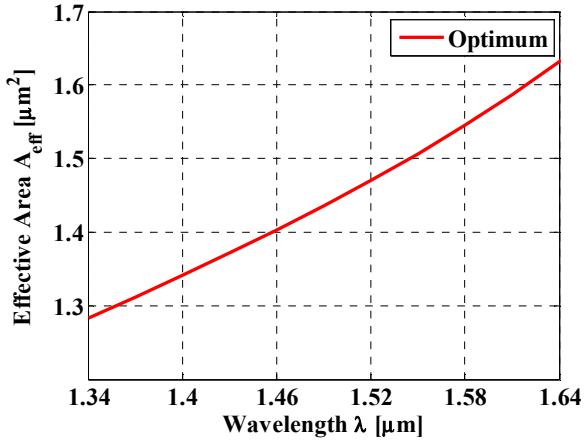



Figure 11. Effective area curve as a function of wavelength for y-polarization and optimum condition $\Lambda$=0.81µm, $d_1/\Lambda$=0.91, $d_2/\Lambda$=0.83, $d_3/\Lambda$=0.75, $d_4/\Lambda$=0.91.

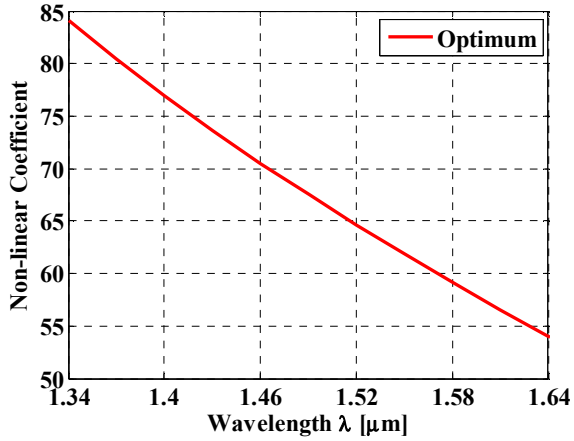



Figure 12. Non-linear coefficient curve as a function of wavelength for y-polarization and optimum condition $\Lambda$=0.81µm, $d_1/\Lambda$=0.91, $d_2/\Lambda$=0.83, $d_3/\Lambda$=0.75, $d_4/\Lambda$=0.91.

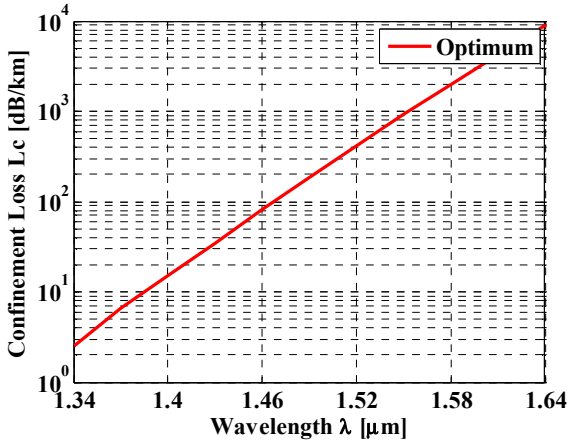



Figure 13. Confinement loss curve as a function of wavelength for y-polarization and optimum condition $\Lambda$=0.81μm, $d_1/\Lambda$=0.91, $d_2/\Lambda$=0.83, $d_3/\Lambda$=0.75, $d_4/\Lambda$=0.91.

## V. CONCLUSION

Here we have proposed a square-lattice PCF that has high negative dispersion coefficient with perfectly matched dispersion slope with conventional SMF and high non-linear coefficient. This proposed MOF offers high negative dispersion coefficient of about -595 to -1288 ps/(nm.km) over S to L-bands and -975 ps/(nm.km) at the operating wavelength 1550 nm and it offers high non-linear coefficient of 61.88 $W^{-1}km^{-1}$ at the operating wavelength 1550 nm also. Another special advantage is that, compared with previously presented square-lattice PCF, the design procedure and geometrical structure is very simple because relatively fewer geometrical parameters are used. So, it can be hope that our proposed PCF will be useful and successful in dispersion compensation for broadband transmission application and nonlinear optics application.

## REFERENCES

[1] S. K. Varshney, N. J. Florous, K. Saitoh, M. Koshiba, T. Fujisawa, " Numerical investigation and optimization of a photonic crystal fiber for simultaneous dispersion compensation over S+C+L wavelength bands," *Optics Communications*, vol 274, pp.74-79, 2007.

[2] Jingyuan Wanga, Chun Jianga, Weisheng Hua, Mingyi Gao, "Modified design of photonic crystal fibers with flattened dispersion," *Optics & Laser Technology*, vol.38, pp.169-172, 2006.

[3] M. Koshiba, K. Saitoh, "Structural dependence of effective area and mode field diameter for holey fibers", *Opt. Exp*., 11 (2003) 1746-1756.

[4] S. G. Li, X. D. Liu, L. T. Hou, "Numerical study on dispersion compensating property in photonic crystal fibers", *Acta Phys. Sin.*, 53 (2004) 1880–1886.

[5] B. Zsigri, J. Laegsgaard, A. Bjarklev, "A novel photonic crystal fibre design for dispersion compensation", *J. Opt. A Pure Appl. Opt*., 6 (2004) 717–720.

[6] TAN Xiao-ling, GENG You-fu, TIAN Zhen, WANG Peng, and YAO Jian-quan, "Study of ultra-flattened dispersion square-lattice photonic crystal fiber with low confinement loss," *Optoelectronics Lerrers,* vol.5, no.2, pp.0124-0127, 2007.

[7] Nejad, S.M.; Ehteshami, N., "A novel design to compensate dispersion for square-lattice photonic crystal fiber over E to L wavelength bands," *Communication Systems Networks and Digital Signal Processing (CSNDSP), 2010 7th International Symposium on* , vol., no., pp.654,658, 21-23 July 2010.

[8] Zinan Wang, Xiaomin Ren, Xia Zhang, Yongzhao Xu and Yongqing Huang, "Design of a microstructure fiber for slopematched dispersion compensation," *J.Opt.A: Pure Appl.Opt*, vol.9, pp.435-440, 2007.

[9] K. Saitoh, M. Koshiba, T. Hasegawa, and E. Sasaoka, "Chromatic Dispersion control in photonic crystal fibers: Application to ultra flattened dispersion,"*Opt. Express*, vol. 11, pp. 843–852, 2003.

[10] S. M. A. Razzak, Y. Namihira, F. Begum, S. Kaijage, N. H. Hai, and N. Zou, "Design of a decagonal photonic crystal fiber with ultra-flattened chromatic dispersion," IEICE Trans. Electron., vol. E90-C, no. 11, pp. 2141-2145, Nov. 2007.

[11] M. Selim Habib, M. Samiul Habib, S. M. Abdur Razzak, and M. anwar Hossain, "Proposal for Highly Birefringent Broadband Dispersion Compensating Octagonal Photonic Crystal Fiber," *Opt. Fiber technol*., (http://dx.doi.org/10.1016/j.yofte.2013.05.014), May 2013.

[12] Feroza Begum, Yoshinori Namihira, S.M. Abdur Razzak, Shubi Kaijage, Nguyen HoangHai, Tatsuya Kinjo, Kazuya Miyagi, Nianyu Zou, "Novel broadband dispersion compensating photonic crystal fibers: Applications in high-speed transmission systems," *Optics & Laser Technology*, vol.41, pp. 679-686, 2009.

1st International Conference on Electrical Information and Communication Technology (EICT 2013)
13-15 February 2014, Khulna-9203, Bangladesh

# Design of a Square Lattice Photonic Crystal Fiber for Dispersion Compensation over Telecom Bands

A. H. Siddique[1,*], Redwan Ahmad[1], Sharafat Ali, M. A. Islam[1], K. M. Nasim[1] , E. Khandker[1], M. Samiul Habib[1]
[1]Department of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
*E-mail: alamgir_7846@yahoo.com

*Abstract*—**In this paper, we numerically demonstrate a large negative dispersion with highly birefringent square lattice photonic crystal fiber (SLPCF) in the entire telecom wavelength bands. Finite element method with perfectly matched layer boundary condition is used to evaluate the modal properties of the fiber. Numerical results reveal that it is possible to obtain a large negative dispersion coefficient of -897 ps/(nm.km), a relative dispersion slope (RDS) close to that of single mode fiber (SMF) of about 0.0036 $nm^{-1}$ and birefringence of the order $1.34\times10^{-2}$ at 1.55 μm. To evaluate the tolerance of fabrication variation of structural parameters from their optimum value is carried out.**

***Keywords—photonic crystal fiber, negative dispersion, effective area, confinement loss.***

## I. INTRODUCTION

Photonic crystal fibers (PCFs) having microscopic array of air channels helps in tuning dispersion slope and controlling confinement losses in a way which was not possible by conventional fiber [1], [2] . Flexibility in tuning dispersion is essential in designing a dispersion compensation fiber which is only offered by PCFs [3], [4]. In long distance optical data transmission system, the dispersion of group velocity in optical links is one of major impairments of pulses. Due to this impairment various dispersion compensation techniques like dispersion compensation gratings, optical phase conjugation and electronic dispersion compensation are extensively required. Among these, only dispersion compensation fiber (DCF) due to its widespread application has been installed worldwide [5], [6].

The magnitude of negative dispersion should be large and size of DCFs should be as short as possible to minimize insertion loss and reduce the cost. On the other hand dispersion and dispersion slope should be compensated in single mode fiber at the same time [7]. It is necessary to consider dispersion, dispersion slope, relative dispersion slope, bandwidth and mode property to design a DCFs [8] .The idea of using PCF for dispersion compensation (DC) purpose was first proposed by Birks et al [9]. PCFs having high birefringence with high negative dispersion coefficient are important in sensing application.

Various designs of single material PCFs have been proposed for achieving high negative dispersion coefficient [10], but they fail to achieve negative dispersion coefficient larger than -600 ps/(nm.km) and a compensation bandwidth wider than C band. The proposed fiber in [11] has a high negative dispersion coefficient of -1455 ps/(nm.km) , but the effective area is very small. However several attempts have been made by another group to attain high effective area but their resultant negative dispersion coefficient was around -588 ps/(nm.km) [12]. A honeycomb structure PCF with a Ge-doped central core has been proposed for a wide compensation bandwidth and a large dispersion coefficient which can reach -1350 ps/(nm.km) , but the doped core will lead to fabrication difficulties [13]. Another drawback of germanium doping in the core of conventional solid DCFs is higher fiber loss and a significant increase in nonlinearity due to high doping concentrations [14]. Recently a proposed genetic algorithm procedure in a PCF for DC design attained a peak dispersion of -500 ps/(nm.km) with 14 air-hole rings [15]. Huttumen et al. [16] proposed a dual-core PCFs with dispersion peak -59,000 ps/(nm.km) and effective area of 10 μm² but no attempts to match the relative dispersion slope (RDS) to that of conventional SMF was made [17] .

In this paper, we propose a PCF which successfully compensate the dispersion in most widely used telecom bands. The designed fiber shows large negative dispersion coefficient of about -897 ps/(nm.km), RDS of about 0.0036 $nm^{-1}$ high birefringence of about $1.34\times10^{-2}$ at 1.55 μm. We numerically investigate the dispersion compensation characteristics of the PCF over various bands. We believe that this proposed PCF will be promising candidate and highly applicable to optical communication and sensing application.

## II. GEOMETRY OF PROPOSED SLPCF

The proposed SLPCF is made of pure silica surrounding with circular air-holes in the fiber cladding. Fig.1 shows the geometry of proposed dispersion compensating fiber with optimized air hole diameter $d_1$, $d_2$, $d_3$ and pitch Λ .The air-hole diameter of first, fourth and fifth rings are same as $d_1$. The structure has one missing air-hole in first ring. The spacing between air-holes on same ring of the square structure is same as the pitch Λ. In the designed SLPCF total number of air-holes for ring 1, 2, 3, 4, 5 are respectively 7, 16, 24, 32 and 40 .This provides a higher air-filling ratio and lower refractive index around core which provide strong confinement ability [12]. The air-holes of the first ring are relatively large in comparison with that of second and third. This increase in diameter of first ring causes high negative dispersion peak [16]. The diameter of air-holes of second ring is very smaller, which causes changes in the slope of the dispersion [18].

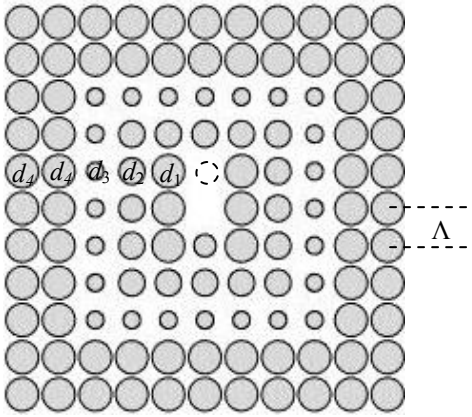



Figure 1. Cross section of the proposed PCF.

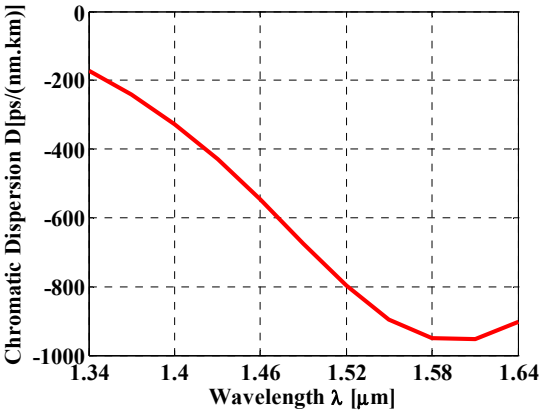



Figure 2. Wavelength response of chromatic dispersion of proposed SPCF for optimum parameter .

The outer rings have insignificat influence on dispersion [19], so those rings are set to have same air-holes diameter.

## III. Simulation Technique & Equations

The finite element method (FEM) with circular perfectly matched boundary layers (PML) is used to calculate chromatic dispersion, confinement loss, residual dispersion and effective area of SLPCFs. To make zero reflection at boundary, an efficient boundary condition is essential to use [12] . PMLs are the most efficient absorption boundary conditions for this purpose [20] .Once the modal effective index $n_{eff}$ can be calculated by solving an eigenvalue problem drawn from the Maxwell equations using FEM. The chromatic dispersion $D(\lambda)$, confinement loss $L_c$, and effective area $A_{eff}$ can be calculated given by [4] . PCFs with polarizing maintaining (PM) properties are essential in various applications such as in sensing applications, in stabilizing the operation of optical devices and in eliminating the effect of polarization mode dispersion (PMD). In PCFs birefringence properties are more imperative for polarization maintaining application.

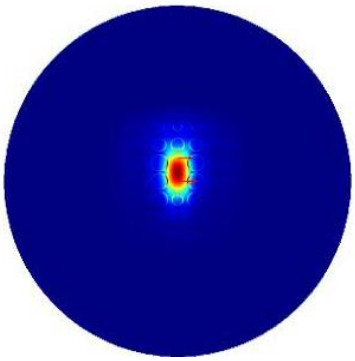


Figure 3. Mode profile of proposed PCF at 1550 nm.

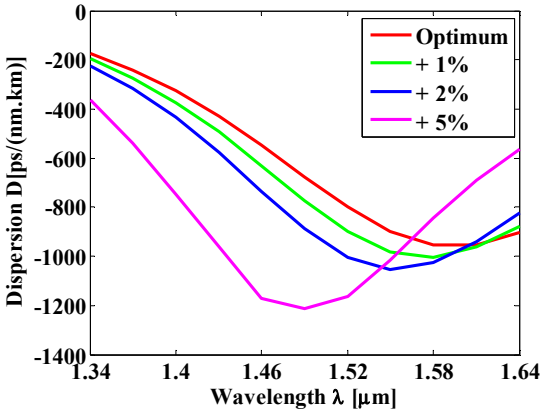



Figure 4. Effect of $d_1$ on dispersion behavior.

The term birefringence can be defined as [21].

$$B = \left| n_x - n_y \right| \qquad (1)$$

Where $n_x$ and $n_y$ represent the mode indices of the two orthogonal polarization modes.

The PCF cross sections, with a fixed number of air-holes are divided into homogeneous subspaces where Maxwell's equations are solved by accounting for the adjacent subspaces. These subspaces are triangles that permit a good approximation of the PCF structures. From Maxwell's curl equations we can obtain the following vectorial equation [22]

$$\nabla \times ([s]^{-1} \nabla \times E) - k_0^2 n_{eff}^2 [s] E = 0 \qquad (2)$$

where $E$ is the electric field vector, $k_0$ is wave number in the vacuum, $n$ is the refractive index, $[s]^{-1}$ is an inverse matrix of $[s]$ and $\lambda$ is the operating wavelength. The effective refractive index is given as $n_{eff} = \beta/k_0$, where $\beta$ is the propagation constant, $k_0 = 2\pi/\lambda$ is the free-space wave number.

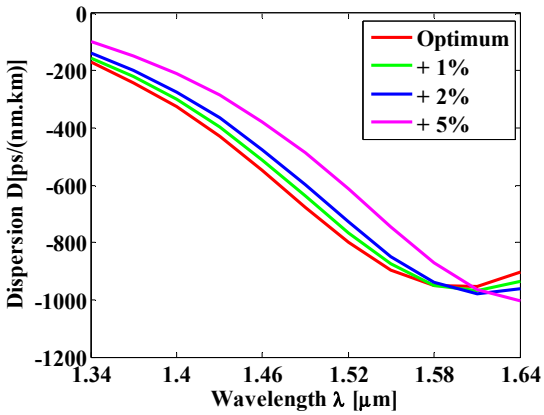



Figure 5. Effect of $d_2$ on dispersion behavior.

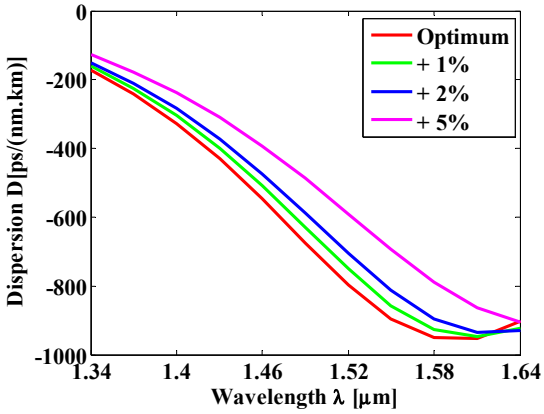



Figure 6. Effect of $d_3$ on dispersion behavior

A long distant optical fiber transmission system usually has a positive dispersion of about 10 ps/(nm.km) . As there stays a positive dispersion and dispersion slope at the SMF, the elementary need of a DCF for WDM operation is the negative dispersion as large as possible to reduce the length of DCF which leads to a reduced cost. To compensate the dispersion of the SMF over a particular range of the wavelengths, the mentioned expression must be satisfied [23],

$$RDS = S_{SMF}(\lambda)/D_{SMF}(\lambda) = S_{DCF}(\lambda)/D_{DCF}(\lambda) \qquad (3)$$

Where, $S_{SMF}(\lambda)$ and $S_{DCF}(\lambda)$ are the dispersion slope of the SMF and DCF respectively. The relative dispersion slope RDS of SMF is 0.0036 $nm^{-1}$ and unit of RDS is $nm^{-1}$. The design of the DCF can be corroborated when RDS of the proposed DCF is exactly equal or very close to that of SMF.

## I. Simulation Result

Varying the different structural parameters the dispersion properties of proposed SLPCF are observed. It is shown in fig. 1, there are four degrees of freedom ($d_1$, $d_2$, $d_3$ and $\Lambda$) in proposed design procedure. Influence of those four parameters on dispersion curve is investigated by adjusting them separately. The dispersion coefficient and dispersion slope at

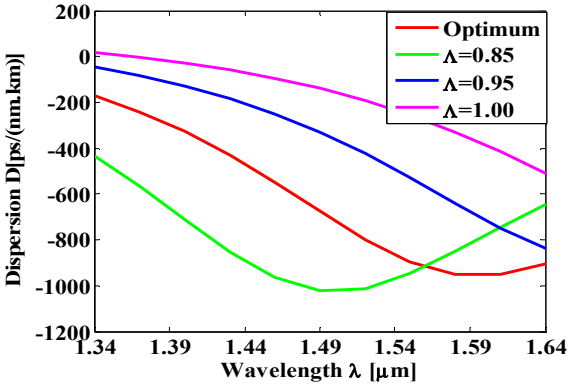



Figure 7. Optimum dispersion and effects of changing pitch Λ.

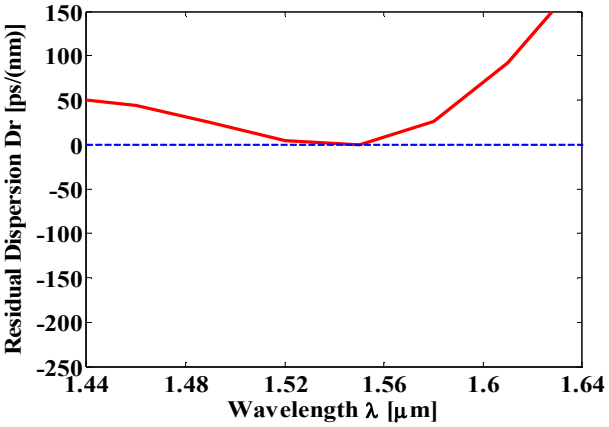



Figure 8. Optimum dispersion and effects of changing pitch Λ.

1.55 µm of SMF is 17 ps/(nm.km) and 0.06 ps $nm^{-1}$ $km^{-1}$ respectively and the calculated RDS is 0.0036 $nm^{-1}$ [22]. Optimizing these parameters negative dispersion coefficient of -897 ps/(nm.km) is obtained at 1.55 µm. Fig. 3 shows the fundamental mode profile of the proposed fiber at 1.55 µm. The positions of air holes and the mode field confinement are shown in this figure. Firstly the optimum dispersion is obtained according the way described in fig. 2 , then we have checked the dispersion accuracy of the proposed design . It is well known that in a standard fiber draw , ± 1% variations in fiber global diameter may occur [23] during fabrication process. That's why roughly an accuracy of ± 2% may require to ensure dispersion tolerance [19] . Fig. 2 shows the effect of first ring diameter ($d_1$) on dispersion behavior with $\Lambda$ = 0.90 µm , $d_1 / \Lambda = 0.889$, $d_2 / \Lambda = 0.711$, $d_3 / \Lambda = 0.22065$ . The air-hole diameter $d_1$ is varied up to +5% from its optimum value and the Corresponding dispersion curves are shown in fig. 4. With the varied values of $d_1$, at 1.55 µm the calculated dispersion coefficient is -897 ps/(nm.km) , -982 ps/(nm.km) , -1052 ps/(nm.km) , -1014 ps/(nm.km) and the RDS is 0.0036 , 0.0028 , 0.0015 , 0.0048 $nm^{-1}$ with +1% , at 1.55 µm the

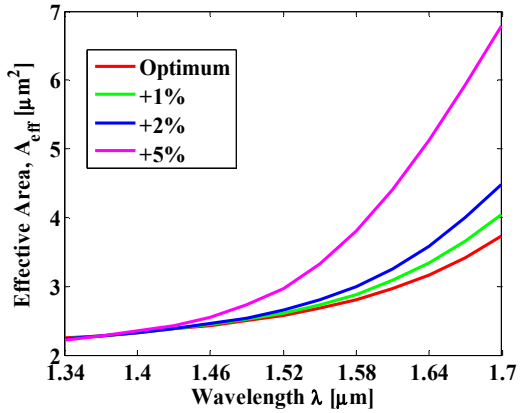
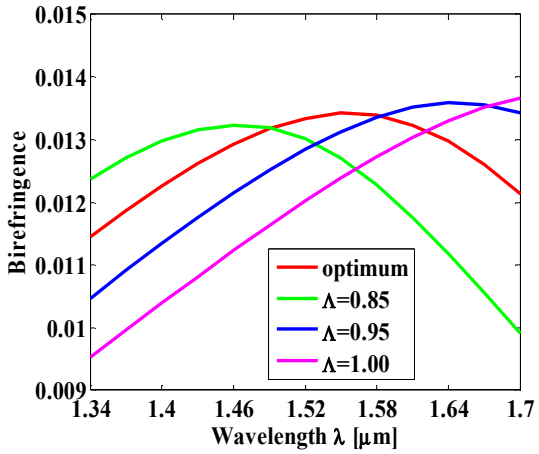



Figure 9. Effective area with air-hole diameter of first ring varying up to




+5%.

Figure 10. Birefringence with pitch, Λ variations.

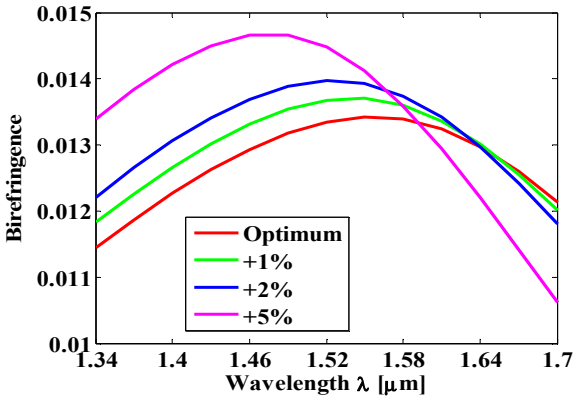



Figure 11. Birefringence with $d_1$ variation.

dispersion coefficient is -876 ps/(nm.km) , -849 ps/(nm.km) , -747 ps/(nm.km) and RDS is 0.0042 , 0.0047 , 0.0059 $nm^{-1}$ for +1%, +2% & +5% variation respectively from their optimum values. In fig. 6 the effects of variation of air-holes diameter of third ring on the dispersion characteristics are shown. Effect of variation of diameters up to +5% on dispersion are investigated and shown in this figure. With varied value of $d_3$, at 1.55 μm the dispersion coefficient is -856 ps/(nm.km) , -814 ps/(nm.km) , -694 ps/(nm.km) and RDS is 0.0041 , 0.0044 , 0.0050 $nm^{-1}$ for +1%, +2% & +5% variation respectively. Fig. 7 shows the effects of pitch, Λ variation on dispersion of proposed PCF. The pitch varies from 0.85 to 1.00 μm with optimum value 0.90 μm. The corresponding results are shown in fig.7. Fig. 8 shows the residual dispersion obtained after the dispersion compensation by 0.774 km long Fiber for dispersion accumulated in one span (40 km) of PCF. It will cover the S, C, L band. The effective area can be calculated using [21] which is shown in fig. 9. Here the optimum parameter is used and first ring diameters are varied from +1% to +5% . At 1.55 μm the calculated effective area is 2.7 μm². It changes dynamically for +5% changes in parameters. Though its mode field diameter (MFD) is very small, it must show difficulties input and output coupling of light [12] .To reduce splicing problem, the PCFs have to splice with conventional SMFs in a specially constructed manner [23]. Fig. 10 and fig.11 shows birefringence with the variation of pitch Λ and air-hole diameters of first ring .With optimum parameter birefringence is very high and in order of $1.34\times10^{-2}$ at 1.55 μm. Birefringence is lower for pitch rather than optimum but it dynamically increases with varying d1 to 5%. This high birefringence is very much useful in stabilizing the operation of optical devices [24]. So the proposed PCF with modest number of design parameter, high negative dispersion, high birefringence and large effective area may help in various application of optics.

## V. Conclusions

A highly negative dispersion SLPCF with high birefringence has been proposed which can be efficiently used in the entire telecom bands. It has been shown through numerical simulation results that, a five ring SLPCF can assume a negative dispersion in order of- 897 ps/(nm.km) along with high birefringence of $1.34\times10^{-2}$ at 1.55 μm. Moreover, this fiber has a modest number of design parameters, five rings, three air-hole diameters and an air-hole pitch. Due to large negative dispersion and high birefringence, the proposed fiber could be potential candidate for optical communication and sensing applications.

## References


[1] J. C. Knight, ‘‘Photonic crystal fibers,’’ *Nature* 424(2003) 847-851.

[2] J. C. knight, T. A. Birks, P. S. J. Russell, and D. M. Aktin, “All-silica single-mode optical fiber with photonic crystal cladding,” *Opt. Lett.* 21(1996) 1547-1549.

[3] S. M. A. Razzak, and Y. Namihira, “Proposal for highly nonlinear dispersion-flattened octagonal photonic crystal fibers,” IEEE Photon. Technol. Lett. 20 (2008) 249-251.

[4] K. Saitoh, M. Koshiba, T. Hasegawa, and E. Sasaoka, “Chromatic dispersion control in photonic crystal fibers: application to ultra-flattened dispersion,” *Opt. Exp.* 11(2003) 843-852.

[5] Z. Wang, X. Ren, X. Zhang, Y. Xu, and Y. Huang, “Design of a microstructure fiber for slope-matched dispersion compensation,” J. Opt. A: Pure Appl. Opt. 9(2007) 435-440.

[6] L. Gruner-Nielson , M. Wandel, P. Kristensen, C. Jorgensen, L. V. Jorgensen, B. Edvold, B. Palsdottir, and D. Jakobsen, "Dispersion –compensating fibers ," *J. Lightw . Technol*. 23(2005) 3566-3579.

[7] S. G. Li, X. D. Liu, and L. T. Hou , "Numerical study on dispersion compensating property in photonic crystal fibers," Acta Phys. Sin. 53 (2004) 1880–1886.

[8] B. Zsigri, J. Laegsgaard, and A. Bjarklev, "A novel photonic crystal fibre design for dispersion compensation," *J. Opt. A Pure Appl. Opt.* 6 (2004) 717–720.

[9] T. A. Birks, D. Mogilevtsev, J. C. Knight, and P. St. J. Russell, "Dispersion compensation using single-material fibers," *IEEE Photon. Technol. Lett.* 11 (1999) 674-676.

[10] B. Zsigri, J. Lægsgaard and A. Bjarklev, "A novel photonic crystal fibre design for dispersion compensation," *J. Opt. A: Pure Appl. Opt.* 6(2004) 717-720.

[11] M. Selim Habib, M. Samiul Habib, S. M. A. Razzak, M. I. Hasan, R. R. Mahmud, and Y. Namihira, "Microstructure holey fibers as wideband dispersion compensating media for high speed transmission system ," *Optik*, (doi:10.1016/j.ijleo.2013.03.128).

[12] M. Selim Habib, M. Samiul Habib, S. M. A. Razzak, and Md. Anwar Hossain ,"Proposal for highly birefringent broadband dispersion compensating octagonal photonic crystal fiber," *Optical Fiber Technol*., (doi.org/10.1016/j.yofte.2013.05.014), May 2013.

[13] S. Yang, Y. Zhang, L. He, and S. Xie, "Broadband dispersion-compensating photonic crystal fiber," Opt. *Lett*., 31 (2006) 2830–2832.

[14] P. J. Roberts, B. J. Mangan, H. Sabert, F. Couny, T. A. Birks, J. C. Knight and P. St. J. Russel l, "Control of dispersion in photonic crystal fibers," J. Opt. Fiber Commun ., *Rep*. 2 (2005) 435-461.

[15] T. Fujisawa, K. Saitoh, K. Wada, and M. Koshiba, "Chromatic dispersion profile optimization of dual concentric-core photonic crystal fibers for broadband dispersion compensation," *Opt. Exp.* 14 (2006) 893-900.

[16] F. Poli, A. Cucinotta, S. Selleri, and A. H. Bouk, "Ailoring of flattened dispersion in highly nonlinear photonic crystal fibers," *IEEE Photon. Technol. Lett.*, 16 (2004) 1065-1067.

[17] S. Yang, Y. Zhang, X. Peng , Y. Lu, S. Xie, J. Li , W. Chen , Z. Jiang , J. Peng, and H. Li, "Theoretical study and experimental fabrication of high negative dispersion photonic crystal fiber with large area mode field," *Opt. Exp.*, 14 (2006) 3015-3023.

[18] F. Gérôme, J.-L. Auguste, and J.-M. Blondy, "Design of dispersion-compensating fibers based on a dual- conce ntric-core photonic crystal fiber," *Opt. lett*., 29 (2004) 2725-2727.

[19] F. Poletti, V. Finazzi, T. M. Monro, N. G. R. Broderick, V. Tse, and D. J. Richardson, "Inverse design and fabrication tolerances of ultra-flattened dispersion holey fibers," *Opt. Exp.*, 13(2005) 3728–3736.

[20] S. Guo, F. Wu, S. Albin, H. Tai, and R. S. Rogowski, "Loss and dispersion analysis of micro structured fibers by finite difference method," *Opt. Exp.*, 12 (2004) 3341–3352.

[21] S. F. Kaijage, Y. Namihira, N. H. Hai, F. Begum, S. M. A. Razzak, T. Kinjo, K. Miyagi, and N. Zou, "Broadband dispersion compensating octagonal photonic crystal fiber for optical communication applications," *Japanese J. of Appl. Phy.*, 48 (2009) 052401-052408.

*[22]* K. Saitoh, M. Koshiba, Full-vectorial imaginary-distance beam propagation method based on finite element scheme : Application to photonic crystal fibers, *IEEE J. Quantum Electron, 38 (2002) 927-933.*

*[23]* K. Kaneshima, Y. Namihira, N. Zou, H. Higa, Y. Nagata, Numerical investigations of octagonal photonic crystal fibers with strong confinement confinement field, *IEICE Trans, Electron, E89-C (2006) 830-837* .

[24] M. Selim Habib, M. Samiul Habib, S. M. A. Razzak, Y. Namihira, M. A. Hossain, and M. A. G. Khan, "Broadband dispersion compensation of conventional single mode fibers using microstructure optical fibers," *Optik-Int. J. Light Electron Opt.* (2013) http://dx.doi.org/10.1016/j.ijleo.2012.12.014.

[25] W. H. Reeves, J. C. Knight and P. S. J. Russell, "Demonstration of ultra-flattened dispersion in photonic crystal fibers," *Opt. Exp.,* 10 (2002) 609–613.

[26] M. Y. Chen, "Polarization and leakage properties of large-mode-area micro structured-core optical fibers," *Opt. Exp*., vol. 15, no. 19, pp. 12498–12507, 2007.

# Enhanced Tolerance of Rayleigh Backscattering in WDM-PONs by using Coded RZ Modulation

P. K. Choudhury
*Dept. of Electronics and Communication Engineering*
*Khulna University of Engineering & Technology (KUET)*
*Khulna-9203, Bangladesh.*
*E-mail: pallab@ece.kuet.ac.bd*

***Abstract*- Rayleigh backscattering (RB) noise is a fundamental limitation for a centralized loopback WDM-PONs based on RSOA in ONU side. By exploring the RSOA chirp induced spectral boarding with DC-balanced line coding, here a novel modulation format called as Coded RZ is proposed to enhance the system tolerance against RB. The results showed that the signal-to-noise (SCR) is improved by about 8 dB and 2.5 dB from the conventional NRZ and RZ modulation formats respectively to maintain the power penalty ≤ 2dB.**

***Keywords- Reflective Semiconductor Optical Amplifier, Rayleigh Backscattering Noise, Passive Optical Networks.***

## I. INTRODUCTION

To meet the growing bandwidth demand from end users, fiber based access solution called as Passive Optical Networks (PONs) have attracted much attention recently [1]. The network development is mostly driven by cost effectiveness in the case of system design, maintenance and future up-gradation. One of the well-investigated network architecture is colorless optical network unit (ONU) based PON system [2]. Due to its colorless property, ONU does not require any wavelength specific optical source, which helps to deploy an identical module in each access unit and also reduces the system implementation complexity. A promising candidate for the design of colorless PONs is the Reflective Semiconductor Optical Amplifier (RSOA) based single fiber loopback network. Under this network scenario as shown in Fig. 1, a CW light transmits from the central office (CO), propagates over the fiber, modulated and reflected by ONU and finally sends back to the CO. In fact, RSOA itself is a key device in this reflective ONU architecture due to its properties of optical gain, wide optical bandwidth, and integration capability. However, as the downstream CW signal and upstream modulated data share the same fiber, the system performance is inherently limited by Rayleigh Backscattering (RB) noise [3]. The interference between upstream data signal and RB components produce the optical beating noise in CO which is highly coherent near the DC frequency as also shows in Fig. 1. This in-band coherent noise severely degrades the system performance if the Signal to Crosstalk Ratio (SCR) reach below the acceptable label.

In recent years, several proposals have been investigated to enhance the system tolerance against this RB noise [4-8]. The major concern about all of the above

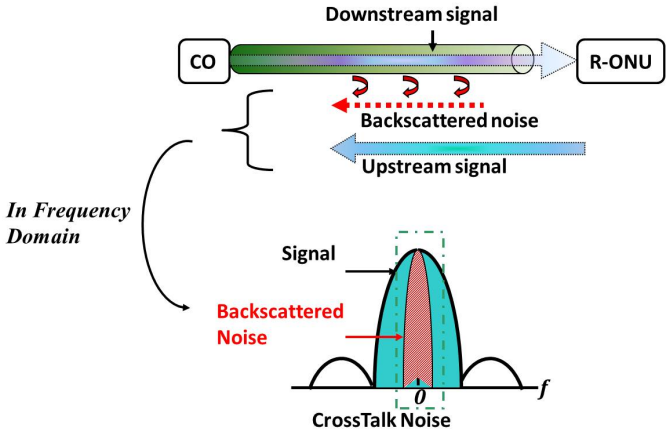



Fig. 1. Reflective ONU based PON architecture with spectral shape of crosstalk noise

proposals was to reduce the coherence time of the interfering signals so that crosstalk becomes incoherent and finally cancel out. Use of bias-dithering of RSOA can improve the system performance in the presence of RB noise [4], at the expense of maintaining the dithering frequency twice of data rate. Another simplified approach is to use of post detection high pass electrical filter [5] and with the addition of DC-balanced line coding [6] can also ensure a good noise tolerance. However, both the techniques require a suitable choice of filter cut-off frequency that depends on data rate and line-width of seeding signal [7]. Furthermore, in [8] a scheme based on external phase modulation in ONUs has exploited the spectral broadening of upstream signal to minimize the overlapping between signal and noise in optical domain. However, the technique substantially increases the complexity of ONU, which eventually affects the cost of network deployment. While all of the above investigations were used the conventional Non Return to Zero (NRZ) pulse format for RSOA modulation, recently an effective approach has been proposed by using the Return-to-Zero (RZ) format [9]. As the number of phase transitions of RZ format is higher than the NRZ format, it improves the chirp induced spectral broadening for direct modulation of RSOA. This behavior ultimately ensures the incoherent interference between signal and noise to enhance the system performance. However, the above solution still provide a limited performance as the conventional PRBS coded RZ format has significant amount of low frequency components, which can increase the system

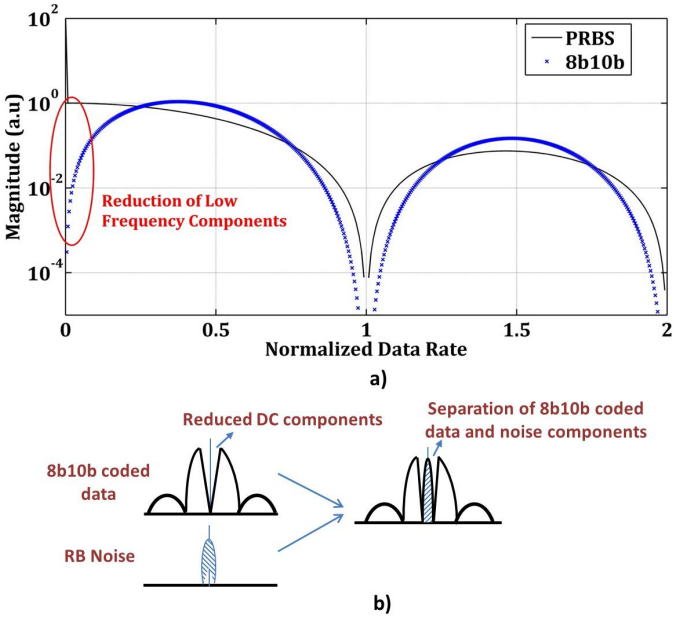



Fig. 2. a) Power spectral density of PRBS and 8b10b coded data. b) Crosstalk noise reduction process by separating the noise and 8b10b coded data stream.

penalty in the presence of DC like RB noise.

To overcome the above limitation, here a modified RZ modulation format, called as Coded RZ (CoRZ), is proposed by using 8b10b coded data stream with RZ pulse shape. This novel modulation format not only ensures an effective reduction of low frequency crosstalk components but also preserves a sufficient spectral boarding to make the interfering signals incoherent. The system performance has been studied by comparing the NRZ, RZ and CoRZ modulation formats. The result shows that the proposed CoRZ modulation format can effectively increase the tolerance of SCR label by about 8 dB from NRZ and 2.5 dB from RZ to maintain the power penalty $\leq$ 2dB.

## II. Working Principle

According to [10], the power spectral density (PSD) of RB noise signal is approximately identical to source signal. Therefore, the spectral shape of the noise shows a narrow line-width with high amount of low frequency components when it is generated from CW seeding source. These frequency components are mainly responsible to generate a concentrated noise near DC (as from Fig. 1) if the conventional PRBS coded, as it has almost 50% power in DC, data format is used. However, an alternative coded sequence is useful in this case like 8b10b coding. The 8b10b line coding was invented by IBM corporation and popularly used in fiber channel as well as Gigabit-Ethernet (GbE). Due to its DC balance property, the aforementioned line coding can significantly reduce the frequency components near DC compare to PRBS coded data stream as shown in Fig. 2(a). Therefore, the 8b10b coded data stream effectively separates the noise from information contents to minimize the crosstalk noise as illustrated in Fig. 2(b). On the other hand, instead of NRZ pulse shape, RZ pulse shows a higher chirping effect for RSOA modulation, which further decrease the coherence time of signal as already investigated in earlier studies. Thus, a combination of 8b10b coded data stream with RZ pulse has a good potential to enhance the system resilience against RB noise limited optical system.

## III. System Model for Crosstalk Tolerance

The Fig. 3 shows the schematic used for the measurement of system penalty against different crosstalk levels in the presence of RB noise from CW source. The transmission system design and analysis were carried out by using optical system design simulation tool VPITransmissionMaker- 9.0®. The CW light from DFB laser (Center frequency=193.1THz, 10 MHz line-width) was divided into two paths, firstly the path (top) for RB noise and secondly, the path (bottom) for modulated data signal. The RB noise was extracted from optical circulator (OC-1) when the CW light passed through the 30 Km of fiber span. Since the optical power at fiber input was fixed to 0 dBm, the RB noise label was found to a value of -33 dBm as maintained constant throughout this work. A polarization controller (PC) was also used in the crosstalk path to maximize the noise interference (degree of polarization for RB signal is 33%). To generate the modulated data signal, the bottom path was used for RSOA modulation with 3 types of data signals i.e., Non Return to Zero (NRZ), Return-to-Zero (RZ) and propose Coded Return to Zero (CoRZ).

In the simulation model, RSOA input saturation power and

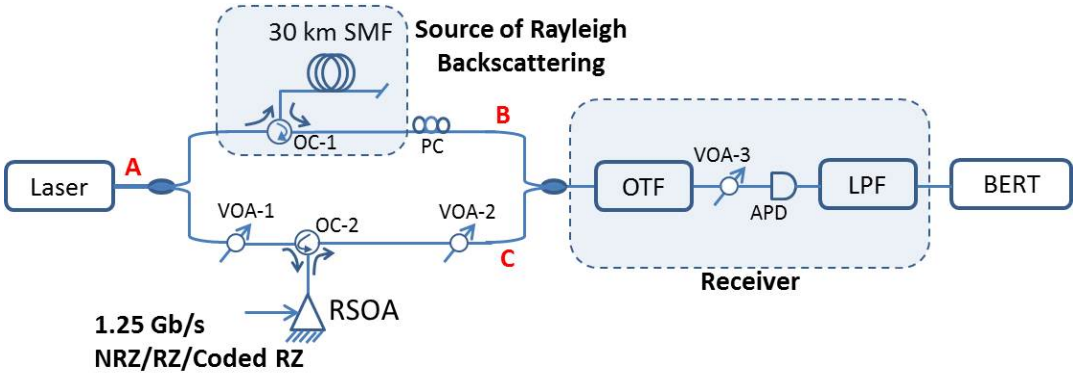



Fig. 3. Schematic to analysis the system crosstalk tolerance; OC = Optical circulator, VOA = Variable optical attenuator, PC = Polarization controller, OTF = Optical tunable filter, LPF = Low pass filter, APD = Avalanche photodiode

gain were approximately -15 dBm and 21 dB, respectively. In addition, E/O modulation bandwidth was around 1.5 GHz. The line-width enhancement factor (chirp parameter) was 5 (five), a typical value for RSOA device [11]. The RSOA bias current was 40 mA and driven by 1.25 Gb/s PRBS sequence with NRZ or RZ (50% duty cycle) pulse shape. For coded RZ modulation, 8b10b coded sequence was used with RZ pulse shape at 1.25 Gb/s data rate. The extinction ratio (ER) was fixed at 6 dB for all modulation formats. The input power of RSOA was fixed at -18 dBm by using Variable Optical Attenuator (VOA-1). Note that the above choice of RSOA input power ensures the RSOA operation in linear region, which eventually helps to maintain good power budget in downstream transmission for real deployment. The output modulated signal of RSOA was extracted from OC-2 and further attenuated by VOA-2. As the RB noise power was fixed to -33 dBm, different levels of signal to crosstalk ratio (SCR) can be achieved by controlling the optical power of modulated signal through VOA-2.

Finally, both the signal components were combined and proceed to optical receiver for BER measurement. The optical receiver consists of optical tunable filter (Gaussian shape, bandwidth 25 GHz), VOA-3, Avalanche Photodiode (10 dB gain and sensitivity of -33.5 dBm at 1.25 Gb/s) and electrical Low Pass Bessel Filter (Bandwidth of 75% of data rate). The OTF was tuned to center frequency of laser and useful to avoid the out of band ASE noise that comes from RSOA. By controlling the VOA-3, the system power penalty can be measured at different SCR values to reach the BER=$10^{-9}$.

## IV. Results and Discussion

The Fig. 4 shows the optical spectrums at different positions of schematic in Fig. 3. As predicted [10], the backscattered noise signal preserves the same shape of incident CW light source but with reduced power level of 33 dB compare to laser power. Moreover, the crosstalk noise power concentrates more near DC frequency and extends up to the line-width of the laser source. The Fig. 4(c) shows the modulated spectrums for NRZ, RZ and CoRZ formats. Due to the residual phase modulation of RSOA, RZ modulation shows a significant spectral boarding compare to conventional NRZ format. And thus, this chirp induced spectral boarding can effectively reduce the coherence between the noise and data signals to improve the tolerance against crosstalk. However, the conventional PRBS coded RZ format has significant low frequency components and therefore, limit the further improvement of system performance. To overcome this limitation, the proposed CoRZ shows the expected wide spectral shape with desire reduction of low frequency components. Therefore, the CoRZ modulation format gives better resilience against the RB noise which can be justified by the eye diagrams as shown in Fig. 5. The eye diagrams were taken for NRZ, RZ, and CoRZ modulation format with different SCR values i.e., 21, 18, 15 dB. The SCR values were determined by varying (use of VOA-2) optical power of modulated signals while by keeping the backscattered noise power fixed at -33 dB. It shows that RZ modulation format greatly improved the system performance compare to

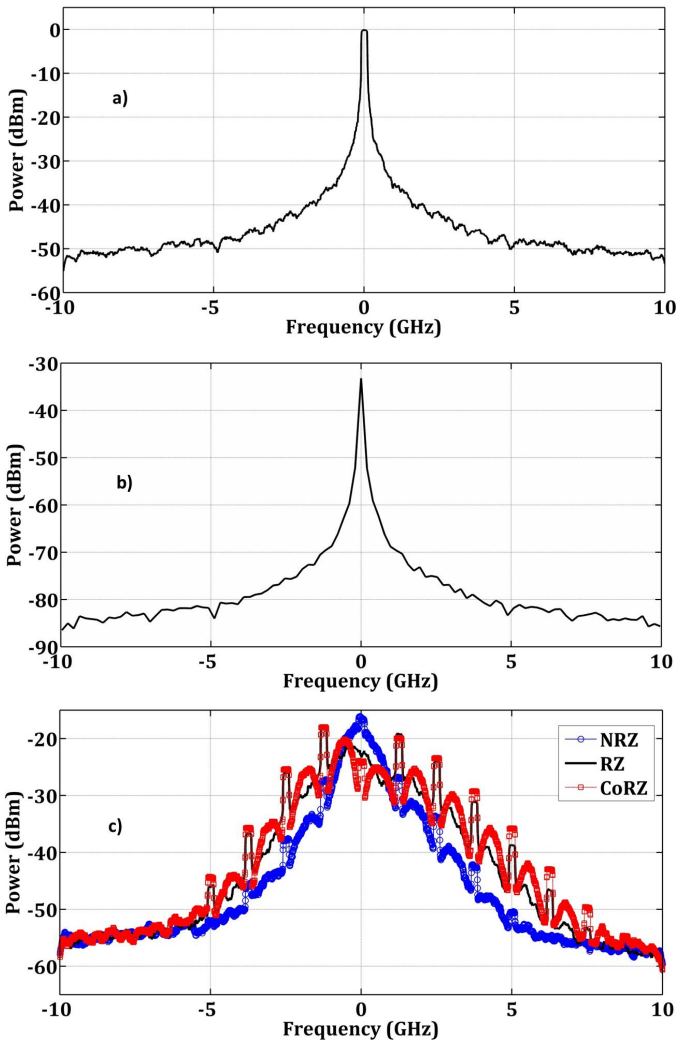



Fig. 4. Optical spectrum at different positions of Fig. 3; a) Laser output (point-A), b) Backscattered signal (point-B), c) Modulated spectrum (point-C).

conventional NRZ modulation even at low SCR value of 18 dB. However, the propose CoRZ outperforms both the RZ and NRZ formats and shows much wider eye opening at a very low SCR value of 15 dB. Therefore, the proposed CoRZ format is suitable to use in RB noise limited optical system.

To get into further insight of propose method, the system performance was also measured in terms of power penalty at BER=$10^{-9}$ for various crosstalk levels. The obtained results are shown in Fig. 6. When the conventional NRZ modulation format was used for RSOA modulation, the system power penalty was rapidly increased and reaches on error floor as SCR close to 20 dB. This behavior agrees well with previous study on rayleigh backscattering noise limited system [12]. With the application of RZ modulation, system tolerance against crosstalk was improved by around 4 dB and shows the BER floor when SCR approaches to 16 dB. As mentioned earlier, this improvement comes from the interplay between RSOA chirp and phase transitions of RZ format. However, the system shows further improvement when the proposed Coded RZ signal was applied.

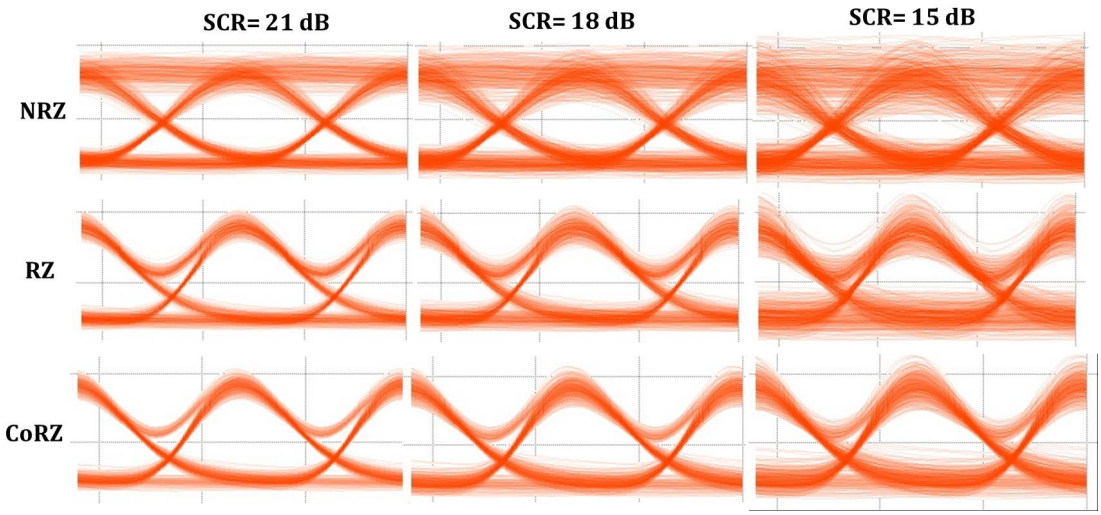



Fig. 5. Eye diagrams of different modulation formats (NRZ, RZ, CoRZ) at the SCR levels of 21 dB, 18 dB and 15 dB for 1.25 Gb/s transmission.

With the application of CoRZ modulation format, error free transmission was achieved at SCR down to 13 dB. Moreover, at a given power penalty of 2 dB, the system SCR can be now reduced by 8 dB (from NRZ penalty) and 2.5 dB (from RZ penalty) due to application of CoRZ. This confirms the effectiveness of the proposed solution to reduce in-band crosstalk generated by rayleigh back-scattering in CW seeded reflective PON architecture.

## V. CONCLUSIONS

A novel modulation format is demonstrated for reducing the RB noise in RSOA based loopback WDM-PONs. The proposed solution is obtained for chirped RSOA modulation by 8b10b coded data with RZ pulse shape. The obtained esults show that the proposed technique gives substantial improvement in recovery of power penalty compare to conventional NRZ and RZ modulation formats.

## ACKNOWLEDGMENT

The work was supported by the HEQEP project of University Grants Commission of Bangladesh and World Bank under the grant CP-2378.

## REFERENCES

[1] E. Wong, "Next-generation broadband access networks and technologies," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 597-608, 2012.

[2] J. Prat, C. Arellano, V. Polo, and C. Bock, "Optical network unit based on a bidirectional reflective semiconductor optical amplifier for fiber-to-the- home networks," *IEEE Photon. Technol. Lett.*, vol. 17, no.1, pp. 250-252, 2005.

[3] Arellano, C., Langer, K., and Prat, J.: 'Reflections and multiple Rayleigh backscattering in WDM single-fiber loopback access networks', *J. Lightwave Technol.*, vol. 27, no.1, pp. 12-18, 2009.

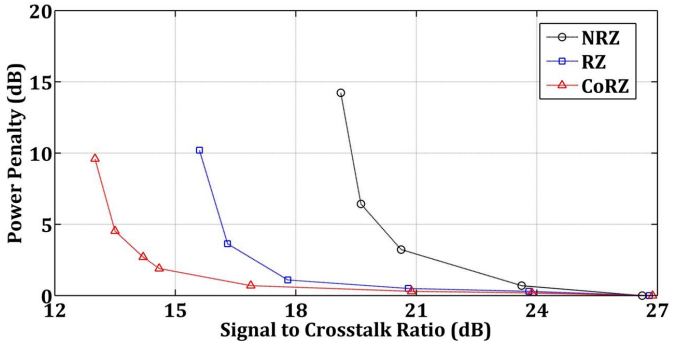



Fig. 6. Power penalty as a function of crosstalk levels for the modulation formats NRZ, RZ, and CoRZ.

[4] Urban, P.J., Koonen, A.M.J., Khoe, G.D., and de Waardt, H.: 'Interferometric crosstalk reduction in an RSOA-based WDM passive optical network', *J. Lightwave Technol.*, vol. 22, pp. 4943, 2009.

[5] Marki, C.F., Marki, A.F., and Esener, S.C.: 'Reduction of interferometric optical crosstalk penalty via DC blocking', *Electron. Lett*., vol. 43, no. 11, pp. 644-646, 2007.

[6] Chiuchiarelli, A. *et al.*, 'Enhancing resilience to Rayleigh crosstalk by means of line coding and electrical filtering', *IEEE Photonics Technol. Lett.,* vol. 22, no. 2, pp. 85-87, 2010.

[7] D. Jorgesen, C. F. Marki, and S. Esener, "Improved high pass filtering for passive optical networks," *IEEE Photon. Technol. Lett.,* vol. 22, no.8, pp. 1144–1146, 2010.

[8] A. Chowdhury *et al.*, "Rayleigh backscattering noise-eliminated 115-km long-reach bidirectional centralized WDM-PON with 10-Gb/s DPSK downstream and remodulated 2.5-Gb/s OCS-SCM upstream signal," *IEEE Photon. Technol. Lett.*, vol. 20, no. 24, pp. 2081-2083, 2008.

[9] L. Banchi, R. Corsini, M. Presi, F. Cavaliere, and E. Ciaramella, "Enhanced reflection tolerance in WDM-PON by chirped RZ modulation," Electron. Lett., vol. 46, no. 14, pp. 1009 -1011, 2010.

[10] P. Gysel and R. K. Staubli, "Statistical properties of Rayleigh backscattering in single-mode fibers," *J. Lightwave. Technol.*, vol. 8, no. 4, pp.561–567, Apr. 1990.

[11] I. Cano, M. Omella, J. Prat, P. Poggiolini, "Colorless 10Gb/s extended reach WDM PON with low BW RSOA using MLSE," in *Proc. OFC 2010*, Paper OWG2.

[12] G. Talli, D. Cotter, and P. D. Townsend," Rayleigh backscattering impairments in access networks with centralised light source," *Electron. Lett.*,vol.42,p.877,2006.

# Design of Hybrid Photonic Crystal Fibers for Tailoring Dispersion and Confinement Loss

Md. Sharafat Ali[1,*], Aminul Islam[1], Redwan Ahmad[1], A. H. Siddique[1], K M Nasim[1], M A G Khan[1], M Samiul Habib[1]
[1]Dept of Electrical and Electronic Engineering, Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh.
[*] E-mail: sharafat.ali@ieee.org

*Abstract*— **This paper presents the proposal of a hybrid cladding photonic crystal fiber offering flat dispersion and low confinement operating in the Telecom bands. Simulation results reveal that near zero ultra flattened dispersion of 0 ± 1.20 ps/(nm.km) is obtained in a 1.25 to 1.70 µm wavelength range i.e. 450 nm flat band along with low confinement losses which is less than $10^{-2}$ dB/km at operating wavelength 1.55 µm. Moreover, the sensitivity of the fiber dispersion properties to a ±1% to ±5% variation in the optimum parameters is studied for practical conditions.**

***Keywords— photonic crystal fiber, confinement loss, chromatic dispersion, effective area.***

## I. INTRODUCTION

Photonic crystal fibers (PCFs) also known as holey fibers (HFs) are usually glass fibers having minuscule array of air channels running alone its entire length that makes the low index cladding covering the undoped fused silica core [1]. Unlike the traditional fiber both the core and cladding are made from the same material in PCFs and light can be well confined and guided properly through the fiber by the mechanism of either total internal reflection (TIR) or photonic band gap (PBG) [2]. Main advantages of using PCFs over conventional fiber is flexibility of controlling chromatic dispersion, nonlinearity, birefringence by tuning different parameters of the holey fiber [3].

Several designs for the PCFs have been proposed to achieve the nearly zero ultra-flattened chromatic dispersion properties and low confinement. The decagonal fiber in [4] has a flat band of 200 nm, but fabrication of decagonal structure is challenging hence less attractive in telecom applications, while design in [5] has 350 nm flat bands which is much lower than our designed fiber. Here our proposed design has 450 nm flat band which is much higher than the previously proposed designs.

In this paper, we suggest a hybrid design of PCF that shows better dispersion accuracy having a flat dispersion in a 450 nm band along with ultra-flattened dispersion of 0 ± 1.20 ps/(nm.km) and confinement loss less than $10^{-2}$ dB/km in the entire band of interest.

## II. DESIGN METHODOLOGY

We propose a hybrid six ring structured dispersion flattened Holey fiber with optimized air-hole diameter $d_1$, $d_2$, $d$ and pitch Λ. The first and second air-hole diameter is assigned as $d_1$ and $d_2$ while rest of rings diameter are same and denoted by $d$. The air-hole pitch is Λ. The first two rings are arranged in a hexagonal shape, on the other hand rest of rings are put in circular respect. The refractive index of fiber silica is $n_s$=1.45 and refractive index of air-hole is $n_a$=1 at 1.55 µm. At the last four circular rings the air-holes are put respectively $30^0$, $22.5^0$, $15^0$ & $15^0$ apart.

In this structure there are four parameters for controlling dispersion behavior, namely $d_1$, $d_2$, $d$ and Λ. We kept the diameter of first and second ring air holes lower with respect to outer four circular rings for proper dispersion behavior and air hole diameter of the outer circular four rings is kept large for better field confinement and for reducing confinement loss.

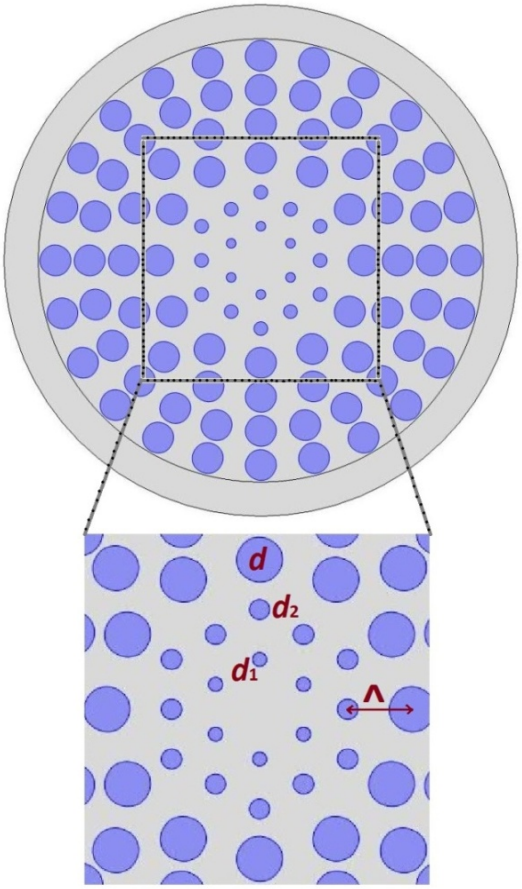



Figure 1. Air-hole distribution of the proposed Hybrid Structure with number of ring, $N_r$ = 6.

## III. SIMULATION TECHNIQUE AND EQUATION

We have used COMSOL 4.2 software to investigate the guiding properties this PCF. Effective refractive index, chromatic dispersion, effective area, confinement loss are calculated by using finite element method (FEM) with perfectly matched boundary layers (PML). The FEM directly solves the Maxwell equations to get an approximate value of the effective refractive index. Once the effective refractive index, $n_{eff}$ of the model is obtained by solving Maxwell equations using the software, effective area, chromatic dispersion and confinement loss of PCF's can be calculated easily.

Effective area $A_{eff}$ can be obtained using the following equation [6]

$$A_{eff}=(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|E|^2dxdy)^2/(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|E|^2dxdy) \quad (1)$$

Where, E is the electric field derived by solving the Maxwell equations.

The confinement loss $L_c$ is obtained from the imaginary part of $n_{eff}$ as follows [2]

$$L_c = 8.686\times k_0\, Im[n_{eff}] \quad (2)$$

Where, $Im[n_{eff}]$ is the imaginary part of the refractive index

$k_0 = 2\pi/\lambda$ is the wave number in the free space.

Chromatic dispersion D can be calculated from the equation [2].

$$D(\lambda)=-\lambda/c(d^2Re[n_{eff}]/d\lambda^2) \quad (3)$$

Where $Re[n_{eff}]$ is the real part of effective refractive index, $n_{eff}$, $\lambda$ is the wavelength, c is the velocity of light in vacuum.

## IV. SIMULATIONN RESULTS AND DISCUSSION

In this proposed Hybrid PCF design, effects on chromatic dispersion and confinement losses are investigated by changing geometrical parameters, such as air hole diameters of inner two rings and pitch. Here we first set air-hole diameters of the inner first hexagonal ring at $d/\Lambda$ = 0.38, inner second hexagonal ring at $d/\Lambda$ = 0.4, outer four circular rings at $d/\Lambda$ = 0.9 where the pitch $\Lambda$ is 1.66 µm. To optimize the dispersion flatness, we tune four structural parameters throughout out our simulation.

Fig. 2 shows the effect of changing $d_1/\Lambda$ on the dispersion behavior for +/- 1%, +/- 2%, +/-5% change from optimum value. For a fixed air-hole pitch $\Lambda$=1.66 µm, second ring $d_2/\Lambda$ = 0.4 and outer ring normalized diameter $d/\Lambda$ = 0.9. It is reported that dimension of the first ring's air holes is particularly important for the overall dispersion-flatness with required accuracy [7].

In Fig. 2 there is an increasing positive dispersion slope as we increase $d_1/\Lambda$, but the slope changes significantly from the positive slope to the negative one near the wavelength 1.35µm.

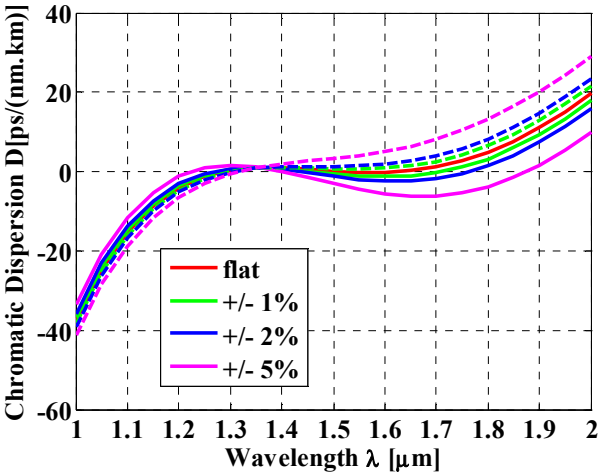



Figure 2. Effect on dispersion due to $d_1/\Lambda$ variation for +/- 1%, +/- 2%, +/-5%.

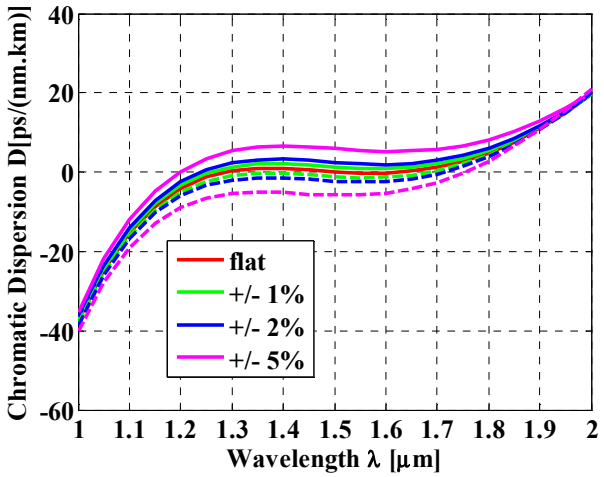



Figure 3. Effect on dispersion due to $d_2/\Lambda$ variation for +/- 1%, +/- 2%, +/-5%.

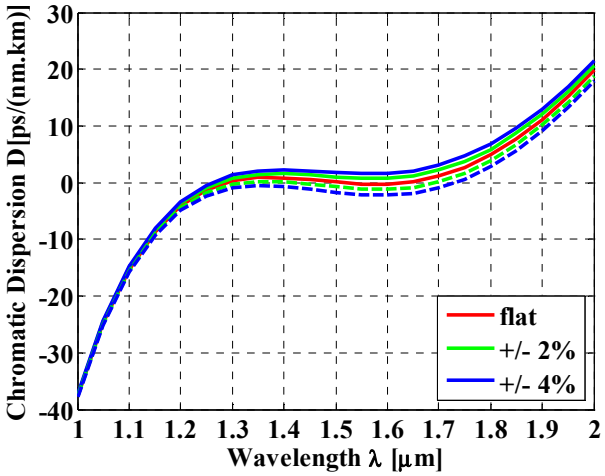



Figure 4. Effect on dispersion due to $d/\Lambda$ variation for +/- 2%, +/- 4%.

Again, we see decreasing positive dispersion slope as we decrease $d_1/\Lambda$, but the slope changes significantly from the negative slope to the positive near the wavelength 1.35µm.

Fig. 3 shows the effect of changing $d_2/\Lambda$ on the dispersion behavior for +/- 1%, +/- 2%, +/-5% change from optimum value. For a fixed air-hole pitch $\Lambda$=1.66µm, first ring $d_1/\Lambda$ = 0.38 and outer ring air filling fraction $d/\Lambda$ = 0.9. It has been seen that there is very less effect on the dispersion slope due to change the air-filling fraction $d_2/\Lambda$.

Fig. 4 shows the effect of changing $d/\Lambda$ on the dispersion behavior for +/- 2%, +/- 4% change from optimum value. For a fixed air-hole pitch $\Lambda$=1.66µm, first ring $d_1/\Lambda$ = 0.38 and second ring $d_2/\Lambda$ = 0.4. It has been seen that there is insignificant effect on the dispersion slope due to change the air-filling fraction $d/\Lambda$.

It is known that in a standard fiber draw, ±1% variations in fiber global diameter may occur [8] during the fabrication process. From Fig. 5 we observe that pitch dominantly influence the dispersion level but little impact on the dispersion slope. With increasing pitch from 1.6 µm to 1.9 µm, the level of dispersion increases significantly and shifts the dispersion to the right band.

Fig. 6 shows the effective area for optimum parameters $d_1/\Lambda$= 0.38, $d_2/\Lambda$ = 0.4, $\Lambda$= 1.66 µm, $d/\Lambda$ = 0.9. The effective area of the fiber at 1.55µm is 7.08E-12 µm$^2$. Simulation result reveals that effective area increases with the increase of wavelength due to optical field spread towards the cladding with wavelength.

Fig. 7 is showing that there is a very little effect on the variation of the effective area due to the changing of $d/\Lambda$. According to simulation it is seen that the effect area increases slightly for decreasing the value of $d/\Lambda$

Fig. 8 shows the change in confinement loss Lc due to the variation in the wavelength. This graph shows that the confinement loss is increasing smoothly with the wavelength.

Fig. 9 shows an optimum dispersion curve of the proposed PCF is in which ultra flattened dispersion of 0 ± 1.20 ps(nm.km) is obtained in a 1250 to 1700 nm wavelength range for $d_1/\Lambda$ = 0.38, $d_2/\Lambda$ =0.40, $d/\Lambda$ = 0.90 and $\Lambda$ = 1.66µm.

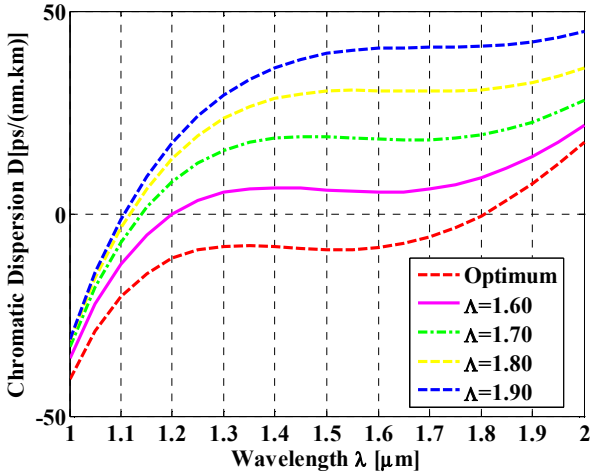


Figure 5. Effect on dispersion due to pitch, Λ variation.

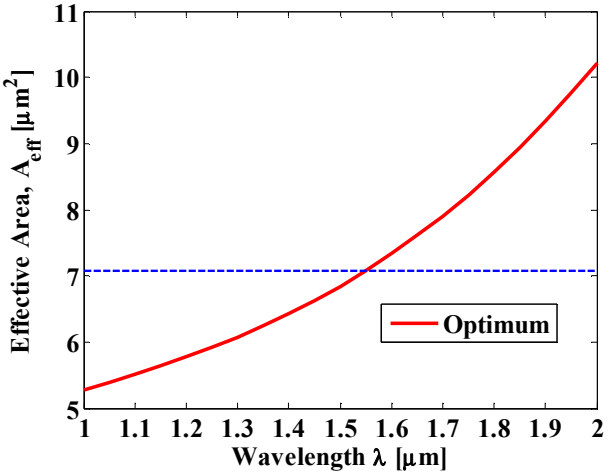


Figure 6. Effective area for optimum parameters.

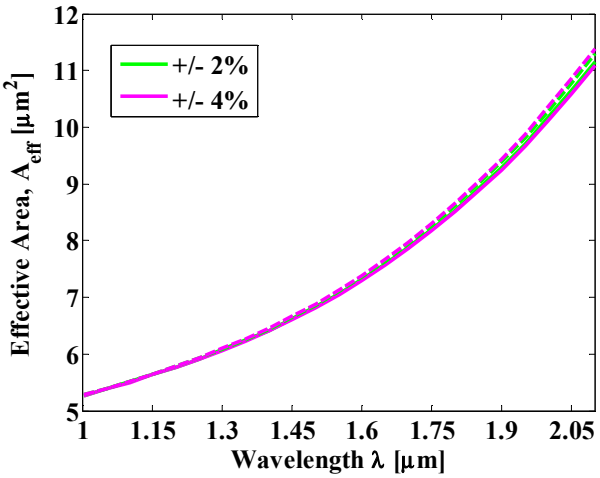


Figure 7. Effect on effective area due to $d/\Lambda$ variation for +/- 2%, +/-4%.

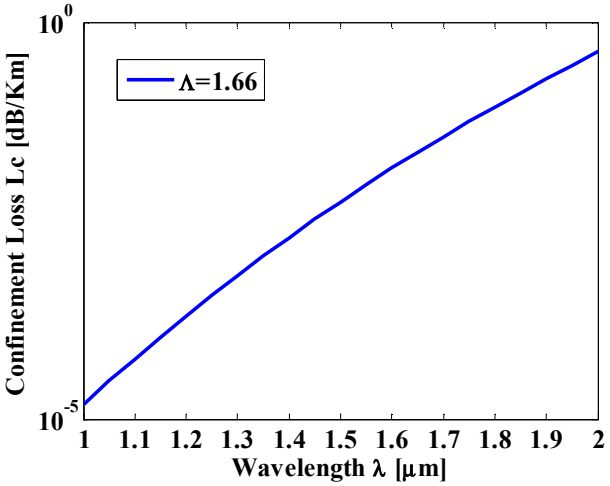


Figure 8. Wavelength dependence of the PCF's confinement loss for optimum parameters.

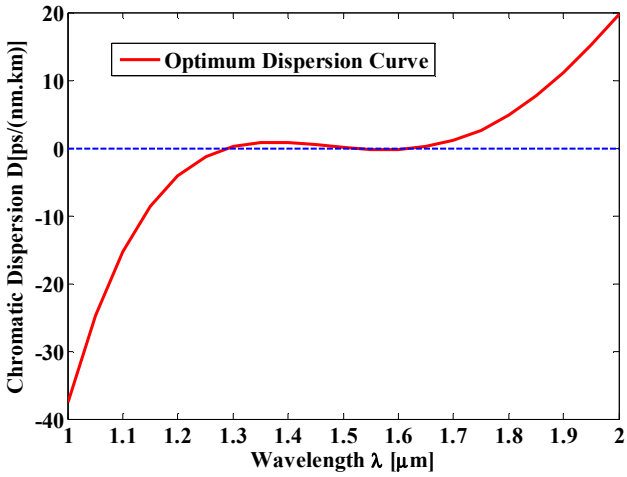



Figure 9. Chromatic dispersion for optimum parameters: $d_1/\Lambda$ = 0.38, $d_2/\Lambda$=0.40, $d/\Lambda$ = 0.90 and $\Lambda$ = 1.66μm.

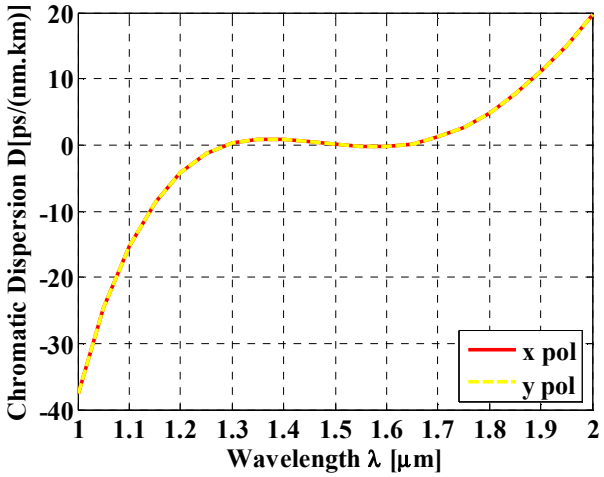



Figure 10. Chromatic dispersion for x and y polarization due to optimum parameters.

TABLE I. COMPARISON OF MODAL PROPERTIES BETWEEN PROPOSED DESIGN AND OTHER DESIGNS

| PCF | Comparison of modal properties | | | |
|---|---|---|---|---|
| | ***ΔD (ps/nm/km)*** | ***Flat band (nm)*** | ***Lc (dB/km)*** | ***NDP ($N_r$, $N_\Lambda$, $N_d$)*** |
| Ref. [1] | 0 ± 0.26 | 200 | $\leq 10^{-6}$ | 5,1,2 |
| Ref. [4] | 0 + 4.8 | 350 | 0.1 | 8,2,2 |
| Ref. [5] | 0 ± 0.5 | 200 | 0.06 | --- |
| Ref. [9] | 0 ± 0.69 | 390 | --- | --- |
| Proposed Design | 0 ± 1.2 | 450 | $\leq 10^{-2}$ | 6,1,3 |

Now we investigate the dispersion behavior for both polarizations. From Fig. 10 it is seen that dispersion for both polarization is almost same due to structural symmetry.

Finally, a comparison is made between properties of the dispersion flat PCF and some other fibers designed for telecom and nonlinear optics applications is shown in the table I taking flat dispersion magnitude and range, number of rings ($N_r$), number of pitch ($N_\Lambda$), number of parameters ($N_d$) in the cladding as comparing parameters.

## V. CONCLUSION

In conclusion, we have designed a hybrid cladding PCF that ensures near zero flattened dispersion of 0 ± 1.2 ps/(nm.km) in a 450 nm flat band and low confinement loss in a wide wavelength range from 1.25 to 1.70 μm. The dispersion can be reduced by further investigations. This fiber has a modest number of design parameters, and also the design is easy to fabricate as there are only circular air holes used. Due to its noteworthy guiding properties, the designed PCF may be excellent candidate for telecom applications.

## REFERENCES


[1] J. C. Knight, T. A. Birks, P. S. J. Russell, and D.M. Atkin, "All-silica single-mode optical fiber with photonic crystal cladding," *Opt. Lett.* 21, 1547-1549 1996.

[2] J. Broeng, D. Mogilevstev, S. E. Barkou, and A. Bjarklev, "Photoniccrystal fibers: A new class of optical waveguides," *Opt. Fiber Technol.*, vol. 5, pp. 305–330, 1999.

[3] K. Saitoh , M. Koshiba , T. Hasegawa and E. Sasaoka, "Chromatic dispersion control in photonic crystal fibers: application to ultra-flattened dispersion,"*Opt. Exp.*, 11 (2003) 843-852

[4] S. M. A. Razzak, Y. Namihira, F. Begum, S. Kaijage, T. Kinjo, J.Nakahodo, K. Miyagi, and N Zou, "Decagonal photonic crystal fibers with ultra-flattened chromatic dispersion and low confinement loss," *Optical Fiber Communication Conference/ National Fiber Optic Engineers Conference (OFC/NFOEC)* JTha91 Anaheim, USA,pp. 1-6, Mar. 25-29: 2007.

[5] T. Matsui, J. Zhou, K. Nakajima, and I. Sankawa, "Dispersion flattened photonic crystal fiber with large effective area and low confinement loss," *J. Lightwave Technol.,* vol. 23, no. 12, pp. 4178-4183, Dec 2005.

[6] S. M. A. Razzak, Y. Namihira, F. Begum, S. Kaijage, N. H. Hai, and N. Zou, "Design of a decagonal photonic crystal fiber with ultra-flattened chromatic dispersion," *IEICE Trans. Electron*, vol. E90-C, no. 11, pp. 2141-2145, Nov. 2007.

[7] F. Poletti, V. Finazzi, T. M. Monro, N. G. R. Broderick, V. Tse, and D. J. Richardson, "Inverse design and fabrication tolerances of ultra-flattened dispersion holey fibers," *Opt. Express*, vol. 13, no. 10, pp. 3728–3736, May 2005.

[8] W. H. Reeves, J. C. Knight, P. S. J. Russell and P. Roberts, "Demonstration of ultra-flattened dispersion in photonic crystal fibers," *Opt. Exp.*, vol. 10, no. 14, pp. 609–613, July 2002.

[9] S. M. A. Razzak, and Y. Namihira, "Proposal for highly non-linear dispersion flattened octagonal photonic crystal fibers," *IEEE Photon. Technol. Lett.*, vol. 20, no.4, pp. 249-251, Feb. 2008.

# Protection of WDM PON Systems based on Modified 2-OLT Architecture

Fahmida Rawshan and Youngil Park
Department of Electronics Engineering, Kookmin University, Seoul, South Korea
E-mail: fahmida_rawshan@yahoo.co.in, ypark@kookmin.ac.kr

*Abstract*—**In this paper, we propose a novel protection architecture of passive optical network (PON) system. Two optical line terminals (OLTs) are connected to each other as well as to the access area. This network can sustain services to access area and restores normal operation in faulty conditions. Subscribers are divided into two groups to reflect priority based bandwidth allocation. WDM access increases the efficiency of the system by operating both OLTs using different wavelengths in normal condition. During fault, a Shared-bandwidth allocation scheme is employed in operating the survived OLT. Performance analysis shows the improvement of the proposed scheme over an existing scheme.**

***Keywords—Passive optical network, multi-OLT, WDM, Self-similar traffic.***

## I. INTRODUCTION

A multi-OLT PON is a good solution to meet recent demands of telecommunication industry as it provides large capacity, user freedom and protection facilities [1-2]. Protection is very important issue since the business and home services demand intact access. Protection mechanism via dual-homed PON is already proposed by some research groups [3]. This operation in revertive mode keeps a standby OLT idle in normal condition. To improve the bandwidth utilization by using all OLTs in normal operation, we propose a WDM based architecture of 2-OLT PON system in our previous study [4]. In this paper, we propose a modified architecture to improve the performance further. In normal status, two OLTs transmit data to different ONU groups using different wavelengths. In failure status, the survived OLT transmits in two wavelengths in turn, thereby replacing the failed OLT. Direct connectivity between two OLTs is the improved feature of this modified structure which decreases protection time.

A modified version of control protocol is suggested to operate the proposed 2-OLT PON system. Users are divided into two groups to distinguish home and enterprise services. Besides, we suggest a Shared-bandwidth DBA (SBA) scheme for downstream transmission to further improve the performance of the 2-OLT PON in faulty situation [4].

The rest of this paper is organized as follows. Section II introduces system architecture and operation algorithm of the WDM based 2-OLT PON system. DBA schemes for both normal operation and faulty condition are also described here. Section III provides simulation results to demonstrate network performance using the suggested DBA schemes. Finally, Section IV concludes our work.

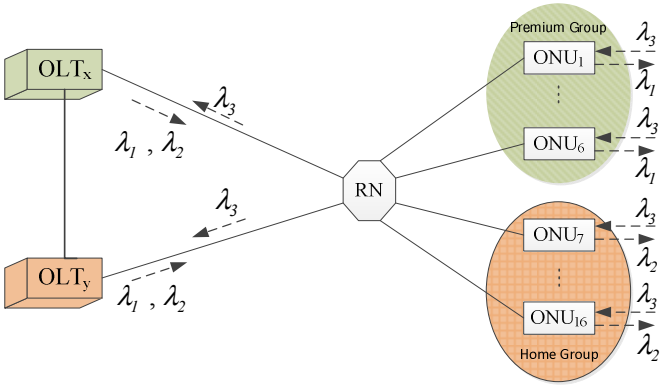



Figure 1. Proposed network architecture.

## II. OPERATION OF THE MODIFIED TWO-OLT WDM NETWORK

In the proposed two-OLT PON, two OLTs are connected to the access area via a passive optical distribution network. The ONUs are divided into two groups 'Premium' and 'Home'. Premium group has lower number of ONUs and receives higher bandwidth facilities than the other group. Each OLT works as a backup OLT for the other one. In normal situation, each OLT is responsible for data transmission of its corresponding ONU group. In the modified structure, OLTs are connected to the other ONU group and works for it during failure condition. Figure 1 shows the architecture of the proposed network for two OLTs and sixteen ONUs. $OLT_x$ manages data transmission to Premium Group ($ONU_1$ to $ONU_6$) in wavelength $\lambda_1$ , while $OLT_y$ manages it for ONUs in Home Group ($ONU_7$ to $ONU_{16}$) in wavelength $\lambda_2$. As two OLTs are connected directly, the exchange of control packets to measure RTT for ranging between $OLT_x$ and Home ONU Group, and between $OLT_y$ and Premium Group is unnecessary. Each OLT measures RTT of the corresponding group and shares it with the other OLT which results in short protection time.

During failure of one OLT the other OLT is capable of transmitting data to both groups. Each OLT can transmit in one of two wavelengths, $\lambda_1$ or $\lambda_2$ at a time while receives in $\lambda_3$.

ONUs in Premium and Home Group receive in $\lambda_1$ and $\lambda_2$, respectively, while any of them transmit in $\lambda_3$.

During the protection of PON systems, however, the efficiency of bandwidth utilization should be maximized since the two ONU groups share the network; the Fixed Bandwidth Allocation (FBA) is not a good scheme in this case as it wastes huge bandwidth in bursty traffic environment. Therefore, a new DBA scheme is suggested for downstream transmission. It is a Shared Bandwidth Allocation which reserves a guaranteed downstream bandwidth toward each group, but allows transfer of unused bandwidth between two groups [4]. The transfer is made when one group encounters slot overflow while the other owns underutilized slots. For example, bandwidth provision from Premium to Home Group is decided by the OLT when this Group has adequate excess bandwidth ($L_1<R_1$ and $G_1<R_1$) and Home group needs additional bandwidth ($L_2>R_2$). Released bandwidth from the 1st group, $BW_{r,1}$ , is decided by eq. (1) depending on the total network load, $L$ [4].

$$BW_{r,1} = \begin{cases} BW_1 \times (1 - G_1 / R_1) \times y_1 & \text{for } L \le 1 \\ BW_1 \times (1 - L_1 / R_1) \times y_1 & \text{for } L > 1 \end{cases} \quad (1)$$

, where $L$ is the normalized total downstream load, $L_{1(2)}$ the load towards $\text{Group}_{1(2)}$, $G_{1(2)}$ the load generation ratio towards $\text{Group}_{1(2)}$, $BW_{1(2)}$ the reserved bandwidth towards $\text{Group}_{1(2)}$, $R_{1(2)}$ the bandwidth reservation ratio for $\text{Group}_{1(2)}$ which is made by service level agreement (SLA), and $y_1$ is the restriction factor.

## III. Performance Analysis

The performance of the two-OLT PON during both normal and protection status are investigated for a system with sixteen ONUs that are divided into two groups. The Premium and the Home group contain 6 ONUs and 10 ONUs, respectively. Each ONU has 1 MB queue, while each OLT has two separate 10 MB queues toward two groups of ONUs. Packets from queues are sent in Ethernet frames. The queues are simple first-in first-out (FIFO) type. Each OLT to corresponding ONU group enjoys 1Gbps in normal condition which makes network capacity of 2Gbps. During protection the single survived OLT carries data transmission toward two groups in 1 Gbps capacity. Performance of normal and protection is presented by comparing FBA and SBA schemes. The cycle time is considered to be 2msec and the guard time be 1 µs.

Matlab® is used for the simulation where data packets are generated by a self-similar traffic model [5-6]. The resulting traffic is similar to the real IP data traffic reflecting recent applications such as P2P, instant messenger, video and VoIP [7].

To reflect the SLA during protection, we reserved 50% of total downstream bandwidth towards each group, Premium and Home, respectively. Average packet delay and throughput are analyzed and plotted with respect to total downstream load. It is seen that FBA shows small delay when the load generation ratio matches the bandwidth reservation ratio. The FBA delay, in the other hand, rises abruptly as generation ratio deviates more from the bandwidth reservation ratio. However, the delay of Shared scheme is not much affected by load variation since bandwidth is shared in efficient way by the two OLTs. Delay of each group is illustrated separately in Fig. 2 and Fig. 3 for FBA and SBA scheme, respectively. In this figures, 'Normal' indicates average delay of two OLTs in normal condition. P(SBA, A%) indicates average delay toward Premium using Shared scheme when A% of total downstream load are transmitted to this Group. Delay becomes very critical for FBA scheme when the reserved bandwidth is less than the real traffic load; this problem is eased with Shared scheme as it controls bandwidth according to the real load. The sharp delay slope of P(SBA,70%) is found at load 0.9, much higher than with FBA scheme, while H(SBA,30%) shows some more delay than FBA, which is negligible. It indicates that efficient transfer of unused bandwidth is made from one group to the other.

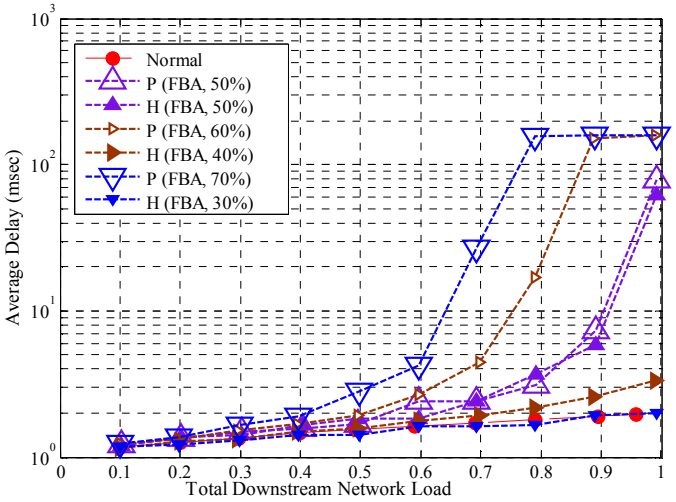



Figure 2. Comparison of downstream delay toward individual group with FBA scheme at bandwidth reservation ratio of 0.5:0.5.

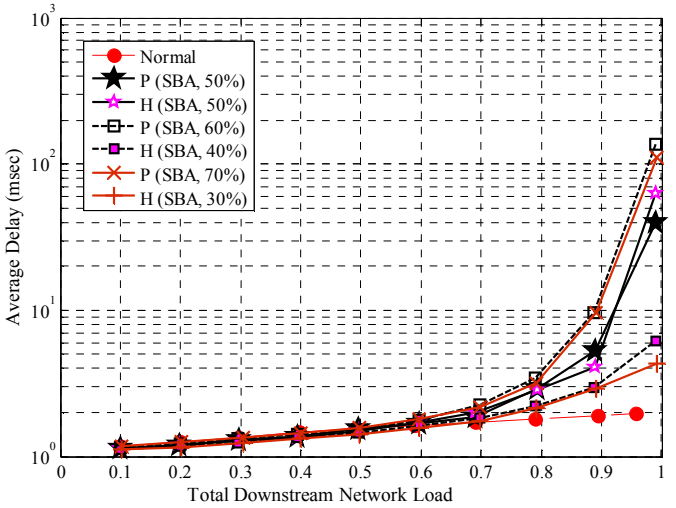



Figure 3. Comparison of downstream delay toward individual group with SBA scheme at bandwidth reservation ratio of 0.5:0.5.

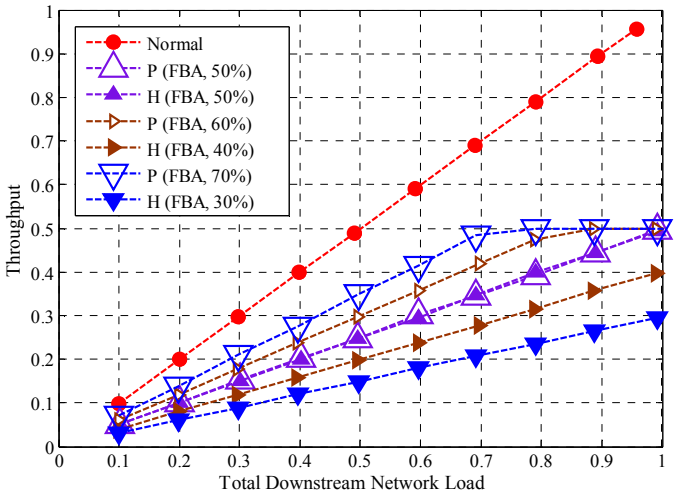



Figure 4. Comparison of downstream throughput toward individual group with FBA scheme at bandwidth reservation ratio of 0.5:0.5.

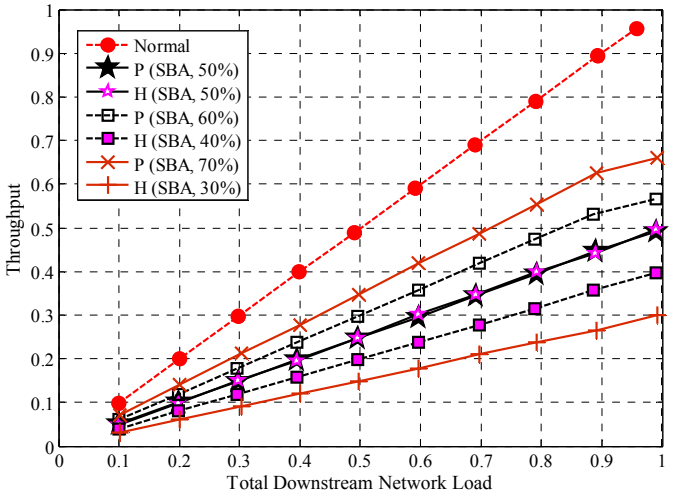



**Figure 5.** Comparison of downstream throughput toward individual group with SBA scheme at bandwidth reservation ratio of 0.5:0.5.

The Shared scheme can increase the bandwidth utilization efficiently. It provides more than 10% of throughputs in unexpected traffic distribution like 70%:30% toward P:H compared to FBA. Throughput toward each group of ONUs is shown in Fig. 4 and 5 for FBA and Shared schemes, respectively. It is observed that throughputs of both P(SBA,70%) and H (SBA,30%) are increasing until load of 1.0; after then throughput of Premium Group decreases since Home Group restores its own bandwidth from it to guarantee its 50% share.

## IV. CONCLUSION

In this paper, we proposed a 2-OLT PON system based on WDM. Both OLTs are used in the data transmission in normal condition, while one of them works for the other OLT during fault condition. Therefore, no OLT is in idle status normally and this increases the bandwidth efficiency compared to other protection scheme such as dual homing. The direct connectivity between the OLTs makes easy information exchange and quick protection. The ONUs are divided into two groups to get different service levels, i.e., home and enterprise services. An appropriate bandwidth allocation algorithm as well as a wavelength assignment scheme is suggested. The Shared-bandwidth scheme is used to improve the bandwidth efficiency during protection and its performance is analyzed by simulation. It is expected that the proposed protection scheme will enhance the survivability of the next generation high-speed PON systems in business enterprises.

## ACKNOWLEDGEMENT

This work was supported by Basic Science Research Program through NRF funded by the MEST, Korea (No.2011-0007107).

## REFERENCES

[1] F. Rawshan and Y. Park, "Architecture of Multi-OLT PON Systems and Its Bandwidth Allocation Algorithms," *Photonic Network Communications,* vol. 25, no. 2. pp. 95-104. April 2013.

[2] F. Rawshan and Y. Park, "A Study on Downstream Dynamic Bandwidth Allocation Algorithm in Multi-OLT PON System," *The Journal of Korea Information and Communications Society (KICS)*, vol.36, no.10, pp. 1143-1149, October 2011.

[3] J. Kang, M. Wilkinson, K. Smith, and D. Nesset, "Restoration of Ethernet Services over a Dual-Homed GPON System – Operator Requirements and Practical Demonstration," in *OFC/NFOEC, Optical Fiber Communication/National Fiber Optic Engineers Conference*, San Diego, CA, February 2008.

[4] F. Rawshan and Y. Park, "Protection/Restoration of PON Systems Using WDM based 2-OLT Structure," *The Journal of Korea Information and Communications Society (KICS)*, vol.37, no. 12, pp. 1168-1173, December 2012.

[5] G. Kramer, *Ethernet Passive Optical Networks*. New York: McGraw-Hill, 2005, pp. 59-63.

[6] W. Leland, M. Taqqu , W. Willinger , and D. Wilson, "On the Self-Similar Nature of Ethernet Traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp.1-15, February 1994.

[7] G. Xie, G. Zhang, J. Yang, Y. Min, V. Issarny, and A. Conte, "Survey on Traffic of Metro Area Network with Measurement On-Line," in *Proc. ITC, 20th international teletraffic conference on Managing traffic performance in converged networks,* Ottawa, June 2007, pp. 666-677.