ICCIE

26-27 November, 2015

1ST INTERNATIONAL CONFERENCE ON COMPUTER & INFORMATION ENGINEERING

ICCIE 2015

Organizer
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh

RUET

1ST INTERNATIONAL CONFERENCE ON COMPUTER & INFORMATION ENGINEERING

ICCIE 2015

26-27 November, 2015

RUET

Organizer

Department of Computer Science & Engineering

Rajshahi University of Engineering & Technology

Rajshahi-6204, Bangladesh

The proceeding mainly contain the abstracts of the accepted and presented papers of ICCIE 2015 which are the original contributions of renown researchers around the country and abroad in the field of computer science and information technology. All the papers have gone through a peer review process by the strong technical program committee to assess the suitability for inclusion in this conference. The editor, organizers and sponsors do not take any liabilities of the opinion or idea reflected from the papers.

Editor
Dr. Md. Al Mamun
Joint Technical Secretary, ICCIE 2015
Associate Professor
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology

Publisher
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204
Bangladesh.

MESSAGE

It is a great pleasure for me to know that the department of Computer Science & Engineering (CSE) of Rajshahi University of Engineering & Technology (RUET) is going to arrange an International Conference on Computer & Information Engineering (ICCIE 2015) during 26-27 November 2015.

As Far as I know, an International Conference of such type will certainly foster innovation at the International level. With the help of ICT, human being has entered into a new type of civilization, widely known as digital age. I strongly believe that this prestigious International Conference will become an important congregation for ICT researchers and professionals from all over the world. People from industries along with the academia will get an opportunity to come across the innovations in engineering, science and technology.

I am confident that the Conference will help provide new and creative ideas that will be helpful to institutionalize the dream of digital Bangladesh, I would like to extend my heariest congratulation to the Computer Sicence & Engineering Department of RUET, Organizers and to everyone involved in this event.

I wish the International Conference a success.

Abdul Mannan 08/11/15

Professor Abdul Mannan
Chairman
University Grant Commission of Bangladesh

MESSAGE

Welcome everyone to International Conference on Computer & Information Engineering (ICCIE 2015) proudly organized by Computer Science and Engineering (CSE) department, Rajshahi University of Engineering & Technology (RUET). Aiming at becoming a part of global knowledge and research society, the Rajshahi University of Engineering & Technology (RUET) is now entering into new era with its mandate of providing world class quality education, dynamic and outstanding faculty as a Centre of Excellence in academics.

Rajshahi University of Engineering & Technology (RUET) is dedicated to build quality graduates in line with the national and global needs for the development through equipping them with updated and broad range of research, technical knowledge, competency and relevant industry skills. RUET always contributes in national and global through excellence in scientific, technical, research and management. And with this view, CSE department of RUET always support the co-curricular activities to empower the knowledge with the practices of ICT in the form of conferences, work-shops, research and related events. The department is always committed to utilize the best technical resources for the students, educators, researchers, consultants, training managers, policy makers, curriculum developers, and entrepreneurs in the development of ICT to realize the digital Bangladesh.

I understand that this prestigious International Conference will bring together many ICT researchers and professional from all over the world with their ideas and perceptions. The event will augment and strengthen the researches and knowledge base significantly by presenting innovative and original thinking, I believe.

My sincere appreciation goes to the advisory committee, technical commit-tee and other members of the program for their encouragement and support for the conference. I should also thankful to all the participants for their contributions in making the conference successful. I hope that this confer-ence will promote collaborations among researchers in the field of Com-puter and Information Technology in Bangladesh and abroad.

I wish a grand success of ICCIE 2015

Prof. Dr. Mohd. Rafiqul Alam Beg,
Vice Chancellor
Rajshahi University of Engineering & Technology (RUET)
Rajshahi

MESSAGE

Bangladesh is making rapid advancement in the use and application of information and communication technologies. Since the declaration of Vision 2021: Digital Bangladesh seven years back, numerous ICT initiatives taken by the Government proved to be highly effective in materializing this technology-led development. However, in order to make this development sustainable, we need to nurture both our intellectual as well as non-intellectual capacity in the area of ICT.

It is a great pleasure to see that the Department Computer Science & Engineering (CSE) of Rajshahi University of Engineering and Technology (RUET) is hosting an event called International Conference on Computer & Information Engineering (ICCIE 2015) in November 2015 to provide a platform for intellectual exercises in the area of ICT. I believe that this sort of conferences not only provide individuals with opportunity of creating and sharing new ideas and knowledge, but also develop national capacity. It is even more important for a nation which is striving to become digital by 2021.

As the Executive Director of BCC, I take immense pride to be associated with ICCIE 2015. I recall BCC's contribution to the development of ICT in Bangladesh including promoting knowledge sharing platforms like ICCIE 2015. Hope this conference will bring ICT researchers and experts from home and abroad in one platform and foster greater cooperation in the days ahead.

My best wishes for success of ICCIE 2015.

S. M. Ashraful Islam
Executive Director
Bangladesh Computer Council

MESSAGE

It is my pleasure to inform you that IEEE Bangladesh Section (IEEE-BDS) provided Technical Co-sponsorship for 2015 International Conference on Computer and Information Engineering (ICCIE 2015) to be organized by the Department of CSE of Rajshahi University of Engineering and Technology (RUET). As a Technical Co-sponsor, IEEE-BDS provided different essential guidelines and support to enhance the quality of the conference.

IEEE is world's largest professional association and dedicated to advancing technological innovation and excellence for the benefits of humanity. From its inception in 1993, IEEE-BDS is working to promote engineering education and advanced research in Bangladesh. Currently it has around 1500 members, chapters of four technical societies (communication society (COMSOC), power and energy society (PES), electron device/solid state circuit society (EDS/SSCS)), Engineering in Medicine and Biology Society (EMBS), two affinity groups (women in engineering (WIE) and young professional (YP)), two student chapters (Industrial Application Society (IAS) and Microwave Theory and Techniques (MTT) society), two SIGHT groups (FLASH and CARG) and 21 student branches from 21 engineering universities. Petition for opening of IEEE Computer Society is at the final stage. IEEE-BDS and its societies, affinity groups, chapters, and student branches regularly organize technical seminars, workshops, congress, contests, Olympiads, short courses, industrial tours, project competitions, outreach programs, humanitarian activities, robotic challenges, award programs, industrial, educational and professional activities. Because of outstanding performance, our WIE Chair has own prestigious awards from region 10 and international WIE. This year some of our dedicated volunteers have been selected in different R10 committees, which is a great honor for us. IEEE BDS has won project funds from R10 in six different categories through very tough competitions. This year, volunteers from IEEE BDS have won prizes in IEEE R10 SYW Congress held in Sri Lanka. My best wishes to all award winners and project grant winners. Last year total number of members was about 1000 and in this year, it is now more than 1500. This year IEEE-BDS is awarded “2015 Outstanding Section Membership Recruitment Performance”. I hope participants in this conference will join in IEEE if not yet joined and actively participate in our upcoming events.

This year apart from ICCIE 2015, several International Conferences, namely ICCIT, ICEEICT, ICAEE, EICT, NSYSS, and ICEEE are technically co-sponsored by IEEE-BDS and IEEE ICTP is sponsored by COMSOC BD Chapter. For the first time in history, IEEE-BDS is going to organize WIE technical conference IEEE WIECON-ECE 2015 in collaboration with WIE Affinity Group of IEEE-BDS. For the first time, IEEE-BDS has launched IEEE ProTalks 2015, a workshop focusing technical leadership, management and entrepreneurship. This year IEEE-BDS has successfully organized the IEEE Region 10 Meet 2015 where 110 foreign delegates from 17 countries including USA, Australia, Japan, India, and China participated. Since its inception, this is the first time IEEE BDS has got this opportunity.

The conference organizers were requested to follow the guidelines provided by IEEE and IEEE-BDS. In this regard, I appreciate their willingness and effort. I believe, because of hard work of all the members of the tracks and concerned reviewers, including checking of plagiarism and conflict of interest, the quality of the review process is enhanced. It is expected from the organizer that, being an International Conference, a significant number of authors from abroad will attend the conference for presenting their papers. I express my sincere gratitude to all the authors, speakers, track committees, reviewers, advisers and other members whose sincere efforts are the key factors for the success of this conference.

I appreciate feedback and suggestions from all the participants regarding the technical issues of this conference, which will help us to achieve more success in enhancing the quality of the upcoming conferences. I wish ICCIE 2015 all the success.

Dr. Shaikh Anowarul Fattah
Chair, IEEE Bangladesh Section
Professor, Dept. of EEE, BUET

MESSAGE

I am delighted to know that the department of Computer Science and Engineering (CSE), Rajshahi University of Engineering & Technology (RUET) is going to organize 1st International Conference on Computer & Information Engineering (ICCIE 2015) in Rajshahi on 26-27 November 2015.

The widespread applications of Computer & Information Engineering can greatly benefit the economy of our country. We need to enhance the technological knowledge base of our society. Substantial advancement in technological arena is needed particularly to give impetus to Digital Bangladesh programme. By virtue of this conference, the knowledge of computer and information engineering shall be disseminated throughout the country. This type of conference provides opportunities to bridge gap in knowledge between national and international participants and to update new knowledge through sharing the outcome of their studies and research findings.

I expect that this ICCIE 2015 conference will produce positive results in terms of sharing scientific, technical and engineering ideas for scientist, engineers, researchers and students from all over the world.

On my behalf and that of ECE faculty, we express our appreciation to the organizers of this conference for their efforts in holding this conference on a regular basis.

I wish ICCIE 2015 Conference a grand success.

Prof. Dr. Md. Rafiqul Islam Sheikh
Dean, Faculty of Electrical & Computer Engineering
Rajshahi University of Engineering & Technology, Bangladesh.

MESSAGE

As Chairman of technical committee, I am delighted to welcome delegates from all over the world to the International Conference on Computer & Information Engineering (ICCIE 2015) to be held at Department of Computer Science & Engineering of Rajshahi University of Engineering & Technology on 26-27 November 2015.

This exciting event is entirely dedicated to recent developments in Computer Science and related technologies. As we are living in the modern era, the Computer and Information Engineering are inseparable from each other and with the demand of services at door-steps today, the application of ICT has become phenomenon. Besides the Digital Bangladesh added new dimension in catering digital services in many sectors in a cost saving manner.

The objective of ICCIE 2015 is not only exchange ideas and share knowledge among the Scientists, Scholars, Engineers, Educators and Technologists who have gathered here for the presentation on different aspects of the computer Science and Engineering along with the Information Technology but also to transform the scholastic work for the usefulness of the mankind through the government, industry and academia.

I would like to congratulate all the participants for presenting their papers from home and abroad. I am thankful to all supporting partners, authors, reviewers and different functional committee of ICCIE 2015.

I wish all the success for the event.

Prof. Dr. Md. Shahid Uz Zaman
Chairman
Technical Committee, ICCIE 2015
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology

MESSAGE

Welcome everyone to the 1st International Conference on Computer & Information Engineering (ICCIE 2015) to be held at department of Computer Science & Engineering of Rajshahi University of Engineering & Technology.

I am very much delighted to be here at CSE department, RUET on the occasion of ICCIE 2015. I want to thank all of you from the deepest part of my heart.

Information and Computer Engineering has been evolved enormously and over the last decade countries all over the world has achieved a notable progress in Information and Computing skills. Bangladesh is no exception and has made a tremendous progress in the sector of Information and Computer Engineering.

I believe this kind of Conference creates a great opportunity for the researchers to enrich their skills, experience and also share their research and experience with others. I hope all the researchers will try to show their best innovations and idea in this platform. It is surely a glorious chance for all of you to catch this chance and make it a reality. However I am very immense to attend here. I warmly thank everyone of CSE Department, RUET.

I wish all the success for the event.

Dr. Boshir Ahmed
Secretary, Technical Committee, ICCIE 2015
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology.

MESSAGE

It is a great pleasure for me to welcome you on behalf of the organizing committee, to the 1st International Conference on Computer and Information Engineering (ICCIE 2015) to be held at the Department of Computer Science & Engineering at Rajshahi University of Engineering & Technology. It is indeed a great honor for Rajshahi University of Engineering & Technology to host the ICCIE 2015 at the Department of Computer Science & Engineering during 26-27 November, 2015.

Computer and Information Engineering plays an enormous role in promoting knowledge and technology which is essential for the educators, researchers, industrial and commercial concerns in the present digital age. Being a core part of this conference from the beginning, I myself feel very much enthusiastic about the conference and I hope that we all will get benefit academically through mutual collaborations.

As a secretary of the organizing committee , I appreciate the entire functional arrangement of ICCIE 2015 through the devoted faculty members along with different sub-committees. It is not possible to arrange 1st ICCIE 2015 without the valuable contributions from the supporting partners such as Bangladesh Computer Council (BCC), CISCO Networking Academy RUET and IEEE Bangladesh section.

I wish all delegates a very fruitful and enjoyable moments in ICCIE 2015.

Professor Dr. Md. Rabiul Islam
Secretary
Organizing Committee, ICCIE 2015
Rajshahi University of Engineering & Technology, Bangladesh

MESSAGE

I extend my most sincere welcome to all attendees of this year's 2015 International Conference on Computer & Information Engineering (ICCIE 2015) in the Department of Computer Science & Engineering of Rajshahi University of Engineering & Technology.

I hope that the 1st International Conference on Computer & Information Engineering (ICCIE 2015) is an attempt to expose the recent advancement and innovation throughout the vast area of computer and information science. It is expected to be an intellectual platform to share ideas and present the latest findings and experiences in the mentioned areas. World's leading creative educators, researchers, consultants, training managers, policy makers, curriculum developers, entrepreneurs, and others in Computer Science education are likely to participate in the spark discussion.This brings a plurality of interests and perspectives to a single location.

I hope you take advantage of this opportunity and contribute, through presentations, discussion and interaction, to the development of new ideas and new directions in research and applied technology.

Finally, my thanks goes out to the paper reviewers and the keynote speakers as well as invited speakers and authors who have helped to make this Conference a successful one. I would like to extend a special thanks to the sponsors, specifically Bangladesh Computer Council (BCC), Dhaka, CISCO Academy, RUET, IEEE Bangladesh Section and other organizing committee members who have taken time out of their busy schedules to help organizing this year's Conference.

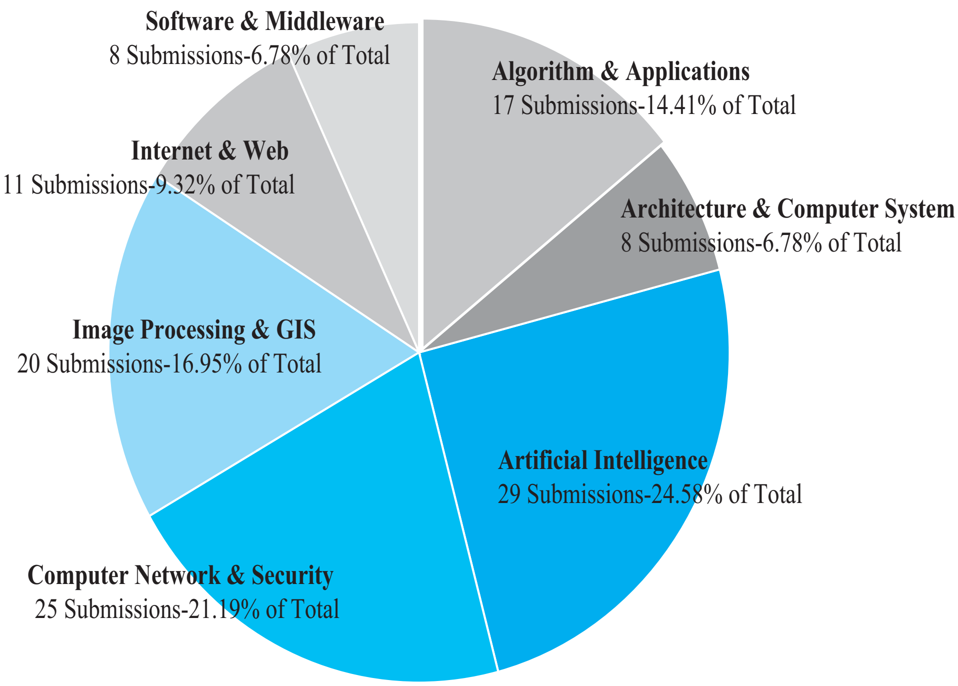
Prof. Dr. Md. Nazrul Islam Mondal
Organizing Chair
ICCIE 2015 and
Head
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology

INDEX PAGE

About ICCIE 2015

Track Wise Submission

Track Name	Submission		Total Number of Paper
Software & Middleware	8 Submissions-	6.78% of Total	118
Internet & Web	11 Submissions-	9.32% of Total	
Image Processing & GIS	20 Submissions-	16.95% of Total	
Computer Network & Security	25 Submissions-	21.19% of Total	
Artificial Intelligence	29 Submissions-	24.58% of Total	
Architecture & Computer System	8 Submissions-	6.78% of Total	
Algorithm & Applications	17 Submissions-	14.41% of Total	



Country-wise Submission

Country	Number of Submission
Bangladesh	110
Malaysia	2
Turkey	1
Germany	1
Iran	1
Australia	2
Japan	1

About ICCIE 2015

Foreign Universities Participated in ICCIE 2015
Islamic Azad University South Tehran Branch, Iran
Saitama University, Japan
Fırat University, Turkey
International Islamic University, Malaysia
Universiti Kebangsaan, Malaysia
University of New South Wales, Australia
University Kassel, Germany

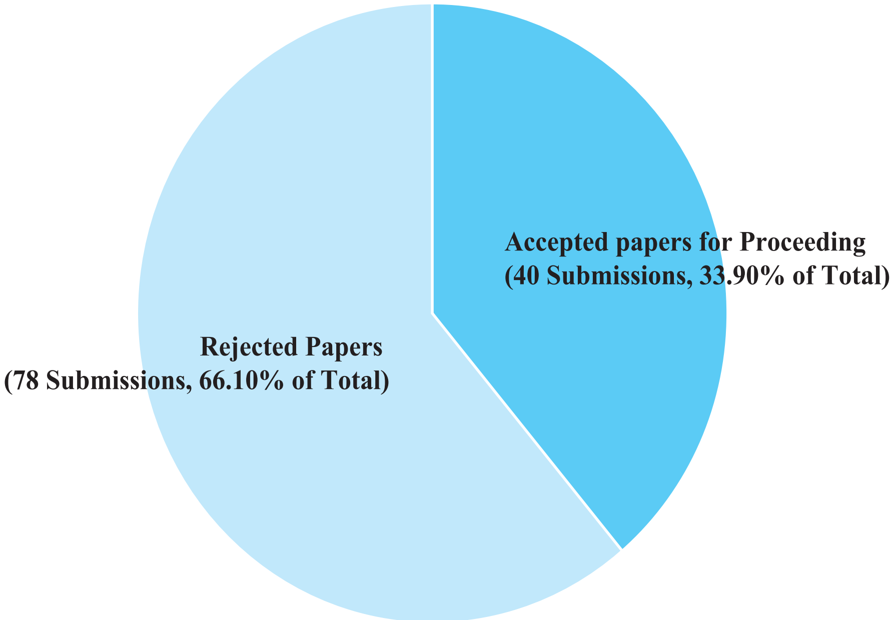
Local Universities/Organizations Participated in ICCIE 2015	
Bangladesh University of Engineering & Technology	International University of Business Agriculture and Technology
Rajshahi University of Engineering & Technology	The University of Liberal Arts
Khulna University of Engineering & Technology	North South University
Chittagong University of Engineering & Technology	Jahangirnagar University
Dhaka University of Engineering & Technology	Khulna University
Shahjalal University of Science and Technology	University of Asia Pacific
Pabna University of Science & Technology	North South University
East West University	Independent University
Primeasia University	Southern University
American International University Bangladesh	Military Institute of Science & Technology
Ahsanullah University of Science & Technology	University of Science and Technology Chittagong
Mawnlana Bhashani Science &Technology University	Northern University
International Islamic University, Chittagong	Brac University
Islamic University	Bangladesh Power Development Board
Grameenphone Ltd.	Institute of Business Administration, University of Dhaka
United International University	Premier University, Chittagong
Uttara University	Bangabandhu Sheikh Mujibur Rahman Maritime University
University of Chittagong	Software Global Consultancy
University of Information Technology & Science	Jatiya Kabi Kazi Nazrul Islam University

About ICCIE 2015

Acceptance and Rejection Details

Total Number of Submission	Number of Paper Accepted for Final Proceeding	Number of Paper Rejected	Acceptance Rate	Rejection Rate
118	40	78	33.90%	66.10%

Acceptance and Rejection Rate



Track-wise Acceptance and Rejection (For Final Proceeding)

Track Name	Total Number of Paper	Number of Paper Accepted	Number of Paper Rejected
Software & Middleware	8	3	5
Internet & Web	11	3	8
Image Processing & GIS	20	10	10
Computer Network & Security	25	11	14
Artificial Intelligence	29	10	19
Architecture & Computer System	8	0	8
Algorithm & Applications	17	3	14

Reviewer List

Name	Organization
Dr. Abhay	Amity University, India
Dr. Md Jahangir Alam	UNSW Canberra, Australia
Dr. Alpana	Amity University, India
Dr. Md. Waselul Haque Sadid	Concordia University, Canada
Prof. Muhammad Channa	Department of Information Technology Quest Nawabshah, Pakistan
Dr. Ruhul A. Sarker	UNSW Canberra, Australia
Dr. Kaushik Roy	North Carolina A &T State University, USA
Dr. Abul Haque	North South University, Bangladesh
Dr. Md. Aminul Haque Akhand	Khulna University of Engineering & Technology, Bangladesh
Prof. Dr. A. K. M. A Hossain	Rajshahi University, Bangladesh
Prof. Dr. A. K. M. Fazlul Haque	Daffodil International University, Bangladesh
Prof. Dr. Syed Akhter Hossain	Daffodil International University, Bangladesh
Dr. Md. Ali Hossain	Rajshahi University of Engineering & Technology, Bangladesh
Dr. A. M. M. Amanat Khan	Dhaka University, Bangladesh
Dr. Asaduzzaman	Chittagong University of Engineering & Technology , Bangladesh
Dr. Md. Ashikur Rahman Khan	Noakhali Science and Technology University, Bangladesh
Dr. G. M. Atiqur Rahman	Khulna University, Bangladesh
Dr. Boshir Ahmed	Rajshahi University of Engineering & Technology, Bangladesh
Dr. Md. Mahfuzur Rahman	Eastern University, Bangladesh
Prof. Dr. Chowdhury Mofizur Rahman	United International University, Bangladesh
Dr. Md. Al Mamun	Rajshahi University of Engineering & Technology, Bangladesh
Dr. Dipankar Das	Rajshahi University, Bangladesh
Prof. Dr. Md. Fayzur Rahman	Daffodil International University, Bangladesh
Prof. Dr. Mohammod Abul Kashem	Dhaka University of Engineering & Technology, Bangladesh
Prof. Dr. Md. Nasim Akhtar	Dhaka University of Engineering & Technology, Bangladesh
Dr. Sheak Rashed Haider Noori	Daffodil International University, Bangladesh
Prof. Dr. Md. Ekramul Hamid	Rajshahi University, Bangladesh
Dr. Md. Selim Hossain	Rajshahi University of Engineering & Technology, Bangladesh
Dr. Md. Faruk Hossain	Rajshahi University of Engineering & Technology, Bangladesh
Group Captain Md. Afzal Hossain	Military Institute of Science and Technology, Bangladesh
Prof. Dr. Md. Khadem Islam Molla	Rajshahi University, Bangladesh
Prof. Dr. Kamrul hasan Talukder	Khulna University, Bangladesh

Reviewer List

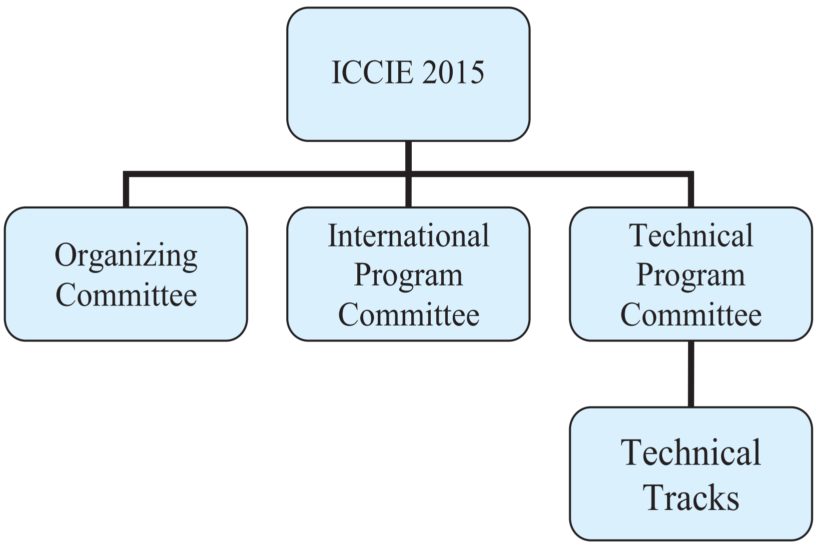
Name	Organization
Dr. Md Mokammel Haque	Chittagong University of Engineering & Technology, Bangladesh
Dr. M. S. Anower	Rajshahi University of Engineering & Technology, Bangladesh
Dr. Md. Obaidur Rahman	Dhaka University of Engineering & Technology, Bangladesh
Dr. Manjur Hasan	Chittagong University of Engineering & Technology, Bangladesh
Prof. Dr. Muhammad Masroor Ali	Bangladesh University of Engineering & Technology, Bangladesh
Prof. Dr. Mohammed Moshiul Hoque	Chittagong University of Engineering & Technology, Bangladesh
Dr. Motiur Rahman	Mawlana Bhashani Science and Technology University, Bangladesh
Dr. Mohammad Shamim Kaiser	Jahangirnagar University, Bangladesh
Prof. Dr. Md. Nazrul Islam Mondal	Rajshahi University of Engineering & Technology, Bangladesh
Dr. Pranab Dhar	Chittagong University of Engineering & Technology, Bangladesh
Prof. Dr. Md. Rabiul Islam	Rajshahi University of Engineering & Technology, Bangladesh
Dr. Rafiqul Islam	Dhaka University of Engineering & Technology, Bangladesh
Dr. Kazi Shah Nawaz Ripon	Khulna University, Bangladesh
Prof.Dr. Md. Rafiqul Islam Sheikh	Rajshahi University of Engineering & Technology, Bangladesh
Prof. Dr. Kazi Md. Rokibul Alam	Khulna University of Engineering & Technology, Bangladesh
Dr. Md. Rokanujjaman	Rajshahi University, Bangladesh
Prof. Dr. Md. Rokunuzzaman	Rajshahi University of Engineering & Technology, Bangladesh
Prof. Dr. Md. Sheikh Sadi	Khulna University of Engineering & Technology, Bangladesh
Dr. Sajjad Waheed	Mawlana Bhashani Science and Technology University, Bangladesh
Dr. Mohammad Arefin	Chittagong University of Engineering & Technology, Bangladesh
Dr. Ajay Krishno Sarkar	Rajshahi University of Engineering & Technology, Bangladesh
Prof. Dr. Shamim Ahmad	Rajshahi University, Bangladesh
Dr. Md. Shohrab Hossain	Bangladesh University of Engineering and Technology, Bangladesh
Dr. Md. Islam	Jahangirnagar University, Bangladesh
Prof. Dr. S. M. Abdur Razzak	Rajshahi University of Engineering & Technology, Bangladesh
Prof. Dr. Mohammad Shah Alam	Bangladesh University of Engineering and Technology, Bangladesh
Prof. Dr. Md. Shahid Uz Zaman	Rajshahi University of Engineering & Technology, Bangladesh
Dr. Touhid	Daffodil International University, Bangladesh
Dr. Tanvir Rahman	Daffodil International University, Bangladesh
Dr. Md. Yousuf	Jahangirnagar University, Bangladesh
Mr. Enamul Haque	University of British Columbia, Canada
Mr. Mahmudul Haque	The Arctic University of Norway, Norway
Mr. Iqbal Sarker	Swinburne University of Technology, Australia

Technical Tracks Members

Technical Track 1 Algorithm & Application	**Chair:** Prof Dr. Ruhul A. Sarker, UNSW, Canberra, Australia **Co-Chair:** Prof. Dr. Shaikh Anowarul Fattah, BUET, Bangladesh **Member:** Prof. Dr. Kamrul hasan Talukder, KU, Bangladesh
Technical Track 2 Architecture & Computer System	**Chair:** Prof. Dr. Engr. M. N. I Mondal, RUET, Bangladesh **Co-Chair:** Prof. Dr. Md. Shamsul Arefin, CUET, Bangladesh
Technical Track 3 Artificial Intelligence	**Chair:** Prof. Dr. Muhammad Masroor Ali, BUET, Bangladesh **Co-Chair:** Prof. Dr. Md. Rabiul Islam, RUET, Bangladesh
Technical Track 4 Software & Middleware	**Chair:** Prof. Dr. Syed Akhter Hossain, DIU, Bangladesh **Co-Chair:** Dr. Mohammad Shamim Kaiser, JU, Bangladesh
Technical Track 5 Computer Networks & Security	**Chair:** Prof. Dr. Md. Nasim Akhter, DUET, Bangladesh **Co-Chair:** Dr. Mohammad Shah Alam, BUET, Bangladesh **Member:** Dr. Md. Al Mamun, RUET, Rajshahi
Technical Track 6 Image Processing & GIS	**Chair:** Prof. Dr. Md. Shahid Uz Zaman, RUET, Bangladesh **Co-Chair:** Dr. Kaushik Roy, North Carolina A &T State University, USA **Member:** Dr. Md. Ali Hossian, RUET, Bangladesh
Technical Track 7 Internet & Web	**Chair:** Prof. Dr. Mohammed Moshiul Hoque, CUET, Bangladesh **Co-Chair:** Dr. Boshir Ahmed, RUET, Bangladesh

Committees-ICCIE 2015

Committee Organization



Chief Patron

Prof. Dr. Mohd. Rafiqul Alam Beg
Vice-Chancellor, RUET

International Program Committee

Prof. Dr. Koji Nakano
Department of Information Engineering, School of Engineering, Hiroshima University, Japan

Prof. Dr. Jacir L. Bordim
Department of Computer Science, University of Brasillia, Brasil

Prof. Dr. Narayanan Kulathuramaiyer
Faculty of Computer Science & Information Technology, University Malaysia Sarawak, Malaysia

Prof. Dr. Noor Alamshah Bolhassan
Faculty of Computer Science & Information Technology, University Malaysia Sarawak, Malaysia

Prof. Dr. Ruhul A. Sarker
School of Engineering and Information Technology (SEIT), UNSW@ADFA, University of New South Wales, Australia

Prof. Dr. Morikazu Nakamura
Dept. of Information Engineering, University of the Ryukyus, Japan

Dr. Shahin Hashtrudi Zad
Department of Electrical and Computer Engineering, Concordia University, Canada

Dr. Guchuan Zhu
Department of Electrical Engineering, Ecole Polytechnique de Montreal, Canada

Dr. Mohd. Hasan Ali
Electrical & Computer Engineering, Memphis University, USA

Dr. M. Ali Akber Dewan
School of Computing and Information Systems, Athabasca University, Canada

Dr. Xiuping Jia
School of Engineering and Information Technology (SEIT), UNSW@ADFA, University of New South Wales, Australia

Dr. Kaushik Roy
Department of Computer Science, North Carolina A&T State University, USA

Dr. Jahangir Hossain
Department of Electrical & Electronic Engineering, Griffith School of Engineering, Australia

Dr. Apel Mahmud
Department of Electrical and Electronic Engineering, Swinburne University, Australia

Dr. Mohammad Zahidul Hasan Bhuiyan
Finnish Geospatial Research Institute (FGI), Finland

Technical Program Committee

Prof. Dr. Md. Shahid Uz Zaman, CSE, RUET, Bangladesh (Chair)
Prof. Dr. M. Kaykobad, CSE, BUET, Bangladesh
Prof. Dr. Muhammad Masroor Ali, CSE, BUET, Bangladesh
Prof. Dr. Pran Kanai Shaha, EEE, BUET
Prof. Dr. Quazi Deen Mohd Khosru, EEE, BUET, Bangladesh
Prof. Dr. Md. Saidur Rahman, CSE, BUET
Prof. Dr. A. B. M. Harun-Ur-Rashid, EEE, BUET, Bangladesh
Prof. Dr. Md. Shah Alam, BUET, Bangladesh
Prof. Dr. Md. Aynal Haque, EEE, BUET, Bangladesh
Prof. Dr. M. M. A. Hashem, CSE, KUET, Bangladesh
Prof. Dr. Kazi Khairul Islam, EEE, IUT, Bangladesh
Prof. Dr. Md. Ashraful Hoque, EEE, IUT, Bangladesh
Prof. Dr. Md. Ruhul Amin, EEE, IUT, Bangladesh
Prof. Dr. Abdul Mottalib, CSE, IUT, Bangladesh
Prof. Dr. Anwarul Abedin, EEE, DUET, Bangladesh
Prof. Dr. Md. Nasim Akthar, CSE, DUET, Bangladesh
Prof. Dr. M. A. Kashem, CSE, DUET, Bangladesh
Prof. Dr. Mohammed Moshiul Hoque, CSE, CUET, Bangladesh
Prof. Dr. Kamrul Hasan Talukder, CSE, KU, Bangladesh
Prof. Dr. Md. Mortuza Ali, EEE, EU, Bangladesh
Prof. Dr. Shamim Ahmad, CSE, RU, Bangladesh
Prof. Dr. Syed Akhter Hossain, CSE, DIU, Bangladesh
Prof. Dr. Md. Rafiqul Islam Sheikh, EEE, RUET
Prof. Dr. S. M. Abdur Razzak, EEE, RUET, Bangladesh
Prof. Dr. Engr. M. N. I Mondal, CSE, RUET, Bangladesh
Prof. Dr. Md. Rabiul Islam, CSE, RUET, Bangladesh
Prof. Dr. Md. Zahurul Islam Sarker, EEE, RUET
Prof. Dr. Shaikh Anowarul Fattah, EEE, BUET, Bangladesh & Chair, IEEE Bangladesh Section
Prof. Dr. Muhammad Abdul Goffar Khan, EEE, RUET & Vice Chair, IEEE Bangladesh Section
Prof. Dr. Satya Prasad Majumder, EEE, BUET, Bangladesh, & Chair, IEEE Communication Society, Bangladesh Chapter
Dr. Boshir Ahmed, CSE, RUET, Bangladesh (Secretary)
Dr. Md. Al Mamun, CSE, RUET, Bangladesh (Joint Secretary)
Dr. Ajay Krisno Sarker, EEE, RUET
Dr. Shamim Anower, EEE, RUET
Dr. Md. Rabiul Islam, EEE, RUET
Dr. Md. Selim Hossain, EEE, RUET
Dr. Md. Masud Rana, EEE, RUET
Dr. Md. Faruk Hossain, EEE, RUET
Dr. Mohammad Shamim Kaiser, CSE, JU, Bangladesh, Student Activities, IEEE Bangladesh Section
All Tachnical Track chairs, Co-chairs and Members

Local Organizing Sub-Committees

Organizing Committee

Prof. Dr. Md. Nazrul Islam Mondal, Head, CSE, RUET	Chair
Prof. Dr. Md. Rabiul Islam, CSE, RUET	Secretary
All teachers of the Department of CSE, RUET	Member

Technical Committee

Prof. Dr. Md. Shahid Uz Zaman, CSE, RUET	Chair
Dr. Boshir Ahmed, CSE, RUET	Secretary
Dr. Md. Al Mamun, CSE, RUET	Join Secretary
Prof. Dr. Md. Nazrul Islam Mondal, Head, CSE, RUET	Member
Prof. Dr. Md. Rabiul Islam, CSE, RUET	Member
Prof. Dr. Md. Rafiqul Islam Sheikh, EEE, RUET	Member
Prof. Dr. S. M. Abdur Razzak, EEE, RUET	Member
Prof. Dr. Zahurul Islam Sarker, EEE, RUET	Member
Prof. Dr. Muhammad Masroor Ali, BUET, Bangladesh	Member
Prof. Dr. Shamim Ahmed, RU, Bangladesh	Member
Prof. Dr. Ruhul Amin, IUT, Bangladesh	Member
Prof. Dr. M. A Hashem, KUET, Bangladesh	Member
Prof. Dr. Md. Anwarul Abedin, EEE, BUET	Member
Prof. Dr. Md. Ibrahim Khan, CSE, CUET	Member

Publication Sub-Committee

Dr. Md. Al Mamun, CSE, RUET	Chair
Rizoan Toufiq, CSE, RUET	Secretary
Julia Rahman, CSE, RUET	Member
Sadia Zaman Mishu, CSE, RUET	Member
Abu Sayeed, CSE, RUET	Member

Registration, Kit Distribution and Finance Sub-Committee

Dr. Md. Ali Hossain, CSE, RUET	Chair
Emrana Kabir Hashi, CSE, RUET	Secretary
Prof. Dr. Md. Rabiul Islam, CSE, RUET	Member

Sponsorship Sub-Committee

Prof. Dr. Md. Shahid Uz Zaman, CSE, RUET	Chair
Dr. Boshir Ahmed, CSE, RUET	Secretary

Invitation Transportation, Accommodation and Foreign Delegates Management Sub-Committee	
Prof. Dr. Md. Nazrul Islam Mondal, CSE, RUET	Chair
Mumu Aktar, CSE, RUET	Secretary

Session Management Sub-Committee	
Dr. Boshir Ahmed, CSE, RUET	Chair
Biprodip Pal, CSE, RUET	Secretary
Barson Sen, CSE, RUET	Member

Reception and Refreshment Sub-Committee	
Prof. Dr. Md. Rabiul Islam, CSE, RUET	Chair
Md. Kalimuzzaman, Senior System Analyst, CSE, RUET	Secretary

Website, Information Management Sub-Committee	
Biprodip Pal, CSE, RUET	Chair
Shihan Arafin, Programmer, CSE, RUET	Secretary

Inauguration and Closing Ceremony Management Sub-Committee	
Prof. Dr. Md. Rabiul Islam, CSE, RUET	Chair
Shyla Afroge, CSE, RUET	Secretary

Security Sub-Committee	
Prof. Dr. Md. Nazrul Islam Mondal, Head, CSE, RUET	Chair
Security Officer, RUET	Secretary

Venue Management Sub-Committee	
Prof. Dr. Md. Nazrul Islam Mondal, CSE, RUET	Chair
Firoz Mahmud, CSE RUET	Secretary
Shyla Afroge, CSE, RUET	Member
Mumu Aktar, CSE, RUET	Member

Best Paper Award

A review of all papers which will receive high review scores will be undertaken by ICCIE 2015 Technical Program Committee to nominate best paper awards. Prof. Dr. Fatema Rashid Best Paper Award will be given to three best papers based on recommendation of the ICCIE Technical Program Committee. Only the papers presented in the ICCIE 2015 will be considered for this awards. Each award contains prize money and a certificate.

- 1st prize : BDT 10,000 and a certificate
- 2nd prize : BDT 6,000 and a certificate
- 3rd prize : BDT 4,000 and a certificate

PROGRAM SCHEDULE

<table>
<tr><th colspan="2">26 November, 2015 (Thursday) : Day 1</th></tr>
<tr><td>2.00 PM
To
3.00 PM</td><td>Registration and Kit Distribution
Venue: Auditorium, RUET</td></tr>
<tr><th colspan="2">03.00 PM – 4.00 PM: Inaugural Ceremony
Venue: Auditorium, RUET</th></tr>
<tr><td>3:00 PM
3:05 PM
3:10 PM
3:20 PM
3:28 PM
3:35 PM
3:42 PM
3:50 PM
4:00 PM
4:05 PM
4:35 PM</td><td>Chair, Guest and Participants take their seats
Welcome speech by the Organizing Chair, Dr. Md. Nazrul Islam Mondal, Head & Professor, CSE Dept., RUET
Speech by the Technical Chair, Dr. Md. Shahid Uz Zaman, Professor, CSE Dept., RUET
Speech by the Special Guest, IEEE, Bangladesh Section
Speech by the Special Guest, Engr. Mohammad Enamul Kabir, Director (Training), BCC, Dhaka
Speech by the Special Guest, Dr. Anirban Mukhopadhyay, Professor, University of Kalyani, India
Speech by the Special Guest, S.M. Ashraful Islam, Executive Director, BCC, Dhaka
Speech by the Chief Guest, Prof. Dr. Mohd. Rafiqul Alam Beg, Vice-Chancellor, RUET
A Short Presentation about the CSE department by Joint Technical Secretary, Dr. Md. Al Mamun, CSE Dept., RUET
Vote of thanks by the Organizing Secretary, Dr. Md. Rabiul Islam, Professor, CSE Dept., RUET
Invited Speech, Engr. Mohammad Enamul Kabir, Director (Training), BCC, Dhaka</td></tr>
<tr><th colspan="2">4:10 PM to 4:30 PM: Prayer & Tea Break</th></tr>
<tr><td>4:35 PM
To
5:30 PM</td><td>Invited Speech
Venue : Auditorium, RUET
Session Chair : Prof. Dr. Md. Shahid Uz Zaman, CSE, RUET
Session Co-Chair : Dr. Boshir Ahmed, CSE, RUET
Title : Building an Eco-System for Digitalization Using National Enterprise Architecture
Speaker : Engr. Mohammad Enamul Kabir
Director (Training), BCC, Dhaka</td></tr>
<tr><th colspan="2">27 November, 2015 (Friday) : Day 2</th></tr>
<tr><td>8.30 AM To
9.00 AM</td><td>Registration and Kit Distribution
Venue: Conference Lobby, Department of CSE, RUET</td></tr>
<tr><td>9:00 AM
To
09:40 AM</td><td>Keynote Session I
Venue : Auditorium, RUET
Session Chair : Prof. Dr. Md. Nasim Akhter, CSE, DUET
Session Co-Chair : Prof. Dr. Md. Nazrul Islam Mondal, CSE, RUET
Title : Semantic Web-Basics and Applications
Speaker : Prof. Dr. Muhammad Masroor Ali, CSE, BUET</td></tr>
</table>

9:45 AM To 10:30 AM	**Keynote Session II** **Venue** : Auditorium, RUET **Session Chair** : *Prof. Dr. Md. Mamun-or-Rashid,* CSE, DU **Session Co-Chair** : *Dr. Md. Ali Hossain,* CSE, RUET **Title** : Does Sequence of Presentation matters? Study of Learning Phenomena by a Computational Model of Semantic Network Growth via Episodic Comprehension **Speaker** : *Prof. Dr. Javed I. Khan,* CS, Kent State University, USA

10:30 AM to 10:50 AM: Tea Break

27 November, 2015 (Friday)
Technical Session I (10.50 AM – 01.00 PM)

Venue: Conference Room , Dept. of CSE, RUET **Track: Artificial Intelligence**

Session Chair : *Prof. Dr. Kaushik Deb,* CSE, CUET
Session Co-Chair : *Dr. Md. Ali Hossain,* CSE, RUET

PI 5	A Vision Based Guide Robot System: Initiating Proactive Social Human Robot Interaction in Museum Scenarios *Md. Golam Rashed, Ryota Suzuki, Antony Lam, Yoshinori Kobayashi, Yoshinori Kuno*
PI 28	Phonetic Features Enhancement for Bangla Automatic Speech Recognition *Rasel Shuvro, Fariha Nusrat, Foyzul Hassan, Foysal Ahamed, Khondaker Mamun, Mohammad Huda*
PI 32	Bangla Pronunciation Error Detection System *Selina Parveen, Farhana Sarker, Rasel Shuvro, Khondaker Mamun, Mohammad Huda*
PI 75	A Belief Rule Based Expert System to Control Traffic Signals under Uncertainty *Mohammad Shahadat Hossain, Rashed Mustafa, Hoimonty Sinha*
PI 80	Review of integrated applications with AIML based chatbot *Md. Shahriare Satu, Md. Hasnat Parvez, Shamim-Al Mamun*
PI 87	Character Recognition System: Performance Comparison of Artificial Neural Networks and Genetic Algorithm *Shahazan Ali, Md. Nazrul Islam Mondal*

PI 119	Effects of Caffeine Doses on Cardiac Activity using Laser Doppler Flowmetry *Md. Uddin, M Reza, Mohiuddin*
PI 120	An Empirical Framework for Parsing Bangla Assertive, Interrogative and Imperative Sentences *Mohammed Safayet Arefin, Lamia Alam, Shayla Sharmin, Mohammed Hoque*
PI 145	Land Cover Classification for Satellite Images based on Normalization Technique and Artificial Neural Network. *Boshir Ahmed, Md. Abdullah Al Noman*
PI 159	An Empirical Analysis of Attribute Skewness over Class Imbalance on Probabilistic Neural Network and Naïve Bayes Classifier *Nazmul Shahadat, Biprodip Pal*

27 November, 2015 (Friday)
Technical Session II (10.50 AM – 01.00 PM)

Venue: Seminar Room, Dept. of CSE, RUET **Track: Computer Networks & Security**

Session Chair : Prof. Dr. Md. Nasim Akhter, CSE, DUET
Session Co-Chair : Dr. Md. Al Mamun, CSE, RUET

PI 27	Performance Analysis and Redistribution among RIPv2, EIGRP & OSPF Routing Protocol *Kanti Dey, Md. Mobasher Ahmed, Kazi Tanvir Ahmmed*
PI 39	A Novel Elliptic Curve Cryptography Scheme Using Random Sequence *Fatema Akhter*
PI 41	VANET Topology Based Routing Protocols & Performance of AODV, DSR Routing Protocols in Random Waypoint Scenarios *Aditi Roy, Bijan Paul, Sanjit Kumar Paul*
PI 54	Performance Analysis of DCS-Based Limited Wavelength Interchanging Cross-Connects in WDM Network *Punab Chandra Kundu, Dr. Md. Rabiul Islam, Pejush Chandra Sarker, Satya Prasad Majumder*

PI 59	A Comparative Study of Different Dispersion Compensation Techniques in Long Haul Communication *Md. Talha, Kazi Salam, Hasan Zaman*
PI 62	vSIM : The Next Generation Mobile Technology *M. M. Rahman, A. B. M. A. Islam*
PI 72	A Secret Key-Based Security Architecture for Wireless Sensor Networks *Md Mokammel Haque, Anuva Chowdhury, Farzana Alam, Shanta Chowdhury*
PI 91	Cross-correlation Based Approach of Underwater Network Size Estimation with Unequal Sensor Separation *D.K. Mondal, S. A. H. Chowdhury, Q. N. Ahmed, M. S. Anower*
PI 117	Blind Audio Watermarking Based on Fast Walsh-Hadamard Transform and LU Decomposition *Pranab Kumar Dhar*
PI 125	Paradigm Shift towards Cloud Computing for Banking Sector *Mohammad Anik Islam, Md. Khaled Ben Islam, Md. Nazmus Shakib Beg*
PI 160	Cooperative Game Theory based Load Balancing in Long Term Evolution Network *Subarno Saha, Rajkin Hossain, Muhidul Islam Khan*

01.00 PM – 2.20 PM: Prayer & Lunch Break
Lunch: Cafeteria, RUET

27 November, 2015 (Friday)
Technical Session III (2.30 PM – 4.30 PM)

Venue: Conference Room, Dept. of CSE, RUET **Track: Image Processing & GIS**

Session Chair : Prof. Dr. Md. Kazi Khairul Islam, EEE, IUT
Session Co-Chair : Prof. Dr. Md. Rabiul Islam, CSE, RUET

PI 17	Performance Comparison of Three Optimized Alternative Pulse Shaping Filters with the Raised Cosine Filter for Wireless Applications *Tushar Kanti Roy, Monir Morshed, Md Ferdous Pervej*

PI 24	A Heuristic Solution of the Vehicle Routing Problem to Optimize the Office Bus Routing and Scheduling using Clarke & Wright's Savings Algorithm *Emrana Kabir Hashi, Rokibul Hasan, Md. Shahid Uz Zaman*
PI 26	Interaction with Large Screen Display using Fingertip & Virtual Touch Screen *Dejan Chandra Gope, Md. Shafiqul Islam*
PI 33	An Adjustable Novel Window Function with its Application to FIR Filter Design *Hrishi Rakshit, Muhammad Ullah*
PI 71	Development of a Smart Learning Analytics System Using Bangla Word Recognition and an Improved Document Driven DSS *Wali Mohammad Abdullah, Kazi Lutful Kabir, Nazmul Hasan, Sharmin Islam*
PI 89	Histogram based Water Quality Assessment in Satellite Images *Mumu Aktar, Md Al Mamun, Md. Shamimur Rahman Shuvo, Md. Ali Hossain*
PI 106	A Real- Time Face to Camera Distance Measurement Algorithm Using Object Classification *Md Hossain, Md Mukit*
PI 139	Closest Class Measure based Subspace Detection for Hyperspectral Image Classification *Md. Ali Hossain, Md. Shahid Uz Zaman, Md Al Mamun, Md. Nazrul Islam Mondal*
PI 162	Bangladeshi Road Sign Detection Based on YCbCr Color Model and DtBs Vector *Kaushik Deb*
PI 163	GIS-based Extraction of Open Space in Rajshahi City *Md. Shahid Uz Zaman, Fatima Jahan Sarmin*

27 November, 2015 (Friday)
Technical Session IV (2.30 PM – 3.20 PM)

Venue: Seminar Room, Dept. of CSE, RUET **Track: Algorithm & Application**

Session Chair : Prof. Dr. Muhammad Masroor Ali, CSE, BUET
Session Co-Chair : Dr. Pranab Kumar Dhar, CSE, CUET

PI 64	Smart Vehicle Accident Detection and Alarming System Using a Smartphone *Ahmed Imteaj, Mahfuzulhoq Chowdhury, Adnan Faiz*

PI 143	Multi-temporal FFT Regression *Md. Al Mamun, Md. Nazrul Islam Mondal, Boshir Ahmed, Md. Shahid Uz Zaman, Shyla Afroge*
PI 149	A Novel Idea of Tackling Hard Bit-vector Problems in a Beneficial Way by Eager Solver *Chowdhury Sajadul Islam, Md. Nahid Hasan*
27 November, 2015 (Friday) Technical Session V (3.25 PM – 4:15 PM)	
Venue: Seminar Room, Dept. of CSE, RUET	**Track: Internet & Web**
Session Chair: Prof. Dr. Md. Shahid Uz Zaman, CSE, RUET Session Co-Chair: Dr. Boshir Ahmed, CSE, RUET	
PI 3	Issues in Entry Creation for an Educational Institution and Searching in Semantic Web *Syeda Nyma Ferdous, Jannatut Tabassum, Muhammad Masroor Ali, Abdullah Al Noman*
PI 68	Developing a Framework for Analysing Web Data to Generate Recommendation *Tanjila Khanam, Mohammad Arefin, Md. Islam*
PI 153	Phishing Attack Detection Using Taxonomy Model *Chowdhury Sajadul Islam*
4:20 PM to 4:40 PM: Prayer & Tea Break	
27 November, 2015 (Friday) Technical Session VI (4.40 PM – 5.20 PM)	
Venue: Seminar Room, Dept. of CSE, RUET	**Track: Software & Middleware**
Session Chair : Prof. Dr. Shamim Ahmed, CSE, RU Session Co-Chair : Prof. Dr. Md. Shamsul Arefin, CSE, CUET	
PI 38	Telemedicine in South Asia for Rural People: Current Scenario and Future Recommendations *Uzzal Prodhan, Muhammad Rahman, Israt Jahan*
PI 92	Introducing Spatial Orientation Sensitive Cellphone Messaging System for Blind People *Parijat Prashun Purohit, Fardina Fathmiul Alam, S. M. Mostaq Hossain*

<table>
<tr><td>PI 115</td><td>Smartphone based Teacher-Student Interaction Enhancement System
Ahmed Imteaj, Saad Sajjad</td></tr>
<tr><td colspan="2">5:20 PM to 5:40 PM: Prayer & Tea Break</td></tr>
<tr><td colspan="2">27 November, 2015 (Friday) : Day 2</td></tr>
<tr><td>5:40 PM
To
06:20 PM</td><td>Keynote Session III
Venue : Auditorium, RUET
Session Chair : Prof. Dr. S. M. Abdur Razzak, EEE, RUET
Session Co-Chair : Prof. Dr. Md. Rabiul Islam, CSE, RUET
Title : Iris Biometric and its Application in Cyber Identity
Speaker : Dr. Kaushik Roy, North Carolina A&T State Univesrity, USA</td></tr>
<tr><td>6:25 PM
To
07:00 PM</td><td>Keynote Session IV
Venue : Auditorium, RUET
Session Chair : Prof. Dr. Md. Shahid Uz Zaman, CSE, RUET
Session Co-Chair : Prof. Dr. Md. Shamsul Arefin, CSE, CUET
Title : Multiobjective Genetic Algorithms for Data Clustering
Speaker : Dr. Anirban Mukhopadhyay, CSE, University of Kalyani, India</td></tr>
<tr><td colspan="2">Day 2: 27 November, 2015 (Friday)
07:30 PM – 8:00 PM: Closing & Award Night
Venue: Master Chef Ball Room, Rajshahi</td></tr>
<tr><td>7:30 PM
7:35 PM
7:45 PM
7:47 PM
7:50 PM</td><td>Chair, Guest and Participants take their seats
Speech by the Special Guest, S.M. Ashraful Islam, Executive Director, BCC, Dhaka
Best Paper Award by the Chief Guest, Prof. Dr. Mohd. Rafiqul Alam Beg, Vice-Chancellor, RUET
Speech by the Chief Guest, Prof. Dr. Mohd. Rafiqul Alam Beg, Vice-Chancellor, RUET
Closing Speech by the Organizing Chair, Dr. Md. Nazrul Islam Mondal, Head & Professor, CSE Dept., RUET</td></tr>
<tr><td colspan="2">Day 2 : 27 November, 2015 (Friday)
08:00 PM – 10:00 PM: Grand Dinner
Venue: Master Chef Ball Room, Rajshahi</td></tr>
</table>

Keynote Speech

Prof. Dr. Javed I. Khan
Computer Science Department
Kent State University, Ohio, USA
Javed@ kent.edu

Keynote Title: Does Sequence of Presentation Matters? Study of Learning Phenomena by a Computational Model of Semantic Network Growth Via Episodic Comprehension.

Abstract: Though classical learning research in computer science has not delved too much into reading as a means of learning it is one of the most common means of learning for modern humans. Text comprehension is one of the major processes by which we modern persons learn throughout our life. Classical learning research has investigated a number of aspects of learning phenomena from its early days. It has given rise to models of artificial learning where computing algorithms can mimic various forms of human learning such classification and learning of hidden functions, Hebb's reinforcement, self-organization, dynamic shift of attention, etc. Yet, human learning is a complex phenomenon way too many aspects are yet to be understood. It seems reading based learning is one of the phenomenon which has received relatively little attention in computer science until now. In this talk we present our work how computing can shed new light into the understanding of this form of learning. We particularly investigate an interesting topic- how the presentation sequence of concepts during reading might impact the semantic network of meaning generated. Text comprehension is postulated as an incremental and episodic process in cognitive science researchers. Concepts which are recognized are subsequently integrated into the background knowledge represented by a semantic network. For analysis, we propose a new computational model of the segmentation and integration (SSI) process and summarize the learning. We present a reading experiment in which groups of readers are presented same sample texts but in varying order of sentences to simulate different sequences of concept presentation. We then compare the resulting summary networks of the reader groups revealing a number of interesting observations. We also share additional complex network experiments on the growth characteristics of the individual semantic networks.

Speaker Biography: Dr. Javed I. Khan is a professor and Chair of Department of Computer Science at Kent State University. Dr. Javed I. Khan's research team specializes in applying multi-area expertise in cross-cutting problems in advanced networking, perceptual engineer, and modeling of cognitive, social and biological systems. His lab is currently working on perceptual comprehension and active & programmable networking. His cross-area research has been funded by various agencies including US Defense Advanced Research Project Agency (DARPA), National Science Foundation (NSF), NASA, AFRL, World Bank and State of Ohio. He has first-authored 100+ reviewed publications. He has served as editor for 100+ journal articles in his area, in national panel on next generation communications. Besides being an active researchers and educator, Dr. Khan is passionately active in international educational technology for mitigating digital divide. He helped planning and designing two national high speed advanced network infrastructures- BDREN and NgREN which are now being implemented by the national governments. He served as the area expert in the Fulbright National Roster of experts as Senior Specialist on high performance education networking and digital divide. Dr. Khan has received his PhD from University of Hawaii at Manoa and B.Sc. from Bangladesh University of Engineering & Technology (BUET) and a Past Open Grants Fellow at East West Center at Hawaii. More information about Dr. Khan's research can be found at medianet.kent.edu

Keynote Speech

Prof. Dr. Ruhul A Sarker
Senior Academic
School of Engineering and IT, University of New South Wales
Canberra, Australia

Abstract: In designing any evolutionary algorithm, it is usually one main search operator is selected to perform the search process within the defined search space. Because of the variability in the mathematical properties of different decision and optimization problems, no single algorithm with one search operator can solve all types of practical problems efficiently. We have recently introduced a general framework that allows multiple search operators and/or multiple evolutionary algorithms to work under a single algorithm structure for solving decision and optimization problems. The proposed framework has added a significant problem solving capacity for covering a wide range of decision and optimization problems, and at the same time it has shown consistent superior performance over the existing algorithms. This pioneering research has opened up a new research direction for evolutionary algorithm design. In this talk, different adaptive and dynamic configuration aspects for evolutionary algorithms will be highlighted and our experiences in solving different practical problems will be shared.

Speaker Biography: Ruhul Sarker received his Ph.D. in 1991 from Dalhousie University, Halifax, Canada, and B.Sc. Eng.(1982) and M.Eng.(1984) from BUET, Dhaka. He is currently a senior academic in the School of Engineering and IT, University of New South Wales, Canberra, Australia. From 2011 to 2014, he was the Deputy Head (Research) in the school. His main research interests are Evolutionary Computation, and Applied Operations Research, and their interfaces. He is the lead author of the book 'Optimization Modelling: A Practical Approach' published by Taylor & Francis, USA. He has published more than 270 refereed articles in the international journals, edited books, and conference proceedings. He has edited /co-edited 8 books on specialized topics in computational Intelligence including one on 'Evolutionary Optimization'. He is currently an associate editor of Memetic Computing Journal, Flexible Service and Manufacturing Journal, and Journal of Industrial and Management Optimization. Prof. Sarker was a technical co-chair for IEEE CEC'2003 and a Proceedings Co-chair for IEEE WCCI'2012. He is a member of IEEE CIS Task Force on Differential Evolution and Evolutionary Multiobjective Optimization. He has delivered invited and keynote speeches in a number of major conferences.

Keynote Speech

Kaushik Roy, Ph.D.
North Carolina A&T State University, Greensboro, NC, USA

Keynote Title: Iris Biometric and Its Application in Cyber Identity

Abstract: Person identification systems based on iris biometrics have received huge attention due to its applicability to many areas, including national border control, forensics, secure financial transactions, and cyber identity. Most of the existing iris recognition algorithms mainly depend on the iris images that are captured in a cooperative environment to ensure the higher recognition accuracy. However, in many cases, iris image acquisition process is affected by different noise factors such as illumination variations, noncooperation of persons, head rotations, gaze directions, and camera angles. These nonideal factors in data acquisition process may result in motion blurs, reflections, eyelash and eyelid occlusions, and pupil center deviation and further hamper the iris localization performance. Therefore, iris recognition in a nonideal situation still remains a challenging issue.

This talk will begin with an overview of the non-cooperative iris recognition. The remainder of the talk will focus on multibiometric systems. Application of iris recognition to biometric-based access control will also be discussed.

Speaker Biography: Kaushik Roy received his PhD from Concordia University, Montreal, QC, Canada in 2011 in Computer Science. He completed his MS degree in Computer Science from the Concordia University in 2006 and B.Sc. degree in Computer Science and Technology from University of Rajshahi, Bangladesh in 2000. Kaushik Roy is currently an Assistant Professor at the Department of Computer Science and Director of the Cyber Identity and Biometric (CIB) lab, North Carolina A&T State University, USA. Previously, he worked as a postdoctoral fellow in the Department of Electrical and Computer Engineering, University of Waterloo, ON, Canada during 2011-2012. He also taught at Rajshahi University of Engineering and Technology (RUET) as a lecturer in the Department of Computer Science and Engineering during 2001-2004. His research interests include biometrics, cyber identity, biometric based cyber security, game theory, information fusion, and machine learning. Dr. Roy is the author or co-author of over 60 articles in journals, book chapters and conference proceedings. Recent grants are for research on cyber identity framework, author identification, multispectral iris recognition, large data network, and cloud identity.

Keynote Speech

Dr. Anirban Mukhopadhyay
Department of Computer Science and Engineering
University of Kalyani, Kalyani, West Bengal, India

Keynote Title: Multiobjective Genetic Algorithms for Data Clustering

Abstract: Clustering is an important data mining technique where a set of patterns, usually vectors in multidimensional space, are grouped into clusters based on some similarity or dissimilarity criteria. Clustering techniques aim to find a suitable grouping of the input dataset so that some criteria, such as compactness, separation, and connectivity are optimized. A straightforward way to pose clustering as an optimization problem is to optimize some cluster validity index that reflects the goodness of the clustering solutions. All possible partitionings of the dataset and the corresponding values of the validity index define the complete search space. Traditional partitional clustering techniques, such as K-means and fuzzy C-means, employ greedy search techniques over the search space to optimize the compactness of the clusters. These algorithms often get stuck at some local optima depending on the choice of the initial cluster centers. Moreover, they optimize a single cluster validity index (compactness in this case), and therefore do not cover different characteristics of the datasets. To overcome the problem of local optima, some evolutionary global optimization tools such as Genetic Algorithms (GAs) have been widely used to reach the global optimum value of the chosen validity measure. Conventional GA-based clustering techniques use some validity measure as the fitness value. However, no single validity measure works equally well for different kinds of datasets. Thus it is natural to simultaneously optimize multiple such measures for capturing different characteristics of the data. Simultaneous optimization of multiple objectives provides improved robustness to different data properties. Hence it is useful to utilize multiobjective GAs (MOGAs) for clustering. In this talk, I will first describe some preliminaries of multiobjective optimization and Pareto optimality. Subsequently, I will discuss a multiobjective GA-based clustering algorithm. Finally I will demonstrate some applications of multiobjective clustering technique in remote sensing and bioinformatics.

Speaker Biography: Dr. Anirban Mukhopadhyay is an Associate Professor and former Head of the Department of Computer Science and Engineering, University of Kalyani, Kalyani, West Bengal, India. He did his B.E. from National Institute of Technology, Durgapur, India, in 2002 and M.E. from Jadavpur University, Kolkata, India, in 2004, respectively, both in Computer Science and Engineering. He obtained his Ph.D. in Computer Science from Jadavpur University in 2009. Dr. Mukhopadhyay is the recipient of the University Gold Medal and Amitava Dey Memorial Gold Medal from Jadavpur University in 2004 for ranking first class first in M.E. He received Erasmus Mundus fellowship in 2009 to carry out post-doctoral research at University of Heidelberg and German Cancer Research Center (DKFZ), Heidelberg, Germany during 2009-10. Dr. Mukhopadhyay also visited I3S laboratory, University of Nice Sophia-Antipolis Nice, France in 2011 as a Visiting Professor, and University of Goettingen, Germany, as a Visiting Scientist with DAAD scholarship in 2013. He has received Institution of Engineers, India (IEI) Young Engineers Award in Computer Engineering Discipline in 2014, and Indian National Academy of Engineering (INAE) Young Engineer Award in 2014. He has coauthored one book and over 125 research papers in various reputed International Journals and Conferences. He is a senior member of Institute of Electrical and Electronics Engineers (IEEE), USA, and member of Association for Computing Machinery (ACM), USA, and International Association of Engineers (IAENG), Hong Kong. Dr. Mukhopadhyay is currently the Secretary of IEEE Computational Intelligence Society, Kolkata Chapter. His research interests include soft and evolutionary computing, data mining, multiobjective optimization, pattern recognition, bioinformatics, and optical networks.

Keynote Speech

Dr. Muhammad Masroor Ali
Professor
Department of Computer Science and Engineering,
Bangladesh University of Engineering & Technology

Keynote Title: Semantic Web-Basics and Applications

Abstract: The semantic web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries [1]. We can also define semantic web as, a collection of technologies and standards that allow machines to understand the meaning (semantics) of information on the Web [2]. The term "semantic web" was coined by World Wide Web Consortium (W3C) director Sir Tim Berners-Lee, and was formally introduced to the world in his seminal paper [3] in 2001. From the current web, we want to formulate a smart data integration agent. For a given resource, we would like to know everything that has been said about it. More specifically, we would like to accomplish this goal by collecting as much information as possible about this resource. We will then understand it by making queries against the collected information [2]. In order to accomplish this goal, we have to understand what exactly is there in a web page. If we generalize our goal a bit, and want to have a web understandable to machines, we observe that the three major activities we perform on the web are, searching, information integration and web data mining. For this, the changes we need to have on each web site include, but not limited to, paraphrasing from [2], standard model to represent knowledge on the web, acceptance of this model as a standard by all the web sites, methodology to create statements on each web site, usage of some common terms and relationships for a given domain and being able to define these common terms and relationships. We will want to build models, calculate with knowledge, and exchange information [4]. In order to attain the aforementioned goals, we have a number of components in the semantic web available. Resource Description Framework (RDF) is a standard model for data interchange on the Web. RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed. RDF extends the linking structure of the Web to use URIs to name the relationship between things as well as the two ends of the link (this is usually referred to as a "triple"). Using this simple model, it allows structured and semi-structured data to be mixed, exposed, and shared across different applications [5]. RDF Schema provides a data-modelling vocabulary for RDF data. RDF Schema is a semantic extension of RDF. It provides mechanisms for describing groups of related resources and the relationships between these resources [6]. SPARQL is an RDF query language, that is, a semantic query language for databases, able to retrieve and manipulate data stored in RDF format [7]. However, the variety of things you can say in RDF and RDF Schema is very limited. RDF is limited to binary ground predicates, and RDF Schema is limited to a subclass hierarchy and a property hierarchy, with domain and range definitions of these properties. However, in many cases we need to express more advanced, more expressive, knowledge. The Web Ontology Working Group and the OWL Working Group, identified a number of characteristic use cases for the semantic web that require much more language features than those that RDF and RDFS have to offer. The resulting language, OWL, for the Web Ontology Language, is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things. It is closely related to a fragment of a family of logics that are specially crafted for representing terminological knowledge. The features of these Description Logics (DL) are well understood by the community. OWL2 is the second iteration of the OWL language [8], [9]. For logic and inference in the semantic web, predicate logic can be used. The languages of RDF and OWL2 profiles can be viewed as specializations of predicate logic [9]. At the end of the presentation, some real life case studies of semantic web applications from the areas of supply chain management, media management, data integration, web search and ecommerce will be explored [9], [10].

REFERENCES

[1] "W3C semantic web activity," http://www.w3.org/2001/sw/.

[2] L. Yu, A Developer's Guide to the Semantic Web, 2nd ed. Springer, 2014.

[3] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," Scientific American, vol. 284, no. 5, pp. 34–43, May 2001.

[4] P. Hitzler, M. Krotzsch, and S. Rudolph, ¨ Foundations of Semantic Web Technologies. Chapman & Hall/CRC, 2009.

[5] "Resource description framework (RDF)," http://www.w3.org/RDF/.

[6] "RDF schema 1.1," http://www.w3.org/TR/rdf-schema/.

[7] E. Prud'hommeaux and A. Seaborne, "Sparql query language for rdf," http://www.w3.org/TR/rdf-sparql-query/, Jan. 2008.

[8] "Web ontology language (OWL)," http://www.w3.org/2001/sw/wiki/ OWL.

[9] G. Antoniou, P. Groth, F. van Harmelen, and R. Hoekstra, A Semantic Web Primer, 3rd ed. The MIT Press, 2012.

[10] "Example semantic web applications," http://www.cambridgesemantics. com/semantic-university/example-semantic-web-applications.

Speaker Biography: Muhammad Masroor Ali received his B.Sc. Engineering degree from BUET in 1988 and M. Engg. and Ph.D degrees respectively in 1994 and 1997 from Kyushu University, Fukuoka, Japan. Currently, he is a Professor in the Department of CSE, BUET where he has been working for more than twenty five years. With current research interest focused in the field of semantic web, Professor Ali has published around forty papers in refereed journals and conferences. He has successfully supervised more than one hundred undergraduate and around ten graduate theses. He has worked as a member in the organizing committee/program committee/reviewer panel in a number of international conferences and journals. He was the Convener, Sectoral Committee for Computer Science and Engineering, Board of Accreditation of Engineering and Technical Education (BAETE) in the period May, 2003 to July 2008. He was also a Member, Committee for eLearning Material Development, under the Alberta-BUET Linkage Project. He was a member of the Editorial Advisory Board for the book Technical Challenges and Design Issues in Bangla Language Processing. Professor Ali has been a successful contributor as a member of BUET panel of experts in various projects at national level. In addition to teaching, he served in a number of administrative positions at BUET. Professor Ali likes to call himself an, "Open Source Software Enthusiast" and wishes he could make everybody use LATEX for all their publications.

Invited Talk

Engr. Mohammad Enamul Kabir
Director (Training)
Bangladesh Computer Council

Title: Building an Eco-System for digitalization using National Enterprise Architecture

Abstract: National Enterprise Architecture (NEA), a framework that the Government of Bangladesh has planned to facilitate digital services to the citizen. In the context of Digital Bangladesh, which the Government of Bangladesh has envisioned to achieve within 2021, the NEA framework provides the foundation of the eco-system for digitalization of the Government. It is broadly designed covering architectural domains of governance, operations, security, and with data, application, interoperability, mobility, technology for business having access and presentation. NEA establishes a service bus architecture with provision for both open source technology and Oracle SOA ensuring interoperability between various services and application layers. The service bus collaborates with different layers covering integration, workflow, rules, processes and notifications. In the eco-system, this framework will ensure end-to-end digital process orchestrations and thus improving quality of life for the citizen through seamless use of e-services across the Government. In the analysis of NEA a systematic approach based on component model is applied to explore the detail orchestration needs of different layers. The research also identified and elaborated detail process interactions for the framework components along with the needs for additional interfaces using the UML modeling of the requirements which is intended to contribute in the test case development for the functional integration of NEA and its service bus.

ABSTRACTS

Paper ID: 03

Issues in Entry Creation for an Educational Institution and Searching in Semantic Web

Syeda Nyma Ferdous[1], Jannatut Tabassum[2], Muhammad Masroor Ali[3], Abdullah Al Noman[4]
[1,2,4] Department of Computer Science and Engineering
Military Institute of Science and Technology, Dhaka, Bangladesh
[3]Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
[1]nymaferdous@gmail.com, [2]sweetytabassum@yahoo.com, [3]mmasroorali@cse.buet.ac.bd, [4]noman14032@gmail.com

Abstract: The main purpose of semantic web is driving the evolution of current web by enabling users to find, share and combine information more easily. In traditional web, it is both tedious and time-consuming to find our required information. But with the development of semantic web, it gets much more efficient and faster to obtain required information ignoring all other unnecessary data. Semantic web understands the meaning of data and can make intelligent decision based on available data. To build up the necessary web information and their relationship or the ontology for the semantic web, creation of valid RDF document is needed. These documents can be deployed and retrieved by the search engine after URL submission and completion of indexation. This paper describes the process of our RDF document creation for an educational institution, deployment, URL submission for indexation in the search engine, Swoogle and retrieval of the document. It emphasizes on identification of the issues or challenges throughout the whole process.

Paper ID: 05

A Vision Based Guide Robot System: Initiating Proactive Social Human Robot Interaction in Museum Scenarios

M. Golam Rashed[1], R. Suzuki[1], A. Lam[1], Y. Kobayashi [1,2,] Y. Kuno[1]
[1]Saitama University, Japan, [2] JST, PRESTO, Japan
[golamrashed, suzuryo, antonylam, yosinori, kuno]@cv.ics.saitama-u.ac.jp

Abstract: This paper presents a guide robot system which observes people's interest and intention towards paintings in museum scenarios and proactively offers guidance to them using a guide robot, if needed. To do that, we utilized multiple USB video camera sensors to support the guide robot in detecting and tracking people's visual focus of attention (VFOA) to-ward paintings. In this study, we consider each person's head orientation and profile information and compute importance values to identify a target-person that may be interested in a particular painting. After identifying the target-person. the guide robot moves autonomously through an appropriate motion path from the so called public-distance to his/her social-distance to explain details about the painting to which s/he is interested. Furthermore, we demonstrate the viability of our guide robot system by experimenting with the Robovie-R3 as a museum guide robot. Finally, we tested the system to validate its effectiveness.

Paper ID: 17

Performance Comparison of Three Optimized Alternative Pulse Shaping Filters with the Raised Cosine Filter for Wireless Applications

T. K. Roy[1], M. Morshed[2], M. F. Pervej[1]
[1]Dept. of Electronic & Telecommunication Engineering
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
[2]Dept. of Information & Communication Technology
Mawlana Bhashani Science & Technology University, Tangail, Bangladesh
[1]roy_kanti@yahoo.com, [2]monirmorshed.mbstu@gmail.com, [3]ferdous.ete09@gmail.com

Abstract: This paper is concerned with the performance and cost analysis of different pulse shaping filters such as raised cosine, finite impulse response (FIR) Nyquist, FIR half-band and infinite impulse response (IIR) linear phase half-band for wireless applications. Note that the pulse shaping is the process of changing the waveform of transmitting pulses to reduce the interferences keeping a signal in an allotted bandwidth. This paper proposes an optimal design for the pulse shaping filters to reduce their effective costs in the area of wireless communication applications. Here, we analyze the performance of pulse shaping filters based on the minimum stop band attenuation required to suppress the inter-symbol interference (ISI) and inter-channel interference (ICI). Finally, we compare the performance of raised cosine, FIR Nyquist, FIR half-band and IIR linear phase half-band filters in terms of bit-error rate (BER) and hardware requirements. Our results demonstrate that the hardware implementation costs for the FIR Nyquist, FIR half-band and IIR half-band filters are about 60% to 90% less than the raised cosine filter.

Paper ID:24

A Heuristic Solution of the Vehicle Routing Problem to Optimize the Office Bus Routing and Scheduling using Clarke & Wright's Savings Algorithm

Emrana Kabir Hashi[1], Md. Rokibul Hasan[2], Md. Shahid Uz Zaman[3]
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh
[1]emranakabir@gmail.com, [2]rakib063049@gmail.com, [3]szaman22@yahoo.com

Abstract: Office bus routing is a rising problem because transportation service should be more efficient, safe and reliable. This research try to answer a question how to facilitate employees with transportation service that would be fast, cost effective, and in a timely manner. This research work discourse the Vehicle Routing Problem and scheduling issues of transportation service for Dhaka City, Bangladesh. Our implementation will help to develop office bus routing and scheduling prototype model. This model will help to design transportation management system of any organization. From this implementation, it will able to find shortest and fastest routes and schedule of office bus, and also allocate bus stops that will use to pick up and drop employees according to their residence. This application has considered cost matrix of time, routes length and the vehicles with same capacity. We have implemented this application with Clarke & Wright's savings algorithm. This paper concludes with the solutions of schedules and visualization of routes with GIS techniques that will helps to develop decision support system for any kind of Vehicle Routing Problem.

Paper ID:26

Interaction with Large Screen Display using Fingertip & Virtual Touch Screen

Dejan Chandra Gope[1], Md. Shafiqul Islam[2]
Computer Science and Engineering
Dhaka University of Engineering and Technology, Dhaka, Bangladesh
[1]dejan.gope23@gmail.com, [2]msislam 80@yahoo.com

Abstract: Human-Computer Interaction (HCI) involves the study, planning, design and uses of the interaction between people (users) and computers. We present a nonintrusive system based on computer vision for human-computer interaction controlled by finger pointing gestures. A new approach for Interaction with Large Screen Display using Fingertip & Virtual Touch Screen by taking into account the location of the user and the interaction area available. We can estimate aninteraction surface virtual touch screen between the display and the user. Userscan use their pointing finger to interact with the display as if it was brought forward and presented directly in front of them, while preserving viewing angle. Users are allowed to walk around in a room and manipulate information displayed on its walls by using their own finger as pointing devices. Once captured and tracked in real-time using stereo vision camera, finger pointing gestures are remapped onto the current point of interest, thus reproducing in an advanced interaction scenario the "drag and click" behavior of traditional mice. It's a simple signal processing on images obtained from a regular laptop web-camera.

Paper ID:27

Performance Analysis and Redistribution among RIPv2, EIGRP & OSPF Routing Protocol

Golap Kanti Dey[1], Md. Mobasher Ahmed[2], Kazi Tanvir Ahmmed[3]
Department of Applied Physics, Electronics & Communication Engineering
University of Chittagong, Chittagong-4331, Bangladesh
[1]d.k.golap@gmail.com, [2]md.mobasher@gmail.com, [3]tanvir@cu.ac.bd

Abstract: In a network topology, it is very usual to use various kinds of routing protocol for forwarding packets. A routing table is used in the memory of a router that keeps the track of routes to particular network destination and the most popular routing algorithms used to forward packets are Routing Information Protocol (RIPv2), Enhanced Interior Gateway Routing Protocol (EIGRP) and Open Shortest Path First (OSPF). The ultimate concentration of this research work is to depict the performance analysis comparison of these three dynamic routing protocols and redistribution among the protocols. Eight Cisco routers and a switch are used in our simulated network topology where four routers with different directly connected with the switch take the responsibility for the redistribution algorithm.

Paper ID:28

Phonetic Features Enhancement for Bangla Automatic Speech Recognition

Sharif M Rasel Kabir[1], Fariha Nusrat[2], Foyzul Hassan[1], Foysal Ahamed[1],
Khondokar Mamun[1], Mohammad Nurul Huda[1]
[1]Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh
[2]Department of Computer Science and Engineering, Military Institute of Science and Technology, Dhaka, Bangladesh
shuvro.mist@gmail.com, fariha1502@gmail.com, foyzul.hassan@gmail.com, thisisfoysal@gmail.com,
mamun@cse.uiu.ac.bd, mnh@cse.uiu.ac.bd

Abstract : This paper discusses a phonetic feature (PF) based automatic speech recognition system (ASR) for Bangla (widely known as Bengali), where the PF features are enhanced. There are three stages in this method where the first step maps Acoustic Features (AFs) or Local Features (LFs) into Phonetic Features (PFs) and the second step incorporates inhibition/enhancement (In/En) algorithm to change the PF dynamic patterns where patterns are enhanced for convex patterns and inhibited for concave patterns. The final step is for normalizing the extended PF vector using Gram–Schmidt algorithm and then passing through a Hidden Markov Model (HMM) based classifier. In our experiment on speech corpus for Bangla, the proposed feature extraction method provides higher sentence correct rate (SCR), word correct rate (WCR) and word accuracy (WA) compared to the methods that not incorporated In/En network.

Paper ID:32

Bangla Pronunciation Error Detection System

Selina Parveen[1], Farhana Sarker[2], Sharif Mohammed Rasel Kabir[1],
Khondokar Mamun[1], Mohammad Nurul Huda[1]
[1] Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh
[2] Department of Computer Science and Engineering, University of Liberal Arts, Dhaka, Bangladesh
selina.parveen.aj@gmail.com, farhana.sarker@ulab.edu.bd, shuvro.mist@gmail.com, mamun@cse.uiu.ac.bd,
mnh@cse.uiu.ac.bd

Abstract: In this paper a pronunciation error detection system was modeled and also tested for large vocabulary, speaker independent and continuous speech recognizer for Bengali language. The recognizer was developed using Hidden Markov Model (HMM); and the Hidden Markov Modeling Toolkit was used to implement it. In the process, a corpusdatabase comprised of 3000 utterances that were used for training and 100 plus sentences for development and evaluation. The data was preprocessed in line with the requirements of the HTK toolkit. In order to support the acoustic models, a bigram language model was constructed. In addition, pronunciation dictionary was prepared and used as an input. Standard experiment for in depth performance analysis of the detector was designed and all the results are analyzed to get the proper insight of the error pattern, both at sentence and word level. The effect of Conjugant words, the effect of first letter of the word, the gender effect on error pattern along with the effect of colloquial dialect or local accent on error pattern is observed. The findings are quite promising and may open new possibilities to design an efficient pronunciation error correcting system for aiding the non-native Bengali speaking people.

Paper ID:33

An Adjustable Novel Window Function with its Application to FIR Filter Design

Hrishi Rakshit[1], Muhammad Ahsan Ullah[2]
Electrical and Electronics Engineering
Chittagong University of Engineering & Technology, Chittagong, Bangladesh
[1]rishirok@gmail.com, [2]ahsan_cuet@yahoo.com

Abstract: A new class of adjustable window function, based on combination of tangent hyperbolic function and a weighted cosine series, is proposed. The proposed window is adjustable in the sense that by changing its controlling parameter, one can change not only its shape but also its spectral characteristics. The spectral characteristic of the proposed window is studied and its performance is compared with Dolph-Chebyshev, Gaussian and Kaiser Windows. Simulation results show that the proposed window yields better ripple ratio and side-lobe roll-off ratio than the above named windows. Moreover, the paper represents the application of the proposed window in finite impulse response (FIR) filter design. The results confirm that the filter designed by the proposed window provides better spectral characteristics than Kaiser window. This paper also highlights an application of the proposed Filtering method in the area of eliminating noise from corrupted ECG signal

Paper ID: 38

Telemedicine in South Asia for Rural People: Current Scenario and Future Recommendations

Uzzal Kumar Prodhan[1], Muhammad Zahidur Rahman[2], Israt Jahan[2]
[1]Department of Computer Science & Engineering,
Jatiya Kabi Kazi Nazrul Islam University, Trishal, Maymensingh, Bangladesh
[2]Department of Computer Science & Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh
[1]uzzal_bagerhat@yahoo.com

Abstract: Telemedicine encompasses methods for electronically transmitting medical information to sustain and/or improve patient's health condition by using information and communication technology. Different countries of South Asiahave tried to introduce telemedicine services from different perspectives for the rural people where health care facilities are poor through its private and public partnership. This paper examines different telemedicine models of South Asia and findsout the potentials difficulties faced by different models and presents the factors which should be assessed carefully for thesuccessful implementation of telemedicine which will be feasible, easily maintainable, sustainable and cost effective solution for the poor people of rural areas of South Asia regions.

Paper ID: 39

A Novel Elliptic Curve Cryptography Scheme Using Random Sequence

Fatema Akhter
IEEE student member
Department of Computer Science and Engineering
Jatiya Kabi Kazi Nazrul Islam University, Bangladesh
fatema.kumu02@gmail.com

Abstract : This paper proposes a new Elliptic Curve Cryptography (ECC) scheme and a data mapping technique on elliptic curve over a finite field using maximum length random sequence generation. While its implementation, this paper also proposes a new algorithm of scalar multiplication for ECC. The proposed scheme is tested on various bits length of prime field and the experimental results show very high strength against cryptanalytic attack like random walk and better performance in terms of computation time comparing with standard approaches.

Paper ID:41

VANET Topology Based Routing Protocols & Performance of AODV, DSR Routing Protocols in Random Waypoint Scenarios

Aditi Roy[1], Bijan Paul[2], Sanjit Kumar Paul[3]
[1]Dept. of Computer Science and Engineering, Shahjalal University of Science and Technology, Bangladesh
[2]Dept of Computer Science, American International University, Bangladesh
[3]Dept of Mathematich, University of Asia Pacific, Bangladesh
[1]aditiroy1710@gmail.com, [2]bijancse@gmail.com, [3]sanjitkpaul@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.

Paper ID:54

Performance Analysis of DCS-Based Limited Wavelength Interchanging Cross-Connects in WDM Network

Punab Chandra Kundu[1], Md. Rabiul Islam[2], Pejush Chandra Sarker[2], Satya Prasad Majumder[3]
[1]Bangladesh Power Development Board, Bangladesh
[2]Department of Electrical and Electronic Engineering,
Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh
[3]Department of Electrical and Electronic Engineering,
Bangladesh University of Engineering & Technology, Dhaka, Bangladesh
punabkundu97@yahoo.com, rabiulbd@hotmail.com, pejush01@yahoo.com, spmajumder@eee.buet.ac.bd

Abstract: Crosstalk originated from various types of sources in delivery and coupling switches (DCS) based limited wavelength interchanging optical cross connect (OXC) architectures has been studied and analytical formulas have been carried out for those crosstalk. A comparative picture of the system performance has been depicted for all cases in DCS based limited wavelength interchanging OXC architectures which will help the system designer to choose the effective one. The system performance has been analyzed in MATLAB environment.

Paper ID: 59

A Comparative Study of Different Dispersion Compensation Techniques in Long Haul Communication

Md. Talha[1], Hasan U. Zaman[2], K. M. A. Salam[3]
Dept. of ECE, North South University, Dhaka, Bangladesh
[1]talhahome@yahoo.com, [2]hasan.zaman@northsouth.edu, [3]kazi.salam@northsouth.edu

Abstract: Optical fiber communication (OFC) provides us a very high bit rate data communication. There are different types of impairments and signal degradation mechanisms involved with this high speed communication system. In case of long haul communication, the most effective impairment is dispersion. It affects the signal very badly when signal travels a long distance. Different techniques are available to compensate dispersion. Not all of them are the same in performance. In this work, we study multiple dispersion compensation techniques and their performance measures and compare them with respect to the BER, Q-factor, Eye height, threshold etc. For this purpose, we used OptiSystem13 simulation software. Based on our analyses, we derive a conclusion on which technique is better for high speed long haul OFC networks.

Paper ID:62

vSIM: The Next Generation Mobile Technology

M. M. Rahman[1], A. B. M. A. A. Islam[2]
Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
[1]engr.mizanbd@ymail.com, [2]alim_razi@cse.buet.ac.bd

Abstract: Global system for mobile communication (GSM) is the most popular mobile network around the world. Subscriber Identity Module (SIM) is must for communicating using GSM network. Recently, virtual SIM technology has been introduced, which runs SIM cards remotely. Both of these technologies, i.e., SIM and virtual SIM, hold many risks for its users and can be attacked in many ways. In this paper we will introduce a new technology named software controlled virtual SIM (vSIM), which will solve these threats and it will provide a secure and reliable communication in GSM network.

Paper ID:64

Smart Vehicle Accident Detection and Alarming System Using a Smartphone

Adnan Bin Faiz[1], Ahmed Imteaj[2], Mahfuzulhoq Chowdhury[3]
Dept. of Computer Science and Engg.
Chittagong University of Engg. & Tech, Chittagong-4349, Bangladesh
[1]asifliveadnan@live.com, [2]imtu_1992@yahoo.com, [3]mahfuz_csecuet@yahoo.com

Abstract: Vehicle accident is the paramount thread for the people's life which causes a serious wound or even dead. The automotive companies have made lots of progress in alleviating this thread, but still the probability of detrimental effect due to an accident is not reduced. Infringement of speed is one of the elementary reasons for a vehicle accident. Therewithal, external pressure and change of tilt angle with road surface blameworthy for this mishap. As soon as the emergency service could divulge about an accident, the more the effect would be mitigated. For this purpose, we developed an Android based application that detects an accidental situation and sends emergency alert message to the nearest police station and health care center. This application is integrated with an external pressure sensor to extract the outward force of the vehicle body. It measures speed and change of tilt angle with GPS and accelerometer sensors respectively on Android phone. By checking conditions, this application also capable of reducing the rate of false alarm.

Paper ID:68

Developing a Framework for Analysing Web Data to Generate Recommendation

Tanjila Aftab Khanam[1], Mohammad Shamsul Arefin[2], Md. Shahedul Islam[3]
[1,2]Department of CSE, CUET, Bangladesh
[3]ICT Center, BSMRMU, Bangladesh
[1]ta_08_32@yahoo.com, [2]sarefin_406@yahoo.com, [3]shahed95@yahoo.com

Abstract: From the advent of the Internet, data contained in the web is increasing exponentially and the number of e-commerce companies is also increasing significantly. Nowadays, we can find that many e-commerce companies selling sameproducts and/or services at different prices. For a customer it is difficult to find an e-commerce company which will be best suited for him. Moreover, it is time consuming and tedious to search for a product in different online sites. For reducing this difficulty, in this paper, we develop a system that can extract web data from different e-commerce sites even though the language of the websites are different. We then analyse the extracted data and recommend best products / services from these sites to the users. For the experimental evaluation of our system, we consider two different languages: English and Bangla and extract books data from different online book stores considering these two languages and recommend books to the users. From the experimental results, we can find that our system can help the users in selection of their products easily and efficiently.

Paper ID:71

Development of a Smart Learning Analytics System Using Bangla Word Recognition and an Improved Document Driven DSS

Wali Mohammad Abdullah[1], Kazi Lutful Kabir[2], Nazmul Hasan[3], Sharmin Islam[4]
Department of Computer Science and Engineering,
Military Institute of Science and Technology, Dhaka, Bangladesh.
[1]shashi.mist@gmail.com, [2]mlutful.02526@gmail.com, [3]nazmulcse10@gmail.com, [4]priyaislam122@gmail.com

Abstract: A smart learning analytics system is proposed here. In the field of research, education learning analytics is manipulated asan optimistic way to render learning. In this proposed system, we put forward an idea regarding the design of learning analytics based on document-driven Decision Support System (DSS) and Bangla handwritten words recognition. The proposed learning analytics system aims at classifying the documents via syntactic analysis as well as categorizing them by semantic analysis of their contents. But the documents may be presented as printed format or hand-written format. To choose the optimal feature vector for hand written words recognition (especially for Bangla basic characters) projection-based features are mostly efficient. Decision makers use DSS for decision making which is actually a computerized information system. We propose a route through the classified observations for providing the explanation which may improve the acceptance of the decision maker with a better correction rate using kNN algorithm. This system can certainly assist researchers for learning as well as decision making.

Paper ID:72

A Secret Key-Based Security Architecture for Wireless Sensor Networks

Anuva Chowdhury[1], Farzana Alam Tanzila[2], Shanta Chowdhury[3], Md. Mokammel Haque[4]
Faculty of Electrical and Computer Engineering
Chittagong University of Engineering and Technoloty, Chittagong-4349, Bangladesh
[4] malin405@yahoo.com

Abstract: Information sharing and transmitting has been increased exponentially in today's world of Information and Communication Technology. Cryptography is a lasting solution to save the sensitive data from unauthorized access. Finding proper cryptographic structure for wireless sensor network (WSN) is almost challenging as there are some limitations of energy and memory resources of WSNs. Hence, this paper presents an efficient secret key based security architecture which aims to contribute robust security to the topic with ideas to man in the middle attacks and many other hacker situations. In this scheme, singular value decomposition (SVD) of pseudo-inverse matrix has been used for key generation. This scheme does not require key handshaking for data transmission and reception between the nodes. The analytical results demonstrate that this proposed scheme shows a significant gain in the rank of security, and is suitable for WSNs of the modern age.

Paper ID: 75

A Belief Rule Based Expert System to Control Traffic Signals under Uncertainty

Mohammad Shahadat Hossain[1], Hoimonty Sinha, Rashed Mustafa
Department of Computer Science and Engineering, University of Chittagong, Bangladesh.
[1] hossain_ms@cu.ac.bd

Abstract: Traffic congestion is a condition on road network that occurs as vehicle increases, and is characterized by slower speeds, longer trip times, and increased vehicular queuing. It effects the economic growth of a country, increases accidents, resource cost and environment pollution. One of the most cost-effective ways, to deal with this problem is by employing traffic control signals at the road intersections. Now-a-days, most signal controls are implemented with either fixed cycle time control or dynamic control. These conventional methods for traffic signal control fails to deal efficiently because they are unable to taking account of the uncertainty associated with traffic flow. Therefore, this paper presents the design, development and application of a belief rule based expert system (BRBES) with the capability of handling uncertainty. The system uses Belief Rule Base (BRB) as the knowledge representation schema and the evidential reasoning as the inference engine. The results generated by the system have been compared with a fuzzy logic based expert system (FLBES). The BRBES's reliability is better than th at of FLBS. The system applied in a number of road intersections of the Chittagong City of Bangladesh.

Paper ID:80

Review of Integrated Applications with AIML based Chatbot

Md. Shahriare satu[1], Md. Hasnat Parvez[2], Shamim-Al-Mamun[3]
Institute of Information Technology, Jahangirnagar university, Dhaka, Bangladesh
[1]shahriar.setu@gmail.com, [2]hasnatiit847@juniv.edu, [3]sam@juniv.edu

Abstract: Artificial Intelligence Markup Language (AIML) is derived from Extension Markup Language (XML) which is used to build up conventional agent (chatbot) artificially. There are developed a lot of works to make possible to use it in low cost, configuration and availability make possible to use it in various application. In this paper, we give a brief review of some application are used AIML chatbot for their conventional service. These application are related to cultural heriage, e-learning, e-government. web based model, dialog model, semantic analysis framework, interaction framework, humorist expert, network management, adaptive modular architecture as well. In this case they are not providing useful services but also interact with customer and give solution of their problems through AIML chatbotinstead of human beings. So, this is popular day by day with entrepreneur and users to provide efficient service.

Paper ID:87

Character Recogntion System: Performance Comparison of Neural Networks and Genetic Algorithm

Md. Shahazan Ali[1], Md. Nazrul Islam Mondal[2]
Dept. of Computer Science & Engineering
Rajshahi University of Engineering Technology, Rajshahi, Bangladesh
[1]kernel.writer@gmail.com, [2]nimbd@yahoo.com

Abstract: This thesis underlying the thesis title "Character Recognition System: Performance Comparison of Artificial Neural Networks and Genetic Algorithm" represents the character recognition using the Artificial Neural Networks (ANN) and Genetic Algorithm (GA) and measurement of variousperformance by changing various selection criteria of both algorithm. We used Backpropagation Learning Neural Network Algorithm (BPN) as the ANN. This system has been taken the character's image as its input. The input images have been filtered by filtering methods of image processing to remove noise and smoothing it and converted to binary image to detect its edges properly and clipped to get the actual image to input to the system. Features of each individual clipped images have been extracted by taking a definite resize binary image value. These extracted features of input images of characters is used by the Backpropagation Learning Neural Network Algorithm and Genetic Algorithm. Thus the network has been trained and creates a knowledge base of recognition. The same procedures have been applied for recognition but with only difference is that the neural network is used the previously learned weights and thresholds to calculate the output for BPN. In this work the performances of changing the various selection criteria of both algorithms have been also measured to learn and recognize the character.

Paper ID:89

Histogram based Water Quality Assessment in Satellite Images

Mumu Aktar[1], Md. Al Mamun[2], Md. Shamimur Rahman Shuvo[3], Md. Ali Hossain[4]
Computer Science & Engineering Department
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
[1]mumu.ruet@gmail.com, [2]cse_mamun@yahoo.com, [3]shuvo1364@gmail.com, [4]ali.ruet03@yahoo.com

Abstract: Water is one of the most precious resources of our environment. This water body is often faced the quality questions because of being polluted by ammonia, chemical wastes, sulfur dioxide from power plants, fertilizers containing nutrients--nitrates and phosphates , sediment, phytoplankton, etc. So, it is very necessary to assess quality of different water bodies. In this study, satellite images have been used for water quality measurement using histogram comparisons. A satellite image has been chosen to use as the original image whose water body has been considered as a standard of clear water body. Then this clear water body has been separated from other features using clustering at first to be used as standard ones and their number of pixels in percentage has been counted. Those images which contain the same percent of water bodies as the perspective standard ones can be verified by comparing their histograms. Euclidean distance has been measured between the standard and tested image's histograms. A tolerance level has been taken to assess the water quality as excellent, better, good, bad and poor. Finally if the distance falls within the tolerance level the water body can be categorized as excellent, better, good, bad or poor based on their degree of purity.

Paper ID: 91

Cross-correlation Based Approach of Underwater Network Size Estimation with Unequal Sensor Separation

D. K. Mondal[1], S. A. H. Chowdhury[2], Q. N. Ahmed[3], M. S. Anower[4]
[1,2,3] Dept. of Electronics & Telecommunication Engineering
[4]Dept. of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
[1]dheeman.ete10.ruet@gmail.com, [2]arif.1968.ruet@gmail.com,
[3]nusabaahmedaurthy@gmail.com, [4] md.shamimanower@yahoo.com

Abstract: It is very difficult to estimate the number and location of the operational nodes in underwater network using conventional protocol based techniques due to long propagation delay, strong background noise, non-negligible capture effect, high absorption and dispersion of underwater environment. So, a unique approach based on cross-correlation is applied for underwater network size estimation. In three-sensor case of this approach, network size is obtained by cross-corelating the Gaussian signals received at equidistant pair of sensors. The assumption of equidistant sensors poses some limitations and requires further investigation. The aim of this work is to obtain a more flexible estimation process by removing this assumption. For this purpose, three-sensor case of cross-correlation based underwater network size estimation technique is investigated in this paper with unequal sensor separation.

Paper ID: 92

Introducing Spatial Orientation Sensitive Cellphone Messaging System for Blind People-Tracking A Blind People in Emergency Situation

Parijat Prashun Purohit[1], Fardina Fathmiul Alam[2] , S. M. Mostaq Hossain[3]
Department of Computer Science & Engineering
[1,2] International University of Business Agriculture &Technology, Dhaka, Bangladesh
[3] Chittagong University of Engineering & Technology, Chittagong 4349, Bangladesh
[1]parijat2009@gmail.com, [2]fardina.mist.cse09@gmail.com, [3]mostaqcse10@gmail.com[3]

Abstract: It could be difficult and challenging for any person (especially blind people) to inform in an unusual situation for instant help. Falling in any critical situation for blind people is very common. One way to assist them in their unexpected situation by using an app of cell phone is proposed and described in this paper. Every smartphone consists of a good number of sensors. We have utilized accelerometer sensor's event and have introduced spatial orientation sensitive messaging system. According to our proposed system,A blind people will shake the cell phone in a specific axis so that message from cellphone will go to nearest helping hand destination. Here we have fixed a threshold and time limit so that we can avoid false messages for usual cellphone movements. Our last effort was to track the person using geodata of received message. Qualitative and quantitative performance of this system is evaluated. Moreover the entire messaging system is tested in various constraints and factors.

Paper ID: 106

A Real-Time Face to Camera Distance Measurement Algorithm Using Object Classification

Hossain[1], Md Asgar[1], Md Mukit[2]
[1]Electrical Communication University Engineering, University Kassel, 34109 Kassel, Germany
[2]Electrical and Electronic Engineering, World University of Bangladesh, Dkaka, Bangladesh
[1]Hossain.a.md@ieee.org, [2]smmukit@outlook.com

Abstract: In a human and computer interaction based system, distance estimation by computer vision between camera and human face is a vital operation. To calculate the distance between the camera and face, an estimation method based on feature detection is proposed in this paper, where detection of eyes, face and iris in an image sequence is described. From the estimated iris and the distance between the centroid of the iris, an algorithm is proposed to determine the distance from the camera to face. An architecture for face detection based system on AdaBoost algorithm using Haar feature and canny and Hugo Transform for edge and circular iris estimation is presented here. Wrongly detected faces are removed by analyzing the disparity map. From the estimated face, canny and Hugo transform is used to determine the iris. Later Pythagoras and similarity of triangles are used for distance estimation. The implementation is done in C++ using OpenCV image processing libraries to reduce system overhead.

Paper ID:115

Smartphone based Teacher-Student Interaction Enhancement System

Saad Bin Sajjad[1], Ahmed Imteaj[2]
Dept. of Computer Science and Engineering,
Chittagong University of Engg. & Tech., Chittagong-4349, Bangladesh
[1]saad2617@gmail.com, [2]imtu_1992@yahoo.com

Abstract: Smartphone applications have found their way into almost every avenue in our lives. This research aims to develop a mobile application based on teacher-student interaction system. As the efficiency and success of an academic class is somewhat dependent on the level and scope of teacher-student interaction in class, this paper emphasizes on how our mobile application can effectively contribute in increasing interaction and provide the necessary feedback to make the present and future classes better ones.

Paper ID:117

Blind Audio Watermarking Based on Fast Walsh-Hadamard Transform and LU Decomposition

Pranab Kumar Dhar
Department of Computer Science and Engineering,
Chittagong University of Engineering and Technology, Chittagong-4349, Bangladesh.
pranabdhar81@gmail.com

Abstract: Digital watermarking is identified as a major technique for copyright protection of multimedia contents. This paper presents an audio watermarking algorithm based on fast Walsh-Hadamard transform (FWHT) and LU decomposition (LUD). Firstly, we preprocess the watermark data to enhance the security of the proposed algorithm. Then, the original audio is segmented into non-overlapping frames and FWHT is applied to each frame. LUD is applied to the FWHT coefficients represented in a matrix form. Watermark data is embedded into the largest element of the upper triangular matrix obtained from the FWHT coefficients of each frame. Experimental results indicate that proposed algorithm is considerably robust and reliable against various attacks without degrading the quality of the watermarked audio. Moreover, it shows more excellent results than the state-of-the-art methods in terms of imperceptibility, robustness, and data payload.

Paper ID:119

Effects of Caffeine Doses on Cardiac Activity using Laser Doppler Flowmetry

Md. Bashir Uddin[1], M Nahid Reza[2], and Mohiuddin Ahmad1[3]
[1]Department of Biomedical Engineering
[2,3]Department of Electrical and Electronic Engineering,
Khulna University of Engineering & Technology, Khulna, Bangladesh,
[1]bas.u.kuet@gmail.com, [3]mohiuddin.ahmad@gmail.com

Abstract: In this work, the effects of different caffeine doses on cardiac activity have been evaluated by frequency domain analysis of laser Doppler flowmetry signal. Using different data acquisition units, the blood perfusion on human forearm middle finger tip has been recorded. Blood perfusion has been recorded efore and immediately after the consumption of all caffeine doses. Different caffeine doses contain different amount of caffeine but approximately same amount of sugar. Recorded and pre-processed data have been analyzed using frequency spectrum within the frequency range of cardiac activity. It is found that the consumption of caffeine free dose increases the cardiac activity. Besides, the consumption of caffeinated doses decreases the cardiac activity with respect to normal condition (before consumption). As the amount of caffeine consumed increases, the cardiac activity improves. Result shows 42.9%, 32.2% and 39.4% decrement in cardiac activity due to the consumption of 27 mg, 48 mg, and 64 mg caffeine respectively. The consumption of 80 mg caffeine dose significantly increases the cardiac activity. The outcome indicates the reduction of heart functions due to theconsumption of lower caffeine dose.

Paper ID:120

An Empirical Framework for Parsing Bangla Assertive, Interrogative and Imperative Sentences

Mohammed Safayet Arefin[1], Lamia Alam[2], Shyla Sharmin[3], Mohammed Moshiul Hoque[4]
Dept. of Computer Science & Engineering,
Chittagong University of Engineering & Technology, Chittagong-4349, Bangladeh
[1]safayet_arefin@yahoo.com, [4]moshiulh@yahoo.com

Abstract: To interpret language we need to determine a sentence structure. To do this we know the rule of how sentences of a language are organized and have an algorithm to analyze sentences given those rules. Parsing serves in language to combine the meaning of words and phrases. Parsing a sentence then involves finding a possible legal structure for sentence. This paper proposes a set of context-sensitive grammars (CSG's) to parse the Bangla sentences including assertive, interrogative and imperative. Experimental result reveals that the proposed framework can parse Bangla of sentences with over 80% accuracy.

Paper ID: 125

Paradigm Shift towards Cloud Computing for Banking Sector

Mohammad Anik Islam[1],Md. Khaled Ben Islam[2], Md. NazmusShakib Beg[3]
Department of Computer Science and Engineering
Pabna University of Science and Technology, Pabna, Bangladesh
[1]anik0054@gmail.com,[2]mdkhaledben@gmail.com,[3]shakibpstu@gmail.com,

Abstract: Cloud computing is one of the buzzwords of recent time. Cloud computing can be defined as a new style of computing in which dynamically scalable and often virtualized resources are provided as a service over the Internet. Unfortunately, though the Banking sector which works as a financial intermediary and one of the major industries for monetary activities is still deprived of the advantages of cloud. The major reason behind this is the lack of confidentiality and security. As a solution, we are proposing private cloud architecture for banking sector. So their confidentiality will be maintained alongside their data will be safe and sound. We are proposing a simple yet efficient load allocation algorithm to match the homogeneous nature of the banking tasks. We are also proposing a modular banking system. This solution is particularly helpful for developing countries, rural and remote places where setting banking infrastructure is not possible or feasible.

Paper ID:139

Closest Class Measure based Subspace Detection for Hyperspectral Image Classification

M. A. Hossain[1], S.U. Zaman[2], M. A. Mamun[3] , M. N. I. Mondal[4]
[1]University of New South Wales Canberra, Australia
[2,3,4] Rajshahi University of Engineering and Technology Rajshahi, Bangladesh
[1]ali.ruet03@yahoo.com

Abstract: The objective of this study is to develop a hybrid nonlinear subspace detection technique in which Kernel Principal Component Analysis (KPCA) is combined with a Closest Class Pair (CCP) measure for the task of hyperspectral image classification. In the proposed approach, KPCA is applied first to generate the new features from original dataset then a the CCP is applied to rank the features that are able to separate the complex or overlapping classes. Finally, the two ranked score such as KPCA and CCP is combined to select a subset of features that are relevant and able to better discrimination the input classes of interest. Experiments are performed on a real hyperspectral image acquired by the NASA Airborne Visible Infrared Imaging Spectrometer (AVIRIS) sensor and it can be seen that the proposed approach obtained the best classification accuracy 84.58%.

Paper ID:143

Multi-temporal FFT Regression

Md. Al Mamun, Md. Nazrul Islam Mondal, Boshir Ahmed, Md. Shahid Uz Zaman, Shyla Afroge
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh
{cse_mamun, nimbd, boshir_bd, szaman22, shyla_cse05}@yahoo.com

Abstract: Sequential data transmission regarding multi-temporal image analysis is mainly dependent upon prediction or forecasting. The transmission time can be substantially reduced by properly exploiting the temporal correlation. Multi-temporal images are often affected by sensor, and illumination variations, non-uniform attenuation, atmospheric absorption and other environmental effects which render system changes in them. Most of these changes are gradual and incremental. So a recent image can be predicted from a previous sequence of images if the amount of real land-cover change is limited. Regression based prediction is the most appropriate one in this case as it can quantify the relationships between images obtained by different measurement systems in different environments. FFT regression based temporal prediction is proposed in this paper whereby the least-squares minimization is conducted on the amplitude matrices of the readings via the FFT. For a given model, the value of squared coefficient of determination (R) is always increased beyond the value obtained by conventional regression which is a common quality measure of the chosen model.

Paper ID:145

Land Cover Classification for Satellite Images based on Normalization Technique and Artificial Neural Network

Boshir Ahmed[1], Md. Abdullah Al Noman[2]
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
[1]boshir_bd@yahoo.com, [2]noman093039@gmail.com

Abstract: The Satellite images and the extracted thematic maps provide higher-level information for the recognize, monitoring and management of natural resources. It is very difficult to identify land cover classification manually from a satellite image. The remotely-sensed images are invaluable sources of information for various investigations since they provide spatial and temporal information about the nature of earth surface materials and objects. This study aims to determine the level of contributions of multi-temporal and multi-sensor data together with their principal components for Artificial Neural Network classifiers. The suitability of Back Propagation Neural Network (BPNN) for classification of remote sensing images is explored in this paper. Automatic image classi fi cation is one of the challenging problems of recent year. BPN is self-adaptive dynamic system which is widely connected with the large amount of neurons. It can solve the regular problem arise from remote sensing images. This paper discusses about the BPNN method to improve the high resolution remote sensing image. The principle and learning algorithm of BPNN is analyzed and high resolution imagery of Beijing has been used. Back Propagation Neural Network classifies the remote sensing image into the classified image of their pattern recognition.

Paper ID:149

A Novel Idea of Tackling Hard Bit-vector Problems in a Beneficial Way by Eager Solver

Chowdhury Sajadul Islam[1], Md. Nahid Hasan[2]
[1]Department of Computer Science & Engineering, Independent University, Banglasesh
[2]Department of Electrical and Electronic Engineering
Rajshahi University of Engineering & Technology, Rajshahi-6204, Bangladesh
[1]chow.cth@gmail.com, [2]nahidxy@gmail.com

Abstract: The theory of fixed-width bit-vectors is a sequence of zero or one bits. The bit-vector theory offers a natural way of encoding the semantics of operations that manipulate binary data, the building block almost all modern computer systems. The predominant approach to deciding the bit-vector theory is via eager reduction to our propositional logic. Bit-vector satisfiability can be reduced to SAT by replacing function symbols by their hardware circuit representation expressed in propositional logic. An eager approach that takes full advantage of the solving power of off the shelf propositional logic solvers. In this paper focuses on the features of the eager bit-vector solver cvcE implemented in the satisfiability modulo theories (SMT) solver CVC4. The goal of paper explores to efficiently solving bit-vector constraints where eager solver cvcE is now an essential option for tackling hard bit-vector problems and evaluated their performance. We proposed a new technique of reducing the size of the bit-blasted formula by factoring out isomorphic sub-circuits to solve the constraints and enhance performance.

Paper ID:153

Phishing Attack Detection Using Taxonomy Model

Chowdhury Sajadul Islam
Department of Computer Science & Engineering
Uttara University, Dhaka, Bangladesh
chow.cth@gmail.com

Abstract: The objective of this paper is to detect phishing attacks using proposed phishing threat taxonomy model. To detect current phishing attack we collected the information by applying diverse methodology and constructing an intellectual data set. We have done the experiment on the basis of data sets by our proposed equations which produced results on the predicted rate of a phishing threat parameters (i.e. method, origin, component and target) in respect of predicted number of threats. From the numerical results it is apparent that 16.07% phishers use spare phishing 14.85%, fake web 12.2%, link manipulation and pharming attack techniques which are the highest rate of attack in subtotal 53.9% and rest of the threats use 44.61%. The results suggest while the accuracy of taxonomy model is a critical factor, reducing the success rates of phishers require other considerations such as improving tools, update dataset instantly and enhancing users' knowledge of phishing. Finally, we got a phishing threat taxonomy model from the collected data set, proposed equation, numerical results and graphical representation which is the main goal or novelty of this paper. Given the prevalence of phishing-based web fraud, the findings have important implications for individual and enterprise security. This is to provide a clear vision of the challenges that should be worked onto ensure next generation phishing security for cyber world.

Paper ID:159

An Empirical Analysis of Attribute Skewness over Class Imbalance on Probabilistic Neural Network and Naive Bayes Classifier

MD. Nazmul Shahadat[1], Biprodip Pal[2]
Department of Computer Science and Engineering
Rajshahi University of Engineering and Technology, Bangladesh
[1]nazmul.ruet@gmail.com, [2]biprodip.cse@gmail.com

Abstract: Many real world data are subjected to skewness or imbalance. Often class distribution is imbalanced, while several attribute or feature skewness is also frequent. Skewness affectsthe classification of the dataset samples. While classskewnessbiases the classification towards majority classes, skewed features may also bias the classification as they are significant for few classes. The purpose of this paper is to find out the impact of skewed feature variation in the training dataset for the Naïve Bayesian Classifier (NBC) and Probabilistic Neural Network (PNN). The experiment was carried out on six KEEL dataset which are skewed in terms of class distribution. This work looked for skewed features in those dataset and analysed the classification performance with and without the skewed features. The confusion matrix analysis illustrates that NBC is performing well compared to PNN.

Paper ID:160

Cooperative Game Theory Based Load Balancing in Long Term Evolution Network

Subarno Saha[1], Rajkin Hossain[2], Md. Muhidul Islam Khan[3]
[1]Dept. of Electrical and Electronic Engineering,
[2,3]Dept. of Computer Science and Engineering,
Brac University, Dhaka, Bangladesh.
[1]subarno.saha@gmail.com, [2] rajkin69@gmail.com, [3]muhidul.khan@bracu.ac.bd

Abstract: Long term evolution (LTE) network, incompatible with 2G and 3G networks is the most promising technology for wireless communication with higher speed and capacity. Self-organized load balancing is an important research issue for the wireless networks. Game theory provides an efficient way to provide self-organizing properties in a distributed environment like LTE networks. Load balancing means to assign users from highly loaded cells to neighbor lower loaded cells. The amountof load needs to be offloaded or accepted by a particular cellis not really specified and currently totally vendor specified. Inour proposed cooperative game theoretic approach, each cell is considered as a player where they trade the load by forming a coalition by satisfying the overall performance of the network. Simulation results show that our proposed method provides better performance in terms of satisfied users and adjusted load values.

Paper ID:162

Bangladeshi Road Sign Detection Based on YCbCr Color Model and DtBs Vector

Soumen Chakraborty[1], Kaushik Deb[2]
Dept. of CSE, Chittagong University of Engineering & Technology, Chittagong, Bangladesh
[1]soumen0123@yahoo.com, [2]debkaushik99@cuet.ac.bd

Abstract: Detection of road sign from a road image is very crucial for intelligent transportation system for awareness of drivers and blind pedestrians. In this regard, a framework has been proposed in this paper for detection of the road sign from road images. For detection of the road sign, two natural properties of a Bangladeshi road sign is utilized, they are – border color rim of the road sign and the shape of the road sign. Based on these ideas initially, YCbCr color model is used to eliminate the illumination sensitiveness for segmentation. In the second step, statistical threshold value is used for color segmentation. Next, labeling and filtering is used to extract the shapes. Finally, Distance to Borders (DtBs) vector is used to verify the region of interest (ROI) for detecting the Bangladeshi road sign (BRS). Various road images are used with a variety of conditions to test the proposed framework and results are presented to prove its efficiency.

Paper ID:163

GIS-based Extraction of Open Space in Rajshahi City

Fatima Jahan Sarmin[1], Md. Shahid Uz Zaman[2]
Department of Computer Science and Engineering,
Rajshahi University of Engineering and Technology, Rajshahi, Bangladesh
[1]sarmin_ruet_cse@yahoo.com, [2]szaman22@yahoo.com

Abstract: In this paper we present a way to extract features and information from a digitize GIS-based map and for obtaining the open space inRajshahi city. As we know the environment plays a most vital role for our living, the unplanned use of land can be a great threat to our environment and life. In this respect Geographic information system (GIS) gives us a way to keep track of land use through a digital map. We can capture, digitize, and store information and data in a GIS map. This paper aims to find out the useful data from the GIS based digitize map and analyze the information. After extracting the information a yearly based analysis is made and shows how terribly the number of building in Rajshahi area is increasing day by day since 1931 to 2000. And finally it is tried to make a forecast about the increasing the effect of it.

ABOUT ICCIE 2017

The International Conference on Computer & Information Engineering (ICCIE) organized by the Department of Computer Science & Engineering of Rajshahi University of Engineering & Technology is an attempt to expose the recent advancement and innovation throughout the vast area of computer and information science. It is expected to be an intellectual platform to share ideas and present the latest findings and experiences in the mentioned areas. World's leading creative educators, researchers, consultants, training managers, policy makers, curriculum developers, entrepreneurs, and others in Computer Science Education are likely to participate in the spark discussion. Thus the 2nd International Conference on Computer & Information Engineering (ICCIE 2017) will be organized during 23 – 24 November 2017, to focus the rapid development and technology advancement in computer and information science.

COMMITTEES

Organizing Chair
Head, Dept. of CSE, RUET

Organizing Secretary
To be Announced

Technical Chair
Ex-Head, Dept. of CSE, RUET

Technical Secretary
To be Announced

SCOPE

ICCIE 2017 invites papers from all of the Computer and Information Engineering areas. The scopes of the conference include but not limited to:

- Algorithms
- Artificial Intelligence & Expert System
- Bio-informatics
- Biometrics
- Cloud Computing
- Combinatorial Optimization
- Computer Architecture
- Cyber Crime & Ethics
- Computer Forensics
- Computer Graphics & Animations
- Data Mining & Warehousing
- Distributed and Parallel Computation
- Digital Signal and Image Processing
- Evolutionary Computing
- Geographical Information System
- Human-computer interaction
- Internet and Web Applications
- Mobile Computing
- Natural language processing
- Network Security and Cryptography
- Neural Network &Fuzzy Logic
- Pattern recognition & Machine Learning
- Robotics & Computer Vision
- Social network analysis
- Software Engineering
- VLSI
- Wireless Communications

PAPER SUBMISSION

The ICCIE MS Word or Latex template will be available at the conference website (http://cse-events.ruet.ac.bd/iccie/ Full Paper (PDF format not exceeding four A4 size pages) prepared in the prescribed format will be submitted using the online submission system available at conference website. The manuscript will be accepted on the basis of peer review without any provision for revision.

PUBLICATIONS

The proceedings of the conference will be published in printed book (abstract) and CD which are available to the delegates during the conference. All accepted and presented papers will be submitted to IEEE X-plore Digital Library.

CONTACT ADDRESS

Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi 6204, Bangladesh.

Tel.	:	+880721-750838
Fax.	:	+880721-750742-3 Ext. 503
		+880721-750838
Email	:	iccie@ruet.ac.bd
		iccie.cse.ruet@gmail.com
Web	:	http://cse.ruet.ac.bd/iccie

TECHNICAL CO-SPONSOR

IEEE
BANGLADESH SECTION

SPONSOR

CISCO
Networking Academy, RUET

CO-ORGANIZER

ICT DIVISION
Mistry of poste Telecommunication and Information Technology

Bangladesh Computer Council

ISBN NUMBER

978-984-33-9844-4

Issues in Entry Creation for an Educational Institution and Searching in Semantic Web

Syeda Nyma Ferdous
Department of Computer Science and Engineering
MIST, Dhaka, Bangladesh
nyma.ferdous@gmail.com

Jannatut Tabassum
Department of Computer Science and Engineering
MIST, Dhaka, Bangladesh
sweetytabassum@yahoo.com

Muhammad Masroor Ali
Department of Computer Science and Engineering
BUET, Dhaka, Bangladesh
mmasroorali@cse.buet.ac.bd

Abdullah Al Noman
Department of Computer Science and Engineering
MIST, Dhaka, Bangladesh
noman14032@gmail.com

***Abstract*—The main purpose of semantic web is driving the evolution of current web by enabling users to find, share and combine information more easily. In traditional web, it is both tedious and time-consuming to find our required information. But with the development of semantic web, it gets much more efficient and faster to obtain required information ignoring all other unnecessary data. Semantic web understands the meaning of data and can make intelligent decision based on available data. To build up the necessary web information and their relationship or the ontology for the semantic web, creation of valid RDF document is needed. These documents can be deployed and retrieved by the search engine after URL submission and completion of indexation. This paper describes the process of our RDF document creation for an educational institution, deployment, URL submission for indexation in the search engine, Swoogle and retrieval of the document. It emphasizes on identification of the issues or challenges throughout the whole process.**

***Keywords*—Semantic web; Semantic web annotation; search engine; Swoogle; ontology**

I. Introduction

Semantic web [1], [2], the meaningful web, is the movement to reach the goal of more easily processed web than the current web by the World Wide Web Consortium. The word *semantic* stands for meaning as well as the study of the significant meaning. So, semantic web is the extension of current web which allows the machine to understand the data on the web. It is done by adding metadata in the documents. The main purpose of semantic web is driving the evolution of current web by enabling users (including machines) to find, share, and combine information more easily.

The prerequisite for the development of semantic web are semantics, metadata and ontologies [1]–[4]. Ontology [3]–[6] is the exact description of web information and relationship among them. To create and implement this ontology, the basic building block is RDF [1]–[4], [7], [8]. OWL is also a part of the semantic web vision. The W3C Web Ontology Language (OWL) is a semantic web language designed to represent knowledge about things, groups of things, and relations between things [5], [6]. Both RDF and OWL are written in XML. XML is the basic syntax for these and RDF holds the semantics. XML Schema, XML Namespace, RDF Schema [1]–[4] are also associated with semantic web.

A web definitely needs a search engine. Semantic web search engines crawl the web of linked data by following their RDF links, and prepare their indexations accordingly. While there are many search engines for traditional web, there are also several search engines for semantic web for human eyes and applications, including Falcon, Sindice, SWSE and Swoogle [1], [9]. We have used Swoogle as the semantic web search engine in our case study.

In our case study, we created an entry for searching in the semantic web for an educational institution. RDF documents for an educational institution have been created using RDF creation tool. Then the documents have been deployed in the web and the corresponding URL has been submitted in the search engine for indexation. After indexation is complete, documents can be retrieved by any semantic web search engine. In our case study, we tested it with Swoogle.

Contributions of our paper are, identifying the issues in creation and deployment of a semantic web page and subsequent detection of the same page in semantic web search engine. We have also pointed out the design and technical challenges one might face when creating a semantic web page and how to overcome them. The whole gamut of validations presented along with their technical significance will also be beneficial to a prospective implementer.

Organization of rest of the paper is follows. Section II discusses the related technologies and standards for semantic web as well as semantic web search engine. Section III discusses the actual implementation carried out by us. Section IV presents the issues identified while we carried out our implementation and study. Finally, the paper ends in section V which puts the concluding remarks as well presents some open problems and suggestions on the creation and deployment of semantic web pages.

II. Semantic Web World

While an exhaustive description of semantic web is out of the scope of this paper, we present below a brief introduction of the related technologies and standards as are required for successful implementation of a semantic web site.

A. RDF

RDF [1]–[4], [8] stands for Resource Development Framework and it is a standard recommended by W3C for describing information or metadata of web resources. RDF expresses data using statements expressed as triples. Each statement is composed of a subject, a predicate, and an object. RDF imposes formal structure on XML to support the consistent representation of semantics [7], [10]. Dublin Core is also associated with this structure. Dublin Core is a set of predefined properties for describing documents [11]. RDF uniquely identifies property-types by using the XML namespace mechanism.

B. Ontology and OWL

Ontologies [3]–[6] are considered as one of the pillars of semantic web. An ontology is a formal specification of a shared conceptualization. Ontologies are implemented using OWL (Web Ontology Language) [3]–[6]. OWL allows users to write explicit, formal conceptualizations of domain models. The main characteristics are, a well-defined syntax, efficient reasoning support, a formal semantics, sufficient expressive power and convenience of expression [3].

C. Semantic Web Search Engine, Swoogle

As we have mentioned before, we have used Swoogle [1], [9] as the semantic web search engine in our case study.

Swoogle's architecture can be broken into four major components: SWD discovery, metadata creation, data analysis and interface [9]. This architecture is data centric and extensible.

In Swoogle, a general user can query with keywords, and the documents that match those keywords are returned in ranked order. For advanced users, an advanced search interface is provided. This allows them to fill in the constraints to a general SQL query on the underlying database.

III. Implementation of Semantic Web Technology

A. Generation of RDF Page

In our case study, to implement semantic web technology, at first, we tried machine generation of RDF documents. We tried the most well known and well used metadata element set, Dublin Core (DC) [12]. DC Metadata tool can be found at `http://www.ukoln.ac.uk/metadata/dcdot/`. We submitted the page `http://www.mist.ac.bd` to this tool. The tool read the page and generated the corresponding RDF.

To make our resource more meaningful, flexible and at the same time machine understandable, we created our own RDF file. We used Apache Jena [13], a free, open-source Java platform for applications on semantic web for this purpose. Jena is now believed to be the most used Java toolkit for building applications on semantic web.

In our specified RDF file, there are several elements that's worth some explanation. In our RDF file, in addition to the standard namespace, we used two more namespaces — `home` and `dept`. We used several elements such as `objective`, `courses`, `email`, `address` etc. The document is about an educational institute — its departments, location, contact information etc.

B. Deployment of RDF Document on Web

In order to deploy our created web page, we needed to buy a domain name. For this purpose, we used the free domain service at `http://www.000webhost.com` and registered a domain for our page. We uploaded the aforementioned RDF file to this site.

We verified whether our upload was technically successful or not. We tested it with HttpRequester and then we discovered our document with Semantic Radar (discussed further later on).

To elaborate further on the above mentioned issue, as noted in [14] the following two are very important steps in deployment of the RDF Document on Web.

1) *Dereferencing HTTP URIs:* URI Dereferencing is the process of looking up a URI on the Web in order to get information about the referenced resource.
   The information resources can be dereferenced by generating a new representation, a new snapshot of the information resource's current state, and sending it back to the client.
   This issue has been discussed further in section III-D.
2) *Content Negotiation:* HTML browsers usually display RDF representations as raw RDF code, or simply download them as RDF files without displaying them. Therefore, it is necessary to serve a proper HTML representation in addition to the RDF representation. This can be achieved using an HTTP mechanism called content negotiation.

C. Indexing URL in Semantic Search Engine and Retrieval of Document

After deployment, URL of the website that links these documents needed to be indexed in a semantic web search engine. After a certain period of days, the search engine was able to find the documents.

Swoogle employs a term frequency-inverse document frequency (TF/IDF) model with a standard cosine similarity metric. It indexes discovered documents by using either character N-Gram or URIrefs as keywords to find relevant documents and to compute the similarity among a set of documents. A number of algorithms, namely, Optimized Linear Instance Retrieval, Binary Instance Retrieval, Dependency Based Instance Retrieval, Static Index based Instance Retrieval, Dynamic Index based Instance Retrieval are used in this process [9], [15].

We tested whether the RDF creation and search engine indexing of our document was successful. We tried to retrieve the RDF documents by searching in Swoogle. The indexing was successful and it was able to find the RDF.

```
GET http://nymaferdous.host56.com/misthome.rdf
```

(a) Command issued.

```
-- response --
200 OK
Date:  Tue, 20 Jan 2015 13:45:14 GMT
Server:  Apache
Last-Modified:  Tue, 20 Jan 2015 13:44:43 GMT
Accept-Ranges:  bytes
Content-Length:  1381
Connection:  close
Content-Type:  application/rdf+xml
```

(b) Successful result returned. First line is the status line while the rest are response headers.

Fig. 1: Command issued in and result returned by HttpRequester.

D. W3C Markup Validation

In order to check whether our generated file is as per W3C standards and also to check the triples, we had our file parsed at `http://www.w3.org/RDF/Validator/`.

E. Other Validations on Information Retrieval from the Created Page

In order to further test our created semantic web page, we tried using several tools.

1) HttpRequester: HttpRequester [16] is used for making HTTP requests like GET, POST and viewing the responses. We tested our RDF submission with HttpRequester. It successfully recognized our RDF document. See Figure 1 for details.

2) RDF Auto-discovery: RDF auto discovery [17] implies extracting metadata from a site. For RDF auto discovery, we have used `<link>` tag in the `<head>` section of our page. Use of `<link rel="alternate" type="application/rdf+xml" title="RSS 1.0" href=`*`site address`*`>` enables the semantic agents to detect our underlying RDF file. This link is used to let the semantic agents know that machine readable version is associated with the site. Users don't have to exercise any additional effort for this process.

3) Semantic Radar: We also discovered our document with Semantic Radar [17], [18]. It discovers the existence of RDF documents in a site. If any RDF document is present in the site, semantic radar shows the URL and displays a status bar icon to indicate presence of Semantic Web (RDF) data in the web page. This also allows one to explore information in more detail. We entered our site and clicked on the semantic radar. It indicated a successful presence of our RDF document.

4) Vapour: We validated our document with Vapour [19], [20], a semantic linked data validator. Vapour is a validation service which is used for checking whether the semantic data is published on the web correctly or not. Vapour checks this based on two test requirements: i) dereferencing URI (requesting XML/RDF) and ii) dereferencing URI (without content negotiation). It also checks for meaningful triples in the response documents, links to popular semantic web browsers, and conclusions on the type of the resources. A document will be deemed to be published correctly if it shows the test results "Passed". In case of our document, we submitted the URL in Vapour and it showed the result "Passed". It also marked our document as an information resource.

5) DBpedia: Semantic crawlers can crawl different interlink connections. So, we have linked our document to DBpedia [21], [22]. In this way, search engines can find our document easily.

6) Sindice Inspector tool: We found our document with Sindice Inspector tool [23], [24]. Sindice is a lookup index for semantic web documents [25]. Sindice indexes the semantic web and can tell one which sources mention a resource URI, IFP, or keyword. We used Sindice on our page to find relevant RDF sources and it was successful.

IV. Implementation Issues and Challenges

In order to successfully implement a semantic web page and make it available in search engines, there are some specific technical issues one needs to take care of. We elaborate on these issues and challenges in this section.

A. Design Issue

We have avoided blank nodes when designing the RDF file. Blank nodes are nodes that have no name. RDF parsers locally generate a name for these kinds of nodes. Blank nodes make it difficult to link and merge data from different sources. We have also avoided `<rdf:Bag>` container for this purpose. See [26] for details.

B. RDF Deployment Issue

Our generated RDF documents have to be uploaded to a website. This website then has to be deployed on the web. We have to be careful about the server so that it recognizes xml+rdf MIME types.

C. Document Indexation on Search Engine

Generating and deploying RDF documents on the web is not enough. We have to make sure that the created RDF is indeed correct and can be found when searched. For this, we submitted our website link to Swoogle for indexing. But Swoogle can not index a document just after the URL of that document has been submitted. For Swoogle to index the document and create an entry on its index table, it takes some time. So, one needs to wait some time to be sure, whether the effort to find a RDF document when searched, is indeed successful. We can find our semantic web document by searching keyword misthome or url:"http://nymaferdous.host56.com/misthome.rdf".

V. Discussion and Conclusion

A. Discussion

Current web search engines such as Google and all the web are not very suitable for documents encoded in the semantic web languages. These systems have been developed to work with natural languages and expect documents to contain unstructured text composed of words. They are rather inefficient in scanning semantic web documents. They fail to understand conventions such as those involving namespace and similar entities. Moreover, they do not understand the structural information encoded in the documents and are thus unable to take advantage of it.

Our created semantic web page enables the semantic web search engines search semantically. This enables the web crawlers to understand the semantics of these data, integrate and correlate different activities to automate the process to make it more effective and efficient.

We have presented the issues one will face when implementing a semantic web page and make it available in search engines. The issues have been discussed in the context of a complete design, development, implementation, retrieval and validation of a semantic web page built from scratch. The context of an educational institution makes it particularly relevant to the educators and researchers.

B. Open Problems/Future Works

In addition to our identified issues, we think that solving the following open problems will facilitate further creation and deployment of semantic web pages.

Automatic annotation of semantic web pages: With the massive volume of information available now-a-days, manual annotation is prohibitively expensive and time-consuming. So, automatic annotation of web pages [27], [28], preferably in an unsupervised way, is essential.

Creation of multilingual semantic web pages: Many web pages like news items, encyclopedia are available in multiple languages and resource construction in several languages [29], [30] by integrating lexicographic and encyclopedic knowledge plus cross-lingual ontology mapping is necessary.

Quality assurance of created pages: With the creation of semantic web pages, good quality ontology, metadata and above all, correct annotations [31], [32] are very important.

References

[1] L. Yu, *A developer's guide to the semantic web*. Springer, 2011.

[2] "Semantic web," http://www.w3.org/standards/semanticweb/, (last accessed Feb 17 2015).

[3] G. Antoniou, P. Groth, F. van Harmelen, and R. Hoekstra, *A Semantic Web Primer*, 3rd ed. The MIT Press, 2012.

[4] P. Hitzler, M. Krötzsch, and S. Rudolph, *Foundations of Semantic Web Technologies*. Chapman & Hall/CRC, 2009.

[5] P. Hitzler, M. Krotzsch, B. Parsia, P. F. Patel-Schneider, and S. Rudolph, "Owl 2 web ontology language primer," http://www.w3.org/TR/owl2-primer/, (last accessed Feb 17 2015).

[6] M. K. Smith, C. Welty, and D. L. McGuinness, "Owl web ontology language guide," http://www.w3.org/TR/2004/REC-owl-guide-20040210/, Feb. 2004, (last accessed Feb 17 2015).

[7] G. Schreiber and Y. Raimond, "RDF 1.1 primer," http://www.w3.org/TR/rdf11-primer/, (last accessed Feb 17 2015).

[8] "Resource description framework," http://www.w3.org/RDF, (last accessed Feb 17 2015).

[9] L. Ding, T. Finin, A. Joshi, R. Pan, R. S. Cost, Y. Peng, P. Reddivari, V. Doshi, and J. Sachs, "Swoogle: a search and metadata engine for the semantic web," in *Proceedings of the thirteenth ACM international conference on Information and knowledge management*. ACM, 2004, pp. 652–659.

[10] Z. A. Shaikh, N. Zaman, and Z. H. Khuhro, "Semantic web & its content creation process," *Journal of Information & Communication Technology*, vol. 3, no. 2, pp. 87–89, 2009.

[11] "Web services tutorial — RDF," http://www.w3schools.com/webservices, (last accessed Feb 17 2015).

[12] J. Kunze and T. Baker, "The Dublin Core metadata element set," RFC 5013, Aug. 2007.

[13] "Apache Jena," https://jena.apache.org/, (last accessed Feb 17 2015).

[14] C. Bizer, R. Cyganiak, T. Heath *et al.*, "How to publish linked data on the web," (last accessed Feb 18 2015).

[15] V. Haarslev and R. Möller, "Optimization techniques for retrieving resources described in OWL/RDF documents: First results," in *Ninth International Conference on the Principles of Knowledge Representation and Reasoning*, 2004, pp. 163–174.

[16] "HTTP requester," http://www.apycom.com/httpreq/, (last accessed Feb 17 2015).

[17] U. Bojars, A. Passant, F. Giasson, and J. Breslin, "An architecture to discover and query decentralized RDF data," in *Proc. the 2nd European Semantic Web Conference*, vol. 1, no. 1, 2007.

[18] U. Bojars and S. Fernandez, "Semantic radar 1.1," https://addons.mozilla.org/en-US/firefox/addon/semantic-radar/, (last accessed Feb 17 2015).

[19] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data-the story so far," *Special Issue on Linked Data, International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 5, no. 3, 2009.

[20] "Vapour linked data validator," http://www.w3.org/2001/sw/wiki/Vapour, (last accessed Feb 18 2015).

[21] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, *Dbpedia: A nucleus for a web of open data*. Springer, 2007.

[22] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann, "Dbpedia-a crystallization point for the web of data," *Web Semantics: science, services and agents on the world wide web*, vol. 7, no. 3, pp. 154–165, 2009.

[23] G. Tummarello, R. Delbru, and E. Oren, *Sindice. com: Weaving the open linked data*. Springer, 2007.

[24] E. Oren, R. Delbru, M. Catasta, R. Cyganiak, H. Stenzhorn, and G. Tummarello, "Sindice. com: a document-oriented lookup index for open linked data," *International Journal of Metadata, Semantics and Ontologies*, vol. 3, no. 1, pp. 37–52, 2008.

[25] "Sindice," http://www.w3.org/2001/sw/wiki/Sindice, (last accessed Feb 18 2015).

[26] "Issue-24: Should we deprecate RDF containers (alt, bag, seq)?" http://www.w3.org/2011/rdf-wg/track/issues/24, (last accessed Feb 20 2015).

[27] S. Mukherjee, G. Yang, and I. V. Ramakrishnan, "Automatic annotation of content-rich html documents: Structural and semantic analysis," in *The Semantic Web-ISWC 2003*. Springer, 2003, pp. 533–549.

[28] D. Sánchez, D. Isern, and M. Millan, "Content annotation for the semantic web: an automatic web-based approach," *Knowledge and Information Systems*, vol. 27, no. 3, pp. 393–418, 2011.

[29] R. Navigli and S. P. Ponzetto, "Babelnet: Building a very large multilingual semantic network," in *Proceedings of the 48th annual meeting of the association for computational linguistics*. Association for Computational Linguistics, 2010, pp. 216–225.

[30] B. Fu, R. Brennan, and D. O'sullivan, "Cross-lingual ontology mapping and its use on the multilingual semantic web." *MSW*, vol. 571, pp. 13–20, 2010.

[31] S. Currier, J. Barton, R. OBeirne, and B. Ryan, "Quality assurance for digital learning object repositories: issues for the metadata creation process," *Association for Learning Technology Journal*, vol. 12, no. 1, pp. 5–20, 2004.

[32] A. Zaveri, A. Maurino, and L.-B. Equille, "Web data quality: Current state and new challenges," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 10, no. 2, pp. 1–6, 2014.

A Vision Based Guide Robot System: Initiating Proactive Social Human Robot Interaction in Museum Scenarios

M. Golam Rashed[1], R. Suzuki[1], A. Lam[1], Y. Kobayashi[1,2], Y. Kuno[1]
[1] Saitama University, Japan, [2] JST, PRESTO, Japan
Email: [golamrashed, suzuryo, antonylam, yosinori, kuno]@cv.ics.saitama-u.ac.jp

***Abstract*—This paper presents a guide robot system which observes people's interest and intention towards paintings in museum scenarios and proactively offers guidance to them using a guide robot, if needed. To do that, we utilized multiple USB video camera sensors to support the guide robot in detecting and tracking people's visual focus of attention (VFOA) toward paintings. In this study, we consider each person's head orientation and profile information and compute importance values to identify a target-person that may be interested in a particular painting. After identifying the target-person, the guide robot moves autonomously through an appropriate motion path from the so called *public-distance* to his/her *social-distance* to explain details about the painting to which s/he is interested. Furthermore, we demonstrate the viability of our guide robot system by experimenting with the *Robovie-R3* as a museum guide robot. Finally, we tested the system to validate its effectiveness.**

***Keywords*—VFOA, gaze direction, guide robot.**

I. Introduction and Motivation

During the last decades, computer vision has offered the ability for service robots to understand human behavior. One of them is the visual focus of attention (VFOA) of a person, which is the behavioral and cognitive process that indicates where and at what a person is looking, and that can be determined by eye gaze and/or head pose dynamics [1]. Thus, VFOA is an important key to understand human behaviors. The VFOA of people can be used by service robots to estimate information such as their interests and intentions in relation to the environment [2]. If any defined VFOA related to their interests and intentions are found from their gaze direction, then service robots can serve them accordingly. For example, with such information in an art museum, a human museum guide could guess which paintings patrons consider more attractive. Such types of information can be very helpful for developing and adapting various services for people in public spaces where service robots can exclusively provide services to them. Although dealing with such types of situations for a service robot is quite difficult, the purpose of this research is to develop such a robotic system for serving people who are in need in a proactive manner as opposed to the conventional reactive approach where robots wait until people exclusively request them for their service.

The main research goal here is to find people that may want the service robot's help. To find such people, we estimate their VFOA from a computer vision point of view. We believe that we can find such information by detecting and tracking either their eye gaze, head orientation, face profile information, or some combination of them. In this work, we demonstrate our solution through a museum guide robot application. There are several guide robots that have been proposed (for example, [3]) that can provide a guided tour to people in a museum scenario. However, here we consider a guide robot system that can find a person who seems to be interested in a particular painting. In that case, the guide robot can proactively provide more information about the painting to the person to make his/her visit to a museum a more enjoyable.

When trying to find out where people are looking around in a museum, a first solution would be measure their VFOA which can be estimated by their gaze direction [4] because gaze direction typically follows the focus of his/her interest in the environment [5]. Todorovic et al. in [6] defined the gaze direction with respect to some references. There are several such references, and thus several different alternatives of the notion of gaze direction. Some of among them are *looker-related*, *observer-related* and *environment-related* gaze direction. Details about those people's gaze direction with respect to some references can be found in [4]. In this study, we consider people's *environment-related* gaze direction to estimate their interest and intention to the painting inside the museum because in a museum scenario a human guide will most likely not be concerned with where the people look with respect to themselves, but rather what s/he looks at in the environment. In our experiments, environmental low cost USB video camera sensors are used as such references in the museum.

Many studies (for example, [7]) have found that eye gaze tracking based information is not always feasible in obtaining people's VFOA due to several environmental constraints in museum like environments. For this reason, in this study, we propose a guide robot system in which the guide robot determines the general gaze direction of people towards paintings, but not the exact gaze point, using the head orientation information obtained from video information of the camera sensor. We believe that our guide robot system can perform robust real-time detection of people's interest and intention toward paintings using camera sensors in a museum scenario for providing proactive guidance to the museum's patrons. For any real museum environment, a large viewing volume is desired for any guide robot system so that people in it can be tracked as they move from painting to painting.

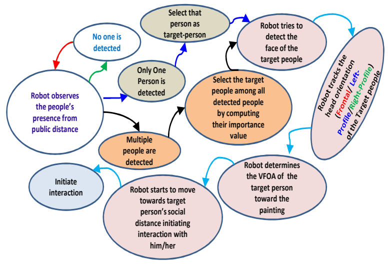


Fig. 1. States of our proposed Human-Robot Interaction System.

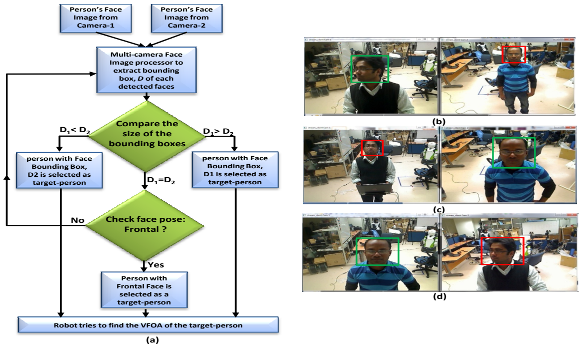


Fig. 2. (a) Flow diagram to obtain target-person using importance value. Two face trackers are shown to select target-person, though more could be possible. (b)-(c) People with greater green colored bounding box are treated as target-person, (d) Person with same bounding box size as other person but person with the frontal face is treated as the target-person.

Multiple camera sensors afford this ability by providing a large combined viewing volume by which people's attention can be tracked. The advantage of our proposed robotic system is that multiple camera sensors cover a much larger field of view to observe people than that covered by a single sensor in our museum scenario.

II. Our Approach

The states of our guide robot system is illustrated in Fig. 1. In our proposed approach, we consider a number of paintings (P_1, P_2, P_3, etc.) as exhibits and the *Robovie-R3* as the guide robot in a museum where paintings are hung on a wall at equally separated distances and the *Robovie-R3* is placed at a *public-distance* from the paintings and people's painting viewing regions[1]so that its presence may not interfere with each person's movements and attention. The objective of the guide robot is to provide more information to people when their interest towards any painting is detected by the system.

To provide more information about any particular painting to any person to which s/he seems to be interested, first the guide robot system tries to determine the presence of people within the museum's vicinity using USB video camera sensors placed just above the paintings. If no person is detected by the sensors, then the system will continue its task to detect people. If only one person is detected by any one of the sensors then the system will treat that person as the target-person. If multiple people are detected, then the system will compute the importance value (described in Sect. II-2) of each of the detected persons to select one as the target person. Thereafter, the system will obtain the head orientation information of the target person from the video frames and thereby extract gaze directional information to ultimately estimate his/her VFOA. The system will then trigger the guide robot to get the face profile information from the head orientation information to select the appropriate motion path to move from the *public-distance* to the *social-distance* of the target-person to proactively approach him/her and explain the painting.

1) Visual Detection and Tracking of People: Our employed people visual detection and tracking system is based on the 3D head tracking method presented in [8]. They used a Cascaded classifier based on the AdaBoost algorithm and Haar-like features for human head tracking. Originally, this idea was developed by Viola and Jones [9] to reliably detect faces without requiring a skin color model. This method works quickly and yields high detection rates [10]. Thus, with the 3D head tracking method, we can easily detect and track people in terms of tracking head orientation accurately in wider angular views in real time from video frames delivered by the USB video camera sensors where image sequences are captured at VGA resolution at a video frame rate about of 15 fps and processed by one PC (Intel(R) Core(TM) i5-2400 CPU @3.10GHz, Memory 4 GB).

2) Target-Person Selection Procedure: In our guide robot system, if only one person is detected by the video camera at any time then s/he is automatically selected as the target-person to track his/her VFOA. If more than one persons are detected by the cameras then our system computes their importance values individually. The importance value allows the system to choose the one target-person among all detected persons. In our present study, it currently depends on two parameters (**a**) the distance of the persons from the paintings, and (**b**) the face profile information. Here, the distance between the persons from the paintings is estimated by considering the size of the bounding box, D of their detected faces on the image planes. A person with his/her face closer to the paintings, will have a larger bounding box for his/her face and vice-versa. The procedure to compute the importance value of only two persons is illustrated in Fig. 2, though more could be possible. Anyone with a greater sized bounding box gets a higher importance value over others. Our system focuses its attention on the person who has the highest importance value. These scenarios are depicted in Fig. 2(b) and (c). When the bounding box of both of the persons are the same in size the person with the most frontal face profile will be considered as the target-person. This scenario is depicted in Fig. 2(d).

3) Recognition of Target Person's VFOA : Initially, our guide robot system classifies the target person's yaw angle of his/her head into three fundamental angular regions: Left Near Peripheral Field of View (LNPFV), Central Field of View (CFV), and Right Near Peripheral Field of View (RNPFV). The frontal view, which covers 30° (75° to 105°) is defined as the CFV. The right side angular region with respect to CFV, which covers about 45° (30° to 75°) is defined as the RNPFV. And the LNPFV is defined on the left side with respect to CFV, covers about 45°(105° to 150°). Fig. 3(a) illustrates these head orientation classifications with example face images when the camera is situated in the CFV angular region of any person. Once the target-person is selected, to measure the level

[1]The regions from where a person typically views exhibits in a museum.

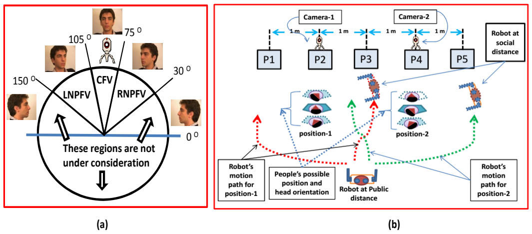


Fig. 3. (a) Head orientation classification into three angular regions in our system, (b) Robot's motion path and position planning scenario.

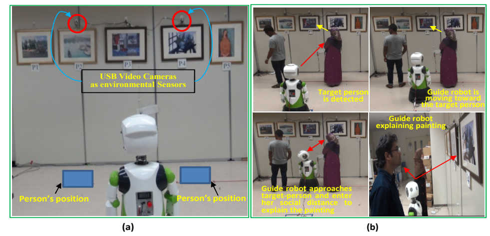


Fig. 4. (a) Our experimental environment, (b) Some snapshots of our experimental scenes.

of interest to any particular paintings, our system measures the head orientation stability by calculating the average head yaw angle for next the 30 consecutive video frames captured by the USB video camera sensor. If the target person's head orientation is stable within any particular angular region, then the guide robot will consider the orientation of his/her gaze direction toward that region and thereby measure the stable VFOA of the target-person toward any particular painting. For example, if the stable VFOA of the target-person is found in the CFV angular region, the guide robot system will detect only the frontal face of the target-person. However, in the other defined angular regions, the guide robot system may detect two face patterns which are either the $45°$ degree of right profile face or the $45°$ degree left profile face in the LNPFV and RNPFV angular region, respectively.

4) Guide Robot's Motion Path Planning: In this study, we restricted the possible motion paths for our guide robot, *Robovie-R3* to only that of two people's possible predefined positions. Our guide robot itself does not have any built-in artificial intelligence, but in the implementation of our guide robot's behaviors, we programmed it to perform various kinds of actions. For each position, we programmed two motion paths and positions for the *Robovie-R3* to reach the target person's *social-distance* from a *public-distance* and to initiate conversation. The schematic representation of the guide robot's motion paths for the people's two locations is depicted in Fig. 3(b).

For position-1, we define the left-side motion path and position for LNPFV region of camera-1, and the straight motion path and position for RNPFV and CFV region of camera-1. For position-2, we define the right-side motion path and position for RNPFV regions for camera-2, and the straight motion path and position for the LNPFV and CFV region of camera-2. Thus, we designed a total of four possible motion paths and positions. As soon as the head stability of the target-person is found to reside in any of the angular regions described in Sect. II-3, the guide robot starts to move from the *public-distance* through any of the preprogrammed motion paths to reach his/her *social-distance* to approach and explain that painting to which his/her VFOA is detected.

III. System Evaluation

We performed experiments to verify that our proposed system is useful in our guide robot scenario. In our experiment, we used 10 people (8 males, average age 29.2 years) from our campus. Each person participated in our experiments four times in a total of two phases. In the first phase, two experiments were conducted to verify the flexibility of using multiple video camera sensors over a single video camera sensor based robotic system to detect and track people's VFOA towards the paintings, and in the second phase, two experiments were conducted to verify the effectiveness of our proposed system using *Robovie-R3* as a guide robot to proactively initiate interaction with target people to offer a explanation about any particular painting. In every case, people were asked to observe the exhibits from possible predefined positions randomly.

5) Experiment Design: In our experiments, we hung five paintings (P_1, P_2, P_3, P_4, P_5), separated equidistant from their neighboring paintings and all at the same height. Two USB video cameras were mounted as environmental sensors on the top of paintings P_2 and P_4 to detect and track people's VFOA. Paintings P_1, P_2 were placed in the LNPFV, and CFV areas of camera-1, respectively whereas paintings P_4, P_5 were placed in the CFV and RNPFV areas of camera-2, respectively. Finally, painting P_3 was placed in between the RNPFV area of camera-1 and the LNPFV area of camera-2. The *Robovie-R3* was used as our guide robot and was initially placed at the *public-distance* from the people's possible positions (usually far from exhibits and behind person's positions). The experimental setup for the museum scenario is illustrated in Fig. 4(a). Finally, we designed a total of four possible motion paths and positions (as described in Sect. II-4) so that the guide robot could move to suitable positions for approaching a person. Fig. 4(b) shows some snapshots of our experiments. One video camera was also used in a appropriate position to document all experimental activities.

6) Experimental Cases: To validate the effective of our robotic system, we conducted two modes of experiments and compared them.

•*Multi-Sensor Mode (MSM) (Proposed mode).* In the current implementation of our system, we observe the performance of our system using two USB video cameras to select the target-person and track his/her VFOA toward the paintings while other people could also be viewing the paintings.

•*Single-Sensor Mode (SSM).* Under this system, we use only one USB video camera to detect the target-person and track his/her VFOA toward all the paintings while other people could also be viewing the paintings.

Additionally, we also conducted experiments with the following two modes within our proposed approach to measure the effectiveness of our guide robot in dealing with guiding the target-person.

•*Single-Person Mode (SPM).* Only one person observes the paintings randomly where vision data from the two USB

TABLE I. PEOPLE'S IMPRESSION FOR VARIOUS QUESTIONNAIRES

Questionnaire	Modes	Mean	SD	$F(1,9)$	$P-value$	η^2
Q1	*SSM* *MSM*	2 6	0.67 0.45	240	0.00000008	0.88
Q2	*SPM* *MPM*	4.4 6.3	1.16 0.9	19.23	0.001759	0.49
Q3	*SPM* *MPM*	4.2 6.0	3.067 2.22	12.35	0.00656	0.25

video cameras will be processed to detect him/her as the target-person and track his/her VFOA toward the paintings by guide robot.

•*Multi-Person Mode (MPM).* In this mode, multiple person will observe all the paintings randomly where the data of the two USB video cameras will be processed simultaneously to determine the target-person by the guide robot.

7) Measurements: In all of our experiments, we evaluate the following two cases:

•*People's Impression*: We asked participants to fill out a questionnaire for each mode after completion of experiments. The measurement was a simple rating on a Likert scale of 1 to 7. There were three items in that questionnaire: (**Q1**) Did the robotic system estimate your gaze directions (VFOA) while you were viewing all the paintings (P_1 to P_5)?, (**Q2**) Did the robotic system effectively detect and track your VFOA while you were viewing any of the paintings?, (**Q3**) Was the *Robovie-R3* effective in approaching you in a timely manner to start interaction with you?

•*Success Rate*: From the documented videos and experimental site, we observed how many times the guide robot detected the target-person's VFOA and established a successful interaction under our proposed approach. The success rate was measured by the ratio of the number of successful interactions to the total number of attempts that the guide robot made.

8) Results: The experiment was conducted in a within-participant design, and the order of all experimental trails was counterbalanced. We performed the repeated measures ANOVA for all measures.

a) People's Impression:: Subjective measures for all the modes are shown in Table I. The result shows the questionnaires measure which represents the means (M), and standard deviations (SD), F, $p-value$, and η^2 in each condition for the participants.

For **Q1** (Table I, 2nd row), the ANOVA analysis shows statistically significant differences between the modes (MSM and SSM) ($F(1,9) = 240, p < 0.01$, $\eta^2 = 0.88$) for the people. Thus, these results reveal that the performance of the multi-camera sensor based robotic system (our proposed approach) outperforms the single-camera sensor based robotic system in detecting people and tracking their VFOA during their viewing of all the paintings.

For **Q2** (Table I, 3rd row), the differences between the two modes were statistically significant ($F(1,9) = 19.23, p < .01$,$\eta^2 = 0.49$). Thus, these results mean that our proposed approach is more effective in detecting and tracking the target-person's VFOA when more than one person is detected by the guide robot.

For **Q3** (Table I, 4th row), significant differences between the two modes ($F(1,9) = 12.35, p < 0.01$, $\eta^2 = 0.25$) were found under our proposed robotic system. Most of the participants felt that our guide robot was more effective in the case of *MPM* over *SPM* to approach the target-person accurately to explain the painting to which s/he was interested.

b) Success Rate: 10 participants experienced a total of 48 trials as the target-person under the *MPM* settings. Among 48 trials, our robotic system was able to detect and track their VFOA and finally make successful interactions at a rate of 87.5%. Thus we can say that our proposed system is effective for initiating interaction with the target-person.

Thus, we experimental results revealed that our proposed robotic system is more effective for target-person detection from multiple people and tracking of his/her VFOA using multiple USB video cameras to obtain his/her interest and intention toward the painting. It is also revealed that our proposed robotic system is much more operative at proactively initiating interaction with the target-person with the autonomously movable guide robot in the museum scenario.

IV. CONCLUSION AND FUTURE WORK

We developed a vision based museum guide robot system that can detect the target-person and track his/her VFOA toward any paintings using his/her head orientation information and finally employed a guide robot to initiate proactive interaction with him/her to offer more explanation about the painting. In the future, we would like to employ laser sensor technology in addition to a vision system in our robotic system to know the interests and intentions of the people in museum-like environments more precisely, which will provide more information to establish effective social human robot interaction. These are left for our future research.

This work was supported by JST, CREST, JAPAN.

REFERENCES

[1] D. Das *et al.*, "Supporting human robot interaction based on the level of visual focus of attention," *Human-Machine Systems, IEEE Transactions on*, vol. PP, no. 99, pp. 1–12, 2015.

[2] S. O. Ba, "Joint head tracking and pose estimation for visual focus of attention recognition," Ph.D. dissertation, STI, 2007.

[3] T. Oyama *et al.*, "Tracking visitors with sensor poles for robot's museum guide tour," in *Human System Interaction (HSI), 2013 The 6th International Conference on*, June 2013, pp. 645–650.

[4] R. Stiefelhagen and J. Zhu, "Head orientation and gaze direction in meetings," in *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '02, 2002, pp. 858–859.

[5] T. Yokoyama *et al.*, "Attentional capture by change in direct gaze," *Journal of Perception*, vol. 40, pp. 785–797, 2011.

[6] D. Todorovic, "Geometrical basis of perception of gaze direction," *Vision Research*, vol. 46, no. 21, pp. 3549 – 3562, 2006.

[7] M. G. Cline, "The perception of where a person is looking," *The American Journal of Psychology*, vol. 80, no. 1, pp. 41–50, 1967.

[8] Y. Kobayashi *et al.*, "3d head tracking using the particle filter with cascaded classifiers," in *In: BMVC*, 2006.

[9] P. Viola and M. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[10] M. Bennewitz *et al.*, "Towards a humanoid museum guide robot that interacts with multiple persons." in *Humanoids*. IEEE, 2005, pp. 418–423.

Performance Comparison of Three Optimized Alternative Pulse Shaping Filters with the Raised Cosine Filter for Wireless Applications

T. K. Roy, M. F. Pervej
Dept. of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
roy_kanti@yahoo.com, ferdous.ete09@gmail.com

M. Morshed
Dept. of Information & Communication Technology
Mawlana Bhashani Science & Technology University
Tangail, Bangladesh
monirmorshed.mbstu@gmail.com

***Abstract*—This paper is concerned with the performance and cost analysis of different pulse shaping filters such as raised cosine, finite impulse response (FIR) Nyquist, FIR half-band and infinite impulse response (IIR) linear phase half-band for wireless applications. Note that the pulse shaping is the process of changing the waveform of transmitting pulses to reduce the interferences keeping a signal in an allotted bandwidth. This paper proposes an optimal design for the pulse shaping filters to reduce their effective costs in the area of wireless communication applications. Here, we analyze the performance of pulse shaping filters based on the minimum stop band attenuation required to suppress the inter-symbol interference (ISI) and inter-channel interference (ICI). Finally, we compare the performance of raised cosine, FIR Nyquist, FIR half-band and IIR linear phase half-band filters in terms of bit-error rate (BER) and hardware requirements. Our results demonstrate that the hardware implementation costs for the FIR Nyquist, FIR half-band and IIR half-band filters are about 60% to 90% less than the raised cosine filter.**

Keywords*—*FIR Nyquist filter; FIR half-band filter; IIR half-band filter; pulse shaping filter; raised cosine filter.

I. Introduction

Recently, the wireless application has grown much more rapidly than wire-line services. High quality and high-speed are the greatest demands in wireless communications. But one of the major difficulties in the way of errorless digital communication is ISI. There are several methods exist in the literature among these methods, pulse shaping filtering is one of the method which is used to solve this problem. Pulse shaping filters are widely used as the heart of many modern data transmission systems (e.g. mobile phones, high definition TV) to reduce the interferences by keeping a signal in an allotted bandwidth. When the pulses are sent by the transmitter, then the goal at the receiver is to sample the received signal at an optimal point in the pulse interval to maximize the probability of an accurate binary decision. This implies that the fundamental shape of the pulse is such that they do not interfere with one another at the optimal sampling point [1]. In order to achieve this property, it needs to satisfy certain criteria. These criteria are known as Nyquist criteria which states that to achieve an ISI free transmission, the impulse response of the pulse shaping filter should have zero crossings at multiples of the symbol period. Therefore, the overall impulse response should satisfy the Nyquist criterion to minimize the ISI of transmitting pulses [2]. Moreover, the pulse shaping filters have two important properties to ensure the non-interference in the radio frequency (RF) wireless communication systems. First one is that high stop-band attenuation is required to reduce the ICI as much as possible. Second one is that in order to achieve the minimum ISI (arising from multi-path signal reflections) the BER is as low as possible. The source of bit errors in the base band data transmission system is the channel noise. Moreover, the pulse shaping filter plays a crucial role in spectral shaping in the modern RF wireless communication to reduce the spectral bandwidth. Hence, the pulse shaping is also a spectral processing technique by which the fraction out of band power is reduced for low cost, reliable and spectrally efficient mobile radio communication systems [3]. Again, when a signal is transmitted at higher modulation rate through a band-limited channel, it can create an ISI [3]. As we know, when the modulation rate increases, the signal bandwidth increases, then the signal bandwidth becomes larger than the channel bandwidth. As a result, the channel starts to introduce distortion to the signal and this distortion is usually seen as an ISI. Pulse shaping filters are being used widely in wireless communication [4-6] and several techniques exist in the literature for the design of pulse shaping filter to solve this problem [7-9]. To analyze the performance of a pulse shaping raised cosine filter under different quadrature amplitude modulation (QAM) in terms of error rate and signal to noise ratio (SNR), a vector signal generator and a vector signal analyzer are proposed in [10]. Nyquist filter is a special class of filter which is very useful for the Multirate implementations of the filter. It is also used in digital communications as a pulse shaping filter [4]. When $L = 2$ the Nyquist filter is also called a half-band filter in this case the cutoff frequency is π/L and the impulse response is zero for every L-th sample. The additional advantage of FIR design is that the every other coefficient is equal to zero. IIR design can achieve quasi-linear phase and they can offer a great cost savings and can achieve an extremely low pass-band ripples. On the other hand, the half-band filter is widely used in multirate signal processing applications when interpolating/decimating factor is two. The half-band filters are implemented efficiently in the polyphase form, because half of its coefficients are approximately equal to zero [11]. As a result, it reduces the memory locations requirement to store these coefficients. Due to multirate applications of half-band filter, it is widely used in digital signal processing, such as cell phones, digital receivers, television, compact disc (CD) & digital versatile disc (DVD) players etc. Furthermore, when multiple octaves of reduction are needed, a cascade of half-band filter is commonly used.

Therefore, this paper is concerned for the performance analysis of pulse shaping filter by using the FIR Nyquist, FIR half-band and IIR linear phase half-band in terms of BER to reduce the effect of ISI and ICI for errorless wireless

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

communication. We also focus on the hardware requirements to reduce their effective costs in the area of wireless communication applications.

The rest of the paper is organized as follows. Section II briefly introduces the theory of different filters. The performance and cost effective analysis of the different design filters are analyzed in Section III. Finally, the paper is concluded in Section IV.

II. THEORY

This section presents the different theory of different pulse shaping filters.

A. Raised Cosine Filter

The raised cosine filter is frequently used as a pulse shaping filter in digital modulation due to its ability to minimize the ISI. The vestigial symmetry is one of the most important property of raised cosine filter which means that if the frequency characteristic has odd symmetry at the cutoff frequency, the impulse response will have zeros at uniformly spaced intervals. The frequency-domain description of this filter is given below,

$$H(f)=\begin{cases} T & |f|\le \frac{1-\beta}{2T}, \\ \frac{T}{2}\left[1+\cos\left(\frac{\pi T}{\beta}\right)\left[|f|-\frac{1-\beta}{2T}\right]\right], & A<|f|\le B, \\ 0 & otherwise, \end{cases} \tag{1}$$

where, $A=\frac{1-\beta}{2T}$, $B=\frac{1+\beta}{2T}$ and β is the roll-off factor.

The impulse response of the raised cosine filter is given by,

$$h(t)=\frac{\sin \pi t/T}{\pi t/T}\,\frac{\cos\left(\frac{\pi\beta\, t}{T}\right)}{1-\frac{4\beta^2 t^2}{T^2}} \tag{2}$$

From equation (2) it is clear that it has a sinc term which ensures zero crossing like an ideal low pass filter. In addition, it has another term which decays in time.

B. FIR Nyquist Filter

Nyquist filter is a special type of filter with a transfer function which has certain zero value coefficients. Due to the presence of these zero-valued coefficients this filter is computationally more efficient than the other low pass filter of the same order. The relation between the output and input of the interpolator is given by,

$$Y(z)=H(z)X(z^L).$$

If the interpolation filter *H(z)* is realized in the Lth-band polyphase form, then we have,

$$H(z)=\sum_{i=0}^{L-1} z^{-i}E_i(z^L)\cdot$$

Assume that the k-th polyphase component of the *H(z)* is constant, i.e., $E_k(z)=\alpha$. Then, we can express *Y(z)* as

$$Y(z)=\alpha\, z^{-k}X(z^L)+\sum_{\substack{l=0\\ l\ne k}}^{L-1} z^{-l}E_l(z^L)X(z^L)\cdot \tag{3}$$

As a result, $y(Ln+k)=\alpha\, x(n)$. Thus the input samples appear at the output without any distortion for all values of n, in between (L-1) output samples determined by interpolation. A filter with the above property is called Nyquist filter or Lth-band filter. Now the impulse response of an ideal Lth-band low pass filter with a cutoff at $\omega_c=\frac{\pi}{L}$ is given by,

$$h_{LP}(Ln)=\frac{\sin(n\pi/L)}{n\pi},\ -\infty\le n\le\infty\,. \tag{4}$$

Hence, the coefficient condition of the Lth-band filter is given below,

$$h(Ln)=\begin{cases}\alpha, & n=0\\ 0, & otherwise.\end{cases}$$

It can be seen from the above that

$h(n)=0$ for n=±L, ±2L……

The transition width requirement can be deduced from the roll-off factor and interpolation factors as follows;

Transition bandwidth=2 x roll-off factor/interpolation factor.

C. Half-Band Filters

A half-band filter is a special type of Nyquist filters which can be used for the interpolation factor of 2. Half-band filters have two important characteristics, first one is that the pass-band and stop-band ripples must be the same, and the second one is that the pass-band-edge and stop-band-edge frequencies are equidistant from the half-band frequency π/2. These characteristics make it possible to design an FIR filter whose every other coefficient is zero, and whose non-zero coefficients are symmetrical about the center of the impulse response. Again the linear phase half-band filter has one important property that half of the filter coefficients are zero-valued which makes the design computationally efficient. Beside this, it plays an important role in multirate digital signal processing. FIR half-band filters are easily designed to exhibit exactly linear phase, but IIR half-band filters provide higher computation speed at the cost of the phase nonlinearity. Hence, the advantage of IIR half-band filters is evident in the applications where the computation speed, low power consumption and miniaturization are the main requirements. A half-band filter transfer function has to satisfy the pass-band and stop-band symmetry conditions. This implies the constraints in the choice of the filter parameters, pass- stop-band ripples and pass- stop-band-edge frequencies.

Since an Lth-band filter for L=2 is called a half-band filter, thus the transfer function of the linear phase half-band filter is given by,

$$H(z)=\alpha+z^{-1}E_1(z^2)\,. \tag{5}$$

The condition $H(z)=\alpha+z^{-1}E_1(z^2)$ reduces to H(z)+H(-z)=1 by assuming $\alpha=0.5$. If H(z) has real coefficient, then $H(-e^{j\omega})=H(e^{j(\pi-\omega)})$. Hence, $H(-e^{j\omega})+H(e^{j(\pi-\omega)})=1.$

The impulse response of this filter is given by

$$h(2n)=\begin{cases}\alpha, & n=0\\ 0, & otherwise.\end{cases} \tag{6}$$

From the above equation, we can observe that the odd indexed coefficients are zero-valued except for the central coefficient. This enables us to avoid approximately half the number of multiplications in the time of implementing this kind of filter.

III. Performance Analysis And Cost Analysis

In this section, we analyze the performances of pulse shaping filter by using FIR Nyquist, FIR half-band and IIR linear phase half-band in terms of BER to minimize the effect of ISI and ICI for wireless communication. We also analyze the hardware requirements which reduce their effective costs in the area of wireless communication applications. In fact the stop-band attenuation is a very important constraint and it must be minimized in order to reduce the ICI and ISI. To suppress the ICI and ISI effects of the pulse shaping filter the minimum desirable stop-band attenuation is about 60dB to 80dB. A raised cosine filter is typically used to shape and oversample a symbol stream before transmission. The width of the transition band of a raised cosine filter depends on the roll-off factor. Practical digital communication systems use a roll-off factor between 0.1 and 0.5. In order to evaluate the numerical simulation, we choose the oversampling or interpolation factor 8 and roll-off factor 0.5. The comparative magnitude responses of four pulse shaping filter are shown in the Fig.1.

From Fig.1, it can be observed that the FIR Nyquist, FIR half- band and IIR linear phase half-band filters have stop-band attenuation larger than 60dB. Hence we can say that three alternative pulses shaping filters provide the favorable solution to suppress the interferences in place of the raised cosine filter, because to achieve this minimum stop-band attenuation (60dB) the required number of is required 260 for the raised cosine filter. The magnitude and impulse responses of all the pulse shaping filters are discussed separately below.

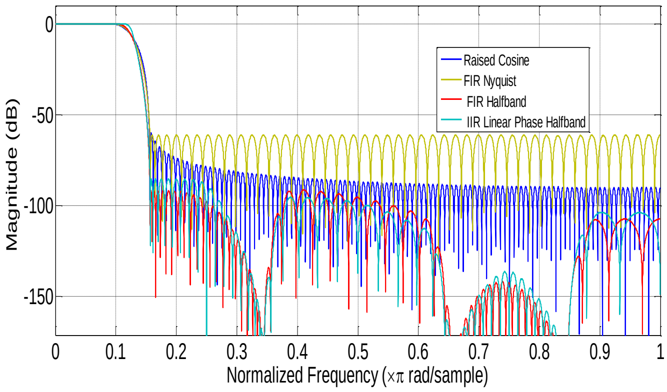


Fig.1 Transfer function magnitude comparison of all the designs filters

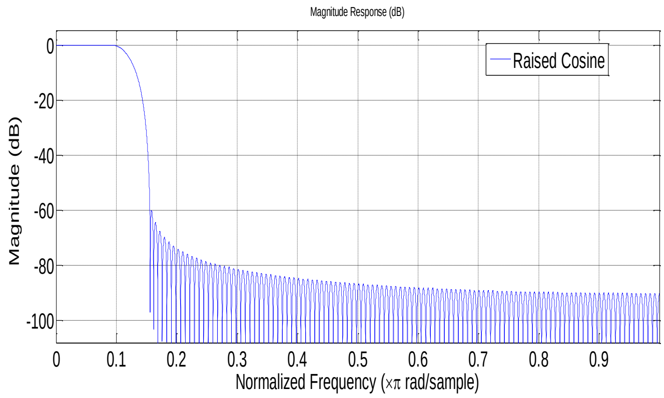


Fig. 2: Transfer function magnitude of the raised cosine filter

The magnitude and phase response of the raised cosine filter are shown in the Figs.2 and 3, respectively. From the magnitude and impulse response curves of raised cosine filter, it is clear that to attain a minimum stop-band attenuation of 60 dB the required number of filter order is 260 which is almost double than other three filters. So, the raised cosine filter is not optimal in any sense.

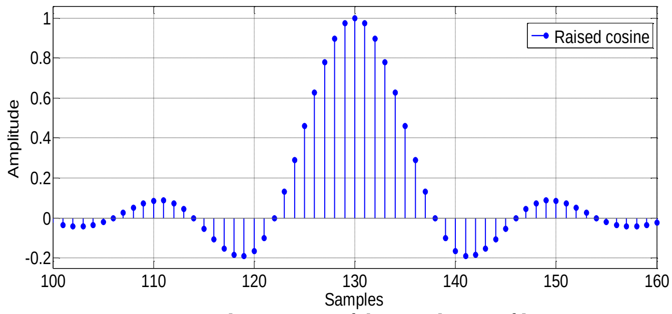


Fig. 3: Impulse response of the raised cosine filter

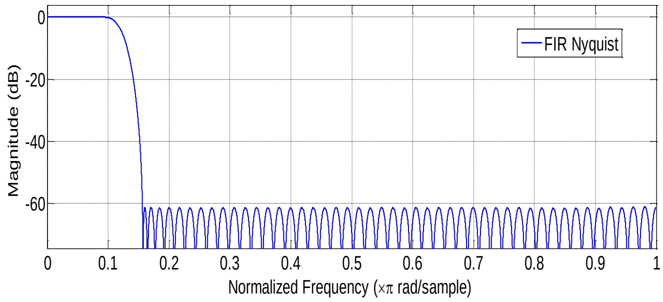


Fig. 4: Transfer function magnitude of the FIR Nyquist filter

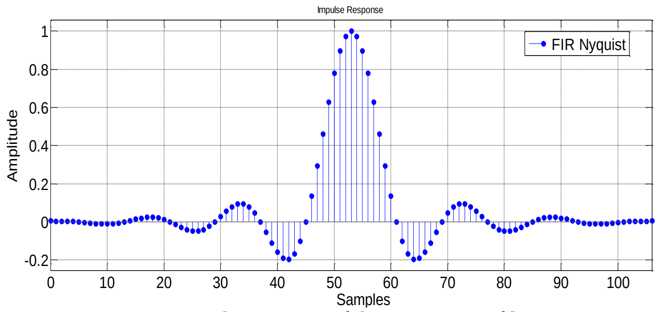


Fig. 5: Impulse response of the FIR Nyquist filter

The raised cosine filter can be replaced with FIR Nyquist filter for a fraction of the cost because they have an optimal equiripple response. The magnitude and phase response of the FIR Nyquist filter are shown in the Figs.4 and 5, respectively. The same stop-band attenuation and transition width is obtained in a lower order which is 106.

The magnitude response of the FIR half-band and IIR half-band filters are shown in the Figs.6, and 8, respectively. From these figures, it is clear that they have stop-band attenuation larger than 85dB.

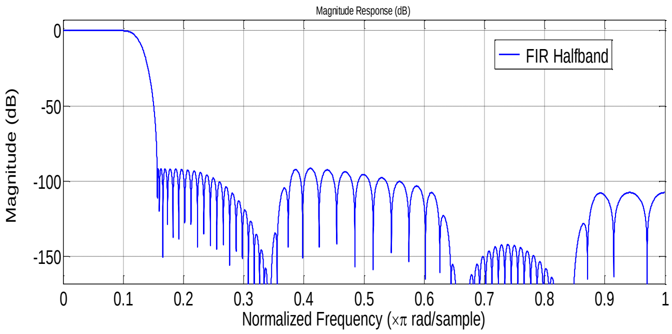


Fig. 6: Transfer function magnitude of the FIR half-band filter

The impulse responses of 8^{th} band pulse shaping raised cosine filter, FIR Nyquist filter, FIR half-band and IIR linear phase half-band filter are shown in Figs.3, 5, 7, and 9, respectively. From these figures, it is clear that the value of impulse responses for every 8^{th} sample is equal to zero (except at the center).

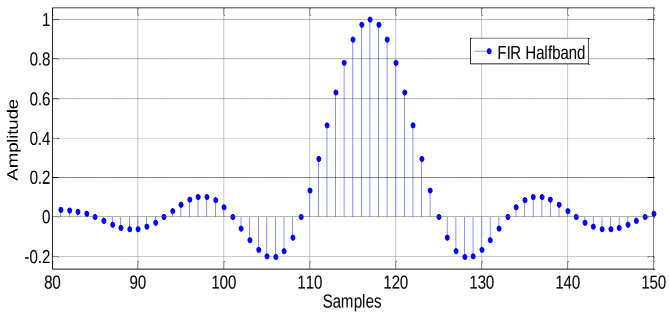


Fig. 7: Impulse response of the FIR half-band filter

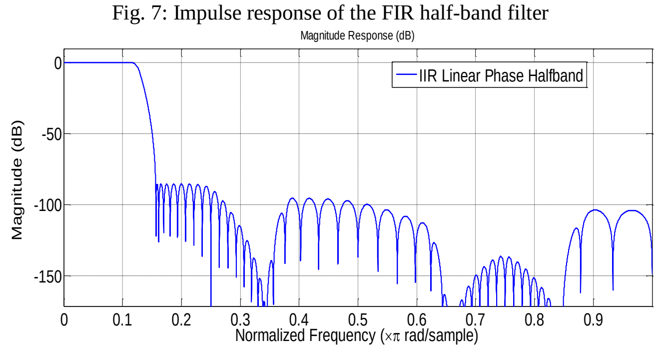


Fig. 8: Transfer function magnitude of the IIR linear phase half-band filter

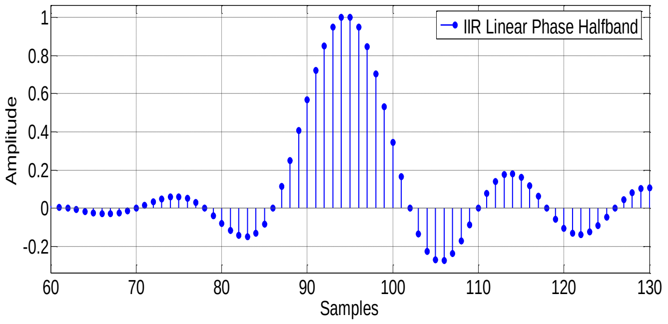


Fig. 9: Impulse response of the IIR linear phase half-band filter

From the magnitude response of the four pulse shaping filters, it is clear that the minimum stop-band attenuation is 60dB. So, we can say that the pulse shaping filters can reduce the ISI and ICI to a satisfactory level. To obtain the BER of the four pulse shaping filters we use a 16-QAM modulation scheme with an additive white Gaussian noise channel. The results show that the BER obtain for the four pulse shaping filters is very similar which is shown in table I.

TABLE I
BIT ERROR RATE (BER) COMPARISON

Filters	10dB SNR	15dB SNR	20dB SNR
Raised cosine	0.0586	0.0041	0
FIR Nyquist	0.0605	0.0046	0
FIR-Half band	0.0602	0.0046	0
IIR-Half band	0.0634	0.0055	0

TABLE III
IMPLEMENTATION COST COMPARISON

Filters	Multipliers	Adders	Multipliers per input sample	Adders per input sample	Number of States
Raised cosine	260	253	260	253	32
FIR Nyquist	94	87	94	87	13
FIR-half band	32	29	60	53	29
IIR-half band	12	24	22	44	29

From table I, we see that the performances of the four pulse shaping filters are almost identical, but their implementation cost varies greatly, which is shown in table II. From table II, we can observe that three pulse shaping filters (i.e. FIR Nyquist, FIR half-band and IIR half-band filters) provide significant savings both in terms of hardware and operations per sample as compared to raised cosine filter.

IV. CONCLUSION

In this paper, we focus on the performance comparison of three optimized alternative pulse shaping filters (i.e. FIR Nyquist, FIR half-band and IIR half-band filters) with the raised cosine filter. Based on our numerical evaluations and observations, we can conclude that the performance of optimized pulse shaping and raised cosine filters are almost identical in terms of the BER, but the cost of their hardware implementation in terms of operation per sample requirements varies greatly. The hardware implementation costs for the FIR Nyquist, FIR half-band and IIR half-band filters are about 60% to 90% less than the raised cosine filter. Although the raised cosine filter attracts a great attention to its user, for the advantages of minimizing ISI and ICI, but it is not cost-effective compared to the optimized pulse shaping filters proposed in this paper. Hence, we believe that the proposed optimized pulse shaping filters can be used to provide the cost-effective solution for the wireless communication applications.

REFERENCES

[1] Rajesh Mehra and Swati Singh, "FPGA Based Gaussian Pulse Shaping Filter for Error Removal," *International Journal of Research & Innovation in Computer Engineering*, Vol 1, Issue 1, August 2011.

[2] Massimiliano Laddomada, Letizia Lo Presti and Marina Mondin, "Digital pulse-shaping FIR filter design with reduced intersymbol and interchannel interference," *European Transactions on Telecommunications*,14:423–433, 2003

[3] A. S Kang and V. Sharma, "Pulse Shape Filtering in Wireless Communication-A Critical Analysis," International *Journal of Advanced Computer Science and Applications*, Vol. 2, No.3, Mar. 2011.

[4] Pervej, M.F.; Sarkar, M.Z.I.; Roy, T. K., "PAPR Reduction Analysis of DFT-SCFDMA System using Different Pulse Shaping Filters," in Strategic Technology (IFOST), 2014 9th International Forum on, Cox's Bazar, 21-23 Oct. 2014.

[5] Pervej, M.F.; Roy, T.K.; Koli, M.N.Y., "Impact analysis of input and output block size of DCT-SCFDMA system," in Computer and Information Technology (ICCIT), 2014 17th International Conference on, vol., no., pp.440-445, 22-23 Dec. 2014.

[6] Pervej, M. F.; Sarkar, M. Z. I.; Islam, Md Toufiqul, "Impact Analysis of Different Input and Output Block Size of DFT-SCFDMA System," presented at the 2nd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Dhaka, May 2015.

[7] Behrouz Farhang-Boroujeny, "A Universal Square-Root Nyquist (M) Filter design for Digital Communication Systems," *Proceeding of the SDR 06 Technical Conference and Product Exposition*, 2006.

[8] A S Kang and Vishal Sharma, "Data Communication using Pulse Shaping techniques in wireless signal processing approach-In depth review," *IJCEM International Journal of Computational Engineering & Management*, Vol. 12, April 2011.

[9] V.Sridhar, T. Nagalaxmi, M.V.B. Reddy and M. S., "Design and Implementation of Optimal Pulse Shaping Filters for Digital Radio Systems," *International Journal of Advanced Research in Computer Engineering & Technology*, Volume 1, Issue 5, July 2012

[10] Mahua Pal, "Algorithm for sharpening Raised Cosine Pulse Shaping Digital Filter and analysis of Performance of QAM system when subjected to sharpened Raise Cosine Filter, " *International Journal of Scientific and Research Publications*, Vol. 2, Isse.1, January 2012.

[11] S K Mitra, "Digital Signal Processing,: Tata Mc Graw Hill, Third Edition.

A Heuristic Solution of the Vehicle Routing Problem to Optimize the Office Bus Routing and Scheduling using Clarke & Wright's Savings Algorithm

Emrana Kabir Hashi, Md. Rokibul Hasan
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
{emranakabir, rakib063049}@gmail.com

Md. Shahid Uz Zaman
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
szaman22@yahoo.com

***Abstract*—Office bus routing is a rising problem because transportation service should be more efficient, safe and reliable. This research try to answer a question how to facilitate employees with transportation service that would be fast, cost effective, and in a timely manner. This research work discourse the Vehicle Routing Problem and scheduling issues of transportation service for Dhaka City, Bangladesh. Our implementation will help to develop office bus routing and scheduling prototype model. This model will help to design transportation management system of any organization. From this implementation, it will able to find shortest and fastest routes and schedule of office bus, and also allocate bus stops that will use to pick up and drop employees according to their residence. This application has considered cost matrix of time, routes length and the vehicles with same capacity. We have implemented this application with Clarke & Wright's savings algorithm. This paper concludes with the solutions of schedules and visualization of routes with GIS techniques that will helps to develop decision support system for any kind of Vehicle Routing Problem.**

Keywords—vehicle routing problem; heuristic solution; clarke & wright savings algorithm; GIS

I. INTRODUCTION

The Vehicle Routing Problem (VRP) is one kind of NP-hard problem; the goal is to serve customers with vehicles where customers are geographically apart. From this, have to determine routes with well developed algorithms that will minimize operational cost. In this problem, objectives are contradictory that will maximize customer service and minimize operational cost. The VRP can be stated as a set of customer with known location and demand are to be supplied from a depot by delivery vehicles with known capacity subjected to all customer demand being met, vehicle capacity not being exceeded and total trip length not exceeding some specified level [1]. The capacitated vehicle routing problem (CVRP) was initially introduced by Dantzig and Ramser (1959) in their article on a truck dispatching problem and, consequently, became one of the most important and widely studied problems in the area of combinatorial optimization [2]. To find the shortest routes and optimized the traveling cost is the main concern of VRP. Each vehicle starts from a depot and after completing the service cycle ends in the same depot. Clarke and Wright (1964) develop a heuristic solution method which becomes known as the savings method and was the first algorithm that becomes widely used [1]. The Clarke and Wright savings algorithm (CW) is the most widely applied heuristic for solving CVRP due to its simplicity of implementation and efficient calculation speed [3].

In this paper, we focus on one of GIS-based scheduling applications and GPS technologies for the vehicle routing problems [4]. Here Geographic Information System (GIS) is used to visualize the route on the Google map that shows the latitude and longitude of all bus stops with travel length and time.

II. BACKGROUND AND LITERATURE RESEARCH

Since CVRP was first proposed in 1959, it has received much attention from researchers and practitioners [2]. Human beings meet with VRP in different areas in daily life. Some of them: various types of food and drink, clothing, heating material transportation, and garbage collection, postal service, personnel and student transportation. In this point of view, VRP separates into two main areas: human and freight transportation [5]. Recent years, various location-based services are becoming popular. As in [6], spatial data, digital road map and moving vehicle records using various mobile sensors, are integrated and stored into the spatial temporal database systems. There are many research papers about spatial data mining technologies of probing and streaming data, and there exit various applications, for instance, optimization of travel plan, visualization of traffic jam, GPS assisted navigation, road design, and ICT-assisted traffic congestion information systems [6].

VRP has been studied widely in different optimization problem. VRP study works efficiently in different distributed network for minimizing cost and maximizing service. There are many paper works for solving school bus routing problem, this paper try to solve office bus routing problem and evaluate with Google map.

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

A. *Vehicle Routing Problem (VRP)*

Customers i=1,...,n with demands of a product must be served using a fleet of vehicles for the deliveries. The vehicles, with given maximum capacities, are situated at a central depot (or several depots) to which they must return. Determine a routing schedule that minimizes the total cost of the deliveries. Graph formulation for a single depot problem:
Input: Graph G=(V,E), with costs (or distances) C_{ij} on the edges and demands d_i on the vertices, |V|=n, m vehicles with capacities K_j, j=1,...,m. A central depot is at vertex 0. Problem: Determine cycles R_1, ..., R_m for the vehicles that start from vertex 0 and service all vertices such that the load of vehicle j doesn't exceed its capacity and the total cost of the cycles is minimized [7].

III. Objectives and Problem Statement

A. *Objectives*

The objective of this paper is to define the office bus routing problem and find the solution of this problem using the Clarke and Wright's savings algorithm. The outputs will be shown in the Google map to clear the visualization of the resultant routes and schedules.

B. *Problem Description*

In Bangladesh Office bus service is popular as school bus service. There are some difficulties (cost, time, reliability etc.) to provide office bus service in Dhaka City, based on employee numbers. There are lots of office wish to provide this transportation service to their employees. For our VRP application we have considered a software company in Dhaka called "Tasawer Interactive" who provided transportation facility to their employees and currently stopped because of those difficulties. All employees of this company live in different place in Dhaka City as scattered. We are going to propose a transportation system that will include routes and schedule; this will minimize their cost and time and maximize the reliability.

C. *Model Formulation*

Consider the office bus transportation problem to pick up and drops the employees between home and office. Employees are lived in geographical area around the office and the buses that carried them are same capacities. Here the following notations have used to give a formal description of the problem [8]:

$N = \{0, 1, \ldots, n\}$: a set of locations where 0 indicates the depot(office) and j (for $j = 1, \ldots, n$) is the index of a location where one or more employees live

$M = \{1, \ldots, m\}$: a set of buses

$R_i = \{ r_i(1), \ldots, r_i(n_i) \}$: the route for bus i, where $r_i(j)$ is the index of the jth location visited and n_i is the number of locations in the route. We assume that every route finishes at the office, i.e., $r_i(n_i+1) = 0$

t_{jk}: the direct traveling time from location j to location k, for $j = 0, \ldots, n$ and $k = 0, \ldots, n$ and $t_{jk} = 0$ for $j = k$

c_i : the capacity of a bus i

q_j: the number of employees to be picked up at location j, for $j = 1, \ldots, n$

$length(i)$: the length of route i (which is also the maximum traveling time corresponding to the employees picked up at the first location).

The bus routing problem is to find a set of routes in order to:

Minimize m (1)

and $length\ (i)$ (2)

Objective function (1) attempts to minimize the number of buses while objective function (2) minimizes the maximum time in the bus. Note that all the employees must be picked up and that a given location cannot be assigned to more than one route. When we have found the resultant route then we will schedule the time table of bus for our transportation system.

IV. Methodology

Most heuristic methods can be classified into two categories as follows:

• Constructive methods: Regarding capacities and costs, routes are made by adding nodes to partial routes or combining sub routes. Example: Clarke & Wright algorithm.
• Two-phase methods: Consist of 1) clustering of vertices and 2) route construction. The cycle of these stapes may be "route first, clustering later" or "clustering first, route later". Example: Sweeping heuristics.

In this work, we have used Clarke & Wright algorithm as constructive heuristics. Though we are going to design routes for office, so Clarke & Wright algorithm would be best choice for us, because of lower employee numbers. If employees' number will increase then it will better to modify the algorithm.

A. *Clarke & Wright savings heuristics*

Algorithmic approach is given bellow [8].

Assume that a complete, undirected graph with symmetric costs and an unlimited number of identical vehicles with capacity K.

Step1. Form sub tours i-0-i for i=1,2,...,n. (Each customer is visited by a separate vehicle)
Step2. Compute savings $s_{ij}=c_{0i} + c_{0j} - c_{ij}$ for all i,j and $i\neq j$.
Step3. Identify the node pair (i,j) that gives the highest saving s_{ij}.
Step4. Form a new subtour by connecting (i,j) and deleting arcs (i,0) and (0,j) if the following conditions are satisfied

a) Both node i and node j have to be directly accessible from node 0,

b) Node i and node j are not in the same tour,

c) Forming the new sub tour does not violate any of the constraints associated with the vehicles,

Step5. Set $s_{ij}=-\infty$, which means that this node pair is processed.
Step6. Go to Step 3, unless all node pairs with $s_{ij} \geq 0$ are processed.

B. Numerical example

The actual way the algorithm works is illustrated in the following by means of a numerical example.

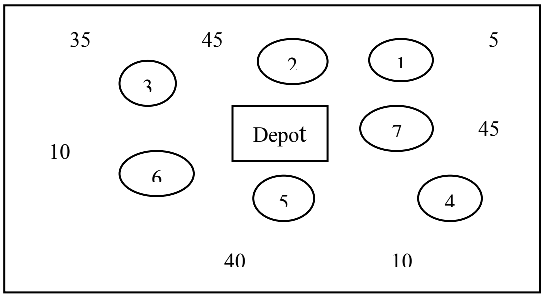


Fig. 1: Customer with their demands.

In Fig.1, we have considered a problem with 7 customers. The transportation costs between all pairs of points are shown in TABLE I, where 0 represents the depot (the costs are symmetric, and for that reason only the upper half of the table is filled in).

TABLE I. TRANSPORTATION COSTS BETWEEN ALL PAIRS OF NODES

i/j	0	1	2	3	4	5	6	7
0	-	20	57	51	50	10	15	90
1		-	51	10	55	25	30	53
2			-	50	20	30	10	47
3				-	50	11	60	38
4					-	50	60	10
5						-	20	90
6							-	12
7								-

The customers' demands that must be delivered from the depot are given in TABLE II.

TABLE II. THE CUSTOMERS' DEMANDS

i	1	2	3	4	5	6	7
di	5	45	35	10	40	10	45

Capacity of a vehicle: K = 60

Solution with Clarke & Wright:

Savings $s_{ij} = c_{0i} + c_{0j} - c_{ij;}$ Symmetry: $s_{ij} = s_{ji}$

The savings S_{ij} are calculated to the following values (only the upper half of the table is shown, since the savings are symmetric due to symmetric costs) are shown in TABLE III:

TABLE III. THE SAVINGS S_{IJ}

i/j	1	2	3	4	5	6	7
1	-						
2	26	-					
3	61	58	-				
4	15	87	51	-			
5	5	37	50	10	-		
6	5	62	6	5	5	-	
7	57	100	103	130	10	93	-

Result: TABLE IV describes the resultant routes with correspond load and cost.

TABLE IV. SOLUTIONS

Route	Load	Cost
0-2-0	45	114
0-5-0	40	20
0-4-7-0	55	150
0-1-3-6-0	50	105

Total cost = 389

The resultant routs can be shown in the following figure.

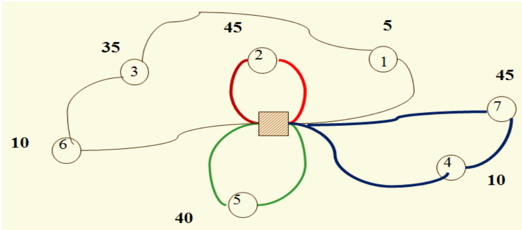


Fig. 2: The resultant routes.

V. IMPLEMENTATION

A. Data Collection

We have implemented this savings algorithm based on a software Company (Tasawar Interactive) in Dhaka, Bangladesh. At first we have collected all employees address and built location database. Then we have chosen a pick up point for employees in common area. For this reason we have chosen eight pick up points for fifty employees. The locations with symmetric cost have shown in TABLE V. This company used minibus with capacity 20 seats. We have calculated times and distance among all points including office and each pick up points using GPS and Google map. We have built a time matrix based on number of employees have to pick up in each points. We have considered time as cost because office starts at 10:00 AM, and have to drop employees before start the office. Then we have determined savings matrix and in the next step we determine the resultant routes based on the algorithm. Finally plot those routes on Google map though deriver can drive as predefined routes. In this application, there will be different options such as edit the location, cost matrix, demand and bus capacity.

TABLE V. LOCATIONS LIST WITH COSTS & DEMANDS

Locations	Symmetric Costs (minutes)	Demands
Mohakhali DOHS	0,18,20,10,13,15,18,16,13	0
Mirpur-10	18,0,22,15,13,11,17,11,14	3
House B. Bus Stop	20,22,0,17,29,30,33,31,28	10
Uttar Badda	10,15,17,0,17,22,25,23,20	2
Shegunbagicha	13,13,29,17,0,11,19,16,15	9
Dhanmondi	15,11,30,20,11,0,10,10,12	11
31/1 Azimpur Road	18,17,33,25,19,10,0,15,18	4
Mohammadpur	16,11,31,23,16,10,15,0,18	6
Kazipara Bus Stop	13,14,28,20,15,12,18,18,0	5

B. Implementation Tools

We have developed our algorithm with language Ruby. The designs of our application have implemented with html and CSS. Finally we have used jquery, Google Map Api for showing the result on the Google map.

C. Routing & Scheduling on Google Map

For 50 employees 3 (three) minibus are needed. Our resultant 3 routes are shown in TABLE VI:

TABLE VI. RESULTS

Resultant Route	Capacity (employees)	Times (minutes)
Mohakhali DOHS<-->Dhanmondi<-->31/1 Azimpur Road<-->Kazipara Bus Stop<-->Mohakhali DOHS	20	56.0
Mohakhali DOHS<-->Mirpur-10, Begum Rokeya Avenue<-->House Building Bus Stop<-->Uttar Badda<-->Mohakhali DOHS	15	36.0
Mohakhali DOHS<-->Shegunbagicha<-->Mohammadpur<-->Mohakhali DOHS	15	45.0

The resultant routes are described below:

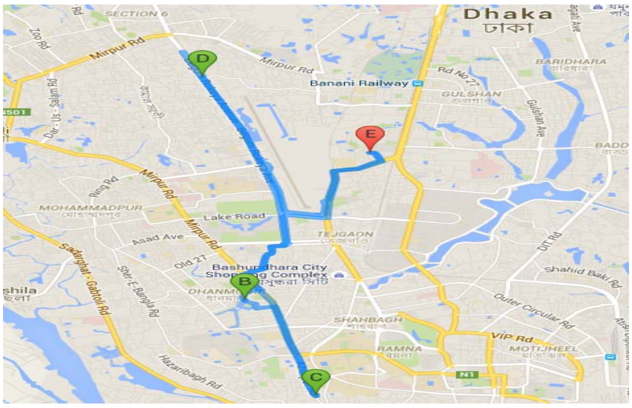


Fig. 3: Route 1 for bus 1.

Bus 1 able to pick up 20 employees and time is needed 56.0 minutes (without considering traffic jam and others delay).

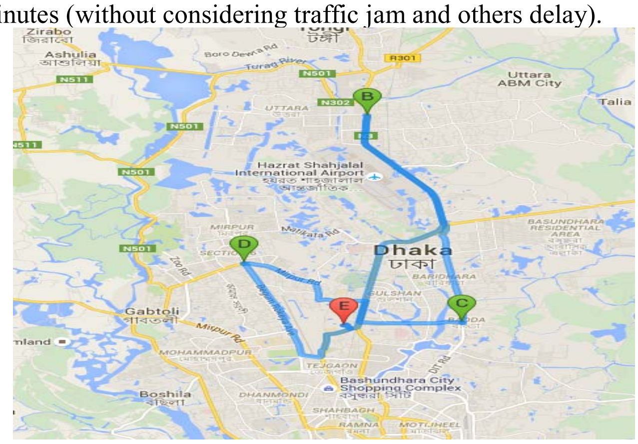


Fig. 4: Route 2 for bus 2.

Bus 2 able to pick up 15 employees and time is needed 36.0 minutes (without considering traffic jam and others delay).

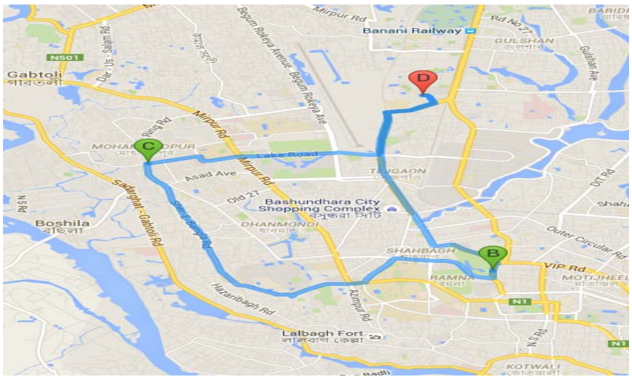


Fig. 5: Route 3 for bus 3.

Bus 3 able to pick up 15 employees and time is needed 45.0 minutes (without considering traffic jam and others delay).

VI. CONCLUSION

In the result, we are able to utilize the maximum capacity of every bus and minimize the number of buses. In past, four minibuses were used in this transportation service. We have an improvement that took low cost and time than before. Our application have minimized length, time, number of stops and buses, and maximized the service quality. This application are fully automated and anyone can use it from anywhere with compatible device and internet connection. Number of employees can be adjusted any time with a simple database entry and the optimized routes can be determined. Our application can work for any kind of organization (schools, offices) database for determining routes and schedules based on time. It is also flexible for bus capacity; number of employees in a route can be adjusted dynamically by setting bus capacity. We will try to improve the existing algorithm for optimization of routing problem.

REFERENCES

[1] K. Graham, “The life and times of savings method for Vehicle Routing Problem,” Lancaster University Management School, United Kingdom, vol. 25(2), 2009, pp. 125-145.

[2] G. B. Dantzig, J. H. Ramser, “The truck dispatching problem,” Institute for Operations Research and the Management Science, October 1959, pp. 80–91.

[3] G. Clarke, J. W. Wright , “Scheduling of vehicles from a central depot to a number of delivery points,” Operations Research, 1964, pp.568–81.

[4] H. Yu, S.L. Shaw, “Exploring potential human activities in physical and virtual spaces: A spatiotemporal GIS approach,” International Journal of Geographical Information Science, 2008, pp.1-45.

[5] M. Fatih Demiral, I. Güngör, K. Oğuzhan, "Optimization at service Vehicle Routing and a case study of Isparta, Turkey," Süleyman Demirel University, Turkey, pp. 374-387.

[6] K. Huroyuki, K. Masahiro, “Multi-agent scheduling and routing problem with time windows and visiting activities,” Nanzan University, Hyogo University of Health Sciences, 2009, pp. 442-447.

[7] W. Min, “Rich Vehicle Routing Problems and applications”, Department of Management Engineering, Denmark, 2010, pp. 1-50.

[8] Angel C., Elena F., Manuel L., Rafael M., “Heuristic solution to the problem of Routing School buses with multiple objectives,”

[9] Jens L., “ Clarke & Wright's saving Algorithm,”

Interaction with Large Screen Display using Fingertip & Virtual Touch Screen

Dejan Chandra Gope
Computer Science and Engineering (CSE)
Dhaka University of Engineering and Technology (DUET)
Dhaka, Bangladesh
Email: dejan.gope23@gmail.com

Md. Shafiqul Islam
Computer Science and Engineering (CSE)
Dhaka University of Engineering and Technology (DUET)
Dhaka, Bangladesh
Email: msislam_80@yahoo.com

ABSTRACT:- **Human-Computer Interaction (HCI) involves the study, planning, design and uses of the interaction between people (users) and computers. We present a nonintrusive system based on computer vision for human-computer interaction controlled by finger pointing gestures. A new approach for Interaction with Large Screen Display using Fingertip & Virtual Touch Screen by taking into account the location of the user and the interaction area available. We can estimate an interaction surface virtual touch screen between the display and the user. Users can use their pointing finger to interact with the display as if it was brought forward and presented directly in front of them, while preserving viewing angle. Users are allowed to walk around in a room and manipulate information displayed on its walls by using their own finger as pointing devices. Once captured and tracked in real-time using stereo vision camera, finger pointing gestures are remapped onto the current point of interest, thus reproducing in an advanced interaction scenario the "drag and click" behavior of traditional mice. It's a simple signal processing on images obtained from a regular laptop web-camera.**

Keywords:- **Human Computer Interaction (HCI), Hand Gesture Recognition, Virtual Touch Screen, Hand Pointing, Pointing Accuracy, Computer Vision, Fingertip Interaction, Hand-posture Recognition, Contour Tracking, Convex Hull.**

I. INTRODUCTION

HCI (human-computer interaction) is the study of how people interact with computers [4] and to what extent computers are or are not developed for successful interaction with human beings. Typically, humans interact with computers using mouse and keyboard. Outside of computer interaction, we are used to interacting with the world using our hands, body, and voice. Interfaces based on interaction with hands are a natural and intuitive way to interact with computers[13]. Such an interface could be used for robot [8] and human collaboration, virtual reality, scientific visualization, geographic information systems (GIS), or games. Microsoft Kinect [5] becoming popular, complicated background is no longer a insurmountable barrier and a real-time hand segmentation can be attained based on depth information given by the sensor. In the depth image, each value of the pixel refers to the distance from the person to Kinect and the hand can be segmented as the nearest region from sensor. Even though Kinect successfully solved the intractable segmentation problem, most of the depth image based approaches are built on the strong assumption that the hand remain unmoved[1]. Unfortunately, when the hand is moving or has small but sudden rotation, the boundary of the hand becomes so unstable that it is hard to recognize the hand as ideal circumstances. Virtual touch screens (VTS) [1] are used primarily for entertainment and research purposes. A popular virtual touch screen system is the Kinect system made commercially available by Microsoft. The Kinect system is primarily used for consumer entertainment such as playing games and using the menus found on the Xbox system. However, in recent years the Kinect has been used for research projects in the field of robotics and3-D interactions outside of the field of games such as browsing the web, and medical practices. There is a potential for the technology where the 3D [10], [11] visualization and physical interaction with the objects is necessary, such as teaching students introductory calculus. Virtual touch screen technology can also be applied to non-touch displays into interactive, touch-capable surfaces using a Kinect sensor in combination with a projector.

In medical practices, virtual touch interaction has been applied to allow users to view and manipulate digital data such as 3D images used in medical scans, allowing medical staff to interact with the data without any physical contact, thus avoiding the need for re-sterilization. Additionally, Microsoft's Kinect sensor has also employed the use of virtual touch screen technology to help stroke patients recover and improve motor function in their limbs through a game-like system involving interpretation of gestures combined with adjustable difficulty levels based on the user's performance.

Finger and hand gesture recognition with the Kinect is still an open problem due to its low resolution (640*480), especially considering how hands occupy a much smaller portion of the full body image. Traditional vision-based hand gesture recognition methods are far from satisfactory due to the limitations of the optical sensors used and dependency on lighting conditions and backgrounds. Data gloves can be used for precise accuracy, but require the user to wear a special glove; this may hinder the naturalness of the hand gesture. The Microsoft Kinect is a commodity hardware device that can be used for designing natural gesture based interfaces.

In order to address this problem, we propose a method to perform hand tracking in a mostly static environment with only the hands in the range of visiblity using OpenCV. We apply background elimination using simple techniques, and apply a set of filter to get the hand's contour, on which we create a convex hull and calculate the defect points in the hull which is used to detect the no of fingers, center and size of the palm which is a challenge.

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

II. VIRTUAL TOUCH SCREEN

A virtual touch screen (VTS) is a user interface system that augments virtual objects into reality either through a projector or optical display using sensors to track a person's interaction with the object. For instance, using a display and a rear projector system a person could create images that look three-dimensional and appear to float in midair. A virtual touch screen is constructed for human computer interaction [1], which does not need estimate pointing direction. Many researchers pay much attention to human body skeleton, so hand-arm or hand-eye is always used to determine the pointing direction. Fig-1 shows that how user interact virtually with the large screen display.

Virtual touch screen (VTS) is used primarily for entertainment and research purposes. A popular virtual

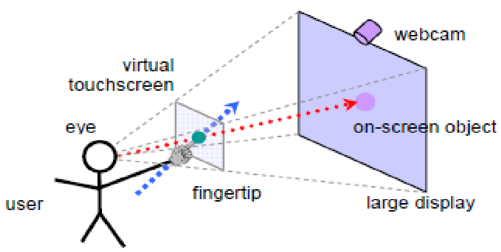


Fig-1: Fingertip Interaction System.

touch screen system is the Kinect system made commercially available by Microsoft. The distance between the operator and the virtual touch screen remains unchanged namely the arm's length is assumed to be constant [1]. Apparently, it's not suitable for different operators. A virtual touch screen is constructed in this paper which is scalable and flexible. Define z coordinate of the virtual screen as z one of user's pointing fingertip. It will move forward or backward as the pointing arm extends or contracts the screen which is more natural and suitable for different operators.

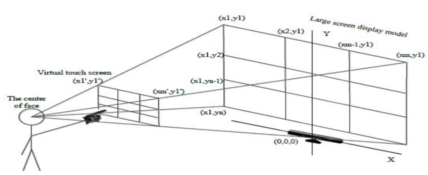


Fig-2: Virtual Touch Screen Interaction.

When an operator stands in front of the Kinect sensor and interacts with a large screen. The Kinect coordinate is defined as following shown in Fig-2. Axis X is upturned, axis Y is rightward and axis Z is vertical. The Kinect can capture the depth of any objects in its workspace.

The 3D coordinates of head and pointing index fingertip are used to construct a virtual touch screen [2] as Fig-2. One can note that the large screen and Kinect are in the same plane, which means its z coordinate is zero. Assume the large screen is divided into $m \times n$ modules, and the x coordinates of vertical demarcation points are $x1$, $x2$, … xm, the y coordinates of the horizontal demarcation points are $y1$, $y2$, …, yn. Suppose the corresponding values on virtual screen are $x'i$, $y'j$, i=1, 2, …, m, j=1, 2, …, n.

$$\frac{x_h - x_i'}{x_h - x_i} = \frac{y_h - y_j'}{y_h - y_j} = \frac{z_h - z'}{z_h - z} \qquad (1)$$

where z is 0, and z' equals to the pointing fingertip's z coordinate, x_h, y_h and z_h refer to the 3D coordinates of operator's head joint.

The corresponding coordinates on virtual touch screen can be figured out as (2), (3),

$$x_i' = \frac{(x_h - x_i)\times(z_f - z_h)}{z_h} + z_h \qquad (2)$$

$$y_j' = \frac{(y_h - y_j)\times(z_f - z_h)}{z_h} + y_h \qquad (3)$$

III. HAND SEGMENTATION

The RGB images and depth images are captured by the Kinect sensor which is fixed in front of the operator [3]. Human hand tracking and locating is carried out by continuously processing RGB images and depth images of an operator who is performing the pointing behavior to interact with large screen. Microsoft Kinect has the ability to track the movements of 24 distinct skeletal joints on the human body, wherein head, hand and elbow points have been used in this experiment. When operator held out the hand to perform pointing gesture [6][11], the hand is closer than the other one. The 3D coordinates of pointing hand joint can be obtained from skeletal map.

$$H_p = \begin{cases} H_r, & z_r < z_l \\ H_l, & else \end{cases} \qquad (4)$$

Where Hp is the pointing hand, Hr refers to right hand and Hl refers to left hand, zr and z_l present the z coordinates of right and left hand respectively.

In order to catch the hand motion used for controlling the large screen, it is need to separate the hand from the depth image. A depth image records the depth of all the pixels of the RGB image. The depth value of hand joint can be figured out through skeleton-to-depth conversion [9], and it is taken as a segmentation threshold which is used to divide the hand region $Hd(i, j)$ from the raw depth image.

$$H_d(i,j) = \begin{cases} 255, & d(i,j) < T_d + \varepsilon \\ 0, & else \end{cases}, d(i,j) \in D; i = 1,2,\cdots,m; j = 1,2,\cdots n \qquad (5)$$

Where $d(i, j)$ is the pixel of depth image D, Td is the depth of the pointing hand's joint, e refers to a small range around threshold Td, m is the width of D, and n is the height of D.

IV. FINGERTIP DETECTION

In this paper, pointing fingertip with more precise is detected instead of hand to interact with large screen. Most of existing methods for pointing gesture recognition use the operator hand-arm motion to determine the pointing direction. To further study hand gestures, some approaches for fingertips detection have been proposed. The hand's minimum bounding rectangle makes it easy to extract the pointing fingertip [2].

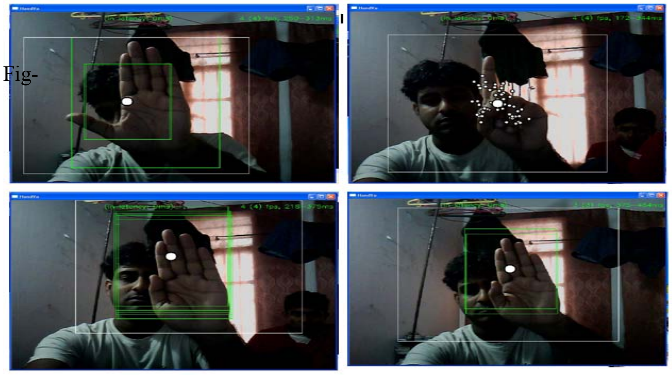


Fig-

3: Hand Detection.

The rules for pointing fingertip detection especially index fingertip are as follows:

(1) Extract the tracked hand region using the minimum bounding rectangle by skeletal information as well as the corresponding elbow.

(2) When the operator's hand is pointing to left, his or her pointing hand joint's coordinate in the horizontal direction (x coordinate) is less than the corresponding elbow's.

At this moment, if the hand joint's coordinate in vertical direction (y coordinate) is also less than elbow's and the width of minimum bounding rectangle is less than its height, it's not difficult to find that index fingertip moves along the bottom edge of bounding rectangle;

If the pointing hand joint's x coordinate is larger than the corresponding elbow's and the width of the minimum bounding rectangle is less than its height, the index fingertip moves along the up edge of bounding rectangle; If the minimum's width is larger than its height, the index fingertip moves along the left edge of bounding rectangle.

(3) When the operator's hand is pointing to right, his or her pointing hand joint's x is larger than that of the corresponding elbow's one, the distinguishing rule is the same as step (2).

The distinguishing rules above are to find the pointing fingertip in depth image, and its 3D coordinates can be figured out through depth-to-skeleton conversion.

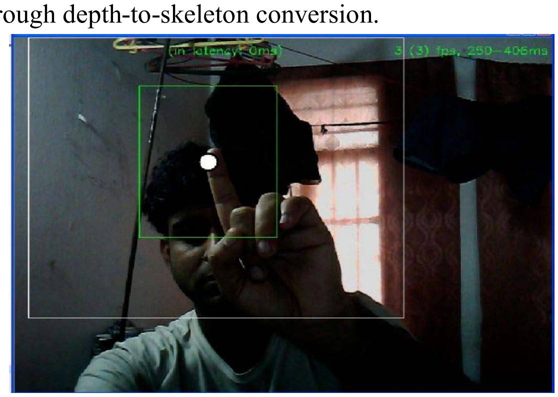


Fig-4: Fingertip Detection.

V. FINGERTIP TRACKING

A Kalman filter [14] is used to record the detected fingertip motion trajectory the features of pointing gestures (including fingertip's position and speed) can be extracted through this tracking method. In some cases, the fingertip will be detected by mistaken due to the surrounding environmental interference. To avoid false detection, a tracking technique is introduced to track the feature points [13].

Motion of pointing fingertip can be described by a linear dynamic model consisting of a state vector *x(k)* and a state transition matrix *Φ(k)*. The state vector containing the position, velocity in all three dimensions is described as (6).

$$x(k) = [x, y, z, v_x, v_y, v_z] \quad (6)$$

where *x, y, z* presents the image coordinates of the detected fingertip, and v_x, v_y, v_z presents its displacement.

The Kalman filter model assumes the true state at time *k* is evolved from the state at *(k−1)* according to (7).

$$x(k) = \Phi(k)x(k\text{-}1) + W(k) \quad (7)$$

where *Φ(k)* is the state transition model which is applied to the previous state *x(k−1); W(k)* is the process noise which is assumed to be drawn from a white Gaussian noise process with covariance *Q(k)*.

At time *k* an observation (or measurement) *z(k)* of the true state *x(k)* is made according to (8).

$$z(k) = H(k)z(k) + V(k) \quad (8)$$

where *H(k)* is the observation model which maps the true state space into the observed space and *V(k)* is the observation noise which is assumed to be zero mean Gaussian white noise with covariance *R(k)*.

In our method, the Kalman filter [14] is initialized with six states and six measurements, the measurements correspond directly with *x, y, z* in the state vector.

VI. GESTURES RECOGNITION

Gestures recognition [15] refers to detecting and extracting meaningful gestures from an input video. It is crucial how to recognize pointing gesture since the recognition procedure is only performed for the detected pointing gestures. Human motion is a continuous sequence of actions or gestures and non-gestures without clear-cut boundaries. When a person makes a pointing gesture, the whole motion can be separated into three phases including non-gesture (hands and arms drop naturally, it's unnecessary to find fingertip), move-hand (the pointing hand is moving and its direction is changing), point-to (the hand is approximately stationary). Among the three phases, only the point-to phase or the corresponding pointing gesture is relevant to target selection [13].

To recognize pointing gestures, the three phases must be distinguished. When operator moves his hand, the velocity at time *t* is estimated by (9)

$$V_x = x_t - x_{t-1},\ V_y = y_t - y_{t-1},\ V_z = z_t - z_{t-1} \quad (9)$$

where *x, y, z* represents the position of pointing fingertip in 3D space, respectively. If *vx, vy* and *vz* maintains in a small specific range v_T, it is approximately stationary. Assuming this stationary state lasts for *n* frames, the initial value of *n* is 0

$$n = \begin{cases} n+1, & |v_x| < v_T \ \& \ |v_y| < v_T \ \& \ |v_z| < v_T \\ 0, & \text{else} \end{cases} \quad (10)$$

When n reaches a certain amount, it presents the operator is pointing at the interaction target, & is logical AND operation. The effective gesture recognition is Rp.

$$R_p = \frac{M_r}{M_a} \quad (11)$$

Where Rp presents the rate of pointing gesture recognition, Mr refers to the frame number of recognized pointing gesture and Ma refers to the actual pointing frame number.

$$R_c = \frac{N_c}{N_p} \times 100\% \quad (12)$$

Where Rc refers to the correct target selection rate, Np is the times of pointing to one target and Nc is the times of correct target selection.

VII. EXPERIMENT RESULTS

In this section we describe the accuracy of our hand tracking and gesture recognition algorithm. The application has been implemented in C++ using proposed methodology the OpenCV libraries, OpenNI and NITE [12]. The application has been tested on a Microsoft Windows 7. The images have been captured using a A4 Tech WebCam with USB connection. The camera provides 640x480 images at a capture and processing rate of 30 frames per second. For the performance evaluation of the hand detection and gesture recognition, the system has been tested 6 times respectively on a set of 6 users. Each user has performed a predefined set of 6 movements and therefore we have 360 gestures to evaluate the application results. It is natural to think that the system's accuracy will be measured controlling the performance of the desired user movements for managing the calculator.

	*L*1	*L*2	*L*3	*L*4	*L*5	*L*6
*L*1	**98.4%**	1.6%	0	0	0	0
*L*2	0.5%	**99.5%**	0	0	0	0
*L*3	0	0	**99.6%**	0.4%	0	0
*L*4	0	0	0	**100%**	0	0
*L*5	0	0	0	0	**98.4%**	1.6%
*L*6	0	0	0	0	0.5%	**99.5%**

Table-1: Results of Target Selection with Kalman Filter.

VIII. CONCLUSION

Hand gesture recognition for real-life applications is very challenging because of the requirements on its robustness, accuracy and efficiency. In this paper, we presented a robust real-life hand gesture recognition system using the Kinect sensor. Another contribution of this paper is the real-life HCI applications we built on top of our hand gesture recognition system. It shows that with hand gesture recognition technique we can mimic the communications between human, and involve hand as a natural and intuitive way to interact with machines.

Consequently we can benefit our daily life in many aspects such as providing aids for the hearing impaired, and maintaining absolute sterility in health care environments using touch less interfaces via gestures. Our future research will focus on exploring a more efficient part-based representation and the efficiency drawback of near-convex decomposition based finger detection method. And we will further develop interesting HCI applications of our hand gesture recognition system.

REFERENCES

[1] K. Cheng and M. Takatsuka, "Estimating virtual touch screen for fingertip interaction with large displays", Proceedings of ACM International Conference Proceeding Series, **(2006)**, pp. 397-400.

[2] N. Miyata, K. Yamaguchi and Y. Maeda, "Measuring and modeling active maximum fingertip forces of a human index finger", Proceedings of IEEE International Conference on Intelligent Robots and Systems, **(2007)**, pp. 2156-2161.

[3] M. Van Den Bergh and L. Van Gool, "Combining RGB and ToF cameras for real-time 3D hand gesture interaction", Proceedings of 2011 IEEE Workshop on Applications of Computer Vision, **(2011)**, pp. 66-72.

[4] Dejan Chandra Gope, Md. Nasir Uddin, Md. Anisur Rahman, "Fingertip Interaction With Large Screen Display System Using Virtual Touch Screen", Vol. 5, No. 2, Mar. 2013.

[5] J.L. Raheja, A. Chaudhary, K. Singal, "Tracking of Fingertips and Centre of Palm using KINECT ", In proceedings of the 3rd IEEE International Conference on Computational Intelligence, Modelling and Simulation, Malaysia, 20-22 Sep, 2011, pp.248-252.

[6] J. R. Michael, C. Shaun and J. Y. Li, "A multi-gesture interaction system using a 3-D iris disk model for gaze estimation and an active appearance model for 3-D hand pointing", IEEE Transactions on Multimedia, vol. 3, no. 13, **(2011)**, pp. 474-486.

[7] Kelvin Cheng "Direct interaction with large displays through monocular computer vision " 2008.

[8] L. Cheng, Q. Sun, H. Su, Y. Cong and S. Zhao, "Design and implementation of human-robot interactive demonstration system based on Kinect", Proceedings of 24th Chinese Control and Decision Conference, (2012), pp. 971-975.

[9] H.-I Suk, B.-K. Sin and S.-W. Lee, "Hand gesture recognition based on dynamic Bayesian network framework", Pattern Recognition, vol. 9, no. 43, (2010), pp. 3059-3072.

[10] M. Raj, S. H. Creem-Regehr, J. K. Stefanucci and W. B. Thompson, "Kinect based 3D object manipulation on a desktop display", Proceedings of ACM Symposium on Applied Perception, **(2012)**, pp. 99-102.

[11] Y. Li, "Hand gesture recognition using Kinect", Proceedings of IEEE 3rd International Conference on Software Engineering and Service Science, **(2012)**, pp.196-199.

[12] N. Villaroman, D. Rowe and B. Swan, "Teaching natural operator interaction using OpenNI and the Microsoft Kinect sensor", Proceedings of ACM Special Interest Group for Information Technology Education Conference, **(2011)**, pp. 227-231.

[13] Pan Jing, Guan Ye-peng "Human-computer Interaction using Pointing Gesture based on an Adaptive Virtual Touch Screen", Vol.6.No.4. August, 2013, pp.81-92.

[14] X. Li, T. Zhang, X. Shen and J. Sun, "Object tracking using an adaptive Kalman filter combined with mean shift", Optical Engineering, vol. 2, no. 49, **(2010)**, pp. 020503-1-020503-3.

[15] R. Kehl and L. V. Gool, "Real-time pointing gesture recognition for an immersive environment", Proceedings of 6th IEEE International Conference on Automatic Face and Gesture Recognition, (2004), pp. 577-582.

Performance Analysis and Redistribution among RIPv2, EIGRP & OSPF Routing Protocol

Golap Kanti Dey, Md. Mobasher Ahmed,
Department of Applied Physics, Electronics & Communication Engineering
University of Chittagong, Chittagong-4331, Bangladesh
Email: d.k.golap@gmail.com,md.mobasher@gmail.com

Kazi Tanvir Ahmmed
Department of Applied Physics, Electronics & Communication Engineering
University of Chittagong, Chittagong-4331, Bangladesh
Email: tanvir@cu.ac.bd

Abstract— **In a network topology, it is very usual to use various kinds of routing protocol for forwarding packets. A routing table is used in the memory of a router that keeps the track of routes to particular network destination and the most popular routing algorithms used to forward packets are Routing Information Protocol (RIPv2), Enhanced Interior Gateway Routing Protocol (EIGRP) and Open Shortest Path First (OSPF). The ultimate concentration of this research work is to depict the performance analysis comparison of these three dynamic routing protocols and redistribution among the protocols. Eight Cisco routers and a switch are used in our simulated network topology where four routers with different protocols directly connected with the switch take the responsibility for the redistribution algorithm.**

Keywords— RIPv2, EIGRP, OSPF, redistribution, dynamic routing protocols etc.

I. INTRODUCTION

It is possible to exchange the routing information between routers through the routing protocols. Routing protocols allow routers to share information about remote networks dynamically and add this information to their routing tables automatically.

To recognize the best path to each network routing protocols are used and added to the routing table. The fundamental advantage of using dynamic routing protocol is that whenever there is topology change routers exchange routing information which permits routers to certainly learn about new networks as well as to find alternate paths if there is a link failure to a running network.

In comparison with static routing, less administrative overhead is required in dynamic routing protocols. However, the expense of using dynamic routing protocols is dedicating part of a router's resources for protocol operation including CPU time and network link bandwidth. Besides, to meet the demands of changing network requirements dynamic routing protocols have evolved over several years. Though several organizations have shifted towards more recent routing protocols such as Enhanced Interior Gateway Routing Protocol (EIGRP) and Open Shortest Path First (OSPF), many of the earlier routing protocols, such as Routing Information Protocol (RIP), are still in use today.Figure-1 shows the classification of dynamic routing protocols.

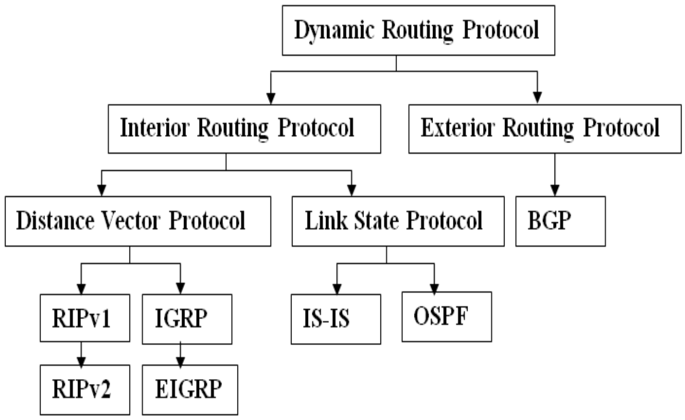


Fig. 1. Classification of Dynamic Routing Protocols

Since the early 1980s dynamic routing protocols have been used. Sheela Ganesh Thorenoor [1] used OPNET modeler for dynamic routing protocol implementation decision between EIGRP, OSPF and RIP. Multipath routing algorithm based on OSPF routing protocol [2] have been developed. Alex Hinds [3] did the evaluation for (OSPFv3) and (EIGRPv6) and compare the changes these protocols have undergone to support IPv6. Reference [4] worked on link recovery comparison between OSPF and EIGRP. Besides RIPv2 routing protocol authentication, discussed by Li Xiaohua [5], which can effectively prevent the router from receiving unauthorized or malicious routing updates, thereby improving network safety. Several investigation and research works are also conducting now-a-days by laureate researchers.

In our research work we will investigate comparative performance analysis of selected interior gateway dynamic routing protocols such as RIPv2, EIGRP and OSPF. Packet tracer simulation software is used here to show how to transmit data among different networks running different routing protocols by using route redistribution systems. Each of these dynamic routing protocols has different strengths and weaknesses- one protocol may have fast convergence, while another may be very reliable. In general dynamic routing has better scalability, robustness, and convergence. However, the cost of these added benefits include more complexity and some overhead -bandwidth that is used by the routing protocol for its own administration and route redistribution allows routes from one routing protocol to be advertised into another routing protocol.

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

II. Performance Analysis

A. Metrics

Routing protocols use metric value to decide which route is the best path. When a routing protocol learns of more than one route to reach the same destination a metric is a value which is used by routing protocols to identify costs to reach distant networks. To determine which path is most preferable when there are multiple paths to the same remote network the metric is used. Each of the routing protocol computes it's metric in a different way. For example, hop count is used in RIP, combination of bandwidth and delay are used in EIGRP, and the OSPF uses bandwidth.

1) RIPv2 - Hop count

For the metric value in RIPv2 protocol the Hop count is used. The hop count refers to the number of routers a packet must cross to reach the destination network. Best path is chosen by the route with the lowest hop count.

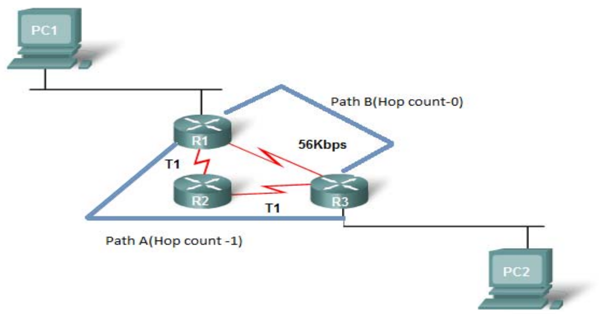


Fig. 2. Metric value for the RIPv2 protocol

To exchange the data from router R1 to R3 in the figure-2, there are two paths path A and path B. In the path A hop count is one and for path B hop count is zero. So path B is best path between R1 and R2.

2) EIGRP: Bandwidth, delay, reliability, and load

EIGRP uses the following values in its composite metric to calculate the preferred path to a network: Bandwidth, Delay, Reliability, Load. Best path is chosen by the route with the smallest composite metric value calculated from these multiple parameters. By default, only bandwidth and delay are used. The EIGRP routing[6] update takes the hop count into account though EIGRP does not include hop count as a component of composite metrics. The total delay and the minimum bandwidth metrics can be achieved from values which are put together on interfaces and the formula used to compute the metric is followed by:

Complete Composite Formula EIGRP have 5 composite to calculate: K1 = Bandwidth, K2 = Load, K3 = Delay, K4 = Reliability, K5 = MTU(Maximum Transmission Unit)

The weighting is as follows: K1=K3=1 and K2=K4=K5=0, Then substitutes all K parameters into the equation as follows:

Metric = 256*[K1*bandwidth + (K2*bandwidth) / (256 - load) + K3*delay] * [K5 / (K4 + reliability)]

Default Metric Calculation: Metric= 256*(BW + Delay).

3) OSPF: Cost

Best path is chosen by the route with the lowest cost. The Cisco implementation of OSPF uses bandwidth to determine the cost. The path cost of an interface in OSPF is called metric that indicates standard value such as speed. The cost of an interface is calculated on the basis of bandwidth. Cost is inversely proportional to the bandwidth. Higher bandwidth is attained with a lower cost [7].

$$\text{Cost}=\frac{10^8}{\text{Bandwidth in bps}}$$

Where the value of 10^8 is 100000000 in bps is called reference bandwidth based on by default.

B. Convergence

When the routing tables of all routers are at a state of consistency it is called convergence. The network is said to be converged when all routers have complete and accurate information about the network. Total time requires by the routers to calculate best paths, update their routing tables and to share information is known as convergence time. To make a network perfectly operable the network must be converged. Convergence is both collaborative and independent. The speed of propagation of routing information and the calculation of optimal paths are included in convergence properties. The faster the convergence, the better the routing protocol. Generally, RIP is slow to converge, whereas EIGRP, OSPF are faster to converge.

C. Network Throughput

In small and condensed networks RIPv2 has better performance than others. For medium-sized and scattered networks OSPF and EIGRP [8] show excellent execution. Overall in both small and relatively large networks EIGRP is more stable and balanced.

Traffic throughput of a network is regulated by the routing protocol in activities, and the hardware of routers, which is a key point for many network administrators. EIGRP utilizes the limited network bandwidth better than the OSPF. Based on the estimation of protocols performance [9], it can be said that EIGRP could function both as a distance vector and link state protocol. In comparison to OSPF through the intelligent use of metrics within the DUAL algorithm EIGRP performed better CPU utilization and bandwidth control. However the throughput of the protocols is comparable with simulation results [10] from finding OSPF providing greater network throughput than EIGRP, these results may differ due to the different network topology used in testing.

D. Preventing Routing Loops

Routing loops, which may be a short-lived, can be extremely harmful for the performance of a network. Hold-down timers and split horizon are used in RIPv2 to prevent routing loops. The key way that EIGRP check routing loops by using DUAL algorithm.

On the other hand OSPF doesn't have any special feature to avoid loops but its architecture is modeled in the way that it occupies instinctive loop prevention mechanism.

E. Authentication

The possibility of accepting invalid routing updates is a security concern of any routing protocol. An attacker could be the source of this invalid traffic malevolently trying to disturb the network or attempting to get packets by misleading the router into sending its updates to the wrong destination. A misconfigured router could be another source of invalid updates. Or may be a host is attached to the network and unknown to its user - the host is running the routing protocol of the local network. It is necessary to authenticate routing information transmitted between routers. To authenticate routing information, RIPv2, EIGRP, OSPF can be configured. This practice provide routers will only accept routing information from other routers that have been configured with the right password or authentication information.

F. When to use

1) RIPv2

Routing Information Protocol version 2 supports subnet mask and reduces broadcast load. It is validated for updates as well as used for multivendor environment.

2) EIGRP

Enhanced Interior Gateway Routing Protocol is used in very large and complex networks as well as for fast convergence.It supports VLSM, multiprotocol.

3) OSPF

Open Shortest Path First protocol is preferred for large hierarchical networks, fast convergence,complex networks, Multivendor and VLSM.

III. SIMULATION WORK

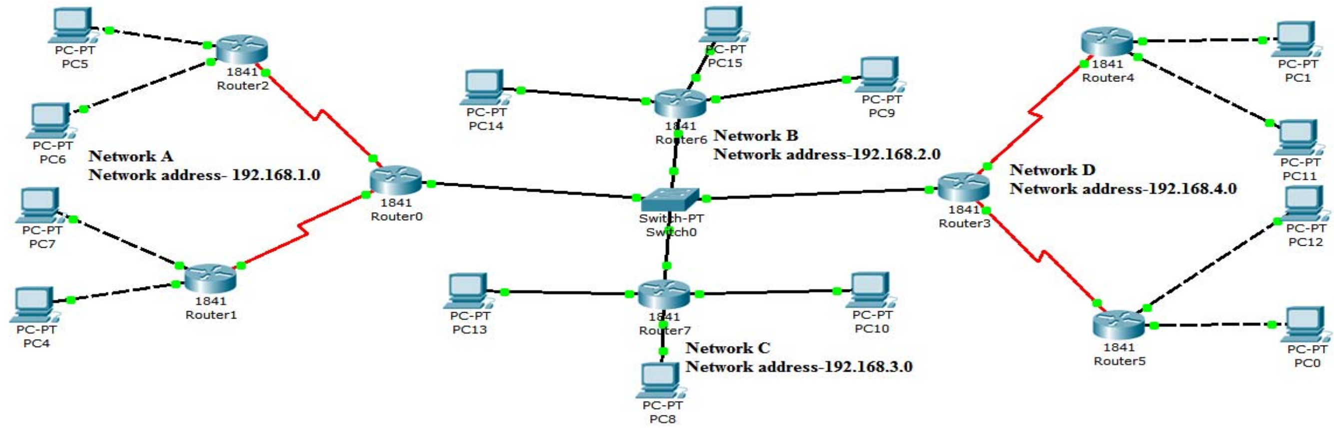


Fig.3. Simulated topology

In our simulated work we have used total eight routers where router 0, router 3, router 6, router 7 were directly connected with a switch.

The network A consists of router 0, router 1, router 2 with network address 192.168.1.0 performing EIGRP routing protocol. Network B and Network C containing network address 192.168.2.0 and 192.168.3.0 respectively performing RIPv2 routing protocol where Network D consists of router 3,router 4,router 5 containing network address192.168.4.0 respectively performing OSPF routing protocol . Within each individual network every end user can communicate with one another. But end users of two different networks can not transmit data among them. As for example- PC4 and PC6 of network A can ping each other but PC4 of network A cannot ping PC12 of network D. Now for successful communication between end users of different network s, running different networking protocols, route redistribution is used among router 0, router 3, router 6, and router 7.

IV. REDISTRIBUTION

The adaptation of a routing protocol to announce routes that are accomplished by some other means, for instance by another routing protocol, static routes, or directly connected routes, is called redistribution[11]. Multi−protocol routing is common for a number of reasons, such as company mergers, multiple departments managed by different network administrators and multi−vendor environments though running a single routing protocol throughout your entire IP internetwork is desirable. Running multiple routing protocols is often part of a network design. Redistribution is required for the environment of having multiple protocols.Through the router redistribution [12], routes from one routing protocol will be revealed into another routing protocol. Received redistributed routes are marked as external in the routing protocol. Logically-originated routes are usually more preferred than external routes.

A. Redistributing into RIP

Following command shows how a RIP router 6 in figure-3 redistributing EIGRP and OSPF.

```
router rip
network 192.168.2.0
redistribute eigrp 1
redistribute ospf 1
default−metric 1
```

We are using same of the above commands for redistributing EIGRP and OSPF into RIP router 7 except changing the network address 192.168.3.0. The RIP metric is composed of hop count, and the maximum valid metric for RIPv2 is 15. By defining a metric of 1, we can enable a route to travel the highest number of hops in the domain of a RIP. Though doing this raise the possibility of routing loops if there are several redistribution points and a router acquire knowledge about the network with a preferable metric from the redistribution point than from the original source. Hence it is necessary that the metric is neither too high, restraining it from being advertised to all the routers, or too low, guiding to routing loops when multiple redistribution points are presented.

B. Redistributing into EIGRP

EIGRP is a hybrid routing protocol that, by default, uses a composite of bandwidth and delay as its distance metric. EIGRP can additionally consider Reliability, Load, and MTU for its metric. An EIGRP router 0 in the figure-3 redistributing Open Shortest Path First (OSPF) and RIP through the commands-

```
router eigrp 1
network 192.168.1.0
redistribute ospf 1
redistribute rip
default−metric 10000 100 255 1 1500
```

TABLE I. EIGRP METRIC VALUES IN THE DEFAULT METRIC COMMAND

Metric	Value
Bandwidth	In units of kilobits per second; 10000 for Ethernet
Delay	In units of tens of microseconds; for Ethernet it is100 x 10 microseconds = 1 ms
Reliability	255 for 100 percent reliability
Load	Effective load on the link expressed as a number from 0 to 255 (255 is 100 percent loading)
MTU	Minimum MTU of the path; usually equals that for the Ethernet interface, which is 1500 bytes

C. Redistributing into OSPF

OSPF is a standardized Link-State routing protocol that uses cost ,based on bandwidth, as its link-state metric. To show an OSPF router 3 in the figure-3 redistributing RIP [13] and EIGRP we need -

```
router ospf 1
network 192.168.4.0 0.0.0.255 area 0
redistribute rip metric 200 subnets
redistribute eigrp 1 metric 100 subnets
```

The OSPF metric is a cost value based on 10^8/ bandwidth of the link in bits/sec. For example, the OSPF cost of Ethernet is 10: $10^8/10^7 = 10$. If a metric is not specified, OSPF puts a default value of 20 when redistributing routes from all protocols except Border Gateway Protocol (BGP) routes, which get a metric of 1.

V.CONCLUSION

Performance analysis of selected interior gateway dynamic routing protocols such as RIPv2, EIGRP and OSPF and their different performance issues have been investigated in this article. We have also presented a simulated work and the performance of redistribution command to establish communication between end users of different networks with different routing protocol. Route redistribution technology between diverse routing protocols has significant importance. Route redistribution is certainly easily realized and cost effective technique. Through using it we can also settle Tactical Internet Communication. Besides, comparative analysis among several routing protocol shows that the EIGRP protocol is better than the OSPF and RIPv2 routing protocol. But sometime EIGRP is held back by its proprietary features and costs. OSPF is better than other in large networks where its hierarchical nature increases scalability. And RIPv2 is useful in local and small area network. The redistribution command shows the way to communicate with different routing protocols.

REFERENCES

[1] Thorenoor, S.G. ,"Dynamic Routing Protocol Implementation Decision between EIGRP, OSPF and RIP Based on Technical Background Using OPNET Modeler", 2010 Second International Conference on Computer and Network Technology (ICCNT), pp. 191-195,2010.

[2] Li Lan , Li Li , Chen Jianya , "A Multipath Routing Algorithm Based on OSPF Routing Protocol" 2012 Eighth International Conference on Semantics, Knowledge and Grids (SKG), pp. 269-272,2012.

[3] Alex Hinds, Anthony Atojoko, Shao Ying Zhu ," Evaluation of OSPF and EIGRP Routing Protocols for IPv6", International Journal of Future Computer and Communication, Vol. 2, No. 4, August 2013.

[4] Ittiphon krinpayorm , Suwat Pattaramalai," Link Recovery Comparison Between OSPF & EIGRP",International Proceedings of Computer Science & Information Technology, Vol. 27,pp.192-197,2012.

[5] Li, Xiaohua, *et al.* ," Simulation and analysis of RIPv2 routing authentication based on GNS" International Conference on Automatic Control and Artificial Intelligence (ACAI 2012), pp. 1842-1845,2012.

[6] Y.Navaneeth Krishnan , Dr Shobha G , "Performance Analysis of OSPF and EIGRP Routing Protocols for Greener Internetworking" , Proceedings of International Conference on Green High Performance Computing ,pp. 1-4, 2013.

[7] Rick Graziani , Allan Jonson, "Routing protocols and concepts: CCNA exploration companion guide" Pearson Education. London,2008.

[8] Alex Hinds, Anthony Atojoko, and Shao Ying Zhu, "Evaluation of OSPF and EIGRP Routing Protocols for IPv6", International Journal of Future Computer and Communication, Vol. 2, No. 4, August 2013.

[9] S. G. Thorenoor, "Dynamic routing protocol implementation decision between EIGRP, OSPF and RIP based on technical background using OPNET modeller," Proceedings of Second International Conference on Computer and Network Technology, China, pp. 191-195,2010.

[10] M. N. Islam, M. A. U. Ashique,"Simulation based EIGRP over OSPF performance analysis," Master Thesis in Electrical Engineering Emphasis on Telecommunications, No. 4983, Blekinge Institute of Technology, 2010.

[11] Franck Le ,JL Sobrinho,"Interconnecting Routing Instances", IEEE/ACM Transactions on Networking , Vol. 22, Issue. 2, April 2014.

[12] Liang Zhijian," The New Data Exchange method in Different Networks", 2010 International Coriference on Educational and Information Technology (ICEIT 2010),Vol. 2,pp. 69-72, 2010.

[13] Franck Le, Geoffrey G. Xie, Hui Zhang, "Understanding Route Redistribution",IEEE International Conference on Network Protocols (ICNP),pp. 81-92, October 2007.

Phonetic Features Enhancement for Bangla Automatic Speech Recognition

Sharif M Rasel Kabir, Foyzul Hassan, Foysal Ahamed, Khondokar Mamun, Mohammad Nurul Huda
Department of Computer Science and Engineering
United International University
Dhaka, Bangladesh
shuvro.mist@gmail.com, foyzul.hassan@gmail.com, thisisfoysal@gmail.com, mamun@cse.uiu.ac.bd, mnh@cse.uiu.ac.bd

Fariha Nusrat
Department of Computer Science and Engineering,
Military Institute of Science and Technology
Dhaka, Bangladesh
fariha1502@gmail.com

***Abstract*— This paper discusses a phonetic feature (PF) based automatic speech recognition system (ASR) for Bangla (widely known as Bengali), where the PF features are enhanced. There are three stages in this method where the first step maps Acoustic Features (AFs) or Local Features (LFs) into Phonetic Features (PFs) and the second step incorporates inhibition/enhancement (In/En) algorithm to change the PF dynamic patterns where patterns are enhanced for convex patterns and inhibited for concave patterns. The final step is for normalizing the extended PF vector using Gram–Schmidt algorithm and then passing through a Hidden Markov Model (HMM) based classifier. In our experiment on speech corpus for Bangla, the proposed feature extraction method provides higher sentence correct rate (SCR), word correct rate (WCR) and word accuracy (WA) compared to the methods that not incorporated In/En network.**

Keywords—Automatic Speech Recognition; Acoustic Feature; Hidden Markov Model; Inhibition/Enhancement Network; Local Feature; Multilayer Neural Network; Phonetic Feature

I. INTRODUCTION

Although Bangla (widely known as Bengali) is one of the largely spoken languages in the world, but very few literatures found in automatic speech recognition (ASR) for Bangla. Almost 250 million total speakers of Bangla language which make this ranked seventh in the world [1]. A major difficulty to research in Bangla ASR is lack of proper speech corpus, though some efforts are made to develop Bangla speech corpus to build a Bangla text to speech system [2]. For research, the myriad of phonological variations for non-standard dialects were given in [3].

The developments of various hidden Markov model (HMM) based ASR systems can be found in [4]-[8], where Bangla speech processing or Bangla ASR are found developed to some extent. In this case a preprocessed form is used most of the time by these ASR systems, like mel-frequency cepstral coefficients (MFCCs) of the speech signal, which encodes the time frequency distribution of signal energy. However, in real acoustic conditions these MFCC based systems do not provide better recognition performance (See Figure 1(a)). On the contrary, an articulatory features based system (AFs) or phonetic features (PFs) exhibits a higher recognition which is accurate in practical conditions, and models coarticulatory phenomena more naturally [7] (See Figure 1(b)). Few misclassification are output by the PF-based system and it is shown in Figures 1(a) and 1(b).

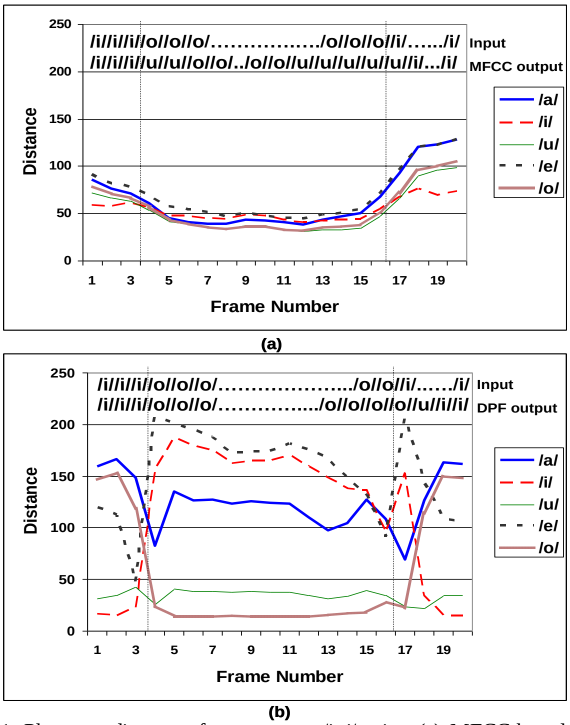


Fig. 1. Phoneme distances for utterance, /ioi/ using (a) MFCC-based system and (b) PF-based system.

A phonetic feature (PF) extraction method is being proposed in this research for use in a Bangla continuous word recognition with high performance in an automatic speech recognition (ASR) system. The method incorporated three stages. A multilayer neural network (MLN) is used by the first stage and it maps continuous acoustic features, local features (LFs), onto PFs. Next is the second stage and it embeds an inhibition/enhancement (In/En) functionalities to discriminate whether the PF dynamic patterns of trajectories are convex or concave. Here convex and concave patterns are enhanced and inhibited respectively. The last one is the third stage, and it normalizes the PF vector before inserting into a hidden Markov model (HMM) based classifier. We prepared an experiment on medium scale speech corpus and the proposed feature extractor was found to provide a higher sentence correct rate (SCR), also word correct rate (WCR) and word accuracy (WA) using context dependent triphone-based HMM which is compared with the method not incorporated In/En network. Besides conventional MFCCs are used in the design of experiments instead of LFs with and without incorporating In/En network. For the In/En network ASR performance is improved as it provides better classification for HMM by enhancing convex pattern and also by inhibiting concave pattern.

II. Phonetic Scheme and Corpus for Bangla

A phoneme can easily be identified by using its unique Distinctive Phonetic Features (DPFs) set [8].

A. *Bangla Phonemes*

There are 14 vowels, which includes seven nasalized vowels and there are 29 consonants in the phonetic inventory of Bangla. In IPA an approximate phonetic scheme is given in [9][10], here only the main 7 vowel sounds are shown, though there are two more long counterpart of /i/ and /u/ denoted as /i:/ and /u:/ respectively.

Table 1 shows the different Bangla words with IPA where the same /a/ has different pronunciation based on succeeding phonemes. Sometimes these pronunciations are long or short. For long and short /a/, two different phonemes /aa/ and /ax/ have been used respectively. Likewise all variations of same phonemes have been considered and consequently, total 51 phonemes are found which does not include beginning end silence (/sil/) and short pause (/sp/).

TABLE I. Some Bangla words with their IPA.

Bangla Word	English Pronunciation	IPA	Our Symbol
আমরা	AAMRA	/a m ɾ a/	/aa m r ax/
আচরণ	AACHORON	/a tʃ ɾ n/	/aa ch ow r aa n/
আবেদন	ABEDON	/a b æ d̪ n/	/ax b ae d aa n/

B. *Bangla Speech Corpus*

From the Bangla newspaper “Prothom Alo” [11] hundred sentences are uttered by 30 male speakers of different regions of Bangladesh are used as training corpus (D1). Likewise from the same newspaper 100 different sentences uttered by 10 different male speakers are used as test corpus (D2). These speakers are all Bangladeshi nationals and native speakers of Bangla: Dhaka (central region), Comilla – Noakhali (East region), Rajshahi (West region), Dinajpur – Rangpur (North-West region), Khulna (South-West region), Mymensingh and Sylhet (North-East region). The age of the speakers ranges from 20 to 40 years.

III. Bangla Phonetic Features

We can easily identify a phoneme by its PFs [12]. That’s why for identification the Bangla PFs for all the phonemes with their international phonetic alphabet (IPA) and also Bangla orthographic transcription were given in [8]. The fifty three Bangla phonemes and their twenty two PFs were also mentioned.

IV. Proposed ASR System using PFs

A PF-based ASR system with an input acoustic vector of LFs using an MLN has been implemented which is shown in Figure 2. This system comprised of three stages: i) input acoustic features, LFs extraction [13] and feeds these LFs to an MLN for extracting 22 PFs these all are dealt in the first stage, ii) the second stage implants an inhibition/enhancement (In/En) functionalities and it discriminates whether the PF dynamic patterns of trajectories are convex or concave, where convex and concave patterns are enhanced and inhibited respectively, and iii) the last stage integrates a triphone-based HMM for generating the output text strings and here we input normalized inhibited/enhanced 22 PF values. The twenty five dimensional acoustic features, LFs which was extracted in the first stage are entered into the MLN with five layers and it also includes three more hidden layers after combining a current frame x_t it is with the other the frames that are three points before and after the current frame (x_{t-3}, x_{t+3}) where the MLN generates twenty two PF values for each input frame of 25x3 features. A total of 400, 200 and 100 units are comprised the three hidden layers. The standard back-propagation algorithm is used for the MLN training.

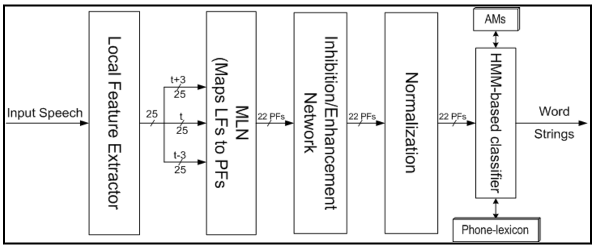


Fig. 2. Proposed PF-based ASR System.

A. *Inhibition/Enhancement Network*

The PF extractor which is actually neural network based is applied in this research. It mainly generates 22 PF patterns for each input frames of 25x3 each (here 25 for current, 25 for proceeding and 25 for succeeding frames). Because these 22 PF patterns may not follow the input PF patterns of a phoneme string accurately, there is an uncertainty among some phonemes for classifying the phoneme which was targeted in the HMM-based classifier. As a result, there are some phonemes which are not correctly recognized. An uncertainty takes place when the value of consecutive PF peaks and PF dips in a PF time pattern of a phoneme string are closer to each other. For example, the values of left peak, middle dip, and right peak which are generated by a PF extractor are <0.7, 0.7, 0.7, 0.7>, <0.4, 0.4, 0.4>, and <0.7, 0.7, 0.7>, respectively. Therefore the classifier faces a problem to decide whether this pattern is <1, 1, 1, 1, 0, 0, 0, 1, 1> or <1, 1, 1, 1, 1, 1, 1, 1, 1, 1>, while input PF pattern was <1, 1, 1, 1, 0, 0, 0, 1, 1, 1>. Here, the value for 0.4 and 0.7 are not same. Value 0.4 is assumed as either zero or one, whereas we consider 0.7 as one. So there are some obvious differences between a PF peak and dip along time axis. If there is any mechanism that improves PF peak values up to a certain level and which can suppress PF dip values accordingly, then we can find a difference between a peak and dip. To acquire this type of effect we have incorporated an In/En network.

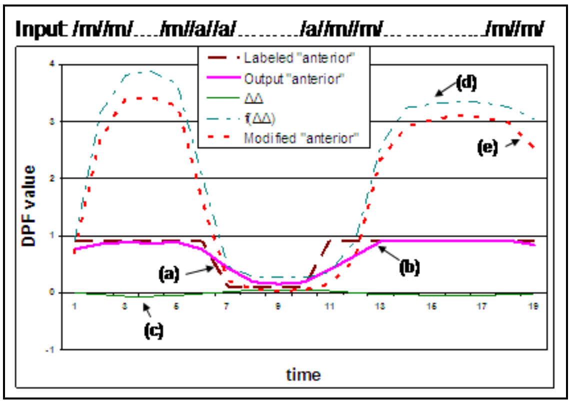


Fig. 3. Working mechanism of the In/En network. Five curves are denoted by (a), (b), (c), (d), and (e), respectively. The curves: a) Labeled "anterior" PF for input utterance, /mam/, b) Output "anterior" PF by MLN, c) ΔΔ for output "anterior", d) f(ΔΔ) for ΔΔ, and e) Modified "anterior" by multiplying curve (b) with curve (d).

Step1: For each element of the PF vector, find the acceleration ΔΔ parameters by using three-point Linear Regression (LR).

Step2: Check whether ΔΔ is positive (concave pattern) or negative (convex pattern) or zero (steady state).

Step3: Calculate f(ΔΔ).

if pattern is convex,

$$f(\Delta\Delta) = \frac{c_1}{1 + (c_1 - 1)e^{\beta\Delta\Delta}}$$

if pattern is concave,

$$f(\Delta\Delta) = c_2 + \frac{2(1 - c_2)}{1 + e^{\beta\Delta\Delta}}$$

if steady state,

$$f(\Delta\Delta) = 1.0$$

Step4: Find modified PF patterns by multiplying the PF patterns with f(ΔΔ).

In Figure 3 we can see the working mechanism of the In/En network by using the "anterior" PF pattern of an input utterance, /mam/ along time axis. In the figure mainly five curves are presented and they represent labeled "anterior" PF for the input utterance, corresponding output "anterior" generated by the MLN, ΔΔ for the output "anterior" values, f(ΔΔ) for ΔΔ values, and modified "anterior" PF, are indicated by (a), (b), (c), (d), and (e), respectively. Here the curve (e) is mainly found by multiplying curve (b) with curve (d). after implementing the In/En network algorithm on curve (b), the PF values of frames 1-6 and 13-19 (convex pattern or PF peak) are improved, and frames 7-11 (concave pattern or PF dip) are inhibited. Here the PF pattern of curve (e) shows a clear distinction between a PF peak and dip which should be noted, and hence, categorizes the PF movement more correctly.

V. EXPERIMENTS

A. *Setup*

For evaluating word recognition performance using an HMM-based classifier correct rate (WCR), word accuracy (WA) and sentence correct rate (SCR) for D2 data set are evaluated. To design Bangla triphone HMMs with five states, three loops, and left-to-right models the D1 data set is used. Input features for the classifier are 39 dimensional MFCCs (12MFCC, 12ΔMFCC, 12ΔΔMFCC, P, ΔP, ΔΔP where P represents raw energy of the input speech signal) and log values of 22 dimensional PFs. The mixture components are set to 1, 2, 4 and 8.

We have designed the following experiments for evaluating the performance of standard MFCC-based method which includes the proposed method.

a) MFCC:dim-39 [Baseline]

b) PF(MFCC) +log10:dim-22 [8]

We have designed following two experiments for comparison purposes using normalized PFs of 22 dimensions. These feature vectors which are normalized, they are inserted into HMM-based classifier.

c) PF(MFCC)+Norm:dim-22

d) PF(LF)+Norm:dim-22

Finally, including In/En network other sets of experiments are designed. Then normalized In/En network is inputted to HMM-based classifier.

e) PF(MFCC)+In/En+Norm:dim-22

f) PF(LF)+In/En+Norm:dim-22 [Proposed]

The non-linear function is a sigmoid from 0 to 1 (1/(1+exp(-x))) for the hidden and output layers, In our experiments of the MLN.

After evaluating some methods for different values of C1, such as 2, 4, and 6, and the value of the steepness coefficient, β, is set to 80, the value of the enhancement coefficient, C1, is set to 4.0 for the In/En network. To keep the values of f (ΔΔ) between 0.25 and 1.0 after observing the PF data patterns the value of inhibitory coefficient, C2, is fixed to 0.25.

B. *Result Analysis and Discussion*

For experiments (a) and (b), evaluation results are given of them in [8]. From [8], we observe that PF-based methods can provide higher WCA, WA and SCR for all mixture components examined except two.

In Table 2 we see the comparison of word correct rates among the examined methods, (b) [8], (c), (d), (e) and proposed (f). Here all the components are combined and among them, the method which was proposed shows higher accuracy in comparison with all other experiments. From experiment (c) and (d), and (e) and (f) we can see that LF-based method is performed better in ASR compared to MFCC-based method. It is mainly because local features incorporate both time and frequency whereas MFCC are in time domain only. Again from experiment (c) and (e), and (d) and (f) we observe that the inclusion of In/En network increases WCRs which is in a certain percentage. It may be stated that these enhancement for the case of MFCC features are remarkable though corresponding increment is near about one percent for LFs.

TABLE II. WORD CORRECT RATE FOR CLOSE (USING TRAINING FILES) AND OPEN (USING TEST FILES) TEST.

Word Correct Rate(%)						
		(b)	(c)	(d)	(e)	(f)
Train	mix1	90.79	89.75	98.01	96.69	99.33
	mix2	89.31	90.54	97.97	95.88	99.39
	mix4	92.90	91.37	98.53	96.54	99.59
	mix8	94.58	92.37	98.80	97.29	99.70
Test	mix1	89.73	88.69	97.14	96.29	98.27
	mix2	88.39	88.72	97.26	95.62	98.57
	mix4	92.19	89.06	96.99	95.99	98.02
	mix8	92.25	89.03	96.96	96.72	98.39

TABLE III. WORD ACCURACY FOR CLOSE (USING TRAINING FILES) AND OPEN (USING TEST FILES) TEST.

Word Accuracy (%)						
		(b)	(c)	(d)	(e)	(f)
Train	mix1	90.43	89.40	97.86	96.55	99.22
	mix2	88.84	90.22	97.79	95.72	99.31
	mix4	92.48	91.00	98.41	96.42	99.55
	mix8	94.21	92.05	98.62	97.21	99.66
Test	mix1	89.45	88.42	96.78	95.96	98.09
	mix2	88.02	88.36	96.87	95.35	98.39
	mix4	91.43	88.39	96.60	95.62	97.81
	mix8	91.64	88.39	96.53	96.50	98.18

TABLE IV. SENTENCE CORRECT RATE FOR CLOSE (USING TRAINING FILES) AND OPEN (USING TEST FILES) TEST.

Sentence Correct Rate (%)						
		(b)	(c)	(d)	(e)	(f)
Train	mix1	90.40	89.37	97.90	96.63	99.27
	mix2	89.17	90.27	97.83	95.77	99.33
	mix4	92.60	90.93	97.83	96.37	99.60
	mix8	94.17	91.77	98.60	97.17	99.70
Test	mix1	89.20	88.00	96.20	95.30	97.70
	mix2	88.20	87.80	96.10	95.10	97.90
	mix4	91.50	87.80	95.90	95.70	97.20
	mix8	91.60	87.60	96.10	96.30	97.90

Here in Table 3, it compares the accuracy of word among the examined method, (b) [8], (c), (d), (e) and proposed, (f). For this evaluation, the method which was proposed shows a performance which is higher compared to all other examined methods. Here in this table (from (c) and (d), and (e) and (f)) it also shows that LFs are better feature for MLN than MFCCs. Moreover, from experiments, (c) and (e), and (d) and (f) we found that the Inhibition/Enhancement of phonetic features in our system provides higher accuracy (brilliant performance for MFCCs and approximately one percent enhancement for LFs) which is for both MFCCs and LFs based method.

Finally, from Table 4 we can get an idea about the sentence correct rates of all the methods we have examined, (b) [8], (c), (d), (e) and proposed, (f) except bottom line. We get much improved performance by (d) and (f) over (c) and (e) due to the acoustic characteristic, LFs in MLN. Here, in experiments (c) and (e), and (d) and (f) the sentence correct rates are also improved because of the Inhibition/Enhancement of phonetic features in our systems. As a result, by the proposed method the highest sentence correct rates among all other experimented methods are obtained.

VI. CONCLUSION

An ASR system for Bangla incorporating Inhibition/Enhancement network has been developed in this paper. The paper concludes the following.

i) WCR, WA and SCR using the proposed method are higher than the MFCC-based method.

ii) LF is better input feature for MLN than MFCC

iii) The performances of ASR are tremendously increased by the incorporation of Inhibition/Enhancement network.

The author would like to design a real time system for hearing impaired people in Bangladesh incorporating the ASR proposed in this study in near future.

REFERENCES

[1] http://en.wikipedia.org/wiki/List_of_languages_by_total_speakers.

[2] S. P. Kishore, A. W. Black, R. Kumar, and Rajeev Sangal, "Experiments with unit selection speech databases for Indian languages," Carnegie Mellon University.

[3] http://en.wikipedia.org/wiki/Bengali_phonology.

[4] S. A. Hossain, M. L. Rahman, and F. Ahmed, "Bangla vowel characterization based on analysis by synthesis," Proc. WASET, vol. 20, pp. 327-330, April 2007.

[5] M. A. Hasnat, J. Mowla, and Mumit Khan, " Isolated and Continuous Bangla Speech Recognition: Implementation Performance and application perspective, " in *Proc. International Symposium on Natural Language Processing* (*SNLP*), Hanoi, Vietnam, December 2007.

[6] A. K. M. M. Houque, "Bengali segmented speech recognition system," Undergraduate thesis, BRAC University, Bangladesh, May 2006.

[7] K. Kirchhoff, et. al, "Combining acoustic and articulatory feature information for robust speech recognition," Speech Commun.,vol.37, pp.303-319, 2002.

[8] Foyzul Hasan, Mohammed Rokibul Alam Kotwal and Mohammad Nurul Huda, "Bangla Phonetic Feature Table Construction for Automatic Speech Recognition," 16th International Conference on Computer and Information Technology (ICCIT 2013), Khulna, Bangladesh, Dec. 22-24, 2013, pp: 51-55.

[9] C. Masica, the Indo-Aryan Languages, Cambridge University Press, 1991.

[10]Ghulam Muhammad, Yousef A. Alotaibi and Mohammad Nurul Huda, "Automatic Speech Recognition for Bangla Digits," ICCIT'09, Dhaka, Bangladesh, December 2009.

[11]Daily Prothom Alo. Online: www.prothom-alo.com

[12] S. King and P. Taylor, "Detection of Phonological Features in Continuous Speech using Neural Networks," Computer Speech and Language 14(4), pp. 333-345, 2000.

[13] T. Fukuda and T. Nitta, "Noise-robust ASR by Using Distinctive Phonetic Features Approximated with Logarithmic Normal Distribution of HMM," Proc. Eurospeech 2003, Vol.III, pp.2185-2188, Sep. 2003.

Bangla Pronunciation Error Detection System

Selina Parveen, Sharif Mohammed Rasel Kabir, Khondokar Mamun, Mohammad Nurul Huda
Department of Computer Science and Engineering
United International University
Dhaka, Bangladesh
selina.parveen.aj@gmail.com, shuvro.mist@gmail.com, mamun@cse.uiu.ac.bd, mnh@cse.uiu.ac.bd

Farhana Sarker
Department of Computer Science and Engineering,
University of Liberal Arts
Dhaka, Bangladesh
farhana.sarker@ulab.edu.bd

***Abstract*— In this paper a pronunciation error detection system was modeled and also tested for large vocabulary, speaker independent and continuous speech recognizer for Bengali language. The recognizer was developed using Hidden Markov Model (HMM); and the Hidden Markov Modeling Toolkit was used to implement it. In the process, a corpus database comprised of 3000 utterances that were used for training and 100 plus sentences for development and evaluation. The data was preprocessed in line with the requirements of the HTK toolkit. In order to support the acoustic models, a bigram language model was constructed. In addition, pronunciation dictionary was prepared and used as an input. Standard experiment for in depth performance analysis of the detector was designed and all the results are analyzed to get the proper insight of the error pattern, both at sentence and word level. The effect of Conjugant words, the effect of first letter of the word, the gender effect on error pattern along with the effect of colloquial dialect or local accent on error pattern is observed. The findings are quite promising and may open new possibilities to design an efficient pronunciation error correcting system for aiding the non-native Bengali speaking people.**

Keywords*— automatic speech recognition; pronunciation error detection; error pattern for local accent; gender effect; conjugant word; sentence correct rates; hidden Markov model

I. INTRODUCTION

A pronunciation detector for Bangla language is very useful tool that will help us to detect the errors those are introduced due to local accent by native Bangla speakers as well as non-native Bangla speakers. The pronunciation detector is based on a robust HMM automatic speech recognition (ASR) system. Various methods have been used to obtain robust ASR system; however, the ASR system that shows satisfactory performance at anytime and everywhere could not be realized now. One of the reasons is that the acoustic models (AMs) of an HMM-based classifier include many hidden factors such as speaker-specific characteristics that include gender types and speaking styles [1-3]. It is difficult to recognize speech affected by these factors, especially when an ASR system comprises only a classifier that made its training by a single type of gender. One solution is to employ a acoustic model for both types of gender. Though the robustness of this acoustic model by utilizing the both gender specific characteristic is limited, but it resolves the gender effects more precisely. On the other hand, only a very few works have been done in ASR for Bangla (can also be termed as Bengali) in spite of one of the largely spoken languages in the world. More than 220 million people speak in Bangla as their native language. It is ranked seventh based on the number of speakers [4]. A major difficulty to research in Bangla ASR is the lack of proper speech corpus. Some developments on Bangla speech processing or Bangla ASR can be found in [7-14]. For example, Bangla vowel characterization is done in [7]; isolated and continuous Bangla speech recognition on a small dataset using hidden Markov models (HMMs) is described in [8]. Again, Bangla digit recognition was found in [15].

In this paper, to evaluate the performance of the pronunciation detector, the HMM based classifier was trained with sufficiently large number of male and female samples from different areas of Bangladesh. Then a well balanced speech corpus of moderate size is used to detect the error in pronunciation comprising of 10 male samples and 10 female samples from different areas of Bangladesh are tested for detection of errors due to local accent and gender effect. In addition to these detections, the effect of conjugant words along with the effect of the first letter of the word is observed both at word level and sentence level.

This paper is organized as follows. Sections II discusses Bangla phoneme schemes, Bangla speech corpus and triphone model. Then the Section III outlines mel frequency cepstral coefficients (MFCCs) extraction procedure. In Section V the proposed pronunciation error detecting technique is explained. Section VI describes an experimental setup, and section VIII analyzes experimental results. Finally, section IX concludes the paper with some future remarks.

II. BANGLA PHONEME SCHEME, TRIPHONE DESIGN AND BANGLA SPEECH CORPUS

Bangla phonetic scheme and IPA (International Phonetic Alphabet) for Bangla were described in [16]. The paper [16] also showed characteristics of some Bangla words by using the spectrogra and triphone model based on HMM were also analyzed for Bangla words.At present, a real problem to do experiment on Bangla phoneme ASR is the lack of proper Bangla speech corpus. In fact, such a corpus is not available or at least not referenced in any of the existing literature.

978-1-4673-8342-4/15/$31.00

Therefore, we develop a medium size Bangla speech corpus, which is described below.

Hundred sentences from the Bengali newspaper "Prothom Alo" are uttered by 30 male speakers of different regions of Bangladesh. These sentences (30x100) are used as male training corpus (D1). On the other hand, 3000 same sentences uttered by 30 female speakers are used as female training corpus (D2).On the other hand, different 100 sentences from the same newspaper uttered by 10 different male speakers and by 10 different female speakers are used as male test corpus (D3) and female test corpus (D4), respectively. All of the speakers are Bangladeshi nationals and native speakers of Bangla. The age of the speakers ranges from 20 to 40 years. We have chosen the speakers from a wide area of Bangladesh: Dhaka (central region), Comilla – Noakhali (East region), Rajshahi (West region), Dinajpur – Rangpur (North-West region), Khulna (South-West region), Mymensingh and Sylhet (North-East region). Though all of them speak in standard Bangla, they are not free from their regional accent.Recording was done in a quiet room located at United International University (UIU), Dhaka, Bangladesh. A desktop was used to record the voices using a head mounted close-talking microphone. We record the voice in a place, where ceiling fan and air conditioner were switched on and some low level street or corridor noise could be heard. Jet Audio 7.1.1.3101 software was used to record the voices. The speech was sampled at 16 kHz and quantized to 16 bit stereo coding without any compression and no filter is used on the recorded voice.

III. PROPOSED SYSTEM

Figure 1 shows the proposed MFCC-based pronunciation detector for Bangla Language. Here, an input speech is converted into MFCCs of 39 dimensions (12-MFCC, 12-ΔMFCC, 12-ΔΔMFCC, P, ΔP and ΔΔP, where P stands for raw energy of the input speech signal), hamming window of 25 ms is used for extracting the feature. The value of pre-emphasis factor is 0.97. Then, these extracted MFCCs are used to train the classifier based on triphone HMM, while the Viterbi algorithm is used for evaluating the test data set for male and female.

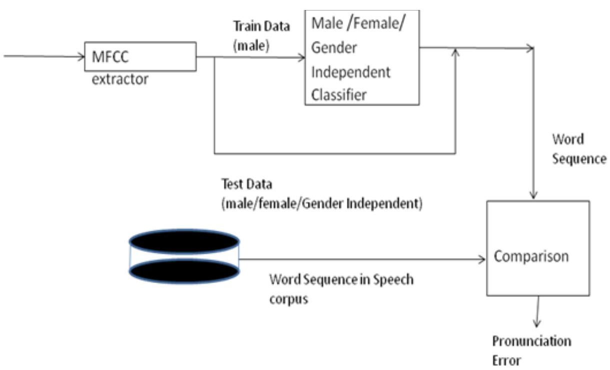


Fig. 1. The Proposed Pronunciation Error Detection System.

In the proposed model, the MFCC extractor followed by the HMM based classifier works together to detect the word sequence. The detected word sequences are then compared with the original word sequences in the speech corpus to detect the error pattern.

IV. EXPERIMENTAL SETUP

For evaluating the performance of Pronunciation Detector using proposed method, we have designed the following experiments:

A. *Experiment-I [Exp-I]*

- MFCC (Train: 3000 male, Test: 1000 male)
- MFCC (Train: 3000 male,Test: 1000 female)
- MFCC (Train 3000 male, Test:10 group of 100 male each)
- MFCC (Train: 3000 male, Test:10 group of 100 female each)

B. *Experiment-II [Exp-II]*

- MFCC (Train: 3000 female, Test: 1000 male)
- MFCC (Train: 3000 female,Test: 1000 female)
- MFCC (Train 3000 female, Test:10 group of 100 male each)
- MFCC (Train: 3000 female, Test:10 group of 100 female each)

C. *Experiment-III [Exp-III]*

- MFCC (Train: 3000 male+3000 female, Test: 1000 male)
- MFCC (Train: 3000 male + 3000 female Test: 1000 female)
- MFCC (Train: 3000 male + 3000 female, Test:10 group of 100 male each)
- MFCC (Train: 3000 male + 3000 female, Test:10 group of 100 female each)

V. EXPERIMENTAL RESULT AND ANALYSIS

We will present the results in following fashion for better comparison as well as better clarifications.

- The results obtained for 1000 male test data for three different experiments Exp-I(A), Exp-II(A) and Exp-III(A)
- The results obtained for 1000 female test data for three different experiments Exp-I(B), Exp-II(B) and Exp-III(B)
- The results obtained for 10 group of 100 male each as test data for three different experiments Exp-I(C), Exp-II(C) and Exp-III(C)
- The results obtained for 10 group of 100 male each as test data for three different experiments Exp-I(D), Exp-II(D) and Exp-III(D)

From the Table I it's obvious that, to minimize the gender effect the detector should be trained with a large set of both male and female data. Then the pronunciation error detector carries out detailed analysis of error pattern both in sentence level and word level to find out the error prone words which will surely give us some idea how to minimize such errors. We will here present the data for Exp I (A) which gives the best results both in sentence level and word level in Table II. The conjugant words are underlined in the Table II.

Then the effects of conjugant words or conjuncts are summarized in the tables III and IV. From these two tables it is evident that, the effect of conjugant on sentence level is quite noteworthy, while not very significant at word level. Noticeable error rate for words starting with B/BH, P, S etc is observed in Figure 2.

TABLE I. THE RESULTS OBTAINED FOR 1000 MALE TEST DATA FOR THREE DIFFERENT EXPERIMENTS EXP-I(A), EXP-II(A) AND EXP-III(A)

Experiment	**Sentence recognition performance (out of 1000)**		**Word recognition performance (out of 3590)**			
	Correctly recognized Sentence, H	**Substitution, S**	**Correctly recognized Words, H**	**Deletion, D**	**Substitution, S**	**Insertion, I**
Exp-I (A) Train: 3000 male	**937**	**63**	**3105**	**40**	**145**	**17**
Exp-II (A) Train: 3000 female	**642**	**358**	**2126**	**343**	**821**	**24**
Exp-II (A) Train: 3000 male+3000 female	**826**	**174**	**2745**	**137**	**408**	**19**

TABLE II. LIST OF WORDS IN ERROR ALONG WITH NUMBER OF ERRORS AND FREQUENCY OF OCCURRENCES IN THE BRACKETS IN EXP_I(A).

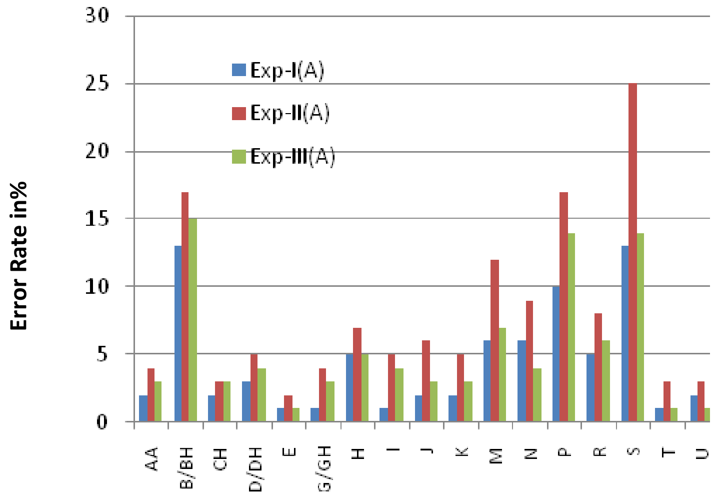
AAGUN(1/10)	CHITRO (1/30)	KARI(5/10)	PHONE(7/30)	SHAPER (1/10)
AALOCHONA (5/20)	CHOLO (1/30)	KARJOKROM (1/10)	PORAMORSHO (1/10)	SHIKHKHA (1/10)
BAABOWSTHAY (5/10)	DESHE (5/10)	MAA (1/10)	POROMANU (6/30)	SINDIKET (2/10)
BABSHA (3/10)	DHAKAY (5/50)	MANER (1/10)	POSHAK (1/10)	SMRITI (1/10)
BAD(1/20)	DITE (1/20)	MEYE(1/10)	POTRIKAY (2/10)	SOBHA(1/10)
BANGLA (1/10)	EGIYE (1/40)	MILI (5/10)	PRAN (3/20)	SOINIK (2/40)
BANGLAY (1/10)	GREPTAR (1/10)	MOBAIL (9/40)	PROSHIKKHON (7/40)	SONGGOTHON (1/10)
BHINNOMOT (1/10)	HASON(1/10)	MUKHI(1/10)	PROTIBEDON (1/10)	SONGIT (2/10)
BIDDUT(1/10)	HIMSHIM(1/10)	NASHER(20)	PURONO(4/10)	SORBORAHE(2/10)
BIDESHI(1/10)	HOBE(1/10)	NIHOTO(1/10)	RAJA(1/10)	SROMIKER(1/10)
BISHAL(1/10)	HRID(1/10)	NIONTRONE(1/10)	RAJPORIBARKE (5/10)	SUNDORBONE (1/10)
BISSHO(1/40)	HUMKI(3/20)	NIRAPOTTA(2/10)	ROBINDRO(2/10)	SURER(1/10)
BONDHO(9/60)	INTERNET (1/10)	NIRBHORJOGGO (1/10)	ROGIDER(1/10)	TERHIN(1/10)
BONDUK(1/10)	JHUDDHE(1/10)	NOTUN(1/10)	ROPTANI(1/30)	UTPADON(1/40)
BRITTI(3/20)	JONOSARTHE (2/10)	PATKOL (9/40)	SANGBADIKDER (1/10)	UTSAB(1/10)

TABLE III. EFFECT OF CONJUNCTS/CONJUGANT WORDS ON WORD ERROR. IN EXPERIMENT WITH 1000 MALE DATA

Experiment	**Error Words**	**Word with conjugant**	**Other words**
Exp-I (A) Train: 3000 male	78	48.7%	51.3%
Exp-II (A) Train: 3000 female	134	51.4%	48.6%
Exp-II (A) Train: 3000 male+3000 female	89	50.1%	49.9%

TABLE IV. EFFECT OF CONJUNCTS/CONJUGANT WORDS ON ERROR SENTENCE FOR 1000 MALE TEST DATA.

Experiment	Error Sentences	Sentence having words with conjugant	Other sentences
Exp-I (A) Train: 3000 male	63	84%	16%
Exp-II (A) Train: 3000 female	305	89.4%	10.6%
Exp-II (A) Train: 3000 male + 3000 female	174	82.1%	17.9%



A/AA	B/BH	CH	D/DH	E	G/GH	H	I	J	K	M	N	P	R	S	T	U
অ/আ	ব/ভ	চ	ড/ঢ	এ	গ/ঘ	হ	ই/ঈ	জ/ঝ	ক/খ	ম	ন/ণ	প/ফ	র	শ/ষ/স	ট/ঠ	উ

Fig. 2. First letter of the word for male only.

TABLE V. THE RESULTS OBTAINED FOR 1000 FEMALE TEST DATA FOR THREE DIFFERENT EXPERIMENTS EXP-I(A), EXP-II(A) AND EXP-III(A)

Experiment	Sentence recognition performance (out of 1000)		Word recognition performance (out of 3590)			
	Correctly recognized Sentence, H	Substitution, S	Correctly recognized Words, H	Deletion, D	Substitution, S	Insertion, I
Exp-I (B) Train: 3000 male	695	305	2337	258	695	35
Exp-II (B) Train:3000 female	867	133	2885	105	300	19
Exp-III (B) Train: 3000 male+3000 female	810	190	2760	142	442	22

TABLE VI. EFFECT OF CONJUNCTS/CONJUGANT WORDS ON WORD ERROR. IN EXPERIMENT WITH 1000 FEMALE DATA.

Experiment	Error Words	Word with conjugant	Other words
Exp-I (B) Train: 3000 male	78	48.7%	51.3%
Exp-II (B) Train: 3000 female	121	47%	53%
Exp-II (B) Train: 3000 male + 3000 female	89	50.1%	49.9%

TABLE VII. EFFECT OF CONJUNCTS/CONJUGANT WORDS ON ERROR SENTENCE FOR 1000 FEMALE TEST DATA.

Experiment	Error Sentences	Sentence having words with conjugant	Other sentences
Exp-I (B) Train: 3000 male	163	81.3%	18.7%
Exp-II (B) Train: 3000 female	133	87.4%	12.6%
Exp-II (B) Train: 3000 male + 3000 female	147	80.1%	19.9%

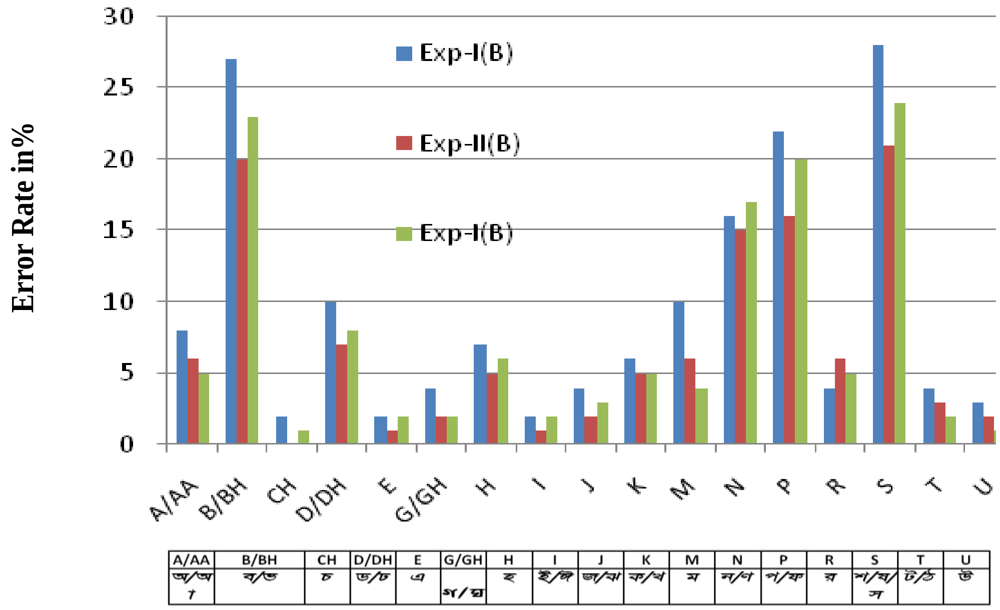
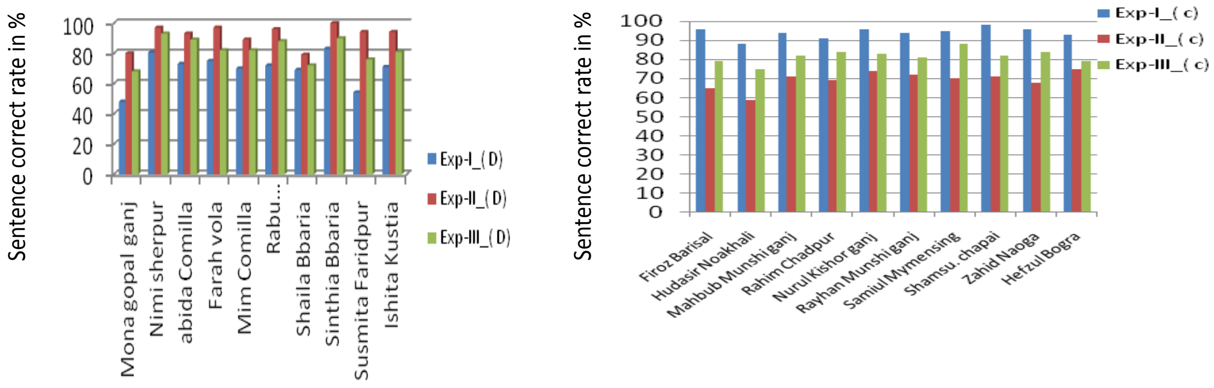


Fig. 3. Error rate vs. the First letter of the word for female only.

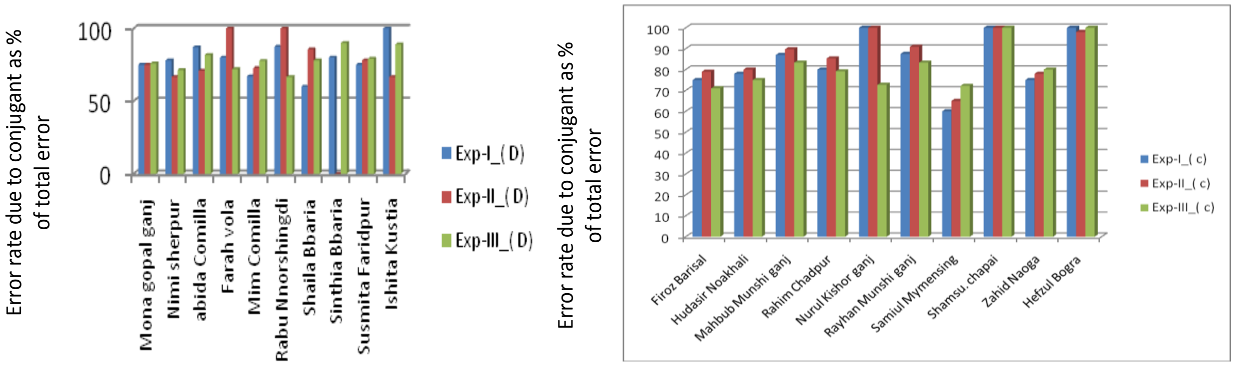
As like as the Table I, we can see that the gender effect and the way to minimize it by training with a large set of male and female data in Table V. Now the effect of conjugant words on word and sentence levels are presented in Table VI and VII respectively.

After observing the effect conjugant words, the effect of first letter on the performance of pronunciation detector is presented in Figure 3. The results in Table VI and VII are Showing that the effect of conjugant word is very prominent in sentence level only.



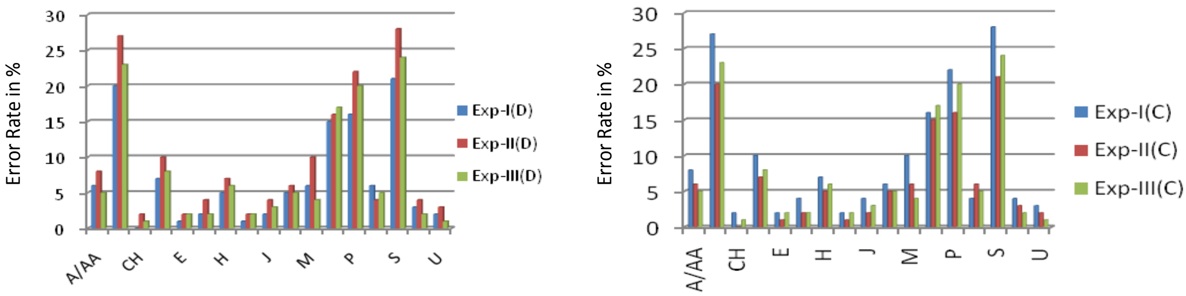
a b

Fig. 4. Comparison of Sentence level and correctness of 10 different (a) female and (b) male test groups.



a b

Fig. 5. Error rate due to conjugant as % of total error vs. Different (a) female (b) male test data.



a b

A/AA	B/BH	CH	D/DH	E	G/GH	H	I	J	K	M	N	P	R	S	T	U
অ/আ	ব/ভ	চ	ড/ঢ	এ	গ/ঘ	হ	ই/ঈ	জ/ঝ	ক/খ	ম	ন/ণ	প/ফ	র	শ/ষ/স	ট/ঠ	উ

Fig. 6. Error rate vs. the First letter of the word for all 10 (a) female (b) male groups.

At this point we are going to analyze the results obtained for 10 groups of 100 male and 100 female data side by side data in each group. First we show the comparison from view of sentence correctness given in Figure 4. The effect of the conjugant words and the first letter in each case of 10 male group 6 respectively. From figure 5 it obvious that the effect of conjugant on error at sentence level is quite notable as in all previous case. Figure 6 shown effect of some specific starting letter on error is quite prominent. Next, to observe the regional effect due to local accent on female test data are presented in foregoing subsections

VI. CONCLUSION

In light of the obtained results and analyzing them with keen awareness we can conclude that the proposed pronunciations detector works best when it is trained using male and female data and tested with male and female respectively. For opposite arrangement it performs the worst. Next to minimize the gender effect the detector should be trained with a large set of both male and female data. Then the effect of conjugant on sentence level is quite noteworthy while not very significant at word level and the effect of first letter is also noticeable, error rate for words starting with B/BH, P, S etc are significantly higher. Moreover, we can see that the conjuncts are mostly converted to a simple non conjunct word. In future, the authors would like to incorporate neural network based systems to improve the performance of the pronunciation detector and develop a guidance system for probable correction for detected pronunciation error.

REFERENCES

[1] S. Matsuda, T. Jitsuhiro, K. Markov and S. Nakamura, "Speech Recognition system Robust to Noise and Speaking Styles," Proc. ICSLP'04, Vol.IV, pp.2817-2820, Oct. 2004.

[2] M. A. Hasnat, J. Mowla, and Mumit Khan, "Isolated and Continuous Bangla Speech Recognition: Implementation Performance and application perspective, " in Proc. International Symposium on Natural Language Processing (SNLP), Hanoi, Vietnam, December 2007.

[3] R. Karim, M. S. Rahman, and M. Z Iqbal, "Recognition of spoken letters in Bangla," in Proc. 5th International Conference on Computer and Information Technology (ICCIT02), Dhaka, Bangladesh, 2002.

[4] http://en.wikipedia.org/wiki/List_of_languages_by_total_speakers, Last accessed April 11, 2009.

[5] S. P. Kishore, A. W. Black, R. Kumar, and Rajeev Sangal, "Experiments with unit selection speech databases for Indian languages," Carnegie Mellon University.

[6] S. A. Hossain, M. L. Rahman, and F. Ahmed, "Bangla vowel characterization based on analysis by synthesis," Proc. WASET, vol. 20, pp. 327-330, April 2007.

[7] A. K. M. M. Houque, "Bengali segmented speech recognition system," Undergraduate thesis, BRAC University, Bangladesh, May 2006.

[8] K. Roy, D. Das, and M. G. Ali, "Development of the speech recognition system using artificial neural network," in *Proc. 5th International Conference on Computer and Information Technology (ICCIT02),* Dhaka, Bangladesh, 2003.

[9] M. R. Hassan, B. Nath, and M. A. Bhuiyan, "Bengali phoneme recognition: a new approach," in *Proc. 6th International Conference on Computer and Information Technology (ICCIT03)*, Dhaka, Bangladesh, 2003.

[10] K. J. Rahman, M. A. Hossain, D. Das, T. Islam, and M. G. Ali, "Continuous bangla speech recognition system," in *Proc. 6th International Conference on Computer and Information Technology (ICCIT03)*, Dhaka, Bangladesh, 2003.

[11] S. A. Hossain, M. L. Rahman, F. Ahmed, and M. Dewan, "Bangla speech synthesis, analysis, and recognition: an overview," in *Proc. NCCPB,* Dhaka, 2004.

[12] S. Young, et al, The HTK Book (for HTK Version. 3.3), Cambridge University Engineering Department, 2005. http:///htk.eng.cam.ac.uk/prot-doc/ktkbook.pdf.

[13] http://en.wikipedia.org/wiki/Bengali_script, Last accessed April 11, 2009.

[14] C. Masica, *The Indo-Aryan Languages*, Cambridge University Press, 1991.

[15] Ghulam Muhammad, Yousef A. Alotaibi, and Mohammad Nurul Huda, "Automatic Speech Recognition for Bangla Digits," ICCIT'09, Dhaka, Bangladesh, December 2009.

[16] Foyzul Hasan, Rokibul Alam Kotwal, Saiful Alam Khan and Mohammad Nurul Huda, "Gender Independent Bangla Automatic Speech Recognition," IEEE/IAPR Internationaol Conference on Informatics, Electronics and Vision (ICIEV) 2012, May 2012, Dhaka, Bangladesh .

[17] T. Nitta, "Feature extraction for speech recognition based on orthogonal acoustic-feature planes and LDA," Proc. ICASSP'99, pp.421-424, 1999.

An Adjustable Novel Window Function with its Application to FIR Filter Design

Hrishi Rakshit
Electrical and Electronics Engineering
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
rishirok@gmail.com

Muhammad Ahsan Ullah
Electrical and Electronics Engineering
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
ahsan_cuet@yahoo.com

***Abstract*—A new class of adjustable window function, based on combination of tangent hyperbolic function and a weighted cosine series, is proposed. The proposed window is adjustable in the sense that by changing its controlling parameter, one can change not only its shape but also its spectral characteristics. The spectral characteristic of the proposed window is studied and its performance is compared with Dolph-Chebyshev, Gaussian and Kaiser Windows. Simulation results show that the proposed window yields better ripple ratio and side-lobe roll-off ratio than the above named windows. Moreover, the paper represents the application of the proposed window in finite impulse response (FIR) filter design. The results confirm that the filter designed by the proposed window provides better spectral characteristics than Kaiser window. This paper also highlights an application of the proposed Filtering method in the area of eliminating noise from corrupted ECG signal.**

Keywords*— *Window shape, main-lobe width, ripple ratio and side-lobe roll-off ratio.

I. INTRODUCTION

The impulse response of an ideal filter is infinitely long. The most straight forward way to make the response finite is to truncate the ideal response. If $h_d(n)$ is the response of an ideal filter, the simplest way to obtain a causal finite impulse response is to multiply the $h_d(n)$ with a finite window function w(n). In engineering term, a window is a finite array, consists of coefficients to satisfy the desirable requirements [1]. In general, the window function can be defined as

$$w[n] = \begin{vmatrix} f(n), & 0 \le n \le N \\ 0, & otherwise \end{vmatrix} \quad (1)$$

and the impulse response of the system is given by

$$h[n] = h_d[n]w[n] \quad (2)$$

Therefore the Fourier transform h[n], denoted by $H(e^{j\omega})$, is the periodic convolution of the $H(e^{j\omega})$, with the Fourier transform of w[n], denoted by $W(e^{j\omega})$. The important spectral characteristics of a window function are i) main-lobe width ii) ripple ratio and iii) its side-lobe roll-off ratio. There are two desirable specifications for any window function. They are smaller main-lobe width and the smaller ripple ratio. However, these two requirements are contradictory [2]. The simplest window is the rectangular window [2].It has a very sharp transition band. However its ripple ratio is the highest among the all commonly used windows. Dolph-Chebyshev and Kaiser windows are the adjustable functions. So by changing the adjustable parameter, one can trade-off between the main-lobe width and the ripple ratio. Dolph-Chebyshev window has the high cost of computation [13]. In the same way, Kaiser window has the computational complexity due to the calculation of Bassel functions [4]. So there has been a great interest to design a new window to meet the desire specification for different applications.

In this paper, a new form of adjustable window function is proposed which has superior performance compared to the several commonly used windows. By adjusting its controlling parameter r, it is possible to control the spectral parameters such as the main-lobe width, the ripple ration and the side-lobe roll-off ratio. For beam forming applications, the higher side-lobe roll-off ratio means that it can reject the far end interference better. It also reduces the energy leak from one band to another for speech processing [3].

II. SPECTRAL PROPERTIES OF WINDOW

A window, w(nT), with a length of N is a time domain function which is nonzero co-efficient for n<=|(N-1)/2| and zero for otherwise. A typical window has a normalized amplitude spectrum in dB as shown in Fig 1.

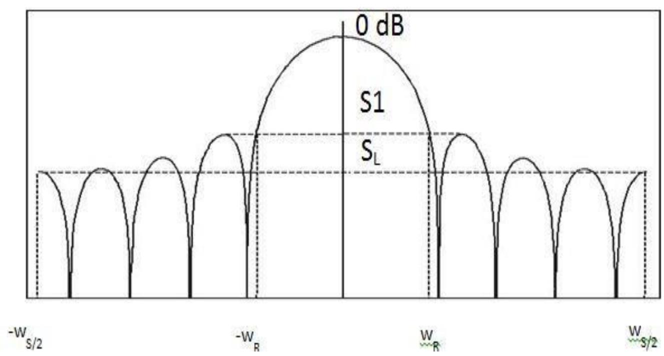


Fig1: A typical window's normalized amplitude spectrum.

Normalized spectrum in Fig1 is obtained by following equation

$$|W_n(e^{j\omega T})| = 20\log_{10}(|A(w)| / |A(w)|_{max}) \quad (3)$$

The spectral characteristics which determine the windows performance are the main-lobe width(w_M), the ripple ratio (R) and the side-lobe roll-off ratio(S). From the Fig. 1, these parameters can be defined as

W_M= Two times half of the main-lobe width=$2w_R$.

R= (Maximum side-lobe amplitude in dB) – (Main-lobe amplitude in dB) = S1.

S= (Maximum side-lobe amplitude in dB) – (Minimum side-lobe amplitude in dB) = S1-SL.

III. Proposed Window

The new proposed window is a combination of tan hyperbolic function and a weighted cosine series. Johansen and Sorensen (1979) combined two hyperbolic tangent functions in frequency domain to construct an analytical expression which is given by as follows [14].

$$p(f) = \frac{1}{2}\tanh[\frac{\pi}{a}(f+\frac{1}{2})] - \frac{1}{2}\tanh[\frac{\pi}{a}(f-\frac{1}{2})] \quad (4)$$

The main disadvantage of the above window function is that there is no time domain representation of the window. So the modified tan hyperbolic function is given by as follows

$$y_1 = \left[\tanh\left\{ \frac{n - \frac{(N-1)}{2} + \cosh^2(\alpha)}{B} \right\} - \tanh\left\{ \frac{n - \frac{(N-1)}{2} - \cosh^2(\alpha)}{B} \right\} \right] \quad (5)$$

The weighted cosine function has been introduced with the tan hyperbolic function to make the ripple ratio better. The weighted cosine function can be expressed as follows

$$y_2 = 0.375 - 0.5\cos\left\{\frac{2\pi l}{N-1}\right\} + 0.125\cos\left\{\frac{4\pi l}{N-1}\right\} \quad (6)$$

Where N is the length of the window, α and B are the constants. Here 'n' is the number of samples. Here the symbol, l =0,1,2,3,……(N-1).

These two equations are combined in the following way.

$$w(n) = (y_1 \times y_2)^{\gamma}, \quad where, \gamma = r^{(\frac{1}{5})} \quad (7)$$

Here r is a variable which controls the shapes and frequency response of the window. So "r" can be named as controlling parameter of the window. The Symbol "×" indicates simple multiplication. In mathematical term, the window can be expressed as

$$w(n) = \begin{vmatrix} (y_1 \times y_2)^{\gamma}, & n \le 0 \le N \\ 0, & otherwise \end{vmatrix} \quad (8)$$

Fig 2 illustrates that the shape of the window varies as the value of the adjusting parameter r changes. With r = 0, the proposed window becomes rectangular shape. Except r = 0, all the window touches X-axis evenly. For r = 1 and r = 1.618, the window shapes are very close to each other. The narrowest window have been found for r = 4.618. Fig 2 also shows that the width of the window gets narrower as the value of the controlling parameter gets bigger.

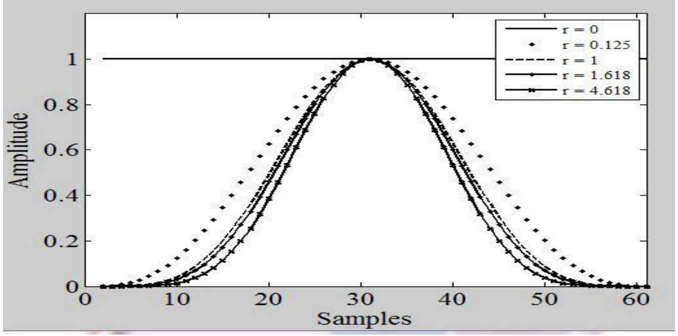


Fig 2: Shapes of proposed window with different values of r.

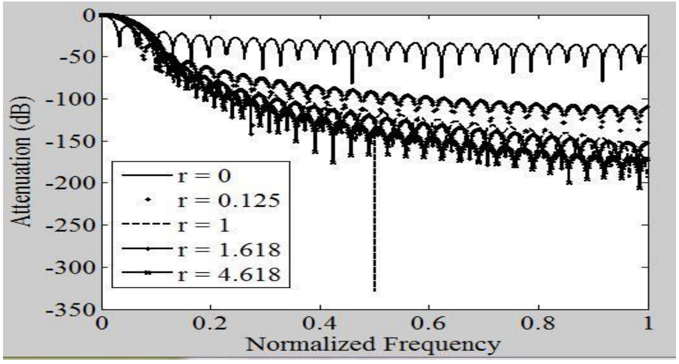


Fig 3: The normalized frequency response of the proposed window with different values of r.

From the Fig 3, it has been observed that frequency response of the proposed window with r=0, is exactly same to the frequency response of the rectangular window. At the same time, Fig 3 shows that as r increases the main-lobe width and the ripple ratio become wider and smaller respectively.

TABLE I : DATA OF THE SPECTRAL PARAMETERS FOR THE PROPOSED WINDOW WITH VARIOUS R.

Adjustable Parameter, r	Main-lobe Width	Ripple Ratio (dB)	Side-lobe Roll-off ratio(dB)
r=0	2π×0.0312	-13.26	24.1055
r=0.125	2π×0.0742	-34.061	78.8125
r=1	2π×0.0977	-40.6225	118.7619
r=1.618	2π×0.1035	-45.4702	111.0602
r=4.168	2π×0.1211	-53.1848	122.9622

Table I shows that as r increases the main-lobe width increases and ripple ratio decreases. At the same time side-lobe roll-off ratio increases with the increment of the controlling parameter r. The proposed window shows its best performance at r = 1.618. Because when r = 1.618, there is a trade-off between the main-lobe width and the ripple ratio. At the same time, the side-lobe roll-off ratio is also very good at r = 1.618.

IV. Performance Analysis

In this section, we compare the shape and the spectral characteristics of the proposed window with several commonly used windows.

A. Dolph-Chebyshev Window:

The zero-phase Dolph–Chebyshev window function $w_0(n)$ is usually defined in terms of its real-valued discrete Fourier transform

$$w_0(k) = \frac{\cos\{N\cos^{-1}[\beta\cos(\frac{\pi k}{N})]\}}{\cosh[N\cosh^{-1}(\beta)]} \tag{9}$$

$$\text{where } \beta = \cosh[\frac{1}{N}\cosh^{-1}(10^{\alpha})]$$

The value of α determines the side-lobes attenuation. The zero phased window function can be calculated by taking the inverse discrete Fourier transform of $W_0(k)$ [5].

$$w_0(n) = \frac{1}{N}\sum_{k=0}^{N-1} w_0(k)e^{\frac{i2\pi kn}{N}}, \quad -\frac{N}{2} \le n \le \frac{N}{2} \tag{10}$$

The normal window function can be written as

$$w(n) = w_0(n - \frac{N-1}{2}), \quad 0 \le n \le N-1. \tag{11}$$

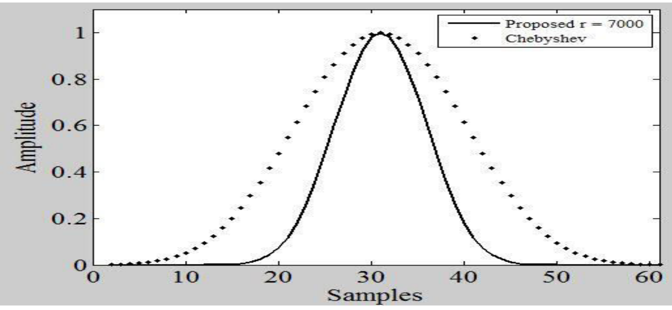


Fig 4: The comparison of Dolph-Chebyshev window and the Proposed window in time domain representation

From the Fig 4 it has been seen that the time domain representation of the proposed window with the adjusting parameter r = 700 is more narrower than the Dolph-chebyshev window.

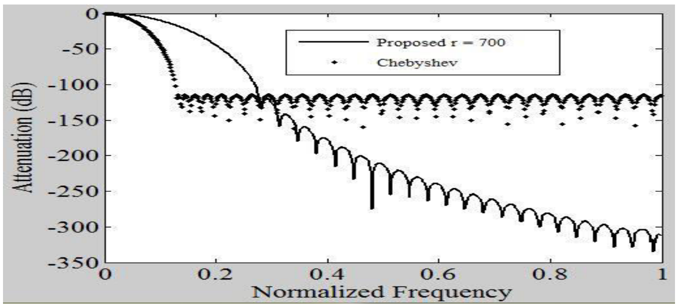


Fig 5: Normalized frequency response comparison of Dolph-Chebyshev and the Proposed windows.

Fig 5 shows that the main-lobe width of the proposed window gets almost double than the Dolph-Chebyshev window. On the other hand, the ripple ratio of the proposed window becomes almost 13.55 dB better than the Dolph-Chebyshev window. In term of side-lobe roll-off ratio the proposed window is approximately 12.8 times better than the Dolph-Chebyshev window. Fig 5 has been summarized in the table II.

TABLE II: SPECTRAL PARAMETERS COMPARISON OF THEDOLPH-CHEBYSHEV AND THE PROPOSED WINDOW.

Window	Main-lobe Width	Ripple Ratio (dB)	Side-lobe Roll-off Ratio(dB)
Dolph-Chebyshev	2π×0.1269	-102.2413	15.1371
Proposed, r=700	2π×0.2519	-115.1503	194.957

B. Gaussian Window:

The coefficients of the Gaussian window function are computed using the following equation.

$$w(n) = e^{\frac{1}{2}(\alpha\frac{n}{N/2})^2} \tag{12}$$

where α is inversely proportional to the standard deviation of a Gaussian random variable. The width of the window is inversely related to the value of α. A larger value of α produces a narrower window [6][7].

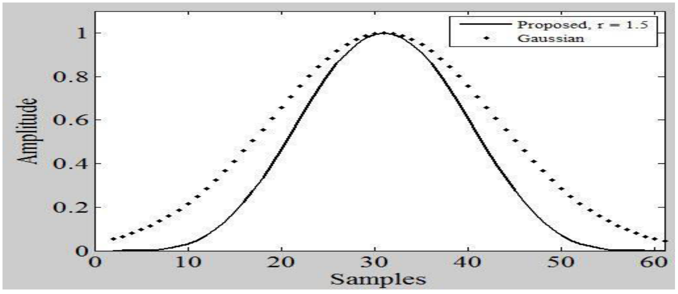


Fig 6: The comparison of Gaussian window and the Proposed window in time domain representation

From the Fig 6 it has been seen that the time domain representation of the proposed window with r =1.5, is narrower than the Gaussian window. Moreover the proposed window touches the X-axis while all the coefficients of the Gaussian window remain above the X-axis.

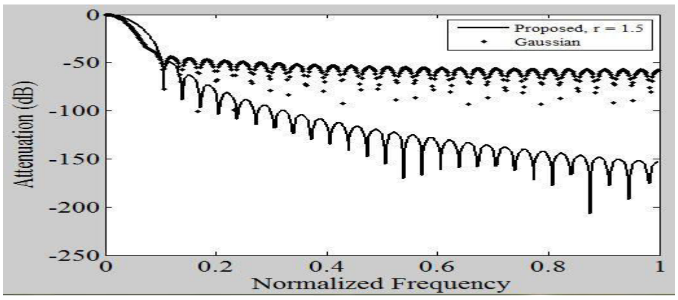


Fig 7: Normalized frequency response comparison of Gaussian and the Proposed windows.

Fig 7 shows that the main-lobe width of the Gaussian and Proposed window is exactly the same while the ripple ratio of the proposed window is 5.11dB smaller than the Gaussian window. In terms of side-lobe roll-off ratio, the proposed window shows superior performance compared to the Gaussian window. Fig 7 has been summarized in the table II.

TABLE III: SPECTRAL PARAMETERS OF GAUSSIAN WINDOW AND THE PROPOSED WINDOW.

Window	Main-lobe Width	Ripple Ratio (dB)	Side-lobe Roll-off Ratio(dB)
Gaussian	2π×0.1035	-44.12	13.44
Proposed, r=1.5	2π×0.1035	-49.23	102.67

C. *Kaiser window:*

The Kaiser window is also known as Kaiser-Bessel window function. The Kaiser window function can be expressed as follows [8][9].

$$w(n)=\frac{I_0(\pi\alpha\sqrt{1-(\frac{2n}{N-1}-1)^2})}{I_0(\pi\alpha)}\ where, 0\le n\le N \qquad (13)$$

where I_0 is the zero-th order modified Bessel function of the first kind [8]. Kaiser window is another very popular adjustable window. The variable parameter α determines the tradeoff between the main-lobe width and the side lobe attenuation.

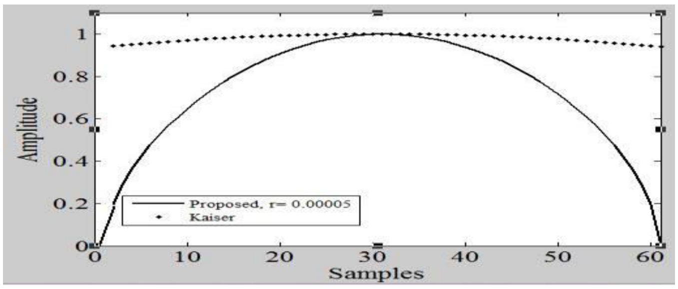


Fig 8: The comparison of Kaiser window and the Proposed window in time domain representation

Although the shapes of the proposed window and the Kaiser windows are different but their spectral performance are very close to each other.

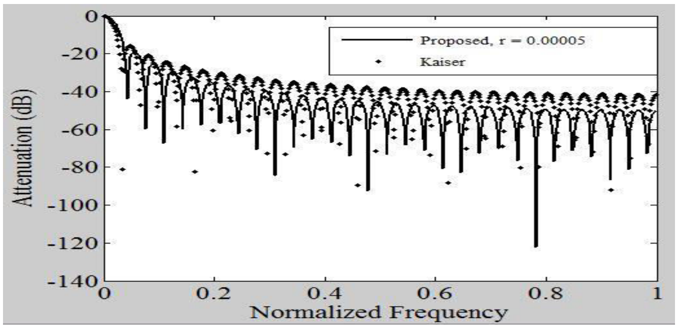


Fig 9: Normalized frequency response comparison of Kaiser and the Proposed windows.

It has been seen that in main-lobe width calculation, Kaiser window is slightly better than the proposed window with controlling parameter r=0.00005. On the other hand, in terms of ripple ratio the proposed window performs finer performance than the Kaiser window. Its side-lobe roll-off ratio is also bigger than the Kaiser window. So it is more frequency selective. Table IV shows the comparative performance of the Kaiser and the proposed window with r=0.00005.

TABLE IV: SPECTRAL PARAMETERS OF KAISER WINDOW AND THE PROPOSED WINDOW.

Window	Main-lobe Width	Ripple Ration (dB)	Side-lobe Roff-off Ratio(dB)
Kaiser	2π×0.0312	-13.0716	30.8469
Proposed, r=0.00005	2π×0.041	-15.5113	37.5349

So from this section, it has been observed that by changing the adjustable parameter of the proposed window, the spectral characteristics can be changed. When r = 1.5, the proposed window yields 5.11 dB and 89.23 dB better ripple ratio and side-lobe roll-off ratio than the Gaussian window for the same main-lobe width. When r = 0.0005, the ripple ratio and the side-lobe roll-off ratio of the proposed window are 2.4397 dB and 6.688 dB better respectively than the Kaiser window, while maintaining almost the same main-lobe width.

V. APPLICATION

In this section we will discuss how the proposed window helps us to design FIR low pass filter and also compare the simulation result with the other adjustable windows.

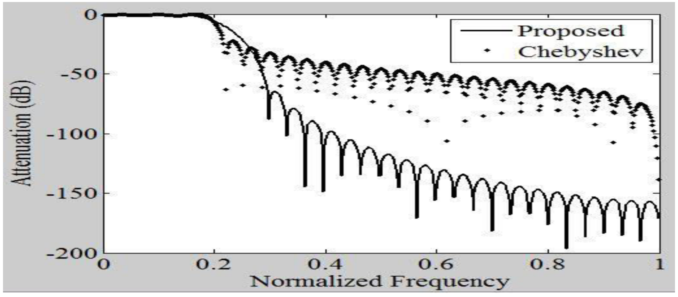


Fig 10(a): Performance analysis of FIR low pass filter between the proposed and Dolph-Chebyshev window.

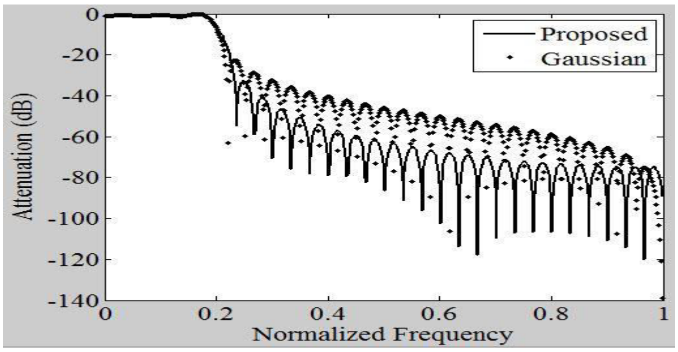


Fig 10(b): Performance analysis of FIR low pass filter between the proposed and Gaussian window.

Having a cut off frequency of ω_c, the impulse response of an ideal low pass filter is given by

$$h_{LP,ideal}[n]=\frac{\sin(\omega_c n)}{\pi n} \qquad (14)$$

By windowing this IIR filter with the windows, discussed in this paper, different FIR filter can be obtained. For $\omega_c = 0.2\pi$ figures 10(a)-10(c) show the normalized frequency response of the FIR filters designed by applying different windows of length 61.

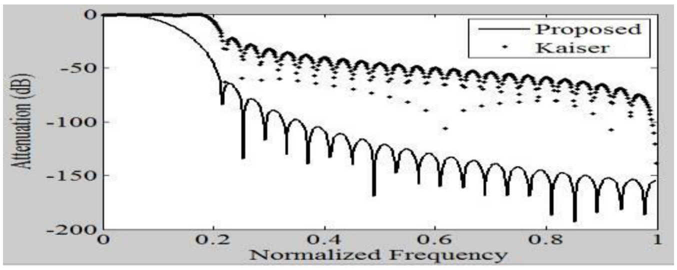


Fig 10(c): Performance analysis of FIR low pass filter between the proposed and Kaiser window.

Our simulation results show that the filter achieve by the proposed window has 42.58 dB, 9.85 dB and 47.48dB less side lobe peaks than Dolph-Chebyshev, Gaussian and Kaiser windows respectively. The following table summarized the total simulation result.

TABLE V: DATA ANALYSIS OF FIR FILTER OBTAINED BY WINDOWING WITH DIFFERENT WINDOWS

Comparing Windows	**Main-lobe Width**	**Ripplr Ratio (dB)**	**Side-lobe Roll-off Ratio**
Proposed Vs Dolph-Chebyshev	$2\pi\times0.2969$	-60.4552	99.2238
	$2\pi\times0.2187$	-17.8744	57.3927
Proposed Vs Gaussian	$2\pi\times0.25$	-27.729	42.7795
	$2\pi\times0.2187$	-17.8744	57.3927
Proposed Vs Kaiser	$2\pi\times0.2129$	-62.92	95.3285
	$2\pi\times0.2187$	-15.4329	59.8342

FIR filter designed with the proposed window performs tremendous performance in terms of ripple ratio and the side-lobe roll-off ratio than Dolph-Chebyshev window. If we compare the proposed window with Gaussian window, than the main-lobe width of the proposed window is very much near with it and ripple ratio is even better than the Gaussian window. The proposed window performs supercilious performance in all respects than Kaiser window.

VI. NOISE ELIMINATION FROM ECG SIGNAL

The electrical activity of heart is called ECG signal which is generated by repolarization and depolarization of the atria and ventricles [9]. The ECG signals are very small in amplitude. The frequency range of the ECG signals lies in between 0.05 to 100 Hz. ECG signals are consist of P-waves, QRS complex and T-waves. The depolarization of atria results P-waves [10]. The QRS complex consists of three waves, sequentially Q, R and S waves. The rapid depolarization of both the ventricles yields the QRS complex. Ventricular repolarization generates T-waves [11]. The standard ECG signal of one period is shown in the Fig 11. Arti-factual signals are generated from various internal and external sources and mixed with ECG signal. The noise generally caused by the signal from muscle contractions, which can be expressed as zero mean Gaussian noise [12].

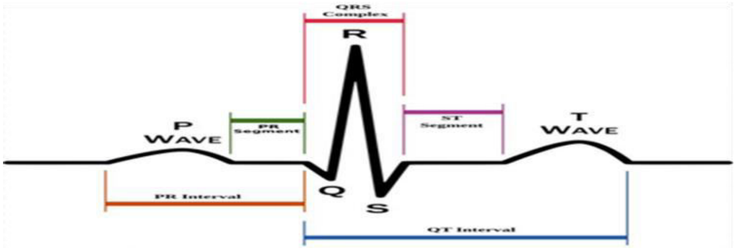


Fig 11: Typical ECG trace.

Now if this ECG signal is corrupted by Additive White Gaussian Noise, then the view of the ECG signal becomes as below.

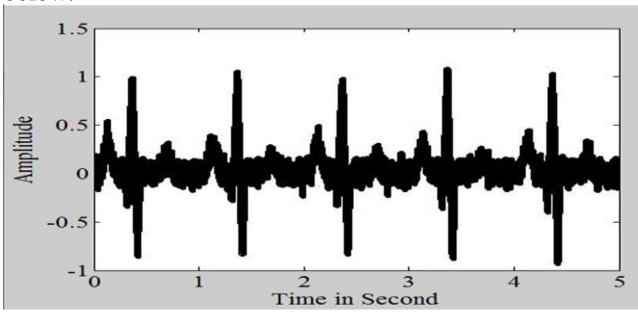


Fig 12: Corrupted ECG signal due to Additive White Gaussian Noise

This corrupted ECG signal is then passed through a low pass FIR filter which is designed using Kaiser window function. The result is shown in the Fig 13.

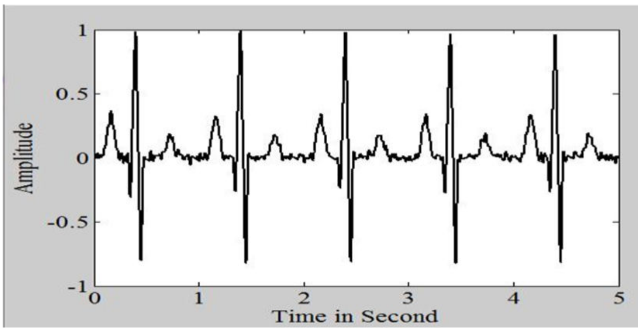


Fig 13: Filtered ECG signal using Kaiser window based FIR low pass filter

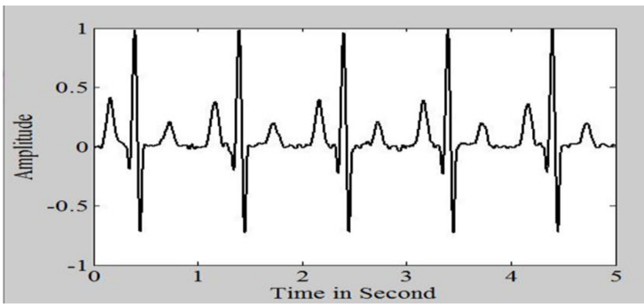


Fig 14 : Filtered ECG signal using the Proposed window based FIR low pass filter

VII. CONCLUSION

A novel highly adjustable window function has been proposed which has adaptive spectral characteristics. One can trade-off between mail-lobe width and ripple ratio by changing its adjusting parameter. The proposed window has very high side-lobe roll-off ratio compared to the commonly used adjustable windows named as Dolp-Chebyshev, Gaussian and Kaiser windows. Another advantages of the proposed windows is that it has only one adjustable parameter. The FIR filter designed using the proposed window achieves less ripple ratio than the above mentioned windows. At the end, the proposed window based FIR low-pass filter reduces Additive White Gaussian Noise (AWGN) from ECG signal more precisely than Kaiser window.

REFERENCES

[1] Hrishi Rakshit and Muhammad Ahsan Ullah "A comparative study on window functions for designing efficient fir filter". In Proceedings of the IEEE 9th International Forum on Strategic Technology (IFOST), ISBN: 978-1-4799-6060-6 DOI: 10.1109/IFOST.2014.6991079 pp 91-96.

[2] Oppenheim, A.V., and R.W. Schafer. Discrete-Time Signal Processing. Upper Saddle River, NJ: Prentice-Hall, 1999, pp. 468-471.

[3] A. Jain, R. Saxena, "A simple alias-free QMT system with near-perfect reconstruction" J. Indian Ins. Sci., Jan-Feb, no. 12,2005,pp 1-10.

[4] Emmanuel C. Ifeacher and Barrie W. Jervis Digital Signal Processing 2nd edition, Pearson Education 2008.

[5] Julius O. Smith ``Spectral Audio Signal Processing'', III, W3K Publishing, 2011, ISBN 978-0-9745607-3-1.

[6] Roberts, Richard A., and C.T. Mullis. Digital Signal Processing. Reading, MA: Addison-Wesley, 1987, pp. 135-136.

[7] Harris, F.J. "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform." Proceedings of the IEEE. Vol. 66, No. 1 (January 1978).

[8] Kaiser, J.; Schafer, R. (1980). "On the use of the I0-sinh window for spectrum analysis". IEEE Transactions on Acoustics, Speech, and Signal Processing 28: 105.

[9] Selected Papers in Digital Signal Processing II, IEEE Press, New York, 1975.

[10] AK.M Fazlul Haque, Md. Hanif Ali, M. Adman Kiber, Md. Tanvir Hasan,"Detection of small variations of ECG features using Wavelets". ARPN ,"Journal of Engineering and Applied Science" ISSN 1819-6608, Vol. 4,No. 6, pp.27-30,Aug2009.

[11] Smith, SW. "Atrial Repolarization Wave Mimicking ST Depression " Retrieved 22 October 2014.

[12] S.A. Chouakri, F. Bereksi-Reguig, S. Ahamaidi, and O. Fokapu, "Wavelet de-noising for the Electrocardiogram signal based on the corrupted noise estimation" Computers in Cardiology, pp. 1021-1024, 2005.

[13] A. Antoniou, "Digital Filters", McGraw-Hall Inc, N.Y. 1993.

[14] Ahmet Tugrul Basokur, 'Digital Filter Design Using the Hyperbolic Tangent Functions', Journal of the Balkan Geophysical Society, vol. 1, no 1, February 1998, pp. 14-18.

Telemedicine in South Asia for Rural People: Current Scenario and Future Recommendations

Uzzal Kumar Prodhan
Computer Science & Engineering Department
Jatiya Kabi Kazi Nazrul Islam University
Trishal, Maymensingh, Bangladesh
uzzal_bagerhat@yahoo.com

Muhammad Zahidur Rahman, Israt Jahan
Computer Science & Engineering Department
Jahangirnagar University
Savar, Dhaka, Bangladesh

***Abstract* – Telemedicine encompasses methods for electronically transmitting medical information to sustain and/or improve patient's health condition by using information and communication technology. Different countries of South Asia have tried to introduce telemedicine services from different perspectives for the rural people where health care facilities are poor through its private and public partnership. This paper examines different telemedicine models of South Asia and finds out the potentials difficulties faced by different models and presents the factors which should be assessed carefully for the successful implementation of telemedicine which will be feasible, easily maintainable, sustainable and cost effective solution for the poor people of rural areas of South Asia regions.**

Keywords—telemedicine; times; dicot; dit; isro; trcl

I. Introduction

Telemedicine is a process to provide health care to the remote areas where there is a scarcity of basic health facilities. The main objective of telemedicine is to provide equal access of medical expertise. It is applicable where medical experts are rare, distances are a major factor and infrastructure is very limited. Most of the experienced doctors and other medical expertise are on urban based than rural areas of South Asia. About 75 to 80% people live in rural areas in South Asia where there are limited medical facilities are available [1]. In order to take proper medical care, all the patients from rural areas will have to travel to major cities where expertise doctors are available to get medical treatment with all the modern facilities. In this paper different telemedicine models of South Asian countries are analyzed through literature review and some recommendations are given for a standard telemedicine model which will be feasible and ideal solution for the rural people in this region.

II. Materils and Methods

In order to carry this research, the different telemedicine models of South Asia are collected from the renowned journals, books and conference papers. We have analyzed these telemedicine models and the characteristics of these models are shown in the result section. The discussion section summarizes the research work and lists the proposed recommendations for standard telemedicine in the region of South Asia.

III. Results

The country based different telemedicine models of South Asia are following:

A. Telemedicine Models in Bangladesh

In Bangladesh, the telemedicine project was started as a model for the health care facilities of the rural people in 1999 [2].

1) Model-1: Center for Rehabilitation of Paralyzed in Savar established telemedicine link with the Royal Navy Hospital, Haslar, UK in 1999. This model used digital camera and satellite telephone for consultation [1,2].

2) Model-2: TRCL demonstrated telemedicine system in the US Trade Show 2001 in Dhaka using Icare software and normal internet connection & started test-run of the system between US and Bangladeshi physicians [1, 2].

3) Model-3: Sustainable Development Network Program (SDNP) Bangladesh started in January 2003 which has four regional nodes in different parts of Bangladesh. These nodes are connected to satellite by VSAT technology. Their sessions provide consultancy and diagnostic support to the physician at the remote end through medical experts at the SDNP head office [3].

4) Model-4: Bangladesh University of Engineering & Technology (BUET) and Comfort Nursing Home had started a telemedicine project with the financial collaboration from European Union (EU) in 2003 through e-mail. Recently the project is not functional [1, 2].

5) Model-5: Bangladesh DNS diagnoses Centre, Gulshan-1 and Comfort Diagnoses & Nursing Home's started a telemedicine centre in 2004. The project was discontinued because of lack of financial viability, patient disinterest and poor market promotion [1, 2].

6) Model-6: Telemedicine model established between Diabetic Association of Bangladesh (DAB), Dhaka and Faridpur General Hospital to access the specialist doctors in 2005 through video conferencing. The cost per consultation was 600 taka and 20 patients per day was the target of the project [4].

7) Model-7: Indian hospitals and Medinova hospital initiated telemedicine services through e-mail and Internet in 2006 [1].

8) Model-8: Telemedicine Service "HealthLine Dial 789" which provides different types of medical information facility, emergency service (SMS based LAB report, ambulance) and real time medical consultation over mobile phone by Telemedicine Reference Centre Ltd. and Grameen phone [5].

9) Model-9: Access to Information project started between Union Information and Service Centers (UISCs) used computer, printer, digital camera, scanner, internet, modem and Skype from 2010. Doctors sitting at the MIS office are giving medical consultation every working day of 4,536 unions of Bangladesh [6].

10) Model-10: Two specialized hospitals (Bangabandhu Sheikh Mujib Medical University and National Institute of Cardiovascular Diseases), 3 district hospitals (Shatkhira, Nilphamari and Gopalganj) and 3 sub-district hospitals (Pirgonj, Dakope and Debhata) used web-camera, Skype or any other video conferencing platform where patients from district and sub-district can get treatment from specialized hospital [6].

11) Model-11: In 2012 ICT Ministry & Grameen Phone implemented telemedicine between TWGBD, Aysha Memorial Specialized Hospital, Dusthya Sathya Kendro and Concern Worldwide for Jessore District. This project used Telemedicine Information, Management and Education System (TIMES), Digital Imaging and Communication in Telemedicine (DICOT), Stethoscope, Sphygmomanometer, Scanner of ECG and X-ray [7].

B. Telemedicine Models in India

Telemedicine activities were started at 1999 in India [8].

1) Model-1: Indian Space Research Organization (ISRO) is operating telemedicine nodes under GRAMSAT (rural satellite) program. It has established a Telemedicine network consisting of 382 Hospitals (51 hospital is Specialists & 306 or more is district or rural level hospital) [9, 10]. It used satellite based connectivity among the nodes.

*2) Model-2:*Department of Information Technology (DIT), Ministry of Communication and IT (MCIT), Government of India has established more than 75 nodes all over India. Project used ISDN, WAN & Satellite connectivity for connecting nodes [9]. It is used for diagnosis & monitoring of tropical diseases in West Bengal.

*3) Model-3:*Ministry of Health and Family Welfare (MoH & FW) was implemented by integrated Disease Surveillance Programme network with the help of ISRO. OncoNET India is a network connecting 25 Regional Cancer Centers and 100 peripheral centers to provide comprehensive cancer treatment facilities and carry out cancer prevention and research activities [9].

4) Model-4: State Governments of Orissa and Uttar Pradesh, Chhattisgarh and Kerala Oncology Network connect its secondary level hospitals to medical colleges and specialty hospitals for Telemedicine. Rajasthan and Karnataka State Government in collaboration with ISRO, has established Telemedicine network for its rural people. Andhra Pradesh and Gujarat carry out telemedicine through "104 services". Punjab and Himachal Pradesh government connected with Indira Gandhi Medical College, Shimla and Postgraduate Institute of Medical Education & Research Chandigarh through ISDN link [10].

5) Model-5: Sanjay Gandhi Postgraduate Institute of Medical Sciences (SGPGIMS), Lucknow model was started in 1999 in project mode [10]. SGPGI Telemedicine network has linked 27 national and international nodes and has been carrying out tele-education and tele-healthcare activities.

6) Model-6: SAARC Telemedicine Network was started in April 2007. The pilot project connects one or two hospitals in each of the SAARC countries with three to four super-specialty hospitals in India [10]. This is being developed as an exemplary model for implementing projects at the regional level.

7) Model-7: The Ministry of External Affairs is implementing Pan-African e network project through Telecommunications Consultants India Ltd. (TCIL) to establish a VSAT-based telemedicine and tele-education infrastructure for African countries in 53 nations of the African Union which provides effective tele-education, telemedicine, Internet, videoconferencing and VoIP services [10].

8) Model-8: Narayana Hrudayalaya's (NH) Telemedicine Service it covers 332 hospitals –299 remote/rural/district hospitals/health centers connected to 33 specialty hospitals located in major cities through the ISRO network,. ECG reports, Audio/Visual data, CT scans, X-rays, MRIs and their analysis are exchanged via the telephone line, broadband connection or satellite provided by ISRO [10].

9) Model-9: Apollo Hospitals Telemedicine project has 150 telemedicine centers in India and abroad [11]. The Apollo Project has opened remote telemedicine centers that link villagers via satellite to specialist services. Apollo's Aragonda project was India's first rural telemedicine station. Today, the Apollo project has expanded with telemedicine centers at Bangladesh and multiple sites in India.

C. Telemedicine Models in Sri Lanka

In Sri Lanka, the telemedicine project was started in 2003.

1) Model-1: Telemedicine activity in the country was started in November 2003 [12]. Initially it was started in 8 hospitals over 5 districts in the country. This is a low cost, store and forward distributed telemedicine system. Sri Lanka Health Telematics (SRLHT) is increased with Voice over Internet Protocol (VoIP) for real time communication during tele-consultation sessions.

2) Model-2: The Telemedicine Consultation Centre in Colombo and Apollo Hospitals in India started telemedicine through health-care information technology in 2011[11]. It offered medical consultation service for rural patients who were seeking consultation with Apollo hospitals in India. Medical information of patients including high definition images of medical reports of X-rays scans and ECGs were stored at a data centre in Hyderabad.

D. Telemedicine Models in Nepal

1) Model-1: Telemedicine program is a newly implemented by the Government of Nepal and it is named as 'rural-telemedicine program' [13]. The rural-telemedicine program was initially implemented in 25 district hospitals of hilly and mountainous districts. In the year 2012, government further extended the program in 5 more districts hospitals so in the present context the Telemedicine program is implemented in total 30 districts out of 75 districts of the country. Rural-telemedicine program is using store and forward method, video-conferencing and telephone based consultation.

E. Telemedicine Models in Bhutan

1) Model-1: The main objectives of the Rural Telemedicine Project in Bhutan are to provide a forum for the doctors at the remote districts to tele-consult cases to the specialists at the national referral hospital [14]. This project was implemented in April 2009 in 14 rural sites with supply of the following equipments: Laptops, Non-invasive vital sign Monitor, Portable ECG Machines and Compatible software to interface the equipment.

F. Telemedicine Models in Maldives

1) Model-1: In 2002 the first Telemedicine project was developed by the government as the 'Health Telematics Project' funded by World Health Organization [15].

2) Model-2: In the year 2004 Telemedicine 2nd project (a component of Integrated Human Development Project (IHDP)) was started and funded by World Bank. Poor project management, improper risk assessment and human factors that contributed to failure of the past telemedicine projects.

G. Telemedicine Models in Pakistan

The Prominent Telemedicine Projects of Pakistan were started from 1998 [16].

1) Model-1: The telemedicine model Telemedpak started in 1998. The area of the project was Resource Center RC-Holy Family Hospital HFH Rawalpindi and Remote Station RS- Taxila, Gilgit. Infrastructure used in this model is dial up modem internet connection. This model used Store and Forward method [17].

*2) Model-2:*The model COMSATS started in 2002 & 2005. The area of the project was RC- Comsats Islamabad and RS – Gujar Khan & Gilgit. Infrastructure used in this project was Internet via landline / VSAT (dial up). It used store and forward method [17].

*3) Model-3:*The project Health Net Suparco & MoIT started in 2007-2010 [16]. Infrastructure used Paksat-1 satellite 500 Kbps Up/Down link connection. This model facilitates tteleconferencing facility.

4) Model-4: Multitasking of e-Health training center HFH was started in 2008.Virtual training lab at HFH provided distant training to nurses & other medical personal from HFH to the RS. This model offers also advanced training, healthcare and rehabilitation facilities [16].

5) Model-5: Aga Khan Health Services started in 2007-2013 [16]. Infrastructure used in this project was Gigabit Ethernet at AKUH Karachi and 2Mbps link with Kabul.

6) Model-6: Jaroka Tele health care started in 2006. The technology used in this project is EDGE mobile internet and VSAT server connection. Store and forward method is used in this project [16].

IV. Discussion

From this study it was found that most of the people of South Asia are living in rural areas and only 15 to 20% people live in urban areas where health facilities are sufficient and moderated. Most of the rural peoples will have to travel to urban areas for their basic treatment. In this respect, telemedicine can play a vital role for the delivery of healthcare services to the huge number of people of rural areas.

In order to find the cost benefit analysis of telemedicine, study shows that telemedicine models like Grameen phone telephone medical advice fees are 15 Taka (around US$0.20) for a three-minute call to a qualified physician, AMCARE membership fees were US$0.60–US$20 per month, Aponjon registered members pay 5 Taka per message, Telemedicine established between DAB and Faridpur need 600 taka per consultation. The cost should be low so that poor people can avail the cost easily.

From the study of different telemedicine models in South Asia, it was found that the expectation of rural people is to access the medical care from the expert doctors with low cost as the expert doctors are not available in rural areas. The rural people will be satisfied if they can get health care services through Telemedicine without going to the urban area. This will save their money, travel difficulties and mental happiness.

The similarities of different models of South Asian countries are most of the countries used store and forward method at the early stages of telemedicine, primary aim of the projects are to enhance the healthcare facilities, most of the models are project basis. The common dissimilarities are the used technology in the models.

The significant contributions of this paper are to review most of the widely used telemedicine models and recommend the attributes of the proposed telemedicine model which will be feasible, easily maintainable and cost effective solution for the poor rural people of South Asian regions.

In order to deliver proper medical facilities to the poor people of South Asian countries, the following recommendation can play a vital role for the effective telemedicine of these countries:

- Selecting the right model: From our analysis we have found that there are different telemedicine models in different countries by which patients are being served but there is no standard model by which all the models can be merged. However it is necessary to develop a viable telemedicine model so that any remote patient can avail better service.
- Choosing telemedicine network and non interruptible service: Telemedicine models used VSAT, Satellite, Optical Fiber, Broadband, Wireless and different connectivity style. Different network has some specific advantages and others have disadvantages for which models are failed to meet the requirements of the smooth consultation. So we have to choose the proper network connectivity for implementation of standard model and the smooth continuation of the consultation should be ensured.
- Financial support and cost effectiveness: The financial investment for the establishment of telemedicine and remote user payment should be low so that it can be used widely.
- User satisfaction: Telemedicine models which are to be implemented need to satisfy their user's requirements. To attract rural population, payment for the service should be lower than the urban counterpart.
- Proper training: As telemedicine services are given between two geographically separated areas so there needs to give proper training facilities to rural people and specialized persons so that both parties can actively participate in the consultation. As a result, the goals of telemedicine can be fulfilled.
- Security of data: Telemedicine data are moved through the network, there is possibility to leakage of personal data of patients. Special security of data should be made on the collected data so that everybody can trust on the system.
- Setup standard equipments for automation: Standard compatible telemedicine equipments need to be setup in both ends so that vital different signs of patients can be measured electronically. All the collected data of patients should be stored in the central server of the telemedicine for future use and processed data digitally to give the service to the patients.

V. CONCLUSION

Recently, telemedicine is playing a vital role in order to develop the health care facilities for the developing countries of South Asia. Through this process, the different telemedicine models are offering various services with a low cost for rural people of South Asia. Therefore, a standard telemedicine infrastructure is essential in order to provide the health care facilities for the rural people of South Asia. We believe that the proposed telemedicine system based on our recommendations can play a vital role in order to get medical health care facilities with a cost effective way for the rural people of South Asian countries.

ACKNOWLEDEGEMENT

The authors would like to express sincere gratitude to the Ministry of Information & Communication Technology, Bangladesh for their financial support.

REFERENCES

[1] M. Sanaullah Chowdhury, Md. Humayun Kabir, Kazi Ashrafuzzaman and Kyung-Sup Kwak "A telecommunication network architecture for telemedicine in Bangladesh and its applicability," International Journal of Digital Contents Technology and its Applications, vol. 3, no. 3, September 2009.

[2] Ahsannun Nessa, Moshaddique Al Ameen, Sana Ullah and Kyung Kwak "Applicability of telemedicine in Bangladesh: current status and future prospects," International Arab Journal of Information Technology, vol. 7, no. 2, Page 138-145, April 2010.

[3] Murelli, E., Arvanitis and Theodoros "E-health & learning: The Bangladesh experience," Technology and Health Care, vol.11, issue 5, pp.341-344, 2003.

[4] Vassallo D., Hoque F., Roberts M., Patterson V., Swinfen P., and Swinfen R., "An evaluation of the first year's experience with a low cost telemedicine link in Bangladesh," Computer Journal of Telemedicine and Telecare, vol. 7, no. 3, pp. 125,138, 2001.

[5] Md. Mahfuz Ashrafi, Noushin Laila Ansari, Bushra Tahseen Malik and Barnaly Rashid. (2011). Evaluating the impact of mobile phone based 'health help line' service in rural Bangladesh [online]. Available FTP: http://www.brainstorm-d.com/secured/publications

[6] DGHS, Bangladesh. Telemedicine Services [online]. Available FTP: http://www.dghs.gov.bd/index.php/en/e-health/our-ehealth-eservices/84-english-root/ehealth-eservice/490-telemedicine-service

[7] Abdullah Mamun.(2013). GP launches Telemedicine services in Jessore [Online]. Available FTP: http://archive.thedailystar.net/beta2/news/gp-launches-telemedicine-service-in-jessore/

[8] Debjit Bhowmik, S.Duraivel, Rajnish Kumar Singh and K.P.Sampath Kumar "Telemedicine- an innovating healthcare system in India," The Pharma Innovation Journal, vol. 2, no.4, 2013.

[9] Mishra, S.K., Basnet, R. and Singh, K. "Current telemedicine infrastructure, network, applications in India," 8th IEEE International Conference on e-health networking, application and services, pp.46-49, 2006.

[10] Saroj Kanta Mishra, Indra Pratap Singh and Repu Daman Chand "Current status of telemedicine network in India and future perspective," The proceedings of the Asia-Pacific Advanced Network, vo. 32, pp. 151-163, 2012.

[11] Krishnan Ganapathy and Aditi Ravindra "Telemedicine in India: The Apollo story: Apollo telemedicine networking foundation," Telemedicine and e-health, vol. 15, no. 6, july/august 2009.

[12] R. B. Marasinghe "Telehealth - bringing healthcare to one's doorstep: how ready is Sri Lanka?," Sri Lanka Journal of Bio-Medical Informatics, vol.1,no.3, pp. 124-38, 2010.

[13] Ramesh R. Subedi, Carrie B. Peterson and Sofoklis Kyriazakos "Telemedicine for rural and underserved communities of Nepal," 15th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics, IFMBE Proceedings vol. 34, pp. 117-120, 2011.

[14] Zhao Y, Nakajima I, and Jujoji H, "Onsite investigation of the early phase of Bhutan health telemetric project," J Med Syst, pp. 67-77, 2002.

[15] Mariyam Nazviya and Subrahmanyam Kodukula "Evaluation of critical success factors for telemedicine implementation," International Journal of Computer Applications, vol. 12, no.10, January 2011.

[16] Khawaja, Wasim Shahid,and Lu Xinhai " A case study of telemedicine for disaster management in underdeveloped remote districts of Balochistan, Pakistan," Journal of Economics and Sustainable Development, vol. 4, no. 20, 2013.

[17] Niamat Ullah, Pervez Khan, Najnin Sultana and Kyung Sup kwak "A telemedicine network model for health applications in Pakistan: current status and future prospects," Journal of Digital Content Technology and its Applications, vol. 3, no. 3, September 2009.

A Novel Elliptic Curve Cryptography Scheme Using Random Sequence

Fatema Akhter
IEEE student member
Department of Computer Science and Engineering
Jatiya Kabi Kazi Nazrul Islam University, Bangladesh
Email: fatema.kumu02@gmail.com

***Abstract*—This paper proposes a new Elliptic Curve Cryptography (ECC) scheme and a data mapping technique on elliptic curve over a finite field using maximum length random sequence generation. While its implementation, this paper also proposes a new algorithm of scalar multiplication for ECC. The proposed scheme is tested on various bits length of prime field and the experimental results show very high strength against cryptanalytic attack like random walk and better performance in terms of computation time comparing with standard approaches.**

Keywords*—*cryptography; elliptic curve cryptography; scalar multiplication; random walk; elliptic curve discrete logarithm problem

I. INTRODUCTION

Recently, ECC [1] has gained popularity for its small key length, low processing cost and robust security comparing to other Public Key Cryptography (PKC) systems. These features engrossed the attentions of manufacturers of small processing devices like smart cards, Raspberry computers, wireless devices, pagers, smart phones and tablets [2]. ECC is mainly used for key exchange, digital signature and authentication [3]. However, it can be applicable to any security applications where computational power and integrated circuit space is limited.

The unique idea of ECC was proposed independently by Koblitz and Miller [4] in 1985. Since then on, a lot of attention has been paid to ECC, it has been studied thoroughly and still there are lots of scopes of research. Abdalhossein Rezai [5] et al. proposed an efficient scalar multiplication algorithm for ECC using a New Signed-Digit Representation. D. Sravana Kumar [4] et al. proposed a new encryption algorithm using Elliptic Curve over finite fields. F. Amounas [6] et al. proposed an algorithm to generate a data sequence and applied it on ECC encrypted message over the finite field GF(p). During this time, cryptanalysis of ECC went on with the same pace.

This paper studies the basics of ECC and some existing algorithms on it to move forward to the proposed approach. This paper proposes a new scheme for ECC and new scalar multiplication algorithm for the proposed scheme. Within our knowledge, the same approach has not yet been reported neither for ECC nor for scalar multiplication. The message to point mapping is done using random sequence generation algorithm which increase the security of the proposed scheme to a higher level.

The rest of the paper is organized as follows: Section II describes the preliminary studies for proposed approach. Section III describes the proposed ECC scheme and scalar multiplication algorithm along with the necessary algorithms needed to implement the proposed ECC scheme. Section IV presents the experimental results and discussions for the proposed approach. Section V provides the conclusion and future work.

II. BACKGROUND THEORY AND STUDY

An elliptic curve E over $\mathbb{F}_p$, for a prime $p > 3$ is defined with the short Weierstrass equation [7]

$$E : y^2 = x^3 + ax + b \ with \ x, y, a, b \in \mathbb{F}_p \tag{1}$$

where a, b are integer modulo p, satisfying: $4a^3 + 27b^2 \neq 0$ mod p, and include a point $\mathcal{O}$ called *point at infinity*. The basic condition for any cryptosystem is that the system is closed, i.e., any operation on an element of the system results in another element of the system. In order to satisfy this condition for elliptic curves, it is necessary to construct nonstandard *addition* and *multiplication* operations.

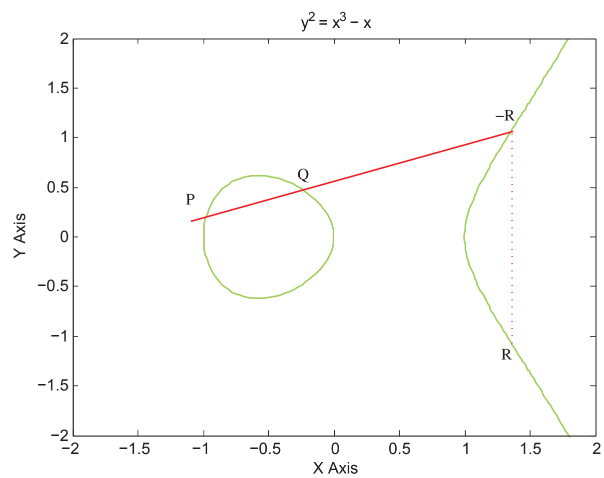


Fig. 1. Description of point addition method on an elliptic curve

A. Geometric Rules of Addition

Let $P(x_1, y_1)$ and $Q(x_2, y_2)$ be two points on the elliptic curve E. The sum $R(x_3, y_3)$ is defined as: first draw a line through P and Q, this line intersects the *elliptic curve* at a

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

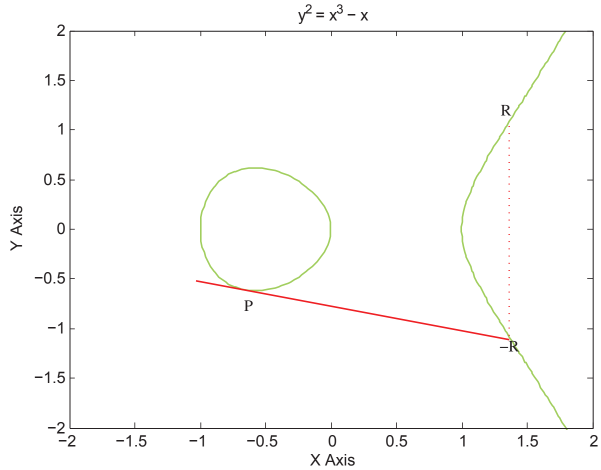


Fig. 2. Description of point doubling method on an elliptic curve

third point. Then the reflection of this *point of intersection* about *X-axis* is R which is the sum of the points P and Q. The same *geometric interpretation* also applies to two points P and P, with the same *X-coordinate*. The points are joined by a *vertical line*, which can be viewed as also intersecting the curve at the *infinity point*. We, therefore, have $P+(-P)=\mathscr{O}$, the identity element which is the *point at infinity*.

B. Point Doubling

First draw the *tangent line* to the *elliptic curve* at P which intersects the curve at a point. Then the reflection of this point about *X-axis* is R. As an example the addition of two points and doubling of a point are shown in Fig.1 and Fig.2 for the elliptic curve $y^2 = x^3 - x$. Point $R(x_3, y_3)$ can be derived as

$$\begin{aligned} x_3 &= \lambda^2 - x_1 - x_2 \\ y_3 &= \lambda(x_2 - x_3)y_2 \end{aligned} \tag{2}$$

where

$$\lambda = \begin{cases} \frac{y_1-y_2}{x_1-x_2} & if P \neq Q \\ \frac{3\times x_1^2+a}{2\times y_1} & if P = Q \end{cases}$$

C. Conventional ECC Scheme

Suppose Alice wants to send a message to Bob and E is an *elliptic curve* over $\mathbb{F}_p$. G is an agreed upon (and publicly known) point on the curve. Bob chooses integer b and calculates $P_b = b\times G$ and makes it public. Alice maps the plaintext m to point M on curve and secretly chooses a random integer k. Alice encrypts M as $C_1 = k\times G$ and $C_2 = M + k\times P_b$. Bob decrypts by calculating $M = C_2 - b\times C_1 = C_2 - b\times kG = M + k\times P_b - k\times P_b = M$.

D. Random Number Generation

Random number can easily be generated using *Linear-Feedback Shift Registers (LFSR)* [8] from maximum length polynomial. For polynomial $f(x) = x^4 + x + 1$, it has a shift register of length $m = 4$. So, it can produce a sequence of length $2^m - 1$, i.e.,15. In the numbers, sequence of bits appears random and has a very long cycle. For the given polynomial, random number sequence can be generated by calculating

$$x^i \ mod \ f(x) \ for \ i = 0,1,2,\ldots,14. \tag{3}$$

The generated random number sequence will be like Fig. 3. Stream of values produced by registers in LFSR is completely determined by its current or previous states and the *Exclusive-OR* operation.

0 0 0 1 (=01)
0 0 1 0 (=02)
0 1 0 0 (=04)
1 0 0 0 (=08)
0 0 1 1 (=03)
0 1 1 0 (=06)
1 1 0 0 (=12)
1 0 1 1 (=11)
0 1 0 1 (=05)
1 0 1 0 (=10)
0 1 1 1 (=07)
1 1 1 0 (=14)
1 1 1 1 (=15)
1 1 0 1 (=13)
1 0 0 1 (=09)

Fig. 3. Random number sequence.

III. Proposed ECC Approach

This section describes our proposed ECC scheme along with proposed and necessary algorithms to implement the approach.

A. Proposed ECC Scheme

Suppose Alice and Bob want to communicate using ECC. They have to agree on some issues related to elliptic curve parameters and base point. The proposed ECC scheme for covert communication is described in Algorithm 1.

Algorithm 1 Proposed ECC Scheme

1: Both agree on curve $E: y^2 = x^3 + Ax + B$ on large prime field $\mathbb{F}_p$ for prime p and a common point G.
2: Alice chooses a *random number* a and a *random point* A on the curve, keeps them as her *private key* $PrA\{a, A\}$. She calculates $PuA_1 = aA$ and $PuA_2 = a(A+G)$ and made them public.
3: Bob chooses *random number* b and a *random point* B on the curve, keeps them as his *private key* $PrB\{b, B\}$. He calculates $PuB_1 = bB$ and $PuB_2 = b(B+G)$ and made them public.
4: If Alice wants to send message M to Bob, then Alice encrypts the message in following way
   a) maps message M to point P_M using *random sequence generation method*, $M \rightarrow P_M$.
   b) generate a *random number* k.
   c) Calculate $C_1 = (a+k)G$ and $C_2 = P_M + (a+k)PuB_2 - (a+k)PuB_1$.
   d) Alice sends $\{C1, C2\}$ to Bob.
5: Bob decrypt the message in the following way
   a) Calculate $P_M = C_2 - bC_1$.
   b) Then message is derived, $P_M \rightarrow M$.

B. Proposed Scalar Multiplication Algorithm

The efficiency of an ECC implementation mainly depends on the way it implements the Scalar or Point Multiplication. Most of the existing algorithms focus on the minimization of Hamming weight [9] of the given value by converting it to binary or sign binary numbers [10]. The proposed algorithm also works with a view to making the hamming weight minimal choosing whatever is suitable between sign binary or binary multiplication without conversion overhead. The proposed Scalar Multiplication algorithm is described in Algorithm 2.

Algorithm 2 Scalar Multiplication, kP

procedure SCALAR MULTIPLICATION(k, P)
  $R \leftarrow 0$
  $S \leftarrow 1$
  while $k > 0$ **do**
    $x \leftarrow \lfloor log_2 k \rfloor$
    if $(k - 2^x) > (2^{x+1} - k)$ **then**
      $R \leftarrow R + (s)2^{x+1}.P$
      $k \leftarrow 2^{x+1} - k$
      $s \leftarrow -1$
    else
      $R \leftarrow R + (s)2^x.P$
      $k \leftarrow k - 2^x$
    end if
  end while
  Return R
end procedure

C. Message to Point Mapping

This section describes how a message M is mapped to point P_M. First random sequences from *maximum length polynomial* are generated. For random sequence, this paper used LFSR technique on the polynomial $x^7 + x^6 + 1$ to generate random sequences. This polynomial has maximum period of 127 values ranging from 1 to 127. So it can represent 127 characters without any repetition. This paper uses only alphanumeric letters where every letter is assigned a value in the order it is generated in the sequence. For letters starts with numbers $[0-9]$, small letters $[a-z]$ and capital letters [A-Z] .

Algorithm 3 Random Sequence Generation

procedure RANDOMSEQUENCEGENERATOR$(x^7 + x^6 + 1)$
  $flag \leftarrow 0x01$
  $data \leftarrow flag$
  $index \leftarrow 1$
  do
    $newbit \leftarrow ((data \gg 6) \wedge (data \gg 5))$ AND $1)$
    $data \leftarrow ((data \ll 1) \vee newbit)$ AND $0x7f$
    $R_{index} \leftarrow data$
    $index \leftarrow index + 1$
  while $data \neq flag$
  return R
end procedure

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

This section presents the computational results of the proposed approach. The approach presented in this paper is coded using C on an Intel laptop with speed of 2.13 GHz and 2GB of RAM under *Ubuntu 14.04 LTS* using *gcc - 4.9* compiler. For the operations of large bits this paper uses GMP library [11] , version-6.0.0a.

A. Experimental Results of Proposed ECC Scheme

Conventional ECC uses only one secret number as a private key and one point on the curve as a public key. On the other hand, the proposed ECC scheme uses one number with a point on the curve as private key and two points on

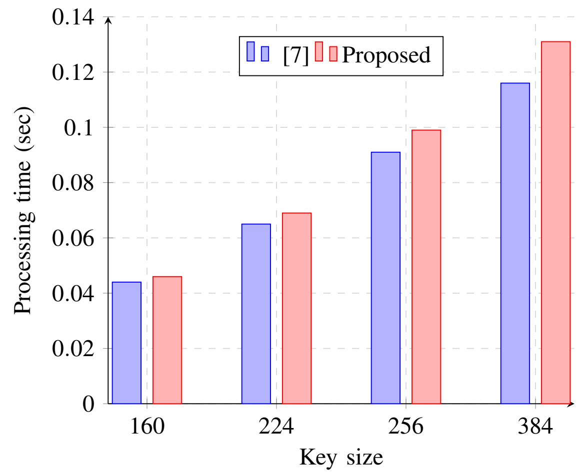


Fig. 4. Comparison of key generation cost between [7] and proposed ECC

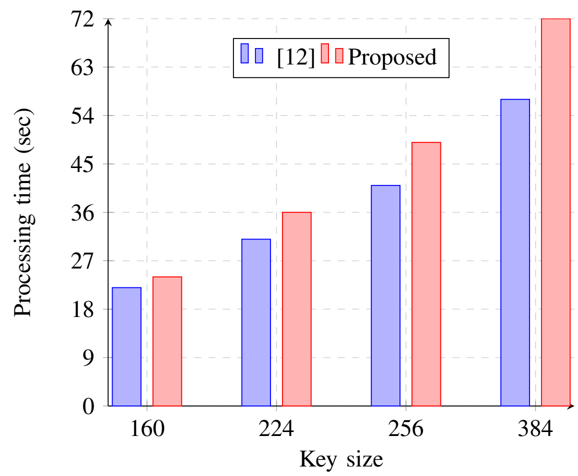


Fig. 5. Comparison of encryption cost between [12] and proposed ECC

the curve as public key. The additional points increase the computational cost but strengthen the security higher than the conventional ECC. The comparison of computational costs of Key generation, ECC encryption and ECC decryption methods of proposed scheme with the conventional ECC scheme are presented in Fig. 4, Fig. 5 and Fig. 6.

B. Experimental Results of Proposed Scalar Multiplication

The efficiency of an ECC scheme depends largely on the scalar multiplication. Most of the existing algorithms have the overhead of converting the scalar number to binary or sign binary presentation to minimize the Hamming Weight. Proposed algorithm doesn't have such overhead except it needs pre-computed doublings and minimizes the Hamming Weight. The computational cost of the proposed Scalar Multiplication is shown in Fig. 7.

C. Resistance against Attack on Proposed ECC Scheme

The standard attack on ECC is *Random Walk* [14] which uses *Pollard-Rho* method for solving *Elliptic Curve Discrete*

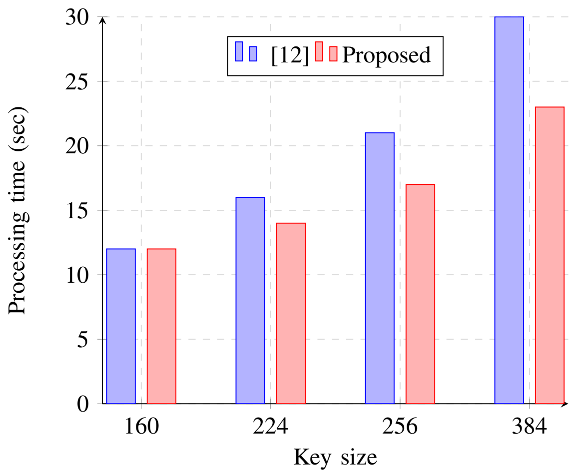


Fig. 6. Comparison of decryption cost between [12] and proposed ECC

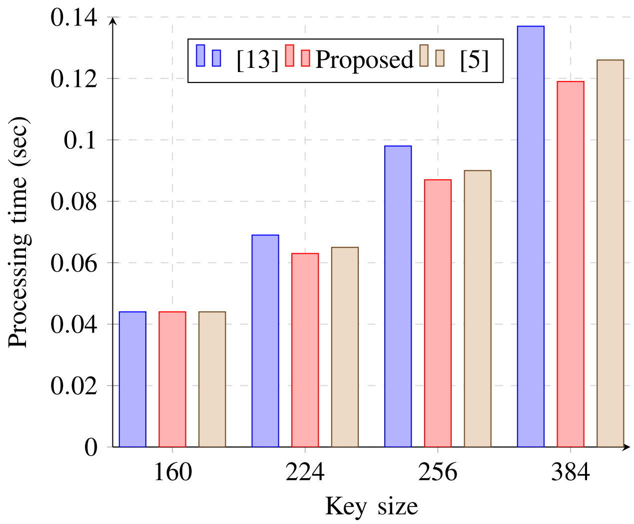


Fig. 7. Comparison of processing cost of scalar multiplication among [13], proposed ECC and [5]

TABLE I. COMPARISON OF COST REQUIRED TO BREAK ECC

Key Size	ECC [15]		Proposed ECC	
	Cost	MIPS years	Cost	MIPS years
160	2^{80}	9.6×10^{11}	2×2^{80}	1.92×10^{12}
186	2^{93}	7.9×10^{15}	2×2^{93}	1.58×10^{16}
234	2^{117}	1.6×10^{23}	2×2^{117}	3.2×10^{23}
354	2^{177}	1.5×10^{41}	2×2^{177}	3.0×10^{41}
426	2^{213}	1.0×10^{52}	2×2^{213}	2.0×10^{52}

Logarithm Problem (ECDLP). Pollard-Rho proved that the expected running time of the method is $\sqrt{(\pi \times n)/2}$ steps, where a step here is an elliptic curve addition. As the proposed algorithm has two different secret keys the expected running time will be twice of the stated cost. So the expected cost of breaking the proposed ECC scheme will roughly be $2 \times \sqrt{(\pi \times n)/2}$ steps. A MIPS (Million Instructions Per Second) year is presented as the computational power of a computer that is rated at 1 MIPS and utilized for one year. The comparison of computation cost required to break the proposed ECC with the conventional ECC [15] are presented in Table I.

V. Conclusion and Future Work

This paper presents a new ECC scheme that provides higher security than the existing scheme. While implementing the proposed scheme, this paper also proposes a new scalar multiplication algorithm that takes the advantages of both binary and signed binary presentation and have less processing cost. Future study aims to integrate the advantages of double scalar multiplications in proposed scheme to minimize the processing cost. Then the proposed scheme will be evaluated on different key sizes against different cryptographic attacks for further improvement of the proposed scheme.

References

[1] J. W. Bos, J. A. Halderman, N. Heninger, J. Moore, M. Naehrig and E. Wustrow, “Elliptic curve cryptography in practice,” Financial Cryptography and Data Security, LNCS vol. 8437, pp. 157-175, Springer Berlin Heidelberg, March 2014.

[2] Z. Liu, J. Groschdl and D. S. Wong, “Low-weight primes for lightweight elliptic curve cryptography on 8-bit AVR processors,” Information Security and Cryptology, LNCS vol. 8567, pp. 217-235, Springer International Publishing, January 2014.

[3] H. L. Yeh, T. H. Chen and W. K. Shih, “Robust smart card secured authentication scheme on SIP using elliptic curve cryptography,” Computer Standards and Interfaces, vol. 36, no. 2, pp. 397-402, February 2014.

[4] D. S. Kumar, C. H. Suneetha and A. Chandrasekhar, “Encryption of data using Elliptic Curve over finite fields,” International Journal of Distributed and Parallel Systems, vol. 3, no.1, January 2012.

[5] A. Rezai and P. Keshavarzi, “An Efficient Scalar Multiplication Algorithm for Elliptic Curve Cryptography Using a New Signed-Digit Representation,” Advanced Science and Technology Letters, vol. 44, pp. 44-48, 2013.

[6] F. Amounas and E. H. El Kinani, “ECC Encryption and Decryption with a Data Sequence,” Applied Mathematical Sciences, vol. 6, no. 101, pp. 5039-5047, January 2012.

[7] D. Hankerson, A. J. Menezes and S. Vanstone, “Guide to Elliptic Curve Cryptography,” Springer-Verlag, 2004.

[8] A. Canteaut, “Linear Feedback Shift Register,” Encyclopedia of Cryptography and Security, pp. 726-729, Springer US, 2011.

[9] R. M. Avanzi, C. Heuberger and H. Prodinger, “Minimality of the Hamming weight of the -NAF for Koblitz curves and improved combination with point halving,” Selected Areas in Cryptography, LNCS vol. 3897, pp. 332-344, Springer Berlin Heidelberg, January 2006.

[10] K. Koyama and Y. Tsuruoka, “ Speeding up elliptic cryptosystems by using a signed binary window method,” Advances in cryptology-CRYPTO92, pp. 345-357, Springer Berlin Heidelberg, January 1993.

[11] J. Liu, Y. H. Lu and C. K. Koh, “Performance Analysis of Arithmetic Operations in Homomorphic Encryption,” ECE Technical Reports, Purdue University, 2010. Online available at: https://gmplib.org/.

[12] I.F. Blake, G. Seroussi and N. P. Smart, “Advances in elliptic curve cryptography, ” vol. 317, Cambridge University Press, 2005.

[13] P.Balasubramaniam and E. Karthikeyan, “Elliptic curve scalar multiplication algorithm using complementary recoding,” Applied mathematics and computation, vol. 190, no.1, pp. 51-56, July, 2007.

[14] F. Zhang and P. Wang, “Speeding up elliptic curve discrete logarithm computations with point halving,” Designs, codes and cryptography, vol. 67, no. 2, pp. 197-208, Springer US, 21 December 2011.

[15] D.B. Johnson and A.J. Menezes, “Elliptic curve DSA (ECDSA): an enhanced DSA,” Proceedings of the 7th conference on USENIX Security Symposium. vol. 7, pp. 13-23, 1998.

VANET Topology Based Routing Protocols & Performance of AODV, DSR Routing Protocols in Random Waypoint Scenarios

Aditi Roy
Department of Computer Science and Engineering, Shahjalal University of Science and Technology, Bangladesh
aditiroy1710@gmail.com

Bijan Paul
Lecturer, Department of Computer Science, American International University Bangladesh
bijancse@gmail.com

Sanjit Kumar Paul
Associate Professor, Department of Mathematics, University of Asia Pacific, Bangladesh
sanjitkpaul@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 discusses briefly about two on demand routing protocols AODV and DSR procedure. Section 3 describes connection types like TCP and CBR. Section 4 presents performance metrics and the network parameters. Section 5 presents our implementation. Section 6 presents experimental analysis. We conclude in Section 7 and with the references at the end.

II. 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 [7] & DSR [8].

A. AODV

AODV [7] 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).

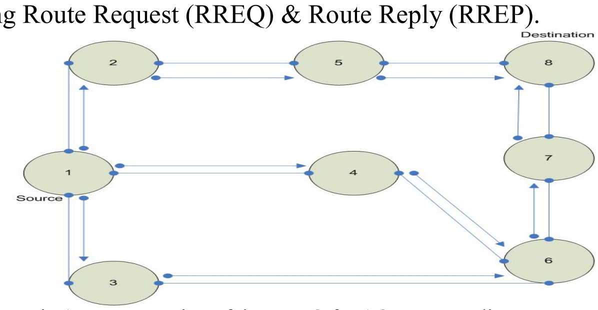


Fig 1. Propagation of the RREQ for AODV route discovery

B. DSR

The Dynamic Source Routing (DSR) [8] protocol has two phases route discovery & route maintenance. In DSR, 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.

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

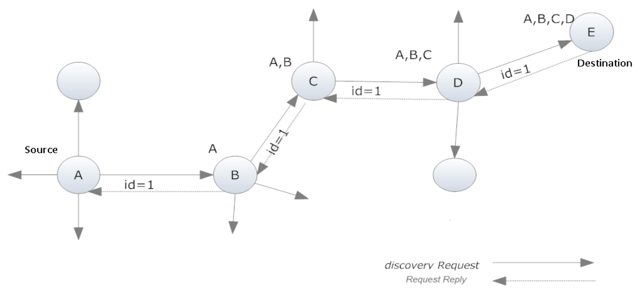


Fig 2. DSR Route Discovery

III. 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.

IV. PERFORMANCE METRICS & NETWORK PARAMETERS

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***

V. OUR IMPLEMENTATION

For implementing our simulation we used Network Simulator NS-2.34 [9, 10]. 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 [10] 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.

PDR and LPR Measure AWK Script

```
START
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
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. In graph, Blue color is for AODV and Red color is for DSR.

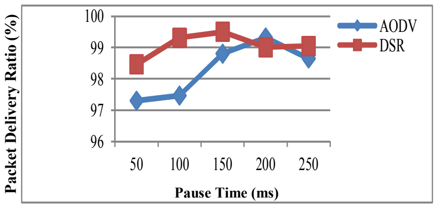


Fig 3. PDR (w.r.t. Pause) of 30 nodes using TCP

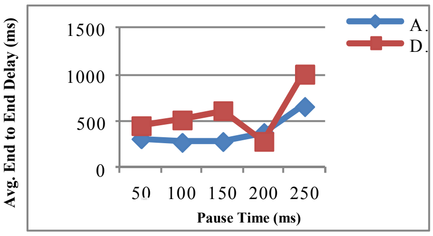


Fig 4. Avg.E2E delay (w.r.t. Pause) of 30 nodes using TCP

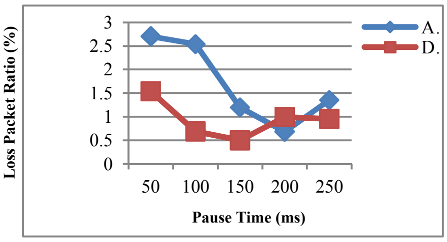


Fig 5. LPR (w.r.t. Pause) of 30 nodes using TCP

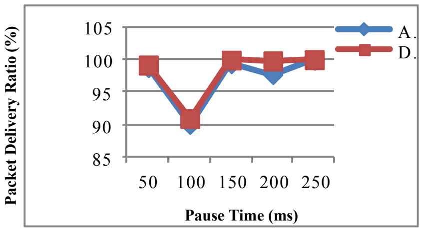


Fig 6. PDR (w.r.t. Pause) of 30 nodes using CBR

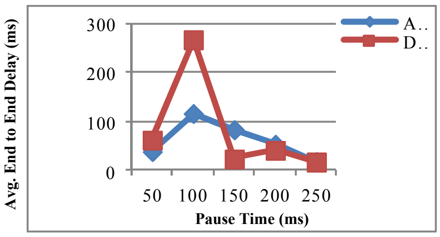


Fig 7. Avg.E2E delay(w.r.t. Pause) of 30 nodes using CBR

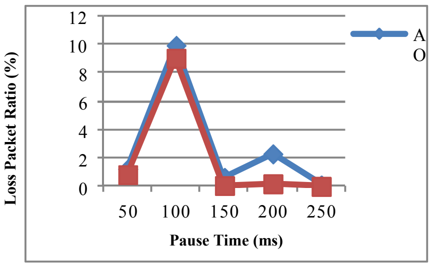


Fig 8. LPR (w.r.t. Pause) of 30 nodes using CBR

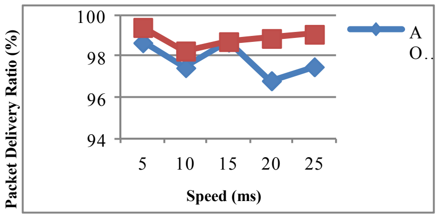


Fig 9. PDR (w.r.t. Speed) of 30 nodes using TCP

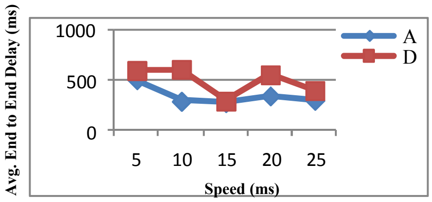


Fig 10. Avg.E2E delay (w.r.t. Speed) of 30 nodes using TCP

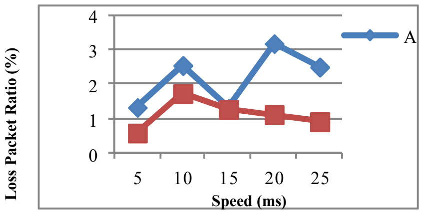


Fig 11. LPR (w.r.t. Speed) of 30 nodes using TCP

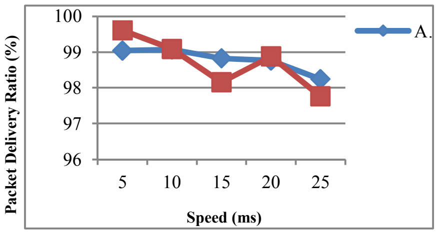


Fig 12. PDR (w. r.t. Speed) of 30 nodes using CBR

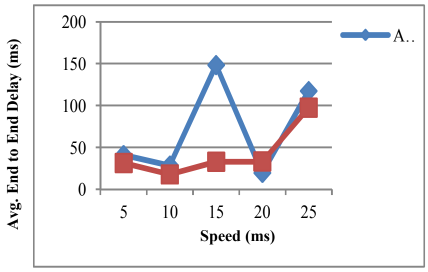


Fig 13. Avg.E2E delay (w.r.t. Speed) of 30 nodes using CBR

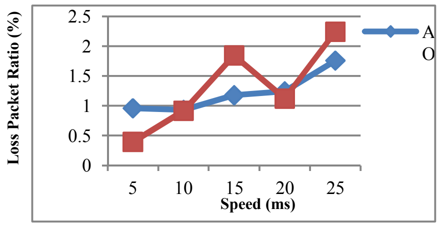


Fig 14. LPR (w.r.t. Speed) of 30 nodes using CBR

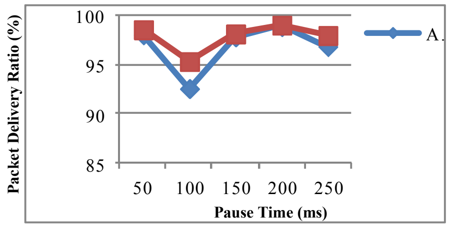


Fig 15. PDR (w.r.t. Pause) of 150 nodes using TCP

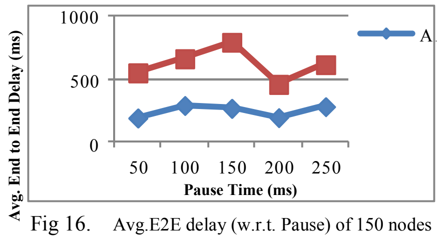


Fig 16. Avg.E2E delay (w.r.t. Pause) of 150 nodes using TCP

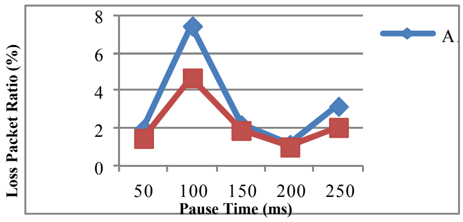


Fig 17. LPR (w.r.t. Pause) of 150 nodes using TCP

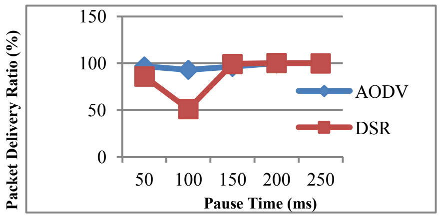


Fig 18. PDR (w.r.t. Pause) of 150 nodes using CBR

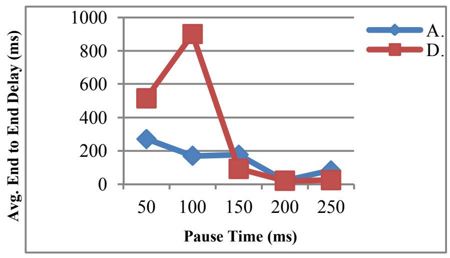


Fig 19. Avg.E2E delay (w.r.t. Pause) of 150 nodes using CBR

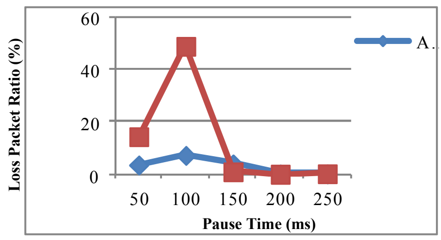


Fig 20. LPR (w.r.t. Pause) of 150 nodes using CBR

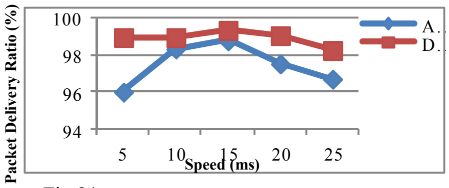


Fig 21. PDR (w.r.t. Speed) of 150 nodes for TCP

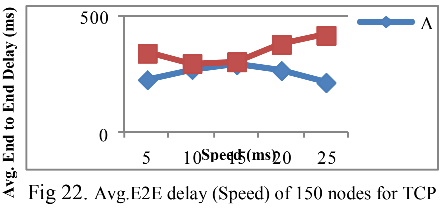


Fig 22. Avg.E2E delay (Speed) of 150 nodes for TCP

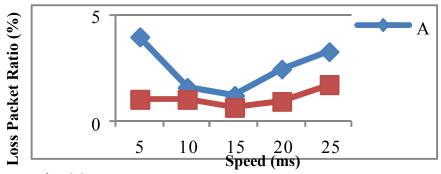


Fig 23. LPR (w.r.t. Speed) of 150 nodes using TCP

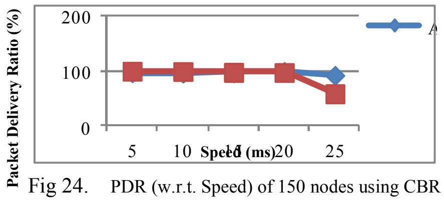


Fig 24. PDR (w.r.t. Speed) of 150 nodes using CBR

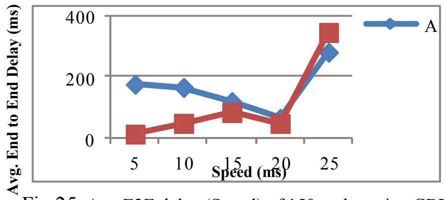


Fig 25. Avg.E2E delay (Speed) of 150 nodes using CBR

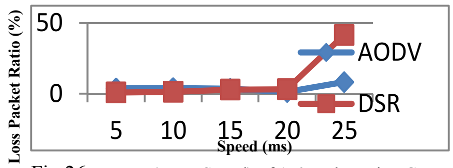


Fig 26. LPR (w.r.t. Speed) of 150 nodes using CBR

VI. Experimental Analysis

A. *Pattern analysis of 30 nodes using TCP and CBR connection comparison*

From our experimental analysis, for TCP connection using pause time as a parameter in low mobility low pause time PDR is average for AODV and high for DSR. But in CBR, PDR of both protocols is high. E-2-E Delay for TCP is average for AODV and high for DSR. But in CBR, PDR of both protocols is low. The LPR for TCP connection is high for AODV and average for DSR. But CBR, it is average for AODV and low for DSR. If the pause time is high the PDR for both routing protocols in is high not only for TCP but also for CBR. E-2-E Delay in TCP is high but low for CBR for both protocols. LPR is low for DSR both TCP and CBR. But for AODV, LPR is average for TCP and low for DSR.

On the other hand, using speed as a parameter in low mobility low speed PDR for both protocols is high for both TCP and CBR connection. Average E-2-E Delay is high for TCP but low for CBR for both routing protocol. LPR is low for both routing protocol not only in TCP but also CBR. But in low mobility high speed for TCP connection, PDR for AODV is average but high for CBR connection. For TCP connection, PDR for DSR is high but average for CBR. E-2-E Delay for TCP is high but low for CBR for both protocols. LPR is average for AODV not only for TCP but also DSR. But for DSR LPR is low for TCP and average for CBR.

B. *Pattern analysis of 150 nodes using TCP connection comparison*

Pause time as a parameter in high mobility low pause time for TCP, PDR for both protocols is high but in CBR it is average for AODV and low for DSR. Both TCP and CBR average E-2-E Delay is average for AODV and high for DSR. For TCP, The LPR is average for both protocols but high for CBR. If the pause time is high the PDR for both routing protocols is average for TCP but high for CBR. For TCP, E-2-E Delay is average for AODV and high for DSR and LPR is high for AODV and DSR. But in CBR, 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, PDR of AODV is average but high for DSR for both TCP and CBR. In TCP E-2-E for AODV & DSR is high, LPR is low for DSR and high for AODV. But for CBR, E-2-E and LPR for AODV is high but low for DSR. If the speed is high in TCP, AODV performs average and DSR performs high. But for CBR, PDR for AODV and DSR is low. Both TCP and CBR, E-2-E is high for both routing protocol. For TCP, LPR of AODV is high but for DSR it is average. LPR of AODV and DSR is high for CBR connection.

VII. 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 & observe that 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.

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] Perkins, C.; Belding-Royer, E.; Das, S. (July 2003)”Ad hoc On-Demand Distance Vector (AODV) Routing”.

[8] 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.

[9] http://www.isi.edu/nsnam/ns/

[10] http://www.isi.edu/nsnam/ns/tutorial/

Performance Analysis of DCS-Based Limited Wavelength Interchanging Cross-Connects in WDM Network

Punab Chandra Kundu[1], Md. Rabiul Islam[2], Pejush Chandra Sarker[2] and Satya Prasad Majumder[3]
[1]Bangladesh Power Development Board, Bangladesh
[2]Department of Electrical and Electronic Engineering, Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh
[3]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering & Technology, Dhaka, Bangladesh
punabkundu97@yahoo.com, rabiulbd@hotmail.com, pejush01@yahoo.com, spmajumder@eee.buet.ac.bd

***Abstract*—Crosstalk originated from various types of sources in delivery and coupling switches (DCS) based limited wavelength interchanging optical cross connect (OXC) architectures has been studied and analytical formulas have been carried out for those crosstalk. A comparative picture of the system performance has been depicted for all cases in DCS based limited wavelength interchanging OXC architectures which will help the system designer to choose the effective one. The system performance has been analyzed in MATLAB environment.**

Keywords—DCS; L-WIXC; WDM; power penalty; multiplexer; demultiplexe; crosstalk

I. Introduction

In wavelength routed all-optical networks, the wavelength sensitive routing function is controlled by optical cross connect (OXC) in an automated manner without having to restore to performing manual patch panel connections [1]–[3]. It offers routing scalability, bit rate and protocol independence, and power saving and increased transport capacity to wavelength division multiplexing (WDM) network [4], [5].

According to [4], improvement of call blocking performance saturates when the number of converters in an OXC is greater than a threshold. So, instead of wave length interchanging cross connect having full conversion capability limited wavelength interchanging cross-connect (L-WIXC) having optimum number of wavelength converters may be used through sharing due to the cost and complexity of wavelength converters [1].

The buildup of crosstalk noise on a certain optical channel due to interference with other signals while propagating through the different element of the WDM network could result in serious problems. Crosstalk due to optical cross connect is one of the basic criteria that characterizes the performance of the WDM network [3], [5]–[7]. Since, optical crosstalk is a major limiting factor, the commercial use of an optical OXC is so far prevented in WDM networks.

Practical implementation of the OXCs often employs multi-stage structures to achieve the required size with less complexity. The architectures of the optical cross-connects have a significant impact on how the unwanted light leaking from the components mix with the actual signal to become crosstalk. In [2], a systematic analysis of such crosstalk has been reported for OXC without wavelength converter, in [3] for OXC with wavelength converter and in [1] for limited wavelength-interchanging cross-connect only for two architectures (share per node & share per link which is based on space switching matrix) but till today, no such analysis has been reported in the literature for delivery and coupling switches (DCS) based limited wavelength OXC. Systematic analysis of crosstalk is a tool for characterizing the crosstalk performance of optical cross connect architectures and enable them to compare base on some key performance indicators. Also, by the analysis, it will be possible to relate the overall crosstalk to the specifications of individual optical components and, hence, enable system architects to specify certain component specifications for achieving a desired crosstalk performance for the chosen OXC architecture. In this paper, a systematic analysis of signal distortion crosstalk and their impact on the desired signal will be considered for DCS based limited wavelength OXC architecture. Here, crosstalk has been modeled as a Gaussian random process.

II. Different Optical Components of DCS based OXC Architecture

A. DCS-1

In DCS-1, the actual signal in question is combined with *M-1* signals at different wavelengths by a star coupler at the output. These signals carry with them crosstalk components having the same wavelength as the actual signal. Due to imperfection, *M-1* crosstalk components at ω are leaked and mixed with the actual signal at the star coupler after passing through the switches. Let X_i ($i=1,2,\ldots\ldots,N$) be the number of crosstalk components at ω that are leakages of the signal entering the OXC at input port i, than, each X_i is an integer satisfying

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

$$\sum_{i=1}^{N} X_i = M-1, \qquad 1 \le X_i \le M-1 \tag{1}$$

In a real DCS based switch, each crosstalk will leak to the unattended outputs including crosstalk to the actual signal. There are *N-1* in band crosstalk components at ω leaking from the first stage switches. These are contributed by signal entering from different input ports.

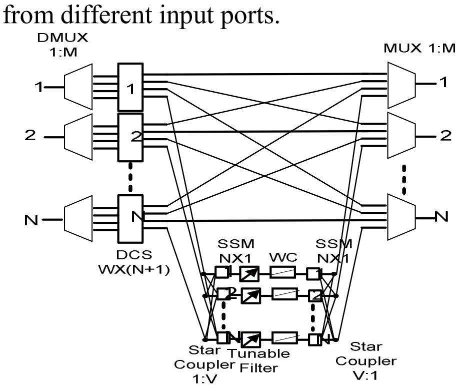


Fig. 1. DCS-1 OXC architecture

In the wavelength converter for a signal converted to ω, there must be another signal originally at ω being converted to another wavelength. Therefore, in the worst case, assuming V converters, there are *K* crosstalk components, with *K*= min (*N-1,* $\lfloor V/2 \rfloor$) which leak from the second stage switches. The converted signal is free from the crosstalk carried with it before wavelength conversion.

Assuming that the OXC is fully loaded, in the worst case, the actual signal will be interfered by *M-1* crosstalk components leaking from the multiplexer and switch imperfection, another *N-1* crosstalk components leaking from the switches and *K* components leaking from the second stage switches. The components traverse the OXC via different paths and, thus, have different propagation delays. Assuming intensity modulation, the electric field, which includes the influence of the crosstalk, is given by

$$\overrightarrow{E(t)} = Eb_1(t)\cos[\omega t + \varphi_1(t)]\ \overrightarrow{P1} +$$
$$\sum_{i=1}^{N}\sum_{j=1}^{Xi} \sqrt{\delta} Eb_i(t-\tau_{ij})\cos[\omega(t-\tau_{ij}) + \varphi_i(t-\tau_{ij})]\ \overrightarrow{Pij}$$
$$+ \sum_{i=2}^{N} \sqrt{\varepsilon} Eb_i(t-\tau_{ix})\cos[\omega(t-\tau_{ix}) + \varphi_i(t-\tau_{ix})]\ \overrightarrow{Pix} +$$
$$\sum_{i=2}^{K} \sqrt{\varepsilon'} Eb'_i(t-\tau'_i)\cos[\omega(t-\tau'_i) + \varphi'_i(t-\tau'_i)]\ \overrightarrow{P'_i} \tag{2}$$

where, τ_{ij}, $\overrightarrow{Pij}$,τ_{ix}, $\overrightarrow{Pix}$,δ and ε are as follows.

τ_{ij} = propagation delay relative to the actual signal of the jth crosstalk component leaking from the signal entering from port *i* of the OXC at the tunable filter.

$\overrightarrow{Pij}$ = the unit polarization vector of crosstalk component τ_{ij}.

τ_{ix} = propagation delay relative to the actual signal of the crosstalk component leaking from the signal entering from port *i* of the OXC at the first stage optical switch.

$\overrightarrow{Pix}$ = the unit polarization vector of crosstalk component τ_{ix}.

τ'_i = propagation delay relative to the actual signal of the crosstalk component leaking from the signal entering from port *i* of the second stage optical switch

$\overrightarrow{P_{i'}}$ = the unit polarization vector of crosstalk component τ'_i.

δ= optical power relative to the actual signal for the crosstalk components leaked at a tunable filter.

ε = optical power relative to the actual signal for the crosstalk components leaked at first stage optical switch.

ε' = optical power relative to the actual signal for the crosstalk components leaked at second stage switch.

As [5] and [9] which can be reduced to the mean *E(j)*=1 and the variance

case 1: $\tau < \tau_{coherent}$ and $\tau \ll T$, bit period or coherent case

$$\max(\sigma^2) = \max\ [\frac{2}{3}\{\sum_{i=2}^{N}(X_i\sqrt{\delta}+\sqrt{\varepsilon})^2 + K\varepsilon'\}]$$
$$= \frac{2}{3}[(M-1)\sqrt{\delta}+\sqrt{\varepsilon}]^2 + \frac{2}{3}(N-2)\varepsilon + \frac{2}{3}k\varepsilon' \tag{3}$$

case 1: $\tau < \tau_{coherent}$ and $\tau \ll T$, bit period or coherent case

$$\max(\sigma^2) = \frac{1}{2}[(M-1)\sqrt{\delta}+\sqrt{\varepsilon}]^2 + \frac{1}{6}(4N-7)\varepsilon +$$
$$\frac{1}{6}(M-1)\delta + \frac{2}{3}k\varepsilon' \tag{4}$$

B. DCS-2

This is the second type of DCS-based OXC. It has the capability of connecting multiple inputs to one output, *M* DCSs are sufficient to avoid blocking of the unconverted light paths. Furthermore, if a wavelength converter is available in this architecture, then any converted light path can find its route through the OXC without the link mismatch problem. In this architecture, the incoming channels are separated by the combination of input demultiplexer and tunable filters. Then, they go through the wavelength conversion process for non-wavelength continuous light paths. For wavelength continuous light path they need not go through the wavelength conversion process. They are then routed by the DCS directly to the proper output [8].

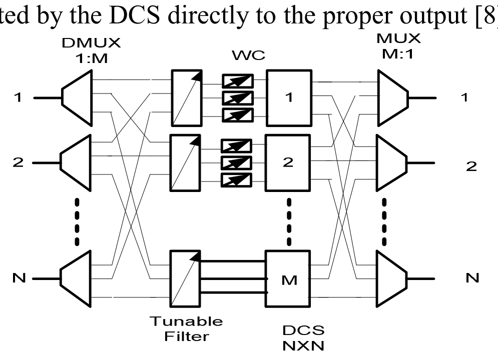


Fig. 2. DCS-2 OXC architecture

As like the previous case, in this architecture, the actual signal is also combined with *M-1* signals at different wavelengths by a multiplexer at the output. These signals also carry with them in band crosstalk components which can be traced back to the tunable filters. Due to imperfect filtering, *M-1* crosstalk components at ω leak through the filters. The signals which go through the conversion process are free from crosstalk. Let X_i ($i=1,2,\ldots\ldots,N$) be the number of crosstalk components at ω that are leakages of the signal entering the OXC at input port *i,* than, each X_i is an integer satisfying

$$\sum_{i=1}^{N} X_i = M - V_n, \qquad 1 \le X_i \le M - V_n. \tag{5}$$

In this case, we also assume that the OXC is fully loaded, in the worst case, the actual signal will be interfered by M-V_n crosstalk components leaking from the tunable filters and another *N-1* crosstalk components leaking from the switches. The components traverse the OXC via different paths and thus, have different propagation delays. Assuming intensity modulation, the electric field which includes the influence of the crosstalk, is given by

$$\overrightarrow{E(t)} = Eb_1(t)\cos[\omega t + \varphi_1(t)]\,\overrightarrow{P_1} + \sum_{i=1}^{N}\sum_{j=1}^{X_i}\sqrt{\delta}Eb_i(t-\tau_{ij})\cos[\omega(t-\tau_{ij}) + \varphi_i(t-\tau_{ij})]\,\overrightarrow{P_{ij}}$$

$$+\sum_{i=2}^{N}\sqrt{\varepsilon}Eb_i(t-\tau_{ix})\cos[\omega(t-\tau_{ix})+\varphi_i(t-\tau_{ix})]\,\overrightarrow{P_{ix}} \tag{6}$$

As [5] and [9] which can be reduced to the mean *E(j)*=1 and the variance

Case 1: $\tau < \tau_{\text{coherent}}$ and $\tau \ll T$, bit period or coherent case

$$\max(\sigma^2) = \frac{2}{3}[(M-V_n)\sqrt{\delta}+\sqrt{\varepsilon}]^2 + \frac{2}{3}(N-2)\varepsilon + \frac{2}{3}(N-2)\varepsilon \tag{7}$$

Case 2: $\tau < \tau_{\text{coherent}}$ and $\tau > T$, bit period or incoherent case :

$$\max(\sigma^2) = \frac{1}{2}[(M-V_n)\sqrt{\delta}+\sqrt{\varepsilon}]^2 + \frac{1}{6}(4N-7)\varepsilon + \frac{1}{6}(M-V_n)\delta \tag{8}$$

III. Results and Discussion

From Fig. 3, it is seen that both the crosstalk power and BER increases as input power or number of wavelength per fiber is increased. In DCS-2 architecture, the signal suffers from more crosstalk as well as BER is high when there is no wavelength converter (WC). But in case of DCS-1 architecture, crosstalk power and BER are high when there are wavelength converters because in this case, a portion of the signals that have been converted by wavelength converters is present in the output. Fig. 4 shows variation of crosstalk power and BER with number of wavelength per fiber *M* for different number of input fiber *N*. Both the crosstalk power and BER increase with both number of wavelength per channel and number of fiber. From the Fig. 4, it is also seen that for a certain throughput (*N* multiplied with *M*) , lowest crosstalk is obtained with large *N* and small *M*.

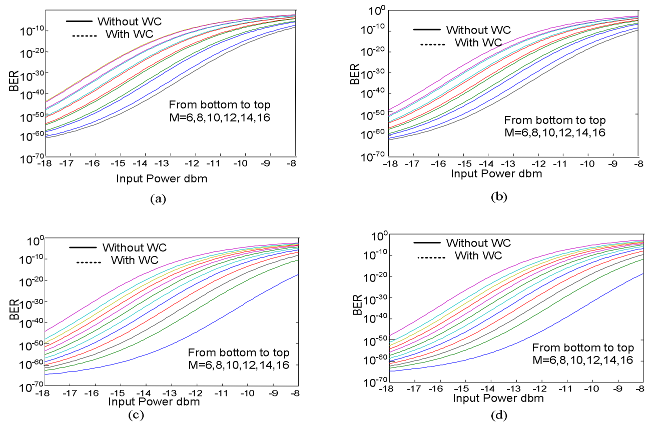


Fig. 3. Variation of BER with input power for different number of wavelength per fiber for (a) DCS-1 OXC architecture coherent case (b) DCS-1 OXC architecture incoherent case (c) DCS-2 OXC architecture coherent case (d) DCS-2 OXC architecture incoherent case.(P_r = -15 dbm, V_n = 4, δ =-20 dbm, ε = -20 dbm and ε′= -20 dbm, N=8).

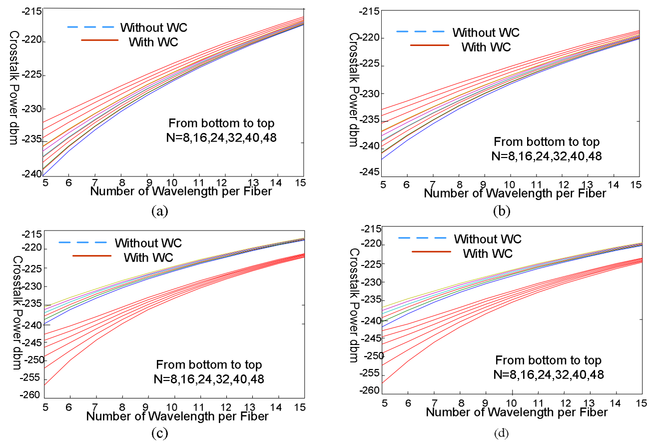


Fig. 4. Variation of crosstalk power with different number of input wavelength per fiber for different Number of input fiber for (a) DCS-1 OXC architecture coherent case (b) DCS-1 OXC architecture incoherent case (c) DCS-2 OXC architecture coherent case and (d) DCS-2 OXC architecture incoherent case.(P_{in} =-18 dbm, P_r = -19 dbm, V_n = 4, δ =-20 dbm, ε = -25 dbm and ε′= -25 dbm.)

Fig. 5 compares variation of crosstalk power and BER with number of input fiber with and without WC. From the figures, it is also seen that crosstalk power and BER increase as the number of input fiber is increased. DCS-1 OXC architecture posses higher crosstalk power and BER with wavelength converter. From that Fig. 6, it is seen that both crosstalk power and BER increase as the number of input wavelength per fiber is increased in both with and without WC cases. These figures also show that when there are wavelength converters DCS-1 architectures suffer from more crosstalk than DCS-2 architecture. So, they involve more BER.

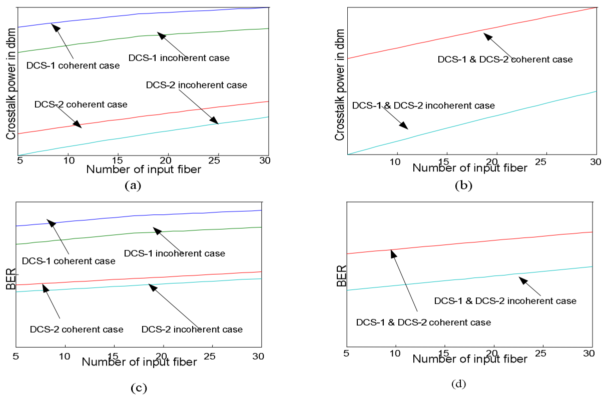


Fig. 5. Comparison of variation of (a) crosstalk power with WC (b) crosstalk power without WC (c) BER with WC (d) BER without WC with number of input fiber for DCS-1 & DCS-2 OXC architectures for both coherent & incoherent cases. (P_{in} = -16 dbm, P_r = -20 dbm, M=8, , V_n = 4 i.e V =32, δ =-20 dbm, ε = -25 dbm and ε′= -25 dbm).

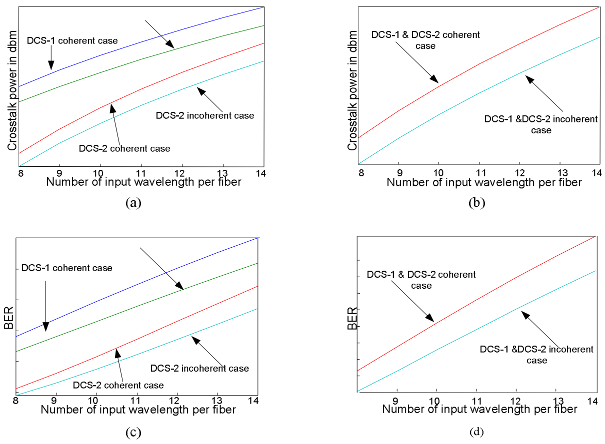


Fig. 6. Comparison of variation of (a) crosstalk power with WC (b) crosstalk power without WC (c) BER with WC (d) BER without WC with number of input wavelength per fiber M for DCS-1 & DCS-2 OXC architectures for both coherent & incoherent cases. (P_{in} = -16 dbm, P_r = -27 dbm; M=8, V_n = 4, δ =-20 dbm, ε = -25 dbm and ε′= -25 dbm).

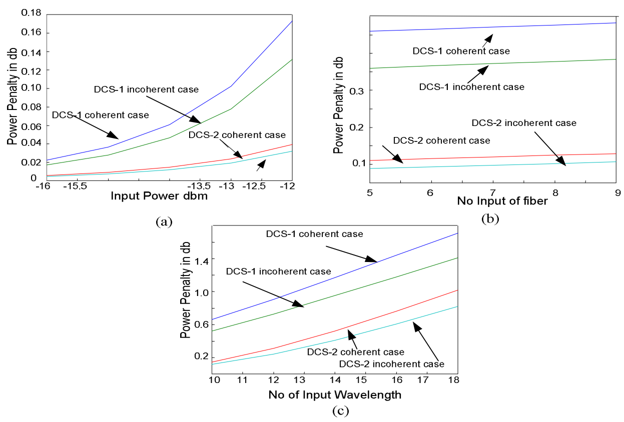


Fig. 7. Comparison of variation of power penalty with (a) input power (b) number of input fiber (c) number of wavelength per fiber for DCS-1 and DCS-2 OXC architectures for both coherent & incoherent cases at BER = 10^{-12}. (P_{in} = -16 dbm, P_r = -27 dbm, N=8, , V_n = 4 i.e V =32, δ =-20 dbm, ε = -25 dbm and ε′= -25 dbm).

Fig. 7 (a) shows the comparison of variation of power penalty with input power, Fig. 7 (b) shows the comparison of variation of power penalty with number of input fiber and Fig. 7 (c) shows the comparison of variation of power penalty with number of input wavelength per fiber with limited number of wavelength converter for the mentioned architecture.

IV. Conclusion

The effect of different system parameters such as input power, number of input wavelength per fiber and number of input fiber on system performance such as crosstalk, BER and power penalty has been investigated. A comparative picture of system performance for those architectures of limited wavelength interchanging cross-connect has been depicted. To find the expression, it has been considered that the signals traverse only one node, signals extinction ratio are infinite, the converted signal is free from the crosstalk carried with it before wavelength conversion and input output characteristic is assumed unity except propagation delay and phase change. We also avoid interferometric intensity noise, relative output noise and amplified spontaneous emission noise. It is seen that both the crosstalk power and BER increase as anyone of the parameter such as input power, number of input wavelength per fiber or number of input fiber is increased. DCS-1 coherent case with wavelength converter suffers from more crosstalk and BER. There is a distinct difference between the amount of crosstalk in coherent cases and incoherent cases. The difference between these two cases indicates the improvement which can be obtained by suppressing the beat term.

References

[1] T. Gyselings, G. Morthier and R. Baets, “Crosstalk analysis of multiwavelength optical cross connects,” *J. Lightwave Technol.*, vol. 17, no. 8, pp.1273–1283, Aug. 1999.

[2] M. S. Islam, S. P. Majumder, N. T. Sim, “Performance limitations of optical cross connect without wavelength converter due to crosstalk,” *J. Opt. Commun*.,vol. 29, pp. 71–75, 2008.

[3] M. S. Islam and S. P. Majumder, “Bit error rate & cross talk performance in optical cross connect with wavelength converter,” *J. Optical Networking*, vol. 6, no. 3, pp. 295–303, March 2007.

[4] K. Lee and V. O. K. Li, “A wavelength-convertible optical network,” *J. Lightwave Technol.*, vol. 11, no. 5/6, pp. 962–970, May/June 1993.

[5] T. Y. chai and T. H. Cheng , “In band crosstalk analysis of optical cross- connect architectures,” *J. Lightwave Technol.*, vol. 23, no. 2, pp. 688–701, Feb. 2005.

[6] S. Yamamoto, T. Yoshimatsu , H. Takara *et al.* “Influence of intrachannel crosstalk with frequency dependence on signal degradation in optical switch network,” *J. Lightwave Technol.*, vol. 27, no. 24, pp. 5716–5722, December 2009.

[7] M. John, D. Downie and A. B. Raffin, “Analysis of signal distortion and crosstalk penalties Induced by optical filters in optical networks,” *J. Lightwave Technol.*, vol. 21, no. 9, pp. 1876–1886, Sept. 2003.

[8] T. Y. Chai, T. H. Cheng, G. shen, S. K. Bose, and C. Lu, “Design and performance of optical cross connect architectures with converter sharing,” *Opt. Net. Mag.* pp.73–83, July/August 2002.

[9] 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.

A Comparative Study of Different Dispersion Compensation Techniques in Long Haul Communication

Md. Talha, Hasan U. Zaman, K. M. A. Salam
Dept. of ECE, North South University, Dhaka, Bangladesh
talhahome@yahoo.com, hasan.zaman@northsouth.edu, kazi.salam@northsouth.edu

***Abstract*—Optical fiber communication (OFC) provides us a very high bit rate data communication. There are different types of impairments and signal degradation mechanisms involved with this high speed communication system. In case of long haul communication, the most effective impairment is dispersion. It affects the signal very badly when signal travels a long distance. Different techniques are available to compensate dispersion. Not all of them are the same in performance. In this work, we study multiple dispersion compensation techniques and their performance measures and compare them with respect to the BER, Q-factor, Eye height, threshold etc. For this purpose, we used OptiSystem13 simulation software. Based on our analyses, we derive a conclusion on which technique is better for high speed long haul OFC networks.**

Keywords— Dispersion, long haul, DCF, FBG, dispersion compensation etc.

I. INTRODUCTION

There are different types of dispersion in OFC [7]. They can be compensated using different techniques. Fiber Brag Grating (FBG) and Dispersion Compensating Fiber (DCF) are two mostly used techniques in long haul communication. We compare the performance analysis between these techniques to find the better compensation technique for long distance optical fiber communication. Some previous study said [1][2], for a 100 km SMF e EDFAs before and after the SMF is enough to compensate the dispersion. But, for a 140 km SMF we need to increase the gain of EDFA. This technique will no longer work for a distance like 288 km SMF. This is because; the attenuation and chromatic dispersion are very high for this type of long distances. For this reason, for a SMF of dispersion parameter (16 ps/nm/km), a DCF of (-80 ps/nm/km) can be used to compensate the dispersion. In Fiber Brag Grating (FBG) compensation, we get a better result than that of no compensation. But using DCF will be more useful for 100 km or 288 km distances. If we look at the Q-factor of the different types of compensation techniques, for DCF at 288 km we get a Q-factor almost same as the Q-factor of FBG compensation at 100 km. That means using DCF we can travel the optical signal 3 times more distance than FBG with the same Q-factor.

II. THEORY

The two different compensation techniques that we considering, control the refractive index of the optical fiber in two different ways. In the FBG compensation technique, the compensating fiber has some grating with different refractive indices; those give different velocities for different modes of light signals. For this reason, delay between two different mode signals compensated and they reach at the receiver at the same time. Depending on the gaps between two consecutive gratings, particular wavelength will be blocked and the rest will go through. The gap which is equal to the wavelength of a particular mode, that gap will block that particular wavelength only.

For a DCF, the core refractive index does not vary only at the core cladding interface; it changes in an irregular fashion to compensate the fiber dispersion. The dispersion parameter is negative and this refractive index pattern changes the dispersion of the signal at reverse direction. If we compare the performance between these methods, we get some decision about which method is better. We can also compare these methods with other compensation techniques [3][4][5].

III. SYSTEM ARCHITECHTURE

In our study, we have worked with two different dispersion compensation techniques. In FBG compensation system shown in Fig. 01, we have used a Brag Grating Fiber with dispersion parameter of -1600ps/nm/km for every 100 km transmission path. SMF of 16 ps/nm/km is used here.

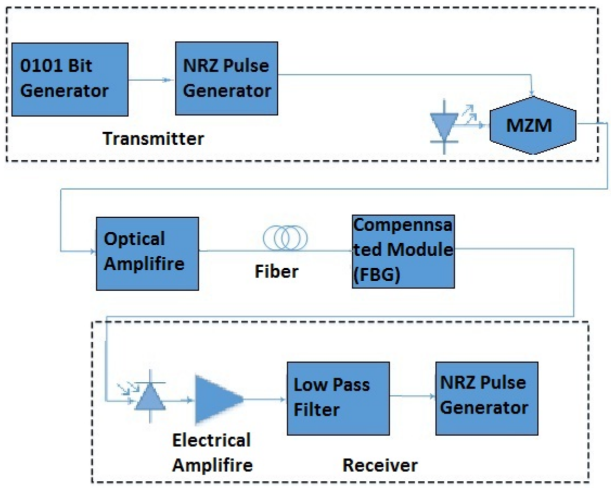


Fig. 1: System Diagram with Compensation Module (FBG).

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

Erbium Doped Fiber Amplifier (EDFA) with gain 20dB and 4 dB noise margin is used. In the DCF system shown in Fig. 2, we have used 20 km DCF of -80 ps/nm/km after every 100 km SMF (dispersion parameter = 16 ps/nm/km). Another EDFA with gain 10 dB and noise margin 4 dB is used after the DCF to compensate the attenuation occurred by the DCF [4].

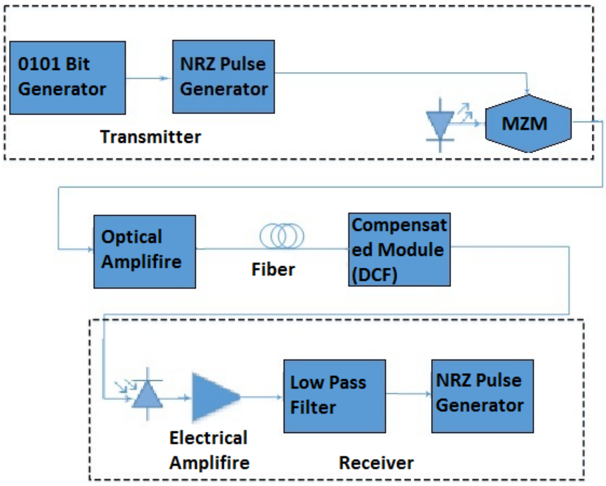
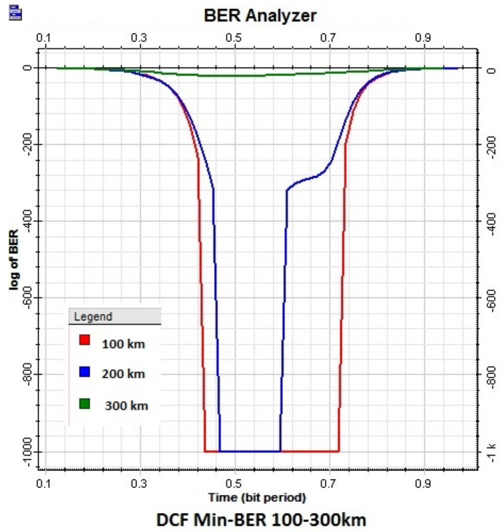


Fig. 2: System Diagram with Compensation Module (DCF).

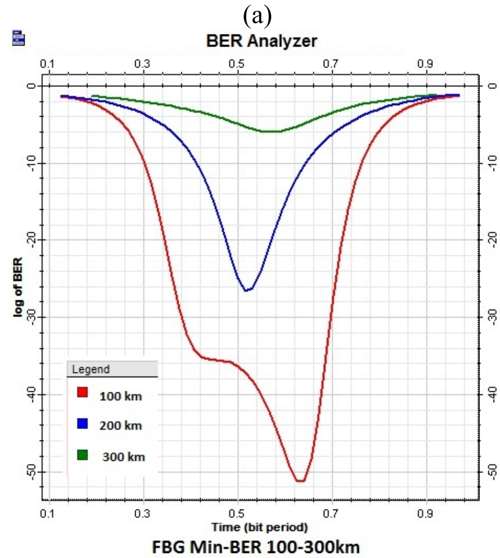
IV. Simulation Results

In case of the FBG compensation, the minimum BER for 100 km SMF is -52.9 when it decreases to -27.45 for a transmission path length of 200 km. If we go further distance with the similar parameters, we get minimum BER of -6 for 300 km SMF. This is shown in Fig. 3b. On the other hand, for DCF based system, we get minimum BER of -1000 for duration from 0.36 to 0.72 for 100 km SMF. If the length increases to 200 km, it still remains the same for duration of 0.38 to 0.54. At SMF length 300 km, we get minimum BER of -23 which is almost the same as the BER of 200 km SMF for FBG compensation. This is shown in Fig. 3a. All the BER values are measured in log of BER. From this BER analysis, we can decide that the DCF technique can provide better performance than the FBG technique in long haul communication.

Another parameter for performance analysis is the Q-factor. In our simulation we measured 3 different values of maximum Q-factor for three SMF lengths. For FBG system, we got maximum Q-factor 15.12, 10.75 and 4.73 for 100, 200 and 300 km, respectively. This is shown in Fig. 4a. For DCF system, maximum measured Q-factor is 177.21, 47.75 and 9.9 for 100, 200 and 300 km, respectively. So, the Q-factor of 300 km using DCF is almost the same as the Q-factor of 200 km for FBG technique. This is shown in Fig. 4b. Thus, we can say that using DCF technique optical signal can travel more distance then that of FBG compensation technique.

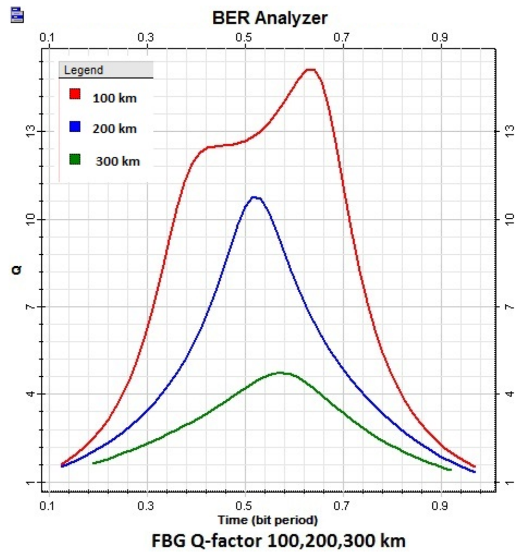


(a)

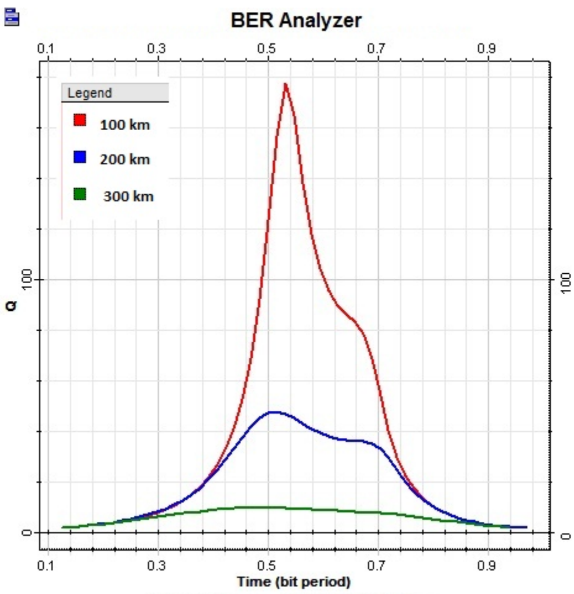


(b)

Fig. 3: Minimum BER of FBG (b) and DCF (a) Compensation for different distances.



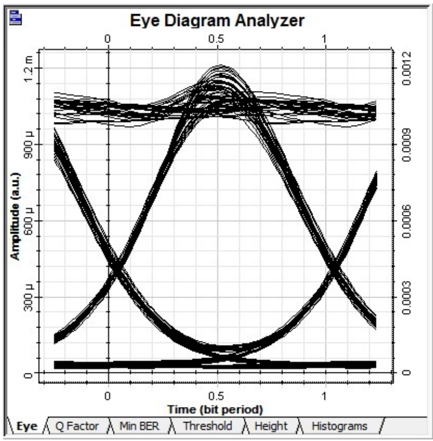
(a)



DCF Q-factor 100,200,300 km

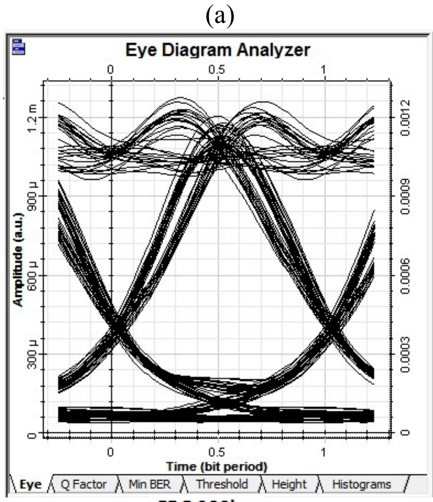
(b)

Fig. 4: Maximum Q-factor of FBG (a) and DCF (b) Compensation for different distances.



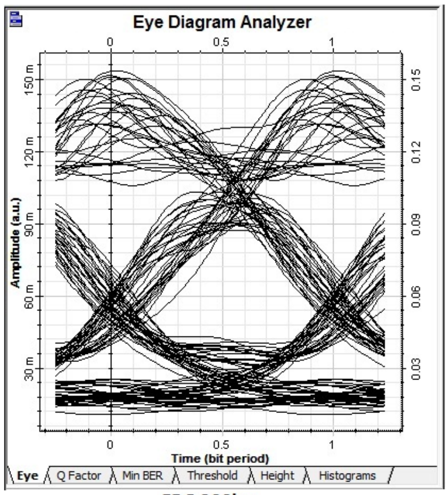
FBG 100km

(a)



FBG 200km

(b)

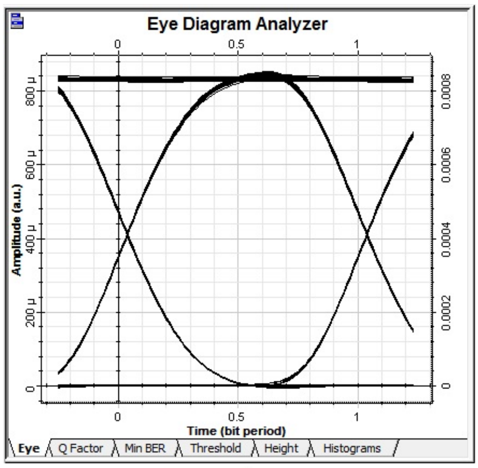


FBG 300km

(c)

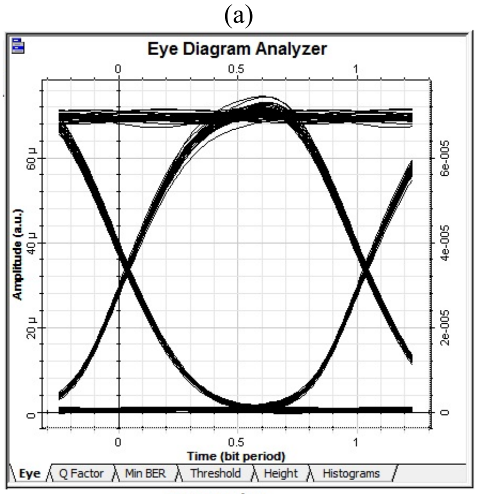
Fig. 5: Eye diagram of FBG compensation for 100km (a), 200km (b) & 300km (c).

Finally, we analyse the Eye diagram for the two mentioned techniques. Eye diagram is the oscilloscope display of a digital signal [1].



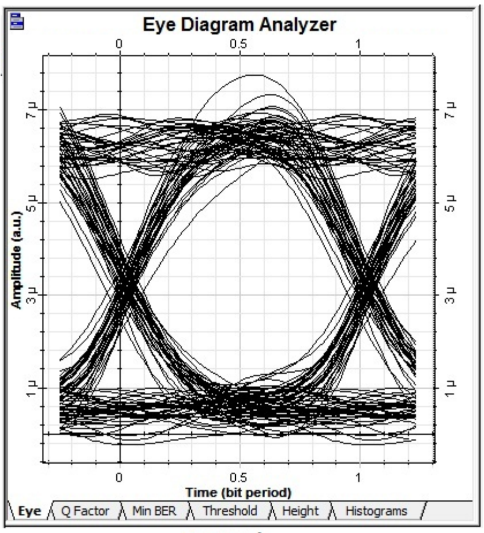
DCF 100 km

(a)



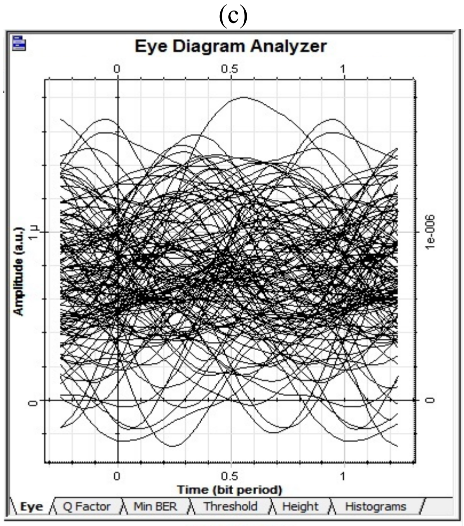
DCF 200 km

(b)



DCF 300km

(c)



DCF 400 km

(d)

Fig. 6: Eye diagram of DCF compensation for 100km (a), 200km (b), 300km (c) & 400km (d).

For the FBG technique, we get acceptable eye shapes for up to 200km. The eye shape is not very good for a distance of 300km. This is shown in Fig. 5.

On the other hand, for DCF compensation we get quite good eye shape for 300 km SMF length. Though it degraded to a poor shape after travelling 400 km. This is shown in Fig. 6. Up to this point we can decide that, DCF technique is a better option than FBG compensation for long distances (e.g. 200km or above).

V. ANALYSIS

After comparing these performance parameters, we can assert that DCF can perform as well as FBG compensation even for a longer distance. Using DCF technique, we can get the Minimum BER and Q-factor for 300 km as those using FBG at 200 km. Even eye height of DCF compensation for 300 km is better than that of FBG at 200 km. So, we can conclude that DCF is a better technique than FBG for long distance high speed optical fiber communication.

VI. CONCLUSION

Different dispersion compensation techniques in Hall communication have been successfully studied. After analysing all the parameters of interest for different compensation techniques using OptiSystem 13.0 simulation software [6], we conclude that the FBG technique is acceptable for up to a certain distance, but after travelling a longer distance, its performance degrades. We cannot make up that degradation by altering any existing component of our system. DCF can be a solution to overcome this difficulty. Even DCF is not sufficiently good after a certain distance. We plan to work on this technique and propose a more effective way of compensating dispersion in long haul OFC.

VII. FUTURE WORK

As we have seen, DCF does not work well for a distance like 400 km or more using our current scheme. We plan to work on finding a more suitable method which can be used for longer distances like 500 km or more. We also plan to work on different DCF techniques and compare them to find a better solution for long distance optical fiber communication.

REFERENCES

[1] Ajeet Singh Verma, A. K. Jaiswal, Mukesh Kumar (2013), "An Improved Methodology for Dispersion Compensation and Synchronization in Optical Fiber Communication Networks." International Journal of Emerging Technology and Advanced Engineering (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 5, May 2013).

[2] Fahad Almasoudi, Khaled Alatawi, Mohammad A. Matin (2013), "Study of OFDM Technique on RoF Passive Optical Network." Optics and Photonics Journal, 2013, 3, 217-224 doi:10.4236/opj.2013.32035 Published Online June 2013.

[3] Anandita joy Agarwal, Mukesh kumar, Rohini Saxena. "Comparison of different techniques of dispersion compensation". International Journal of Electronics and Computer Science Engineering.

[4] Saurabh Kumar, Prof. A. K. Jaiswal, Er. Mukesh Kumar, Er. Rohini Saxena (2013), " Performance Analysis of Dispersion Compensation in Long Haul Optical Fiber with DCF". IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) e-ISSN: 2278-2834,p- ISSN: 2278-8735. Volume 6, Issue 6 (Jul. - Aug. 2013), PP 19-23.

[5] N. K. Kahlon and G. Kaur, "Various Dispersion Compensation Techniques for Optical System: A Survey" Open Journal of Communications and Software Volume 1, Number 1, May 2014.

[6] OptiSystem Component Library, "Optical Communica- tion System Design Software," Optiwave Company, Ot- tawa, 2008.

[7] G. Keiser, "Optical Fiber Communications," 4th Edition, McGraw-Hill Companies, Inc., New York, 2010.

[8] Sandeep Singh, Vaibhav Bhushan Tyagi, Anuj Bhardwaj, Neeraj Gupta (2012), "Performance Analysis of WDM Link Using Different DCF Techniques". International Journal of Advanced Research in Computer Science and Software Engineering. Volume 2, Issue 8, August 2012.

vSIM: The Next Generation Mobile Technology

M. M. Rahman
Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
Email: engr.mizanbd@ymail.com

A. B. M. A. A. Islam
Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
Email: alim_razi@cse.buet.ac.bd

***Abstract*—Global system for mobile communication (GSM) is the most popular mobile network around the world. Subscriber Identity Module (SIM) is must for communicating using GSM network. Recently, virtual SIM technology has been introduced, which runs SIM cards remotely. Both of these technologies, i.e., SIM and virtual SIM, hold many risks for its users and can be attacked in many ways. In this paper we will introduce a new technology named software controlled virtual SIM (vSIM), which will solve these threats and it will provide a secure and reliable communication in GSM network.**

Keywords: - GSM network, SIM card technology, security in mobile communication, virtual SIM.

I. Introduction

GSM was mainly developed for European nations and then being popular. It is now used all over the world. GSM cell phones needs a SIM card to communicate.A SIM card is a portable memory chip. It is inserted in mobile phones and get connected to the GSM network. These cards store different personal information of the account holder like his or her phone number, address book, text messages, and other system data that are needed to manage the GSM system. If a user wants to use new cell phone then he or she usually removes the SIM card from the old one and inserts it into the newer one. [1-3].GSM follows COMPv23 for authentication, which consists of A3 and A8 algorithms and encryption needs A5 algorithm.

The GSM network verifies the identity of the subscriber using a challenge-response mechanism. A 128-bit Random Number (RAND) is sent to the mobile station (MS). The MS generates 32-bit Signed Response (SRES) using the RAND and A3 algorithm using the individual subscriber authentication key (Ki) and send it to the network. The GSM network repeats the calculation and checks whether the SRES value is correct or not. If SRES value is matched, the subscriber is now authenticated. This authentication process repeats after a random period of time. Now SIM generates Kc using this RAND and A8 algorithm using the Ki. This Kc is stored for further encryption using A5 algorithm.

Encrypted voice and data communications between the MS and the network is accomplished by using A5 algorithm with the help of Kc value produced from A8 algorithm. This Kc is used until the next challenge-response mechanism process repeats. [17]

Based on our work we have contributed the followings-

- mentioning limitations and security threats of SIM and virtual SIM.
- proposing a new software controlled GSM communication architecture named vSIM.
- present performance evolution of SIM and virtual SIM.
- analytical model for performance of vSIM.
- performance evaluation of vSIM based on our analytical model.

II. Security Threats of SIM

Karsten Nohl and his research team has developed a system named Osmocon, which stands for Open Source Mobile Communication. SIMtrace is a part of Osmocon, which is able to breach the GSM security and get all data from SIM and decrypt successfully. It has two parts including a Hardware and a Software part.

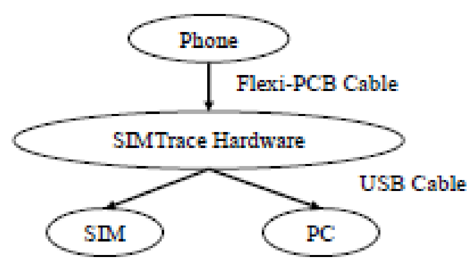


Fig. 1. SIM trace technology [16]

The SIM card is placed into the SIMtrace Hardware and the handset is connected to the SIMtrace Hardware using a cable known as flex cable. The SIMtrace hardware is then connected to a Computer by an USB device from the host and the SIMtrace software will try to find and access this device. The SIMtrace software receives packets from the SIMtrace hardware and can forward them to the IANA assigned GSMTAP port (4729) using the GSMTAP protocol. A modified version of Wireshark is used to analyze the received data form the hardware.

III. Related Work

Some organizations introduced virtual SIM technologies, which is slightly different than our existing mobile system that solves this SIM cloning problem. SIM cards are placed in a

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

Fig. 2. SIM trace technology [16]

server or multiple servers and we access the system using a wireless device attached to our mobile phone. Several existing virtual SIM technologies are implementa, IQ SIM, world SIM, movirtu etc. The key benefits of virtual SIM technology are-

- SIM cards can be replaced in a handset just by software commands, there is no need to change it physically.
- There is no need to have SIM Cards and Mobile Terminals in same location.
- A user can easily use multiple SIM cards in parallel.
- It potentially removes the risk of theft and managing SIM is easier.
- Faster access.
- SIM cards can be allocated based on its location. [8-11]

The main limitations of these technologies are-

- if the server is hacked then all communication is fallen down and all are affected at a time.
- Sniffing is also possible in this network. Therefore, it is possible to monitor other users.

Virtual SIM technology is more secure than common GSM technology, however, it takes more time to initiate a call or sending data while communication. Therefore it is needed to introduce a system, which will provide advantage of virtual SIM with less and effective communication time.

IV. Proposed Architecture

Our proposal is to run SIM card virtually in smart phone. There will be a software in our mobile phone that holds all information that is normally stored in a SIM card. All relevant algorithms i.e., A3, A8 and A5, will also be available on that software. From the turning on of a mobile handset this software will be run in background and communicate with mobile network. This software is use a dedicated portion of our RAM and processor of smart phone. Therefore, a user will never feel the latency or time delay while communicating with others. We are considering the processor frequency of mobile phone is minimum 1GHz that is lowest in smart phones.

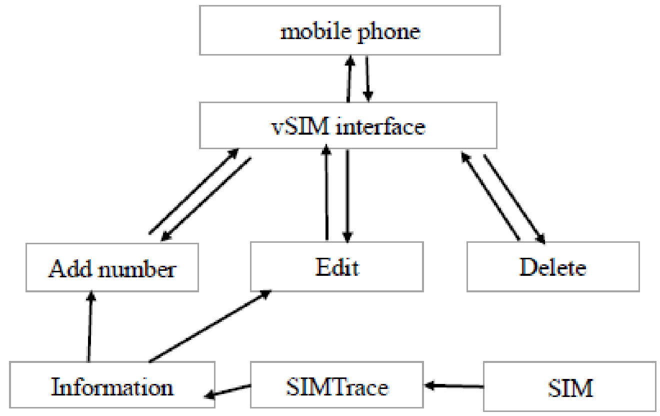


Fig. 3. vSIM technology

As the SIM card information is not publicly opened, we will use SIMtrace technology to read those data and there will be an interface to add a SIM card information and will be stored in our software database. A user will be able switching his or her subscriber number that is already been added there. Therefore, a single user can easily use multiple numbers without changing any SIM cards.

The advantages of this vSIM are-

- The total cost of 6 billion US dollar to produce SIM cards will be reduced to zero.
- No server and wireless device is needed that was needed in virtual SIM technology.
- It will provide faster communication than other technologies i.e., SIM card and virtual SIM, described in section III.
- User will be introduced to a better communication system.

V. Analytical Performance Analysis of vSIM

Time required to establish a communication in one end of GSM network is-

$$T = \frac{I}{fsim} + Ln \quad (1)$$

Where, I= Total number of instruction in COMP v23 or A5 algorithm, fsim= Frequency of SIM, Lsim = Latency of SIM and Ln = Latency of GSM network. In 3G network, fsim=900MHz, Lsim=100ms, LN=250ms and I=3609.

Therefore, time required to send data in one half is 350.004ms.

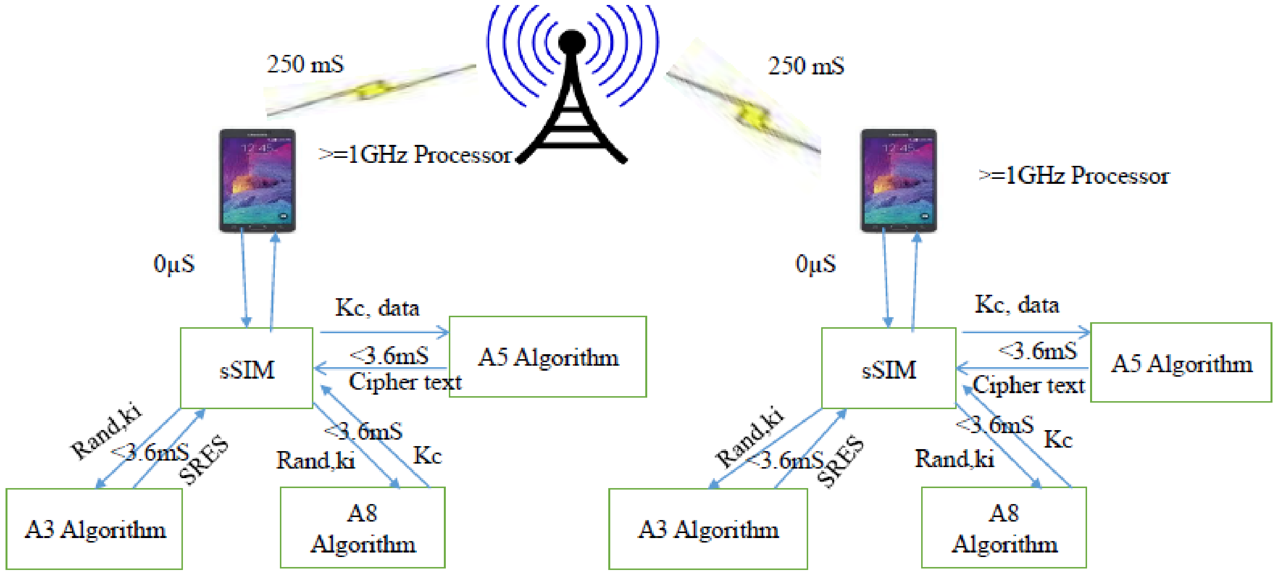


Fig. 4. vSIM technology

On the other hand, time required for virtual SIM technology is-

$$T = \frac{I}{fsim} + Lm + Ln + t \tag{2}$$

Where,Lm = Latency of Mobile Phone, t= Time required to send data from mobile to server. In 3G network, fsim=900MHz, Lsim=100ms, Lm=20ms, LN=250ms, I=3609 and t=10ms.

Therefore, time required to send data in one half is 380.004ms, which is greater than common GSM technology. As mobile is connected to server through wireless connection, the value of t will be increased with respect to the position of mobile phone.

If we replace our SIM technology using vSIM then this communication time will be-

$$T = \frac{I}{fm} + Ln \tag{3}$$

Currently the minimum mobile frequency fm=1GHz, Lm = 20ms, LN = 250ms, I=3609. Therefore, time required to send data in one half is 270.003ms. This means our technology saves communication and serves faster mobile communication. Therefore, vSIM saves at least 160.0074ms and 220.0074ms for SIM and virtual SIM technology respectively in a single byte communication.

VI. Numerical Simulation

We simulated vSIM efficiency in different way. These are-

A. Simulation settings

We considered the following situations-

- what will be the possible time to run A3/A8/A5 algorithm, if the number of instruction increases in ideal situation. It means we introduce more complex algorithms for same device with no time delay and simulated the result.
- our devices are being introduced with new and advanced technologies. So, we simulated considering that A3/A5/A8 algorithm is same however, the processor frequency is variable.
- Third and the last was based the latency of devices. With the advancement of technology different mobile devices will provide different time delay for accessing data. We considered that new technologies will introduce low latency and simulated the effect of latency time on communication where device frequency and algorithms are same.

B. Simulation Results

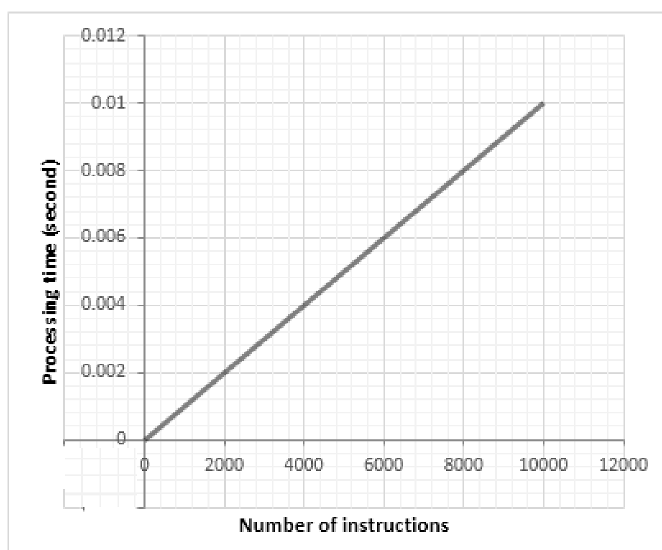


Fig. 5. Instruction processing time for 1GHz processor

Fig. 5 demonstrates the first condition. It indicates that the processing time is totally depended on the number of instruction to be processed if our processor frequency is fixed.

For second condition, the instruction processing time varies with the frequency of the processor and proves that the advancement of technology this time will be reduced to very low time.[Fig. 6]

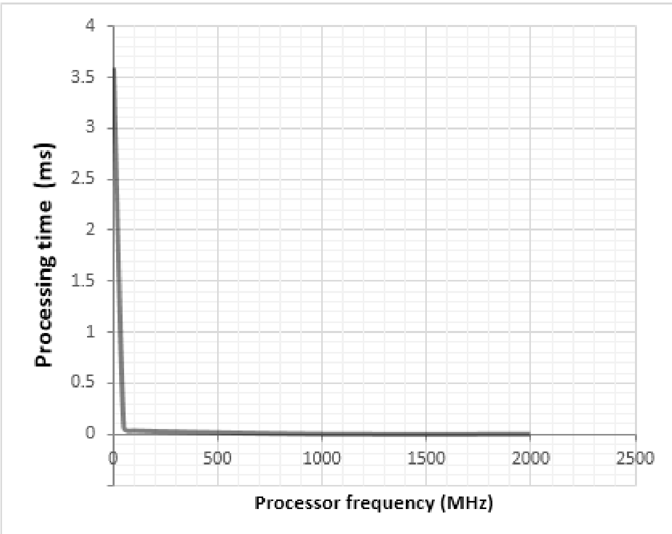


Fig. 6. Present instruction processing time for different processor

Instruction processing time has a great dependency on the latency of mobile phone. So, the communication will be faster if we can reduce this latency time. Fig. 7 demonstrates the third condition.

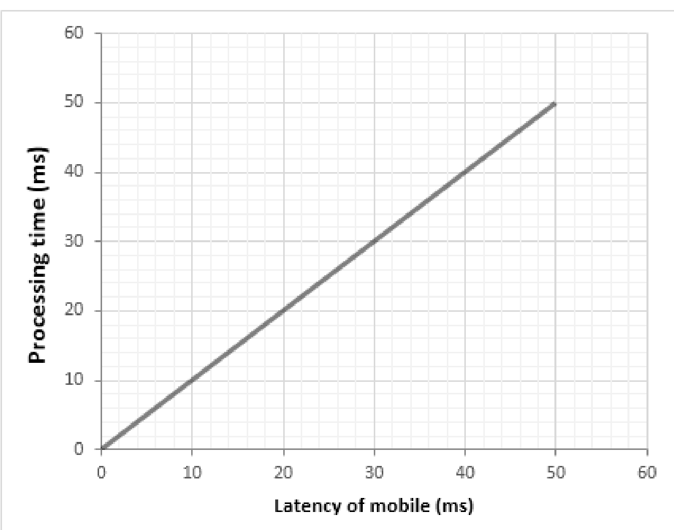


Fig. 7. Processing time with different latency

VII. Future Work

We have simulated our logical necessity of vSIM. Our aim is to modify android OS and integrate our code there so that we can run GSM numbers in our android smart phone without SIM card. As A3/A5/A8 algorithms can be modified or updated, the automatic update will also be available of these algorithms. As information of GSM SIMs can be collected, we will use those subscriber identification number virtually.

VIII. Conclusion

Currently, more that 6 billion US dollar is needed per year to produce SIM cards. Again, virtual SIM needs both SIM cards and server. Therefore, cost is increased in virtual SIM technology. On the other hand, more time is needed for communicating with others. Hence, security is also increased in virtual SIM technology. Our proposed vSIM technology will need half time of those technologies. This will be faster, convenient and reliable communication system.

As this is totally a new work and we hope we will be able to introduce a new vSIM technology. Again, we are working as intruder for Android OS, our aim to find out the bugs of GSM and solving those limitations. Our proposed architecture is fully new and different than other existing technologies.

References

[1] Rehab El Nemr, Imane Aly Saroit Ismail, S. H. Ahmed: Action- Triggered Public-Key cryptography for GSM systems with Phone- Dependent end-to-end encryption, Vol (5), Issue (2), June 2006.

[2] http://www.cbc.ca/news/technology/cellphone-sim-cards-contain-significant-security-flaw-1.1351810

[3] http://www.infoworld.com/article/2611169/mobile-security/sim-cards-vulnerable-to-hacking–says-researcher.html

[4] Jeremy Quirke, ”Security in the GSM system”, May 2004, www.it.iitb.ac.in/ kavita/GSM-Security-Papers/Security in the GSM system 01052004.pdf

[5] Movirtu virtual SIM http://www.movirtu.com/

[6] World SIM https://www.worldsim.com/virtual-phone-number/

[7] IqSIM http://www.iqsim.com/VirtualSIM.en.htm

[8] Implementa virtual SIM http://www.implementa.com/solutions/virtual-sim-platform

[9] Mohsen Toorani, Ali Asghar Beheshti Shirazi: Solutions to the GSM Security Weaknesses, Issue Date: 16-19 Sept. 2008, pp. 576-581, Date of Current Version: 20 January 2009.

[10] H. Welte, Running your own GSM stack on a phone, 27th ChaosCommunication Congress, Berlin, 2010.

[11] K. Nohl and S. Munaut, Wideband GSM Sniffing, 27th Chaos Communication Congress, Berlin, 2010

[12] Sameer Hasan Al-Bakri, M.L. Mat Kiah, A.A. Zaidan, B.B. Zaidan, Gazi Mahabubul Alam: Securing peer-to-peer mobile communications using public key cryptography: New security strategy, Vol.6(4), pp. 932-938, 18 February 2011.

[13] M. Toorani, A.A. Beheshti Shirazi, ”Solutions to the GSM Security Weaknesses”, The Second International Conference on Next Generation Mobile Applications, Services and Technologies 2008, pp. 576-581, 16-19 Sept. 2008.

[14] V. Bocan and V. Cretu, ”Mitigating Denial of Service Threats in GSM Networks”, 1st IEEE International Conference on Availability, Reliability and Security (ARES’06), April 2006.

[15] Osmocon Project, http://bb.osmocom.org/trac/wiki/SIMtrace

[16] D. Spaar and H. Welte, Running your own GSM network, 25th Chaos Communication Congress, Berlin, 2008.

[17] K. Nohl, Attacking phone privacy, in BlackHat 2010 Lecture Notes, 2013.

[18] J. Heider and R. El Khayari, ios keychain weakness faq - further information on ios password protection, Fraunhofer Institute for Secure Information Technology, Tech. Rep., 2012.

Smart Vehicle Accident Detection and Alarming System Using a Smartphone

Adnan Bin Faiz
Dept. of Computer Science and Engg.
Chittagong University of Engg. & Tech.
CUET, Chittagong-4349, Bangladesh
asifliveadnan@live.com

Ahmed Imteaj
Dept. of Computer Science and Engg.
Chittagong University of Engg. & Tech.
CUET, Chittagong-4349, Bangladesh
imtu_1992@yahoo.com

Mahfuzulhoq Chowdhury
Dept. of Computer Science and Engg.
Chittagong University of Engg. & Tech.
CUET, Chittagong-4349, Bangladesh
mahfuz_csecuet@yahoo.com

Abstract— **Vehicle accident is the paramount thread for the people's life which causes a serious wound or even dead. The automotive companies have made lots of progress in alleviating this thread, but still the probability of detrimental effect due to an accident is not reduced. Infringement of spFieed is one of the elementary reasons for a vehicle accident. Therewithal, external pressure and change of tilt angle with road surface blameworthy for this mishap. As soon as the emergency service could divulge about an accident, the more the effect would be mitigated. For this purpose, we developed an Android based application that detects an accidental situation and sends emergency alert message to the nearest police station and health care center. This application is integrated with an external pressure sensor to extract the outward force of the vehicle body. It measures speed and change of tilt angle with GPS and accelerometer sensors respectively on Android phone. By checking conditions, this application also capable of reducing the rate of false alarm.**

Keywords— Vehicle accident detection; emergency alert message; pressure sensor; GPS; accelerometer.

I. INTRODUCTION

The automotive industry around the world has shown a tremendous enhancement in its production over the recent years. Millions of vehicles are being produced annually. But along with these, the accident rates are also getting significantly increased. As a result, even the optimistic nature of people has become worried while going outside. United States Department of Transportation data for 2005 from the Fatality Analysis, Reporting System show that for passenger cars, 18.62 fatal crashes occur per 100,000 registered vehicles. In 2009, 33,808 people died in vehicle traffic crashes only in USA [1].

Most of the accidents occur due to human negligence, such as reckless driving, lack of good infrastructure, etc. An immediate rescue process after an accident can be considered as a tightrope walk between life and death. Any fractional time delay of arriving medical help can cost the life of the victims. A study by Virtanen et al. shows that 4.6% of the fatalities in accidents could have been prevented only in Finland if the emergency services could be provided at the place of the accident at the proper time [2]. As such, an efficient automatic accident detection with an automatic notification to the emergency service with the accident location is a prime need to save the precious human life.

As smartphones become such an important part of our life, it is feasible to use smartphones in a post-accident fatality prevention system. Our application uses the GPS receiver in phone to detect the rapid change of deceleration that occurred at accident time. It also takes the change of pressure from the pressure sensor and the change of tilt from an accelerometer sensor in a Smartphone. By detecting these three conditions as accident detection, this android app send the accident location for emergency help. An emergency switch option also added to this app which provides a chance to driver for sending alert message without checking accident detection condition. Therefore, our contribution can be listed as follows:

- Construction of an efficient automated vehicle accident detection system using Android.
- Develop a framework for reducing false alarm of vehicle accident detection.
- Dispatch automatic emergency accident alert message to relative, nearest police station and hospital.

We have discussed on some relevant papers in section II. In section III we have chronicled the system functionality with technical details. Later we have shown the implementation result with experimental data, functional and subjective evaluation of our application. In the end, we discussed about the challenges and future extension of our work.

II. RELATED WORKS

Mobile devices, especially Smartphones have been deployed as floating traffic probes and sensors in many applications, both academically and commercially [3]. These applications include road conditions survey, traffic conditions monitoring [4] and accident detections [5]. All of these abilities are essential to an Intelligent Transport System (ITS), which aims to reduce traffic congestion and enhance traffic safety [6]. Real-time traffic accident prediction focuses on the change of traffic conditions before an accident occurrence, while traffic incident detection studies are concerned with the change of traffic conditions after an incident occurrence [7]. However, the performance of these detection and prediction system is greatly restricted by the number of monitoring sensor, available fund, the algorithms used to confirm an accident, weather, traffic flow etc.An acoustic accident detection method is proposed by D. A. Whitney and J. J. Pisano [8]. There are possibilities of false alarm in the system and also does not guarantee the occurrence of an accident. S. Amin et al. proposed a method of accident detection and reporting system using GPS, GPRS and GSM Technology[9].

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

C. Saiprasert et al. propounded a reporting system in case of dangerous driving using android [10]. Jung Lee described an accident detection system on a highway by tracking the vehicle [11]. We have used pressure sensor, GPS and accelerometer to detect a real alarm and dispatch alert message to nearby police station and health care center.

III. SYSTEM DESCRIPTION

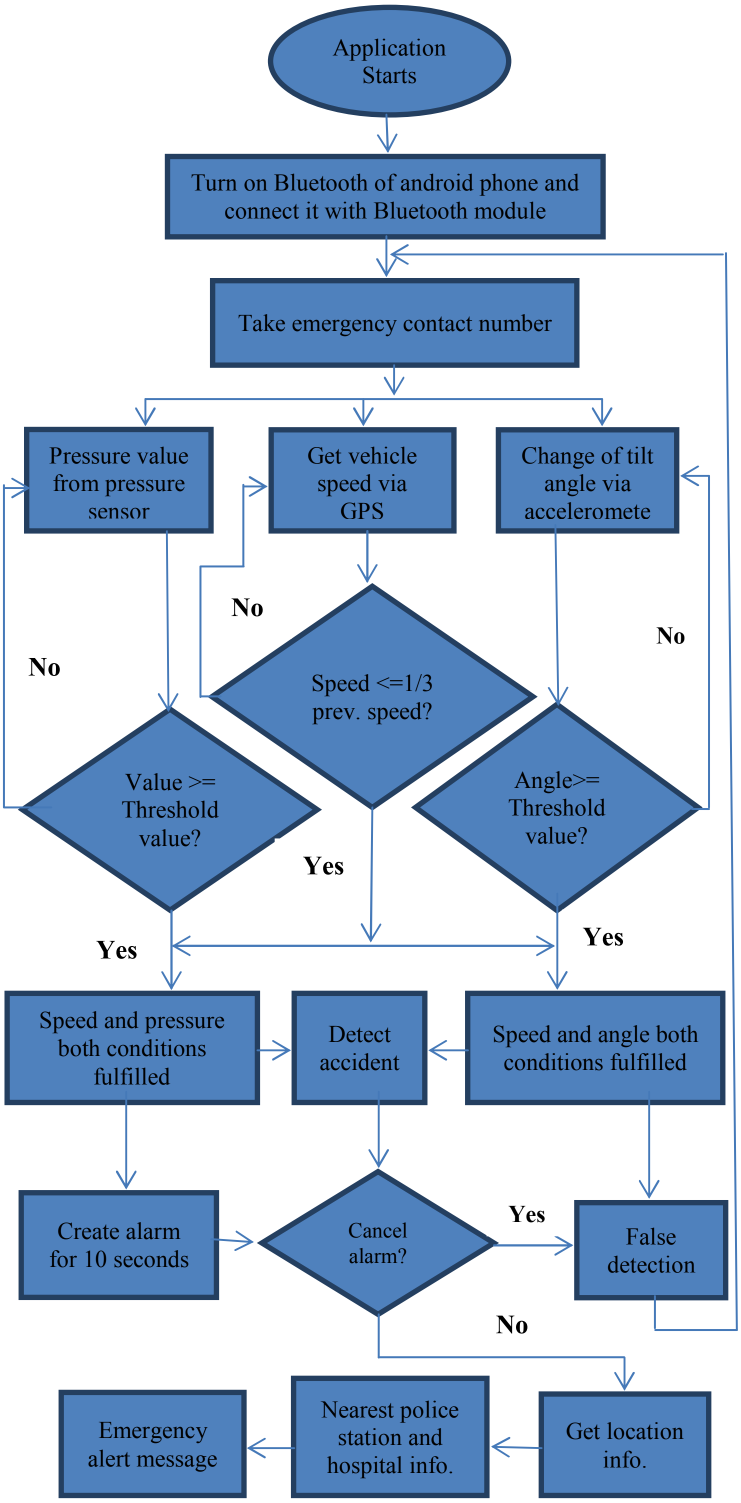


Fig. 1. System architecture

A. *Accident Detection*

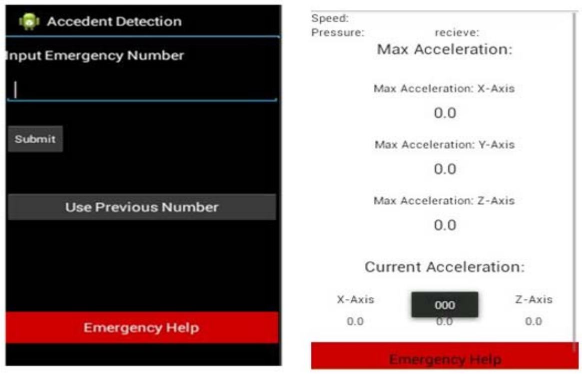
We have considered three parameters to detect a situation as an accident. As we know, when an accident occurs velocity of vehicle decrease rapidly. But in this application, we do not detect the situation as an accident, only when the speed decreases rapidly. If pressure rate or change of tilt angle exceed the defined threshold value and speed of vehicle fault equal or below one-third, only then application detects the situation as an accident.

Fig. 1 exhibited the whole system architecture.

IV. SYSTEM IMPLEMENTATION

A. *Bluetooth Connection, Taking Emergency Contact Number and Measurement of Parameters*

The Bluetooth connection of android phone needs to turn on for receiving pressure sensor data. After that, the application takes an emergency number to send message immediately after detecting accident which is displayed in fig. 2(a). The measured value of parameters which we consider to detect an accident shown in fig. 2(b). Initially, all flag bits are zero, so the flag number shows ‘000’.

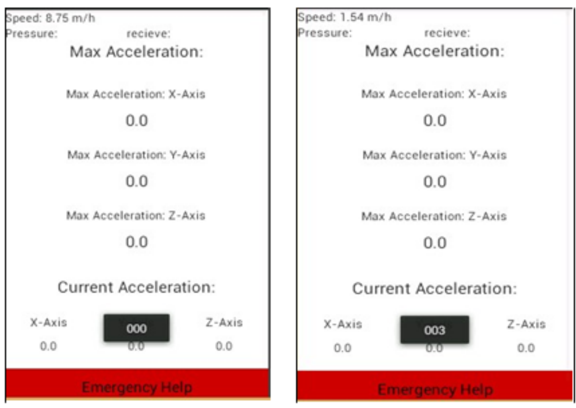


(a) (b)

Fig. 2.(a) Taking emergency contact number (b) Measurement layout

B. *Speed Measurement and Pressure Value*

The application retrieves the speed value at the screen with the help of GPS. Here, speed shows at mile per hour unit. In this stage, the third flag bit that represents speed stage remain zero. So, the flag number shows ‘000’ which is displayed in fig. 3(a).

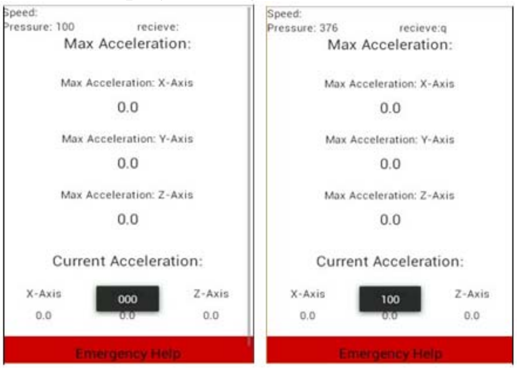


(a) (b)

Fig. 3.(a) Speed measurement (b) Speed exceed threshold

When the speed of the vehicle decreases rapidly at a few time distance, then the third flag bit changed to ‘3’. It actually indicates when the current speed is less than or equal to one-third of previous speed. So, ultimately the flag number turns into ‘003’ which is also given in fig. 3(b).

When pressure sensor receives the external force that exceeds the threshold value that we defined 350, then application changed the pressure flag bit from zero to ‘1’. First flag bit changed to ‘1’. So, ultimately the flag number turns into ‘100’ displayed in fig. 4(b).



(a) (b)

Fig. 4. (a) Pressure measurement (b) Pressure exceeds threshold

C. *Change of Tilt Angle*

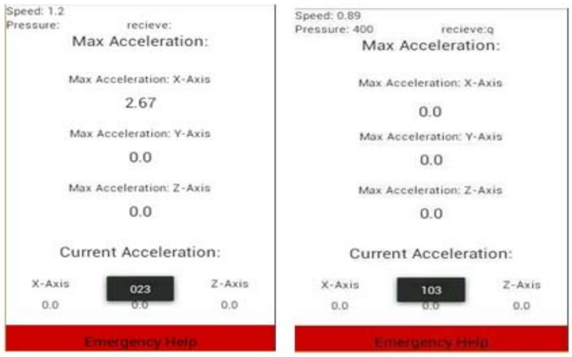
We measured the tilt angle of the vehicle with the road surface by the accelerometer sensor in android phone. When a situation happens like an accident, vehicle makes a bigger tilt angle than before with the road surface. Then a change of tilt angle becomes bigger than before. We consider here the change of tilt angle greater than or equal to 2 of x-axis than previous to detect an accident. Second flag bit changed to ‘2’. So, the flag number turns into ‘020’ showed in fig. 5(b).

(a) (b)

Fig. 5.(a) Change of tilt angle measurement (b) Change of tilt angle exceeds threshold

D. *Accident Detection*

In case 1, if the speed of vehicle dropped down rapidly and a change of tilt angle with road surface increases much, then it detects as an accident. Actually, it means when the threshold values of speed and change of tilt angle of vehicle exceed, then our application considers the situation as an accident. And, this time two flag bits change and flag number turn into ‘023’ which is given in fig. 6(a). In case 2, when the threshold values of speed and pressure of vehicle exceed, it considers the situation as an accident. This time two flag bits change and flag number turn into ‘103’. Fig. 6(b) demonstrated this

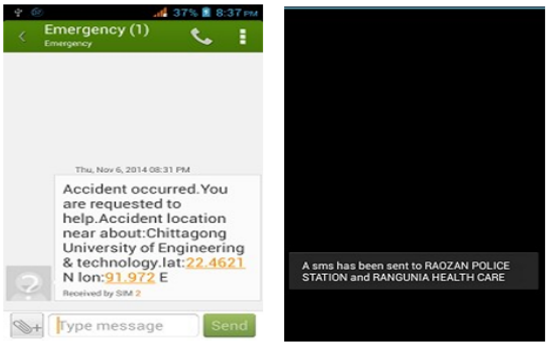


(a) (b)

Fig. 6.(a) Accident detection-case 1 (b) Accident detection-case 2

E. *Alarm and Emergency Alert Message*

When any one of our proposed accident detection method happen, then an alarm arises. User can press ‘Cancel’ button and no emergency alert message will send. This kind of situation is considered as a false alarm. Otherwise, an emergency alert message will send to user’s emergency contact number, nearest police station and hospital.



(a) (b)

Fig. 7.(a) Emergency alert message (b) Notification of sending alert message

V. EXPERIMENTAL RESULTS AND EVALUATION

A. *Experimental Data*

1) Case 1: In this case, Here, we checked a situation considering the speed and tilt angle of a running motorcycle. We changed tilt angle value by rotating sometimes. We also checked speed value by hard break, turning a curve road and slope road. The experimental results of this case are shown in table 1.

TABLE I. EXPERIMENTAL RESULT CASE-1

Experiment no.	**Speed**		**Change of tilt angle (X-axis)**	**Alarm**
	Initial Speed (m/h)	***Changed speed (m/h)***		
1	9.3	3.2	1.67	Off
2	7.9	3.1	1.98	Off
3	8.5	2.4	2.67	On
4	8.3	2.6	2.43	On
5	7.8	2.5	1.74	Off

2) Case 2: In this case, we checked our application by changing the pressure value manually by physical force. We also checked speed value by hard break, turning a curve road and slope road. The details of this case are shown in table 2.

TABLE II. EXPERIMENTAL RESULT CASE-2

Experiment no.	**Speed**		**Pressure value**	**Alarm**
	Initial Speed (m/h)	***Changed speed (m/h)***		
1	8.3	3.2	89	Off
2	8.9	3.1	74	Off
3	8.5	2.5	108	Off
4	9.3	2.8	445	On
5	7.8	2.9	65	Off

B. Functional Evaluation

For this experiment, we take input on a range 0 to 1. And we defined the condition of false alarm up to 0.2. That means, if the input is greater than 0.2 than it indicates accident detection else it was a false alarm. Fig. 8 and 9 displayed these two results.

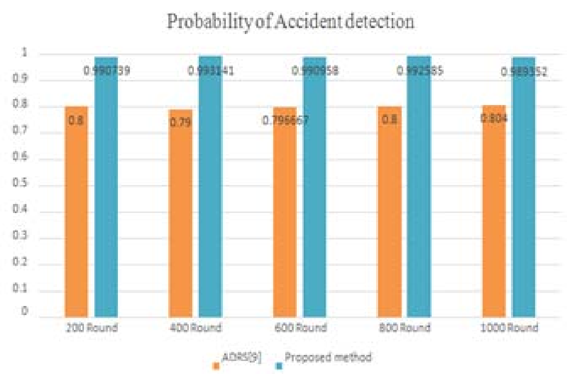


Fig. 8. Comparison of accident detection probability

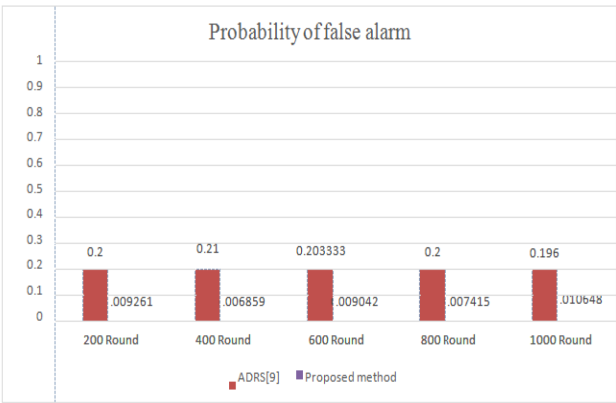


Fig. 9. Comparison of false alarm probability

C. Subjective Evaluation

The evaluation results statistics based on the public opinion are presented graphically in fig. 10.

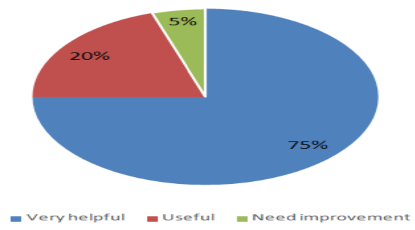


Fig. 10. Review result

VI. CONCLUSION

In this paper, we have shown that road accident can be detected efficiently by using some particular parameters. Our proposed approach capable of deciding whether a situation is an accident or not and if so, then immediately traces nearest police station as well as hospital and send emergency alert message for help. Besides, we have demonstrated the reduction of false alarm in a greater extent compared to other previous works. Though the system requires a continuous Internet connection, but this it is very much cost effective and can be applied significantly in the practical world. In the future, we have a plan to consider more parameters for detecting accident and developing the application in a single device. Hence, the application would play a crucial role in post-accident services and could mitigate the effect due to accident remarkably.

REFERENCES

[1] Highlights of 2009 Motor Vehicle crashes, Traffic Safety Facts, Research Notes, NHTSA (National Highway traffic Safety Administration) [Online], Accessed on 16 October 2011.

[2] N. Virtanen, A. Schirokoff and J. Luom, "Impacts of an automatic emergency call system on accident consequences," 18th ICTCT, Workshop Transport telemetric and safety, 2005.

[3] S. M. Tang and H. J. Gao, "Traffic-incident detection-algorithm based on nonparametric regression," IEEE Transactions on Intelligent Transportation Systems, 2005.

[4] G. Rose, "Mobile Phones as Traffic Probes: Practices, Prospects and Issues," Transport Reviews, 2006.

[5] P.Mohan,V.N.Padmanabhan,and R.Ramjee, "Nericell:richmonitoring of road and traffic conditions using mobile smartphones," in Proceedings of the 6th ACM conference on Embedded network sensor systems, 2008.

[6] D. A. Whitney and J. J. Pisano TASC, Inc., Reading, Massachusetts, "Auto Alert: Automated Acoustic Detection of Incidents," IDEA project, 15 October 2011.

[7] C. Thompson, J. White, B. Dougherty, A. Albright and D.C. Schmidt, "Using Smartphones to Detect Car Accidents and Provide Situational Awareness to Emergency Responders," in Proceedings of 3rd Mobile Wireless Middleware, Operating Systems, and Applications Conference, 2010.

[8] J. Yoon, B. Noble and M. Liu, "Surface street traffic estimation," in Proceedings of 5th International Conference on Mobile Systems, Applications,and Services, 2007.

[9] Md. Syedul Amin, Jubayer Jalil, M. B. I. Reaz, "Accident Detection and Reporting System using GPS, GPRS and GSM Technology," in Proceedings of International Conference on Informatics, Electronics & Vision, 2012.

[10] Chalermpol Saiprasert and Wasan Pattara-Atikom, "Smartphone Enabled Dangerous Driving Report System," in Proceedings of 46th Hawaii International Conference on System Sciences, 2013.

[11] In Jung Lee, "An Accident Detection System on Highway Using Vehicle Tracking Trace," ICT Convergence, 2011.

Developing a Framework for Analysing Web Data to Generate Recommendation

Tanjila Aftab Khanam
Department of CSE,
CUET, Bangladesh
Email: ta_08_32@yahoo.com

Mohammad Shamsul Arefin
Department of CSE,
CUET, Bangladesh
Email: sarefin_406@yahoo.com

Md. Shahedul Islam
ICT Center,
BSMRMU, Bangladesh
Email: shahed95@yahoo.com

***Abstract*—From the advent of the Internet, data contained in the web is increasing exponentially and the number of e-commerce companies is also increasing significantly. Nowadays, we can find that many e-commerce companies selling same products and / or services at different prices. For a customer it is difficult to find an e-commerce company which will be best suited for him. Moreover, it is time consuming and tedious to search for a product in different online sites. For reducing this difficulty, in this paper, we develop a system that can extract web data from different e-commerce sites even though the language of the websites are different. We then analyse the extracted data and recommend best products / services from these sites to the users. For the experimental evaluation of our system, we consider two different languages: English and Bangla and extract books data from different online book stores considering these two languages and recommend books to the users. From the experimental results, we can find that our system can help the users in selection of their products easily and efficiently. Keywords: Web data analysis, data processing, e-commerce**

***Keywords*—Web data analysis, data processing, e-commerce.**

I. INTRODUCTION

Due to the rapid growth of Internet technologies, there are many well known e-commerce companies such as Amazon[1], eBay [2], Olex [3], Agoda [4], TripAdvisor [5], Expedia [6], Rakuten [7], Elance [8] those sell their products and services via Internet. Besides the e-commerce companies, agencies, government organizations, non government organizations provide different types of services via Internet. We can find that these e-commerce companies sell same type of products and services nowadays. However, the quality and price of products and services varies from company to company. As for example, if we want to buy a particular book of an author, we can find that different book selling e-commerce company offer different price for the same book. The same situation arises while booking a hotel as different hotel booking companies offer different prices for the same room type at the same hotel.

There is also the situation where a buyer is looking for a best quality product / service. For example, a student is looking for a Discrete Mathematics book that is most suitable for him whatever the price is. Different e-commerce companies sell different authors' books with the similar titles that makes the selection task of the buyer difficult.

In above type of situations, a buyer needs to search different e-commerce sites for selecting the products / services that wastes time and efforts of buyers. Considering these facts, in this paper, we develop a framework that analyse the data of different e-commerce companies and recommends desired products / services to the buyer that saves his efforts and time. More specifically, the contribution of this paper can be summarized as follows:

- We have extracted data automatically from several websites of two different languages: English and Bangla.
- We then analyse the extracted data to recommend best products / services to the users.
- We have evaluated the performance of our framework using both subjective and objective measures.

The remainder of this paper is organized as follows. Section II provides a brief review of related work. In section III, we detail the computation framework of our proposed approach. Section IV presents the experimental results. Finally, we conclude and sketch future research directions in Section V.

II. RELATED WORK

There have been numerous number of extraction programs discovered in the previous years. Some of them are automated and some are manual. The most straightforward approach is wrapper generation. Moreover, some machine learning algorithms and pattern discovery are also used. They are better than wrapper generation processes but are very difficult to implement. Furthermore, an extensive degree of training is required for accurate information extraction by them.

Hammer et al. [9] developed a configurable tool for extracting semistructured data from a set of HTML pages. It can also convert the extracted information into database objects. Their tool at first takes input from the user. It then converts these data using Object Data Model (OEM) as OEM is well suited for representing semistructured data. In their system, it is necessary to update the specification of the website by the user if there are changes in the website. It has an additional capability of allowing users to specify one or more possible patterns. This is useful for sources that has dynamic structure. Shaker et al. [10] also proposed a framework for extracting and classifying semi-structured web data. They have extracted Nokia products from several websites and converted it into text files. Then, duplicates or less important files were removed and classified all important files.

Knoblock et al. [11] proposed a framework for accurately and reliably extracting data from the web. They developed an

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

algorithm based on machine learning that learns extraction rules based on examples, labelled by the user. This algorithm incrementally builds extraction rules from the examples. Soper [12] proposed a framework for automated web business intelligence system which is based on content acquisition and knowledge creation. In content acquisition phase, they extracted relevant information from web using data agent and in knowledge creation they used neural network, pattern discovery, regression for decision making. However, it is not a fully automated system as none of the current ranking algorithms can be fully trusted to accurately determine the contextual relevance of a web source. The tool [13] developed by Baumgartner et al. is used to combine internal and external data in which the user can select each of the items he wants to extract by clicking on that item. If the structure of the source changes after specification, the wrapper can still extract data from the source correctly. This is because the system uses intelligent conditions called logical patterns.

Bhawsar et al. build a system [14] where they have given input of web server log files and extracted information of web sessions performed by user. A modelling based approach was proposed by Lerman et al. [15]. The system extracts lists and tables from web pages automatically. In their approach, at first they split each web page into tokens. If two or more example pages are given, they pick the smallest page as seed. Then, they build a sequence of tokens by adding tokens to it. But the condition is that, this tokens should appear exactly once on each page. If any of the list contains more than two rows, the tags specifying the structure of the list will not be added to the token list. Because they appear more than once on that page. Then they extract data between the tags that appears more than once on each page.

Expedia [6] is the largest online travel site which is famous for flight and hotel bookings. It books online tickets, reserve hotels and other vacation packages. The user will give input of his departure and arrival place, departing date etc. Then the site will show information about airways such as price from lowest to highest order, available seat, time needed for travelling etc.

In this paper, we developed a framework for extracting and analysing web data for identifying best products from competitors websites considering that the contents can be either of the two languages: English or Bangla. We then recommend such products to the users.

III. System Design and Architecture

The flow graph of the proposed system is given in Fig. 1. At first, the system gets input website information from a file. Then it connects to particular website using URL. Then, it extracts data from websites. The system eliminates unnecessary tag and data from the extracted data. This unnecessary tags are HTML or CSS tags. Then it stores corresponding data into files.

At the second phase, it reads stored pre-processed data from files. Then, it performs the task of secondary data elimination which is noise removal from saved data. This noises are different types of words, page numbers, IP addresses etc. Then, the system finds page template which is finding similarity between web pages. After finding page template, it stores this template data. Then it compares all the web pages with this template data and finds differences. These founded differences are the data that are extracted and saved into the database.

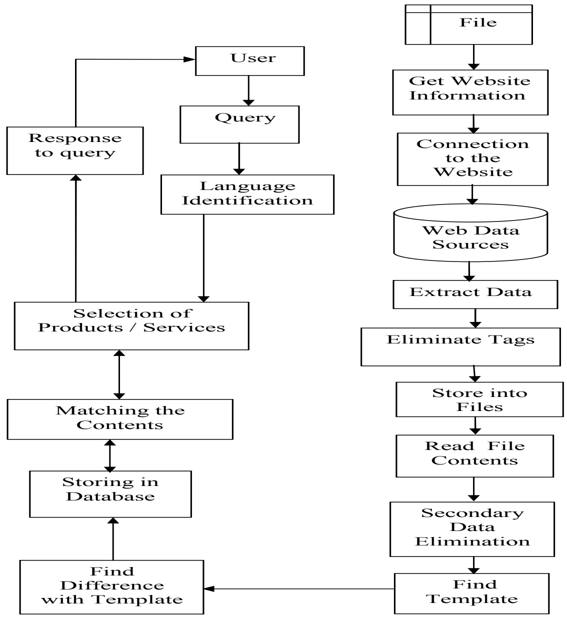


Fig. 1. Flow graph of the proposed framework

The user gives query input to the system. Then, the language of the query is identified. According to user's query, the system either finds the product / service from the website that offers minimum price or the product / service that has highest user rating. For finding the lowest price of a product, the system relies on the unique identifier of the product. This identifier is generated while an user submits his query. For finding a product / a service with highest rating related to user's query, it searches the related products / services and calculates average user rating of those products / services and sort them in descending order of average rating. Then, the system returns the results to the user.

A. Initialization and Connection

In the initialization and connection phase, at first, we store the information of the websites in a file. The algorithm for connecting to the website from file system is given in Fig. 2.

B. Elimination of Tags

A webpage consists different types of tags and unwanted data for making a website user friendly and attractive. However, we need to remove these tags as these tags do not contain important information for us. We define regular expressions to remove the tags. The tags elimination algorithm is given in Fig. 3.

Algorithm ***ConnectWeb***

1. Create table named on website name in the database with attributes information.
2. Read webpage information of each website stored in file system.
3. Initialize counter variable to 1.
4. While counter is not greater than length of the file
   4.1 Connect to the webpage and extract content of the page.
   4.2 Increment the value of counter by 1.

Fig. 2. Algorithm for initialization and connection to the website

```
Algorithm  Elimination_of_Tags
Begin
  for each word do
      match with regular expression syntax
      if word is matched then
          remove the word from content
      else
          do nothing
      end if
  end for
End
```

Fig. 3. Algorithm for elimination of tags

C. Template Finding

For finding template, first we split webpage contents using beginning and ending of each tag. This creates one multidimensional array for each webpage. After removal of unnecessary tags, we compare arrays of two webpages using array intersect. The array data that are similar are page template information. Then, we compare this result with all webpages and find the differences. The algorithm for template matching is given in Fig. 4.

D. Data Conversion

We often find that different e-commerce companies represent same data in different format. As for example, different book selling online stores listed their book price in different format and languages. So we need to convert the prices into single format for further processing. Similarly, we need to perform necessary conversion for ensuring data integrity and making further processing accurate.

```
Algorithm  TemplateFinding
Begin
 for each webpage in the file do
     split using beginning and ending tag
 end for
 set  word := NULL
 for first two webpage do
    find similarity between them
    word :=  similar data
 for each  webpage do
        find difference with word  array
        store difference into another array
    end for
End
```

Fig. 4. Algorithm for finding template

```
Algorithm  Language_Identification
Begin
 get user input
 find length in strlen() and mb_strlen()
 if they match then
        input  is  English
    else
        input is Bangla
    end if
 End
```

Fig. 5. Algorithm for language identification

E. Language Identification

We have the option for user to provide input in two different languages: bangla and english. Our system can automatically detect in which language the user has given the input. For this identification, we have used the strlen() and mb_strlen() function of PHP. If they matches, then the user has given input in English else the Language is Bangla. This is because mb_strlen() can detect bangla but strlen() can not detect Bangla. We provide algorithm for language identification in Fig. 5.

F. Query Processing

The query manager receives the users query through an interface. The user can select either English or Bangla language to enter his query. A user can search according to price or user rating. Then language identification module identifies the language of the user's input. After language identification, we compare requested product with the information stored in the database. If a match is found in a table, we check for that item in other tables as well. Remind that in our system for each online store under consideration, we maintain a table and each table contains exactly one record for a product. As a result we can perform parallel search to improve the search performance. After, checking all the tables, we collect top-*k* products / services those best matches with user's query and present them to the user.

IV. Implementations and Experiments

In this section, we provide the implementation procedure and performance analysis of our developed system.

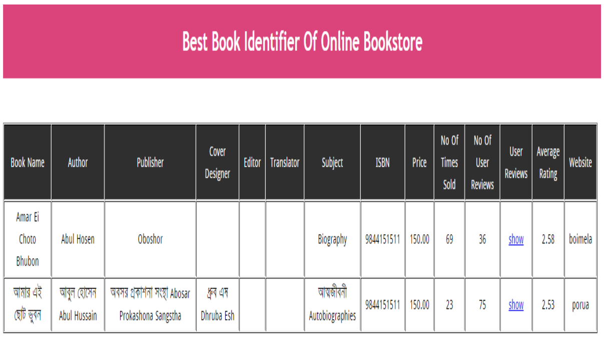
A. Experimental Setup

The book recommendation system has been developed on a machine having the operating system Windows 2007 Professional, 1.80 GHz Quad Core Processor with 6 GB RAM. The system was implemented in HTML5 in the front end and XAMPP 1.8.3 Mysql 5.6.16 in the back-end for storing the text database and related information.

We have extracted data from five different online book stores with different number of books. We have created 5 tables in the books database with same number and type of columns. Among the websites that are used for data extraction, two of them are in Bangla named Annyaprokash [16] and Porua [17] and the other three are in English which are Rubibook [18], Boimela [19] and Jain Book Agency [20]. Table I contains books information of these sites.

TABLE I
BOOKSTORE INFORMATION

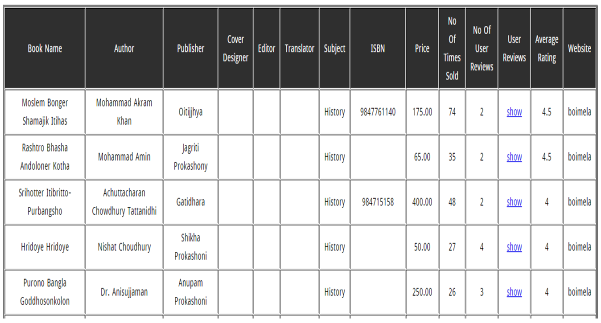
Name	Language	No. of Books
Boimela	English	15,571
Porua	Bangla	4,318
Rubibook	English	2,204
Annyaprokash	Bangla	819
Jain Book Agency	English	1,994



Best Book Identifier Of Online Bookstore

Book Name	Author	Publisher	Cover Designer	Editor	Translator	Subject	ISBN	Price	No Of Times Sold	No Of User Reviews	User Reviews	Average Rating	Website
Amar Ei Choto Bhubon	Abul Hosen	Oboshor				Biography	9844151511	150.00	69	36	show	2.58	boimela
আমার এই ছোট ভুবন	আবুল হোসেন Abul Hussain	অবসর প্রকাশনা সংস্থা Abosar Prokashona Sangstha	ধ্রুব এষ Dhruba Esh			আত্মজীবনী Autobiographies	9844151511	150.00	23	75	show	2.53	porua

Fig. 6. Query result based on price



Book Name	Author	Publisher	Cover Designer	Editor	Translator	Subject	ISBN	Price	No Of Times Sold	No Of User Reviews	User Reviews	Average Rating	Website
Moslem Bonger Shamajik Itihas	Mohammad Akram Khan	Oitijhya				History	9847761140	175.00	74	2	show	4.5	boimela
Rashtro Bhasha Andoloner Kotha	Mohammad Amin	Jagriti Prokashony				History		65.00	35	2	show	4.5	boimela
Srihotter Itibritto-Purbangsho	Achuttacharan Chowdhury Tattanidhi	Gatidhara				History	984715158	400.00	48	2	show	4	boimela
Hridoye Hridoye	Nishat Choudhury	Shikha Prokashoni				History		50.00	27	4	show	4	boimela
Purono Bangla Goddhosonkolon	Dr. Anisujjaman	Anupam Prokashoni				History		250.00	26	3	show	4	boimela

Fig. 7. Query result based on rating

TABLE II
PRECISION AND RECALL FOR RATING BASED RECOMMENDATION

k-values	Precision	Recall
2	1	1
3	1	1
4	0.75	1
5	0.80	1
6	0.67	1

B. Implementation

We have developed a website to facilitate users queries and generate responses of the users. We have made two options available for users' query. One is to search the lowest price for a particular book and the other is selecting a best book on a particular topic based on users' rating.

For example, Fig. 6 shows the results of a query that was issued for the lowest price of the book named "Amar Ei Choto Bhubon". Here, we can see that the query result shows all book data such as bookname, author, ISBN, no of times that book was sold, user rating, user review, website name etc. If we click the show button, then the user reviews of that book will appear. The results will be appeared from lowest price to highest price order. So that the user can find which store offers lowest price for that book and buy it from that book store.

If the user searches for the best book in a particular subject, our system will consider average rating of the users to select best product from the websites. For example, while searching for a good book on "History", our system responded with the results as shown in Fig. 7.

In this case, the book that has highest user rating represents better book. If several books have same average rating, our system checks the number of copies that was sold for each book and presents the book with more copies sold at the top of the results. In our system, we also have kept the facility of checking individual user's review on a particular book. For observing the individual's review on a particular book, the user needs to click on the show button at users reviews column.

C. Performance Evaluation

To evaluate recommendation accuracy of the proposed method, at first, we examined the recall and precision of recommendation while recommendation was generated based on average rating. Precision is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. Recall is the ratio of the number of relevant records retrieved to the total number of relevant records in the database. Table II shows the precision and recall results for different k-values.

We also performed a survey on twenty users of a public university to test the applicability of our developed system. Each of these twenty users used our system and answered following five questions with a rating between 1 and 5 where 1 means strongly disagree and 5 means strongly agree. Other rating are 2 for disagree, 3 for neutral, 4 for agree.

- Question 1: Do you think the system is helpful for you?
- Question 2: Do you think the system shows correct output?.
- Question 3: Do you think the system provides timely response?
- Question 4: Do you think the system has a good user interface?
- Question 5: Will you recommend others to use the

TABLE III
SURVEY RESULTS ABOUT THE SYSTEM

User	question 1	question 2	question 3	question 4	question 5
user 1	5	4	4	2	5
user 2	5	4	5	3	5
user 3	4	3	4	3	5
user 4	4	5	4	4	5
user 5	5	5	4	4	5
user 6	4	3	5	4	5
user 7	5	4	3	3	5
user 8	4	4	5	4	5
user 9	4	2	4	4	4
user 10	5	5	4	3	5
user 11	4	5	5	5	5
user 12	5	3	3	3	4
user 13	5	4	5	1	5
user 14	4	4	4	4	4
user 15	4	4	5	3	4
user 16	5	3	5	2	4
user 17	5	5	4	3	5
user 18	5	4	5	4	5
user 19	5	2	3	4	4
user20	4	5	4	2	4
Average Rating	4.55	3.90	4.05	3.25	4.65

system?

The survey results that has been found is given in Table III. From the results shown in Table III, we can see that most of the users strongly agreed with the usefulness of the system and wanted to recommend the system to others. They are also satisfied with the system's performance. However, most of them are not satisfied with the user interface of the system.

V. Conclusion

In this paper, we provided a framework for analysing web data to provide recommendations of products / services to the users. We have implemented the system with Bangla and English online bookstore and recommended books to the users based on price or average rating. Form the experimental evaluation, we found that the developed system can provide efficient recommendation. Although in this paper, we have considered websites of two different languages, it can incorporate websites of other languages easily.

In this paper, we did not analyse text-based comments for getting insight about products/ services. In future, we want to plan to incorporate such facilities in our system to generate more efficient recommendation. Moreover, we plan to experiment the performance of our system over larger domain.

References

[1] Amazon E-commerce Company. Available at: http://www.amazon.com

[2] eBay Inc. Available at: http://www.ebay.com

[3] Olex. Available at: http://www.olx.com

[4] Agoda Company Pte Ltd. Available at: http://www.agoda.com

[5] TripAdvisor. Available: http://www.tripadvisor.com/

[6] Expedia. Available at: http://www.expedia.com

[7] Rakuten: Available at: http://www.rakuten.com

[8] Elance: Available at: https://www.elance.com

[9] J. Hammer, H. Garcia-Molina, J. Cho, R. Aranha, and A. Crespo, "Extracting semistructured information from the web", *In Proc. of the Workshop on Management of Semistructured Data*, Tucson, Arizona, May 1997.

[10] M. Shaker, H. Ibrahim, A. Mustapha and L. Abdullah, "A Framework for extracting information from semi-structured web data sources", *In Proc. of Third International Conference on Convergence and Hybrid Information Technology*, pp. 27-31, 2008.

[11] C. A. Knoblock, K. Lerman, S. Minton, and I. Muslea, "Accurately and reliably extracting data from the web: a machine learning approach", *In Proc. of IEEE Data Engineering Bulletin*, pp. 33-41, 2000.

[12] D. S. Soper, "A Framework for automated web business intelligence systems", *In Proc. of 38th Hawaii International Conference on System Sciences (HICSS)*, Jan. 3-6, 2005.

[13] R. Baumgartner, P. Harz, O. Frolich, M. Herzog, G. Gottlob, and P. Lehmann, "Web data extraction for business intelligence: the lixto approach", *In Proc. of 11th German Biannual Conference on Databases for Business, Technology and the Web(BTW)*, Karlsruhe, Germany, pp. 30-47, 2005.

[14] S. Bhawsar, K. Pathak, S. Mariya, and S. Parihar, "Extraction of business rules from web logs to Improve web usage mining", *In International Journal of Emerging Technology and Advanced Engineering*, ISSN 2250-2459, Volume 2, Issue. 8, pp. 333-340, August, 2012.

[15] K. Lerman, C. Knoblock, and S. Minton, "Automatic data extraction from lists and tables in web sources", *In Proc. of Automatic Text Extraction And Mining Workshop*, 2001.

[16] Annyaprokash. Available at: http://e-anyaprokash.com

[17] Porua. Available at: http://www.porua.com.bd

[18] Rubibook. Available at: http://www.rubibook.com

[19] Boimela. Available at: http://www.boi-mela.com

[20] Jain Book Agency. Available at: http://www.franchisebangladesh.com/Jain-Book-Agency/

Development of a Smart Learning Analytics System Using Bangla Word Recognition and an Improved Document Driven DSS

Wali Mohammad Abdullah, Kazi Lutful Kabir
Department of CSE
Military Institute of Science and Technology
Dhaka, Bangladesh
shashi.mist@gmail.com,mlutful.02526@gmail.com

Nazmul Hasan, Sharmin Islam
Department of CSE
Military Institute of Science and Technology
Dhaka, Bangladesh
nazmulcse10@gmail.com,priyaislam122@gmail.com

***Abstract*—A smart learning analytics system is proposed here. In the field of research, education learning analytics is manipulated as an optimistic way to render learning. In this proposed system, we put forward an idea regarding the design of learning analytics based on document-driven Decision Support System (DSS) and Bangla handwritten words recognition. The proposed learning analytics system aims at classifying the documents via syntactic analysis as well as categorizing them by semantic analysis of their contents. But the documents may be presented as printed format or hand written format. To choose the optimal feature vector for hand written words recognition (especially for Bangla basic characters) projection-based features are mostly efficient. Decision makers use DSS for decision making which is actually a computerized information system. We propose a route through the classified observations for providing the explanation which may improve the acceptance of the decision maker with a better correction rate using kNN algorithm. This system can certainly assist researchers for learning as well as decision making.**

Keywords- Learning analytics; artificial neural network; document-driven DSS; data mining; word recognition.

I. Introduction

The usage of “Big data” can be considered as s growing phenomenon in various fields like computer science, political science, medicine and economics, physics and social sciences. The process of examining these large amounts of data to uncover hidden patterns, unknown correlations and other useful information is referred to as “Big data” analytics.

Analytics is also perceived as a reliable tool for decision-making in the education sector. Analytics in education borrow techniques from different fields. Among these fields Educational Data Mining (EDM), Social Network Analysis, web analytics and Information Visualization facilitate the exploration of data coming from educational contexts [1].

In learning analytics system, we have a huge amount of data in the database. Now, from this database the system should provide the exact data or file which is asked by the users. So, we need a decision support system to make clusters of these documents and to categorize them as required. Now-a-days, data grow too fast. It has exceeded human capacity to retrieve and utilize the content of the documents from the database which has become a problem for the decision makers. Decision makers often use decision support system in order to get rid of the plight of this information explosion. A successful DSS depends on the accuracy of the results as well as on acceptance of these results by its users [2]–[6].

In Bangladesh, either of the two languages (i.e. English and Bengali) is used for learning. Besides, the document may be preserved in printed form or in hand written form. Among them, segmenting the Bangla (i.e. Bengali) handwritten words from a document and then identify the exact contents from the database is a challenge. Inclusion of projection-based features yields a greater positive impact on recognition rate rather than of inclusion of other features [7]

A. Our Contribution

In our system we propose to use DSS for categorizing the documents. Another contribution is the consequential identification of the exact documents by using projection based features. The rest of the paper has been organized as follows. In Section II, we discussed the present system for learning analytics and associated preliminary concepts. Our proposed system is described in Section III. Finally, Section IV concludes the paper.

II. Present System

The primary objective of the Learning Analytics (LA) is to improve the efficiency of the distance learning. In the present LA system, the major activities are the collection and analysis of the materials as well as reporting the data to the learners [8]. Most of the cases the LA system is used to identify the necessity and measurement of the improvements of the study materials, evaluate the performance of the students in a particular course and provide the necessary assistance regarding any kind of suggestion, sometimes it also evaluate the performance of the teacher and provide necessary suggestions. Finally LA makes a very efficient and effective relationship among the study materials, learners, and instructors [1].

III. Proposed System

In this section we have presented our proposed system for learning analytics. Our system uses an improved DSS for categorizing the documents and to identify the exact contents from

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

TABLE I
WORST CASE PERFORMANCE OF kNN ALGORITHM

ClassLabels	{3x1 cell}
GroundTruth	[150x1 double]
NumberOfObservations	150
ControlClasses	[2x1 double]
TargetClasses	1
ValidationCounter	1
SampleDistribution	[150x1 double]
ErrorDistribution	[150x1 double]
SampleDistributionByClass	[3x1 double]
ErrorDistributionByClass	[3x1 double]
CountingMatrix	[4x3 double]
CorrectRate	0.9333
ErrorRate	0.0667
LastCorrectRate	0.9333
LastErrorRate	0.0667
InconclusiveRate	0
ClassifiedRate	1
Sensitivity	1
Specificity	1
PositivePredictiveValue	1
NegativePredictiveValue	1
PositiveLikelihood	NaN
NegativeLikelihood	0
Prevalence	0.3333
DiagnosticTable	[2x2 double]

TABLE II
AVERAGE CASE PERFORMANCE OF kNN ALGORITHM

ClassLabels	{3x1 cell}
GroundTruth	[150x1 double]
NumberOfObservations	150
ControlClasses	[2x1 double]
TargetClasses	1
ValidationCounter	10
SampleDistribution	[150x1 double]
ErrorDistribution	[150x1 double]
SampleDistributionByClass	[3x1 double]
ErrorDistributionByClass	[3x1 double]
CountingMatrix	[4x3 double]
CorrectRate	0.9400
ErrorRate	0.0600
LastCorrectRate	0.8667
LastErrorRate	0.1333
InconclusiveRate	0
ClassifiedRate	1
Sensitivity	1
Specificity	1
PositivePredictiveValue	1
NegativePredictiveValue	1
PositiveLikelihood	NaN
NegativeLikelihood	0
Prevalence	0.3333
DiagnosticTable	[2x2 double]

TABLE III
BEST CASE PERFORMANCE OF kNN ALGORITHM

ClassLabels	{3x1 cell}
GroundTruth	[150x1 double]
NumberOfObservations	150
ControlClasses	[2x1 double]
TargetClasses	1
ValidationCounter	1
SampleDistribution	[150x1 double]
ErrorDistribution	[150x1 double]
SampleDistributionByClass	[3x1 double]
ErrorDistributionByClass	[3x1 double]
CountingMatrix	[4x3 double]
CorrectRate	0.9800
ErrorRate	0.0200
LastCorrectRate	0.9800
LastErrorRate	0.0200
InconclusiveRate	0
ClassifiedRate	1
Sensitivity	1
Specificity	1
PositivePredictiveValue	1
NegativePredictiveValue	1
PositiveLikelihood	NaN
NegativeLikelihood	0
Prevalence	0.3333
DiagnosticTable	[2x2 double]

documents, we have utilized projection based features. In the following sections we have presented them elaborately.

A. Improved Document-Driven DSS

The enhancement of the effective communication between the users and DSS is the key element in DSS design. Computer systems are used to solve semi-structured or unstructured problems which are characterized by a complex decision-making process. We have emphasized on one popular algorithm (*kNN* classifier algorithm) to solve this problem. We have used a route through the classified observations for providing the explanation which has improved the acceptance of the decision maker. We can classify all observations into user specified clusters. After implementing the algorithm and then performing accuracy test we have observed that *kNN* algorithm can be used for better correction rate. The worst case, average case and the best case performances of *kNN* algorithm is presented in Table I, II and III respectively.

B. Document Identification System

The documents stored can be either printed or handwritten. Recognition of printed documents won't be a tough task as already developed OCR(for printed) is commercially available for not only Bangla but also many other languages with acceptable recognition rate.

The most challenging job is to recognize the contents from a handwritten Bangla document. In order to accomplish this objective, the first line from the document is segmented as document title. Then, next task is to identify the words of the first line. To recognize the words, step by step character recognition is the suitable approach in this case. Feature extraction is the most important step of character recognition. According to Kabir et. al. for handwritten Bangla character recognition, projection-based features plays a vital role to constitute optimal feature vector [7].

Three projection-based feature sets have been considered to constitute the feature vector. These are: projection profile features (both left and right) and shadow features.

1) Projection-Based Features:

- Left Projection Profile Features: It represents the projection from left side of a particular character. For all rows, the

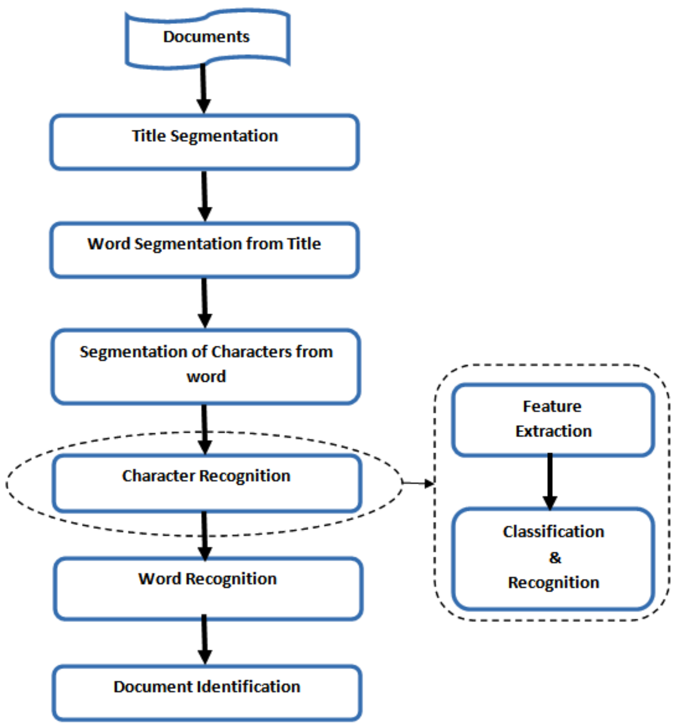


Fig. 1. Document Identification System Architecture

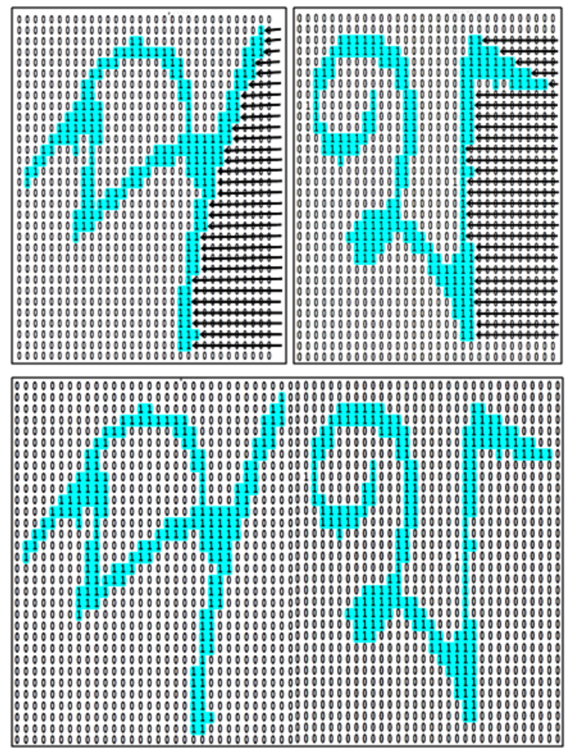


Fig. 2. Illustration of right projection profile features to recognize individual characters and word recognition

total number of non-black pixels until reaching the starting left side portion of the character is counted [7].

- Right Projection Profile: It represents the projection from right side of a particular character. For all rows, the total number of non-black pixels until reaching the starting right side portion of the character is counted [7].
- Shadow Features: Lengths of projections of the character images on the four sides and eight octant dividing sides of the minimal bounding boxes enclosing the character are computed. Total 24 shadow features are extracted from each character as lengths of projections on three sides of each octant are considered [9].

Finally an artificial neural network fulfills the tasks of classifier to complete the recognition process. This character recognition step leads to word recognition. After that, the recognized words are checked against the learned contents. If found, then the whole document is included in the cluster of the matched content(s). If not found, then the database will be updated for new content.

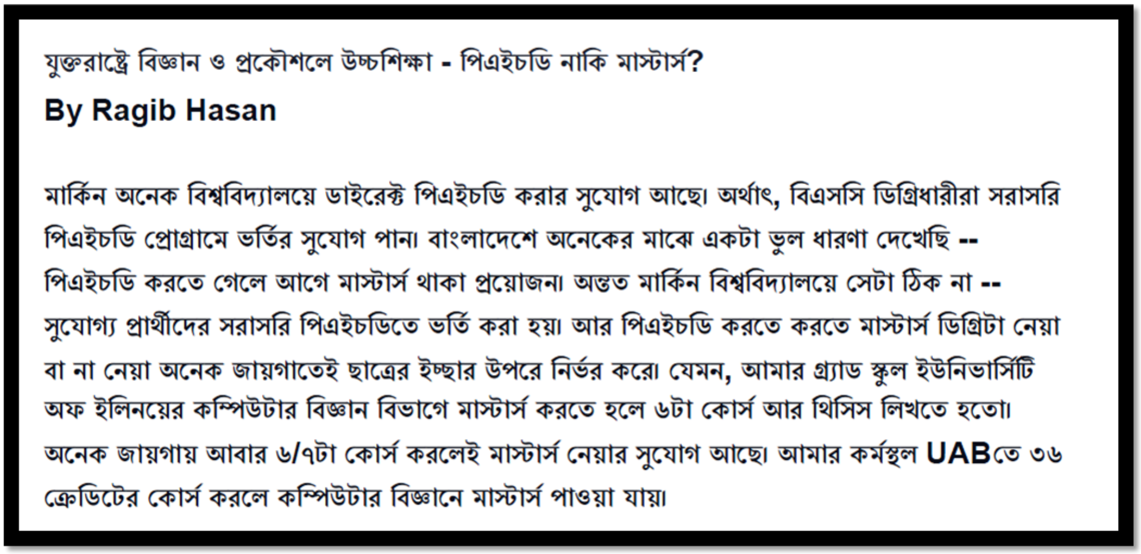
We have tested total 15 documents (Printed-10, Handwritten-05) with different header styles (Some samples are presented in Figure 3 and 4). Our system exhibited flawless performance in title segmentation and word segmentation from a document (Table IV). After that, the segmented words of the titles are checked for matching of contents.

TABLE IV
PERFORMANCE ANALYSIS

Number of Documents Tested	**Title Segmentation**	**Word Segmentation from title**	**Detection and Classification Accuracy**
15 (Printed-10, Handwritten-05)	100% (First line as title)	100%	11 documents (Printed-08, Handwritten-03) has been accurately identified and classified. Accuracy in this particular test set- 73.33%

IV. CONCLUSION

In this paper, learning analytics is recognized as a dominant tool for helping people to reflect on their learning activities and a smart learning analytics system has been proposed. An strategy about the design of learning analytics based on document-driven Decision Support System (DSS) and Bangla handwritten words recognition have been used in our approach. The documents have been classified as well as the contents of the documents have been categorized meaningfully. For learning purpose, the documents may be preserved in printed format or handwritten. The optimal feature vector (consisting of projection-based features) have been constituted for handwritten Bangla words recognition. DSS, which is actually a computerized information system, has also been used for the purpose of decision making.



যুক্তরাষ্ট্রে বিজ্ঞান ও প্রকৌশলে উচ্চশিক্ষা - পিএইচডি নাকি মাস্টার্স?

By Ragib Hasan

মার্কিন অনেক বিশ্ববিদ্যালয়ে ডাইরেক্ট পিএইচডি করার সুযোগ আছে। অর্থাৎ, বিএসসি ডিগ্রিধারীরা সরাসরি পিএইচডি প্রোগ্রামে ভর্তির সুযোগ পান। বাংলাদেশে অনেকের মাঝে একটা ভুল ধারণা দেখেছি -- পিএইচডি করতে গেলে আগে মাস্টার্স থাকা প্রয়োজন। অন্তত মার্কিন বিশ্ববিদ্যালয়ে সেটা ঠিক না -- সুযোগ্য প্রার্থীদের সরাসরি পিএইচডিতে ভর্তি করা হয়। আর পিএইচডি করতে করতে মাস্টার্স ডিগ্রিটা নেয়া বা না নেয়া অনেক জায়গাতেই ছাত্রের ইচ্ছার উপরে নির্ভর করে। যেমন, আমার গ্র্যাড স্কুল ইউনিভার্সিটি অফ ইলিনয়ের কম্পিউটার বিজ্ঞান বিভাগে মাস্টার্স করতে হলে ৬টা কোর্স আর থিসিস লিখতে হতো। অনেক জায়গায় আবার ৬/৭টা কোর্স করলেই মাস্টার্স নেয়ার সুযোগ আছে। আমার কর্মস্থল **UAB**তে ৩৬ ক্রেডিটের কোর্স করলে কম্পিউটার বিজ্ঞানে মাস্টার্স পাওয়া যায়।

Fig. 3. Sample Document (Title left-aligned)

বাংলাদেশ ডেন্টাল কলেজ
বাড়ী নং-৩৫, রোড নং-১৪/এ, ধানমন্ডি, ঢাকা-১২০৯
ফোন ঃ ৯১২০৭৯২-৩, ০১৫৫২৪৪৩৫৫০

১ম বর্ষ বিডিএস কোর্সে ভর্তির বিজ্ঞপ্তি

বেসরকারী পরিমন্ডলে প্রথম সারির মানসম্পন্ন, অত্যাধুনিক যন্ত্রপাতি সমৃদ্ধ, অভিজ্ঞ শিক্ষকমন্ডলী দ্বারা পরিচালিত, বিশ্ব স্বাস্থ্য সংস্থা (WHO) কর্তৃক স্বীকৃতিপ্রাপ্ত, বাংলাদেশ মেডিক্যাল কলেজ সংলগ্ন হওয়ায় মেডিক্যাল বিষয়সমূহে শিক্ষদানের সুব্যবস্থা ও শিক্ষার্থীদের বিনামূল্যে চিকিৎসা সুবিধা সম্বলিত একমাত্র প্রতিষ্ঠান "বাংলাদেশ ডেন্টাল কলেজে" ১ম বর্ষ বিডিএস কোর্সে ভর্তির ফরম বিতরণ করা হচ্ছে। সরকার কর্তৃক গৃহীত ভর্তি পরীক্ষায় মেধা তালিকার ১ হতে ৩৮৫৩২ পর্যন্ত ছাত্র-ছাত্রী ভর্তি ফরম সংগ্রহ করতে পারবেন। ভর্তি ফরমের মূল্য ১,০০০/- (এক হাজার) টাকা মাত্র।

ভর্তি ফরম ক্রয় ও জমাদানের তারিখ ঃ	১৮/১০/২০১১হতে ১৫/১১/২০১১
ভর্তি ফরম ক্রয় ও জমাদানের শেষ তারিখ ঃ	১৫/১১/২০১১ (৮- ২:৩০)
নির্বাচিত প্রার্থীদের তালিক প্রকাশ (কলেজ নোটিশ বোর্ডে) ঃ	১৭/১১/২০১১
ভর্তির তারিখ ঃ	১৯/১১/২০১১ হতে ০৫/১২/২০১১ পর্যন্ত।

অধ্যাপক ডাঃ মুহাম্মাদ আমীরুল ইসলাম
বিডিএস, এমএস, পিডিটি, পিএইচডি, এফআইসিডি(ইউএসএ)
অধ্যক্ষ,বাংলাদেশ ডেন্টাল কলেজ

Fig. 4. Sample Document (From Newspaper)

A route through the classified observations has been used for providing the explanations which may improve the acceptance of the decision maker with a better correction rate using *kNN* algorithm. Finally, this system can certainly assist researchers for learning analytics as well as decision making.

References

[1] E. Durall and B. Gros, "Learning Analytics as a Metacognitive Tool."

[2] D. A. Savić, J. Bicik, and M. S. Morley, "A DSS generator for multiobjective optimisation of spreadsheet-based models," *Environmental Modelling & Software*, vol. 26, no. 5, pp. 551–561, 2011.

[3] H. Kopackova, J. Komarkova, and P. Sedlak, "Decision making with textual and spatial information," *WSEAS Transactions on Information Science and Applications*, vol. 5, no. 3, pp. 258–266, 2008.

[4] M. S. Gönül, D. Önkal, and M. Lawrence, "The effects of structural characteristics of explanations on use of a DSS," *Decision Support Systems*, vol. 42, no. 3, pp. 1481–1493, 2006.

[5] P.-N. Tan, M. Steinbach, V. Kumar *et al.*, *Introduction to data mining*. Pearson Addison Wesley Boston, 2006, vol. 1.

[6] D. J. Power, "A brief history of decision support systems," *DSSResources. COM, World Wide Web, http://DSSResources. COM/history/dsshistory. html, version*, vol. 4, 2007.

[7] K. L. Kabir, M. K. Shafin, T. T. Anannya, D. Debnath, M. R. Kabir, M. A. Islam, and H. Sarwar, "Projection-Based Features: A Superior Domain for Handwritten Bangla Basic Characters Recognition," in *Proceedings of the 9th International Conference on Intelligent Systems and Control(ISCO)*, vol. 2. IEEE, 2015, pp. 1054–1060.

[8] Y. Tabaa and A. Medouri, "LASyM: A Learning Analytics System for MOOCs," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 4, no. 5, 2013.

[9] M. Karim, M. Kaykobad, and M. Murshed, *Technical Challenges and Design Issues in Bangla Language Processing*. IGI Global, 2013.

A Secret Key-Based Security Architecture for Wireless Sensor Networks

Anuva Chowdhury, Farzana Alam Tanzila, Shanta Chowdhury and Md. Mokammel Haque*
Faculty of Electrical and Computer Engineering
Chittagong University of Engineering and Technoloty (CUET)
Chittagong-4349, Bangladesh
Email: *malin405@yahoo.com

***Abstract*—Information sharing and transmitting has been increased exponentially in today's world of Information and Communication Technology. Cryptography is a lasting solution to save the sensitive data from unauthorized access. Finding proper cryptographic structure for wireless sensor network (WSN) is almost challenging as there are some limitations of energy and memory resources of WSNs. Hence, this paper presents an efficient secret key based security architecture which aims to contribute robust security to the topic with ideas to man in the middle attacks and many other hacker situations. In this scheme, singular value decomposition (SVD) of pseudo-inverse matrix has been used for key generation. This scheme does not require key handshaking for data transmission and reception between the nodes. The analytical results demonstrate that this proposed scheme shows a significant gain in the rank of security, and is suitable for sensor nodes of the modern age.**

Keywords—singular value decomposition; pseudo-inverse matrix; secret key; wireless sensor network

I. INTRODUCTION

Security is an obligatory necessity for nearly every sensor set-up purpose. In case of wireless sensor networks (WSNs), security is considered as a standalone constituent and required sufficient consideration. In most function domains, the sensors are employed to accumulate a definite form of data from particular object areas. The accumulated data are repeatedly considered unrevealed and are not aimed for public expose. Thus, proficient and protected systems are required to send out obtained data safely to the right addressees.

WSNs transmit secret information, which, if exposed to adversary parties, might be a reason of catastrophe for forthcoming parties. Mainly in emergency purposes of WSN, it is essential to employ security systems for data transmissions. The efficacy of sensor networks is significantly curtailed, if an enemy spoils the action of the network by altering the transmitted information, ending communication, or by eavesdropping on information. So, it is necessary to ensure a superior rank of protection for this network communication. Cryptography ensures the role of secured secret communication through the WSNs. A variety of algorithms have been introduced to cope up with the upcoming challenges compromising the cost, energy consumption and security level [1].

A bivariate symmetric polynomials has been used for key pre-distribution [2]. A symmetric key distribution algorithm is found [3], which is the combination of the Blom's scheme and the random key pre-distribution method. These innovative efforts have further been expanded, advanced and exploited for different circumstances [4], [5], [6], [7], and [8]. Logic operation has been performed for key generation to increase the security level of WSN's [9]. Likewise, encryption and decryption have been done by the XOR operation of ASCII values [10]. A clustered based architecture using hash function has been proposed to maximize the lifetime of the sensors [11]. Asymmetric key-based architecture (AKA) is also found in the literature to secure data transmission in WSN [12]. Here, AKA applies a linear assembly of pseudo-inverse matrix to produce asymmetric keys to launch a protected communication between connecting nodes and the base-station. AKA is asserted to be more secured than the Diffie-Hellman key exchange protocol (D-H) [13]. The D-H protocol is a symmetric key cryptography that considers the dimension of the group from which the key is being generated. It makes two parties able to create a key with no previous data of each other. However, it is demonstrated that the AKA is not an asymmetric scheme, but a symmetric scheme. It also does not provide authentication semantics and is insecure under impersonation attack [14].

Here, in this paper, we exploit the properties of pseudo inverse matrix for node to node communication, while SVD is employed for key generation. The analytical results illustrate that the proposed scheme provides significant gain in terms of security. In addition, it proves suitable for wireless sensor networks considering the resource constraint of the sensor nodes.

The upcoming parts of this paper are as follows: network assumption and fundamentals are stated in section II, where section III presents performance analysis and assessment, and section IV concludes the methodology highlighting its motives for choosing this scheme with future opportunity.

II. NETWORK ASSUMPTION AND FUNDAMENTALS

It is assumed that the base station is a reliable entity and the sensor nodes are the valid nodes. They have adequate storage as well as power to process and execute the operations. The base station is supposed to have the capacity to store the fingerprints of all authentic nodes and the sensor nodes are assumed to be similar to the current sensor nodes (MICA2) in terms of memory, communication, computational and energy resources. When the sensors are set up in the network, it is assumed that they are static in comparison with their relative positions throughout the communication.

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

A. Pseudo-inverse matrix and Singular value decomposition

The pseudo-inverse matrix, $A^{\dagger}(n \times m)$ is used to get the inversion of a non-square matrix, $A(m \times n)$. The pseudo-inverse matrix is only one of its type over any field. An important property of pseudo-inverse matrix is as follows:

$$AA^{\dagger} = I_m \tag{1}$$

For a general m by n matrix, there might be more than one pseudo-inverse matrix. This property is used to our cryptographic operations. A non-square matrix can be expressed by three matrices from SVD. The SVD of A :

$$A = UDV^T \tag{2}$$

where, $U\,(m \times m)$, and $V^T\,(n \times n)$ are both orthogonal matrices and $D = [S\,0; 0\,0]$. Here, S is a diagonal matrix containing the (positive) singular values of A on its diagonal and D is the singular valued matrix of A.

The pseudo-inverse matrix of A is $A^{\dagger}(n \times m)$ can be defined by:

$$A^{\dagger} = VD^{-1}U^T \tag{3}$$

So, from the equations (1) and (3) it can be written that,

$$AVD^{-1}U^T = I_m \tag{4}$$

B. Our Secret key-Based Security Architecture

The architecture of our proposed methodology is based on SVD decomposition of pseudo-inverse matrix. Two sections have been cast off in the whole scheme. The first part is the key distribution between two valid nodes through base station and the second portion is data transmission and reception between the nodes. Key distribution [4] is the process of sending the secret key from the sender node to the receiver node to decrypt the encrypted message from the sender node. Here, in our approach the process has gone through the base station. The remaining part, data transmission and reception, is the process of sending the chipper text from the sender node and receiving the plaintext by the real receiver to whom the sender wants to send the message. To get the plain text, receiver has to decrypt that chipper text with the correct secret key.

C. Secret Key distribution between two nodes through the base station

First of all, two authentic nodes n_1 and n_2 and a base station in a network are considered.

- Node n_1 generates a random matrix $A_1(m \times n)$ and gets three matrices through SVD; $U_1(m \times m)$,$D_1\,(m \times n)$, $V_1^T\,(n \times n)$.
- Then n_1 calculates three inverse matrices; $U_1^T(m \times m)$,$D_1^{-1} = [S_1^{-1}\,0; 0\,0]$,$V_1(n \times n)$.
- Node n_1 transmits secret key $D_1^{-1}U_1^T$ to the base station in a secured manner.
- When base station receives this key, it transmits the key to node n_2 following the same security arrangement.
- When n_2 receives the key from base station, stores it for further usage to decrypt the encrypted message form node n_1.
- By this time, node n_2 also generates another random non-square matrix, A_2; gets U_2, D_2,V_2^T; calculates U_2^T, $D_2^{-1} = [S_2^{-1}\,0; 0\,0]$ and V_2.Then transmits $D_2^{-1}U_2^T$ to node n_1 through the base station to node n_1. n_1 stores the key for further decryption of the encrypted message sent from node n_2.

The keys are distributed between nodes by the base station. Once the keys are distributed to the corresponding authentic nodes, these keys are stored for future communication between them. So, it is not necessary to distribute the keys repeatedly.

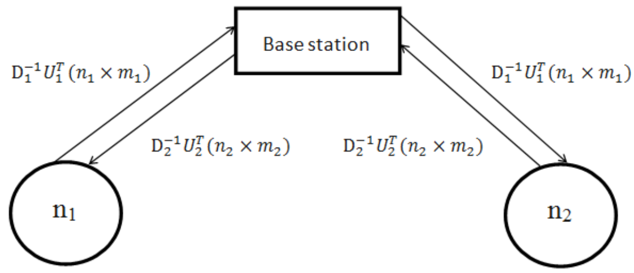


Fig. 1. Key distribution between two nodes through the base station.

Moreover, these keys are generated by the nodes and allocated to the respective nodes by means of the base station in a protected way. As a result, there is a bit chance to get the secret shared keys by any means.

D. Data Transmission and Reception for Node-to-Node Communications

After receiving the keys, the nodes are able to communicate between them. A node can send the encrypted message by multiplying the message with its own matrices. For example, M_1 is the plaintext. Node n_1 wants to send this text to n_2. So, n_1 encrypts this plaintext M_1 by multiplying $M_1$with $A_1$and V_1 and transmits the encrypted message $M_{1(E)}$ to n_2.

$$M_1A_1V_1 = M_{1(E)} \tag{5}$$

Fig. 2. Data transmission and reception between two nodes.

Since, the decryption key ($D_1^{-1}U_1^T$) has already been shared to node n_2 by the base station, n_2 can easily decrypt the message after receiving the encrypted message.

$$M_{1(E)}D_1^{-1}U_1^T = M_1A_1V_1D_1^{-1}U_1^T = M_1A_1A_1^{\dagger} = M_1 \quad (6)$$

The process is shown in Fig. 2. In this proposed scheme, every time, the sender node uses its own matrices to encrypt the message and the receiver node uses the shared secret key to decrypt the message. If the message is exposed to the attacker, it will not get the plaintext until it gets the decryption key. As the key distribution takes place only once during the whole communication, the attacker will not be able to get the plaintext.

III. PERFORMANCE ANALYSIS AND ASSESSMENT

The specifications of MICA2DOT mote platform are considered to assess the performance of the proposed security architecture. This secret key-supported structure is evaluated in terms of security, scalability, energy and memory requirements, and computational cost. In the present section, the detailed analysis of our scheme has been discussed.

Key distribution is done by means of linear matrix operations and matrix multiplications. The complication of matrix multiplication is incredibly low; thus it can be executed very swiftly. In our secret key distribution scheme, node n_1 sends the base station a $n\times m$ matrix, which used nm bits. Base station sends this $n\times m$ matrix to node n_2. It again consists of nm bits. So, the total amount of bits employed to distribute the key from a node to another node through base station is,

$$\begin{aligned} &= nm + nm \\ &= 2nm \text{ bits} \end{aligned} \quad (5)$$

Risks associated with the identification of the involved entities during key sharing has fully been eradicated as; (a) the base station is a reliant being and cannot be compromised in any way, and (b) the identifications of the nodes deployed in the network are checked by the base station before distributing the secret keys. During the time of node-to-node conversation, the receiver node can decrypt the encrypted message by the key received earlier from the sender node through the base station. So, it is not possible for an intruder to get the shared keys of a definite sender-receiver couple.

Let us compare our secret key derivation scheme with asymmetric key based architecture [8]. In the AKA, the security level is 2^{75} for 105 bits key size and the respective value of (m, n, r, k) is (7,12,7,15). While for the same (m, n, k), the level of security provided by the proposed scheme is 2^{84} for 168 bit size. It is found from TABLE I and TABLE II that though a greater key size is required in all cases in our scheme than asymmetric key based scheme for the same value of (m, n, r, k), the security level is superior when the dimension of m is larger (especially m>6). The security level of AKA is based on the probability to crack the scheme. In this scheme, if the key is exposed to eavesdropper by any means, the network node is affected. On the other hand, in our scheme the message is encrypted by the multiplication of the matrices (the random matrix, two inverse matrix found from the SVD of the random matrix) and one matrix (the rest matrix to get the pseudo inverse of the random matrix) is used as the key, shared to the base station in a secured manner. Without the key, the eavesdropper cannot attack the scheme. So, it is not possible to crack the scheme by eavesdropper or man-in-the-middle attacker until they search out the key.

TABLE I. KEY SIZE AND SECURITY LEVEL FOR ASYMMETRIC KEY BASED ARCHITECTURE

(m, n, r, k)	**(4,8,4,12)**	**(5,9,5,15)**	**(6,11,6,14)**	**(7,12,7,15)**
Key size (mk bits)	48	75	84	105
Security level $2^{(n-r).k}$	2^{48}	2^{60}	2^{70}	2^{75}

TABLE II. KEY SIZE AND SECURITY LEVEL FOR SECRET KEY BASED ARCHITECTURE

(m, n, k)	**(4,8,12)**	**(5,9,12)**	**(6,11,14)**	**(7,12,15)**
Key size (nm bits)	32	45	66	84
Security level 2^{nm}	2^{32}	2^{45}	2^{66}	2^{84}

If we compare two schemes in case of the volume of traffic, our scheme will be more feasible than the asymmetric key based scheme. The volume of traffic is related to the total number of bits that has been carried through the whole scheme. In our scheme, total volume of traffic will be 96 bits whereas this value is 304 bits for asymmetric key based scheme. The values of traffic volume are 135, 198 and 252 bits for different (m,n,k) in our scheme. On the other hand, for AKA the values are 417, 546, 657 bits respectively. These values are considered for (m, n, k) of (4, 8, 12), (5, 9, 15), (6, 11, 14) and (7, 12, 15) respectively. So, it is clear that our methodology requires less traffic volume comparing with AKA.

For comparing the communication cost, consider the specification of MICA2DOT motes. The transmission and reception cost of one byte is 59.2 µJ and 28.6 µJ. For simplified calculation, consider total 49 bytes (payload 32 bytes; header 9 bytes; packet ID, source and destination address 8 byte) in a packet. In this mote, execution of microcontroller of 2090 clock cycle requires to transmit one bit. So, the energy required to transmit $49\times59.2 = 2.9$ mJ and to receipt $49\times28.6 = 1.4$ mJ for one packet. In AKA, in case of sender, it has to transmit once to receiver node and when it goes to communicate to the base station, the number of transmission is twice and reception is once. So, the sender consumes $(3\times2.9) +1.4=10.1$ mJ energy. For receiver node, the energy consumption is $2\times1.4+2.9=5.7$ mJ since the total number of transmission is once (base station) and the number of reception is twice; one to sender node and other to base station. On the other hand, in our scheme, the sender has to transmit to base station and receiver node separately by once. So, total power consumption in case of sender node is 2×2.9=5.8 mJ. The receiver receives the key from the base station and encrypted message from the sender node. So, the power consumption in case of receiver node is 2×1.4=2.8 mJ.

Comparing the power consumption, our scheme shows the better performance than the previous one though it requires additional memory requirements.

When any authentic node is added in a wireless sensor network, it transmits its keys to the base station in a secured manner for the existing nodes. The base station distributes the keys to the nodes and the nodes store the keys and generate key for the newly added node.

TABLE III. COMPARISON OF POWER CONSUMPTION WITH ASYMMETRIC KEY BASED ARCHITECTURE

Name of scheme	Communication Costs	
	Sender (mJ)	**Receiver(mJ)**
Asymmetric key based architecture [12]	10.1	5.7
Our Secret key based architecture	5.8	2.8

This key sends back to the base station in the same manner. Thus, the nodes communicate when an authentic node is added in WSN.

IV. CONCLUSION

To smooth the progress of holistic security construction for WSNs, in this paper we propose a symmetric cryptographic technique using the properties of SVD and pseudo inverse matrix. The cryptanalysis ensures that the proposed architecture is indeed a strong one considering the enhanced security level comparing with AKA. Another contribution of this paper is the reduced communicating cost comparing with the similar types of architectures. Our secret key-based architecture does not need certificate authority at all and thus it keeps away from managing and validating massive number of extra communications due to the certificates. In future we plan to merge our work with other security protocols related with different function levels, to assemble a total security plan for WSNs.

REFERENCES

[1] C. Xiangqian, M. Kia, Y. Kang, and N. Pissinou, “Sensor Network Security: A Survey,” IEEE Communications Surveys & Tutorials, vol. 11, no. 2, Second Quarter 2009.

[2] D. Liu, P. Ning, and R. Li “Establishing pairwise keys in distributed sensor networks,” ACM Journal Name, vol. , no. , 20, Pages 1–35, 2003.

[3] W. Du, J. Deng, Y.S. Han, P.K. Varshney, J. Katz, and A. Khalili, “A pairwise key pre-distribution scheme for wireless sensor networks,” ACM Trans. Inf. Syst. Secur., vol. 8, no.2, pp. 228-258, May 2005.

[4] T. Ito, H. Ohta, N. Matsuda, and T. Yoneda, “A key pre-distribution scheme for secure sensor networks using probability density function of node deployment,”SASN, pp.69-75, ACM, Oct. 2005.

[5] A. Mohaisen, and D. Nyang, “Hierarchical grid-based pairwise key predistribution scheme for wireless sensor networks,” EWSN, pp. 83-98, 2006.

[6] A. Mohaisen, Y. Maeng, and D. Nyang, “On grid-based key pre-distribution: Toward a better connectivity in wireless sensor network,” PAKDD Workshops, pp. 527-537, 2007.

[7] A. Mohaisen, D. Nyang, and T. AbuHmed, “Two-level key pool design-based random key pre-distribution in wireless sensor networks,” TIIS, vol. 2, no. 5, pp. 222-238, 2008.

[8] A. Mohaisen, D. Nyang, Y. Maeng, K. Lee, and D. Hong, "Grid-based key pre-distribution in wireless sensor networks," TIIS, vol. 3, no. 2, pp. 195-208, 2009.

[9] M. Soltani, “A New Secure Cryptography Algorithm Based on Symmetric Key Encryption,” ISSN 2090-4304, Journal of Basic and Applied Scientific Research.

[10] Charru, P. Singh, S. Rani, “Efficient Text Data Encryption System to Optimize Execution Time and Data Security,” International Journal of Advanced Research inComputer Science and Software Engineering, Volume 4, Issue 7, July 2014.

[11] K. Banerjee, H. Sharma, and B. K. Chaurasia, “Secure Communication for Cluster Based Wireless Sensor Network,” Sixth International Conference on Computational Intelligence and Communication Networks,2014, IEEE.

[12] M. M. Haque, A.-S.K. Pathan, C.S. Hong, and E. Huh, "An asymmetric key-based ecurity architecture for wireless sensor networks," TIIS, vol. 2, no. 5, PP. 265-279, 2008.

[13] W. Diffie and M. Hellman, "New Directions in Cryptography," IEEE Transactions on Information theory, vol. 22, no. 6, pp. 644-654,1976.

[14] A. Mohaisen, J. W. Choi, and D. Hong, "On the Insecurity of Asymmetric Key-based Architecture for Wireless Sensor Networks," TIIS, vol. 3, no. 4, PP. 376-384, 2009.

A Belief Rule Based Expert System to Control Traffic Signals under Uncertainty

Mohammad Shahadat Hossain
Department of Computer Science and Engineering, University of Chittagong, Bangladesh
hossain_ms@cu.ac.bd

Hoimonty Sinha
Department of Computer Science and Engineering, University of Chittagong, Bangladesh
hoimo_18@yahoo.com

Rashed Mustafa
Department of Computer Science and Engineering, University of Chittagong, Bangladesh
rashed.m@cu.ac.bd

***Abstract* -Traffic congestion is a condition on road network that occurs as vehicle increases, and is characterized by slower speeds, longer trip times, and increased vehicular queuing. It effects the economic growth of a country, increases accidents, resource cost and environment pollution. One of the most cost-effective ways, to deal with this problem is by employing traffic control signals at the road intersections. Now-a-days, most signal controls are implemented with either fixed cycle time control or dynamic control. These conventional methods for traffic signal control fails to deal efficiently because they are unable to taking account of the uncertainty associated with traffic flow. Therefore, this paper presents the design, development and application of a belief rule based expert system (BRBES) with the capability of handling uncertainty. The system uses Belief Rule Base (BRB) as the knowledge representation schema and the evidential reasoning as the inference engine. The results generated by the system have been compared with a fuzzy logic based expert system (FLBES). The BRBES's reliability is better than that of FLBS. The system applied in a number of road intersections of the Chittagong City of Bangladesh.**

Keywords---Traffic Signal Control; Expert System; Belief Rule Base; Uncertainty, Evidential Reasoning;

I. INTRODUCTION

Rapid increase of vehicles and the limited resources usually increases traffic congestion in a country. To reduce traffic congestion we can construct new roads and highways and expand the existing roads which can increase the capacity and reduce the traffic congestion. However, this is a highly expensive and time consuming process to construct new roads or expanding them. Even when construction projects are feasible, it requires temporary closing of roads, which disrupts traffic on the road network. The deployment of traffic control signals at the various intersections of the route network is considered as a candidate solution to reduce traffic congestion. Traffic Signal Control is a system for synchronizing the timing of any number of traffic signals in an area, with an aim of reducing stops and overall vehicle delay or maximizing throughput. Effective traffic signal can provide a more efficient use of existing roadway capacities as well as it can harmonize traffic flows [1].

There are two types of traffic signal controller used; one is called fixed cycle time control and the other is called dynamic control . A pre-timed controller or fixed cycle time controller repeats signal at a fixed time. Dynamic controller detects transports every time and provides signal using sensors. These methods are failed to deal efficiently with the complex, varying and uncertain changing traffic situations. They are modeled without any analysis of traffic situation.

Traffic congestion may get influenced by a variety of factors including number of vehicles, hour of the day, road capacity, entry or exit point of the area etc. [2]. These factors cannot be measured with 100% certainty during the signal control process. Analyzing the causes of the traffic congestion, we consider some major factors of traffic control system. Major factors are- Road Size, Number of Vehicles/traffic flow, Time (Busy hour), Number of Intersections.

Therefore, in order to assign time to the traffic signals whether to the green, blue or red, this uncertainty phenomenon associated with the factors should need to be considered. Since this is an example of complex problem involving uncertainty, algorithmic solution can't be adopted. An expert system can be considered as an appropriate solution in this regard.

The remaining of the paper is structured as follows. Section two presents the related works. Section three provides an overview of RIMER methodology. Section four presents the architecture, design and implementation of the proposed belief rule based expert system (BRBES). Section five presents the results and discussions. Section six concludes the paper.

II. LITERATURE REVIEW

Many techniques have been used for traffic signal control system. However, these systems have a high equipment cost and their accuracy depends on environment conditions. The first attempt made to design Fuzzy Traffic Controller was in 70s by Pappis and Mamdani [3]. There exists a number of fuzzy logic based traffic signal control systems. After that Niittymaki, Kikuchi, Chui and other researchers [4] developed different algorithms and logic controllers to normalize traffic flow. Kelsey and Bisset [5] also designed a simulator for signal controlling of an isolated intersection with one lane. They observed that Fuzzy Controller reduces the vehicle delay when traffic volume was heavy. Niittymaki and Kikuchi [4] developed Fuzzy based algorithm for pedestrians, crossing the road. Nakatsuyama, Nagahashi and

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

Nishizuka [6] have applied fuzzy logic to control two adjacent intersections on an arterial with one-way movements. In recent years, Lin Zhang and Honglong Li [7] also worked on designing Fuzzy Traffic Controller for oversaturated intersections. The developed acyclic real-time control model consists of two components including: an Improved Genetic Algorithm (IGA)-based signal optimization module and a microscopic traffic simulation module. The IGA-based signal optimization module is designed to optimize the phase sequence and phase length with the aim to minimize the delay of both transit and general vehicles. Although, these systems can handle uncertainties due to ambiguity, vagueness or imprecision, but they are unable address other types of uncertainty such as ignorance, incompleteness, ignorance in fuzziness that may exist with the factors associated with traffic control system.

Belief rule-base is a new knowledge representation schema uses a belief structure where belief degrees are embedded with all possible terms of the consequent part of a rule [8]. It can handle various types of uncertainties such as vagueness, ambiguity, imprecision, ignorance and incompleteness. Therefore, the system presented in this paper considered BRB to develop knowledge base and the ER as the inference methodology; combination of BRB and ER, known as RIMER.

III. OVERVIEW OF RIMER

Recently, a generic rule-based inference methodology using the evidential reasoning approach (RIMER) has been proposed [9], which provides a flexible and effective framework to represent not only precise data but also vagueness and ignorance in knowledge, as well as a rigorous inference procedure to deal with such complicated uncertain information. This section presents the building blocks of RIMER, consisting of knowledge representation schema as well as inference procedures. The inference procedures consists of input transformation, rule activation weight calculation, belief update and rule aggregation. The methodology will be introduced by taking account of the traffic signal control problems addressed in this paper.

A. *Knowledge Representation*

Belief rule base (BRB) which is an extension of IF-THEN rule is used to represent uncertain knowledge. Uncertainty due to incompleteness is addressed in the BRB. Below is an example of BRB.

R_k: $if\,(S_{k1}$ is $A_1^k) \wedge (S_{k2}$ is $A_2^k) \wedge \ \ldots\ \wedge (S_{kT_k}$ is $A_{T_k}^k)$,
$then\{(D_1\ is\ \beta_{1k}), (D_2\ is\ \beta_{2k}), \ldots, (D_N\ is\ \beta_{Nk})\}$
….. (1a)

IF (Road Size is Medium) and (Vehicle number is High) and (Busy hour is Yes) and (Intersection is Medium),THEN Traffic Signal is {(Green, 0.7), (Yellow, 0.1), Red, 0.2)} ……. (1b)

B. *Inference Procedures*

Input transformation, rule activation weight calculation, belief update and rule aggregations are the components of inference procedures. These will be discussed with example.

Input Transformation and Rule Activation Weight Calculation

Input transformation is concerned with the distribution of an input value of an antecedent attribute over its different referential values as shown in Eq. (1...5).

$(U_i) = \{(A_{ij}); j = 1 \ldots \ldots \ldots J_i\}$ …… (1)
$S(U_i, \varepsilon_i) = \{(A_{ij}, \alpha_{ij}); j = 1, \ldots \ldots \ldots \ldots J_i\}$ i = 1 …T ….(2)
$S(A_i^*) = \{(h_{3i}, \gamma_{3i}), (h_{2i}, \gamma_{2i})\}$ ……. (3)
$\gamma_{2i} = (h_{3i} - A_i)/(h_{3i} - h_{2i})$ ….. (4)
$\gamma_{3i} = 1 - \gamma_{2i}$ .. (5)

Belief Update

There could be the case that input value of any one of the factors associated with traffic control signal system may not be available.

Hence, the degree of belief associated with each of the referential value of the consequent attribute should need to be updated by using Eq.6

$$\beta_{ik} = \bar{\beta}_{ik} \frac{\sum_{t=1}^{T_k} (\tau(t,k). \sum_{j=1}^{J_i} \alpha_{tj})}{\sum_{t=1}^{T_k} \tau(t,k)} \ldots\ldots\ldots\ldots(6)$$

Rule Aggregation

The final output of the system is to obtain the degree of belief associated with the referential value of the consequent attribute.

$$\beta_j = \frac{\mu \times \left[\prod_{k=1}^{L}\left(\omega_k \beta_{jk} + 1 - \omega_k \sum_{j=1}^{N} \beta_{jk}\right) - \prod_{k=1}^{L}\left(1 - \omega_k \sum_{j=1}^{N} \beta_{jk}\right)\right]}{1 - \mu \times \left[\prod_{k=1}^{L} 1 - \omega_k\right]} \ldots..(7)$$

IV. BRBES TO CONTROL TRAFFIC SIGNALS

This section presents the architecture, design and implementation of the Belief Rule Based Expert System. In addition, the procedure of knowledgebase construction along with the system interface is also discussed.

A. *Architecture, Design and Implementation of BRBES*

The proposed belief rule based architectural framework to develop traffic signal control system adopts three-layer architectural style as shown in Fig 1.

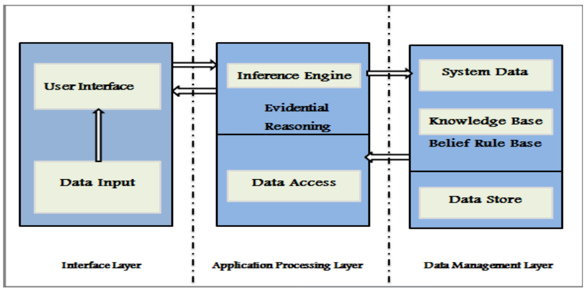


Fig. 1: BRBES Architecture

B. *Knowledge base Construction*

In order to construct the knowledge base for the BRBES, a BRB framework is required as shown Fig 2. Fig 2 has been developed by taking account of the factors associated with traffic control signal.

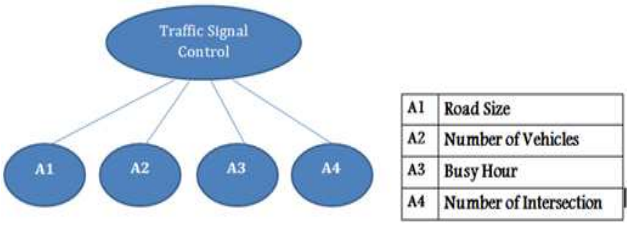


Fig. 2: The BRB Framework to Control Traffic Signal

Table 1shows the initial rule base of the BRBES, which consists of 54 rules according to RIMER method.

TABLE 1: INITIAL BELIEF RULE BASE

Rule ID	Rule Weight	IF (Antecedent)				THEN (Consequent)		
		A1	A2	A3	A4	Green	Yellow	Red
1	1	Large	Heavy	Yes	High	1	0	0
2	1	Large	Heavy	Yes	Medium	0.8	0.2	0
3	1	Large	Heavy	Yes	Low	0.8	0	0.2
4	1	Large	Heavy	No	High	0.6	0.4	0
	..	..	..	..	..	..	..	..
54	1	Small	Low	No	Low	0	0.2	0.8

An example of a belief rule taken from Table 1 is illustrated below.

R52: IF Road Size is Small AND Vehicles is High AND Busy Hours is NO AND Number of Intersections is High THEN Traffic Signal Control{(Green,0),(Yellow,0.02),(Red,0.8)}.

In the above belief rule, the belief degrees are attached to the three referential values.

C. *BRBES Interface*

The user interface (UI) is the visual part of computer application or system through which a user interacts with a computer or software. Fig 3 illustrates the interface of the BRBES.

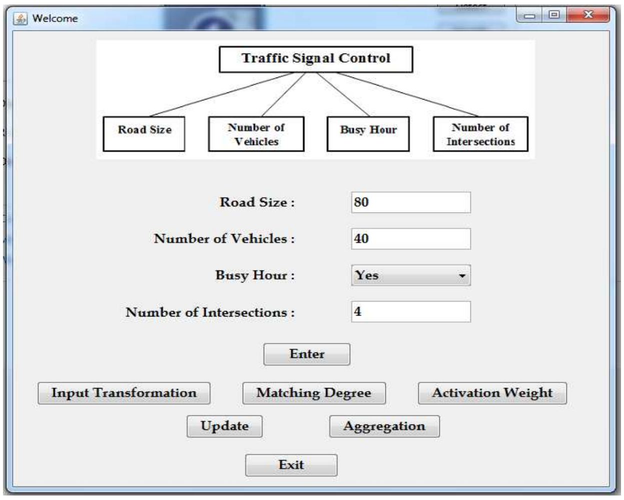


Fig. 3: BRBES's Input Interface

Having received the input values while this functions run, the consequent attribute's referential values corresponding belief degrees, which are green, yellow and red, are displayed in other window.

V. RESULT AND DISCUSSION

The BRBES presented in this paper has been applied at one of the important route intersections of Chittagong City, Bangladesh, known as Muradpur. Fig.4 shows the picture of this intersection. The intersection consists of four roads, each with double line. The four traffic factors data including road size (A1), number of vehicles (A2), busy hour (A3) and number of intersections (A4) have been collected as shown in Table 2. Table 2 shows one set of traffic data related to four roads of the intersection. However, in order to test the reliability of the BRBES's output in comparison to expert opinion as well as fuzzy rule based expert system (FRBES) fifty set of data of this intersection have been collected in different period of a month.

Fig. 4: Muradpur Intersection, Chittagong City, Bangladesh

It can be observed that the green signal of the road number 4 should be given more priority than from the other roads since its degree of belief is higher. Hence, according to the BRBES's generated results at first, signal of the road number 4 of the intersection will be green, then the road number 3, then the road number 2 and then road number 1 will be made green. In this way, by using the BEBES the traffic congestion at the intersection can be managed.

TABLE 2: DATA FOR A CASE STUDY

Road	A1	A2	A3	A4	Green	Yellow	Red
Road 1	62	39	Yes	4	46.706%	51.034 %	2.258 %
Road 2	40	19	Yes	4	47.399%	40.179 %	12.421 %
Road 3	74	41	Yes	4	53.097%	44.86%	2.041 %
Road 4	81	59	Yes	4	60.852%	39.147 %	0.0%

A fuzzy rule based expert system (FRBES) developed in MATLAB environment by using the same traffic factors and the results generated by this system using the same data are illustrated.

Therefore, it has been considered in this research to test the accuracy of the BRBES' output against expert opinion and FRBES's output by taking account of benchmark data. If the expert's opinion or perception on the level traffic is greater than 50%, then outcome is considered as one otherwise zero and this data have been considered as the baseline.

Fig. 5 shows the ROC curve for the above data. It shows three ROC curves. The ROC curve is represented by the blue, green, gray line. The blue line represents BRB expert system, green line represents fuzzy based system and gray line represents expert's opinion. AUC of BRBES is 0.926 (95% confidence intervals (0.853 – .999) FRBES's AUC is .859 (95% confidence intervals 0.742 – .976) and the experts AUC is 0.93 (95% confidence intervals 0.851 – .999). From the AUC, it can be observed that the AUC of BRBES is greater than that of expert opinion and FRBES. This implies that the result generated by the BRBES is better and reliable.

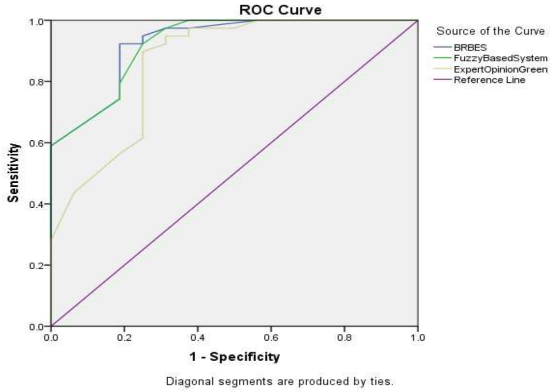


Fig. 5: ROC curve

VI. CONCLUSION

The design, development and application of a belief rule based expert system (BRBES) is presented in this paper. Belief rule base as the knowledge representation schema has been used to develop the knowledge base of the system since it has the capability to handle various types of uncertainty that exist with the traffic data. Evidential reasoning has been used as the inference engine, which is able to aggregate the rules under uncertainty to generate degree of belief of the referential values of the consequent attribute which are green, yellow and blue in traffic signal control. It has been demonstrated that the BRBES presented in this paper is more reliable and better than from fuzzy based expert system and expert opinion. Hence, the system can be applied to control traffic signal of a city as we have demonstrated in our case study.

REFERENCES

[1] M. G. Smadi, (2001) A Knowledge-Based Traffic Signal Control Application, Graduate Faculty of the North Dakota State University of Agriculture and Applied Science.

[2] P. Chan (2011), Advanced Transportation Management Technologies.: Pablo Chan. Online] Available: http://www.scribd.com/doc/47702578/Advance-Transportation-Management-Technologies-Abridged#scribd

[3] Pappis, C. P., and Mamdani, E. H. (2007), *A Fuzzy Logic Controller for a Traffic Junction*. IEEE Systems, Man, and Cybernetics Society, Vol 7, No. 9

[4] Chiu, S. (1993). Adaptive Traffic Signal Control Using Fuzzy Logic. *IEEE International Conference on Fuzzy Logic.* San Francisco, CA

[5] Robert L. K., and Keith R.B. (1993), *Simulation of Traffic Flow and Control Using Fuzzy and Conventional Methods*.: Prentice-Hall, Inc

[6] Nakatsuyama, M., Nagahashi, H., and Nishizuka, N. (1984). *Fuzzy Logic Phase Controller for Traffic Junctions in the One-Way Arterial Road.* In Proc IFAC 9th Triennial World Congress Pergamon Press.

[7] Zhang, L., Li, H., and Prevedouros, P. D. (2008), Signal Control for Oversaturateandd Intersections Using Fuzzy Logic. *Transportation and Development Innovative Best Practices*, ASCE.

[8] Rahman, S. and **Hossain, M.S.,** (2013) "A Belief Rule Based Clinical Decision Support System to Assess Suspicion of Heart Failure from Signs, Symptoms and Risk Factors" *IEEE Technically Co-Sponsored 2nd International Conference on Informatics, Electronics & Vision*, May, 2013, Dhaka, Bangladesh, Print ISBN: 978-1-4799-0397-9.

[9] Hossain, M.S., Zander, PO., Kamal, S. and Chowdhury, L., (2015), *Belief Rule Based Expert Systems to Evaluate E-Government: A Case Study*, Expert Systems: The Journal of Knowledge Engineering, Vol. 32, No.5Jhon Wiley & Sons Ltd.,

Review of integrated applications with AIML based chatbot

Md. Shahriare Satu
Institute of Information Technology
Jahangirnagar University
Dhaka, Bangladesh
Email: shahriar.setu@gmail.com

Md. Hasnat Parvez
Institute of Information Technology
Jahangirnagar University
Bangladesh
Email: hasnatiit847@juniv.edu

Shamim-Al-Mamun
Institute of Information Technology
Jahangirnagar University
Dhaka, Bangladesh
Email: sam@juniv.edu

***Abstract*—Artificial Intelligence Markup Language (AIML) is derived from Extensible Markup Language (XML) which is used to build up conversational agent (chatbot) artificially. There are developed a lot of works to make conversational agent. But low cost, configuration and availability make possible to use it in various applications. In this paper, we give a brief review of some applications which are used AIML chatbot for their conversational service. These applications are related to cultural heritage, e-learning, e-government, web base model, dialog model, semantic analysis framework, interaction framework, humorist expert, network management, adaptive modular architecture as well. In this case, they are not only providing useful services but also interact with customers and give solution of their problems through AIML chatbot instead of human beings. So, this is popular day by day with entrepreneur and users to provide efficient service.**

Keywords*—*AIML, chatbot, Alice

I. Introduction

A chatbot is a conversational agent where a computer program designed to simulate an intelligent conversation [1]. There are invented various kinds of conversational agent. But from those agents, AIML (Artificial Intelligence Markup Language) based chatbot are most popular because they are lightweight and easy to configure. ALICE is the most popular AIML based chatbot which won Leobner Price three times (2000, 2001, and 2004). So, now a day, various kinds of organizations are interested to implements AIML based chatbot to get conversation with customers with minimum configuration and cost. In this paper, we focus on several applications whose implements AIML based chatbot with additional software packages to develop efficient applications.

In this paper, section I is introduced about topic and section II is defined about Artificial Intelligence Markup Language (AIML) and categories and pattern matching schemes of it. Then, section III is given some description about integrated system whose are used AIML based chatbot in their application. Finally, section IV is drawn conclusion of this conversational agent with future scope.

II. Artificial Intelligence Markup Language

Artificial Intelligence Markup Language (AIML) is derivation of Extensible Markup Language (XML). It has class of data object called an AIML object which describes the behavior of computer programs. It contains of units called topics and categories. Categories are basic unit of knowledge in AIML. Each category consists of pattern which contains input and template which contain answer of chatbot. Besides, there are containing some optional context called "that" and "topic". <that>contain chatbot last utterance and <topic>contain a collection of categories together.

AIML consists of words, underscore symbols and wildcard symbol like _and *. It is also case invariant. There are three type of categories. [2]

- **Atomic categories:** These categories are those patterns whose have no wildcards.
- **Default categories:** These categories are those patterns whose have some wildcards _like *. They have to match any input but differ with their alphanumerical order.
- **Recursive categories:** These categories are whose having <srai>and <sr>tag which simply refer to recursive and symbolic reduction.
  1) **Symbolic Reduction:** It reduces complex grammatical forms to simpler ones.
  2) **Divide and Conquer:** It split an input into two or more subpart and combines the response to each.
  3) **Synonyms:** It return similar answer in a pattern for nearest user pattern of it.

A. Preparation of Pattern Matching

Before starting pattern matching algorithm, each input to the AIML interpreter must pass through two processes.

- **Normalization Process:** It is involved 3 phases. Substitution Normalization is a heuristic applied for an input which tries to retain information in the input. Otherwise it would be lost during sentence splitting or pattern fitting normalization. In Sentence splitting Normalization, The sentence is split into two or more sentences using "break sentence at periods" rules. Then, In Pattern fitting Normalization, It removes punctuation from the input and convert it uppercase.

- **Producing Input Path:**

  Normalized-Input<that>Tvalue <topic>Pvalue.

  Input path has three parts. First part is called Normalized Input. Then, <that>tag followed by Tvalue that holds previous root answer and <topic>tag followed by Pvalue that holds topic name if it exists _or * otherwise.

B. Pattern Matching Algorithm

AIML interpreter tries to match word by word to gain longest pattern matching and try to find which the best one is. This behavior can be described with Graphmaster set of files and directories containing a set of nodes which is called nodemaster and branches represents first words of all patterns and wildcard symbols.

III. Integrated Applications with AIML based Chatbot

There are several works which are implemented AIML chatbot with additional package.

- **Multimodal Pervasive Virtual Guide for cultural heritage sites:** [3] Core of the system is a conversational agent which is able to talk to tourist in natural languages. It is based on two main parts: chatbot module and multimodal interface.

  First, chatbot technology is implemented through Alice Technology and it is linked with OpenCyc Technology for common sense ontology knowledge base. Cyc organizes concept and assertions into microtheories into new theory. Besides, chatbot can interact ontology through JAVA API which helps chatbot more fluent dialogue.

  Next, multimodal interface composed of two submodules. First submodule deals with vocal interaction. It has been implemented a new markup language called XHTML+Voice. Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) technologies are embedded into multimodal browser. Speech process is carried out through an ad-hoc built in grammar according to W3C Speech Recognition Grammar specification. Second submodule is RFID module which is estimated PDA position within the environment. To hold data into a device, there are using a device called RFID tags. RFID tags categorizes into active and passive. This is used passive RFID tags which operate without external power resource and each having unique ID.

- **Natural language management for e-learning platform:** [4] It is an integrated platform consist of with a basic AIML knowledge, PABX (Private Automated Branch eXchange) for VoIP phone management and Servers Control daemon.

  In this case, chatbot is called Tutorbot because it is functionality backing of didactics done in e-learning environments. It relies its main features on a description of knowledge in AIML and contains some features like natural language management, presentation of contents, and interaction with search engine. Besides, the whole e-learning platforms working is linked to indispensable services in particular to web service.

  It has been created continuous monitoring service on e-learning platform servers which is another controlling machine called Daemon.

  If there are any anomalies on the server that contain e-learning platform, a demand of emergency call to a fast intervention of technical support that manages PABX. The software of PABX is chosen by Asterisk which allows the management regards some essential practices to our aim.

- **Integrated E-Government System:** [1] This system is utilized the power of chatterbots to serve interactive support system in enterprise applications. It is showcased a reference model to utilize GIS in E-Governance system and consist of core Land Record Maintenance and Extension module. It added ASE (Artificial Support Entity) module as a External Extension that communicates with core IEGS (Integrated E-Government System) using SOA protocols.

  The chatterbot is composed of a set of AIML files which is a stimulus-response model. It is based on Program E and uses PHP as front end stores AIML patterns in MYSQL database and AJAX uses UI application. It consist of Bot User Interface, an interface and AIML database. It also tries to mimic real human conversations using Persona-AIML architecture that consists of Categories Base, Personality Component, Dialogue Log, Reasoning Component.

  IEGS has data outlet modules which work as data extractors that match parameters and return data related objects. Implementation ASE to IEGS requires program E front-end to query IEGS handler using URL parameters. User queries are Formal queries and Business queries.Program E front end utilizes to identify type of queries and decides course of action from classification.IEGS also utilize powerful Google maps API to make location to coordinate and dynamically generate javascripts which drives Google maps objects for extra functionality.

- **Intelligent educational system:** [5] Intelligent educational system (INES) is a functional prototype of an online learning platform, which combines three essential capabilities, Learning Management System (LMS), Learning Content Management System (LCMS),Intelligent Tutoring System (ITS). It consists of different tools and technologies. CHARLIE(CHAtteR Learning Interface Entity) is AIML based intelligent chatterbot and based on program D and uses AJAX technology to maintain Asynchronous communication which enables interactive application or RIA. Intelligent agent based on BDI (Believes, Desires, and Intentions) which acts as brain of the system. Besides, an Inference engine based on JESS (a rule engine based on Java platform) and decides what is allowed to do. At last, Ontology is a part of INES which semantically defines contents of the courses. Semantic managing users, contents tools are also used in this system.

- **Web-based voice chatbot:** [6] This Web-based system is used intelligent voice recognition capabilities and black box approaches. It controls communication structure with web service and consists of three components: client, server, content acquisition.
  The server is a simple object access protocol (SOAP) aware internet application which isolates client from interacting inner workings. All messages are formatted in an XML and encapsulated in SOAP text based message pack.
  The client contains voice recognition processing module. This application can easily accessible by client browser and user is allowed to register and login to the system.
  When the user logs the system, the system will greet the user by default and prepare to receive question or statements. ALICE bot engine is used to fulfill this purposes which contain a set of AIML files. To assist the training, Training modules, Third party expert system, Ultimate Research Assistant were used but to refine query self-processing AIML-assisted categories are used.
  Besides, there are implemented client applet to provide better service. It consists of two options: text-input, voice-input. It also requires a number of libraries for processing voice inputs and secured communication with web service. In applet management process, there are several components which are responsible for client interruption. Control unit of applet can activate to event to trigger either interface or voice recognition module or SOAP communication module.
  So, the client send and receive plain text. Web service processes received queries all response generation modules.

- **Framework for Agent-based Semantic-aware Interoperability:** [7] Framework for Agent-based Semantic-aware Interoperability (FRASI) is streamlining of the chatbot knowledge base and leading to sentence analysis process. This question-answering system uses pattern matching chatbot technology. It consists of atomic, default, ontology based, and ultimate default. But this contains some limitations. To reduce limitations, it proposed three modules to use with chatbot.
  Enhanced Symbolic reduction (ESRAI) preprocesses sentences and reduced them into simpler structure.
  A second module, named bootstrapping module can exploits ontologies in order to automatically extend Chatbot KB.
  Third module is CYD module which makes it possible to use ontologies to make dynamically answers generated by inference processing of user question.

- **Automatically Extracting Dialog Models:** [8] This system is works as task-oriented natural language dialog system which is interact with user to accomplish a given objective. In this case, there is proposed unsupervised, apriori like algorithm which extracts subtasks and their valid orderings. In order to extract subtask from a corpus, this model is responsible to do this.
  First, it is needed to cluster and canonicalize similar utterance semantically to replace all the name entity by their type and generate a feature vector for each utterance. Feature vectors are clustered to group together semantically with similar utterance. This process is called Utterance Normalization.
  Second, there are discovered subtask form normalization calls then finding precondition (utterance pre condition, flow pre condition) for each subtask. This process is called mining subtask from call transcripts.
  Third, Extracted task structure is encoded as an AIML construct to build task-oriented chatbot. This process is called AIML Generation.

- **Multimodal Interaction Framework with Embodied Contextual Understanding:** [9] This framework is based on simulation platform named SIGVerse. It is divided into Server System, Client System, Service providers.
  In Server System, SIGServer is the center of simulations and it runs core application and agent controller is works as agent intelligence.
  Human user interfaces with simulation is happened via SIGViewer and HRI interaction is happened via various user interfaces in the client system.
  Then, in Service Providers, it offers additional features. In this case, Dialogue engine and behavior recognition module are developed and facial expression, speech emotion is under development. Besides, Dialogue Engine utilizes AIML and behavior recognition module is used Mimesis model with Hidden Markov Model.

- **Humorist Chatbot System:** [10] Humorist bot is conversational agent which is provided sense of humor by telling humorous anecdote to user and capable to listening jokes, trying to understand humorous level. It has a set of standard Alice categories which allow holding a general conversation with user. It has also another set of categories which aimed at humorist sentence generation. Next, it contains a set of Chatbot to recognize humorous intent with external resources and CMU pronunciation dictionary. This agent is mainly tried to humorous interact with user.

- **Dorothy Network Management chatbot:** [11] Dorothy chatbot is not only interacts with user but also manage network via management protocols. It is used based on ALICE solution. ALICE integrated with Dorothy by generating 5 parameters which are topic, address, operation, time and situation. This system consists of three modules: central module, PBO module and network history information module. Central module is responsible to receive all information generated by ALICE. Network history information module consists of collect module, history information base. Collect module collect information of network and history information database. History information base is used to collect measurement of network between intervals. PBO module is based on case-based reasoning and problem solving.

- **Sub-symbolic layer in Intuitive chatbot:** [12] This is combined Latent Semantic Analysis with common sense and traditional knowledge representation to improve ialogue capabilities. In this case, agent area is consist of two area: rational area, associative area. Rational area composed of AIML KB and CYC ontology. On the other hand, asssociative area is obtained with mapping CYC concept as vectors into semantic space built by means of Latent Semantic Analysis. Given a specific microtheory which is collection of concepts and facts of particular domain and semantic space inferred from corpus text. Corpus is created both ad hoc extracted pages from Wikipedia repository with specific CYC microtheory. Ecah concept projected on the space and reciprocal geometric distance between concepts defines sub-symbolic relationship net which is seen as sub symbbolic semantic layer automatically added CYC ontology. So, chatbot can interact with user exploiting its standard KB and CYC ontology but it can use of its assertive reasoning area to retrive semantic concepts between ontological concepts that are easily reachable through assoicative sub-symbolic paths.
- **Adaptive Modular Architecture based chatbot:** [13] This conversational agent is based on modular knowledge representation and proof-of-concept prototype. It is hybrid model and consist of different knowledge representation and reasoning capabilities which composed of different components: dialogue engine, dialogue analyzer, corpus callosum. The goal of dialogue engine to split traditional monolithic knowledge base into different components called modules. Each module of the dialogue engine has its own inference engine and knowledge base composed of Alice technology. Besides, The framework is general purpose. Dialogue analyzer is capable to capture particular feature related to context of dialogue and manage them as variables. There are mainly four kinds of context variables: topic, speech act, user, goal. Corpus callosum is used as module selector and a planner. Modular selector can enable or disable dialogue module to select more appropriate module. Besides, a planner uses a specific context variables in order to define temporal evolution of each module.

IV. Conclusion

Integrated Systems are played important role in various field of client service. It is very difficult for a visitor to understand the system easily without any guidelines. In this case, chatbots give these additional service to the viewer.

In this paper, we illustrate some integrated systems which are added AIML based chatbot to their system to make interaction with user. Various different API and package with lightweight AIML files make this system more flexible and interactive to use in various field. Besides, automated conversational agent based system is played a significant role to interact with user.So, service provider has spend less cost to provide automated conversational agent instead of human conversational agent. Besides, user get unlimited full time chatting service which make interest user to use this particular service. So, The integration of this chatbot service give service provider extra business facilities in their particular service. In this case, it has contain a great business field to manage customer using machine intelligence with more flexible and efficient ways.

References

[1] R. Mahapatra, N. Sharma, A. Trivedi, and C. Aman, "Adding interactive interface to e-government systems using aiml based chatterbots," in *Software Engineering (CONSEG), 2012 CSI Sixth International Conference on.* IEEE, 2012, pp. 1–6.

[2] B. A. Shawar and E. Atwell, "Using dialogue corpora to train a chatbot," in *Proceedings of the Corpus Linguistics 2003 conference*, 2003, pp. 681–690.

[3] A. Augello, A. Santangelo, S. Sorce, G. Pilato, A. Gentile, A. Genco, and S. Gaglio, "A multimodal interaction guide for pervasive services access," in *Pervasive Services, IEEE International Conference on.* IEEE, 2007, pp. 250–256.

[4] D. De Orlando and F. Giovanni, "An integrated system, with natural language management, for the monitoring activities in e-learning environments," in *Complex, Intelligent and Software Intensive Systems, 2008. CISIS 2008. International Conference on.* IEEE, 2008, pp. 965–970.

[5] F. Mikic, J. C. Burguillo, M. Llamas, D. Rodríguez, E. Rodríguez *et al.*, "Charlie: An aiml-based chatterbot which works as an interface among ines and humans," in *EAEEIE Annual Conference, 2009.* IEEE, 2009, pp. 1–6.

[6] S. du Preez, M. Lall, and S. Sinha, "An intelligent web-based voice chat bot," in *EUROCON 2009, EUROCON'09. IEEE.* IEEE, 2009, pp. 386–391.

[7] A. Augello, G. Pilato, A. Machi, and S. Gaglio, "An approach to enhance chatbot semantic power and maintainability: Experiences within the frasi project," in *Semantic Computing (ICSC), 2012 IEEE Sixth International Conference on.* IEEE, 2012, pp. 186–193.

[8] S. Negi, S. Joshi, A. Chalamalla, and L. V. Subramaniam, "Automatically extracting dialog models from conversation transcripts," in *Data Mining, 2009. ICDM'09. Ninth IEEE International Conference on.* IEEE, 2009, pp. 890–895.

[9] J. T. C. Tan and T. Inamura, "Extending chatterbot system into multimodal interaction framework with embodied contextual understanding," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction.* ACM, 2012, pp. 251–252.

[10] A. Augello, G. Saccone, S. Gaglio, and G. Pilat, "Humorist bot: Bringing computational humour in a chat-bot system," in *Complex, Intelligent and Software Intensive Systems, 2008. CISIS 2008. International Conference on.* IEEE, 2008, pp. 703–708.

[11] M. D. Leonhardt, L. Tarouco, R. M. Vicari, E. R. Santos, and M. D. S. d. Silva, "Using chatbots for network management training through problem-based oriented education," in *Advanced Learning Technologies, 2007. ICALT 2007. Seventh IEEE International Conference on.* IEEE, 2007, pp. 845–847.

[12] G. Pilato, A. Augello, G. Vassallo, and S. Gaglio, "Sub-symbolic semantic layer in cyc for intuitive chat-bots," in *Semantic Computing, 2007. ICSC 2007. International Conference on.* IEEE, 2007, pp. 121–128.

[13] G. Pilato, A. Augello, and S. Gaglio, "A modular architecture for adaptive chatbots," in *Semantic Computing (ICSC), 2011 Fifth IEEE International Conference on.* IEEE, 2011, pp. 177–180.

Character Recogntion System: Performance Comparison of Neural Networks and Genetic Algorithm

Md. Shahazan Ali
Dept. of Computer Science & Engineering (CSE)
Rajshahi University of Engineering Technology
Rajshahi, Bangladesh
kernel.writer@gmail.com

Dr. Md. Nazrul Islam Mondal
Professor, Dept. of Computer Science & Engineering (CSE)
Rajshahi University of Engineering Technology
Rajshahi, Bangladesh
nimbd@yahoo.com

***Abstract*— This thesis underlying the thesis title "Character Recognition System: Performance Comparison of Artificial Neural Networks and Genetic Algorithm" represents the character recognition using the Artificial Neural Networks (ANN) and Genetic Algorithm (GA) and measurement of various performance by changing various selection criteria of both algorithm. We used Backpropagation Learning Neural Network Algorithm (BPN) as the ANN. This system has been taken the character's image as its input. The input images have been filtered by filtering methods of image processing to remove noise and smoothing it and converted to binary image to detect its edges properly and clipped to get the actual image to input to the system. Features of each individual clipped images have been extracted by taking a definite resize binary image value. These extracted features of input images of characters is used by the Backpropagation Learning Neural Network Algorithm and Genetic Algorithm. Thus the network has been trained and creates a knowledge base of recognition. The same procedures have been applied for recognition but with only difference is that the neural network is used the previously learned weights and thresholds to calculate the output for BPN. In this work the performances of changing the various selection criteria of both algorithms have been also measured to learn and recognize the character.**

Keywords—Character; Recognition; Character Recognition System; ANN; GA; BPN; Pattern; Learning; Generation; Population; Selection; Mutation; Crossover; Chromosome; Survivial of Fitness;

I. INTRODUCTION

It is often useful to have a machine perform recognition. In particular, machines that can read symbols are very cost effective. A machine that reads banking checks can process many more checks than a human being in the same time. This kind of application saves time and money, and eliminates the requirement that a human performs such a repetitive work. It has opened a new era for the mankind to enter into a new world, commonly known as the technical world. It has been a part of everyday life. It is a long cherished dream of mankind to create intelligence outside the human body. As a result of the emergence of the character recognition system, the computer is able to classify the character as human can identify them.

Character recognition means matching a pattern as part of a defined character set of defined language. In present time people used to convert paper copy to soft copy of data to better implementation of storage device. And some machine needs to detect character sequence. Total work of character detection is divided into two parts: one, low and mid-level processing of a give character and two, recognizing the pattern. First part contents image acquisition, sampling, segmentation and key generation and second part is image recognizing by specific computation system which may be three types: neural, non-neural and statistical. Accuracy depends on choosing and implementing both efficient low and mid-level processing and efficient recognition system.

In the character Recognition System multiple preprocessing and processing states have been employed. The complete System can be described in two stages, one is learning stage and other is testing stage.

II. ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks are relatively crude electronic models based on the neural structure of the brain. These neural networks are used to implement the behavior of the biological neurons artificially in the computer.

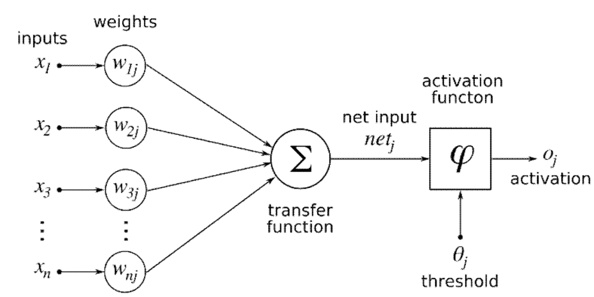


Fig. 1. A Neuron

These networks also learn from the some trained values as human and other animals learn from example and previous experience. If an ANN is trained, it can be used to recognize an input.

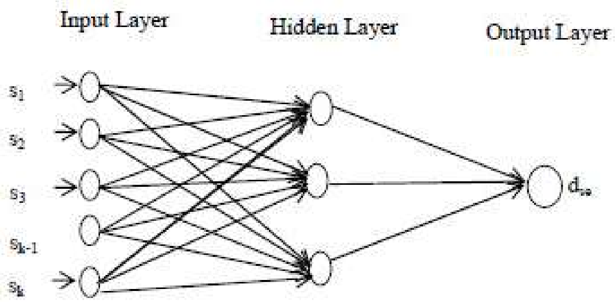


Fig. 2. A Neural Network

A simple model as shown in figure can capture the behavior of a neuron. Every component of the model bears a direct analogy to the actual constituents of a biological neuron and hence is said artificial neuron. Here x1, x2, ..., xn are the inputs to the artificial neuron and w1, w2, …, wn are weights attached to the input links. Hence the total input I, received by the soma of the artificial neuron is, I = w1x1 + w2x2+ … + wnxn. To generate the final output y, the sum is passed to a nonlinear filter α called activation function or transfer function which releases the output. Some of the neurons interfaces with the real world to take input, some does processing and some interface with the real word to produce output. This output might be the particular character that the network can think that has scanned or the particular images it thinks it being viewed. As shown in fig 1 and fig 2, all the neurons connected to process the algorithm is called neural networks.

III. GENETIC ALGORITHM

GA (Genetic Algorithm) is an optimization and search techniques based on the principles of Genetics and Natural Selection. Natural selection always tends to pick the fittest individuals dominating over the weaker ones and it always favors the positive adaptation resulting into the best one to survive in the long run. GA is a part of the evolutionary optimizing computing inspired by Darwin's survival of the fittest. GA is an adaptive search heuristics that mimics the process of natural evolution which uses techniques like Selection, Crossover, Inheritance and Mutation. The basic working flow of GA is that it starts with a set of solution to a problem (initial population) and then applies the selection, crossover, mutation and its other operators to produce another set of population. The algorithm uses 'survival of fittest' technique to choose individuals from the input population. It repeats the same procedure over and over (like human generation) until a certain criteria meets.

IV. IMPLEMENTATION OF CHARACTER RECOGNITION SYSTEM

There are three parts in the recognition system. For the first part, the image processing and extracting patterns/features from it, we used Matlab software by Matwork. The extracted features/patterns of the character images is stored in a file. The second part of the system, learning phase, is to feed those extracted features in to BPN and GA algorithms to learn the extracted patterns. Feature Extraction: In the image processing part i) images are acquired/collected and saved to disk, ii) convert it to a grayscale image, iii) remove noises by median filtering, iv) detect edges by Kenny algorithm, v) detects boundary of the edges and resize the image to minimum size so that it holds the whole character. As BPN works with same size of images and for faster learning, we used 15x15 images after boundary detection. So we need 225 input nodes in BPN network.

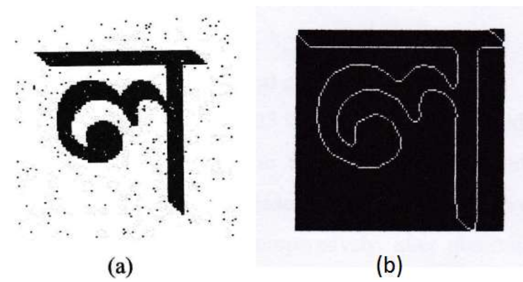


Fig. 3. (a) is the input to image processing part and (b) is the output after boundary detection

A. Learning in BPN

We used 225 input layer nodes, 10 hidden layer nodes and 5 output layers node. We used sigmoidal gain of 0.09, learning coefficient of 0.15 both for hidden and output layer with error tolerance rate 0.05 for faster and efficient learning after studying of those parameter changes in the BPN algorithm.

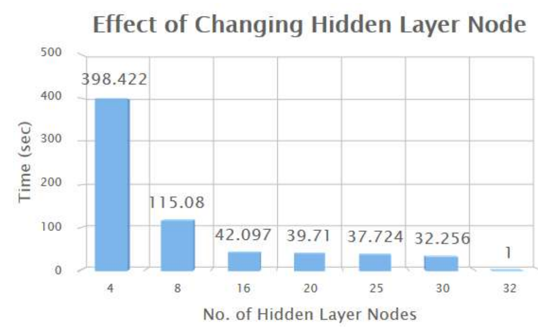


Fig. 4. Effect of changing sigmoidal function value while keeping others constants

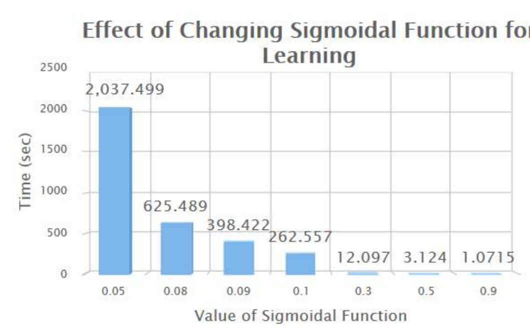


Fig. 5. Effect of changing number of hidden layer nodes while keeping others constants

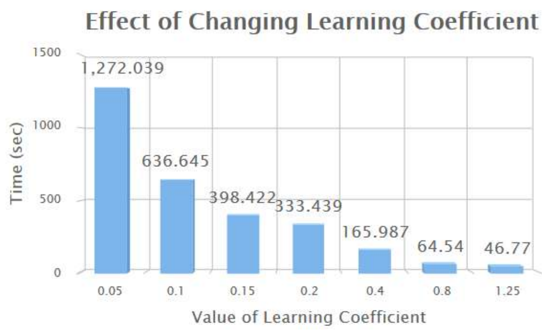


Fig. 6. Effect of changing learning coefficient value while keeping others constants

B. Learning in GA

In GA, the final processed image's binary matrix is used as the chromosome. We applied crossover and mutation operation.

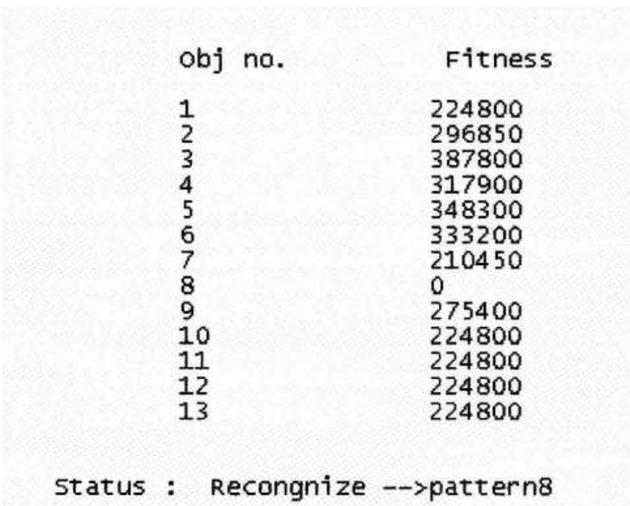


Fig. 7. A sample output of the GA program we wrote

For selection, all the individuals is copied to next generation as we have very new number of chromosomes. At the end of every generation the absolute difference of corresponding pixels is summed up and this is added with the fitness of the previous generation. The fitness at generation zero is zero. This fitness value is used to find out the character that better match with the given unknown by comparing this minimum fitness value with a predefined threshold value. If the minimum fitness value is greater than that threshold value, the character is considered as unknown; otherwise it is considered as known.

V. Analysis and Comparison of the Algorithms

In this character verification system, the Error Backpropagation and Genetic Algorithm is used. The neural network is first initialized for processing by removing any run-specific data from the previous iteration. This involved resetting any cached values, including deltas, and activation levels, to a random value. An input pattern is then presented to the network. In this work, we have studied various techniques to prepare input pattern from given character and identified the better one for our needs. We have used gray-scale images for the steps and that binary prepossessed images is then fed to BPN and GA. In learning phase, we set the error rate is set less than 0.05 for this network. The number of iterations in which the network reached the specified error goal is 300000. The learning rate of the network is set to 0.9 and the spread factor is set to 0.15. It is expected from a character verification system to acquire high accurate recognition rate while the unrecognition rate and false recognition rate should be too low. In later case is opposite, that is, if unrecognition rate and false recognition rate are high, the security questions arises. Following are some comparison between the algorithms based on some parameters.

TABLE I. Time required by BPN for error rate 0.09 and 0.05 with various no. of pattern

Generation	**No. of Pattern**	**Time {sec}**
10	**5**	**0.00**
10	**8**	**0.0**
10	**10**	**1.672**
10	**13**	**1.79**
15	**5**	**0.55**
15	**8**	**1.19**
15	**10**	**1.18**
15	**13**	**1.95**

TABLE II. Time required by GA for generation of 10 and 15 with various no. of pattern

Generation	**No. of Pattern**	**Time {sec}**
10	5	0.00
10	8	0.0
10	10	1.672
10	13	1.79
15	5	0.55
15	8	1.19
15	10	1.18
15	13	1.95

TABLE III. Accuracy of BPN – 84.62%

Parameters	**Value**
Total Learned character	13
Execution of different character	30
Successful Recognize	11
Unsuccessful Recognize	2

TABLE IV. Accuracy of GA – 61.54%

Parameters	Value
Total Learned character	13
Execution of different character	30
Successful Recognize	8
Unsuccessful Recognize	5

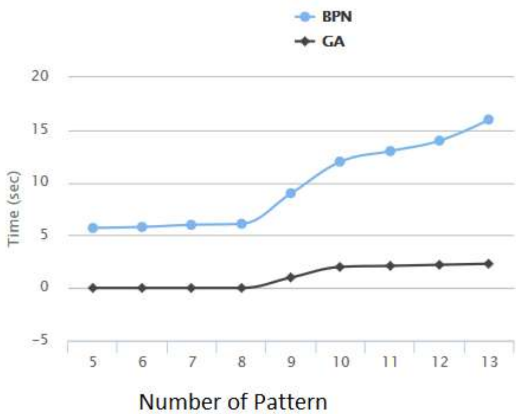


Fig. 8. Time comparison between BPN with error rate 0.09 and GA with generation 10

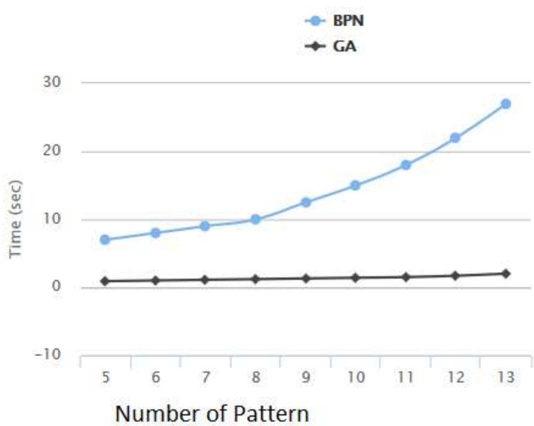


Fig. 9. Time comparison between BPN with error rate 0.05 and GA with generation 15

VI. CONCLUSION

The field of character recognition is an active topic of research and development of many years. The character recognition system is used to identify the Character. Till now, no intelligent system has been developed which gives output without no errors. Always there are occurring some errors due to the variation of input patterns. Character recognition is .a very complicated system because character shape can be various shaped. It is not possible to store all shaped character in database. As a result there are situations like unrecognizing and false recognizing. Moreover due to the performance variation of the input character can't be detected correctly and pattern may change extremely. In this paper we have developed and illustrated a recognition system for character using BPN and GA and shows the various parameters comparison.

REFERENCES

[1] R. Beale and T. Jackson, "Neural Computing and Introduction"

[2] Loone, 'Back-propagation Training with Multiple Pattern"

[3] Noyes James L, "Artificial Intelligence with Common Lisp"

[4] Patric Henry Winston, "Artificial Intelligence"

[5] Rafael C. Conzalez and Richard E. Woods, "Digital Image Processing"

[6] Ramesh Chandra Debnath and A. K. M. Akhtar Hossain, "Pattern Classification by Neural Networks"

[7] Dan W. Patterson, "Introduction to Artificial Intelligence and Expert Systems"

[8] S Rajasekaram and G. A. Vijayalakshami Pai, "Neural Networks, Fuzzy Logic & Genetic Algorithm"

[9] Dwayne Philips, "Image Processing in C"

[10] Jhon G. Proakis, Dimitris G. Manolakis, "Digital Signal Processing Using Matlab"

[11] Dr. K. Gurney, "Neural Net"

[12] James L. Noyes, "Artificial Intelligence with Common Lisp, Fundamentals of Symbolic and Numeric Processing"

[13] J. Soldek, V. Shmerko, Phil Phillips, G. Kukharev, W. Rogers, and S. Yanushkevich "Image Analysis and Pattern Recognition in Biometric Technologies"

[14] N. K. Bose, P. Liang, "Neural Networks Fundamentals with Graphics; Algorithms and Applications", International Edition, 1996

[15] R. Plamondon, "The design of an On-Line character Verification System: From Theory to Practice"

Histogram based Water Quality Assessment in Satellite Images

Mumu Aktar, Md. Al Mamun
Computer Science & Engineering Department
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
mumu.ruet@gmail.com,cse_mamun@yahoo.com

Md. Shamimur Rahman Shuvo, Md. Ali Hossain
Computer Science & Engineering Department
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
shuvo1364@gmail.com,ali.ruet03@yahoo.com

***Abstract*—Water is one of the most precious resources of our environment. This water body is often faced the quality questions because of being polluted by ammonia, chemical wastes, sulfur dioxide from power plants, fertilizers containing nutrients--nitrates and phosphates, sediment, phytoplankton, etc. So, it is very necessary to assess quality of different water bodies. In this study, satellite images have been used for water quality measurement using histogram comparisons. A satellite image has been chosen to use as the original image whose water body has been considered as a standard of clear water body. Then this clear water body has been separated from other features using clustering at first to be used as standard ones and their number of pixels in percentage has been counted. Those images which contain the same percent of water bodies as the perspective standard ones can be verified by comparing their histograms. Euclidean distance has been measured between the standard and tested image's histograms. A tolerance level has been taken to assess the water quality as excellent, better, good, bad and poor. Finally if the distance falls within the tolerance level the water body can be categorized as excellent, better, good, bad or poor based on their degree of purity.**

Keywords—quality assesment; histogram;clustering;Euclidean distance

I. Introduction

Water is a vital element of our environment. But it becomes harmful for us when mixed with various pollutants like sewage, fertilizer, acid rain, wastes from oil industry, etc. These pollutants being mixed with water of lakes, rivers, oceans, aquifers, and groundwater make those water sources being affected. Water of these sources should be tested at regular intervals for measuring their quality to know how much safe they are. In this paper the concern of testing water regularly has become the issue of quality assessment of water body.

Remotely sensed data provides a means of delineating water boundaries over a large area at a given point in time [1].So here remotely sensed images have been chosen for detecting water body and measuring its quality. Commonly three types of techniques are applied to remotely sensed data to separate water bodies from other features: 1. Supervised classification technique 2. Unsupervised classification technique 3. Object-based classification technique. In supervised classification, analysts should have priori knowledge about the data and should provide the network not only the inputs but also desired output that means the network should be trained up first and then it can test new data to classify them into different classes. Artificial neural network can be an example of supervised classification. A cost-effective remote sensing-based methodology was developed to predict water quality parameters namely chlorophyll-a, turbidity, and phosphorus over a large and logistically difficult area before and after ecosystem restoration and during the wet and dry seasons using artificial neural network [2].

In case of unsupervised classification no target output is presented to the network & it learns of its own by discovering and adapting to structural features in the input patterns. Common Unsupervised classification methods are -a. Simple One-Pass Clustering b. K Means c. Fuzzy d. Minimum Distribution Angle e. Adaptive Resonance. The left one is object-based image classification which generates objects of different shapes and scales and these objects can be classified based on texture, context and geometry.

After choosing techniques for detecting water body it is necessary to identify features on the basis of which the decision of separation of water body from others can be taken. Reflectivity can be considered as a feature vector in case of water body classification. Reflectivity of water varies from band to band. The reflectance of water in infrared and short-wave infrared bands is lower than that of other objects like vegetation, buildings, bare soil, and roads and so forth [5]. On the basis of this information water body can be separated from other objects. Composition of two or more bands also gives information about water body. In case of a multispectral image the combination of 3, 2, 1 bands (which is called the nature color combination) provides most water penetration and the composite of bands 7, 4, 2 shows the river in dark blue color also in case of multispectral imagery[6].

Multivariate statistical method including cluster analysis (CA) was used to assess temporal and spatial variations in the water quality of Euphrates River, Iraq, for a period 2008-2009 using 16 parameters at 11 sampling sites [12]. Here in this paper ISODATA clustering (an example of unsupervised classification) algorithm has been used to separate water body from other objects of an image. After applying this clustering technique only water body has been considered for the work

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

rejecting the other objects. Then histograms of the water body of original images have been taken into consideration to compare them with the histograms of test images using Euclidean distance. Finally the distance has been compared to the specified tolerance level of water quality measurement on the basis of which decision about water quality has been taken.

II. CLUSTERING

Clustering is one kind of unsupervised classification technique. It actually means finding natural grouping among objects that is same type of objects are taken into one group. In case of image classification, pixels are grouped based on the reflectance properties of pixels and these groupings are called clustering. We can consider the following figure to understand clustering.

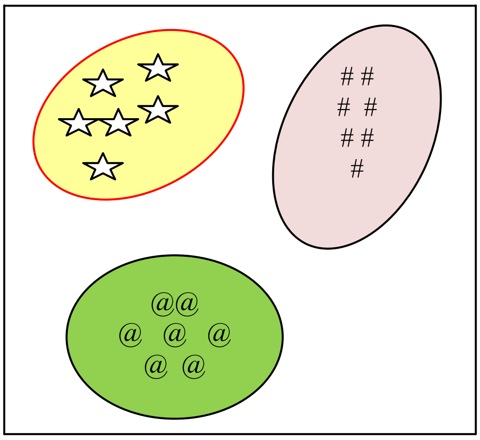


Fig. 1. Example of clustering

☆ : Pixels representing water body

@ : Pixels representing vegetation

: Pixels representing soil

Here from the above figure it can be noticed that pixels representing water body contains same criteria and thus they are considered as one group, the same case for vegetation and soil and this represents the concept of clustering.
Clustering algorithms can be of different types such as single-pass, ISODATA and so on. ISODATA clustering algorithm has been chosen for our work.
ISODATA [8] is a method of unsupervised classification which doesn't need to know much about data beforehand. In case of this algorithm user defines threshold values for parameters and computer runs the algorithm through many iterations until threshold is reached. It has been observed from [8], [9] that ISODATA works based on center method. In this algorithm cluster centers are randomly placed and pixels are assigned based on the shortest distance to center method. The standard deviation within each cluster and the distance between cluster centers is calculated. The ISODATA algorithm splits clusters if one or more standard deviation is greater than the user-defined threshold and the algorithm merges clusters if the distance between them is less than the user-defined threshold. The algorithm performs second iteration with the new cluster centers. ISODATA continues performing iterations until
– i) the average inter-center distance falls below the user-defined threshold,
– ii) the average change in the inter-center distance between iterations is less than a threshold, or
– iii) the maximum number of iterations is reached

In this work, we have used MultiSpecWin32 software to perform the ISODATA clustering. Here to classify each pixel to a specific cluster a threshold value is needed on the basis of which merging or splitting of clusters is done [9]. Merging of clusters is done when the number of pixels in a cluster is less than a certain threshold or if the centers of two clusters are closer than a certain threshold. If the clusters' standard deviation exceeds a predefined value and the number of members (pixels) is twice the threshold for the minimum number of members then clusters are split. For our work '100' has been chosen as the threshold value. This '100' is a standard threshold value which forces the system to assign every pixel in the image to one of the clusters and value less than this specifies the tolerance for assignment of pixels. For this reason '100' has been chosen here so that no pixel remains left without being assigned to any one of the clusters.
Each of the original images has been clustered into seven or more clusters or groups and the number varied from image to image. Then the water body has been separated from the other objects of the images manually.

III. EUCLIDEAN DISTANCE

The histogram matching of the standard image and the tested image has been performed using Euclidean distance measure. Normally it is considered as the distance between two points. Mathematically, we can define the distance as:

$$\mathrm{d}(\mathbf{p},\mathbf{q}) = \mathrm{d}(\mathbf{q},\mathbf{p}) = \sqrt{(q_1-p_1)^2 + (q_2-p_2)^2 + \cdots + (q_n-p_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n}(q_i-p_i)^2}. \quad \text{-------------------(1)}$$

Here in equation (1), d(p,q) is the distance between p & q. $p=(p_1,p_2,\ldots\ldots\ldots p_n)$ and $q=(q_1,q_2,\ldots\ldots\ldots\ldots\ldots q_n)$.

Since the euclidean distance measures the distance between each position of p & q we have decided it to compare the image histograms. To compare the histograms by euclidean distance the pixel positions should be almost same and for that it has been decided to make a database which contains five original images' histograms as standard ones. If any test image contains 40 percent of water body then its histogram will be compared with the standard histogram from original images' database containing 40 percent of water body.

IV. METHODOLOGY

A. Data Collection

Image data has been collected from http://landsat.visibleearth.nasa.gov/. From here at first five images have been selected containing clear water bodies. They are denoted as original image data. These images have been chosen on the basis of their percentage of containing water body. A data table has been created showing the percentage of water body in each image.

TABLE I. DATABASE OF PERCENTAGE OF WATER BODY

Image Name	% of water body
Image 1	15
Image 2	28
Image 3	35
Image 4	50
Image 5	65

One-channel grayscale type has been chosen for this work in case of all images. The five original imges of different sources are shown below.

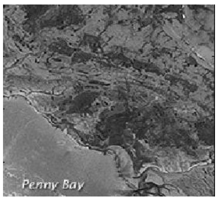


Fig. 2. Image 1

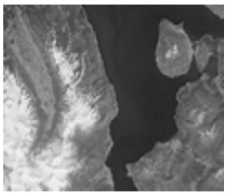


Fig. 3. Image 2

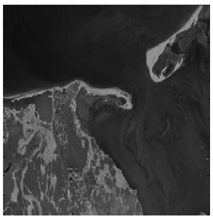


Fig. 4. Image 3

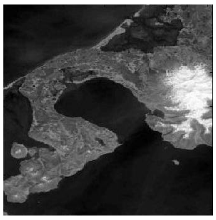


Fig . 5. Image 4

Fig. 6. Image 5

Then from the same source images have been taken to test their quality by comparing with the standard images containing clear water body.

B. Method

The methodology of this work consists of the following steps.

- Clustering: In this study, to determine water quality of different sources water body has been separated at first from other objects of an original image. To cluster pixels containing water body, ISODATA clustering algorithm has been applied and threshold value of 100 has been chosen. Number of clusters has been varied from image to image and the number has been chosen manually after observing images. The clustered image containing 28% of water body is shown below and in the same way the other images have been clustered.

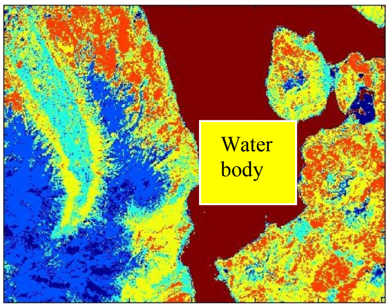


Fig. 7. clustered view of image 2 containing 28% of water body

The images have been clustered using **MultiSpecWin32 software.** Then using the software and also MATAB the number of pixels containing water body in every image has been calculated and thus the percentage of pixels has been found out.

- Database: Since comparing two histograms are illogical if they do not contain almost same number of pixels, so at first the above database (Table: 1) has been prepared using the above five images. Thus if any test image contains nearby 15% of water body then its quality can be measured by comparing its histogram to the original image's histogram containing 15% of water body. The same case can be done for testing images containing 28%, 35%, 50% and 65% water body.

- Histogram: Finally using Euclidean distance the histograms of original and test images have been compared. The formula of equation (1) has been used to compare between every pixel position of the original image and the test image and the distance between each position has been measured. Finally the overall distance has been taken to compare with the defined threshold value to categorize the test image's quality as excellent, better, good, bad or poor. The histograms of original and test images containing 28% of water body are shown below. Actually the original image contains 28.9073% of water body whereas the test image contains 27.1621% of water body.

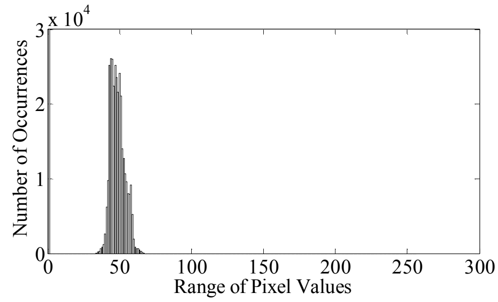


Fig. 8. Histogram of original image containing 28% of water body

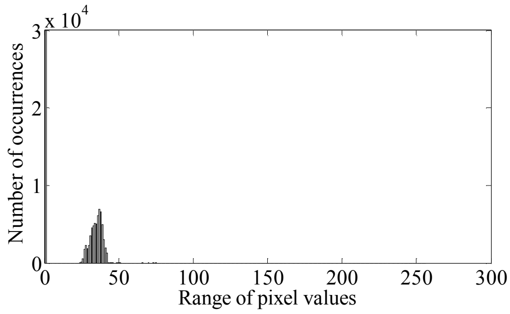


Fig. 9. Histogram of test image containing 28% of water body

- Threshold Define: To compare the histograms of original and test images threshold has been chosen and on the basis of this threshold water quality has been categorized as below.

TABLE II. CATEGORIZATION OF WATER QUALITY

Threshold	Quality of water
1.5e+006	EXCELLENT
2.5e+006	BETTER
3.5e+006	GOOD
4.5e+006	BAD
5.5e+006	POOR

From the above table (Table II) we can decide that if the distance value between original and test image's histograms is less than or equal to the defined thresholds then the water quality can be fell among one of the above declared water qualities. Figure 8 and 9 show the original and test image's histograms containing 28% of water body and here the distance between them is 2.0499e+006 and from Table II we can notice that it is nearest to 2.5e+006 and thus we can declare the test image's water body as of BETTER quality.

V. RESULT AND DISCUSSION

From the methodology and Table II it can be concluded that if the percentage of water bodies in both the original and test images are almost same then from our database of clear water bodies any type of water body can be qualified. The testing has been done for all of the original images containing 15%, 28%, 35%, 50% and 65% of water bodies. If we want to test water quality of an area the only thing we need to check at first is the percentage of water body in that image and if it is almost nearer to the percentage of water bodies contained in our database then it can be easily qualified.

VI. CONCLUSION AND FUTURE WORK

An easy approach has been presented here to measure water quality of different reservoirs. Mainly ISODATA clustering algorithm has been applied to a satellite image to separate its water body from other features. Then histogram matching has been performed between original and test images using Euclidean distance to assess the water body as excellent, better, good and so on. Since in this work we have not considered different types of images such as .tiff, .jpg, .png that means for the comparison we have considered same type of original and test images, so in future the work will be extended so that there is no limitation in case of comparison. Since histogram analysis may not work as appropriate technique for bi-modal and not normally distributed data, so some other matching techniques will be tried to develop in future.

REFERENCES

[1] Paul Shane Frazier, Kenneth John Page, "Water body detection and delineation with Landsat TM data," Photogrammetric Engineering & Remote Sensing Vol. 66, No. 12, December 2000, pp. 1461-1467.

[2] Yirgalem Chebud, Ghinwa M. Naja, Rosanna G. Rivero, Assefa M. Melesse, "Water Quality Monitoring Using Remote Sensing and an Artificial Neural Network", Water Air Soil Pollut (2012) 223:4875–4887.

[3] Vini Malik, Aakanksha Gautam , Aditi Sahai, Ambika Jha, AnkitaRamvir Singh, "Satellite Image Classification Using Fuzzy Logic", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-2, Issue-2, May 2013.

[4] Ms. Venu Shah[1], Ms. Archana Choudhary[2], Prof. Kavita Tewari[3], "River extraction from satellite image", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 2, July 2011.

[5] Kshitij Mishra, P. Rama Chandra Prasad, "Automatic Extraction of Water Bodies from Landsat Imagery Using Perceptron Model", Hindawi Publishing Corporation, Journal of Computational Environmental Sciences, Volume 2015, Article ID 903465, 9 pages.

[6] B.Naga Jyothi[1], Dr. G. R. Babu[2], Dr.I.V.Murali Krishna[3], "Thematic Classification of Multispectral Imagery", International Journal of Electronics and Computer Science Engineering, ISSN-2277-1956/V1N2-181-190.

[7] Nguyen Dinh Duong, "water body extraction from multi spectral image by spectral Pattern analysis",International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B8, 2012 XXII ISPRS Congress, 25 August – 01 September 2012.

[8] Vassilaros, A., "ISODATA-Lectures Notes", 2010.

[9] J. R. Jensen, "Introductory Digital Image Processing--A Remote Sensing Perspective", Prentice Hall, Inc, New Jersey, pp. 197-256, 1996.

[10] S Raja Mohamed, YSSR Murthy, "Identification of Water Bodies from High and Multi-level Resolution Satellite Images Using Novel Feature Extracting Technique", Advances in Water Resource and Protection (AWRP) Volume 1 Issue 2, April 2013.

[11] F.L. Hellweger[a,b,*], P. Schlosser[a,c], U. Lall[a,c], J.K. Weissel[c] , "Use of satellite imagery for water quality studies in New York Harbor", Estuarine, Coastal and Shelf Science 61 (2004) 437–448.

[12] Emad A. Mohammad Salah[1], Ahmed M. Turki[2], Eethar M. Al-Othman[2], "Assessment of Water Quality of Euphrates River Using Cluster Analysis", Vol. 3 No. 12 (2012) , Article ID: 25650 , 5 pages.

Cross-correlation Based Approach of Underwater Network Size Estimation with Unequal Sensor Separation

D. K. Mondal, S. A. H. Chowdhury, Q. N. Ahmed
Dept. of Electronics & Telecommunication Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
dheeman.ete10.ruet@gmail.com
arif.1968.ruet@gmail.com
nusabaahmedaurthy@gmail.com

M. S. Anower
Dept. of Electrical & Electronic Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
md.shamimanower@yahoo.com

Abstract—**It is very difficult to estimate the number and location of the operational nodes in underwater network using conventional protocol based techniques due to long propagation delay, strong background noise, non-negligible capture effect, high absorption and dispersion of underwater environment. So, a unique approach based on cross-correlation is applied for underwater network size estimation. In three-sensor case of this approach, network size is obtained by cross-corelating the Gaussian signals received at equidistant pair of sensors. The assumption of equidistant sensors poses some limitations and requires further investigation. The aim of this work is to obtain a more flexible estimation process by removing this assumption. For this purpose, three-sensor case of cross-correlation based underwater network size estimation technique is investigated in this paper with unequal sensor separation.**

Keywords—cross-correlation function; equidistant sensors; network size estimation; underwater acoustic communication; underwater network; unequal sensor spacing.

I. Introduction

Reliability as well as performance of any wireless sensor network largely depends on the number and location of the active nodes available in the network. Different types of techniques [1]–[5] have been presented in literature to estimate the size of a network. The abovementioned methods ([1]–[5]) do not take into account the characteristics of underwater acoustic channel (UAC) [6]. This limits their performances to estimate the size of underwater network. To overcome the limitation of long propagation delay of UAC, Howlader et al. proposed a delay insensitive estimation process based on ALOHA protocol [7]. Howlader et al. also investigated a node estimation method considering non-negligible capture effect of UAC [8]. These ([7], [8]) are protocol(s) based techniques. Increased protocol complexities of these methods in underwater environment lead to inefficient estimation results.

To address this issue, Anower et al. introduced a cross-correlation based underwater network size estimation technique with reduced protocol complexity using two sensors [9]. This work has been extended by Chowdhury et al. using three sensors [10] to improve the estimation performance of cross-correlation based technique. Two estimation schemes have been proposed using three sensors based on the arrangements of the sensors. They are: sensors in line (SL) scheme and triangular sensors (TS) scheme. In both SL and TS schemes, estimation is performed with equal spacings between the sensors. This may not be possible in some practical cases as it is somewhat troublesome to place the sensors at desired locations among the randomly distributed nodes. It is also problematic to keep the fixed distances between the sensors in deep water due to the waves and sea creatures. So, the constraint of equal sensor spacing makes practical implementation of these schemes difficult. In this paper, underwater network size estimation is performed by cross-correlating the Gaussian signals received at unequally spaced sensors using TS scheme to withdraw this constraint.

II. Background

A review of the size estimation process using TS scheme [10] is discussed in this section to provide sufficient background of this work. Let us consider a three-dimensional spherical network containing N nodes and three sensors, where the nodes are evenly distributed over the whole region as shown in Fig. 1. We assume that, the nodes are the sources of white Gaussian signals in response to a probe request. In TS scheme, the three sensors (H_1, H_2 and H_3) form an equilateral triangle inside the network for estimation purpose, where the centroid of that triangle lies at the centre of the sphere as shown in Fig. 1. So, the distances between the sensors are such that, $d_{\mathrm{DBS}_{12}}$ (distance between H_1 and H_2) $= d_{\mathrm{DBS}_{23}}$ (distance between H_2 and H_3) $= d_{\mathrm{DBS}_{31}}$ (distance between H_3 and H_1) $= d_{\mathrm{DBS}}$ (distance between the equidistant pair of sensors).

Now, if a node emits a Gaussian signal using acoustic wave as carrier (which is recorded at three sensors with the corresponding time delays and attenuations), the CCFs of the Gaussian signals received at each pair of sensors can be expressed individually by a delta function [11]. For every node, we obtain a delta function after cross-correlation. These delta functions occupy any position in the sample space between the corresponding pair of sensors. We can now consider each delta function as a ball and the samples between the corresponding pair of sensors as bins into which the balls may fall. One such CCF obtained with N (=1000) nodes is shown in Fig. 2.

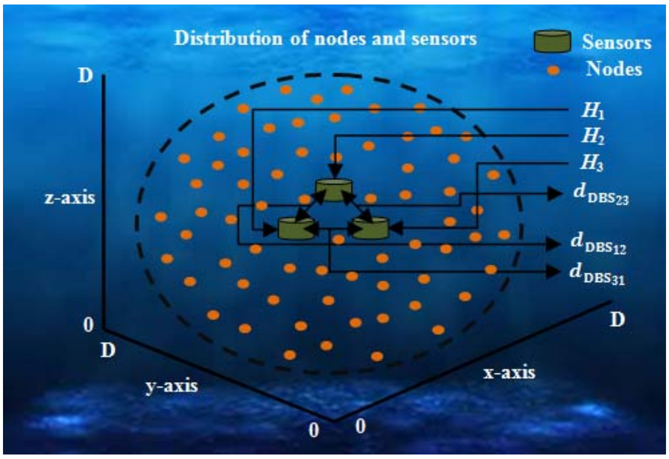


Fig. 1. Distribution of underwater network nodes with N transmitting nodes and three equally separated sensors for TS scheme.

Number of bins, b (as shown in Fig. 2) is defined as twice the number of samples between the sensors, m minus one and can be expressed as [11]:

$$b = \frac{2 \times d_{\mathrm{DBS}} \times S_{\mathrm{R}}}{S_{\mathrm{P}}} - 1 \tag{1}$$

where, S_{R} means sampling rate and S_{P} is the speed of acoustic wave propagation.

In TS scheme, three CCFs, $C_{12}(\tau)$ (formulated by cross-correlating the composite Gaussian signals $S_{\mathrm{r_{c1}}}(t)$ and $S_{\mathrm{r_{c2}}}(t)$ received at H_1 and H_2, respectively), $C_{23}(\tau)$ (formulated by cross-correlating the composite Gaussian signals $S_{\mathrm{r_{c2}}}(t)$ and $S_{\mathrm{r_{c3}}}(t)$ received at H_2 and H_3, respectively) and $C_{31}(\tau)$ (formulated by cross-correlating the composite Gaussian signals $S_{\mathrm{r_{c3}}}(t)$ and $S_{\mathrm{r_{c1}}}(t)$ received at H_3 and H_1, respectively) are used for estimation. Ratio of standard deviation (σ) to the mean (μ), R of CCF is chosen as the estimation parameter of this process, as it requires no prior knowledge of the signal strength from the nodes [12]. Three estimation parameters, R_{12}, R_{23} and R_{31} are derived from three CCFs, $C_{12}(\tau)$, $C_{23}(\tau)$ and $C_{31}(\tau)$, respectively, to calculate the final estimation parameter of TS scheme. Then, the final estimation parameter, $R_{\mathrm{average}}^{\mathrm{3CCF}}$ of TS scheme is obtained by taking the average of R_{12}, R_{23} and R_{31}, and can be expressed as [10]:

$$R_{\mathrm{average}}^{\mathrm{3CCF}} = \frac{R_{12} + R_{23} + R_{31}}{3} = \frac{\frac{\sigma_{12}}{\mu_{12}} + \frac{\sigma_{23}}{\mu_{23}} + \frac{\sigma_{31}}{\mu_{31}}}{3} \tag{2}$$

where, σ_{12}, σ_{23} and σ_{31} denote the standard devation and μ_{12}, μ_{23} and μ_{31} denote the mean of the CCFs, $C_{12}(\tau)$, $C_{23}(\tau)$ and $C_{31}(\tau)$, respectively.

To overcome the difficulty of determining σ and μ of the CCFs using complex statistical expressions, the cross-correlation problem is being reframed into probability problem. After reframing, R of CCF can be expressed as [9]:

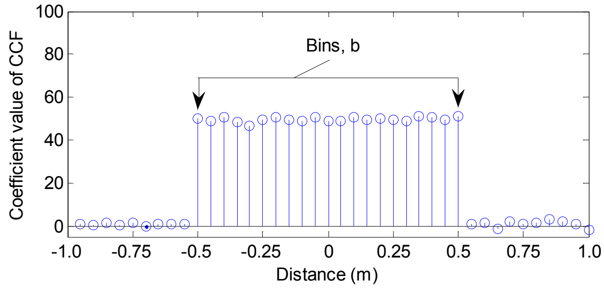


Fig. 2. Bins, b of CCF obtained with N (=1000) nodes.

$$R = \frac{\sigma}{\mu} = \frac{\sqrt{N \times \frac{1}{b} \times \left(1 - \frac{1}{b}\right)}}{N \times \frac{1}{b}} = \sqrt{\frac{(b-1)}{N}} \tag{3}$$

Using this expression, (2) can be written as:

$$R_{\mathrm{average}}^{\mathrm{3CCF}} = \frac{\sqrt{\frac{b_{12}-1}{N}} + \sqrt{\frac{b_{23}-1}{N}} + \sqrt{\frac{b_{31}-1}{N}}}{3} \tag{4}$$

Here, b_{12}, b_{23} and b_{31} represent the number of bins of the CCFs, $C_{12}(\tau)$, $C_{23}(\tau)$ and $C_{31}(\tau)$, respectively. In this case, $b_{12} = b_{23} = b_{31} = b$ according to (1); as the values of S_{R} and S_{P} are fixed during the estimation process and $d_{\mathrm{DBS}_{12}} = d_{\mathrm{DBS}_{23}} = d_{\mathrm{DBS}_{31}} = d_{\mathrm{DBS}}$. So, (4) becomes [10]:

$$R_{\mathrm{average}}^{\mathrm{3CCF}} = \frac{R_{12} + R_{23} + R_{31}}{3} = \sqrt{\frac{(b-1)}{N}} \tag{5}$$

Using (5), we can estimate N, as we know b from (1) and can calculate $R_{\mathrm{average}}^{\mathrm{3CCF}}$ from the CCFs. The total estimation process of TS scheme is represented through a block diagram in Fig. 3.

III. Estimation with Unequal Sensor Separation

In previous section, estimation is performed using TS scheme with equal spacings between the sensors. This section investigates the estimation process considering unequally spaced sensors in triangular orientation as shown in Fig. 4.

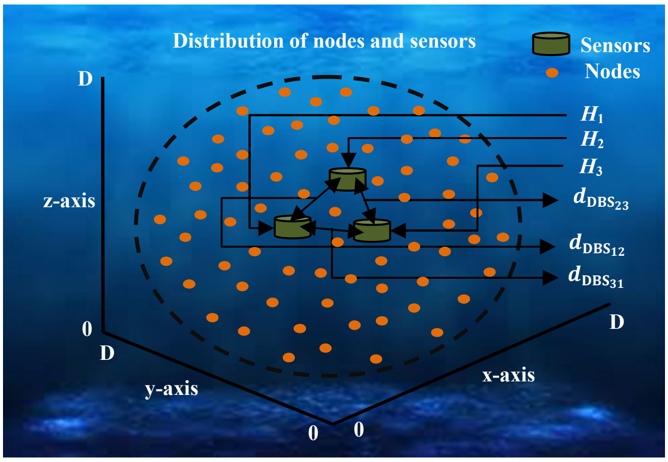


Fig. 4. Distribution of underwater network nodes with N transmitting nodes and three unequally spaced sensors for TS scheme.

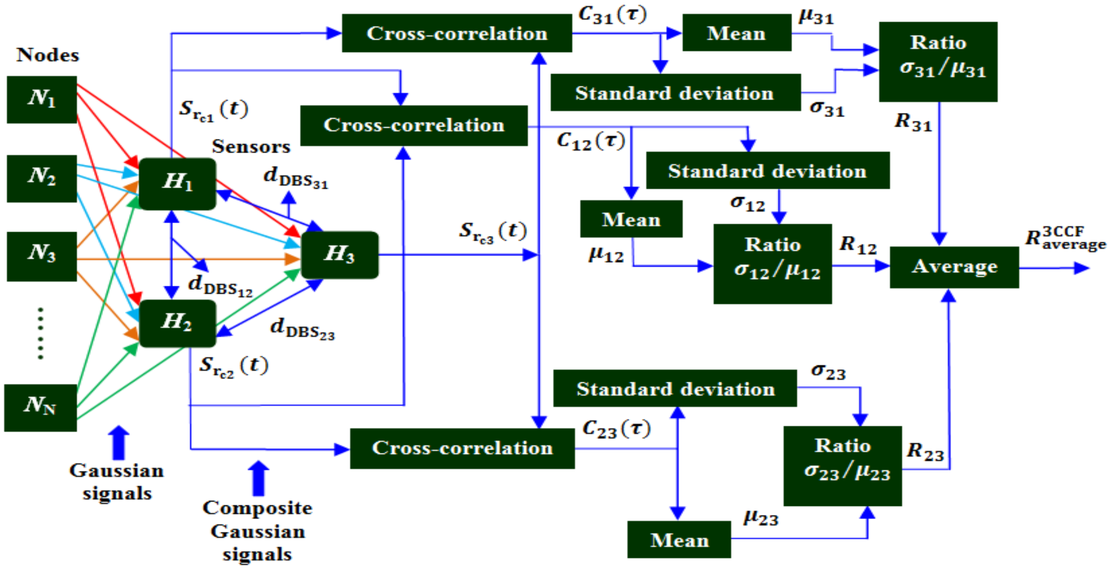


Fig. 3. Block diagram representation of the estimation process for TS scheme.

The position of the triangle formed by three unequally spaced sensors is such that, the centre of the spherical network lies at the centroid of that triangle. Unequal distances between the sensors of this approach suggest that, $d_{DBS_{12}} \neq d_{DBS_{23}} \neq d_{DBS_{31}}$. Estimation procedure of TS scheme with unequal sensor spacing is similar to that of equal sensor spacing as described in Section II. So, estimation parameter of TS scheme with unequal sensor separation, $R^{3CCF}_{unequal}$ is determined similarly as discussed in the previous section by taking the average of the ratios of σ and μ of the CCFs. After reframing the cross-correlation problem into probability problem, $R^{3CCF}_{unequal}$ of CCFs can be expressed using (4) as:

$$R^{3CCF}_{unequal} = \frac{\sqrt{\frac{b_{12}-1}{N}} + \sqrt{\frac{b_{23}-1}{N}} + \sqrt{\frac{b_{31}-1}{N}}}{3} \quad (6)$$

In this case, $b_{12} \neq b_{23} \neq b_{31}$ according to (1), as the values of S_R and S_P are fixed during the estimation process and $d_{DBS_{12}} \neq d_{DBS_{23}} \neq d_{DBS_{31}}$. So, further simplification of (6) is not possible. From (6), network size, N can be estimated using the following expression:

$$N = \left(\frac{\sqrt{b_{12}-1} + \sqrt{b_{23}-1} + \sqrt{b_{31}-1}}{3 \times R^{3CCF}_{unequal}} \right)^2 \quad (7)$$

In order to establish a relationship between the estimation parameters of TS scheme with equal and unequal spacings between the sensors, ratio of $R^{3CCF}_{average}$ and $R^{3CCF}_{unequal}$ is obtained as follows:

$$r = \frac{R^{3CCF}_{average}}{R^{3CCF}_{unequal}} = \frac{3\sqrt{b-1}}{\sqrt{b_{12}-1} + \sqrt{b_{23}-1} + \sqrt{b_{31}-1}} \quad (8)$$

From (8), we can write:

$$R^{3CCF}_{average} = r \times R^{3CCF}_{unequal} \quad (9)$$

This expression relates $R^{3CCF}_{average}$ and $R^{3CCF}_{unequal}$. If r = 1, then $R^{3CCF}_{average} = R^{3CCF}_{unequal}$ according to (9). Now, putting r = 1 in (8), we can obtain the condition of identical estimation parameters for both equal and unequal sensor separation cases as follows:

$$b = \frac{\left(\sqrt{b_{12}-1} + \sqrt{b_{23}-1} + \sqrt{b_{31}-1}\right)^2 + 9}{9} \quad (10)$$

It is noted that, b_{12}, b_{23} and b_{31} are the three different number of bins of the three CCFs for unequal sensor spacing case and b is the number of bins of all three CCFs for equal sensor spacing case.

IV. Results and Discussion

Simulations are performed using TS scheme with unequally spaced sensors in matlab programming environment for validation. Some useful simulation results are presented in this section. Simulated results of $R^{3CCF}_{unequal}$ are plotted against N in Fig. 5 with the corresponding theoretical results (obtained using (6)) for different $d_{DBS_{12}}$, $d_{DBS_{23}}$ and $d_{DBS_{31}}$. The values of $d_{DBS_{12}}$, $d_{DBS_{23}}$ and $d_{DBS_{31}}$ are: 1.5m, 0.5m and 1.0m in Fig. 5(a); and 1.0m, 1.0m and 2.0m in Fig. 5(b), respectively. Other parameters used in the simulations (throughout the work unless otherwise mentioned) are: dimension of the sphere, D = 2000m; S_R = 30kSa/s; S_P = 1500m/s; signal length, $N_s = 10^6$ samples; SNR = 20db. It is obvious from Fig. 5 that, simulated results agree with the theoretical results which validate the estimation process described in Section III.

To obtain similar estimation parameters for both equal and unequal sensor separation cases, simulation is performed using b = 34 (d_{DBS} = 0.875m), b_{12} = 9 ($d_{DBS_{12}}$ = 0.25m), b_{23} = 47 ($d_{DBS_{23}}$ = 1.2m) and b_{31} = 59 ($d_{DBS_{31}}$ = 1.5m) which satisfies (10). Results are shown in Fig. 6. In Fig. 6, the solid lines represent the theoretical results and the circles and stars represent the simulated results with equal and unequal spacings between the sensors, respectively. It is observed from Fig. 6 that, the two simulated results agree with each other and both

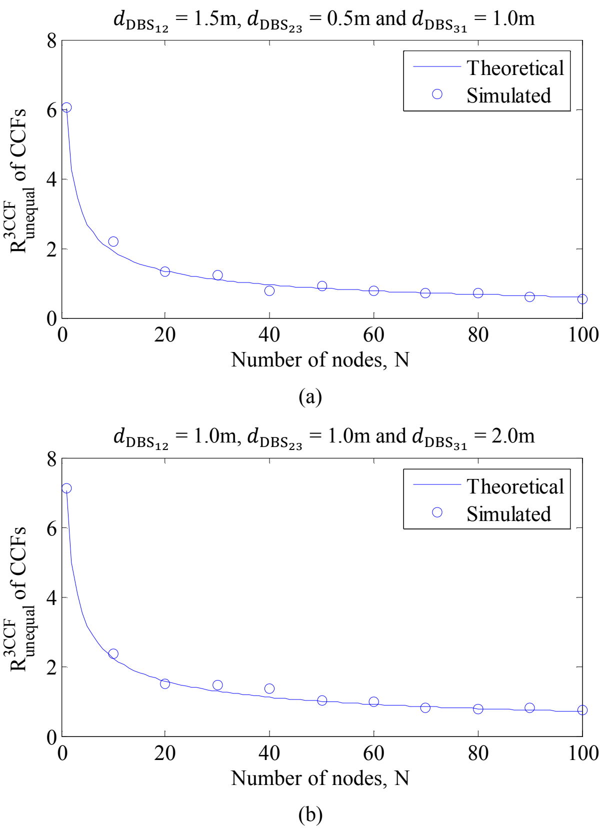


Fig. 5. $R^{3CCF}_{unequal}$ versus N plots for TS scheme with unequally spaced sensors using different $d_{\mathrm{DBS}_{12}}$, $d_{\mathrm{DBS}_{23}}$ and $d_{\mathrm{DBS}_{31}}$.

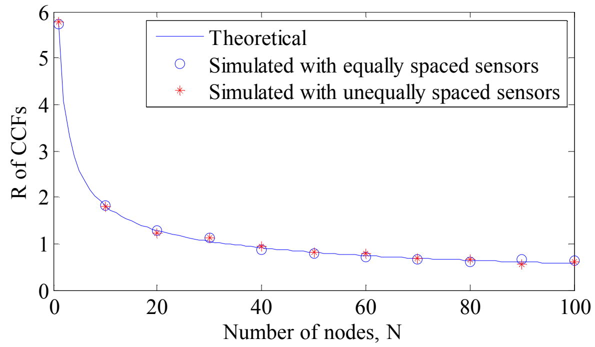


Fig. 6. R versus N plot comparisons of results for theoretical, and simulated with equally and unequally spaced sensors using TS scheme.

follow the theoretical results. This verifies the effectiveness of the condition (10) formulated in the previous section. The simulation results shown in this paper are based on the assumption that, the power received from each node is equal, which can be obtained by applying proper probing technique.

V. Conclusion

This paper focuses on removing the assumption of equal distances between the sensors in three-sensor approach of cross-correlation based underwater network size estimation technique. In this investigation, estimation is performed with unequal sensor spacing using TS scheme to achieve that goal. It is shown that, efficient estimation can be obtained with unequally spaced sensors. This eliminates the constraint of equidistant sensors. A condition is derived in this paper to obtain indistinguishable estimation parameters for both equal and unequal sensor separation cases and verified by simulation. In this estimation process, the sensors are unequally spaced among the randomly distributed nodes which might cause problem in highly densed networks. Current research is going on to estimate the network size by placing the sensors outside the network, which will effectively solve this issue. We also intend to investiagte the estimation process in different shaped network with various distribution of nodes. Finally, conducting experimental estimation using this technique is our ultimate goal.

References

[1] C. Budianu, S. Ben-David, and L. Tong, "Estimation of the number of operating sensors in large-scale sensor network with mobile access," IEEE Transactions on Signal Processing, vol. 54, no. 5, pp. 1703–1715, May 2006. doi: 10.1109/TSP.2006.871973.

[2] O. Sluciak, and M. Rupp, "Network size estimation using distributed orthogonalization," IEEE Signal Processing Letters, vol. 20, no. 4, pp. 347–350, April 2013. doi: 10.1109/LSP.2013.2247756.

[3] D. Varagnolo, G. Pillonetto, and L. Schenato, "Distributed Cardinality Estimation in Anonymous Networks," IEEE Transactions on Automatic Control, vol. 59, no. 3, pp. 645–659, March 2014. doi: 10.1109/TAC.2013.2287113.

[4] J. Cichoń, J. Lemiesz, and M. Zawada, "On cardinality estimation protocols for wireless sensor networks," In Proc. 10th Int. Conf. Ad Hoc Networks and Wireless (ADHOC-NOW), 18–20 July 2011, Paderborn, Germany, pp. 322–331. doi: 10.1007/978-3-642-22450-8_25.

[5] Y. Zheng, and M. Li, "Towards More Efficient Cardinality Estimation for Large-Scale RFID Systems," IEEE/ACM Transactions on Networking, vol. 22, no. 6, pp. 1886–1896, Dec. 2014. doi: 10.1109/TNET.2013.2288352.

[6] S. Climent, A. Sanchez, J. V. Capella, N. Meratnia, and J. J. Serrano, "Underwater Acoustic Wireless Sensor Networks: Advances and Future Trends in Physical, MAC and Routing Layers," Sensors (Basel, Switzerland), vol. 14, no. 1, pp. 795–833, Jan. 2014. doi: 10.3390/s140100795.

[7] M. S. A. Howlader, M. R. Frater, and M. J. Ryan, "Delay-insensitive identification of neighbors using unslotted and slotted protocols," Wirel. Commun. Mob. Comput., vol. 14, no. 8, pp. 831–848, June 2014. doi: 10.1002/wcm.2242.

[8] M. S. A. Howlader, M. R. Frater, and M. J. Ryan, "Estimation in underwater sensor networks taking into account capture," In Proceedings of IEEE Oceans'07, Aberdeen, Scotland, 18–21 June, 2007, pp. 1–6.

[9] S. Anower, M. R. Frater, and M. J. Ryan, "Estimation by cross-correlation of the number of nodes in underwater networks," In Proc. Australasian Telecommunication Networks and Applications Conf. (ATNAC), 10–12 Nov. 2009, pp. 1–6. doi: 10.1109/ATNAC.2009.5464716.

[10] S. A. H. Chowdhury, M. S. Anower, and J. E. Giti, "Effect of sensor number and location in cross-correlation based node estimation technique for underwater communications network," In Proc. 3rd Int. Conf. Informatics, Electronics & Vision (ICIEV), 23–24 May, 2014, Dhaka, Bangladesh, pp. 1–6. doi: 10.1109/ICIEV.2014.6850728.

[11] M. S. Anower, "Estimation using cross-correlation in a communications network," Ph.D. dissertation, SEIT, University of New South Wales at Australian Defense Force Academy, Canberra, 2011.

[12] S. A. H. Chowdhury, M. S. Anower, J. E. Giti, and M. I. Haque, "Effect of signal strength on different parameters of cross-correlation function in underwater network cardinality estimation," In Proceedings of 17th International Conference on Computer and Information Technology (ICCIT 2014), 22–23 Dec. 2014, Dhaka, Bangladesh, pp. 430–434. doi: 10.1109/ICCITechn.2014.7073159.

Introducing Spatial Orientation Sensitive Cellphone Messaging System for Blind People

Tracking A Blind People in Emergency Situation

Parijat Prashun Purohit, Fardina Fathmiul Alam
Department of Computer Science & Engineering
International University of Business Agriculture & Technology
Dhaka, Bangladesh
parijat2009@gmail.com, fardina.mist.cse09@gmail.com

S. M. Mostaq Hossain
Department of Computer Science & Engineering
Chittagong University of Engineering and Technology
Chittagong 4349, Bangladesh
mostaqcse10@gmail.com

Abstract— **It could be difficult and challenging for any person (especially blind people) to inform in an unusual situation for instant help. Falling in any critical situation for blind people is very common. One way to assist them in their unexpected situation by using an app of cell phone is proposed and described in this paper. Every smartphone consists of a good number of sensors. We have utilized accelerometer sensor's event and have introduced spatial orientation sensitive messaging system. According to our proposed system, A blind people will shake the cell phone in a specific axis so that message from cellphone will go to nearest helping hand destination. Here we have fixed a threshold and time limit so that we can avoid false messages for usual cellphone movements. Our last effort was to track the person using geodata of received message. Qualitative and quantitative performance of this system is evaluated. Moreover the entire messaging system is tested in various constraints and factors.**

Keywords—Location Based Services; Android; Spatial Orientation Sensitive Messaging (SOSM); Tracking

I. INTRODUCTION

Cellular phone brought a new dimension in our life. People with variation of ages are dependent on it to some extent. To use a cellular phone in a suitable way one requires visionary and auditory ability. Now it is a burning question whether a blind people can use a cellphone in a usual way. Few research works have been conducted to implement new technology to assist the blind people so that their life become flexible compared to inventions for normal people. Absence of visionary power doesn't let them to utilize a cellphone. If we consider cellphone to be used by blind people then there is difficulity to dial or sending messages. Previously the cellphone contained buttons for each digit as well as there were specific buttons to dial and end a phone call. There was a possibility to understand the digit buttons using fingers. Now a day's smartphones have only one touch sensitive screen and everything is inside there. So it is a little bit tiresome to use the smart phone for them. Moreover a blind people may fall in a worst situation where he/she may ask for emergency assistance. In such scenario having a smartphone makes no sense to be utilized. Voice recognition can be a way to assist blind people but there may be some difficulties to implement such system. In every smartphone there are some sensors included. In Android sensor framework many types of sensors are accessible. Most handset devices have accelerometer. Accelerometer measures the acceleration force in m/s^2 that is applied to a device on all three physical axes (x, y, and z) including the force of gravity [1]. It is commonly used for motion detection purpose. In addition this sensor is available in API level 3 platform. So, availability of this sensor is amiable to implement a sensor event.

This paper is organized in following way. In section II, we will discuss about some relevant works and focus the differences of our work. Section III consists of features, technical issues of our proposed system. Experimental Analysis will be discussed in section IV.

II. RELATED WORK

Some relevant and significant works have been done previously to assist blind or visually impaired. MyVox is introduced in [2] for communication between people: blind, deaf, deaf-blind and unimpaired. Another approach was done for communication among Blind, Deaf and Dumb people in [3]. Al Kalbani et al. represents a bus detection system [4] using RFID technology that aims to ease the traveling and movement of blind people. Zeb et al presented in [5] an indoor auditory navigation system for blind and visually impaired people using computer vision based approach. But an additional webcam is required. Using smartphone's acceleration sensor it is presented a gesture recognition system, providing users with an original, direct form of human-computer interacting way in [6]. Jabnoun et al. propose a system of visual substitution that restores a central function of the visual system which is the identification of surrounding objects [7]. In [8] a braille based mobile communication is introduced where a smart glove translates the Braille alphabet. In addition it is implemented for deaf-blind people. A remote guidance system for blind and visually impaired people via vibrotactile haptic feedback is proposed in by Scheggi et al. [9]. Anam et al presents `Expression' - an integrated assistive solution using Google Glass. The key function of the system is to enable the user in perceiving social signals during a natural dyadic conversation [10]. A design and evaluation of an auditory navigation system for blind and visually impaired is described in [11]. Majority of the works described above are related to assist someone in emergency situation. Some are specific for blind people while some or not. Moreover, some systems seem complex as well as there are simple works too. We have focused in our work some sort of new innovation and a simplified way to assist the blind people in this paper.

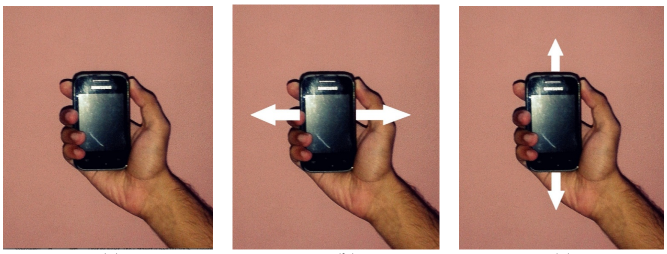
III. Description of Proposed System

This system involves an implementation of a sensor event from sender cell phone. The android application was launched initially. All the time there was a possibility of initialization of sensor event. But we have determined a threshold so that unintentionally the sensor event need not to be initialized. Even, we do not want to run the sensor event continually as there is a possibility to generate false message. So we are going to unregister our listener. When there is an acceleration crossing the threshold then variables are incremented. Within a specified time period if the variable's current value crosses a limit then we register them back on, onResume() of our activity. We have now come to the point to get and use the accelerometer events (x, y, and z). If the shaking activity doesn't cross the acceleration threshold or variables don't reach the boundary within time then the events again become unregistered and variables are initialized as zero. In fact, our target is to save the difference between the values of two consecutive events, because this is the metric that is going to show us, if we have shaken or tilted our mobile device or not. We also have to keep in our mind that this change might be almost imperceptible for our human senses, but our mobile device will calculate even the slightest difference, so that we have to set a low threshold value, that will be considered as noise, and therefore will not be calculated by our device. When the cell phone is shaken in proper axis direction if it meets as well as fulfills all conditions then the sensor event sends a message that mentions sender's required type of help, location's geo data and authentication Code

In backend there were stored some geo data along with cellphone number. When there is a successful sensor event listening, then user's current position's geo data is compared with the rest of the geo data stored in database. Message is sent to the nearest location where the distance is calculated by Haversine formula [12]. Receiver then typed the provided longitude and latitude in Google Maps search option. Then it showed the location of sender. The system is designed for one sender, many receivers.

In very beginning we made a demo application to check the acceleration values by moving the device according to three axes. The demo application contained three variables for three axes which were initialized zero. When the application was launched it showed the values of variables zero. While moving according to axis value of acceleration was changed. We actually implemented SensorEventListener in our main activity of code [13]. This class is used to receive notification from SensorManager. SensorManager accesses the device's sensors. It also gets an instance of this class by calling Context.getSystemService() with the argument SENSOR_SERVICE. Type of sensor was accelerometer.Later the SensorEvent object was created and we received some notifications when sensor values or accuracy changed. These changes were actually the value of acceleration that we received after moving it any axis. E.g.: we moved device towards x axis and we received change of value for x axis. Value of other axis remained static (zero). This made a good sense to initialize a good number of variables. These variables enabled a combination of shaking to launch a new activity.

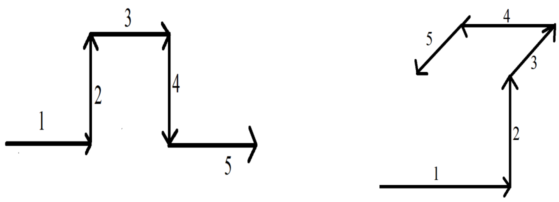
In emergency situation when the application is launched, it provides no user interface and we just need to shake according to predefined direction. The shaking directions are shown with real life example in Fig. 1(b), 1(c) for x and y axis respectively.



(a) (b) (c)

Fig. 1. (a) Cellphone in hand (b) Shaking towards x axis (c) Shaking towards y axis

For three emergency help service, we introduced different shaking combination. Someone may need emergency police help and he will shake the device in specific direction. There may rise a confusion whether message is sent or not. If message is sent a vibration will indicate confirmation of the action. When the device is shaken in a specific direction then variables are incremented each time. Some conditional statement made it possible to retrieve different phone number from database. We initially saved three contacts of emergency police help. After launching the application and shaking for specific number of times, first message was sent. Here actually variable's value was increasing with shaking and when it met an if else condition, it sent message. More shaking in same direction enabled it to send the same message to different numbers. Now there may arrive a confusion that we can only seek three emergency service as there are only three axis exists. As there are three axis and we can make some combination of shaking that can introduce more option to ask help. In Fig. 2(a) and 2(b) two shaking directions are shown that may be used to send message for different help service.



(a) (b)

Fig. 2(a) Alternative Shaking option 1 (b) Alternative Shaking option 2

For the two alternative shaking options showed in Fig. 2, if some variables are introduced then we can see the values of variables in table 1.Variation of variables enables to have more options to send message.

TABLE I: VALUES OF VARIABLES AFTER SHAKING

Variable Name	**Changing Parameter**	**Value for Fig. 6(a)**	**Value for Fig. 6(b)**
X	X axis shaking	3	2
Y	Y axis shaking	2	1
Z	Z axis shaking	0	2

Geographic coordinates are necessary to put inside the text message to track sender's location. As our platform is android, we found android.location API which is available in API level 3. Geocoder is a class for handling geocoding. The term "geocoding" is the process of transforming the description of location into geographic coordinates [14]. Specifically we used two parameters longitude and latitude for this purpose.

Finally we introduced a task that is tracing a sender's location based on his message details. Message with geo data including the latitude and longitude were received that was used later as input in Google Maps. The message that was received as testing our system is showed in Fig. 3. The latitude and longitude were received. Now we attempted to specify the location on a map directly using the geographic coordinates. In search option of Google Map we directly put the latitude and longitude and finally got the position of sender. The map position and real position were approximately accurate. We checked sender's position from different locations.

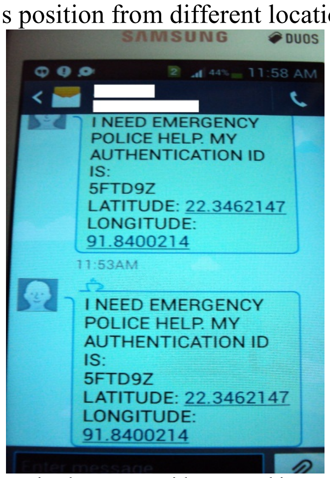


Fig. 3: Received message with geographic coordinates

If we describe our tracking system we can summarize it in a flow chart that is showed in Fig. 4.

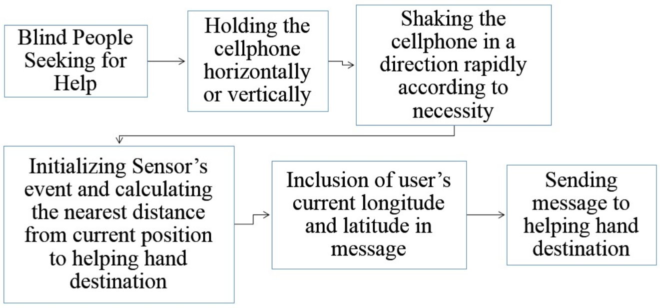


Fig. 4: Flowchart of proposed location tracking system for emergency help

IV. OVERALL EVALUATION

A. *Qualitative Evaluation:*

We've tested our system with 12 blind persons aging between 30 to 60. The feedback from them is graphically represented. After taking the feedback result, 7 persons gave good feedback, 3 persons said the system is moderately good and some modifications may be needed and 2 persons said system need more modification and need to be more user friendly.

B. *Quantitative Evaluation:*

Our messaging system was tested in a quantitative way. We have tested the messaging task in different scenario. Details of the evaluation is listed in table 2.

TABLE II: QUANTITATIVE EVALUATION OF PROPOSED MESSAGING SYSTEM

Physical Condition of Cell phone	Self generated False message	Attempted message through shaking	Successful Sent messages
Static	No	5	5
Walking	No	10	10
Running	Yes(2)	35	31
On transport	No	5	5
Cellphone dropped from hand	No	-	-
Cellphone thrown	No	-	-
Shaking with no intention	Yes(1)	-	-

Most important finding of this paper is determining the threshold acceleration for messaging. If the threshold acceleration value is low then unwanted messages are generated. Again high threshold acceleration will make it difficult to send a message as user have to shake the device with high acceleration. A bar chart in Fig. 5 will show our findings.

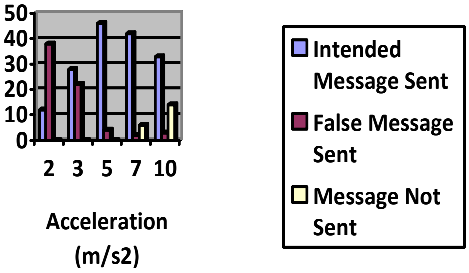


Fig. 5: Relationship between acceleration and sent messages

From above figure it is obvious when the acceleration value is low then various false messages were generated. When it is too high then it was not responsive. So we found an optimum value that is 5 m/s2. Out of 50 messages 46 intended messages were sent during spatial orientation change. Only 4 false messages were generated. Details of the bar chart is shown in Table III.

TABLE III: ACCELERATION AND SENT MESSAGE RELATIONSHIP

	Acceleration (m/s2)				
	2	3	5	7	10
Intended Message Sent	12	28	46	42	33
False Message Sent	38	22	4	2	3
Message Not Sent	0	0	0	6	14

V. CONCLUSION

The challenges faced by blind people in their everyday lives are not expected as they are also a part of our human kind. The most challenging portion is to ask for help in a situation where it is not possible to call or send a message through usual texting via mobile device. This paper represented a taxonomy of the types of questions asked, reported the survey result based on the answers and discussed about how to give a solution to blind people in their critical situation. These results improve our understanding of the problems that blind people face and motivated us to develop an app. We have developed the app based on sensors used in modern cell devices as they had such potential to implement an efficient way for sending message with a view to seeking help. A map is enough to find out one's exact position within short time using geographic coordinates. So, responding will also be possible as soon as the seeker wants. Our tracking system is simpler one to be used in these purposes. Our system is only limited to send geodata to a destination. In future we have plan to include more sensor events along with new features in our application so that blind people can utilize the cell phone as normal people.

REFERENCES

[1] Sensors Overview (July 2015) [Online]. Available:http://developer.android.com/guide/topics/sensors/sensors_overview.html

[2] Ramirez-Garibay, F.; Millan Olivarria, C.; Eufracio Aguilera, A.F.; Huegel, J.C., "MyVox—Device for the communication between people: blind, deaf, deaf-blind and unimpaired," Global Humanitarian Technology Conference (GHTC), 2014 IEEE , vol., no., pp.506,509, 10-13 Oct. 2014

[3] Rastogi, R.; Mittal, S.; Agarwal, S., "A novel approach for communication among Blind, Deaf and Dumb people," Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on , vol., no., pp.605,610, 11-13 March 2015

[4] Al Kalbani, J.; Suwailam, R.B.; Al Yafai, A.; Al Abri, D.; Awadalla, M., "Bus detection system for blind people using RFID," GCC Conference and Exhibition (GCCCE), 2015 IEEE 8th , vol., no., pp.1,6, 1-4 Feb. 2015

[5] Zeb, A.; Ullah, S.; Rabbi, I., "Indoor vision-based auditory assistance for blind people in semi controlled environments," Image Processing Theory, Tools and Applications (IPTA), 2014 4th International Conference on , vol., no., pp.1,6, 14-17 Oct. 2014

[6] Yaodong Huang; Zhiyuan Xu; Ruijin Wang; Dajiang Chen, "GRIB: Gesture Recognition Interaction with Mobile Devices for Blind People," Computer and Information Technology (CIT), 2014 IEEE International Conference on , vol., no., pp.604,609, 11-13 Sept. 2014

[7] Jabnoun, H.; Benzarti, F.; Amiri, H., "Visual substitution system for blind people based on SIFT description," Soft Computing and Pattern Recognition (SoCPaR), 2014 6th International Conference of , vol., no., pp.300,305, 11-14 Aug. 2014

[8] Choudhary, T.; Kulkarni, S.; Reddy, P., "A Braille-based mobile communication and translation glove for deaf-blind people," Pervasive Computing (ICPC), 2015 International Conference on , vol., no., pp.1,4, 8-10 Jan. 2015

[9] Scheggi, S.; Talarico, A.; Prattichizzo, D., "A remote guidance system for blind and visually impaired people via vibrotactile haptic feedback," Control and Automation (MED), 2014 22nd Mediterranean Conference of , vol., no., pp.20,23, 16-19 June 2014

[10] Anam, A.I.; Alam, S.; Yeasin, M., "Expression: A dyadic conversation aid using Google Glass for people who are blind or visually impaired," Mobile Computing, Applications and Services (MobiCASE), 2014 6th International Conference on , vol., no., pp.57,64, 6-7 Nov. 2014

[11] Chun-Hung Yang; Sheue-Ling Hwang; Jan-Li Wang, "The design and evaluation of an auditory navigation system for blind and visually impaired," Computer Supported Cooperative Work in Design (CSCWD), Proceedings of the 2014 IEEE 18th International Conference on , vol., no., pp.342,345, 21-23 May 2014

[12] Haversine Formula (July 2015) [Online]. Available: https://en.wikipedia.org/wiki/Haversine_formula

[13] SensorEventListener (January 2015) [Online]. Available:http://developer.android.com/reference/android/hardware/SensorEventListener.html

[14] Geocoder (January 2015) [Online]. Available:http://developer.android.com/reference/android/location/Geocoder. html

A Real-Time Face to Camera Distance Measurement Algorithm Using Object Classification

Hossain, Md Asgar, M.Sc
Electrical Communication Engineering
University Kassel
34109 Kassel,Germany
Hossain.a.md@ieee.org

Md Mukit
Electrical and Electronic Engineering
World University of Bangladesh (WUB)
Dhaka, Bangladesh
smmukit@outlook.com

Abstract—**In a human and computer interaction based system, distance estimation by computer vision between camera and human face is a vital operation. To calculate the distance between the camera and face, an estimation method based on feature detection is proposed in this paper, where detection of eyes, face and iris in an image sequence is described. From the estimated iris and the distance between the centroid of the iris, an algorithm is proposed to determine the distance from camera to face. An architecture for face detection based system on AdaBoost algorithm using Haar features and Canny and Hugo Transform for edge and circular iris estimation is presented here. Wrongly detected faces are removed by analyzing the disparity map. From the estimated face, canny and Hugo transform is used to determine the iris, and to calculate the distance between the centroid of iris. Later Pythagoras and similarity of triangles are used for distance estimation. The implementation is done in C++ using Intel OpenCV image processing libraries to reduce system overhead.**

Keywords—Haar Classifier; AdaBoost; Estimation; Detection; Hugo Transform; Canny Transform; OprnCV

I. INTRODUCTION

Dynamic development in smart machine vision, robotics, vehicle guidance and other computer vision based applications, distance measurement for computer vision is very important component of modern smart, dynamic and autonomous systems. In a Human and computer interaction based system, distance estimation for computer vision between camera and human face is vital. There have been very limited researches on single camera based distance estimation, which will be faster, reliable and can be used on real time with less computational burden. A 3-D position estimation for human face was proposed on the paper "Face distance estimation from a monocular camera", where location of human's head and face was determined by using motion detection, Hough transform and a statistical color model [4]. Changing pan, tilt and zoom of camera, the face is put on the center of the cameras field of view. Then to measure the distance between human face and an autofocus camera, information taken from focusing the ring are used. A novel method is emphasized in [1] to use devices with monocular camera to determine the depth between the user and the front camera using a back propagation neural network (BPNN). This depth is successively used to calculate the zooming factor for a legible view and to read a document on the display of the mobile device. It is proposed to use frontal facial features acquired from the monocular camera to find the depth information with the use of supervised learning algorithm. An image processing algorithm has been proposed by [2] to measure the distance from the background. A single camera is fixed at a stationary position to capture the real time image of targeted object and determine its distance in contrast to existing and most common vision algorithms of stereo vision. The proposed method is a statistical method. A two-camera system was proposed in [3] to detect the face from a fixed, wide-angle camera, which estimates a rough location for the eye region using an eye detector based on topographic features, and directs another active pan-tilt-zoom camera to focus in on this eye region. In [5], faces are detected in 2D images with a rapid object classifier based on haar-like features and principle component analysis to create an Eigen space. These algorithms are limited to the computational burden and hard to realize in real time. Moreover some of them need extra feature from the cameras. Therefore, we are proposing a new technique to measure the distance from camera based on iris estimation and the distance between them.

In Section II, the human face estimation including the Haar Classifier, Integral Image and Training algorithm is described. In section III, the iris estimation algorithm is presented. TO measure the distance between the face and camera has been described in Section IV. Finally, section V gives the outline of the implementation of the algorithm and a developed software.

II. HUMAN FACE ESTIMATION

The proposed algorithm is based on iris detection. The detection part has been performed by multi stage based processing, which are (i) Estimating the face from image, (ii) Getting the eye of human face, and finally (iii) Iris detection. Inclusion of iris and eye visibility are the prior components of the algorithm of this research. If it fails to find any iris, or it gets both the iris, it will not work. Primarily the paper explains the algorithm to estimate the face from the captured image, where Haar classifier along with Adabost training algorithm is used. In the next, it describes the Hugo and canny transform to identify the iris. Finally in the last stage, the algorithm which will use the distance between the iris will be used to measure the distance between the camera and human face.

A. Haar Classifier

Haar classifier based object detection is based on Haar-like features. Instead of using the intensity of values of a pixel, Haar classifier uses the change in contrast values between adjacent rectangular group of pixel, which makes it much faster for real time detection. The contrast variance between the pixel group is used to determine relative light and dark areas. The value of a two-rectangular feature is the difference between the sum of the pixels within two rectangular regions. The value of any given feature is always the sum of the pixels within clear rectangles subtracted from the sum of the pixels within shaded rectangles.

B. Integral Image

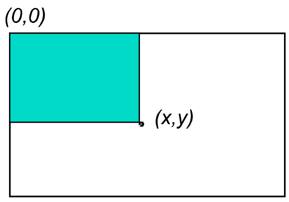
The integral image is an array containing the sums of the pixels' intensity values located directly to the left of a pixel and directly above the pixel at the location (x, y) .

Let, $I(x', y')$ is the original image, and $II(x, y)$ is the integral image. The integral iamge is computed as shown below,

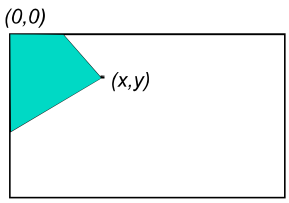
$$II(x,y)=\sum_{x'<x,y'<y} I(x',y') \tag{1}$$

The rotated integral image is calculated by finding the sum of the pixels' intensity values that are located at the forty five degree angle to the left and above for the x value and below for the y value. According to [6], if the rotated image is $IR(x, y)$ and $I(x', y')$ is the original image then,

$$IR(x,y)=\sum_{x'<x,x'<x-|y-y'|} I(x',y') \tag{2}$$



(a) Summed area of horizontal image



(b) Summed area of rotated image

Fig. 1: Integral image

C. Training Algorithm

In this system, AdaBoost is used to select and train the classifier as prescribed in Viola & Jones framework [7]. AdaBoost learning algorithm was to boost the classification performance of a simple training algorithm. It combines a collection of weak classification functions and forms a stronger classifier. A weak classifier $(h(x, f, p, \theta))$ consists of a feature (f), a threshold (θ) and polarity (p) indicating the direction of inequality. The weak classifier, used as a threshold single feature, was also used in [7]

$$h(X)=sign\sum_{j=1}^{M}\alpha_j h_j(X) \tag{3}$$

where each weak classifier is a threshold function based on the feature (f_j).

$$X(\omega) = \begin{cases} -s_j & \text{if } f_j < \theta_j \\ s_j & \text{otherwise} \end{cases} \tag{4}$$

The linear combination in equation [3] forms a strong classifier. A simplified version of learning algorithm is presented in [8].

III. Iris Estimation

Estimating the human face, Canny and Hugo transform was used to estimate the iris. This step consists of two parts, Canny edge detection for estimating the edge of the iris, and Hugo transform to determine the circular pattern. Canny edge detection is a multistage edge detection algorithm. Edge detection is suspectual to noise, therefore, it uses a Gaussian filter for noise reduction. Smoothened image is then filtered with a Sobel kernel in both horizontal and vertical direction to get first derivative in horizontal direction (Gx) and vertical direction (Gy) respectivel. Edge gradient and direction for each pixel can be found from these two images. Gradient direction is always perpendicular to edges. It is rounded to one of four angles representing vertical, horizontal and two diagonal directions. Every pixel, pixel is checked if it is a local maximum in its neighborhood in the direction of gradient. The result is a binary image with "thin edges". This stage decides whether the edges are really edges or not. For this operation, two threshold values are needed- (i) minimum value, and (ii) maximum value. Any edges with intensity gradient more than maximum value are sure to be edges and those below minimum value are sure to be non-edges, and so discarded. Those who lie between these two thresholds are classified edges or non-edges based on their connectivity.

Some data points are there in an image, which are the result of canny edge detection process.

A curve can be detected if it can be expressed as a function of the following form

$$f(a_1, a_2, \ldots, a_n, x, y) = 0.$$

a circle can be represented as

$$(x - a)^2 + (y - b)^2 - r^2 = 0 \tag{5}$$

the parametric equations for a circle in polar coordinates are $x = a + r\cos q$ and $y = b + r\sin q$. Solving for the parameters of the circle, we obtain the equations $a = x - r\cos q$ and $b = y - r\sin q$. Now if the gradient angle q is given at edge point (x, y), then $\cos q$ and $\sin q$ can be computed. These quantities are available as a by-product of edge detection. We can eliminate the radius from the pair of equations above to yield $b = a\tan q - x\tan q + y$.

IV. Distance Measurement

Let us consider that at the calibration stage the centroid of the two estimated iris are $A'(x_1, y_1)$ and $B'(x_2, y_2)$. The estimated distance between the two iris is,

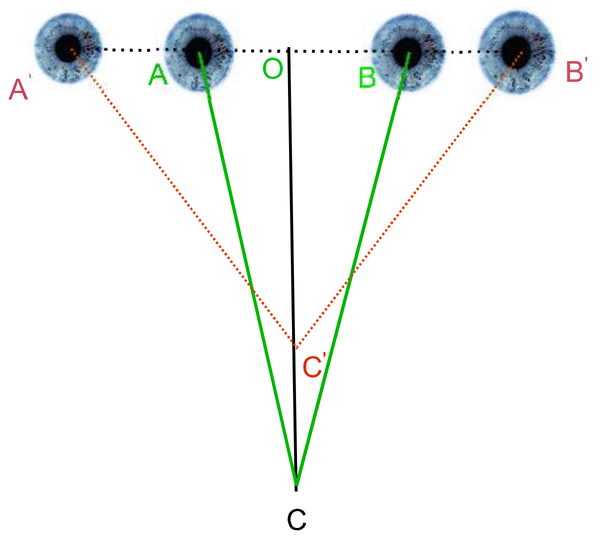


Fig. 2: Proposed algorithm

$$A'B' = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{6}$$

If the middle point of the two centroid is O, then the distance from one iris to the middle point is $A'O = A'B'/2$, and from the initial calibration, the distance between the face and the camera is OC'. From figure 2, for the triangle ΔAOC, the base AO and height OC is calculated.

After calibration, the new position of the human is at C and new estimated position of the centroid of the iris are A and B. Similarly, the distance between the irish is AB. The distance between the middle point O and human face is OC is unknown. Here, AO and OC are inversely proportional,

$$OC \propto \frac{1}{AO} \tag{7}$$

From the diagram it is seen that, ΔA^iOC^i and ΔAOC are similar triangle. By using ΔA^iOC^i and ΔAOC, and the inverse relation (7) between the distance of two iris and the distance between the face and camera we can get,

$$OC = \frac{A'O}{AO} * OC' \tag{8}$$

Here, AO is half of the distance between two iris centroid and OC is the distance between human face and camera during the calibration. Estimating the distance between the iris centroid A^iB^i can be found. Then using the similarity property, the distance between camera and human face can be calculated, as described in equation 8 .

V. Implemantation

The iris image should be rich in iris texture as the feature extraction stage depends upon the image quality. To verify and check the dependency of implemented algorithms for face and iris detection, both high resolution and low resolution images are used here. For the experimental setup, a built in 2MP camera from the laptop and an external 720 P HD web cam were used. The approximate distance between the human face and the camera started from 10 cm. The image acquisition setup is given in Figure 6. To interface this implemented software directly to the webcam *JMF* tool was used [11]

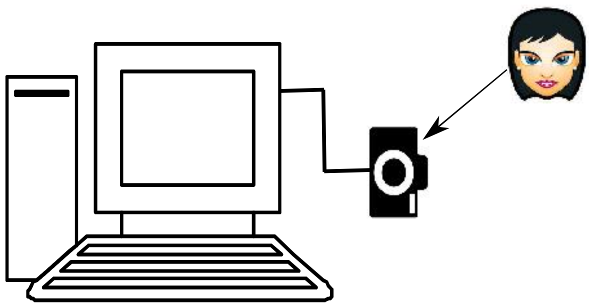


Fig. 3: Image acquisation

Implementaion steps of proposed algorithm are:

1) Initially the software gets connected with the web-cam, reads the input image as described in V and takes the snap shot from live video. It is possible to select the video or image resolution based on the requirement.
2) In the next step, it removes noise using Gaussian filter.
3) Then threshold is applied based on the range of RGB pixels. If pixels value is greater than threshold value, then the background is selected and threshold value is 0 otherwise value is 1.
4) After this preprocessing of image, Harr-like classifier is used with Adboost algorithm to estimate human face. If the human face is absent then it requests for another frame form the video stream. If it estimates a face, it passes the image to the next step.
5) From the estimated face area, canny edge detection algorithm gets the edges of the objects. Afterwards, Hugo transform gets the circular iris. If it does not have both the iris or if eye is not present, it discard the image and asks for another instant.
6) In this step, it finds the position of the centroid of the iris and radius of the circle, and calculates the distance between the iris.
7) Finally the estimated distance between two iris is passed to the algorithm described in IV.

VI. Conclusion

In this paper, a new approach for face to camera distance measurment has been proposed. The estimation method is simple but faster and gives reliable and accurate result for real-time application. With AdaBoost algorithm, the classifier

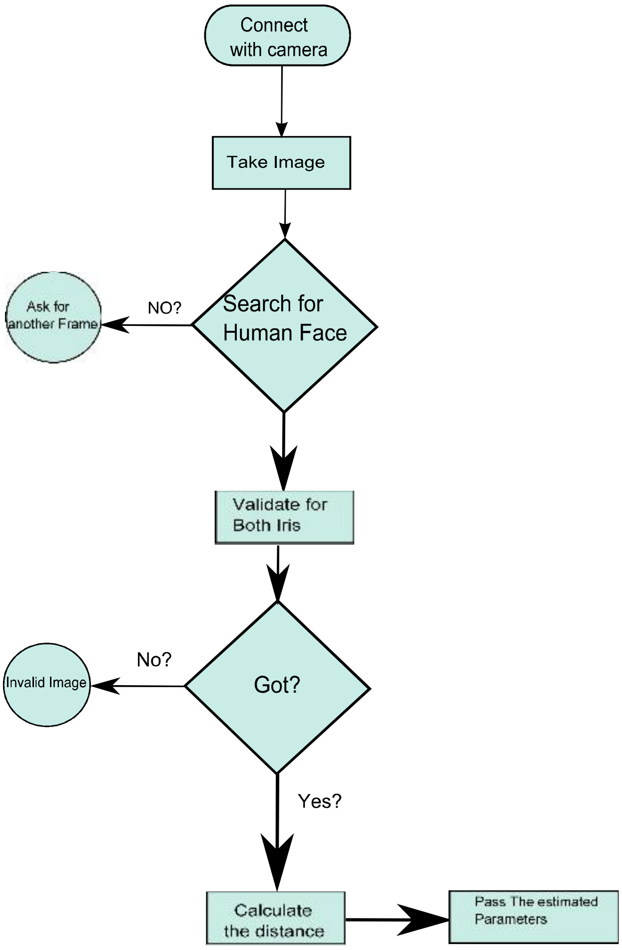


Fig. 4: Work flow

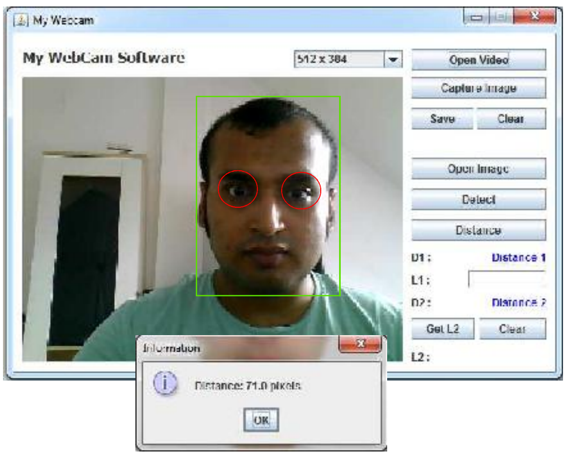


Fig. 5: Initial image acquisition and calibration for a specific user

estimates the face and Hugo transform with canny edge detection gives the estimated eye. The classifier is a onetime factory implement. Therefor, the approach is fast enough for

Fig. 6: Give the distances from camera to human face

real-time application and reduces computational burden. From the estimated iris centroid and the distance between them, the proposed method can estimate the distance from camera to face up to 40 cm. After that, the linearity relation could not be maintained and the error increases. The future work should focus on this. The specific distance and system adaptive experiment verify the feasibility for real-time application.

REFERENCES

[1] Reale, M.; Hung, T.; Lijun Yin, "Pointing with the eyes: Gaze estimation using a static/active camera system and 3D iris disk model," Multimedia and Expo (ICME), 2010 IEEE International Conference on , vol., no., pp.280,285, 19-23 July 2010 doi: 10.1109/ICME.2010.5583014

[2] Ghidary, S.S.; Nakata, Y.; Takamori, T.; Hattori, M., "Human detection and localization at indoor environment by home robot," Systems, Man, and Cybernetics, 2000 IEEE International Conference on , vol.2, no., pp.1360,1365 vol.2, 2000 doi: 10.1109/ICSMC.2000.886043

[3] Kosov, S.; Scherbaum, K.; Faber, K.; Thormahlen, T.; Seidel, H.-P., "Rapid stereo-vision enhanced face detection," Image Processing (ICIP), 2009 16th IEEE International Conference on , vol., no., pp.1221,1224, 7-10 Nov. 2009 doi: 10.1109/ICIP.2009.5413675

[4] Kumar, M.S.S.; Vimala, K.S.; Avinash, N., "Face distance estimation from a monocular camera," Image Processing (ICIP), 2013 20th IEEE International Conference on , vol., no., pp.3532,3536, 15-18 Sept. 2013 doi: 10.1109/ICIP.2013.6738729

[5] Ali, A.M., "A 3D-Based Pose Invariant Face Recognition at a Distance Framework," Information Forensics and Security, IEEE Transactions on , vol.9, no.12, pp.2158,2169, Dec. 2014 doi: 10.1109/TIFS.2014.2362299

[6] Mita, T.; Kaneko, T.; Hori, O., "Joint Haar-like features for face detection," Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on , vol.2, no., pp.1619,1626 Vol. 2, 17-21 Oct. 2005 doi: 10.1109/ICCV.2005.129

[7] Viola, P.; Jones, M., "Robust real-time face detection," Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on , vol.2, no., pp.747,747, 2001 doi: 10.1109/ICCV.2001.937709

[8] Richard Szeliski. 2010. Computer Vision: Algorithms and Applications (1st ed.). Springer-Verlag New York, Inc., New York, NY, USA.

[9] Dimitrios Ioannoua, Walter Hudab, Andrew F. Laine, "Circle recognition through a 2D Hough Transform and radius histogramming" Image and Vision Computing 17 (1999) 1526, Elsevier Science

[10] Naveen Singh, Dilip Gandhi, Krishna Pal Singh, "Iris recognition system using a Canny edge detection and a circular Hough Transfrom ", International Journal of Advances in Engineering & Technology, May 2011, ISSN: 2231-1963

[11] JMF developed by ORACLE,officile website, http://www.oracle.com/technetwork/articles/javase/index-jsp-140239.html

Smartphone based Teacher-Student Interaction Enhancement System

Saad Bin Sajjad
Dept. of Computer Science and Engineering
Chittagong University of Engg. & Tech. (CUET)
Chittagong-4349, Bangladesh
saad2617@gmail.com

Ahmed Imteaj
Dept. of Computer Science and Engineering
Chittagong University of Engg. & Tech. (CUET)
Chittagong-4349, Bangladesh
imtu_1992@yahoo.com

***Abstract*— Smartphone applications have found their way into almost every avenue in our lives. This research aims to develop a mobile application based on teacher-student interaction system. As the efficiency and success of an academic class is somewhat dependent on the level and scope of teacher-student interaction in class, this paper emphasizes on how our mobile application can effectively contribute in increasing interaction and provide the necessary feedback to make the present and future classes better ones.**

Keywords— Teacher student interaction; mobile application; feedback.

I. INTRODUCTION

The success of the education imparted in a classroom depends on the knowledge and delivery of the teacher and the attention of the students and last but not the least the interaction among the teacher and the students.

In this age of computers, almost every sector has seen increased performance employing computer technology. Likewise, this technology can be used to improve the efficiency of a classroom, and one of the ways to do it is to employ computer technologies for increasing the interaction among teachers and students in a classroom.

Many previous researches have proved the fact that employing computer technology in classrooms significantly improved the performance of students. The comparison between two classes in 2004 at Wichita State University, one being a traditional class and the other using traditional materials as well as Blackboard™, a web-based course management software which focused on the teacher student interaction, revealed that the performance of the class using the software was much better [1].

Another research [2] done at MIT in 2007 compared two classes, one being a traditional one and the other employing Tablet PCs and a software called Classroom Learning Partner(CLP) which is based on Classroom Presenter [3]. The results proved that the classroom employing Tablet PCs and CLP was significantly ahead of the traditional class in terms of performance.

So keeping in mind the improved performances of classes using technology for enhanced classroom interaction, we propose a smartphone based solution for increasing the scope and quality of Teacher-Student interaction in class. This application will not need the help of any external database and all the features will be real time.

The rest of this paper is organized as follows. Section II describes the background studies of technology related attempts to help interaction system between teacher and student more efficient. The overall system model and implementation result of the application is presented in Section III. Section IV states the verification of the proposed application and experimental results. Section V provides the conclusion and future scope of the application.

II. RELATED WORKS

Instructors at University of Illinois Chicago(UIC) have devised a method to use cell phones for creating a dynamic classroom with real time interaction employing QR codes and free online polls [4].

Another report encompasses the positive aspects of using smartphones for interactions in classrooms and provides an introduction to a free resource called Study Boost to enhance the level of interaction in classrooms and thus increase the efficiency of a class [5].

Hotseat [6], a software developed at Purdue University facilitates backchannel discussion among students in a classroom or a lecture hall and provides a means of providing near real time feedback. It basically utilizes social networking sites such as Facebook, Twitter and Text Messaging and works on a multiple range of devices from laptops to mobile phones.

III. SYSTEM DESCRIPTION

A. Overview of model

The Entire system will consist of a Smartphone application that will again be divided into two modules, the teacher module and the student module. So the module is to be selected on the basis of the role of the user in the class. There will be also options for publishing results of class tests, quiz, viva and final grade sheet. The notice board feature will help a teacher to send his messages to the students. For evaluating daily class performance, a teacher will be able to ask MCQ (where 4 options will be given) or Short Question (a space will be given to submit written answer). And registered students will respond here with their personal smart phones from this application.

A feature will be given to them to give feedback of the class where they can hide their ID if they want. Whenever the teacher asks MCQ/SQ, they will be able to give answers within a given time. Another attracting feature of our application is taking class attendance. We use Wi-Fi technology for making this feature fast and easy where a teacher can create a session which will be protected by password and will be given only in the classroom. Students present in the class can join that session from their phones with the password provided by the teacher. Successful joining in session will be counted as an attendance and saved by sending it as an email to the teacher's inbox. There will mainly be two parts within the application, one for the teacher and the other for the student. The system model is displayed in fig. 1 and in fig. 2 we have showed the two main modules of the system.

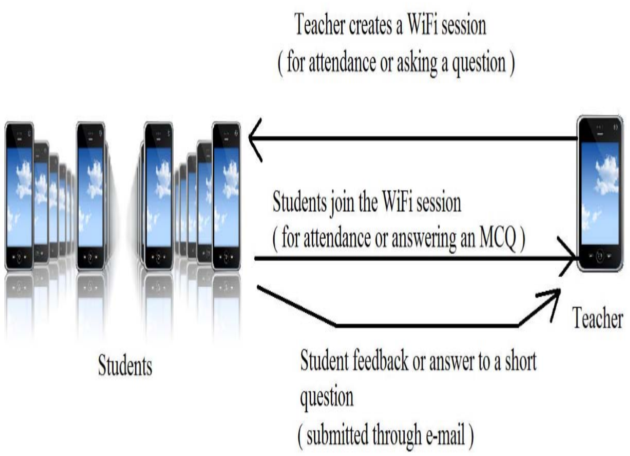


Fig. 1: Proposed System Model

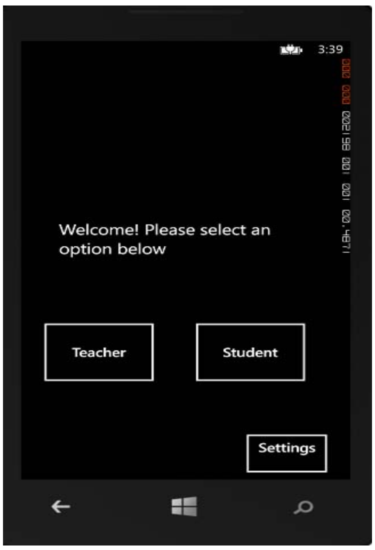


Fig. 2: Home screen

B. Attendance Module

The session will be created using Wi-Fi with a unique username and password and the teacher's email address. The teacher will then announce the username and password in the class so the students in the class, all of whom should have a smart phone with the proposed application installed, can turn on their Wi-Fi and join the session created. The students will have to enter their roll number and a username at their end of the application. The teacher will then be able to see the number of students logged in and thus this feature will enable the teacher to check the attendance in the class without name or roll calling. Once the attendance is counted, the session may be ended. The class attendance module is presented in fig. 3.

Fig. 3: Class attendance interface

C. Question Module

The teacher module will also have the feature of creating MCQs or short questions to be answered by the students in the class. The teacher will have to create another Wi-Fi session and invite the students to join in. Then he/she must select whether the question will be an MCQ or a short question. In case of MCQ, the teacher has to post the question along with 4 options and specify the correct answer. Then the question may be posted, of which the students only see the question with the four options and submit their answers. The teacher will see the active students on his end of the application, and whenever a student will answer the question, a tick mark will appear beside the student name if the answer is correct and a cross sign if it's wrong. In the case of a short question, the teacher only has to provide the question. Students will then write a short answer and submit it, which the teacher will see when he taps on the student name from the list of students.

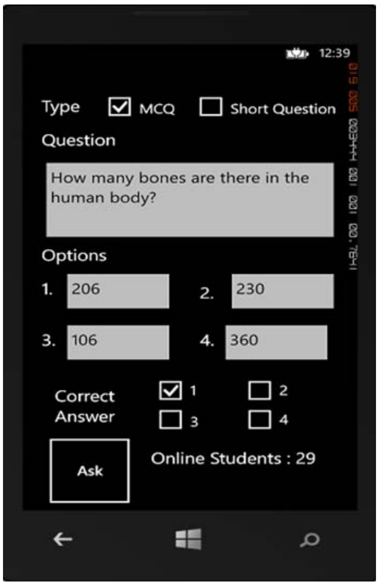


Fig. 4: Question interface of the system

D. *Providing results*

The results will be provided in two ways. First of all, the MCQ questions when asked, the students join the session to view the question and submit an answer. The teacher then may either announce the answer or write it down on the board. Apart from this, the teacher will also have an option to transfer the information to a particular student, whether his/her answer is right or wrong after receiving an answer. The MCQ is currently limited to four options for an answer, so the correct answer will be any one of the four. For the answer to the short questions, after being asked in the class via Wi-Fi session, the student may type the answer and deliver it to the e-mail address of the teacher with the option provided in his/her module of the application. The teacher may check his/her mail and reply accordingly via e-mail.

E. *Feedback Module*

There will be an option for providing feedback from the students. The feedback options will have ratings for the class, a space for writing feedback and suggestions. This feedback will be automatically converted into an email and delivered to the mailbox of the teacher using the address provided at the beginning of the class.

IV. EXPERIMENTED RESULT

After implementing our application, we have craved to perceive user feedback and their persuasion. We have collected feedback from ages of different people. We suggested them to rate this application, whether it is useful or not. 66% users rated this application as very useful. They added that if this system is integrated into the institution, then the educational system would be more digitalized. In fig. 5 we have exhibited the feedback of the users graphically.

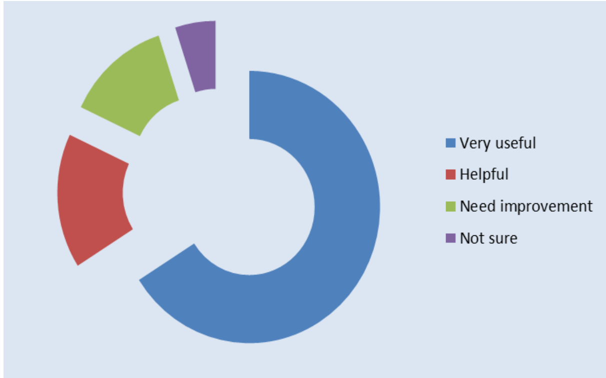


Fig. 5: Feedback of the user

V. CONCLUSION

This paper presents the details of a smart phone application to enhance the productivity and effectiveness of a classroom by enhancing the interaction among the teacher and students. The application thrives on existing established technologies email and Wi-Fi. Further, the application is designed keeping in mind the wide influx and influence of Smart devices such as phones and tabs among this generation of students and teachers, and the immense positive feedback from classes using similar technologies. The application is also designed to be user friendly and intuitive so that there is not much difficulty in operating it among any sort of class. Furthermore, the application opens a new door for collaboration and cooperation among the teacher and students, which aims at an optimal utilization of class time and simple technology to enhance the learning and teaching experience.

REFERENCES

[1] C.Gilmore & G.C.Halcomb,"Technology in the classroom: Investigating the effect on the student-teacher interaction," Software Usability Research Laboratory, July 12, 2004.
http://psychology.wichita.edu/newsurl/usabilitynews/62/online_stats.asp

[2] Koile, Kimberle, and David Singer, "Assessing the impact of a tablet-pc-based classroom interaction system," The Impact of Tablet PCs and Pen-based Technology on Education. Evidence and Outcomes, pp.73-80, 2008.

[3] R. Anderson,"Tablet PC Based Lecture Presentation System in Computer Science Courses", in Proceedings of Special Interest Group on Computer Science Education (SIGCSE), 3-7 March, 2004.

[4] Abigail McMillin, LCSL, "Mobile Technology & Interaction: Using Smartphones to Increase Real Time Engagement in the Classroom," TechTeach,2014. http://2014.techteach.uic.edu/sessions/mobile-technology-interaction-using-smartphones-to-increase-real-time-engagement-in-the-classroom/

[5] M. Zack O'Neill and Grant Legaul, "The Smartphone : The technology of the FUTURE!" unpublished.
http://theegoogle.weebly.com/argumentative-paper-the-smartphone-technology-of-the-future.html

[6] Aagard, Bowen and Olesova, "Hotseat: Opening the Backchannel in Large Lectures," Educause Quarterly, Vol. 33, November 3, 2010.
http://www.educause.edu/EDUCAUSE+Quarterly/EDUCAUSEQuarterlyMagazineVolum/HotseatOpeningtheBackchannelin/213668

Blind Audio Watermarking Based on Fast Walsh-Hadamard Transform and LU Decomposition

Pranab Kumar Dhar
Department of Computer Science and Engineering,
Chittagong University of Engineering and Technology (CUET)
Chittagong-4349, Bangladesh.
pranabdhar81@gmail.com

***Abstract*—Digital watermarking is identified as a major technique for copyright protection of multimedia contents. This paper presents an audio watermarking algorithm based on fast Walsh-Hadamard transform (FWHT) and LU decomposition (LUD). Firstly, we preprocess the watermark data to enhance the security of the proposed algorithm. Then, the original audio is segmented into non-overlapping frames and FWHT is applied to each frame. LUD is applied to the FWHT coefficients represented in a matrix form. Watermark data is embedded into the largest element of the upper triangular matrix obtained from the FWHT coefficients of each frame. Experimental results indicate that proposed algorithm is considerably robust and reliable against various attacks without degrading the quality of the watermarked audio. Moreover, it shows more excellent results than the state-of-the-art methods in terms of imperceptibility, robustness, and data payload.**

Keywords-copyright protection; multimedia contents; Fast Walsh-Hadamard transform; LU decomposition.

I. INTRODUCTION

The growth of the Internet, availability of state-of-the-art digital media production and editing technologies and development of efficient multimedia data compression schemes have led to unauthorized sharing of multimedia content. It is therefore necessary to use robust technologies for protecting copyrighted digital media from illegal sharing and tampering. One of the ways to protect digital data and proof of ownership is digital watermarking. It is the process of embedding digital information known as watermark into any multimedia content such as image, audio or video for authenticity. This technique has several applications such as content authentication, copyright protection, data indexing, broadcast monitoring, fingerprinting, medical safety, and so on. In recent years, a significant number of audio watermarking techniques have been reported. A comprehensive survey on audio watermarking can be found in [1]-[2]. Most audio watermarking methods utilize either a time domain [3]-[4] or a transform domain such as discrete wavelet transform (DWT) [5]-[7], lifting wavelet transform (LWT) [8], and fast Fourier transform (FFT) [9]. Time domain methods are very efficient and easy to implement, however, transform domain methods can provide high robustness. Lie and Chang [3] introduced a method in which group amplitudes are modified to achieve high robustness. Bassia *et al.* [4] presented a watermarking technique in which watermark bits are embedded by modifying the audio samples directly. However, both methods have low data payload. In [5], authors presented an adaptive method using wavelet based entropy, but robustness to re-sampling and low-pass filtering attacks are quite low. Chen *et al.* [6] proposed an algorithm that embeds watermark information by energy-proportion scheme. However, the signal-to-noise ratio (SNR) results of this algorithm are not satisfactory. In [7], authors introduced an optimization-based watermarking scheme which embeds watermark in the lowest-frequency coefficients of DWT. However, the subjective evaluation of watermarked audio signals has not been conducted in this scheme. Erçelebi and Batakçı [8] proposed a watermarking method based on LWT in which a binary image is embedded as watermark. However, from the reported result, robustness to attacks of this method is quite low. Megías *et al.* [9] suggested a watermarking method that embeds watermark in FFT domain, but it has low data payload. Recently, the singular value decomposition (SVD) has been used as an effective technique in digital watermarking [10]-[14]. Dhar *et al.* [10] proposed an efficient audio watermarking algorithm in transform domain based on SVD and Cartesian-polar transform (CPT). However, the detection scheme is non-blind and robustness needs further improvement. In [11], authors proposed a blind SVD based method using entropy and log-polar transform (LPT). However, robustness against re-sampling attack is little low. The methods proposed by Lei *et al.* [12] and Bhat *et al.* [13] provide high robustness against attack, however the data payload of these methods is quite low. Moreover, some other techniques such as empirical mode decomposition (EMD) [14], time spread (TS) echo method [15], and audio histogram [16] techniques are becoming popular in audio watermarking field. The main limitation of the existing audio watermarking techniques is the difficulty to obtain a favorable trade-off among imperceptibility, robustness, and data payload. To overcome this limitation, in this paper, we propose a blind audio watermarking algorithm based on Fast Walsh-Hadamard Transform (FWHT) and LU decomposition (LUD). The main features of the proposed scheme include (*i*) it utilizes the FWHT, LUD, and quantization jointly, (*ii*) it uses a tent map which contains the chaotic characteristic to enhance the confidentiality of the proposed algorithm, (*iii*) watermark is embedded into the largest element of the upper triangular matrix obtained from the FWHT coefficients of each frame by quantization, (*iv*) watermark extraction process is blind, and (*v*) it achieves a good trade-off among imperceptibility, robustness, and data payload. Experimental results indicate that the proposed watermarking algorithm shows high robustness against various

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

attacks such as noise addition, cropping, re-sampling, re-quantization, and MP3 compression. Moreover, it outperforms state-of-the-art methods [5]-[6], [8]-[9], [11]-[12], [14]-[16] in terms of imperceptibility, robustness, and data payload.

The rest of this paper is organized as follows. Background information including FWHT and LUD is described in section II. The proposed watermarking algorithm including watermark preprocessing, watermark embedding process, and watermark detection process is presented in section III. Experimental results of the proposed algorithm are discussed in section IV. Finally, section V concludes this paper.

II. Background Information

A. Fast Walsh-Hadamard Transform

The FWHT is widely used in various applications, such as image processing, signal processing, filtering, and so on. The forward and inverse FWHT can be defined as a linear combination of a set of square waves of different frequencies and are represented by the following equations:

$$\begin{aligned} X_w(k) &= \sum_{n=0}^{N-1} x(n) w_N(k,n) \\ &= \sum_{n=0}^{N-1} x(n) \prod_{i=0}^{M-1} (-1)^{n_i k_{M-1-i}},\ k = 0,1,.....N-1 \end{aligned} \tag{1}$$

$$\begin{aligned} x(n) &= \frac{1}{N} \sum_{n=0}^{N-1} X_w(k) w_N(k,n) \\ &= \frac{1}{N} \sum_{n=0}^{N-1} X_w(k) \prod_{i=0}^{M-1} (-1)^{n_i k_{M-1-i}},\ n = 0,1,.....N-1 \end{aligned} \tag{2}$$

where $x(n)$ is the input signal in time domain, $X_w(k)$ is the transformed signal, $w_N(k,n)$ is the walsh function, $N=2^n$, $M=log_2N$, and n_i is the i-th bit in the binary representation of n.

B. LU Decomposition

Any square matrix A can be decomposed in lower triangular matrix L and upper triangular matrix U. LUD for a 3×3 matrix can be written as follows:

$$A = L \times U = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \times \begin{bmatrix} d_1 & u_{12} & u_{13} \\ 0 & d_2 & u_{23} \\ 0 & 0 & d_3 \end{bmatrix} \tag{3}$$

The diagonal elements of matrix L are 1s, on the other hand, the diagonal elements of U matrix are not 1's, indicating that the LU form is not symmetric. Therefore, in this study, we divide U matrix into a diagonal matrix D and a modified lower triangular matrix U', to make LDU$'$ form symmetric, shown in Eq. (4).

$$A = L \times D \times U' = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \times \begin{bmatrix} d_1 & u_{12} & u_{13} \\ 0 & d_2 & u_{23} \\ 0 & 0 & d_3 \end{bmatrix} \times \begin{bmatrix} 1 & u_{12}/d_1 & u_{12}/d_1 \\ 0 & 1 & u_{12}/d_1 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

III. Proposed Watermarking Algorithm

In this section, an overview of the proposed watermarking algorithm, consisting of watermark preprocessing, watermark embedding and watermark detection processes is given.

Let $X=\{x(n),\ 1 \le n \le L\}$ be an original audio signal with L samples, $W=\{w(i),\ 1 \le i \le I\}$ be a binary watermark data to be embedded into the original audio signal.

A. Watermark Preprocessing

A tent map is utilized which contains the chaotic characteristics to encrypt the binary watermark data for enhancing the confidentiality of the proposed method. It can be defined as follows:

$$y(i+1) = \begin{cases} \frac{1}{\beta} y(i), & 0 \le y(i) < \beta \\ \frac{1}{\beta-1} y(i) + \frac{1}{\beta-1}, & \beta \le y(i) \le 1 \end{cases} \tag{5}$$

where $y(1) \in (0,1)$ and β are real parameters (map's initial condition). Then $z(i)$ is calculated using the following rule:

$$z(i) = \begin{cases} 1 & if\ y(i) > T \\ 0 & otherwise \end{cases} \tag{6}$$

where T is a predefined threshold. Finally $w(i)$ is encrypted by $z(i)$ with the following rule:

$$u(i) = z(i) \oplus w(i), \quad 1 \le i \le I \tag{7}$$

where $\oplus$ is the exclusive-or (XOR) operation and $u(i)$ is the encrypted watermark data. After applying this chaotic encryption technique, the original watermark is encrypted and can not be found by random search. In this study, the value of $y(1)$, a, and b are used as secret key K.

B. Watermark Embedding Process

The proposed watermark embedding process is discussed as follows:

1) The original audio signal X is first segmented into non-overlapping frames $F= \{F_1, F_2, F_3,.......,F_I\}$.

2) The FWHT is applied to each frame F_i to obtain the FWHT coefficients.

3) The FWHT coefficients of each frame F_i are rearranged into an $N \times N$ square matrix H_i. This is done by dividing the coefficient set into N segments with N coefficients.

4) LUD is performed to decompose each matrix H_i into three matrices: upper triangular matrix L_i, diagonal matrix D_i, and lower triangular matrix U_i. The LUD operation is represented as follows:

$$H_i = L_i D_i U_i \tag{8}$$

5) The proposed algorithm embeds watermark bits into the largest element of each matrix D_i using a quantization function to guarantee the robustness and transparency. Let $P_i = round\left(\frac{D_{i(\max)}}{Q}\right)$, where Q is a predefined quantization coefficient and $D_{i(\max)}$ is the largest element of each D_i. The embedding equation is given as follows:

$$P_i' = \begin{cases} P_i + C - (P_i \bmod R), & if\ u(i)=1 \\ P_i + C - ((P_i + C) \bmod R), & if\ u(i)=0 \end{cases} \tag{9}$$

where $R=2C$, C is an integer, mod is the modulo operation.

6) The modified largest element $D'_{i(\max)}$ of each matrix D_i is calculated using the following equation:

$$D'_{i(\max)} = P_i' D \tag{10}$$

where D'_i is the modified diagonal matrix.

7) Each modified largest element $D'_{i(\max)}$ is reinserted into each matrix D_i and inverse LUD is applied to obtain the modified matrix H_i' which is given by

$$H_i' = U_i D_i' L_i \tag{11}$$

Each matrix H_i' is then reshaped to create each frame F_i by performing the inverse operation of step 3.

8) Each matrix H_i' is then reshaped to obtain the watermarked audio frame F_i'.

9) Finally, all watermarked frames are concatenated to calculate the watermarked audio signal X'.

A. Watermark Detection Process

The proposed blind watermark detection process is discussed as follows:

1) The attacked watermarked audio signal X^* is segmented into non-overlapping frames $F^* = \{F^*_1, F^*_2, F^*_3, \ldots\ldots, F^*_I\}$.

2) FWHT is applied to each frame and the FWHT coefficients of each frame F_i^* are rearranged into an $N\times N$ square matrix H_i^*.

3) LUD is performed on each H_i^* and select the largest element $D^*_{i(\max)}$ from each matrix D_i^* of the attacked watermarked audio frame.

4) Calculate P_i^* of each $D^*_{i(\max)}$.

5) Encrypted watermark data is extracted as follows:

$$u^*(i) = \begin{cases} 1; & \text{if } (P_i^* \bmod R) = 1 \\ 0; & \text{otherwise} \end{cases} \tag{12}$$

6) Perform chaotic decryption using the secret key K to find the binary watermark data with the following rule:

$$w^*(i) = z(i) \oplus u^*(i) \tag{13}$$

IV. Experimental Results and Discussion

In this section, several experiments were carried out on four different types of 16 bit mono audio signals (Pop, Jazz, Folk, and Speech) sampled at 44.1 kHz to demonstrate the performance of the proposed algorithm. Each audio file contains 262,144 samples (duration 5.94 sec). Each audio signal is divided into frames of size 256 samples. Thus, the total number of frames is 1024. In each frame of audio signal, we embed one bit watermark data. Here, the selected value for $y(1)$, β, T, B, θ, and Q are 0.6, 0.3, 0.5, 2, 45°, and 0.4, respectively. These parameters were selected in order to achieve a good trade-off among the imperceptibility, robustness, and data payload.

TABLE I. Subjective and Objective Evaluation of fferent Watermarked Sounds

Types of Signal	Subjective Evaluation		Objective Evaluation	
	MOS	**Correct Detection**	**SNR**	**ODG**
Pop	4.90	58%	36.65	-0.53
Jazz	4.90	54%	35.47	-0.58
Classical	4.90	48%	36.39	-0.62
Speech	4.80	46%	35.78	-0.72
Average	**4.88**	**51.5%**	**36.07**	**-0.62**

We have evaluated the performance of the proposed algorithm in terms of data payload, mean opinion score (MOS), correct detection, objective difference grade (ODG), signal-to-noise ratio (SNR), bit error rate (BER), and normalized cross-correlation (NC) [10]-[11].

Perceptual quality of watermarked audio signal can be evaluated using subjective and objective tests. The subjective listening test was carried out by blind ten listeners of different ages and the result is summarized in terms of MOS and correct detection.

TABLE II. SNR and MOS Comparison Between The Proposed and Several Recent Methods

Reference	Algorithm	SNR	MOS
[6]	DWT-based energy proportion	17.95	4.15
[5]	Wavelet-based entropy	22.46	4.38
[14]	EMD	24.12	--
[15]	Patchwork	24.95	4.67
[9]	FFT amplitude modification	25.70	--
[12]	DCT-SVD	32.53	4.71
Proposed	**FWHT-LUD**	**36.07**	**4.88**

TABLE III. Attacks Used in This Study for Watermarked Sound

Attacks	Description
Noise addition	Additive white Gaussian noise (AWGN) is added to the watermarked audio signal.
Cropping	Segments of 200 samples are removed from the watermarked audio signal at five different positions and then these samples are replaced by the watermarked samples attacked with AWGN.
Re-sampling	The watermarked signal originally sampled at 44.1 kHz is re-sampled at 22.050 kHz and then restored by sampling again at 44.1 kHz.
Re-quantization	The 16 bit watermarked audio signal is quantized down to 8 bits/sample and again re-quantized back to 16 bits/sample.
MP3 Compression	MPEG-1 layer 3 compression with 128 kbps is applied to the watermarked audio signal.

TABLE IV. NC and BER of The Extracted Watermark for Different Audio Signals

Audio Signal	Attack Type	NC	BER (%)
Pop	No attack	1	0
	Noise addition	0.9964	0.5254
	Cropping	0.9983	0.2320
	Re-sampling	1	0
	Re-quantization	1	0
	MP3 compression	0.9894	1.2695
Jazz	No attack	1	0
	Noise addition	1	0
	Cropping	0.9926	0.8789
	Re-sampling	1	0
	Re-quantization	1	0
	MP3 compression	0.9918	0.9766
Folk	No attack	1	0
	Noise addition	0.9926	0.8789
	Cropping	1	0
	Re-sampling	1	0
	Re-quantization	1	0
	MP3 Compression	0.9862	1.7465
Speech	No attack	1	0
	Noise addition	0.9951	0.5865
	Cropping	0.9976	0.2930
	Re-sampling	1	0
	Re-quantization	1	0
	MP3 Compression	0.9902	1.1719

TABLE V. A GENERAL COMPARISON BETWEEN THE PROPOSED ALGORITHM AND SEVERAL RECENT METHODS SORTED BY DATA PAYLOAD

Reference	Algorithm	Payload (bps)	Re-sampling BER (%)	Re-quantization BER (%)	MP3 compression BER (%)
Proposed	**FWHT-LUD**	**172.39**	**0 (22.05 kHz)**	**0 (8 bits/sample)**	**1.75 (128 kbps)**
[11]	DCT-SVD-LPT	172.39	1.56 (22.05 kHz)	0 (8 bits/sample)	3.91 (128 kbps)
[5]	Wavelet-based entropy	86.14~172.28	9.1 (22.05 kHz)	--	6.7 (128 kbps)
[14]	EMD	46.9~50.3	3 (22.05 kHz)	0 (8 bits/sample)	1 (32 kbps)
[12]	DCT-SVD	43	0 (22.05 kHz)	0 (8 bits/sample)	3 (32 kbps)
[16]	Histogram	3	0 (--)	0 (8 bits/sample)	15 (128 kbps)
[8]	LWT	--	16.50 (36.750 kHz)	22.09 (8 bits/sample)	51.73 (128 kbps)

The objective test was conducted by calculating the ODG and SNR. The MOS, correct detection, ODG, and SNR values of the different watermarked signals are shown in Table I. It is seen that the MOS, correct detection, ODG, and SNR values range from 4.8 to 4.9, 46% to 58%, -0.72 to -0.53, and 35.47 to 36.65, respectively, indicating that the proposed watermarking algorithm provides good imperceptible watermarked sound. Table II shows a comparison between the proposed algorithm and the several recent methods in terms of SNR and MOS which are based on the reported results in the references [5]-[6], [9], [12], [14]-[15]. From this comparison, it is observed that the proposed algorithm shows better result than the recent watermarking methods in terms of SNR and MOS. In other word, subjective and objective tests prove a high transparency of the proposed algorithm.

Various signal processing attacks shown in Table III were applied to assess the robustness of the proposed algorithm. Table IV shows the NC and BER results of the proposed algorithm against various attacks for different audio signals. We observed that the NC values range from 0 to 0.9862 and the BER values range from 0% to 2%. This clearly indicates a good performance of the proposed algorithm against various attacks.

A general comparison between the proposed algorithm and the several recent methods sorted by data payload, which is based on the reported result in the references [5], [8], [11]-[12], [14] and [16] is presented in Table V. Moreover, Table V shows a comparison for re-sampling, re-quantization, and MP3 compression. From this comparison, we can conclude that the proposed algorithm has higher data payload and lower BER values against various attacks than the state-of-the-art methods. This is because watermark bits are embedded into the largest element of the diagonal matrix obtained from the FWHT coefficients of each frame.

V. CONCLUSION

In this paper, we introduced a blind audio watermarking algorithm based on FWHT and LUD. Experimental results demonstrate that the embedding data are robust against various attacks such as noise addition, cropping, re-sampling, re-quantization, and MP3 compression. In addition, subjective and objective tests show high imperceptibility of the watermarked signals. Moreover, it has high data payload and provides superior performance than the state-of-the-art audio watermarking methods. These results verify that the proposed algorithm can be effectively utilized for audio copyright protection. In future, we will extend our research for image and video watermarking.

REFERENCES

[1] N. Cvejic, T. Seppanen, Digital Audio Watermarking Techniques and Technologies, *Information Science Reference*, USA, 2007.

[2] P. K. Dhar and T. Shimamura, Advances in Audio Watermarking Based on Singular Value Decomposition, Springer, 2015.

[3] W. N. Lie and L. C. Chang, "Robust high quality time domain audio watermarking based on low frequency amplitude modification," *IEEE Trans. Multimedia*, vol. 8, no. 1, pp. 46-59, 2006.

[4] P. Bassia, I. Pitas, and N. Nikolaidis, "Robust audio watermarking in the time domain," *IEEE Trans. Multimedia*, vol. 3, no. 2, pp. 232-241, 2001.

[5] S. T. Chen, H. N. Huang, C. J. Chen, K. K. Tseng and S. Y. Tu, "Adaptive audio watermarking via the optimization point of view on the wavelet-based entropy," *Digital Signal Process.*, vol. 23, no. 3, pp. 971-980, 2013.

[6] S. T. Chen, H. N. Huang, C. J. Chen, and G. D. Wu, "Energy-proportion based scheme for audio watermarking," *IET Signal Process.*, vol. 4, no. 5, pp. 576-587, 2010.

[7] S. T. Chen, G. D. Wu, and H. N. Huang, "Wavelet-Domain Audio Watermarking Scheme Using Optimisation-Based Quantization," *IET Signal Process.*, vol. 4, no. 6, pp. 720-727, 2010.

[8] E. Erçelebi and L. Batakçı, "Audio watermarking scheme based on embedding strategy in low frequency components with a binary image," *Digital Signal Process*., vol. 19, no. 2, pp. 265-277, 2009.

[9] D. Megías, J. Serra-Ruiz, and M. Fallahpour, "Efficient self–synchronized blind audio watermarking system based on time domain and FFT amplitude modification," *Signal Process.*, vol. 90, no. 12, pp. 3078-3092, 2010.

[10] P. K. Dhar and T. Shimamura, "Audio Watermarking in Transform Domain Based on Singular Value Decomposition and Cartesian-Polar Transformation, *Int. J. Speech Technol*., vol. 17, pp. 133-144, 2014.

[11] P. K. Dhar and T. Shimamura, "Blind SVD-Based Audio Watermarking Using Entropy and Log-Polar Transformation, *J. Information Security and Application*, vol. 20, pp. 74-83, 2015.

[12] B.Y. Lei, I. Y. Soon, and Z. Li, "Blind and robust audio watermarking scheme based on SVD-DCT," *Signal Process*., vol. 91, pp.1973-1984, 2011.

[13] V. K. Bhat, I. Sengupta, and A. Das, "An adaptive audio watermarking based on the singular value decomposition in the wavelet domain," *Digital Signal Process*., vol. 20, no. 6, pp. 1547-1558, 2010

[14] K. Khaldi and A. O. Boudraa, "Audio watermarking via EMD," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 21, no. 3, pp. 675-680, 2013.

[15] H. Kang, K. Yamaguchi, B. Kurkoski, K. Yamaguchi, and K. Kobayashi, "Full-index-embedding patchwork algorithm for audio watermarking," *IEICE Trans. Inf. & Syst.*, vol. E91-D, no. 11, pp. 2731-2734, 2008.

[16] S. Xiang and J. Huang, "Histogram based audio watermarking against time scale modification and cropping attacks," *IEEE Trans. Multimedia*, vol. 9, no. 7, pp. 1357-1372, 2007.

Effects of Caffeine Doses on Cardiac Activity using Laser Doppler Flowmetry

Md. Bashir Uddin
Department of Biomedical Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
bas.u.kuet@gmail.com

M Nahid Reza & Mohiuddin Ahmad
Department of Electrical and Electronic Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
mohiuddin.ahmad@gmail.com

Abstract—**In this work, the effects of different caffeine doses on cardiac activity have been evaluated by frequency domain analysis of laser Doppler flowmetry signal. Using different data acquisition units, the blood perfusion on human forearm middle finger tip has been recorded. Blood perfusion has been recorded before and immediately after the consumption of all caffeine doses. Different caffeine doses contain different amount of caffeine but approximately same amount of sugar. Recorded and pre-processed data have been analyzed using frequency spectrum within the frequency range of cardiac activity. It is found that the consumption of caffeine free dose increases the cardiac activity. Besides, the consumption of caffeinated doses decreases the cardiac activity with respect to normal condition (before consumption). As the amount of caffeine consumed increases, the cardiac activity improves. Result shows 42.9%, 32.2% and 39.4% decrement in cardiac activity due to the consumption of 27 mg, 48 mg, and 64 mg caffeine respectively. The consumption of 80 mg caffeine dose significantly increases the cardiac activity. The outcome indicates the reduction of heart functions due to the consumption of lower caffeine dose.**

Keywords—caffeine; energy drink; blood perfusion; laser doppler flowmetry; cardiac activity

I. INTRODUCTION

Caffeine is a stimulant which is the most widely consumed psychoactive drug [1], but unlike many other psychoactive substances, it is legal and unregulated in nearly all parts of the world. Beverages containing caffeine are ingested to relieve drowsiness and supposed to give consumers a short term boost in energy. These beverages known as energy beverages or drinks are very popular among adults. A common and main ingredient in most energy drinks is caffeine. The most common naturally caffeinated beverages are coffee and tea, other beverages are artificially caffeinated. The consumption of caffeinated drinks is often intended entirely or partly for physical and mental effects of caffeine.

Energy drinks have the effects of caffeine and sugar provided, but there is little or no evidence that the wide variety of other ingredients have any effect. Two studies reported significant improvements in mental and cognitive performances as well as increased subjective alertness [2]. Excess consumption of energy drinks may induce mild to moderate euphoria primarily caused by stimulant properties of caffeine [3]. Consumption of a single energy drink will not lead to excessive caffeine intake, but consumption of two or more drinks in a single day can [4]. The drinks may cause seizures due to the “crash” following the energy high that occurs after consumption [5]. Short term side effects have been shown as symptoms of mild caffeine consumption [6]. The effects of having energy drinks with a little amount of caffeine on different physiological activities were studied in [7]. The long term effects of moderate caffeine consumption can be a reduced risk of developing Parkinson's, hepatic and cardiovascular diseases [8]. Caffeine affects cardiovascular responses [9] and study shows that caffeinated drinks have a short-term impact on cardiac contractility. The effects of caffeinated beverage consumption on electrocardiographic parameters were well studied [10]. A recent study reported significant impacts of caffeinated beverage consumption on cardiac function by analyzing spectral components [11].

The effects of caffeine or caffeinated drinks on cardiac functions are well studied and some results are contradictory. No research has been conducted to analyze cardiac activities by consuming different dose of caffeine. In this study, cardiac activities have been investigated by varying the amount of caffeine using laser Doppler flowmetry and frequency domain analysis. It is hypothesized that different dose of caffeine would have different impacts on cardiac function.

II. LASER DOPPLER FLOWMETRY AND CARDIAC ACTIVITY

A. Laser Doppler Flowmetry

Laser Doppler Flowmetry (LDF) is an established and reliable method for measurement of blood perfusion in microvascular research. LDF measurements from the skin reflect blood flow in capillaries, arterioles, venules, and dermal vascular plexus. The principle of laser Doppler flowmetry technique is based on reflected laser light from blood. Low power laser light is used to illuminate tissue using a fiber optic; the light is scattered by the static tissue structures and moving blood cells; the moving blood cells impart a Doppler shift; an adjacent fiber detects light returned from the tissue; this light contains Doppler shifted and unshifted light;

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

the signal is processed to extract the signal related to the moving red blood cells.

The spectral analysis of the LDF signal from human forearm skin has revealed five characteristic frequencies. In addition to the cardiac and respiratory rhythms around 1 and 0.3 Hz, respectively, three frequencies have been detected in the regions around 0.1, 0.04, and 0.01 Hz in human skin [12].

B. Cardiac Activity

Cardiac activity or heart activity can be demonstrated by using Laser Doppler Flowmetry (LDF) and the frequency domain analysis of LDF signal. Frequency domain analysis of LDF signal is an established way to evaluate cardiac activity. Fast Fourier transform (FFT) is a frequency based analysis to compute the frequency component of the non-stationary signal. The discrete Fourier transform of the discrete time signal *x(j)* is noted in (1).

$$X_k = \sum_{j=n_s}^{n_e-1} x(j) e^{-i2\pi \frac{jk}{N}} \quad (1)$$

where, k represents the harmonic number of frequency components of signal and X_k is the discrete frequency domain at different frequency. The output of a FFT will appear in the graph with magnitude plotted against frequencies. From this output plot it is easier to find out the cardiac activities around 1 Hz or in the range 0.6 to 1.6 Hz.

III. Materials and Methods

A. Caffeine Dose Preparation

To vary the amount of caffeine, different drinks were used in this experiment. Direct form of caffeine was not consumed because it was very tough to get. Caffeine is the main ingredient in many drinks and the amount of caffeine varies from brand to brand. Drinks of different brands were used to make different dose of caffeine. Preparation of different dose of caffeine is listed in Table I indicating corresponding amount of caffeine and sugar. The listed amount of caffeine and sugar is calculated according to ingredient list of different brands. All doses of caffeine were prepared using same amount of drinks (250 ml) which was a single type of drinks or combination of two drinks. Single drinks or combination of drinks were used to prepare these doses with different amount of caffeine. Five doses were prepared for consecutive five experimental days as listed in Table I. Each dose was prepared 5 minutes before of consumption.

TABLE I. Different Doses of Caffeine using Soft and Energy Drinks

Days	**Doses of Caffeine (250 ml)**	**Caffeine (mg)**	**Sugar (gm)**
1	Sprite (250 ml)	00	27.55
2	Pepsi (250 ml)	27	29.65
3	RedBull (100 ml) + Pepsi (150 ml)	48	28.59
4	RedBull (200 ml) + Sprite (50 ml)	64	27.11
5	RedBull (250 ml)	80	27.00

B. Experimental Setup and Data Acquisition

Healthy young male subjects free from cardiovascular diseases were enrolled in this study. The subjects had not taken any medication during the week prior to the study. None of the subjects were smokers and refrained from alcohol and caffeine containing drinks at least 20 hours prior to the study. Food intake was totally restricted to a light meal 2 hours prior to the test. After being informed about the summary of the study design, they gave their written consent. Each participant had an initial visit to the experimental laboratory, for a physical examination and a medical history assessment.

The study was performed in a quiet room with the temperature kept constant approximately. The subjects were resting in the supine position throughout the whole experimental period. At first, a familiarization day has been conducted to remove “first-day effect” and the participants were totally blind about experimental design (amount of caffeine consumption) to minimize the effects of different biases. At least 10 minutes were allowed for acclimatization before the LDF measurements were performed on the skin of middle finger tip. Skin blood perfusion was measured before and immediately after the consumption of each dose of caffeine. Using data acquisition unit (MP150), laser Doppler flow amplifier (LDF100C), fiber-optic based probes (TSD140) and AcqKnowledge software, Skin blood perfusion measurement was performed.

C. Signal Pre-processing

The recorded signals contained several noises such as 50Hz frequency component. By filtering, this high frequency component was removed. Smoothing technique was also used to find smoothed signal. Mean value smoothing was done as noted in (2), where “q” is the number of points in the window and “p” is the sample number [13].

$$x_{Sample\ Data} = \sum_{h=p-q/2}^{h=p-(q-1)/2} x_{Measured\ Data}(h)/\,q \quad (2)$$

IV. Data Analysis

A. Recorded and Pre-processed LDF Signal

In this experiment, LDF signal was recorded with duration of 5 minutes and 25 minutes before and immediately after the consumption of each caffeine dose. A typical LDF recording of human forearm middle finger tip before and immediately after the consumption of a caffeine dose are shown in Fig. 1 and Fig. 2 respectively. In Fig. 1, it is observed that the blood perfusion (LDF signal) is oscillating with an average value around 1000 BPU (blood perfusion unit) and blood perfusion is oscillating with an average value around 1200 BPU (Fig. 2). It is seen that, the mean value of blood perfusion is increased around 200 BPU due to the consumption of a dose of 27 mg caffeine. Besides, it is also noticed that, a tiny decrement in peak to peak values of blood perfusion due to the consumption of same dose of caffeine. It is very difficult to evaluate cardiac activity using these data, so further analysis is required.

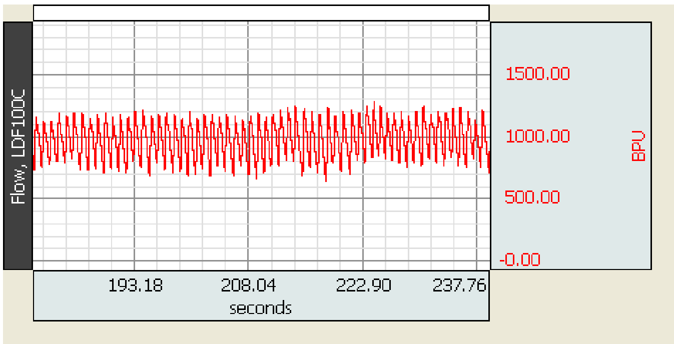


Fig. 1. Recorded LDF signal before consuming a dose of 27 mg caffeine.

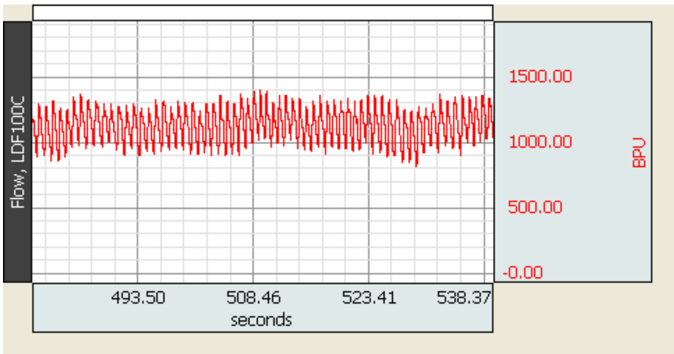


Fig. 2. Recorded LDF signal after consuming a dose of 27 mg caffeine.

B. *Frequency Domain Analysis*

Cardiac activity can be evaluated by frequency domain analysis of LDF signal. Frequency domain or spectral analysis of LDF signal from human forearm middle finger tip is effective to evaluate cardiac activity within the frequency range 0.6-1.6 Hz. In this experiment, fast Fourier transform (FFT) is used as frequency domain or spectral analysis. FFT of LDF signal before and immediately after the consumption of a dose of 27 mg caffeine are shown in Fig. 3 and Fig. 4 respectively. It is seen that the peak magnitudes of cardiac activity are 32.78 BPU and 18.73 BPU respectively before and immediately after the consumption of 27 mg caffeine dose that results a decrement in cardiac activity due to the consumption of that caffeine dose.

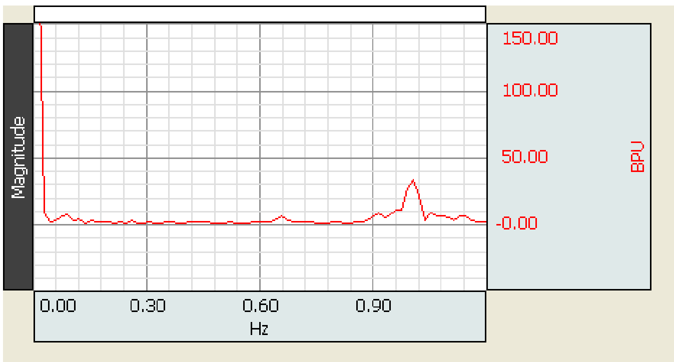


Fig. 3. FFT analysis of LDF before consuming a dose of 27 mg caffeine.

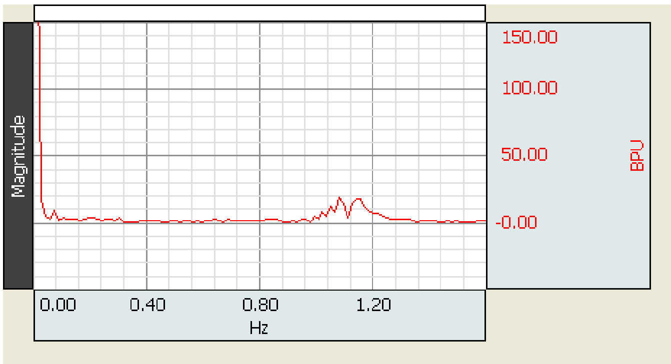


Fig. 4. FFT analysis of LDF after consuming a dose of 27 mg caffeine.

V. Results and Discussion

Cardiac activity is evaluated by frequency domain analysis of LDF signal with varying the amount of caffeine. In this study, a time varying analysis is done to investigate the effects of different amount of caffeine consumed as a dose. Different doses of caffeine contain different amount of caffeine as listed in Table I and these doses were consumed by participant according to the order of days. It is well known that all the effects (physiological and psychological) of drinks (soft and energy drinks) found is due to caffeine and sugar content. It is noticed that all the caffeine doses consumed in this study contain approximately same amount of sugar and increasing order of caffeine (Table I). So the effects due to the consumption of these doses are totally stood for caffeine variation but not for sugar content (amount of sugar is constant in all doses). A time varying analysis of two caffeine doses is shown in Fig. 5.

Fig. 5 shows the comparative effects of 0 mg and 64 mg caffeine doses with time varying plot using frequency domain analysis. It is seen that 0 mg caffeine consumption results an increment in peak magnitude of cardiac activity upto a certain time and then shows decrement to reach its initial condition. Consumption of 64 mg caffeine results a significant decrement in peak magnitude of cardiac activity.

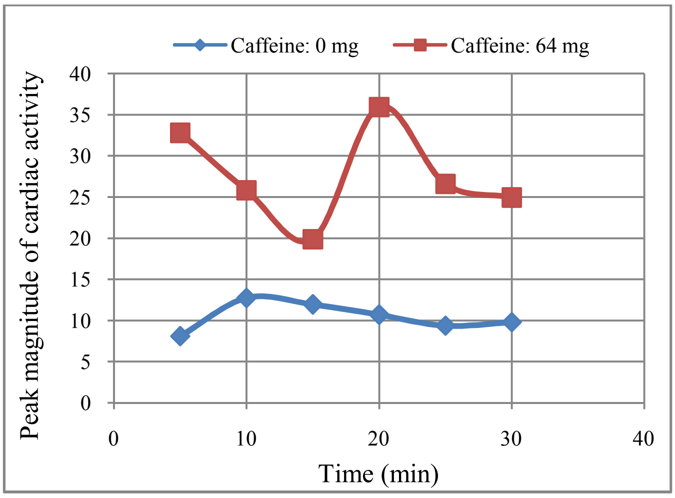


Fig. 5. Cardiac activity variation with caffeine and non-caffeine dose.

The peak magnitude of cardiac activity with different doses of caffeine is listed in Table II. It is noticed that the consumption of 0 mg caffeine results notable increment in peak magnitude of cardiac activity after a certain time (10 min) and then shows decrement with time. The peak magnitude of cardiac activity is 8.11 mV before the consumption of 0 mg caffeine dose. The peak magnitudes of cardiac activities are 12.75 mV, 11.97 mV, 10.72 mV, 9.38 mV and 9.81 mV respectively after 10 min, 15 min, 20 min, 25 min and 30 min due to the consumption of 0 mg caffeine dose. This results an increment in cardiac activity due to the consumption of caffeine free dose. When a 27 mg caffeine dose is consumed, a significant decrement in the peak magnitude of cardiac activity is observed. The peak magnitude of cardiac activity is 32.78 mV before the consumption of 27 mg caffeine dose. The peak magnitudes of cardiac activities are 18.73 mV, 19.22 mV, 24.56 mV, 26.73 mV and 22.58 mV respectively after 10 min, 15 min, 20 min, 25 min and 30 min due to the consumption of 27 mg caffeine dose. Similar decrement in peak magnitudes of cardiac activity is found for 48 mg and 64 mg caffeine dose consumption except 80 mg caffeine dose.

The percentage change of cardiac activity with respect to initial value (before consumption) with different doses of caffeine is listed in Table III. It is noticed that the consumption of 0 mg caffeine results 57.2% increment in peak magnitude of cardiac activity after 10 min and then shows a slow percentage decrement with time. The maximum decrement in peak magnitudes of cardiac activities are 42.9%, 32.2% and 39.4% due to the consumption of 27 mg, 48 mg and 64 mg caffeine consumption respectively. The consumption of 80 mg caffeine dose results some increment in peak magnitude of cardiac activity. All caffeine doses produce significant decrement in cardiac activity except 80 mg dose that indicates the negative impacts of caffeine consumption on heart functions.

TABLE II. VARIATIONS IN PEAK MAGNITUDES OF CARDIAC ACTIVITY WITH TIME BY VARYING THE AMOUNT OF CAFFEINE

Amount of Caffeine (mg)	Peak Magnitudes of Cardiac Activity (mV)					
	Before drink	*After drink*				
	After 5 min	After 10 min	After 15 min	After 20 min	After 25 min	After 30 min
0	8.11	12.75	11.97	10.72	9.38	9.81
27	32.78	18.73	19.22	24.56	26.73	22.58
48	21.08	18.34	26.22	20.90	14.30	23.63
64	32.81	25.83	19.89	35.93	26.61	24.95
80	14.77	22.31	16.13	24.34	27.14	20.31

TABLE III. CHANGES IN PEAK MAGNITUDES OF CARDIAC ACTIVITY WITH TIME BY VARYING THE AMOUNT OF CAFFEINE

Amount of Caffeine (mg)	Changes in Peak Magnitudes of Cardiac Activity (%)					
	Before drink	*After drink*				
	After 5 min	After 10 min	After 15 min	After 20 min	After 25 min	After 30 min
0	0	57.2	47.6	32.2	15.7	20.7
27	0	–42.9	–41.4	–25.1	–18.5	–31.2
48	0	–13.0	24.4	–0.9	–32.2	12.1
64	0	–21.3	–39.4	9.5	–18.9	–24.0
80	0	51.0	9.2	64.8	83.8	37.5

VI. CONCLUSIONS

The effects of different caffeine doses on cardiac activity are investigated in this study. Cardiac activity using frequency spectrum is analyzed with varying the amount of caffeine. The magnitude of cardiac activity increases due to the consumption of caffeine free dose but decreases due to the consumption of caffeinated dose. As the amount of caffeine is increasing in the caffeine doses, cardiac activity increases from the decreasing nature. Consumption of high amount of caffeine results an increment in cardiac activity.

ACKNOWLEDGMENT

This work was partially supported by Higher Education Quality Enhancement Project (HEQEP), UGC, Bangladesh under Sub-project "Postgraduate Research in BME", CP#3472, Dept. of Biomedical Engineering, Khulna University of Engineering & Technology, Bangladesh.

REFERENCES

[1] A. Nehlig, J. L. Daval, and G. Debry, "Caffeine and the central nervous system:mechanisms of action, biochemical, metabolic and psychostimulant effects," Brain Research Reviews, vol. 17, no. 2, pp. 139–70, 1992.

[2] M. A. Howard and C. A. Marczinski, "Acute effects of glucose energy drink on behavioral control," *Experimental and clinical psychopharmacology*, vol. 18, no. 6, pp. 553–61, 2010.

[3] F. Van Den Eynde, P. C. Van Baelen, M. Portzky, and K. Audenaert, "The effects of energy drinks on cognitive performance," *Tijdschrift voor psychiatrie*, vol. 50, no. 5, pp. 273–81, 2008.

[4] A. P. Winston, E. Hardwick, and N. Jaberi "Neuropsychiatric effects of caffeine," *Advances in Psychiatric Treatment*, vol. 11, no. 6, pp. 432–9, 2005.

[5] S. J. Iyadurai and S. S. Chung, "New-onset seizures in adults: Possible association with consumption of popular energy drinks," *Epilepsy & behavior*, vol. 10, no. 3, pp. 504–8, 2007.

[6] R. E. Shapiro, "Caffeine and headaches," *Current Pain and Headache Reports*, vol. 12, no. 4, pp. 311–5, 2008.

[7] Muhammad Muinul Islam, Md. Bashir Uddin, and Mohiuddin Ahmad, "Determination of the Effect of Having Energy Drinks by Analyzing Blood Perfusion Signal," *Journal of Computers*, vol. 9, no. 8, pp. 1817-24, 2014.

[8] K. Sugiyama, Y. Noda, and P. He, "Suppressive effect of caffeine on hepatitis and apoptosis induced by tumor necrosis factor-alpha, but not by the anti-Fas antibody, in mice," *Bioscience, Biotechnology, and Biochemistry*, vol. 65, no. 3, pp. 674–7, 2001.

[9] J. W. Daniels, P. A. Mole, J. D. Shaffrath, and C. L. Stebbins, "Effects of caffeine on blood pressure, heart rate, and forearm blood flow during dynamic leg exercise," *J Appl Physiol*, vol. 85, no. 1, pp. 154-9, 1998.

[10] Md. Bashir Uddin, M. Z. Hossain, M. Ahmad, N. Ahmed, and M. A. Rashid, "Effect of Caffeinated Beverage Consumption on Electrocardiographic Parameters among Healthy Adults," *Modern Applied Science*, vol. 8, no. 2, pp. 69-77, 2014.

[11] Md. Bashir Uddin, M. Ahmad, M. Rizon, N. Yusof, and M. A. Rashid, "Cardiac Function Evaluation Analyzing Spectral Components due to the Consumption of Energy Drinks," *Research Journal of Applied Science, Engineering and Technology*, vol. 7, no. 18, pp. 3906-15, 2014.

[12] A. Stefanovska and P. Kroselj, "Correlation integral and frequency analysis of cardiovascular functions," *Open systems & information dynamics*, vol. 4, pp. 457-78, 1997.

[13] M. Ahmad, A. Islam, T. T. Khan Munia, M. A. Rashid, and T. M. N. Tunku Mansur, "Physiological signal analysis for cognitive state estimation," *Biomedical Engineering: Applications, Basis and Communication*, vol. 24, no. 1, pp. 57-69, 2012.

An Empirical Framework for Parsing Bangla Assertive, Interrogative and Imperative Sentences

Mohammed Safayet Arefin, Lamia Alam
Dept. of Computer Science & Engineering
Chittagong University of Engineering & Technology,
Chittagong, Bangladesh.
Email: safayet_arefin@yahoo.com

Shayla Sharmin, Mohammed Moshiul Hoque
Dept. of Computer Science & Engineering
Chittagong University of Engineering & Technology,
Chittagong, Bangladesh.
Email: moshiulh@yahoo.com

Abstract— **To interpret language we need to determine a sentence structure. To do this we know the rule of how sentences of a language are organized and have an algorithm to analyze sentences given those rules. Parsing serves in language to combine the meaning of words and phrases. Parsing a sentence then involves finding a possible legal structure for sentence. This paper proposes a set of context-sensitive grammars (CSG's) to parse the Bangla sentences including assertive, interrogative and imperative. Experimental result reveals that the proposed framework can parse Bangla of sentences with over 80% *accuracy.***

Index Terms—Natural language processing, context-sensitive grammars. Lexicon, and parse tree.

I. Introduction

Parsing is the most important part as far as natural language processing of Bangla is concerned. To analyze language we must have a good idea about sentence structures. To do these we must know the rules of how language is organized and have an algorithm to language given those rules. We can find out words in sentence related to each other by analyzing the structure of a sentence. The result of parsing is usually a parse tree or structural representation [2]. For Bangla to other language machine translation we need parsing. Most of the previous system has used CFG's to parse the Bangla sentences into English. However, CFG are not sufficient to parse the all types of sentences and hence translation [1, 6]. To parse different kinds of Bangla sentences, we have to use CSG's due to its capabilities to handle agreement between subject-verb and person-class [6, 8]. For example, if we consider the sentence "হাসান কি নদী ভালবাসে?". That can't be parsed by CFG.

Most of the previous parsing systems parse the Bangla sentences into English structurally (i.e., simple, complex, and compound) rather than their function or purpose of the user. Bangla sentences may be classified according to the purpose of the speaker or writer into five categories namely, assertive, interrogative, imperative, optative, and exclamatory sentences. The main contribution of this work is develop a parser that can parse the three types of Bangla sentences such as, assertive, interrogative, and imperative sentences by using a set of CSG rules. The parser output is given in a list in the paper and have verified the system with several types of examples and found that the performance is satisfactory.

II. Previous work

Parsing of Bangla sentences is in rudimentary stage now. A method to translate Bangla sentences into English sentences using context-sensitive grammar rules which accepts Bangla sentences including assertive, interrogative and imperative sentences is implemented in [1]. In their work, they emphasized on machine translation rather than parsing. A parsing technique for simple sentence was implemented using a set of CFG rules in [2]. A comprehensive approach for CFG rules to parse all types of sentences including complex, compound, exclamatory and optative sentences was shown in [4]. Anwar et al. develop a technique to parse Bangla sentences using context sensitive grammar rules which accept almost all types of Bangla sentences including simple, compound and complex sentences is implemented in [5]. It also describes the technique to decompose a complex sentence into a dependent and independent clause and a compound sentence into a simple sentence respectively. Besides in [3], analyzing of the syntax of various types of Bangla sentences and design transformational generative grammar rules for them was shown. A detail explanation of Bangla phrases and different types of sentences by using the TGG was given in [7].

III. Proposed parser module

The schematic representation of proposed Bangla natural language parser module is illustrated in fig. 1. Details description of this module is given in the following subsections.

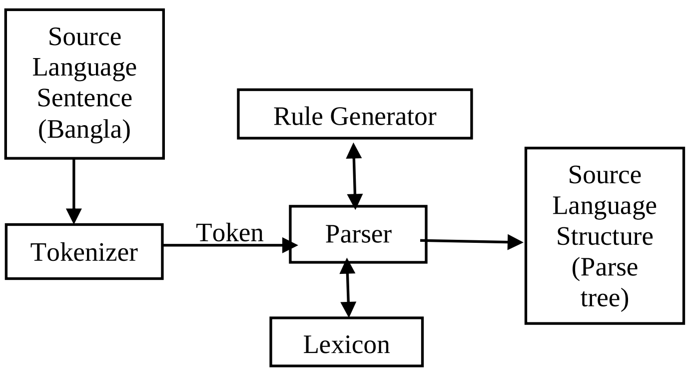


Fig. 1. Proposed Bangla natural language parser

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

A. Input sentence

Bangla sentences are taken as input for the parsing framework. In this system, only assertive, interrogative and imperative Bangla sentences together their negative form are considered as input for implementation.

B. Tokenizer

Tokenizer is the program module that accepts a sentence to be parsed as an unbroken string, breaks into individual words called Tokens. Tokens are stored in the list for further access. The token is then checked into the lexicon for validity, some words, if necessary, should be combined into groups because two or more words may represent a single word type [1]. For example, for input sentence: "হাসান ফুটবল খেলে না" , the output of the tokenizer can be represented as-
Output: ("হাসান", "ফুটবল", "খেল", "এ", "না").

C. Lexicon

The typical entries in the lexicon which we have used in our system are shown in the Table 1.

TABLE 1: TYPICAL ENTRIES IN LEXICON

Bangla	**Features**
আমি	[PR, Per1]
হাসান	[N, Per2]
সে	[PR, Per3]
কে	[IW]
সভা	[N]
নি	[ind]
না	[ind]
আস	[V]
?	[IM]
কি	[IW]
স্কুল	[N]

[Abbreviations: PR: Pronoun, N: Noun, Per1: First person, Per2: Second person, Per3: Third person, IW: Interrogative word, IM: Interrogative marker, ind: Indeclinable, V: Verb]

D. Rule Generator

In this paper, Bangla CSG is used for different kinds of Bangla sentences those are discussed in the following subsections. List of used Bangla CSG to parse the sentences is given in the table 2. While parsing, "Null" is used in every cases when a token in the right side of the rule is not needed to parse the sentence.

TABLE II. CSG RULES OF BANGLA SENTENCES

Rule No	**Bangla CSG's rule**
1	S→AS | IRS | IS
2	AS→ NP VP
3	IRS→ NP VP
4	IS→ NP VP
5	NP→(Qntfr) (PP) N|PN
6	NP→N (Biv) (Adj)
7	NP→ N|PN IW
8	NP→ Null
9	VP→(NP) VF
10	VP→ (NP) VF IM
11	VF→ V (Con) (Aux) (ind)
12	N→ রহিম, হাসান, বই, ফুটবল, স্কুল, জনগণ, ...
13	PN→ আমি, তুমি, সে, ...
14	V→ পড়, খেল, দেয়, ছি, যা, ...
15	Adj→ ভাল, উপস্থিত, ...
16	Biv→ কে, য়, এ, ...
17	Con→ তে, এ, টি, ই, ছে, লাম, ...
18	Aux → পারে, করে, ...
19	ind → না, নি
20	IW→ কি, কে, কোথায়, কখন, ...
21	IM→?
22	Qntfr→ একটি, দুটি, ...

[Abbreviations: S: Sentence, AS: Assertive sentence, IRS: Interrogative sentence, IS: Imperative sentence, NP: Noun phrase, N: Noun, PN: Pronoun, VP: Verb phrase, VF: Verb form, V: Verb, Qntfr: Quantifier, PP: Preposition, Biv: Bivokti (inflection), Adj: Adjective, Con: Concord, Aux: Auxiliary, ind: indeclinable, IW: Interrogative word, IM: Interrogative marker]

1) CSG for Assertive Sentences:

In Bengali assertive-negative sentences, an indeclinable word "না" is used to express the negativity. The position of this word is at the end of the sentence [3]. As an example, we can consider the Bangla sentence "হাসান নদী ভালবাসে না".

In order to get the parse tree we have used CSGs rules from Table 2.

S→AS [Rule no: 1]
→NP VP [[Rule no: 2]
→(Qntfr) (PP) N VP [Rule no: 5]
→Null (PP) N VP VP
→Null Null N VP
→Null Null হাসান VP [Rule no: 12]
→ Null Null হাসান (NP) VF [Rule no: 9]
→ Null Null হাসান N (Biv) (Adj) VF [Rule no: 6]
→ Null Null হাসান নদী (Biv) (Adj) VF [Rule no: 12]
→ Null Null হাসান নদী Null (Adj) VF
→ Null Null হাসান নদী Null Null VF
→ Null Null হাসান নদী Null Null V (Con) (Aux) (ind) [Rule no: 11]
→ Null Null হাসান নদী Null Null ভালবাস (Con) (Aux) (ind) [Rule no: 14]
→ Null Null হাসান নদী Null Null ভালবাস এ (Aux) (ind) [Rule no: 17]

→ Null Null হাসান নদী Null Null ভালবাস এ Null (ind)

→ Null Null হাসান নদী Null Null ভালবাস এ Null না [Rule no: 19]

Structural representation (SR) of this assertive sentence is:

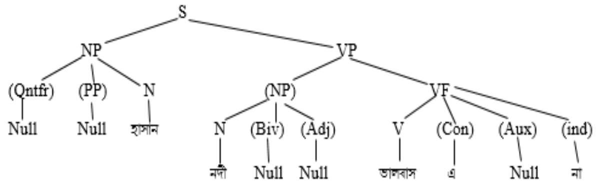


Fig. 2. Structural representation of "হাসান নদী ভালবাসে না"

2) CSG for Interrogative Sentences

In Bengali interrogative sentence we always find an interrogative word (IW) and an interrogative marker (IM) at the end of the sentence. Negative-interrogative sentences are the combination of negative and interrogative rules. In this type of Bangla sentence we find the indeclinable "না" as well as IW and IM in the same sentence.

As an example, we can consider the Bangla sentence "হাসান কি নদী ভালবাসে না?" Here we find the indeclinable "না", IW "কি" and IM "?" in the Bangla sentence.

In order to get the parse tree, CSGs rules from Table 2 is used.

S →IRS [Rule no: 1]

→NP VP [Rule no: 3]

→N IW VP [Rule no: 7]

→হাসান IW VP [Rule no: 12]

→হাসান কি VP [Rule no: 20]

→হাসান কি (NP) VF IM [Rule no: 10]

→হাসান কি N (Biv) (Adj) VF IM [Rule no: 6]

→হাসান কি নদী (Biv) (Adj) VF IM [Rule no: 12]

→ হাসান কি নদী Null (Adj) VF IM

→ হাসান কি নদী Null Null VF IM

→ হাসান কি নদী Null Null V (Con) (Aux) (ind) IM [Rule no: 11]

→হাসান কি নদী Null Null ভালবাস (Con) (Aux) (ind) IM [Rule no: 14]

→ হাসান কি নদী Null Null ভালবাস এ (Aux) (ind) IM [Rule no: 17]

→হাসান কি নদী Null Null ভালবাস এ Null (ind) IM

→ হাসান কি নদী Null Null ভালবাস এ Null না IM [Rule no: 19]

→ হাসান কি নদী Null Null ভালবাস এ Null না ? [Rule no: 21]

SR of this interrogative sentence is:

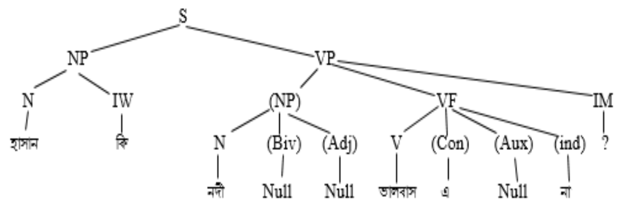


Fig. 3. Structural representation of "হাসান কি নদী ভালবাসে না?"

3) CSG for Imperative Sentences

In imperative sentences, if the subject is second person then the subject may remain hidden. The negative-imperative sentences are same as the normal imperative sentence except that the indeclinable "না" [1]

Let's consider a negative-imperative Bangla sentence "স্কুলে যেও না". There is no noun and pronoun in the first noun. There is no use of second person in this type of sentence.

In order to get the parse tree we have used CSGs rules from Table 2.

S→IS [Rule no: 1]

→NP VP [Rule no: 4]

→Null VP

→Null (NP) VF [Rule no: 9]

→Null N (Biv) (Adj) VF [Rule no: 6]

→Null স্কুল (Biv) (Adj) VF [Rule no: 12]

→ Null স্কুল এ (Adj) VF [Rule no: 16]

→ Null স্কুল এ Null VF

→ Null স্কুল এ Null V (Con) (Aux) (ind) [Rule no: 11]

→ Null স্কুল এ Null যে (Con) (Aux) (ind) [Rule no: 14]

→ Null স্কুল এ Null যে ও(Aux) (ind) [Rule no: 17]

→ Null স্কুল এ Null যে ও Null (ind)

→ Null স্কুল এ Null যে ও Null না [Rule no: 19]

SR of this imperative sentence is:

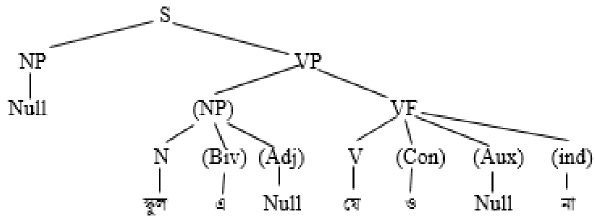


Fig. 4. Structural representation of "স্কুলে যেও না"

E. Parser

. Parsing is a processing of taking tokens of input sentence and producing a parse tree or structural representation (SR) according to CSG rules [1].

1) Parsing algorithm

The input bangla sentence is parsed by using these steps. The steps are given below:

Step 1: We have considered the input of parser is the output of the tokenizer. Tokens are stored in a stack for further access. For example: if the input sentence is "হাসান নদী ভালবাসে না" then the tokens will be ("হাসান", "নদী", "ভালবাস", "এ", "না"). These tokens will store in a stack.

Step 2: The tokens are then checked the lexicon for the validity.For example: the tokens ("হাসান", "নদী", "ভালবাস", "এ", "না") will be enter into the lexicon to find its validity.

Step 3: The tokens are matched with the grammar rules of Bangla. If a rule whose right hand side matches with a token, then the token is assigned with appropriate parts of speech. For example: N→ হাসান, N→ নদী, V→ ভালবাস will produce a partial structure.

Step 4: Starting from left to right hand side of token list, check every rule whose right hand side will match one or more of the parts of speech. If a right hand side of a rule matches with appropriate parts of speech, then we have to select that rule.

Step 5: Repeat step 4, until no more words to generate.

Step 6: If there are no more words to process, then generate SR of the sentence in a list. For example: the parser output of the sentence "হাসান নদী ভালবাসে না" is given below-

S[NP[[N হাসান][VP[NP[[N নদী]]VF[[V ভালবাস][Con ে][ind না]]]]]

2) Parser Output

Fig. 5 represents snapshots of output of the systems for assertive sentence.

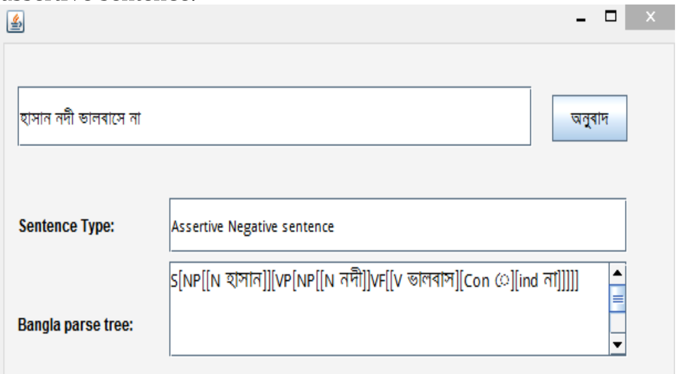


Fig. 5. Parser output of "হাসান নদী ভালবাসে না"

Input: হাসান নদী ভালবাসে না

Sentence Type: Assertive Negative sentence

Bangla parse tree: S[NP[[N হাসান]][VP[NP[[N নদী]]VF[[V ভালবাস][Con ে][ind না]]]]]

IV. EXPERIMENTAL RESULT AND PERFORMANCE

In order to analysis the effectiveness of our system we have tested our system for 420 different sentences and found that total 349 of generated outputs are correct. From our total sentences we have given 30% of the sentences from different sources and 70% is given from the artificially generated sentences.

TABLE III. SUCCESS RATE FOR DIFFERENT TYPES OF SENTENCES

Sentence types	Sentence length	No. of input sentences	No. of correctly translated sentences	Overall success rate (%)
Assertive	3	75	63	81.2
	4	75	61	
	5	20	14	
Interrogative	3	40	35	85
	4	70	60	
	5	80	68	
	6	30	24	
Imperative	2	16	13	80
	3	14	11	
Total	-	420	349	83.09

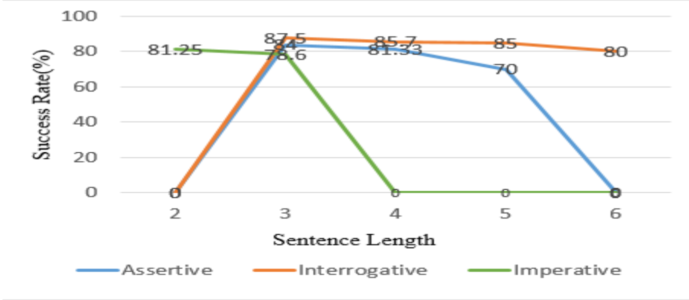


Fig. 8. Success rate vs. sentence length for different types of sentences

Table 3 illustrates the success rate of different types of sentence and Fig. 8 represents the success rate versus sentence length of the system. From the graph, we have found that when sentence length increases then the success rate decreases.

V. CONCLUSION

This paper focused on the Bangla CSG for parsing different kinds of sentences. We have selected very simple and short sentences to evaluate the proposed parsing framework. The parsing algorithm can detect the sentence type and generate the corresponding parse tree efficiently by using CSG's. Experimental results suggest that the framework can parse the sentence with 83% accurately. CSG rules for optative and exclamatory sentences, semantic features and the concepts of voice, narration, infinitive, composition of words can be considered for further improvement.

REFERENCES

[1] M. S. Arefin, M. M. Hoque, M. O. Rahman and M. S. Arefin, "A Machine Translation Framework for Translating Bangla Assertive, Interrogative and Imperative Sentences into English ", In Proc. 2nd International Conference on Electrical Engineering and Information & Communication Technology, ICEEICT'15, Jahangirnagar University, Dhaka, Bangladesh.

[2] M. M. Hoque and M. M. Ali, "A Parsing Methodology for Bangla Natural Language Sentences", In Proc. International Conference on Computer and Information Technology, ICCIT'03, Dhaka, Bangladesh

[3] M. R. Selim and M. Z. Iqbal, "Transformational Generative Grammar for Various Types of Bengali Sentences", SUST Studies, Vol. 12, No. 1, 2010

[4] L. Mehedy, S. M. Arefin and M. Kaykobad, "Bangla Syntax Analysis: A Comprehensive Approach", In Proc. International Conference on Computer and Information Technology, ICCIT'03, Dhaka, Bangladesh, pp. 287-293, vol. 5, 2003

[5] M. M. Anwar, M. Z. Anwar and Md. A. Bhuiyan, "Syntax Analysis and Machine Translation of Bangla Sentences", IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.8, August 2009

[6] M. M. Hoque and M. M. Ali, "Context-sensitive Phrase Structure rules for Structural Representation of Bangla Natural Language Sentences ", In Proc. International Conference on Computer and Information Technology, ICCIT'04, (Dhaka, Bangladesh), PP. 615-620, 2004

[7] M. R. Semil and M. Z .Ikabla, "Syntax analysis of Phrases and Different types of Sentences in Bangla", In Proc. International Conference on Computer and Information Technology, ICCIT' 99", (Sylhet, Bangladesh), pp. 175-186, (1999)

[8] H. Azad, Bakyatattya(বাক্যতত্ত্ব), The University of Dhaka, Dhaka, 1994

[9] H. Mamud, Vasha-shikha (ভাষা-শিক্ষা), 2006.

[10] M. Z. Iqbal , Dipu Number Two, 1996

Paradigm Shift towards Cloud Computing for Banking Sector

Mohammad Anik Islam
Dept. of Computer Science & Engg.
Pabna University of Science & Technology, Pabna, Bangladesh
anik0054@gmail.com

Md. Khaled Ben Islam
Dept. of Computer Science & Engg.
Pabna University of Science & Technology, Pabna, Bangladesh
mdkhaledben@gmail.com

Md. Nazmus Shakib Beg
Dept. of Computer Science & Engg.
Pabna University of Science & Technology, Pabna, Bangladesh
shakibpstu@gmail.com

***Abstract*—Cloud computing is one of the buzzwords of recent time. Cloud computing can be defined as a new style of computing in which dynamically scalable and often virtualized resources are provided as a service over the Internet. Unfortunately, though the Banking sector which works as a financial intermediary and one of the major industries for monetary activities is still deprived of the advantages of cloud. The major reason behind this is the lack of confidentiality and security. As a solution, we are proposing private cloud architecture for banking sector. So their confidentiality will be maintained alongside their data will be safe and sound. We are proposing a simple yet efficient load allocation algorithm to match the homogeneous nature of the banking tasks. We are also proposing a modular banking system. This solution is particularly helpful for developing countries, rural and remote places where setting banking infrastructure is not possible or feasible.**

Keywords*—*cloud computing; private cloud; cloud computing in banking; modular banking

I. INTRODUCTION

Cloud computing can be defined as a new style of computing in which dynamically scalable and often virtualized resources are provided as a services over the Internet [1]. Cloud computing has become a significant technology trend, and many experts expect that cloud computing will reshape information technology (IT) processes and the IT marketplace. Advantages of the cloud computing technology include cost savings, high availability, and easy scalability.

The Banking system is expected to continue to be sensitive to the growth and development needs of all the segments of the society. But till now the banking sector has not been able to embrace the extensive functionalities of the cloud computing. Through our proposed system the banking sector will be able to harness the power of cloud and increase their productivity. Moreover it is not always possible to set up bank in every corner of a country. In our proposed system the modular banks can explore every corner of a country and everyone can get banking facilities. The system has been developed with a view to evolve into a strong, sound and globally competitive financial system, providing integrated services to customers from all segments, leveraging on cloud technology and human resources, adopting the best accounting method.

Banking system is expected to continue to be sensitive to the growth and development needs of all the segments of the society. But till now the banking sector has not been able to embrace the extensive functionalities of the cloud computing. Through our proposed system the banking sector will be able to harness the power of cloud and increase their productivity.

II. BACKGROUND

Cloud Computing is considered as a paragon that render apropos, expeditious network admittance to an earmarked pool of constructible computing resources which can be outfitted and exempted with just nominal diligence and service providers reciprocation. Cloud computing has a great potential and it is regarded as one of the most promising computing infrastructure for the future. A recent survey of 29 senior technology executives at financial institutions found that cloud computing holds the potential to redefine the relationship between corporate tech departments and financial institution business units [2]. More important, the change is coming at a time when costs and regulatory compliance are high priorities; according to the Boston-based Aité Group [2].The Aité Group found that 50% of those surveyed responded that they were likely or highly likely to use private clouds in the next 24 months [2].

The technology consulting firm Gartner has forecasted that, by 2016, poor return on equity will drive more than 60% of banks worldwide to process the majority of their transactions in the cloud [3]. So we can see that banks are very much interested to deploy cloud computing for banking tasks. However the biggest threat in implementing cloud technologies is security. The cloud acts as a big black box, nothing inside the cloud is visible to the clients [4]. Clients have no idea or control over what happens inside a cloud [4]. Even if the cloud provider is honest, it can have malicious system persons who can tamper with the VMs and violate confidentiality and integrity. In a survey conducted by IDC Enterprise Panel 74.6% of the participants identified security as the biggest threat to cloud computing [5]. The system has been developed with a view to eradicate security problems and make the cloud more acceptable to the banks. The system also aims at providing a simple and efficient cloud structure and making baking available to all places by utilizing the potential of cloud computing.

III. PROPOSED SYSTEM

We are proposing a private cloud platform for the banking systems. It will enable the banks to control a worldwide banking network in a centralized and efficient manner. It will also enable the banks to introduce modular banking. So the banks will be able to explore new territories and consumers.

The main concept behind the modular bank is, “If you cannot come to the bank, the bank will come to you”. It will bring business expansion as well as the banks will be ready to face the challenges of this new era.

A. Architectural Components

Our proposed system primarily is composed of User ends, local gateways and cloud central system. The architecture is depicted in Fig. 1.

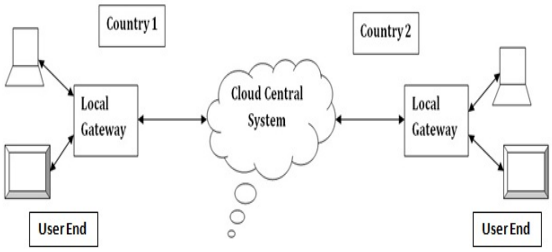


Fig. 1 Cloud Computing Architecture

1) Cloud Central System: The cloud central system is the data center that provides performs two tasks. One is data processing and another is data storage. This central system is the heart of the entire cloud platform. It is connected to the local gateways that are specified for each country. The data to be processed is passed from the user ends via the gateway to the central system. After that the processed data is relayed back to the respective user end. The central system consists of an array of processors that are connected using a master client method.

2) *Local Gateway*: For a bank that is globally operating in different countries, each country will have its own local gateway. The traffic generated from each user end will be filtered through the gateway to maintain authenticity and security. Since different countries have different time slots depending on their geographical location. The same cluster of processors can be used to satisfy different countries.

3) *User End:* The user end will include a thin client that can access the network. It can be possible a battery powered handheld device such as a smart phone, a tablet or a low end notebook computer. The user end will be provided an interface so that the device can use to connect itself to the cloud system. The interface must be accessible through a thin client such as a browser. The required processing of data will be done at the central system as requested by the end user. Then the result will be displayed on the end user through the thin client. So the entire processing will be done at the central system and the end user end will do an absolute minimum processing only to grab input and display the output.

B. Cloud Standards

1) *Private Cloud*: The proposed system will contain a central system as mentioned earlier that will be implemented using private cloud standards. Since the data of a bank is very sensitive and confidential the banks cannot outsource the data to a third party cloud. So they will have to create a cluster of processors in their own premises at a certain location. This cluster will act as the backbone of the private cloud and will perform all the processing tasks. The cloud has been developed in a master slave paradigm. And the central database will be stored in the master that the slaves can readily access. So the data will be safe and secure. The processors will communicate with the master using a private network and connect to an end user using a public network. Whenever an end user is authenticated it will be allocated to a virtualized instance of a physical machine. And the end will have access to a set of predefined resources. It will be available for the user for processing his required data and the necessary update will be done in the database in the master. After completing the desired when the end user logs out from the system the allocated resources will be released. And the released resources will be in a ready state to be allocated to another end user.

2) *Software as a Service (SAAS):* The proposed system implements the Software as a service (SAAS) as the cloud service. Since the banking tasks are fairly simple and the employees only requires a simple set of tools SAAS is the best solution for them. It reduces the complexity of the applications and provides a friendly interface to the users. Initially an employee logs in with his credentials. If his credentials are valid then he is provided a dedicated instance on which he can operate. The interface must be accessible through a thin client. Then the task is carried out by the instance in a virtual machine. Since the instance is virtual in nature it can be booted instantly. As a result there will be minimum lag in the service provided to the employee. The employee utilizes the virtual machine as if it were a physical machine. Moreover a single physical machine can be segmented into a number of virtual machines. So a single machine can serve a number of employees simultaneously. It maximizes the utilization of a machine and reduces the idle time.

C. Proposed Cloud Architecture

1) *Request Initialization Procedure*: According to our propose architecture each user terminal communicates the central system for receiving services from cloud sides. The procedure is depicted in Fig. 2.

The steps are described below.

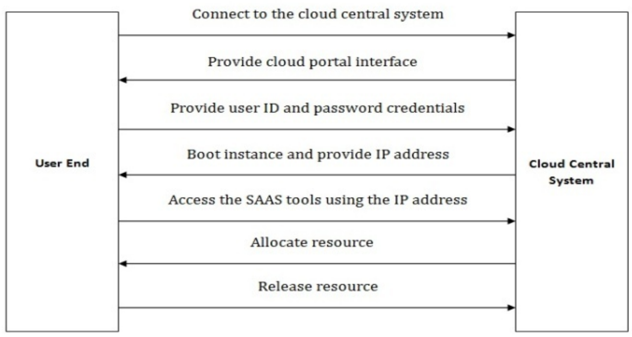


Fig. 2 Steps of Communication between User End and Central System

- First of all the user end device connects to the cloud central system data center using the address of the portal
- After that the controller node of the cloud provides an interface for the user to provide his username and password for authentication
- Using the provided interface the user provides his username and password credentials.
- After confirming user’s authenticity an instance is booted and the native IP address is provided to the user.

- Using the provided IP address the user can access the SAAS tools provided within the instance.
- The instance is preconfigured according to a fixed amount of resources and it is available for the user to perform his desired task.
- After performing the required tasks when the user logs out from the system the resources are released and ready to be allocated to another end user.

2) Resource Mapping: The primary aim of the cloud based banking system is to make proper and efficient utilization of the resources. And for that the resources must be mapped in an orderly and planned manner. Since the instances use virtualization technique we can allocate more virtualized resources with limited amount of physical resources. For example let us assume we have a physical machine containing 2 CPU and 2 GB RAM. If the applications within the cloud require up to 1 VCPU and 1 GB VRAM then we can allocate up to 4 instances in this machine. Since the cloud user virtualization the instances dynamically uses the resources and run smoothly simultaneously. So we can double the utilization.

3) Resource Allocation: In the proposed system the resource allocation follows a first come first served method. Whenever an end user requests access he is provided an instance if instances are free. If no instance is currently free the requesting user has to wait until an instance is released. In terms of data centers a set of processors stacked up horizontally is called a rack of servers. Moreover a set of racks is called a cluster. In our proposed system the allocation of resources follows a sequential manner. For example let us assume a scenario where no instance is running. Whenever a user requests an instance the first rack is powered up and an instance from the first physical machine is allocated. When a physical machine is saturated the next physical machine is powered up and resources are allocated. Similarly when a rack is saturated the next rack is powered up and instances are allocated.

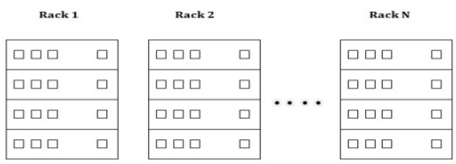


Fig. 3 Cloud Data Center Clusters

This procedure ensures the minimum power requirement and maximum utilization. Likewise if the entire cluster is overloaded additional users may have to wait for resources. In that case we can easily add new cluster to this system as per our requirement and reduce waiting time and ensure smooth service. So this system also exhibits horizontal scalability. It also ensured dynamic and efficient utilization with maximum throughput.

D. Load Balancing

Cloud load balancing is the process of distributing workloads across multiple computing resources. Generally the load distributions of the banking tasks are homogeneous in nature. Since our proposed system includes a SAAS model to provide the functionalities the computational load in almost uniform in nature. Thus if we distribute the processes among different virtual machines the system becomes unnecessarily complex. As a solution we are proposing a simpler and efficient approach to load balancing problem. This solution takes advantage of the virtualization technique. We can allocate a fixed number of virtual machines for each physical machine depending on its actual resources. Then whenever a request comes to the master node it will take decision where to put the request. The master node checks from the first rack to see if there is a free virtual machine. If it finds an unused resource it is allocated to the requesting node. Similarly the next request is allocated to a successive virtual machine and so on. The algorithm is provided below.

Algorithm 1: Load Balancing

Let us assume we have R number of racks where each rack contains Sn number of physical machines. Again each physical server contains Vn number of virtual machines.

procedure Allocation(R, S_n, V_n)
    i : = 1
    while i < = the number of racks, R **do**
        j : = 1
        while j<=the number of physical machines, S_n **do**
            k : = 1
            while k<=the number of virtual machines, V_n **do**
                if V_k is unallocated
                    allocate V_k to the requesting user
                    return
                end if
            end while
        end while
    end while
end procedure

E. Modular Banking

The amazing flexible architecture and powerful processing capabilities of the cloud infrastructure allowed us to propose a new territory in banking called modular banking. This proposed system will be very useful to expand in remote places where banking infrastructure is neither possible nor feasible. It will be particularly useful to the developing and underdeveloped countries.

1) *Proposed System*: Since the employees if a bank is connected to the central cloud server through the Internet there is no need for intranetworking within the employees. As a result the network within a branch becomes very simplified. So we are proposing a mobile modular bank that will be able to reach remote and impracticable places. So the banking facility can to the farthest corners of a country and people of all areas will be able to receive banking services.

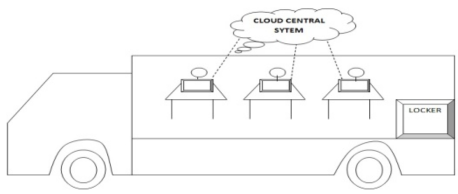


Fig. 4 Modular Bank

2) *System Architecture:* In our proposed system a simple modular bank will be mounted on a moving vehicle possibly

in a truck or a van. All the employees will be connected to the system through a wireless connection. Each employee can access the cloud independently as per his requirement. The employees only need to carry a handheld device to do the banking job. Moreover the handheld devices are battery powered. There is no requirement for a power station.

3) *Time Slot Allowance*: Setting up Modular banks allows us to utilize time slots for different areas. For example if a modular bank has to cover an area as depicted in Fig. 5 we can divide the area into 2 zones. Then we can allow 4 banking hours for each zone. The banking hour starts at 9 am. From 9 am to 1 pm the modular bank will be at a specific place within zone 1 to provide service. From 1 pm to 5 pm the modular bank will be at a specific place within zone 2 to provide service and. So the modular bank will cover an entire area within a day.

IV. Comparative Analysis of the proposed system

A. Private Cloud Ensures Data Confidentiality

The proposed system is based on private cloud architecture. The private cloud architecture keeps the native data within the bank's premises. So the confidentiality of the sensitive data of the consumers is maintained. The data does not need to be outsourced.

B. Less Expensive Compared to Other Systems

In conventional banking system a local area network (LAN) needs to be set up in each branch. But in our proposed system there is no need for any LAN. So the system becomes much simpler. The cabling cost to set up a LAN is eliminated. The entire processing required for the employees of a bank done centrally in the cloud server. So there is no need for setting expensive personalized machine for each employee. When a bank branch is to be established the bank has to invest a lot of capital on buying heavy workstations and a network within them which is not required in this system. So less expense is required.

C. Central Data Processing and Scalability

A globally distributed international bank requires a huge amount of data storage. In conventional method each country or bank branch needs to maintain a separate database. In our proposed system only a single centralized database is required globally. We can add or remove physical machines easily as per our need easily. If the current processing power of the cluster is not enough then we can easily add one more rack of processors to increase ability. This is called horizontal scalability.

D. Increased Data Integrity and Reliability with Less Overhead

In a conventional method the banking data is stored and processed in different ways using different application. So data integrity is lost. But in our proposed system the same application and data structure is used globally. So data integrity is maintained. In a conventional system when the workstation fails it needs to be either repaired or repaired and the employee has to sit idle. Since the machines in the cloud are virtual in nature if a machine fails the user will be instantly provided another machine. Since the banking tasks are uniform we can allocate a user a specific amount of predefined resources. So it produces less overhead compared to other complex algorithms.

E. Efficient Utilization of Resources

The resources are allocated on demand basis. After performing required tasks an employee releases the resources ready to be used by another employee. This ensures maximum utilization. Every country has a different banking hour for geographical reason. For example banking hour in Bangladesh starts at 9 am whereas when it's morning in Bangladesh it is night in USA and the banking hour is closed there. So the same machine can serve both these countries throughout a day. So the resources are utilized round the clock. The applications are accessible through a thin browser client. So connection to this system is possible even for low end devices.

F. Expose Service to Low Earning Consumers

In the developing countries most of the rural and remote people are have a margin of low earning. So coming to the bank at a great distance and deposit their hard earned money is not possible for them. So they are deprived of banking facilities. In proposed system the bank itself is going to them on a regular basis. So the banking services also available even for low earning consumers. And more people will receive banking facilities.

V. Conclusions

Cloud technology is the newest and most advanced technology in IT sector that emerged with a lot of potentials and great future benefits. The paper presents an overview of the architecture that the banks may implement to move towards cloud and be benefited from this.

This system gives a unified combination of the latest cloud technologies to introduce cloud in banking. Employees of the banks can easily access the cloud and get the required tools. This paper also introduces modular banking that will enable to reach out more consumers and provide a convenient method for banking.

References

[1] Manisha B. Jadhav, Vishnu J. Gaikwad, C. V. Patil, G. S. Deshpande, "Cloud Computing Applications in Computer Science", International Journal of Advanced Computer and Mathematical Sciences.Vol 1, Issue 1, Dec, 2010, pp 1-6.

[2] Tom Groenfeldt (2014), "Some Banks Are Heading To The Cloud – More Are Planning To". [Online]. Available: http://www.forbes.com/sites/tomgroenfeldt/2014/06/26/some-banks-are-heading-to-the-cloud-more-are-planning-to.

[3] Rob van der Meulen, Janessa Rivera, "Gartner Identifies Top Vertical Industry Predictions for IT Organizations for 2014". [Online]. Available: http://www.gartner.com/newsroom/id/2607115.

[4] JBV Subramanyam, Kokula Krishna Hari K, "The Proceedings of the International Conference on Information Engineering, Management and Security", ICIEMS 2014.

[5] IDC Enterprise Panel (August 2008).

[6] Shahid Al Noor, Golam Mustafa, Shaiful Alam Chowdhury, Md. Zakir Hossain, Fariha Tasmin Jaigirdar, "A Proposed Architecture of Cloud Computing for Education system in Bangladesh", IJCSNS International Journal of Computer Science and Network Security, VOL.10 No.10, October 2010.

[7] M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, "A Berkeley View of Cloud computing", Technical Report No. UCB/EECS-2009-28, University of California at Berkley, USA, Feb. 10, 2009.

1st International Conference on Computer & Information Engineering, 26-27 November, 2015
Organizer : Dept. of CSE, Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh

Closest Class Measure based Subspace Detection for Hyperspectral Image Classification

M. A. Hossain[1]
Rajshahi University of Engineering and Technology
University of New South Wales Canberra
Rajshahi, Bangladesh
ali.ruet03@yahoo.com[1]

S.U. Zaman, M. A. Mamun and M. N. I. Mondal
Rajshahi University of Engineering and Technology
Rajshahi, Bangladesh

***Abstract*— The objective of this study is to develop a hybrid nonlinear subspace detection technique in which Kernel Principal Component Analysis (KPCA) is combined with a Closest Class Pair (CCP) measure for the task of hyperspectral image classification. In the proposed approach, KPCA is applied first to generate the new features from original dataset then a the CCP is applied to rank the features that are able to separate the complex or overlapping classes. Finally, the two ranked score such as KPCA and CCP is combined to select a subset of features that are relevant and able to better discrimination the input classes of interest. Experiments are performed on a real hyperspectral image acquired by the NASA Airborne Visible Infrared Imaging Spectrometer (AVIRIS) sensor and it can be seen that the proposed approach obtained the best classification accuracy 84.58%.**

Keywords*— *Hyperspectral image, nonlinear feature extraction, feature selection, image classification, statistical distance and kernel principal component analysis.

I. Introduction

The amount of image data that is received from satellite is constantly increasing day by day as a result of advancement in imaging sensors. Imaging sensor in recent days can capture data in hundreds of contiguous narrow spectral bands to produce hyperspectral images. For example, the NASA AVIRIS instrument collects data from 224 channels with spectral resolution 0.01μm spanning the ultraviolet to near infrared. Recently, feature extraction/selection and content-based retrieval have become highly desired goals because this large amount of data presents some complex methodological problems in supervised image classification of remote sensing dataset [1]. For instance, the classification of this large amount of data is time consuming and requires significant computational effort which may not be possible in many applications. Moreover, if the training samples are limited, a reduction in the classification accuracy for the test data is observed due to the poor generalization of the training results. This effect is known as the curse of dimensionality/Hughes phenomenon [2]. On the other hand, hyperspectral sensors capture data in a very close and contiguous spectral range to fill the gaps of multispectral sensors where some bands are redundant and not important for a specific application. The above problem can be solved by keeping the number of features to as few as possible for effective classification. This can be achieved by selecting a subset of features which have high contribution to the separation of spectral classes. In literature [3-5], feature extraction is proposed for dimensionality in which data is transformed to a low dimensional space so that the new subspace is adequate to represent the original meaning. Feature extraction can be supervised and unsupervised, linear and nonlinear [4, 6]. The commonly used feature extraction methods are Linear Discriminant Analysis (LDA) [7] and Principal Component Analysis (PCA) [8]. The LDA is a popular supervised, linear feature extraction method. But it is limited to only normally distributed class data and generates only C-1 new features, where C is the number of input classes of interest. PCA is another frequently used unsupervised, linear feature extraction method. The objective of PCA is to find an optimal subspace where the output features are chosen based on the value of variance from high to low. However PCA is quite limited since it cannot represent nonlinear data structures. In recent years an advanced nonlinear version of PCA has been proposed which is known as Kernel Principal Component Analysis (KPCA). Kernel PCA maps original data to a high dimensional space depending on a suitable kernel function such as Gaussian or Polynomial [9, 10]. The application of linear PCA in this new space is possible and can handle data which is not linearly separable in the original input space. KPCA is also capable of capturing higher order statistics thus provide better representation of the information in the original data set. However, the KPCA approach discussed above depends on global variance but there may have some classes which are small and hardly affect the overall statistics and are not represented by the first few components. In this paper we extended the application of a kernel based data transformation by taking into account the separation between the closest class pair and named this approach as KPCA-CCP. We propose to select a subset of kernel principal components that can provide good separation among the input classes of interest.

II. Feature Extraction Approach

A. Kernel Principal Component Analysis

KPCA is a nonlinear feature extraction technique usually applied when the input data contains linearly inseparable classes [9]. For an input hyperspectral image $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \dots \mathbf{x}_N]$, $\mathbf{x}_i \in R^M$ where M is the number of spectral bands and N is the number of pixels in an image band, the KPCA performs the data transformation by mapping the input data to a new feature space using a nonlinear kernel function such as a Gaussian or Polynomial [10]:

$$\phi : \mathbf{X} \rightarrow \mathbf{H}$$
$$x \rightarrow \phi(x)$$

where is a function that maps the input samples to a new feature space/high dimensional space **H** using a nonlinear kernel in hand (e.g., Gaussian, Polynomial etc.). Each entry k_{ij} of the kernel matrix **K** (size: N x N) is generated by applying the defined kernel function to the input samples $\mathbf{x}_i$ and $\mathbf{x}_j$. Then the linear principal components transform is applied by solving the eigen decomposition of the kernel matrix **K** [9].

$$\lambda \boldsymbol{\alpha} = \mathbf{K}\boldsymbol{\alpha}, \text{ subject to } \|\boldsymbol{\alpha}\|_2 = \frac{1}{\lambda} \qquad (1)$$

where λ is the set of nonzero eigen-values corresponds to the eigen vectors **α.** The i^{th} new feature can be obtained from $\mathbf{z}_i = \boldsymbol{\alpha}_i^t\mathbf{K}$ where $\boldsymbol{\alpha}_i^t$ is the transpose of the i^{th} eigenvector. For the low dimensional projection the first few eigenvectors are chosen based on the egien-values from high to low. The new features generated by the kernel principal component analysis are uncorrelated and ordered based on the higher value of variance [10,11].

B. Supervised Statistical Distance Based Feature Selection

Although KPCA rank the feature based on the high value of global variance, if training samples are available, it can be utilized to select features which provide good separation among the input classes of interest. In this study, a supervised feature selection is performed on KPCA images based on the statistical distance measure between the probability densities of the input classes. To measure the separation among the input classes, the pair-wise statistical distance distance [12] is applied under a Gaussian assumption for the distribution of input classes. The statistical distances among all the class pairs are measured for each kernel principal component. The Jeffries-Matusita distance is used and effective because the input features are already uncorrelated as discussed in section IIA. The proposed feature selection procedure is described as follows.

1. Input feature are ranked based on the smallest value of separation among all the class pairs of each input feature so that each image is examined to determine how well it can separate the complex or overlapping class pairs [4].
2. If the classes are highly overlapped and the largest separation is very small and less than a user defined threshold T, then feature selection is performed based on the mean separation among the classes to aim for a global measure.
3. The rank scores from high variance and closest class pair are combined to obtain the final list of features. Finally the resultant features are selected based on the low value of the final score. Thus, the output features is uncorrelated and ordered based on their class discrimination ability as well.

There are several benefits of applying Kernel Principal Component Analysis before the class pair wise treatment. For example, it uses second-order statistics, and generates the new images based on the combination of original dataset. So the resultant subspace are uncorrelated, contains most of the input variance and good class discrimination ability. This is the effectiveness of the proposed method. The distribution of the input classes can also be examined before feature selection. However, if the data are not normally distributed after transformation, a non-parametric measure such as mutual information can be utilized [3, 13]. The output features selected in this way are highly meaningful and relevant for the classification of input classes.

III. INPUT DATA AND EXPERIMENTAL ANALYSIS

A. Experimental Procedure

To assess the performance of the proposed approach a comparison has been made, using real hyperspectral images, between the proposed method and two other relevant methods taken from the recent literature [14, 15]. This input hyperspectral image was collected with a 20 *m* spatial resolution and 10 *nm* spectral resolution in the wavelength of 0.4μm to 2.5μm. It has 220 channels and each channel contains 145×145 pixels [16]. The original data contains 16 spectral classes. Hence a subset of the input classes are used when supervised methods are tested and a few classes such as Grass/Pasture mowed, Oats, Alfalfa and Stone-steel Tower are not considered as they have insufficient training samples and do not provide overall representative results.

TABLE I: INPUT CLASSES AND THEIR NUMBER OF SAMPLES

Class name	Training samples	Test samples
Hay-windrowed	30	18
Soybean-notill	48	48
Woods	60	60
Wheat	26	26
Grass/Trees	20	20
Corn-notill	28	28
Soybean-min	42	42
Soybean-clean	33	33
Grass/Pasture	30	30

This hyperspectral image was used in this experiment as it contains different types of classes and it is increasingly used in many studies due to the availability of its ground truth reference image [16]. The KPCA approach presented in this study used a Radial Basis Function (RBF) as the kernel [17] with kernel parameter (σ =5.5). The two resultant ranking scores from KPCA and closest class pair distance measure are combined to generate the resultant order of the selected features. The final order of the selected feature for the input hyperspectral image is listed in the following TABLE II. The value of the distances between the most complex or overlapping class pair for each of the first 10 KPCs are shown by the graph in Fig 1. It can be seen that the closest class pair distances are very small even the largest of the minimum is

0.15 which is less than the minimum required value 0.40 (in scale of 2.0) and reflects that the best feature cannot provide the discrimination between the closest class pair. So the input features are ranked based on the mean distance instead from high to low. This global measure can order the features which can provide good separation between the remaining class pairs.

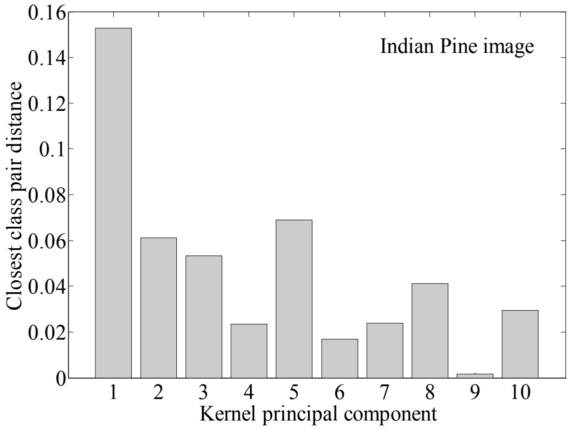


Fig 1. Statistical distance between the closest class pair for each kernel principal component image.

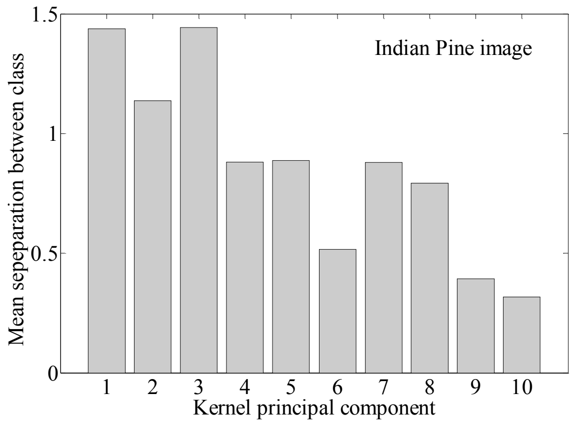


Fig 2. Mean of the pair wise separation for each kernel principal component image.

Finally the two the ranking scores (i.e., based on high variance and closes class pair) are combined to obtain the final rank and a subset of features with lower rank scores is selected. The smallest value of the combined rank score is selected as the best feature in the KPCA-CCP method.

TABLE II: ORDER OF THE SELECTED FEATURES

Method	Order of the input features
KPCA	KPC: 1, 2, 3, 4, 5, 6
KPCA-CCP	KPC: 1, 3, 2, 5, 4, 7
PCA-CCP	PC: 2, 1, 3, 4, 7, 5

The resultant order of the KPCA features for the combined score is shown in TABLE II. It can be seen from TABLE II that kernel principal component 3 is selected as the 2nd best feature instead of kernel principal component 2 for the proposed approach. This is because it has better spatial structure of the input classes and provides good class discrimination. This can also be realized visually from Fig 3 & 4.

B. Experimental Results

The performance of the proposed method KPCA-CCP is compared with the following baseline approach in terms of classification accuracy:

- KPCA: Kernel Principal Component Analysis described in section IIA.
- PCA-CCP: Conventional Principal Component Analysis with Closest Class Pair Distance.

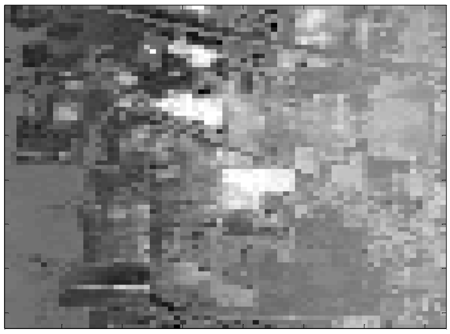


Fig 3. Kernel principal component image 2.

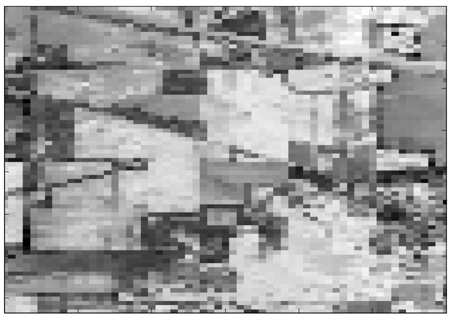


Fig 4. Kernel principal component image 3.

The performances of the feature selection is compared using the classification results obtained by a supervised linear SVM classifier [18]. In the experiment liner SVM is used instead of the kernel SVM in order to investigate whether the data are linearly classified after applying kernel PCA. The input classifier is trained by a cross-validation process using the samples defined on TABLE I. The SVM cost parameter C= 5.12 is selected after 10-fold cross-validation. A comparative result between the proposed and baseline method is presented in TABLE III. For each method, overall mean accuracy and its standard deviation are measured after 5 runs. In each run, 80% of the training samples are selected using the Monte Carlo approach [12] for the training of the SVM classifier.

TABLE III. OVERALL CLASSIFICATION ACCURACIES OBTAINED BY PROPOSED AND BASELINE APPROACHES FOR THE INDIAN PINE AVIRIS HYPERSPECTRAL IMAGE.

Method	Mean accuracy	Standard deviation
PCA-CCP	73.57 %	0.82
KPCA	82.61 %	0.61
KPCA-CCP	84.58 %	0.22

The experimental results show that the proposed KPCA-CCP approach provides best accuracy from the conventional KPCA and PCA-CCP methods. This is because the KPCA-CCP method selects a subset of features which can provide better discrimination among the input classes of interest. An improvement in accuracy is also obtained using KPCA compared to PCA-CCP which is another indication of the effectiveness of the kernel based approach.

IV. CONCLUSION

In this paper the benefits of feature selection using closest class pair measure over KPCA images is demonstrated. The selection criterion for the proposed method is based on maximizing the separation between the input classes of interest. The incorporation of the kernel method in the proposed approach makes it suitable for problems in which the input classes have complex structures and linearly inseparable patterns. Therefore, the proposed methods are more robust and useful when a linear method is inappropriate.

Acknowledgment

The AVIRIS Indian Pine data were provided by Prof. David A. Landgrebe of Purdue University, W. Lafayette, Indiana, USA. Access to this dataset and its Ground Truth reference image is gratefully appreciated.

References

[1] J. A. Benediktsson, J. R. Sveinsson, and K. Arnason, "Classification and feature extraction of AVIRIS data," *IEEE Transaction on Geoscience and Remote Sensing,* vol. 33 no. 5, pp. 1194-1205, 1995.

[2] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory,* vol. It-14 no. 1, pp. 55-63, 1968.

[3] M. A. Hossain, X. Jia, and M. Pickering, "Subspace detection using a mutual information measure for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters,* 2012.

[4] J. Yang, P. Yu, and B. Kuo, "A nonparametric feature extraction and its application to nearest neighbour classification for hyperspectral image data," *IEEE Transaction on Geoscience and Remote Sensing,* vol. 48 no. 3, pp. 1279-1293, 2010.

[5] X. Jia, B. Kuo, and M. M. Crawford, "Feature mining for hyperspectral image classification," *Proceedings of the IEEE,* vol. 101, pp. 676-697, 2013.

[6] S. Tajudin and D. A. Landgrebe, "Covariance estimation with limited training samples," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 37, pp. 2113-2118, 1999.

[7] K. Fukunaga, *Introduction to Statistical Pattern Recognition*: Academic Press, New York 1990.

[8] L. Ying, G. Yanfeng, and Z. Ye, "Hyperspectral feature extraction using selective PCA based on genetic algorithm with subgroups," in *International Conference on Innovative Computing, Information and Control, ICICIC '06*, 2006.

[9] B. Schölkopf, A. Smola, and K. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation,* vol. 10, no. 5, pp. 1299-1319, 1998.

[10] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Kernel principal component analysis for classification of hyperspectral remote sensing data over urban areas," *EURASIP Journal on Advances in Signal Processing,* vol. 2009, pp. 1-14, 2009.

[11] M. A. Hossain, J. Xiuping, and M. Pickering, "Incorporating spatial properties in subspace detection," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, 2013, pp. 3986-3989.

[12] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis, 4th Edition*: Springer-Verlag Berlin Heidelberg, Germany, 2006.

[13] M. A. Hossain, M. Pickering, and X. Jia, "Unsupervised feature extraction based on a mutual information for hyperspectral image classification," *IGARSS 2011, Vancouver,* pp. 1720-1723, 2010.

[14] F. R. L. Bruzzone and S. B. Serpico, "An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection," *IEEE Transaction on Geoscience and Remote Sensing,* vol. 33 no. 6, pp. 1318-13121, 1995.

[15] X. Guorong, C. Peiqi, and W. Minhui, "Bhattacharyya distance feature selection," in *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*.

[16] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*: A John Wiley and Sons Publication, Hoboken, New Jersy, 2003.

[17] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 43, no. 6, 2005.

[18] C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," *Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm,*2001.

Multi-temporal FFT Regression

Md. Al Mamun, Md. Nazrul Islam Mondal
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
{cse_mamun, nimbd}@yahoo.com

Boshir Ahmed, Md. Shahid Uz Zaman, Shyla Afroge
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
{boshir_bd, szaman22, shyla_cse05}@yahoo.com

***Abstract*— Sequential data transmission regarding multi-temporal image analysis is mainly dependent upon prediction or forecasting. The transmission time can be substantially reduced by properly exploiting the temporal correlation. Multi-temporal images are often affected by sensor, and illumination variations, non-uniform attenuation, atmospheric absorption and other environmental effects which render system changes in them. Most of these changes are gradual and incremental. So a recent image can be predicted from a previous sequence of images if the amount of real land-cover change is limited. Regression based prediction is the most appropriate one in this case as it can quantify the relationships between images obtained by different measurement systems in different environments. FFT regression based temporal prediction is proposed in this paper whereby the least-squares minimization is conducted on the amplitude matrices of the readings via the FFT. For a given model, the value of squared coefficient of determination (*R*) is always increased beyond the value obtained by conventional regression which is a common quality measure of the chosen model.**

Keywords—multi-temporal; fft regression; temporal prediction; entropy

I. INTRODUCTION

In many remote sensing studies, Regression analysis is the most common approach used for the calibration of multi-temporal remote sensed images to quantify the relationships between data obtained by different measurement systems in different environments. The quantification of the functional relationship between images of a given target that were acquired by different sensors is quite problematic because of the high percentage of cross-noise. Typically, the correlation coefficient is quite low, even when the compared images look alike [1][2]. The authors come up with FFT based regression method to increase the coefficient of determination of linear fitting. It can be extended to model and analyse multi-temporal variables for temporal prediction. The key issue in quantifying the relationship between two temporal images is to select a suitable regression model. In this case, the image which has already been transmitted is used for prediction. For a given regression model, the parameters need to be estimated. The most common regression analysis finds parameters to predict the data that best fits the current data in a least-squares sense. This optimisation is based on minimising the sum squared error (SSE). The SSE-based least square is often preferred over other methods due to its computational simplicity and good performance in terms of distortion [3]. The Fourier Transform plays a key role in several major image processing issues. First, the equivalence between continuous band-limited images and discrete images has consequences on image sampling and the aliasing phenomenon, image interpolation, geometrical transforms, etc [4]. Second, many filtering issues are better understood and processed in Fourier domain, leading to classical applications in image smoothing, image deblurring, and pattern recognition [5]. Third, the Fourier Transform, due to its global nature, can be used in various issues like image registration, image quality assessment, watermarking etc [6][7][8].

The modelling or prediction part of any compression approach is the fundamental one as it de-correlates the data and fully de-correlated data are the best candidates for the encoding stage. A FFT based regression model for temporal prediction is developed in this paper for the images which are mainly affected by system or environmental noise. The least-squares minimization is conducted on the amplitude matrices of the images via the FFT. For a given model, The squared coefficient of determination, *R* (ratio of variation due to the model with total variation) is always increased beyond the value obtained by conventional regression at the expense of variance increase of the residual. The FFT based method is a radical departure from classical statistics and has the potential of significantly improving statistical inference in remote sensing [1][2].

II. FFT BASED REGRESSION

A. Temporal Correlation and Regression

The correlation coefficient summarises the strength of the linear association between variables and actually quantifies how consistently two variables vary together. The correlation coefficient between two dates' image data sets $X = \{x\}$ and $Y = \{y\}$, is defined as

$$r_{XY} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{1}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of X and Y, respectively, and n the total number of observations. The regression function is mainly chosen according to the relationship indicated by the temporal correlation [9].

B. Temporal Prediction Model

Given two data sets, $Z = \{X = x_i, Y = y_i \mid i = 1,\ldots,n\}$, where X and Y are taken at time t and $t+1$, respectively, and n is the total number of observations, a regression function, $\psi(.)$, and a residual or error function, $d(.) \geq 0$, regression analysis can solve the following minimisation problem[10].

$$\psi^{opt} = \underset{\psi \in \lambda}{\operatorname{argmin}} \sum_{i=1}^{n} d(\psi(x_i), y_i) \tag{1}$$

where $d(\psi(x_i), y_i) = (\psi(x_i) - y_i)^2$ is the squared difference (residuals) between the modelled result and the true value, and λ a set of candidate regression functions. Considering λ as a general polynomial function, the first-order polynomial is taken for multi-temporal images so the candidate regression function becomes

$$\psi(x_i) = \alpha x_i + \beta \tag{2}$$

The optimisation process (Eq. (1)) is to find the model parameters α and β. In other words, a solution for the unknown parameters, α and β, that will minimise the distance between the actual and predicted values of the dependent variable, Y, needs to be determined. Confidence in the model parameters for explaining and predicting the future outcome can be determined by the coefficient of the determination, R^2. The performance of the temporal prediction can actually be determined by the level of explained variability in the model and R^2 is the indicative.

C. FFT based Prediction

The basic premise of regression analysis requires that the reference ground data must be precise and noiseless. Since in most remote sensing studies this condition is not met, classical regression is not an efficient tool for discovering the true functional relation between remotely-sensed data and ground observations. A new prediction method is proposed whereby the least-squares minimization is conducted on the amplitude matrices of the images via the FFT transformation. The alternative FFT regression method presented herein comprises a two-stage combine approach, whereby the initially low correlation between X and Y is increased and the residuals are dramatically decreased. Pair-wise image transformation is applied to the both image X and Y whereby the correlation coefficient is increased. A predicted image $\psi'(x_i)$ is then derived by least squares minimization between the amplitude matrices of X and Y, via the 2D FFT.

$$\psi^{opt} = \underset{\psi \in \lambda}{\operatorname{argmin}} \sum_{i=1}^{n} d(\psi'(x_i), y_i') \tag{3}$$

where,

$$\psi'(x_i) = abs(fft2(\alpha x_i + \beta)) \tag{4}$$

$$y_i' = abs(fft2(y_i)) \tag{5}$$

and

$$d(\psi'(x_i), y_i') = \left(\psi'(x_i) - y_i'\right)^2 \tag{6}$$

D. Measuring the Regression Confidence by Coefficient of Determination

The coefficient of determination, R^2, defines the accuracy of the model for explaining the data. In statistics, it is used in the context of statistical models with its main purpose being the measurement of how well future outcomes are likely to be predicted by a model based on the previous data. In the case of simple linear regression, it is a measure of variations between the outcome and the values being used for prediction, and is defined as

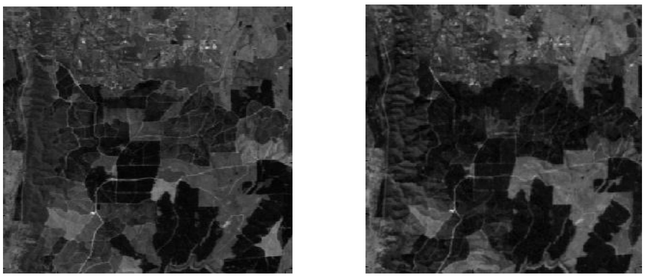
$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i' - \psi'(x_i))^2}{\sum_{i=1}^{n} (y_i' - \bar{y})^2} \tag{7}$$

The value of R^2 provides an insight into the goodness of the fit of the statistical model to the data by explaining any variations.

III. Experimental results

The experiments were performed on geostationary satellite images. The images given in Fig. 1 (a) and (b) are two subsets of Landsat Enhanced Thematic Mapper Plus (ETM+) data recorded over Australia, in 2000 and 2001, respectively. All the six bands with similar resolution are used. The 2D scatter plot (Fig. 2) between the pixel values of the two dates with the overlaid conventional regression approach is shown. The geometrically registered sequential images will be highly correlated when the imaged area contains few land cover changes, with most of the differences being primarily due to variations in the sensing process and atmosphere. The temporal correlation by using Eq. 1 depicts the high dependency. The practical utility of FFT regression will eventually hard to show. Still the model with the scatter plot is shown in Fig. 4 and the corresponding residual resulted from the model is shown in Fig. 5. Entropy and SSE are used as the regression measure as it is an indicator of the potentially lowest number of bits required to transmit image after prediction.

Finally, complete results with variation of R is shown in TABLE I where increase in model fitness is mentionable.



(a) (b)

Fig. 1. ETM+ satellite images taken over Australia in the year 2000 (a) and 2001 (b) of band 3.

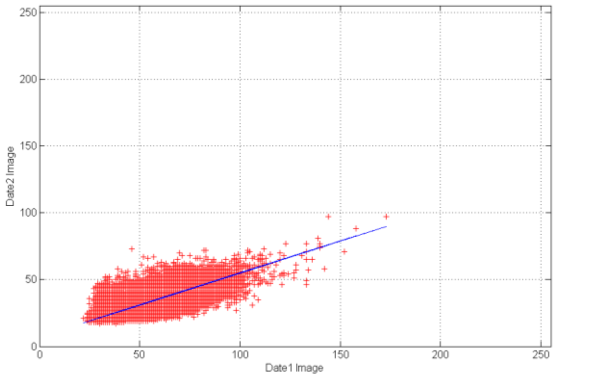


Fig. 2. 2D scatter plot between the two images with regression model overlaid.

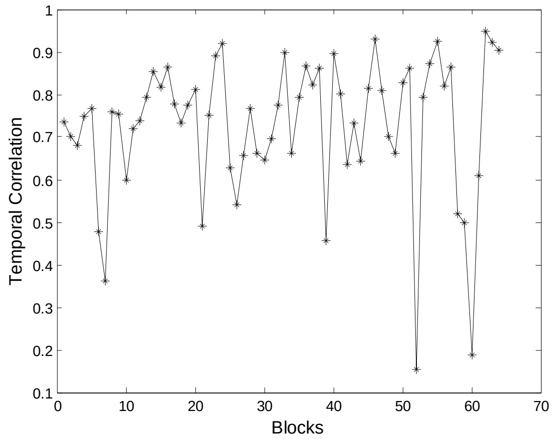


Fig. 3. Temporal correlation between the two images.

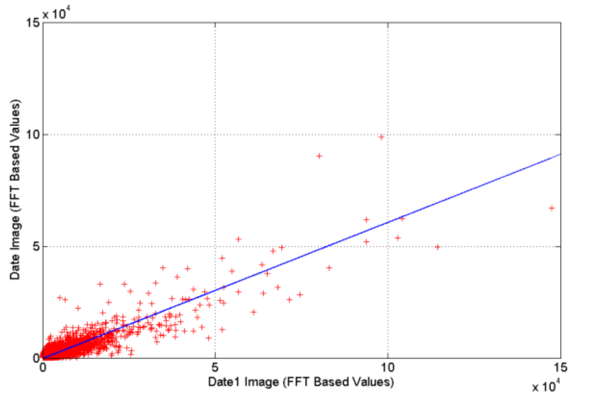


Fig. 4. FFT based prediciton.

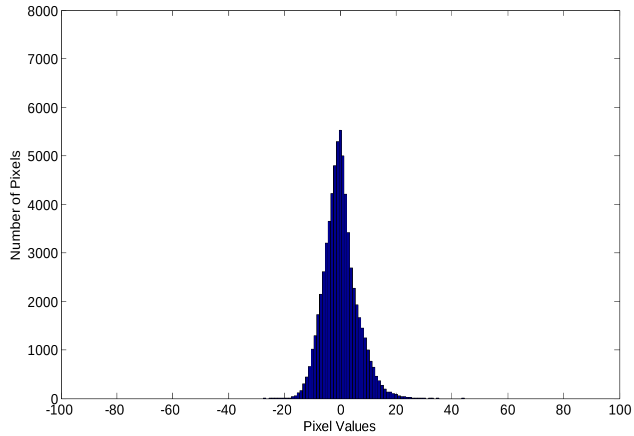


Fig. 5. Residual of the predicted image.

TABLE I. COMPARISON BETWEEN FFT BASED REGRESSION WITH NORMAL METHOD

	Model Parameters (α and β)				**Comparison**					
					Normal Regression			***FFT Regression***		
	Normal		FFT		*E*	*SSE*	*R*	*E*	*SSE*	*R*
Band1	0.4579	14.9783	0.6318	2.7801	3.5796	581913	0.5931	3.7284	696190	1.1260
Band2	0.4761	9.0017	0.6422	0.1152	3.8757	868255	0.6446	4.0122	1061897	1.1671
Band3	0.4788	7.0351	0.4903	7.0121	4.5838	2317270	0.6351	4.5885	2355541	0.6547
Band4	0.6440	19.4211	0.7785	11.4972	4.4659	1952615	0.5221	4.5045	2046854	0.7626
Band5	0.5912	-2.5791	0.5216	4.9494	5.6720	10616166	0.7064	5.6807	11098115	0.5522
Band7	0.5361	2.0450	0.3640	13.1277	5.2812	6297350	0.6171	5.2314	7526017	0.2970

IV. CONCLUSION AND DISCUSSION

The new method proposed whereby entropy is the minimum descriptive complexity of a random variable. The coefficient of determination *R* has increased due to FFT regression compare to regular regression. That means the regression sum of squares or explained sum of squares becomes closer to the total sum of squares which is proportional to the sample or reference data variance. So the variance of the predicted data becomes almost equal to the variance of the original reference data. But in this process the residual data entropy or SSE doesn't improve much, so as entropy. For further illustration to reduce the size of the residual, the phase of the reference data can be locked with the amplitude data from the FFT predicted data [1][2] which will be investigated in future.

V. ACKNOWLEDGMENT

We want to acknowledge Dr. Xiuping Jia for providing us the LANDSAT images.

VI. REFERENCES

[1] K. Peleg, FFT based calibration in remote sensing, *International Journal of Remote Sensing*, No. 19:12, Vol.2301–2315, 1998.

[2] K. Peleg and G. L. Anderson, FFT regression and cross-noise reduction for comparing images in remote sensing, *International Journal of Remote Sensing*, No. 23.10, pp.2097 — 2124, 2002.

[3] Boltz, S., Wolsztynski, E., Debreuve, E., Thierry, E., Barlaud, M., and Pronzato, L. (2006). "A Minimum-Entropy Procedure for Robust Motion Estimation." *IEEE International Conference on Image Processing*, 1249-1252.

[4] C. E. Shannon, “Communication in the presence of noise”, *Proceedings of the Institute of Radio Engineers* 37:1, pp. 10-21, 1949.

[5] Bugao Xu, “Identifying Fabric Structures with Fast Fourier Transform Techniques”, *Textile Research Journal* 66:8, pp. 496-506, 1996.

[6] A. Desolneux, S. Ladjal, L. Moisan, J.-M. Morel, “Dequantizing image orientation”, *IEEE Transactions on Image Processing* 11:10, pp. 1129-1140, 2002.

[7] D. He, Q. Sun, “A Practical Print-scan Resilient Watermarking Scheme”, *Proceedings of the International Conference on Image Processing (ICIP)*, pp. 257-260, 2005.

[8] S. Leprince, S. Barbot, F. Ayoub, J.-P. Avouac, “Automatic and precise orthorectification, coregistration, and subpixel correlation of satellite images, application to ground defor- mation measurements”, *IEEE Transactions on Geoscience and Remote Sensing* 45:6, pp. 1529-1558, 2007.

[9] M. A. Mamun, X. Jia and M. A. Hossain, “Reconstruction of Satellite Images by Multi-temporal Gradient based Sequential Prediction”, IEEE Geoscience and Remote Sensing Symposium (IGARSS ‘14), pp. 1616-1618, Quebec, Canada, July 2014.

[10] M. A. Mamun, X. Jia and M. Ryan, “Non-linear Elastic Model for Flexible Prediction of Remote Sensed Multi-temporal images”. IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 5, pp. 1005 - 1009, May 2014.

[11] C. Lin and Y. Hwang, "An efficient lossless compression scheme for hyperspectral images using two-stage prediction." *IEEE Geoscience and Remote Sensing Letters*, vol. 7(3), pp. 558-562, 2010..

[12] E. Magli, G. Olmo and E. Quacchio, "Optimized onboard lossless and near-lossless compression of hyperspectral data using CALIC." *IEEE Geoscience and Remote Sensing Letters*, vol. 1(1), pp. 21-25, 2004.

Land Cover Classification for Satellite Images based on Normalization Technique and Artificial Neural Network

Boshir Ahmed
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
boshir_bd@yahoo.com

Md. Abdullah Al Noman
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204, Bangladesh
noman093039@gmail.com

***Abstract*— The Satellite images and the extracted thematic maps provide higher-level information for the recognize, monitoring and management of natural resources. It is very difficult to identify land cover classification manually from a satellite image. The remotely-sensed images are invaluable sources of information for various investigations since they provide spatial and temporal information about the nature of earth surface materials and objects. This study aims to determine the level of contributions of multi-temporal and multi-sensor data together with their principal components for Artificial Neural Network classifiers. The suitability of Back Propagation Neural Network (BPNN) for classification of remote sensing images is explored in this paper. Automatic image classification is one of the challenging problems of recent year. BPN is self-adaptive dynamic system which is widely connected with the large amount of neurons. It can solve the regular problem arise from remote sensing images. This paper discusses about the BPNN method to improve the high resolution remote sensing image. The principle and learning algorithm of BPNN is analyzed and high resolution imagery of Beijing has been used. Back Propagation Neural Network classifies the remote sensing image into the classified image of their pattern recognition.**

Keywords*— *landcover classification; artificial neural networks; remote sensing; normalization.

I. Introduction

Satellite images are one of the most powerful and important tools used by the meteorologist. Satellite images have many applications in meteorology, oceanography, fishing, agriculture, biodiversity conservation, forestry, landscape, geology, cartography, regional visible color planning, education, intelligence and warfare [1]. Images can be in and in other spectra. There are also elevation maps, usually made by radar images. Interpretation and analysis of satellite imagery are conducted using specialized remote sensing applications. Like other machine learning methods-systems that learn from data- neural networks have been used to solve a wide variety of tasks that are hard to solve using ordinary rule-based programming, that including computer vision and speech recognition [2]. The basic element of artificial neural networks (ANNs) is the processing node that corresponds to the neuron of the human brain. Each processing node receives and sums a set of input values, and passes this sum through an activation function, providing the output value of the node, which in turn forms one of the inputs to a processing node in the next layer of ANNs. Neural networks are similar to biological neural networks in performing functions collectively and in parallel by the units, rather than there being a clear delineation of subtasks to which various units are assigned [3]. The term "neural network" usually refers to the models employed in. Back Propagation Neural Network is a supervised learning technique for this reason output defined previously. On the basis of output the remote sensing is classified. In this paper we classify the remote sensing image into five categories [3].

II. Material and Study Area

A. Remote Sensing

Remote sensing is the achievement of information about a phenomenon without making physical contact with the object and thus in contrast to on site observations. In modern usage, the term generally refers to the use of aerial sensor technologies to detect and classify objects on Earth (both on the surface, and in the atmosphere and oceans) by means of the propagated signals (E.g. Electromagnetic radiated) [4]. Remote sensing makes it possible to collect data on dangerous or inaccessible areas. Remote sensing applications include monitoring deforestation in areas such as the Amazon Basin, glacial features in Arctic and Antarctic regions, and depth sounding of coastal and ocean depths [4]. Military collection during the Cold War made use of the standoff collection of data about dangerous border areas. Remote sensing also replaces costly and slow data collection on the ground, ensuring in the process that areas or objects are not disturbed [1].

B. Data Processing

Generally speaking, remote sensing works on the principle of the inverse problem. While the object or phenomenon of interest (the state) may not be directly measured, there exists some other variable that can be detected and measured (the observation), which may be related to the object of interest

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

through the use of a data-derived computer model [5]. The common analogy given to describe this is trying to determine the type of animal from its footprints [6].

C. Remote Sensing Region

In this paper Quickbird image is used, which have four multi-spectral bands and a high resolution panchromatic band that contains a plenty of spectral, spatial characteristics [7]. The study area is selected Beijing. Specifically, Haidain district has been used in this paper to improve the classification precision. Take spectral and texture features of images as an input data source of the classification. The image is classified by Back Propagation Neural Network (BPNN) algorithm [8].

III. METHODOLOGY

Like other machine learning methods - systems that learn from data - neural networks have been used to solve a wide variety of tasks that are hard to solve using ordinary rule-based programming, including computer vision and speech recognition [9]. The basic element of artificial neural networks (ANNs) is the processing node that corresponds to the neuron of the human brain [10]. Each processing node receives and sums a set of input values, and passes this sum through activation the function providing the output value of the node, which in turn forms one of the inputs to a processing node in the next layer of ANNs.

TABLE I. INPUT VECTORS AND EXPECTED OUTPUT VECTORS

Input Vectors	Category	Expect Vectors
Sample X_1'	Category 1	(1……..……….0)
Sample X_2'	Category 2	(0,1,...............0)
Sample X_3'	Category 3	(0,0,1,............0)
Sample X_4'	Category 4	(0,0,0,1,….......0)
Sample X_5'	Category 5	(0,0,0,0,1,.…..0)
Sample X_6'	Category 6	(0,0,0,0,0,1,....0)
…	…	…
Sample X_n'	Category n	(0,0,0,………..1)

In this paper, the ground object is classified into five categories. That means n is 5, T is the expect output vector.

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{array}{l} Road \\ Water \\ Vegetation \\ Building \\ Vacant \quad land \end{array}$$

A. Construct a BP Neural Network

Firstly, classify the image into five ground objects so that output neurons are five. After that, determine the number of the neurons in the hidden layer and finally, through some experiments, increase or decrease the number of hidden layer neurons is a suitable way. In this experiment 24 hidden layer neurons have been used [11].

B. The Training Process of BP Neural Network

Assuring the precession statistical property and respective of the sample, select the training sample for classify by visual interpretation and field survey [8]. At the beginning image texture has been used to interpret each ground object and it is corresponding to Table II.

TABLE II. IMAGE INTERPRETATION OF EACH GROUND OBJECT CATEGORY

Ground objects category	Image feature	Ground object description
Road	Light gray	Urban road
Water	Black	River or pond, or sea
Vegetation	Crimson or pink	Flower bed or green field
Building	Bright gray	Urban building
Vacant land	Gray block	Vacant land

C. Training parameter of BP Neural Networks

After repeated training and studying of the samples, 0.8 as the momentum constant value is used to ensure a faster learning the speed, while the value range of the momentum constant is between 0.7 and 0.9. After that set different values for hidden layer, learning layer and output layer, while the hidden layer learning rate is greater than the output layer learning rate [12]. The parameters of the BP neural Network are set as Table III.

TABLE III. TRAINING PARAMETERS OF BPN NETWORKS

Training parameters	Values
Hidden layer neurons	24
Momentum constant	0.7-0.9
Training error	0.05
Maximum iterations	10000

Set the training error as 0.05, maximum number of iterations 10000. If the training error is less, than 0.05 or maximum iterations times reach to 10000, training finish. If the hidden layer increases and the training error decrease, then the result is more accurate, but the time consuming is greater than the experimental time. Same case for the iteration number if increases the iteration number, then the result is getting after few times. However, in the latter circumstance, its training error is still no less than 0.05, then the network structure should be modified with learning rate, momentum constant, or reselect training samples [13].

D. Flow chart of BP Neural Network

In this paper, BP Neural Network Algorithm is divided into two steps. The first step is the training BP Neural Network and the second step is to use the trained BP network to classify the remote sensing image below the figure [14]. Back Propagation Neural Network requires a known input and, desired output for each input value in order to calculate the loss function gradient. It is therefore usually considered to be a supervised learning method, although it is also used in some unsupervised networks such as auto encoders. It is a generalization of the delta rule to multi-layered feed-forward networks [15].

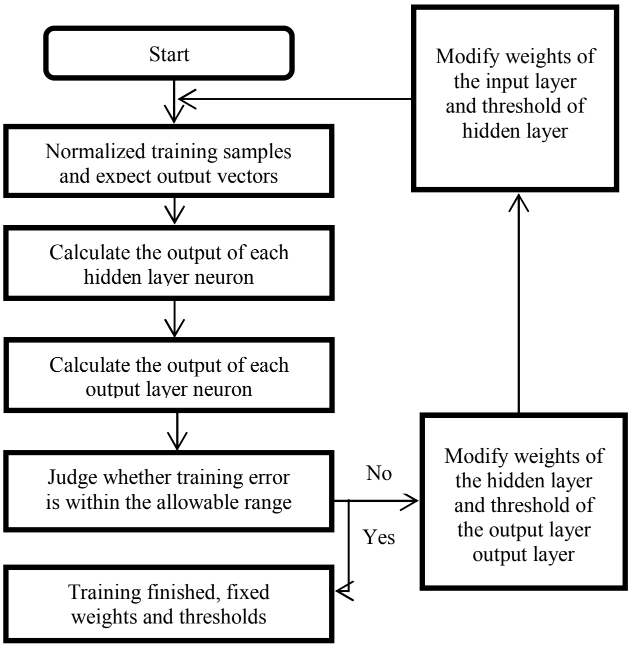


Fig. 1: The first step of the BP classification algorithm.

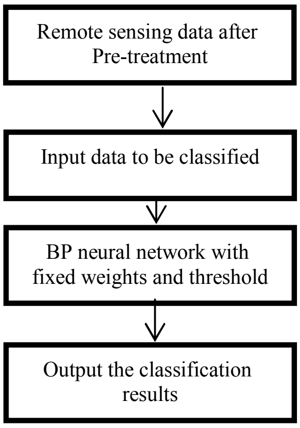


Fig. 2: The second step of the BP classification algorithm.

In neural network toolbox of Matlab, there are many convenient neural network functions the program realizations of the BP neural network are as follows [16]

1. Normalized training samples.
2. Build a BP neural network model and set the parameters.
3. Training BP neural network model and training result.

IV. Experimental Result

The training samples are input to the Back Propagation Neural Network. The Back Propagation algorithm aims to find the set of weights that minimizes the error.

A. The Training Error Curve of BP Neural Network

The training sample is input into the Back Propagation neural Network, which is built by Matlab and then get the training result of remote sensing image. The training curve is shown in below

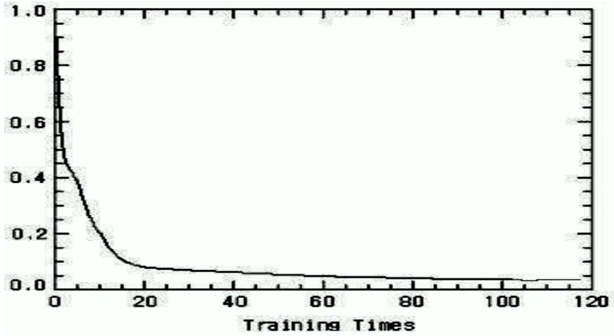


Fig. 3: Training error curve of the BP neural network.

The error curve can be seen after the almost 120 times training, it is not static. It is vary every run time. Training time is found almost around 115 to 150.

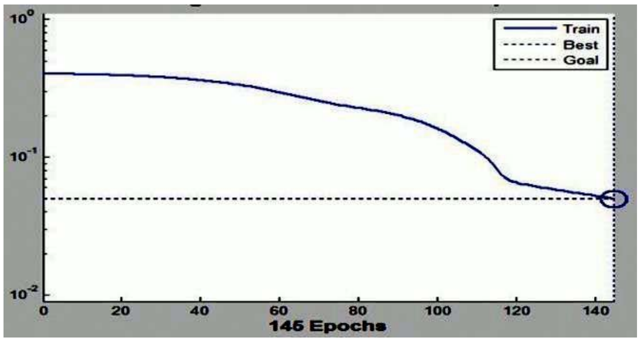


Fig. 4: Training performance

After a number of Epoch 145 the training has been finished. That means in this point threshold, weights are the convergence of learning. The number of Epoch is not fixed. It varies every time. But it is almost around a point. Number of Epoch 145 the learning rate is 11.8154. The learning rate depends on the Epoch number. Vary by Epoch number the learning rate also varies.

B. Classification Results of BP Neural Network

When the Back Propagation Neural Network training success, a Back Propagation Neural Network is got in a accordance with the study area of the image (the fixed number of neurons in each layer, the modified weights and threshold of network, the selected learning rate of each layer, momentum constant).The coverage area of the image is large, so it is clipped out a sub-range of the whole image after analysis. The image size is 600× 600 pixels. The image is given below.

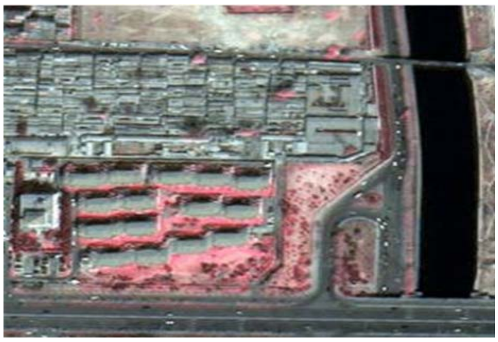


Fig. 5: The image to be classified

Before the input image passes through the Back Propagation Neural Network the remote sensing image converts the grayscale image, then the gray scale image passes through the Back Propagation Neural Network. RGB image size is large and the matrix size is also large. For this reason this is stuff to classify the remote sensing image. For easy to classify first the remote sensing image converts it into a grayscale image. Gray scale image is normalized according to the target or expect output. If the expected output is classified into five categories so the normalized image reduced its size according to the expect output. According to the sampling of the training and learning the network classified the different ground object category image through the memory.

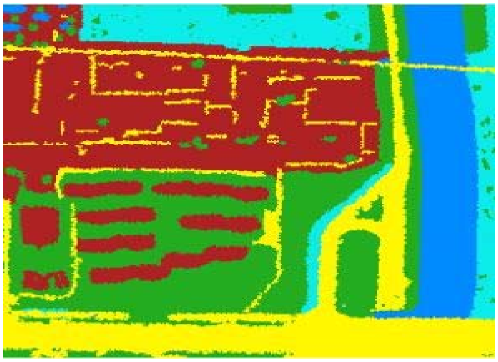


Fig.5: Training simulation of output result

V. Conclusion and Discussions

Manual a satellite image classification is very hard to develop in a region. Since the Back Propagation Neural Network has a limitation, but it is better than other algorithm that is used to classify the high resolution remote sensing image. There are many proposed algorithms to classify land cover classification. Such as SVM (Support Vector Machine), K-means clustering, MLC (Maximum likelihood Classifier), KSON (Kohonen Self Organizing Network), ANNs (Artificial Neural Networks) of Back Propagation Algorithm. But one of them the Back Propagation algorithm is best satellite image classification because it is self-adaptive dynamic system which is widely connected with the large amount of neurons. Artificial Neural Networks (ANNs) have gained increasing popularity as of Remote Sensing Data. An alternative is to statistical methods for classification.

References

[1] N.Musaoglu. Classification of multi-spectral, multi-temporal and multi-sensor images using principal components analysis and artificial neural networks: Beykoz case.

[2] Ashutosh, S.2002. Principle component-based algorithm on multispectral remote sensing data image for spectral discrimination of tree cover form other vegetation types. Current Science, 82(1), pp.67-6.

[3] Civco, D. L., 1993. Artificial neural networks for land cover classification and mapping. International Journal of Geographical Information Systems, 7(2),pp. 173-186.

[4] Morgan, M., 2002. Soil classification in arid lands with thematic mapper data. Terra, 20(2), pp.89-100.

[5] Gopal, S. And Woodcock, C., 1996. Pruning artificial neural networks: an example using land cover classification of multi-sensor images. 20(14), pp. 2787-2803.

[6] Roland Schwaiger, Reinhold Huber and Helmut A. Mayer. Land Cover Classification of Landsat Images Using problem-Adapted Artificial Neural Networks, pp.31-41.

[7] D.Q.Zhu. Principle and Application of Artificial Neural Networks [M]. Science Press, Beijing, 2006, pp.18-20.

[8] C.F.Li. Intelligent processing of remote sensing [M]. Publishing house of electronics industry Beijing, 2007, pp. 3-5.

[9] Y.Xi. Investigation on classification of high resolution satellite image of Back Propagation Neural Network [J], 2006, pp.2-4.

[10] C.C.Wang, W.B.WU, J.P.Zhang. Classification for remote sensing image base on BP neural network [J]. Journal of Liaoning Technical University (Natural Science), 2009,pp.33-35.

[11] Y.S.Zhang,D.C.Gong. Application of high resolution remote sensing satellite imaging model, processing algorithms and application technology [M]. Science Press, Beijing, 2004,pp.23-24.

[12] Martin T. Hagan, Howard B. Demuth, Mark H. Bcale. Neural Network Design [M]. Chin Machine Press, Beijing, 2002, pp.227-228.

[13] Mather, P.M. 1999. Computer Processing of Remotely-Sensed Images: An Introduction, Second Edition, John Wiley & Sons, Chichester, pp.89-90.

[14] Fredric M.Ham Ivica Kstanic. Principle of Neuro-computing for science & Engineering [M]. China Machine Press, Beijing, 2007,pp.77-80.

[15] Y.H.Jia. Application of artificial neural network to classification of multi source remote sensing imagery [J]. Bulletin of surveying and mapping, 2000(07).pp 7-8.

[16] Y.H.Jia, ChunSenZhang, AiPingWang. Classifying of multi sources remote sensing imagery base on BP neural network [J]. Xi' a University of science & Technology Journal,2001,(01). Pp.58-60.

A Novel Idea of Tackling Hard Bit-vector Problems in a Beneficial Way by Eager Solver

Chowdhury Sajadul Islam
Department of Computer Science & Engineering
Independent University, Bangladesh
Dhaka, Bangladesh
chow.cth@gmail.com

Md. Nahid Hasan
Department of Electrical & Electronic Engineering
Rajshahi University of Engineering Technology
Dhaka, Bangladesh
nahidxy@gmail.com

***Abstract*—The theory of fixed-width bit-vectors is a sequence of zero or one bits. The bit-vector theory offers a natural way of encoding the semantics of operations that manipulate binary data, the building block almost all modern computer systems. The predominant approach to deciding the bit-vector theory is via eager reduction to our propositional logic. Bit-vector satisfiability can be reduced to SAT by replacing function symbols by their hardware circuit representation expressed in propositional logic. An eager approach that takes full advantage of the solving power of off the shelf propositional logic solvers. In this paper focuses on the features of the eager bit-vector solver cvcE implemented in the satisfiability modulo theories (SMT) solver CVC4. The goal of paper explores to efficiently solving bit-vector constraints where eager solver cvcE is now an essential option for tackling hard bit-vector problems and evaluated their performance. We proposed a new technique of reducing the size of the bit-blasted formula by factoring out isomorphic sub-circuits to solve the constraints and enhance performance.**

Keywords—bit-vector; binary data; software systems; satisfiability modulo theories (SMT); security, bit-vector

I. Introduction

The process of encoding of a bit-vector formula into propositional logic is called bit-blasting. The satisfiability of bit-vector formulas often rely on eager reduction to propositional logic and do not fit in the conflict-driven clause learning (CDCL) (***T***) framework of general- purpose SMT solvers. The eager solver cvcE by-passes the CDCL(***T***) infrastructure, and encodes the entire Tbv-formula into propositional logic. It then relies on a second SAT solver to decide its satisfiability. Reduction to SAT is appealing as it offers a unified way of reasoning about both the Boolean abstraction and the bit-vector constraints. This enables using the SAT solver as a black box and readily taking advantage of progress in SAT solving. We will call solvers that decide a T_{bv}-formula by bit-blasting the entire formula before starting the SAT search *eager bit-blasting solvers*. These solvers usually do employ algebraic reasoning, but only as a preprocessing step by first apply sophisticated word-level simplifications before bit-blasting.

II. Background

Boolector is a specialized solver for pure bit-vector formulas, it first employs word-level rewriting before bit-blasting the problem into AIG format, followed by conversion to CNF [1]. Recent work improves this approach by using the notion of relevancy to only refine the relevant part of the counter-example [2]. For bit-vector constraints, STP2 first applies word-level simplifications and then encodes the problem to CNF via AIG [3]. It uses the abc [4] AIG package to apply AIG rewriting to the Boolean abstraction of the formula and then encodes the formula to CNF [5]. For example, the following equality $a_{[8]}\&b_{[8]} = c_{[7]}$ o $1_{[1]}$ propagates that the least significant bit of a and b have to be 1: $b[0 : 0] = a[0 : 0] = 1_{[1]}$. Although Z3 is also SMT solver which bit-vector formulas is most similar to that of eager solvers. Z3 first applies word-level rewrite techniques followed by eagerly encoding the formula into propositional logic via bit-blasting [6]. Z3 also uses a technique known as *relevancy* to reduce the size of the problem that is being reasoned about [7].

III. Architecture

Given a quantifier-free T_{bv} input formula ψ, cvcE decides its satisfiability by first applying the preprocessing techniques. The preprocessed formula is converted to propositional logic by the bitblaster and then asserted to a CDCL-style SAT solver. Fig. 1 shows the architecture of the eager bit-vector solver cvcE. The preprocessing module is extended with a special pass, bvToBool, that attempts to lift bit-vector operations over bit-vectors of width 1 to operations over Booleans. For example, consider the formula:

$$\left(\sim (\mathrm{x}[\mathrm{i}:\mathrm{i}])\ \&\ \mathrm{ite}(\mathrm{c}, \mathrm{x}_{[1]}, 0_{[1]}) = 1_{[1]}\right) \wedge \varphi \tag{1}$$

where φ and c are arbitrary formulas, and x is an arbitrary bit-vector term. It is equivalent to the following:

$$\neg(\mathrm{x}[\mathrm{i}:\mathrm{i}] = 1_{[1]}) \wedge \mathrm{ite}(\mathrm{c}, \mathrm{x}_{[1]} = 1_{[1]}, \perp) \wedge \varphi. \tag{2}$$

We can learn $x[i : i] = 0_{[1]} \wedge c$ after applying Boolean circuit propagation. The cvcE solver has very limited support for theory combination via *Ackermannization* [1]. This ackermannize pre-processing pass reduces constraints over the combination of T_{bv} and theory of T_{uf} to T_{bv} as follows:

- For each application $f(\overline{\mathrm{x}})$ where $\overline{\mathrm{x}} = (x_1, . , x_n)$, introduce a fresh variable $f(\overline{\mathrm{x}})$ and use it to replace all occurrences of $f(\overline{\mathrm{x}})$.

• For each $f(\overline{x})$ and $f(\overline{y})$ with $\overline{x} = (x_1, . . , x_n)$ and $(\overline{y}) = (y_1, . . , y_n)$ occurring in the input formula, add the following lemma:

$$\left(\bigwedge_{i=1}^{n} x_i = y_i\right) \Rightarrow f_x = f_y. \tag{3}$$

If the AIG path is enabled, the abc AIG rewriting package [4] is used to simplify the formula and convert it to CNF. If the Boolean path is enabled, the formula is bit-blasted to Nodes and CVC4's Boolean rewrites are applied. The bit-blasted formula is then encoded in CNF using a Tseitin-style conversion [8]. Finally, the CNF formula, obtained through either path, is asserted to the CDCL-style SAT solver SAT_{bb}. Our implementation is based on MiniSAT 2.2.0 [9], but the framework allows for plugging in another SAT solver. To distinguish between the two, we will refer to the CDCL($\boldsymbol{T}$) SAT solver as SAT_{main}.

An alternative implementation could communicate the bit-blasted formula to SAT_{main} by asserting the bit-blasted clauses as $\boldsymbol{T}_{bv}$-lemmas and adding them alongside the clauses that model the Boolean abstraction. In this scenario the $\boldsymbol{T}_{bv}$ solver would be just a "shell" that pushes the bit- blasted formula to SAT_{main}. Our first prototype implementation used this idea. First, using a separate SAT solver for bit-vector reasoning allows for configuring the solver heuristic for bit-vector constraints. Second, SAT_{main} is limited in the number of SAT preprocessing techniques it can employ soundly: it is not usually sound to eliminate a $\boldsymbol{T}$-literal. Using SAT_{bb} allows for applying the MiniSAT preprocessing techniques, such as subsumption and variable elimination [9, 10]. Lastly, SAT_{main} requires non-trivial modifications to be integrated in CDCL($\boldsymbol{T}$). These modifications can add unnecessary overhead and restrict the functionality of the solver.

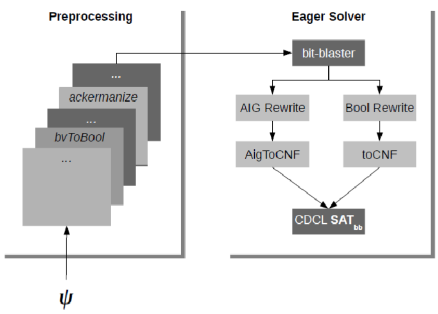


Fig. 1. cvcE architecture

IV. Bit-blasting

In CVC4, bit-blasted terms are represented as a n-tuple of Boolean formulas, each representing the corresponding bit:

$$bbTerm\left(t_{[n]}\right) \equiv \langle b_{n-1}, \ . \ . \ . \ , \ b_0 \rangle \tag{4}$$

where b_i is the i^{th} bit of $t_{[n]}$. For a variable term, each bit is a fresh Boolean variable. For constant terms zero bits are false and one bits are true. Algorithm 1 shows the pseudo-code for bit-blasting the bit-wise and operator &:

Algorithm 1: Bit-blasting bit-wise and.

Input: $t_{[n]} \& t'_{[n]}$

$\langle b_{n-1}, \ . \ . \ . \ , \ b_0 \rangle \leftarrow bbTerm\left(t_{[n]}\right);$

$\langle b'_{n-1}, \ . \ . \ . \ , \ b'_0 \rangle \leftarrow bbTerm\left(t'_{[n]}\right);$

return $\langle b_{n-1} \wedge b'_{n-1}, ..., b_0 \wedge b'_0 \rangle;$

Algorithm 2: Bit-blasting equality.

Input: $t_{[n]} = t'_{[n]}$

$\langle b_{n-1}, \ . \ . \ . \ , \ b_0 \rangle \leftarrow bbTerm\left(t_{[n]}\right);$

$\langle b'_{n-1}, \ . \ . \ . \ , \ b'_0 \rangle \leftarrow bbTerm\left(t'_{[n]}\right);$

return $\bigwedge_{i=0}^{n-1} b_i \Leftrightarrow b'_i;$

to atom a is bbAtom(a). For example, equality is bit-blasted as shown in Algorithm 2. The final bit-blasted formulais obtained by asserting that the bit-blasted definition of each $\boldsymbol{T}_{bv}$-atom is equivalent to the literal a occurring in the input formula:

$$bitblast(\psi) \equiv \psi \bigwedge_{atom\ a \in \psi} (a \Leftrightarrow bbAtom(\mathrm{a})). \tag{5}$$

Bit-blasted terms and formulas are cached. Therefore, bit-blasting $x[i : 0] \times y[i : 0]$ after having already bit-blasted $x \times$ y should not add any new clauses. Because we represent bit-blasted terms as tuples of bits and do not explicitly introduce a fresh variable for each bit until CNF conversion, our bit-blasting procedure does not introduce fresh variables for *bit-propagating* operations as described in [11].

V. And-Inverter-Graphs (AIG) Rewriting

AIGs are a restricted form of Boolean formulas that only use the ∧-nodes and ¬ Boolean operators. An AIG is a DAG in which each node has either 2 or 0 children. Our implementation uses by default the simplifications implemented by the abc command "*balance; rewrite*", but it can be configured to use user-provided abc scripts. For each 4-input cut of an AIG node, it checks a pre-computed table for equivalent ways to express the same Boolean function. After applying the abc simplifications, we rely on abc's own CNF conversion algorithm to convert the AIG to clauses and assert them to SAT_{bb}. Their algorithm has been optimized for converting AIGs to CNF.

A. Experimental Evaluatio

Experiments in this section were run on the StarExec [12] cluster infrastructure with a timeout of 900 seconds and a memory limit of 50GB. Experiments were ran on the queue all *.q* consisting of Intel(R) Xeon(R) CPU E5-2609 0 @ 2.40GHz machines with 268 GB of memory. All scatter plots are on a log-scale, the x and y-axis represent CPU time in seconds, unless otherwise specified. Fig. 2 compares the CVC4 eager solver performance by bit-blasting to AIG and applying AIG simplifications (cvcE+AIG) to bit-blasting to Node and applying Boolean rewrites (cvcE). The results are mixed, with the performance of cvcE+AIG being worse overall. Fig. 3 shows the impact of AIG rewriting on the brummayerbiere*

(*brummayerbiere*, *brummayerbiere2*, *brummayerbiere3*, and *brummayerbiere4*) SMT-LIB 2015-01-02 benchmark families. AIG rewriting seems to have a very negative impact on problems shows in Fig. 4. On these problems, AIG rewriting does not take a long time, but in some cases bit-blasting to AIG increases the number of literals in the bit-blasted formula by 20% compared to bit-blasting to Nodes. Fig. 5 shows the performance on the larger industrial benchmark family spear (1695 bench.). The plot shows how on easier problems the overhead of AIG rewriting hurts performance while for harder problems it is outweighed by the reduction in solving time.

VI. Refactoring Isomorphic Circuits

To take advantage of this structured practical applications we developed a preprocessing pass aimed at identifying isomorphic formulas and factoring them out by introducing fresh variables. Consider the following constraint:

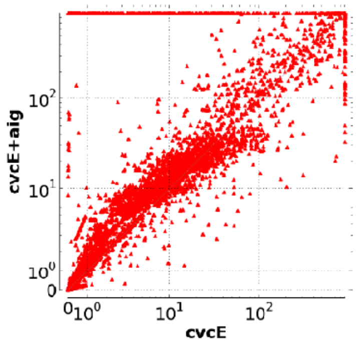


Fig. 2. cvcE vs cvcE +AIG.

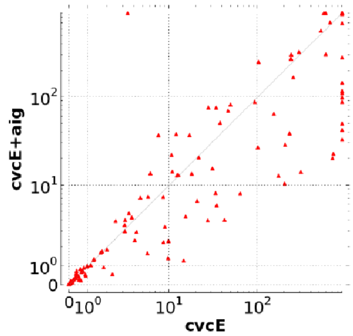


Fig. 3. cvcE vs cvcE+AIG on brummayerbiere*.

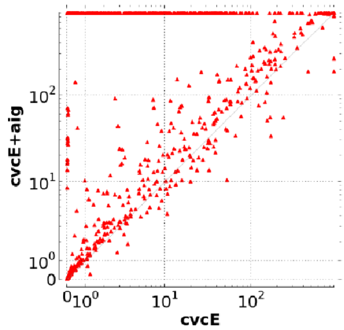


Fig. 4. cvcE vs cvcE+AIG on bruttomesso.

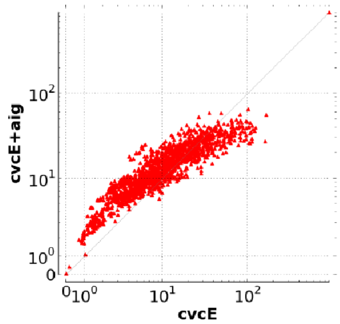


Fig. 5. cvcE vs cvcE+AIG on spear.

$$\begin{pmatrix} x_0 = 2\times y_0 + y_1 \vee \\ x_0 = 2\times y_1 + y_2 \vee \\ x_0 = 2\times y_2 + y_0 \end{pmatrix} \wedge \begin{pmatrix} x_1 = 3\times y_0 + 2\times x_0 + 5 \vee \\ x_1 = 3\times y_1 + 2\times x_0 + 5 \vee \\ x_1 = 3\times x_0 + 2\times y_2 + 5 \end{pmatrix} \tag{6}$$

The disjuncts in each conjunct implement the same function applied to different arguments:

$$\begin{pmatrix} x_0 = f(y_0, y_1) \vee \\ x_0 = f(y_1, y_2) \vee \\ x_0 = f(y_2, y_0) \vee \end{pmatrix} \wedge \begin{pmatrix} x_1 = g(y_0, x_0) \vee \\ x_1 = g(y_1, x_0) \vee \\ x_1 = g(x_0, y_2) \vee \end{pmatrix} \tag{7}$$

where $f(x, x') = 2 \times x + x'$ and $g(x, x') = 3 \times x + 2 \times x' + 5$. During bit-blasting each application of f and g will be bit-blasted to a different but isomorphic set of clauses. For example, in the $x_0 = 2 \times y_0 + y_1$ equality, x_0 and y_1 must have the same last bit, since the last bit of $2 \times y_0$ is 0. The similar clauses encoding this fact, for each application of f:

$$x_0[0:0] = 0 \vee y_1[0:0] = 1$$
$$x_0[0:0] = 0 \vee y_2[0:0] = 1$$
$$x_0[0:0] = 0 \vee y_0[0:0] = 1.$$

We can exploit this structure by introducing fresh variables for the arguments of f and g and pushing the disjunction to equalities over the arguments as follows:

$$\begin{array}{r} x_0 = f(s_0, s_1) \wedge \begin{pmatrix} s_0 = y_0 \wedge s_1 = y_1 \vee \\ s_0 = y_1 \wedge s_1 = y_2 \vee \\ s_0 = y_2 \wedge s_1 = y_0 \vee \end{pmatrix} \wedge \\ x_1 = g(s_0', s_1') \wedge \begin{pmatrix} s_0' = y_0 \wedge s_1' = x_0 \vee \\ s_0' = y_1 \wedge s_1' = x_0 \vee \\ s_0' = x_0 \wedge s_1' = y_2 \vee \end{pmatrix} \end{array} \tag{8}$$

Algorithm 3 shows the refactoring isomorphic circuits pass which first identifies top-level disjunctions φ and looks for a recurring pattern within the disjuncts. The *patternOf* procedure computes the pattern for each disjunct by replacing each variable v by a placeholder γ_n. For example:

$$\begin{array}{r} \mathit{patternOf}\ (3\times y + 2\times x + 5) = 3\times\gamma_0 + 2\times\gamma_1 + 5 \\ \mathit{patternOf}\ (x\times x + 5) = \gamma_0\times\gamma_1 + 5 \end{array} \tag{9}$$

We say a pattern p' is a generalization of pattern p $p(p' \succeq p)$ if there is a substitution that can be applied to p' that makes p syntactically equal to p'. For example:

$$\begin{aligned} \gamma_0 + 5 &\le \gamma_0 + \gamma_1 \\ \gamma_0 + 5 &<\neq \gamma_0 + \gamma_0. \end{aligned} \tag{10}$$

The procedures *getGeneralization* and *setGeneralization* ensure that the patterns used for refactoring are the most general ones to maximize the number of disjuncts refactored. The *getArgs(φ_i, p)* procedure returns the arguments p needs to be instantiated with to re-factor φ_i. Finally, *skolemizeArgs* introduces the disjunction of equalities and collapses disjuncts with the same pattern into one instantiation of the pattern.

A. Experimental Evaluation

The most significant performance gain was observed on the mcm family which encodes the multiple constant multiplication problem: synthesize an optimal sequence of operations that result in multiplying a given input by a fixed set of constants. Fig. 6 is a scatter plot comparing the size of the bit-blasted formula before applying circuit refactoring (cvcE) and after (cvcE +RIC). Each point is a benchmark and the x and y axis represent the thousands of literals in the bit-blasted formula [13]. Fig. 7 shows the performance impact of the reducing the bit-blasted formula size. The x-axis is number of problems solved, and the y axis is time taken to solve that number of problems.

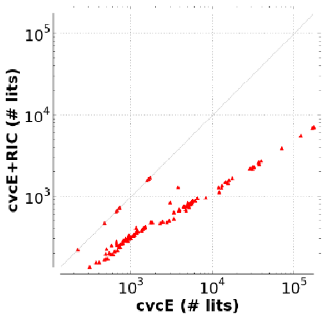


Fig. 6. Literals in bit-blasted formula cvcE vs cvcE+RIC on mcm.

<u>Algorithm 3</u>: Refactoring isomorphic circuits.

Input: ψ
$\psi' \leftarrow \psi$;
for $\varphi \in \psi$ **and** φ *top-level* **do**
  if $\varphi = \vee\, \varphi_i$ **then**
    P ← ∅;
    for $\varphi_i \in \varphi$ **do**
      p ← pattern Of (φ_i);
      P ← P ∪{p};
    for p, p' ∈ P **do**
      if p ≤ p' **then**
        setGeneralization(p, p');
    for $\varphi_i \in \varphi$ **do**
      p ← getGeneralization(patternOf (φ_i));
      args(p) ← args(p) ∪ getArgs(φ_i, p);
    φ' ← skolemizeArgs(φ, args);
    ψ' ← $\psi'[\varphi \leftarrow \varphi']$;
return ψ';

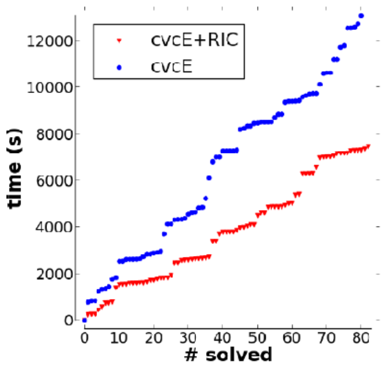


Fig. 7. Run-time cvcE vs cvcE+RIC on mcm.

VII. Conclusion

The performance of formal verification tools is tightly coupled with that of their backend theorem proving engines. Verifying many protection critical systems require reasoning about fixed-width bit-vectors. This paper investigated new techniques of solving bit-vector constraints, and evaluated. The eager bit-vector solver leverages AIG generalization their performance techniques and SAT reasoning to proficiently resolve bit-vector constraints that need low level of bit reasoning. The technique for factoring out isomorphic circuits discovers redundancy in the input formula and permits for reducing the size of the bit-blasted trouble.

References

[1] W. Ackermann, "Solvable cases of the decision problem", North-Holland, Amsterdam, San Jose, CA, U.S.A, 1962.

[2] A. Niemetz, M. Preiner, and Biere, "A. Turbo-charging lemmas on demand with don't care reasoning", Formal methods in computer-aided design (FMCAD) Inc., Switzerland, October 2014.

[3] V. Ganesh, and D. L. Dill, "A decision procedure for bit-vectors and arrays", "19th international conference on Computer aided verification", pp. 519-531, Springer-Berlin, 2007.

[4] R. Brayton and A. Mishchenko,"Abc: An academic industrial-strength verification tool",In computer aided verification, pp.24-40, 2010.

[5] T. Hansen, "A constraint solver and its application to machine code test generation", PhD thesis, University of Melbourne, 2012.

[6] L. D.Moura and N. Bjørner, "Z3: An efficient SMT solver", 14th international conference on tools and algorithms for the construction and analysis of systems, pp. 337-340, Springer-Berlin, 2008.

[7] L.D. Moura and N. Bjørner, "Relevancy propagation", Tech. Rep. MSR-TR-2007-140, Microsoft research, 2007.

[8] G. S. Tseitin, "On the complexity of derivation in propositional calculus", In automation of reasoning, pp. 466-483, Springer, 1983.

[9] N. Sörensson and N. Eén, "Minisat 2.1 and Minisat++ 1.0 - SAT race", Theory and applications of satisfiability testing, pp. 31-32, UK, 2009.

[10] N.Eén, and A. Biere, "Effective preprocessing in SAT through variable and clause elimination", 8th international conference on theory and applications of satisfiability testing, pp. 61-75, Springer, Berlin, 2005.

[11] A. Nadel, "Handling bit-propagating operations in bit-vector reasoning", In Work- shop on satisfiability modulo theories, pp. 243–277, 2013.

[12] A. Stump, G. Sutcliffe, and C. Tinelli, "Starexec: a cross-community infrastructure for logic solving", 7th International joint conference on automated reasoning, IJCAR, pp. 367-373, Vienna, Austria, 2014.

[13] N.P. Lopes, L. Aksoy, V. Manquinho and J. Monteiro, "Optimally solving the MCM problem using pseudo-boolean satisfiability", arXiv preprint arXiv:1011.2685, 2010.

Phishing Attack Detection Using Taxonomy Model

Chowdhury Sajadul Islam
Department of Computer Science & Engineering
Uttara University
Dhaka, Bangladesh
chow.cth@gmail.com

***Abstract*— The objective of this paper is to detect phishing attacks using proposed phishing threat taxonomy model. To detect current phishing attack we collected the information by applying diverse methodology and constructing an intellectual data set. We have done the experiment on the basis of data sets by our proposed equations which produced results on the predicted rate of a phishing threat parameters (i.e. method, origin, component and target) in respect of predicted number of threats. From the numerical results it is apparent that 16.07% phishers use spare phishing 14.85%, fake web 12.2%, link manipulation and pharming attack techniques which are the highest rate of attack in subtotal 53.9% and rest of the threats use 44.61%. The results suggest while the accuracy of taxonomy model is a critical factor, reducing the success rates of phishers require other considerations such as improving tools, update dataset instantly and enhancing users' knowledge of phishing. Finally, we got a phishing threat taxonomy model from the collected data set, proposed equation, numerical results and graphical representation which is the main goal or novelty of this paper. Given the prevalence of phishing-based web fraud, the findings have important implications for individual and enterprise security. This is to provide a clear vision of the challenges that should be worked onto ensure next generation phishing security for cyber world.**

Keywords—phishing attack; phishing origin; melware attack; phishing target; taxonomy model

I. Introduction

Over the years there have been conducted number of experiments, studies, researches have been conducted and many tools have been developed to combat the phishing threats but still the rate of phishing attacks are increasing day by day [1]. There are lots of phishing threats available in the phishing world which are: *fake website*, *link manipulation*, *pharming*, *in-session, spear, phone call*, e*vil twins*, *malware-based*, *search engine, internationalized domain names* (*IDN*) *and in session phishing*, etc[1] [2][4]. Although, all threats are based on websites but first four threats are mainly web based and other threats use email, phone call and software attack [3][5]. In FIFA world cup 2014,[1] phishers send e-mails, links, messages to victims telling them that they had won a World Cup game ticket and when they response, phishers snatch their credentials. In [6] conducted a comparative experiment of phishing and non-phishing URLs using following parameter such as IP addresses, URL, and non-phishing URLs are different in length and misuse of free hosting services by phishers. Phishing URLs was classified using CANTINA, a tool used for analyzing the content of the webpage [8] and finally developed eight discriminatory features and proposed CANTINA+ [8, 9]. The similarities of our research in [7] approach is to classify phishing emails by incorporating key structural features employing different algorithms, dataset for the classification process but in our work we classified the whole phishing threats to propose a taxonomy model. In [4] determines the phishing taxonomy model for web based mobile devices but they are following the way which is not enough. Here we did work on the root of phishing which is classified into four parameters: method, origin, component and target. Each parameters also classified into different entity and each entity has many attributes. If anyone wants to find the threat of any parameters, just put it in our proposed equation it will provide classified result, then get root of threat and permanently stop that class of attack.

II. Analysis and Experiment on Taxonomy Model

The theme of threat taxonomy model of phishing attack is obtained from biology. The functional descriptions of threat taxonomy model's parameters are given below:

Method- refers how the attack happens or the weakness that may have allowed the attack to occur.

Origin- root of phishing attacks, such as soft. and physical.

Component-can be tools, apps, language, API, and people.

Target- is to make financial gain, and fooling the user.

In this paper, we have experimented the methodology on various data sets including both phishing and legitimate sites [4,6]. Data sets have been retrieved from various databases such as ACM digital library[9][10], Springer link [7], IEEE xplore [11], Google and Yahoo. We also have holistically examined phishing attacks in various stakeholders and their countermeasures, and by surveying experts'[4][13] opinions about the current and future threats [1] to build data set. Then we proposed equations, to find the rate of four parameters. using statistic formula which are given below:

A. PhishingMethods Experiment

To find the rate of methods used in phishing attack, using the following equation:

$$\alpha_P = \left(\sum_{i=1}^{n} \frac{\alpha_{TiM}}{T_t \times \alpha_t} \times C \right) \times \left[\frac{C}{\sum_{s=1}^{s=20} \left\{ \sum_{i=1}^{m} \left(\sum_{i=1}^{n} \frac{C\alpha_{TiM}}{T_t \times \alpha_t} \right) \times X_\alpha \right\}} \right] \tag{1}$$

Let α_p denote the predicted rate (%) of method used in our twelve threats (T_t), α_{TiM} threats used in method, C is a constant (100), m is an arbitrary number and X_α is similar number of methods, and *s* no. of methods used in model. We have done the numerical experiment on the basis of our dataset and the numerical results are shown in the following table I.

B. Phishing Origin Experiment

We used the following equation to calculate the rate (%) of origin used in phishing attack:

$$\beta_P = \left(\sum_{i=1}^{n} \frac{\beta_{TiO}}{T_t \times \beta_t} \times C\right) \times \left[\frac{C}{\sum_{u=1}^{u=8} \left\{ \sum_{i=1}^{m} \left(\sum_{i=1}^{n} \frac{C\beta_{TiO}}{T_t \times \beta_t} \right) \times Y_\beta \right\}}\right] \tag{2}$$

In (2), denoted β_p as the predicted rate of origin used in total threats (T_t), number of threats used in origin β_{TiO}, total used origins β_t, constant C, highest number of threats used in origins n, an arbitrary number m (i.e. 1, 2, . . ,100) and similar number of origins used Y_β, and number of origins used u. The entries (2) are used for calculating phishing origin rate (%) which is shown in table II.

C. Experiment on Phishing Component

We can calculate the percentage of a component by using the following equation:

$$\gamma_P = \left(\sum_{i=1}^{n} \frac{\gamma_{TiC} \times C}{T_t \times \gamma_t}\right) \times \left\{\frac{C}{\sum_{v=1}^{v=13} \left(\sum_{i=1}^{m} \left(\sum_{i=1}^{n} \frac{C\gamma_{TiC}}{T_t \times \gamma_t} \right) \times Z_\gamma \right)}\right\} \tag{3}$$

TABLE I. PHISHING METHOD EXPERIMENTAL RESULTS

Method(α_t)	**Total Number of Threats (T_t)**												
	Fake Web.	**Link manip.**	**Spear Ph.**	**Pharming**	**In session**	**Fake Web. Valid.**	**Fake popup**	**Phone Ph.**	**Malware Ph.**	**Search Engine**	**Evil twins**	**Data Theft**	**Methods %)**
Email	3.04		2.80	2.79					2.95	2.70			14.28
DNS	3.02	2.88		2.64		2.90							11.44
URL	3.00	2.82	2.76										8.58
Malicious File		2.98		2.87							2.73		8.58
IDN	2.98	2.75											5.73
Popup					2.65		3.08						5.73
Instant messages	3.05									2.68			5.73
SEO								2.85		2.88			5.73
Social Engineering			2.85										2.85
Double barrel Att.			2.85										2.85
SMS			2.85										2.85
IP	2.85												2.85
M-i-M attack	2.85												2.85
Comp. host files				2.85									2.85
Trojan Horse				2.85									2.85
Session sniffing					2.85								2.85
Tabnabbing					2.85								2.85
Voice Phishing								2.85					2.85
Wireless AP											2.85		2.85
Compromised Serv.												2.85	2.85

TABLE II. PHISHING ORIGIN EXPERIMENTAL RESULTS

Priority Origin	**Total Number of Threats**												
	Fake Web.	**Link manip.**	**Spear Ph.**	**Pharming**	**In session**	**Fake Web. Valid.**	**Fake popup**	**Phone Ph.**	**Malware Ph.**	**Search Engine**	**Evil twins**	**Data Theft**	**Origin (%)**
Website	7.43	6.8				5.02	4.18			2.89			26.32
Human Factor	5.4	4.89						7.01		3.76			21.06
Browser		6.02	5.42		4.35								15.79
Software	4.27								6.26				10.53
OS				5.26									5.26
Malicious file			5.26										5.26
Fake calls								5.26					5.26
Insecure PC												5.26	5.26
Wireless router											5.26		5.26

TABLE III. PHISHING COMPONENT EXPERIMENTAL RESULTS

Threat	**Total Number of Threats**												
	Fake Web	**Link manip.**	**Spear Ph.**	**Pharming**	**In session**	**Fake Web. Valid.**	**Fake popup**	**Phone Ph.**	**Malware Ph.**	**Search Engine**	**Evil twins**	**Data Theft**	**Component(%)**
Scripting languages	6.13	4.41	5.6	4.42	5.23	2.2	2.15		4.21	1.9		2.1	38.35
DNS		4.46	2.87	4.13									11.46
Email		2.21	5.56										7.77
Popup					2.11		5.66						7.77
Auto validation						3.85							3.85
Host file poisoning				3.85									3.85
Web browsers		3.85											3.85
Malicious USBs			3.85										3.85
Wireless Access Point											3.85		3.85
Telephone & VoIP								3.85					3.85
RAT			3.85										3.85
Attachments with file			3.85										3.85
File download									3.85				3.85

where γ_P is predicted rate (%) of a component used in predicted number of total threats (T_t), other term γ_{TiC}, γ_t, n, m, Z_γ represent the number of component, total component, highest number of threat used in a component, same number of component and any arbitrary number, v=1 to no. of component. The table III shows the result of threat component used in phishing attack.

D. *Phishing Target Experiment*

To calculate the rate of target used in phishing attack by:

$$\delta_P = \left\{ \sum_{i=1}^{n} \frac{\delta_{TiT} \times C}{T_t \times \delta_t} \right\} \times \left\{ \frac{C}{\sum_{w=1}^{w=10} \left(\sum_{i=1}^{m} \left(\sum_{i=1}^{n} \frac{C\delta_{TiT}}{T_t \times \delta_t} \right) \times Q_\delta \right)} \right\} \quad (4)$$

In (4) δ_P stands for predicted rate(%) of a target used in predicted number of total threats (T_t), δ_{TiT} no. of threats used in a target, n is highest number of threat used in a target, m is arbitrary number, w=1 to no. of components and Q_δ similar numbers of targets.

III. Results and Discussion

The texts, tables and figures are organized in a way to facilitate easy interpretation of the findings. To implement and test our approach, we have used datasets collected from [5-6], the survey were first preprocessed to create a unified dataset where we found 3,180 web sites out of 3,973 are phishy. In each of these taxonomy parameter's plots, the x-axis is the number of parameter, considered as phishing attacks; the y-axis is the rating of parameter, calculated as the percentage of phishing attacks detected. From the result of calculation of phishing parameters phishing threats can be grouped into twelve categories. The graph in Fig. 1 shows that a single method is used in diverse number of threats and experimented on twenty methods where this graph represents email (14.28%) is the main method which is used to theft user credentials. First four methods (*eg. email, DNS, URL, mal. file)* having the risk level high and sub total is 42.88%. The next four methods (*eg. IDN, popup etc.*) having medium level of risk and accumulation is 22.92%. Last methods show lowest number of risk and overall accumulation is 34.20%.At this point the graph becomes a straight line. Fig. 3 demonstrates the attacks on thirteen phishing threats components of operation. A single component, scripting languages *(eg. js, php, htm etc.)* are the most frequently used having the high risk level which is 38.35%. The second most frequently used components which have the medium risk level 27% and then graph becomes suddenly step down. The low

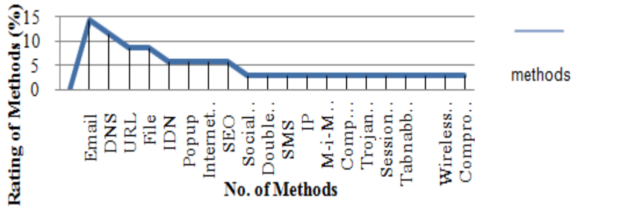


Fig. 1. No. of threats used in a method to do phishing attack

In Fig. 2 depicts heterogeneous frequencies of the origin occupancy in the website is 26.32% and human factor is 21.06% origins have high effect on the overall variation in the number of threat origin for any phishing attack. The threat of origin 15.79% is generated from browser and 10.53% origin is coming from malicious software. All other origins are 5.26% sequentially OS, malicious file, fake calls, insecure PC, wireless router etc.

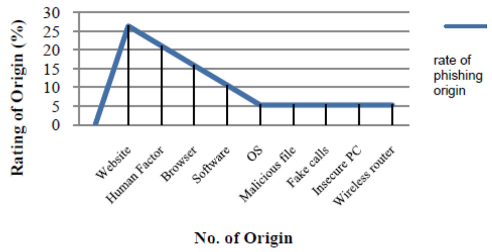


Fig. 2. No. of threats used in a origin to do phishing attack

risk level components (*eg. auto validation toolbar, host file poisoning, web browsers etc.)* individually have the same weight of risk and graph becomes a straight line and the accumulated risk levels are 34.65%. Fig. 4 shows the high prevalence of phishing targets are user credentials which is 34.78% and redirect to fake site is 17.39%. The mid-level targets of phishers are to capture host file 8.70% email attachment/link 8.69% and SMTP 8.69%. Rest of the phishers' target sequentially; human asset, exploiting browser

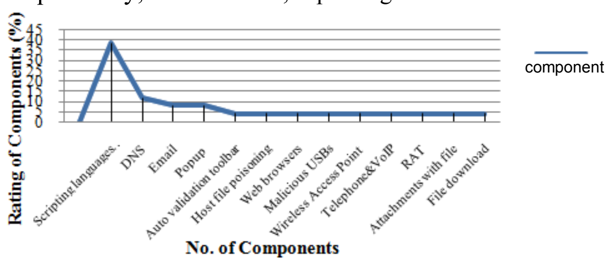


Fig. 3. Phishing components result using different phishing threat

TABLE IV. Phishing Target Experimental Results

Threat Target	Total Number of Threats												
	Fake Web.	Link manip.	Spear Ph.	Phar-ming	In session	Fake Web. Valid.	Fake popup	Phone Ph.	Malware Ph.	Search Engine	Evil twins	Data Theft	Target (%)
User credentials as PIN etc.	5.61		5.29		4.25	3.67		5.98		3.21	3.25	3.52	34.78
Page to redirect to fake site	5.21	4.75				4.88	2.55						17.39
Host file	4.57						4.13						8.70
Email Attachment/link			4.89	3.8									8.69
SMTP			3.71	4.98									8.69
Human Asset									4.35				4.35
Browser vulnerabilities				4.35									4.35
DNS table modification									4.35				4.35
Data theft				4.35									4.35
Take control of PC												4.35	4.35

TABLE V. PROPOSED PHISHING ATTACK DETECTION TAXONOMY MODEL

Threat	Method	Origin	Threat Component	Target
Fake website (14.85%)	URL request Email message Instant messages DNS record Fake DNS	Software Website Human factor	HTML java script PHP Cross-site scripting Malicious code	URL Fake messages Human factor Redirecting to fake site
Link manipulation (12.21%)	Human factor Sub domains Link anchoring IP instead of DNS Image file IDN	Long URL Sub-domain Anchor text Human Factor Browser Fake Website	Email Anchor text, Web browsers DNS	Redirecting to fake site Click on fake site Deceived user Website.
Fake popup (5.44%)	Popup windows Capture user credentials	Website	HTML, popup, java script, php	Redirect to fake site Human emotion.
Fake web. validn. (5.63%)	Registering similar DNS Man-in-the-middle attack.	Website	HTML, java script, auto validation toolbar	Redirecting to fake site User credentials
Pharming (12.29%)	E-mails, upload file Compromised host files DNS table modification, trojan hosts	OS DNS	Host file poisoning HTML and Java script DNS cache poisoning	SMTP Host file DNS server
In session (6.07%)	Phony popup, session hijacking, sniffing, tabnabbing	Browser	popup software, javaScript	User credentials
Spear phishing (16.07%)	Email attachment, hyperlink, double barrel, social engg. SMS phishing ("smishing")	Malicious browser	RAT, scripting, email link and attachments, malicious USBs, DNS	Email doc., fake link Confidential data
Phone ph. (6.95%	Voice phishing (Vishing)	Fake calls	Telephone, VoIP	Account numbers, PIN
Evil twins (4.49%)	Eavesdropping to AP, Bogus AP, snooping hotspot traffic.	RF, Router	Wireless access point (WAP)	Steal user credentials
Malware-Ph. (6.49%)	Email attachment, downloadable file	Mal.software	Email, Mal. code, file download	browser vulnerabilities, PC
Search E Ph. (5.05%)	SEO	Fake web site	Fake anti- malware program	User credentials
Data Theft (18.08%)	Compromised Server	Insecure PC	Malicious code	Data theft

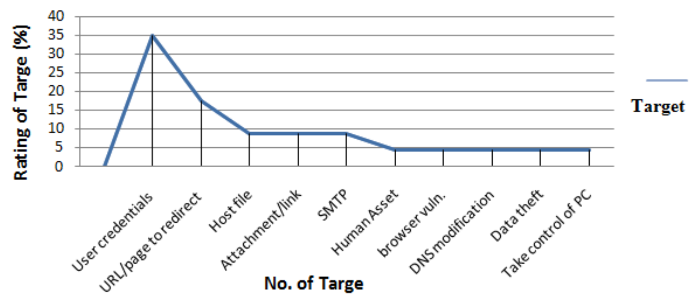


Fig. 4. No. of threats used in a target to do phishing attack

vulnerabilities, DNS server table modification, data theft and taking control of PC is 4.35%. The above experimental results and discussion facilitate us to build the innovative phishing taxonomy model which is shown in table V with the rate of threats (%). By using this model we can easily detect the behavior and characteristics of phishing threat. This model is capable of detecting the parameters of phishing attacks which ultimately suggests the solution to protect phishing attacks.

IV. CONCLUSION

This paper holistically examines on phishing taxonomy of various stakeholders and their countermeasures, and by surveying experts' opinions about the current and future threats and calculating the percentages of the parameters of phishing taxonomy. Our experimental result shows four key findings on four parameters, twelve threats and 624 ways of phishing attack detection. In graphical representation each parameter of taxonomy model have used for better understanding of phishing threat detection. In future by using this model, we are going to build a software which can efficiently detect the taxonomy of phishing attacks; till now we are experimenting on latest threats in every day to add it in our model based software. As a future work, we plan to use phishing taxonomy parameters auto update algorithms, to compare accuracy rates in real time. We also plan to do a thorough feature ranking of parameters and selection on the different data set to come up with the set of features that produces the best accuracy consistently by all the classifiers of taxonomy. In future we will extend not only an optimum anti-phishing elucidation, but also a permutation of anti-phishing and anti-hacking taxonomy approaches to defend diverse web attack.

REFERENCES

[1] G. Aaron, R. Rasmussen, A. Routt, " Global Phishing Survey: Trends and Domain Name Use in 2H2013 ". Last Anti-Phishing Working Group APWG survey report, 15 April 2014.

[2] M. Aburrous, M.A. Hossain, F.Thabatah, K.Dahal, "Intelligent phishing website detection system using fuzzy techniques", Damascus, 7-11 April 2008, pp.1 - 6, IEEE-2008.

[3] S. Vijayalekshmi,S.A. Rabara, "Fending financial transaction from phishing attack", Anna Univ., Chennai, pp.171 - 175, IEEE- Dec. 2010.

[4] E. O. Y. Boateng, P. M. Amanor, "Phishing, SMiShing & Vishing: An assessment of threats against mobile devices", Journal of Emerging Trends in Computing and Info. Sciences, Vol. 5, pp. 297-307, 2014.

[5] M. Bhati, R. Khan, "Prevention approach of phishing on different websites" International Journal of Engineering and Technology, vol. 2 No. 7, UK, July, 2012.

[6] L. Bing, S. Ruifeng, F. Xin, L. Xue, C. Wei-hao," Emergent challenges and IPDS for anti-phishing attack," IEEE conference on IT Convergence and Security (ICITCS), Beijing, pp. 1–4, 2014

[7] R. Basnet, S. Mukkamala, A. H. Sung "Detection of phishing attacks: A machine learning approach" Soft computing applications in industry, studies in fuzziness, springerlink, vol. 226, pp. 373-383, 2008.

[8] S. Gupta, P. Kumaraguru. "Emerging phishing trends and effectiveness of the anti-phishing landing page", Cybersecurity education and research centre (CERC), IIIT-Delhi, 1406.3682v1 [cs.CY], 14 Jun 2014.

[9] G. Xiang, J. Hong, C. P. Rose, and L. Cranor. "Cantina+: a feature-rich machine learning framework for detecting phishing web sites"; ACM Transactions on Information and System Security , pp.1-26, 2011.

[10] I. R. A. Hamid, J. H. Abawajy, "An approach for profiling phishing activities," ACM Journal, Vol. 45, pp. 27-41, April, 2014.

[11] Z. Xu, H. Wang, S.Jajodia, "An emergency line of defense against phishing attacks," IEEE 33rd International Symposium on Reliable Distributed Systems (SRDS), pp. 11 - 20, Oct. 2014.

[12] L. Pollacia, Y. Z.Ding, S. Yang, "Why Phishing Works: Project for an Information Security Capstone Course 2014", Information Systems Educators Conference, Maryland USA, Vol. 31, 2014.

An Empirical Analysis of Attribute Skewness over Class Imbalance on Probabilistic Neural Network and Naïve Bayes Classifier

MD. Nazmul Shahadat, Biprodip Pal
Department of Computer Science and Engineering
Rajshahi University of Engineering and Technology, Bangladesh
nazmul.ruet@gmail.com, biprodip.cse@gmail.com

***Abstract*—Many real world data are subjected to skewness or imbalance. Often class distribution is imbalanced, while several attribute or feature skewness is also frequent. Skewness affects the classification of the dataset samples. While class skewness biases the classification towards majority classes, skewed features may also bias the classification as they are significant for few classes. The purpose of this paper is to find out the impact of skewed feature variation in the training dataset for the Naïve Bayesian Classifier (NBC) and Probabilistic Neural Network (PNN). The experiment was carried out on six KEEL dataset which are skewed in terms of class distribution. This work looked for skewed features in those dataset and analysed the classification performance with and without the skewed features. The confusion matrix analysis illustrates that NBC is performing well compared to PNN.**

Keywords*—*Data Mining, Classifier performance, Skewed Data analysis, Naïve Bayesian Classifier, Probabilistic Neural Network.

I. INTRODUCTION

Gaining information from the data is an important part of Data mining and Machine learning field. Basically, they focus on discovery of unknown properties in the data. But here Supervised Learning is used. Supervised learning is a method of learning from examples where the learner is provided with two sets of data, a training set and a test set. In machine learning and statistics, Classification is the process of identify-ing on the basis a labelled training set to which of a set of categories a new observation belongs [1]. The classification process is often affected with several parameters. Biased or Skewed data is one of them. The classifier performances are subject to the data presented to it during training session. Often the class distributions in the Data are skewed due to data collection faults. Besides several data are intrinsically biased like fraud detection data. But the attributes or features of the data are often skewed too. They also hamper the learning process.

Previous attempts to analysis the skewness effect are mainly on class distribution skewness. Lei et. al. investigated effect of feature selection metrics on frequently used classifiers in the context of highly skewed data[1]. Several techniques are discussed and analysed to handle class skewness in review papers [2]. Maria et. al. Analysed the relationship between cost-sensitive learning and class distributions. They also experimented the limitations of accuracy and error rate to measure the performance of classifiers in this regard [3]. Phuaet. al. Proposed a new method to deal with skewed data in their paper[4]. Other works include importance weighting to optimize average recall from skewed datasets. Besides, Oversampling techniques are famous to deal with skewness of classes [5].

This paper analyses Naïve-Bayesian Classifier (NBC) and Probabilistic Neural Network (PNN). These classifiers are an example of supervised learning. Naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem. The induction of these classifiers is extremely fast and can give better predictive accuracy than other well-known methods [6]. A probabilistic neural network (PNN) is a feed-forward neural network, which is a supervised neural network and a statistical algorithm called Kernel Fisher discriminant analysis that is widely used in the area of pattern recognition. It is predominantly a classifier to map any input pattern to a number of classifications and can be forced into a more general function approximation.

The purpose of this work is to check the impact of attribute skewness on class imbalance. The skewed attributes are looked and checked to see if skewed attributes has any effect on the classification performance that are significantly compared to the impact of class skewness.

II. STATISTICAL PARAMETER

There are several well-known processes measuring the central tendency. The mean is the most popular from them. To find the arithmetic mean of a set of n numbers, add the numbers in the set and divide the sum by n,

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{1}$$

On the other hand, the Standard deviation is the square root of the variance which is used to measure spread means a measure of the dispersion of a set of data from its mean. It's symbol is σ which is denoted by-

$$\sigma = \sqrt{\frac{\sum(X - \mu)^2}{N}} \tag{2}$$

Where, $\sigma\ is$ the Standard Deviation, $X\ is$ each value in the population, $\mu\ is$ the mean of the values, and $N\ is$ the number on values in population. The Gaussian distribution is sometimes informally called the bell curve. A normal Gaussian distribution is

$$f(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3)$$

The parameter μ in this definition is the mean or expectation of the distribution. The parameter σ is its standard deviation; its variance is therefore σ^2 [7]. Skewness strongly used in this paper is a measure of the degree of asymmetry of a distribution. For a sample of *n* values, a natural method of moment's estimator of the population skewness is

$$Skewness = \frac{\mathrm{m}_3}{\mathrm{s}^3} = \frac{1/n\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{3/2}} \quad (4)$$

Where $\bar{x}$ is the mean, s is the sample standard deviation, and the numerator m_3 is the sample third central moment.

III. Naïve Bayes Classifier

For a sample $x = \{x_1, x_2, \dots\dots\dots, x_n\}$, where the component set is consisted of $\boldsymbol{n}$ attributes, x is considered "evidence" and y is hypothesis. The value of $P(y|x)$ can be expressed in terms of probabilities $P(y), P(x|y)$, and $P(x)$ as(According to Bayes' theorem),

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)} \quad (5)$$

Where p(y) is the priori probability, p(x) is the evidence, p(x|y) is the likelihood and p(y|x) is the posteriori probability [7]. It may be expressed when there are multiple output class labels such as that the data x belongs to specific output class labels $y \epsilon \{1,\dots,C\}$. For an instance, X and m classes, $C_1, C_2, \dots, C_m$, the naïve Bayesian classifier predicts that tuple X belongs to the class C_i if and only if

$$P(C_i|X) > P\left(C_j|X\right) \quad (6)$$

for $1 \le j \le m, j \ne i$. Thus $P(C_i|X)$is maximized. The class C_i for which $P(C_i|X)$ is maximized is called the maximum posteriori hypothesis. By Bayes' theorem,

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \quad (7)$$

As $P(X)$ is constant for all output classes, only $P(X|C_i)P(C_i)$ needs to be maximized as to maximize the posterior probability. For N input attributes $X = (x_1, x_2, \dots, x_n)$the class conditional independence is [8][9]-

$$P(X) = \prod_{i=1}^{n} P(X_i) \quad (8)$$

Which can be considered independent both conditionally and unconditionally of one another for the given classC_iSo,

$$P(X|C_i) = \prod_{k=1}^{n} P(X_k|C_i) \quad (9)$$

The probabilities $P(X_1|C_i), P(X_2|C_i), \dots, P(X_n|C_i)$ can be easily estimated from the training tuples. As attributes are continuous-valued, then the calculation is pretty straight forward. A continuous-valued attribute is typically assumed to have a Gaussian distribution with a mean μ and standard deviation σ defined by-

$$g(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (10)$$

In order to predict the class label of $X, P(X|C_i)P(C_i)$ is evaluated for each class C_i.The classifier predicts that the class label of tuple X is the class C_i, if and only if

$$P(X|C_i)P(C_i) > P\left(X|C_j\right)P\left(C_j\right) \quad (11)$$

For $1 \le j \le m, j \ne i$. In this process the Expectation Maximization(EM) algorithm is used to evaluate the equation(10) that is, whose posterior probability is greater [9].

IV. Probabilistic Neural Network

The probabilistic neural network (PNN) paradigm basically uses the Parzen-Cacoulos estimator to obtain the corresponding Probability Density Function's (PDF). PNN uses a supervised training set to develop probability density functions within a pattern layer [10].

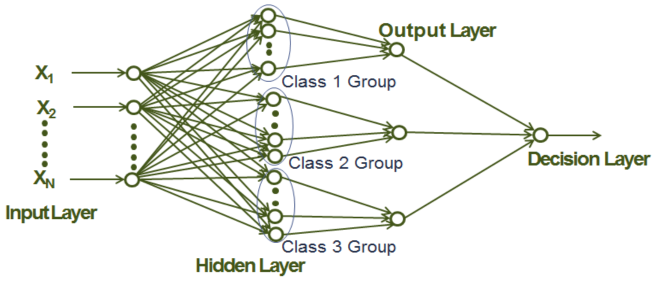


Fig. 1. Basic design of a Probabilistic Neural Network

The architecture of a PNN is limited to four layers: the input layer, the hidden layer, the summation layer or the output layer and the decision layer. The first or lowest layer is called the input layer where external information enters the network. The input layer does not perform any computation and simply distributes the input to the neurons in the pattern layer. In hidden layer, sort the input into the class K(Number of output nodes in output layer (k = 1, 2 or 3 here)). Let there be P exemplar feature vectors $\{x^{(p)}: p = 1, \dots\dots\dots, P\}$ labeled as Class 1 and let there be Q exemplar feature vectors $\{y^{(r)}: r = 1, \dots\dots\dots, R\}$ labeled as Class 2. In the hidden layer there are P nodes in the group for Class 1 and R nodes design the group for Class 2. The equations for each Gaussian centred on the respective Class 1 and Class 2 points $x^{(p)}$ and $y^{(q)}$ (feature vectors) are, for any input vector x

$$g_1(x) = \frac{1}{\left(\sigma\sqrt{2\pi}\right)^N} e^{-\frac{\left\|x-x^{(p)}\right\|^2}{2\sigma^2}} \quad (12)$$

$$g_2(y) = \frac{1}{\left(\sigma\sqrt{2\pi}\right)^N} e^{-\frac{\left\|y-y^{(q)}\right\|^2}{2\sigma^2}} \quad (13)$$

The σ values can be taken to be one-half the average distance between the feature vectors in the same group or at each exemplar it can be one-half the distance from the exemplar to its nearest other exemplar vector. The kth output node sums the values received from the hidden nodes in the k[th] group, called mixed Gaussians or Parzen windows. The sums are defined by

$$f_1(x) = \frac{1}{(\sigma\sqrt{2\pi})^N}\left(\frac{1}{P}\right)\sum_{(p=1,P)} e^{-\frac{\|x-x^{(p)}\|^2}{2\sigma^2}} \quad (14)$$

$$f_2(y) = \frac{1}{(\sigma\sqrt{2\pi})^N}\left(\frac{1}{Q}\right)\sum_{(q=1,Q)} e^{-\frac{\|y-y^{(q)}\|^2}{2\sigma^2}} \quad (15)$$

Where x is any input feature vector, σ_1 and σ_2 are the spread parameters for Gaussians in Classes 1 and 2, respectively , N is the dimension of the input vectors, P is the number of centre vectors in Class 1 and R is the number of centres in Class 2, $x^{(p)}$ and $y^{(r)}$are centres in the respective Classes 1 and 2, and $\|x - x^{(p)}\|$ is the Euclidean distance between x and $x^{(p)}$. Any input vector x is put through both sum functions $f_1(x)$ and $f_2(x)$and the maximum value of $f_1(x)$ and $f_2(x)$ decides the class. For K >2 classes the process is analogous.

V. EXPERIMENTAL RESULTS

For this experiment, It was compared these two algorithms on several datasets in different ways. From these results discussed six of them.

A. Used Datasets

The Thyroid disease data set (Dataset I) was used from UCI machine learning repository [11]. It consists of 3772 learning and 3428 testing examples each with 22 attributes. In the Page Blocks dataset (Dataset II), all attributes are numeric [12]. The Pima Indians Diabetes Data Set(Dataset III) was used from the same repository as above [13]. It consists of 614 learning and 154 testing examples.The ecoli2 Data set(Dataset IV), glass-0-1-6_vs_2(Dataset V) and wisconsin-5Data set(Dataset VI) were used which can be found in the Keel Repository [14]. The ecoli2 dataset contains 336 instances, the glass dataset contains 192 instances and the Wisconsin dataset contains 683 instances and these datasets contain different number of attributes.

B. Implementation

At first the skewness of all attributes for each of the mentioned datasets were calculated.

TABLE I.ATTRIBUTES SKEWNESS VALUE OF DIFFERENT DATA SETS

Dataset	Skewness values of each Attribute
Thyroid disease data set	-0.1987 0.8576 2.2923 8.6031 9.2063 4.7836 8.2585 8.1783 7.9504 3.6318 3.6508 13.9851 10.5519 6.06156 1.4003 4.1637 14.6198 1.8946 1.2890 1.2602 2.2820 -3.9399
Page Blocks dataset	22.0413 2.0867 20.5267 6.8812 1.6147 -0.8512 60.8137 12.9658 10.5035 5.4139 2.6279
Pima Diabetes Data Set	0.9280 0.0605 -1.8102 0.1972 2.1967 -0.4616 1.8891 1.0666 0.6357
ecoli2 Data set	-0.1065 0.7738 5.9424 16.2789 0.0014 0.3063 0.4985 1.9280
glass-0-1-6_vs_2	2.0508 -0.0302 -1.4483 0.7605 -0.9693 2.0257 2.8580 3.1016 3.3903 2.9769
wisconsin-5 Data set	0.6100 1.1951 1.1337 1.5482 1.6976 0.9785 1.1082 1.4538 3.6883 0.6298

According to the work flow that mentioned on the above description, the data set “Thyroid disease data set” is applied on PNN and NBC with varying the training data set examples and the output accuracy are given below,

TABLE II. OUTPUT ACCURACY OF FULL AND REDUCED ATTRIBUTES FOR *DATASET I*

#	Training Data Size	Testing Data Size	Accuracy of Full Attributes (%)		Reduced Attributes (%)	
			NBC	***PNN***	***NBC***	***PNN***
1	3772 × 22	3428 × 22	95.21	96.96	99.41	97.25
2	3000 × 22	3428 × 22	95.42	97.19	99.38	97.22
3	2000 × 22	3428 × 22	95.85	98.24	99.29	97.25
4	1000 × 22	3428 × 22	96.47	98.54	99.62	99.54

The Confusion matrix with output accuracy of both algorithm are shown below,

TABLE III. CONFUSION MATRIX OF NBC & PNN FOR *DATASET I*

Actual	Predicted Data							
	NBC				***PNN***			
	1	***2***	***3***	***Accuracy (%)***	***1***	***2***	***3***	***Accuracy (%)***
1	3274	0	0	99.96%	3275	0	0	100%
2	0	3	4	42.85%	0	5	2	71.42%
3	1	14	131	89.72%	3	93	54	36.98%

According to the workflow that mentioned on the above description, the data set, “Page Blocks” are applied on PNN and NBC varying the training data set examples and the output accuracy are given in table IV,

TABLE IV.OUTPUT ACCURACY OF FULL AND REDUCED ATTRIBUTES FOR *DATASET II*

#	Training Data Size	Testing Data Size	Accuracy of Full Attributes (%)		Reduced Attributes (%)	
			NBC	***PNN***	***NBC***	***PNN***
1	4377 × 11	1095 × 11	95.06	92.51	98.44	95.43
2	3400 × 11	1095 × 11	96.80	92.78	98.08	95.43
3	2400 × 11	1095 × 11	97.80	92.78	97.99	95.34
4	1400 × 11	1095 × 11	96.44	92.60	97.53	95.15

The Confusion matrix with output accuracy of both algorithms are given in table V,

TABLE V. CONFUSION MATRIX OF NBC & PNN FOR *DATASET II*

Actual	Predicted Data					
	NBC			***PNN***		
	1	***2***	***Accuracy (%)***	***1***	***2***	***Accuracy (%)***
1	982	1	99.89%	983	0	100%
2	41	71	63.69%	83	29	25.89%

And according to the workflow that mentioned on the above description, the data set, “Pima Diabetes Data Set”, “ecoli2 Data set”, “glass-0-1-6_vs_2” and “wisconsin-5 Data set” are applied on PNN and NBC with varying the training data set examples and the output accuracy are,

TABLE VI. OUTPUT ACCURACY OF FULL AND REDUCED ATTRIBUTES

Data Sets	Training Data Size	Testing Data Size	Accuracy of Full Attributes (%)		Reduced Attributes (%)	
			NBC	***PNN***	***NBC***	***PNN***
Pima	614 × 11	154 × 11	89.61	81.81	91.55	88.31
Ecoli	268 × 8	68 × 8	98.52	94.11	98.52	94.11
glass	153 × 10	39 × 10	94.87	92.3	97.43	92.3
wisconsin	268 × 8	268 × 8	89.61	81.81	91.55	88.31

The Confusion matrices with output accuracy of both algorithms for these Datasets are shown below,

TABLE VII. OUTPUT ACCURACY OF NBC & PNN FOR *DATASET III* IN CONFUSION MATRIX

Actual	Predicted Data					
	NBC			PNN		
	1	*2*	*Accuracy (%)*	*1*	*2*	*Accuracy (%)*
1	99	1	99%	100	0	100%
2	12	42	77.77%	16	36	66.67%

TABLE VIII. OUTPUTACCURACY OF NBC & PNN FOR *DATASET IV* IN CONFUSION MATRIX

Actual	Predicted Data					
	NBC			PNN		
	1	*2*	*Accuracy (%)*	*1*	*2*	*Accuracy (%)*
1	56	1	98.24%	983	0	100%
2	0	11	100%	83	29	25.89%

TABLE IX. OUTPUTACCURACY OF NBC & PNN FOR *DATASET V* IN CONFUSION MATRIX

Actual	Predicted Data					
	NBC			PNN		
	1	*2*	*Accuracy (%)*	*1*	*2*	*Accuracy (%)*
1	34	1	97.14%	35	0	100%
2	0	4	100%	3	1	25%

TABLE X. OUTPUT ACCURACY OF NBC & PNN FOR *DATASET VI* IN CONFUSION MATRIX

Actual	Predicted Data					
	NBC			PNN		
	1	*2*	*Accuracy (%)*	*1*	*2*	*Accuracy (%)*
1	89	0	100%	89	0	100%
2	2	46	95.83%	47	1	2.08%

C. Result Analysis

Attributes extraction involves reducing the amount of resources required to describe a large set of data. In TABLE I where the skewed attributes value for each datasets are given and there may have the skewness in each datasets. Not only the attribute, this skewness may also appear in each datasets that can understand from confusion matrices. These skewness are really a special case of comparing distributions and a skew in class distributions can pose difficulty for both classifier training and testing. From the above datasets, the numbers of training and testing examples for Class-1, Class-2 and/or Class-3 have huge difference that means the skewed data are appeared in both training and testing datasets. For this, these datasets are classified into two categories- a majority class and a minority class where each category are defined by each class samples. By calculating the output accuracy according to NBC and PNN for all datasets it is found the TABLE II, IV, VI and also found some confusion matrices that are displayed in different tables. From the TABLE II, V and VIII it can be said, the output accuracy of NBC is better than the output accuracy of PNN in all cases. But in all confusion matrices, when it is compared individual classes accuracy then noticed that PNN perform well for majority classes and NBC perform well for minority classes. For each minority classes, the output accuracy for PNN is comparatively low but for majority classes is comparatively good enough. For this reason the total output accuracy of PNN is lower than NBC. There also have an important point, when the skewed attributes are removed from the datasets then overall output accuracy is increased in most datasets for NBC and PNN. Here also NBC has better output accuracy than PNN and it is shown that this assumption unchanged for different field of datasets. According to this result analysis, the classification accuracy rate of NBC is better than of PNN.

VI. CONCLUSION

This paper presents analytical results for choosing NBC and PNN classifier when data are imbalance. When it is hard to appropriately handle class imbalance in data then this analysis is helpful to select appropriate classifier from NBC and PNN if skewed attributes are handled properly.

REFERENCES

[1] Tang, Lei, and Huan Liu. "Bias analysis in text classification for highly skewed data." Data Mining, Fifth IEEE International Conference on. IEEE, 2005.

[2] He, Haibo, and Edwardo Garcia. "Learning from imbalanced data." Knowledge and Data Engineering, IEEE Transactions on 21.9 (2009): 1263-1284.

[3] Monard, Maria Carolina, and Gustavo EAPA Batista. "Learmng with Skewed Class Distrihutions." Advances in Logic, Artificial Intelligence, and Robotics: LAPTEC 2002 85 (2002): 173.

[4] Phua, Clifton, DammindaAlahakoon, and Vincent Lee. "Minority report in fraud detection: classification of skewed data." ACM SIGKDD Explorations Newsletter 6.1 (2004): 50-59.

[5] Mohanty, Banyabaishali, and Subhendu Kumar Pani. "Classification of Skewed Data: A Comparative Analysis of the Performance of Select Classifiers." IUP Journal of Information Technology 10.1 (2014): 18.

[6] Cheng, Jie, and Russell Greiner. "Comparing Bayesian network classifiers. "Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., 1999.

[7] Ribeiro, Maria Isabel. "Gaussian probability density functions: Properties and error characterization." Institute for Systems and Robotics, Lisboa, Portugal(2004).

[8] Han, Jiawei, and MichelineKamber. Data Mining, Southeast Asia Edition: Concepts and Techniques. Morgan kaufmann, 2006.

[9] Tsuruoka, Yoshimasa, and Jun'ichiTsujii. "Training a naive bayes classifier via the em algorithm with a class distribution constraint." Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4. Association for Computational Linguistics, 2003.

[10] Demuth, Howard, Mark Beale, and Martin Hagan. "Neural network toolbox™ 6." User's guide (2008).

[11] Ross Quinlan.(1987) On the use of Thyroid Disease Data Set [Online] . Available http://archive.ics.uci.edu/ml/datasets/thyroid+disease File: ann-train.data&ann-test.data.

[12] DonatoMalerba.(1995) On the use of Blocks Classification [Online] . Available http://repository.seasr.org/Datasets/UCI/arff/page-blocks.arff . File: page-blocks.arff.

[13] Vincent Sigillito.(1990, Last update 2011) On the use of Pima Indians Diabetes Data Set [Online] . Available https://archive.ics. uci.edu/ml/ machine-learning-databases/pima-indians-diabetes/pima-indiansdiabetes.data. File: pima-indians-diabetes.data.

[14] J. Alcalá-Fdez, A. Fernandez, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera. KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework. Journal of Multiple-Valued Logic and Soft Computing 17:2-3 (2011) 255-287

Cooperative Game Theory Based Load Balancing in Long Term Evolution Network

Subarno Saha
Electrical and Electronic
Engineering Department
Brac University
Dhaka, Bangladesh
Email: subarno.saha@gmail.com

Rajkin Hossain
Computer Science and
Engineering Department
Brac University
Dhaka, Bangladesh
Email: rajkin69@gmail.com

Dr. Md. Muhidul Islam Khan
Computer Science and
Engineering Department
Brac University
Dhaka, Bangladesh
Email: muhidul.khan@bracu.ac.bd

***Abstract*—Long term evolution (LTE) network, incompatible with 2G and 3G networks is the most promising technology for wireless communication with higher speed and capacity. Self-organized load balancing is an important research issue for the wireless networks. Game theory provides an efficient way to provide self-organizing properties in a distributed environment like LTE networks. Load balancing means to assign users from highly loaded cells to neighbor lower loaded cells. The amount of load needs to be offloaded or accepted by a particular cell is not really specified and currently totally vendor specified. In our proposed cooperative game theoretic approach, each cell is considered as a player where they trade the load by forming a coalition by satisfying the overall performance of the network. Simulation results show that our proposed method provides better performance in terms of satisfied users and adjusted load values.**

Keywords- cooperative game theory; long term evolution network; eNodeBs; load balancing; physical resource blocks.

I. INTRODUCTION

Long term evolution network (LTE) [1], [2] refers to the leading edge technology for mobile data communication with higher speed and capacity. With the recent advancement of 4G LTE networks, there has been increasing interest on the performance and power characteristics [3], [4].

Load balancing is the most efficient way to improve the performance of a network. Unequal load scenario makes some cells overloaded while, other cells have much less users. In this scenario, the performance deteriorates a lot due to the lack of proper resource distribution. Typical cellular networks employ load balancing techniques by periodically exchanging information between neighboring cells which is totally vendor specific [3].

Recent research trend is to employ self-organizing algorithms to provide the proper load balancing. Most of the existing works consider the player as rational which tries selfishly to maximize its own reward. This non cooperative way is not suitable for the network for its overall performance. Our goal is to provide a cooperative and self-organized way for load balancing in a LTE network.

Game theory provides the most efficient way to balance the load autonomously [5]. We propose a cooperative game theory based algorithm for load balancing in a LTE network to provide better performance in terms of satisfied users and properly distributed loads.

The rest of this paper is organized as follows. Section II discusses related work. Section III explains our system model. In Section IV we present our proposed method. Section V discusses simulation results for a LTE network. Section VI concludes this paper with a brief summary.

II. RELATED WORKS

In distributed networks, the goal is to build a collaborative environment for facilitating the effective usage of resources [6]. Load balancing is the key factor to maintain the resources of the network in a way that the performance is maximized. Most of the existing load balancing techniques does not provide self-organizing properties.

In [3] and [7], game theoretic approaches are proposed for the load balancing in a homogeneous network. In these methods, cells are considered as a rational player who deals with the load using game theoretic approach. But this non cooperative method makes the cell selfish to maximize its own reward. So, it is often possible that the performance of a particular cell may be improved but the overall performance of the network can be deteriorated.

Cooperative method helps to provide cooperation among cells to maximize the overall performance of the network. We propose a cooperative game theoretic approach, Shapely value [5] by forming a coalition among cells of a LTE network.

This paper proposes a cooperative game theoretic approach for load balancing which helps to balance the load by calculating the Shapely value among the neighboring cells, forming a coalition among them. Simulation results show that the performance is improved in terms of satisfied users by improving the load balancing across the network.

III. SYSTEM MODEL

The eNodeB allocates the physical resource blocks (PRBs) to all users by using a scheduling function.This scheduling function will be based on signal-to-noise ratio (SINR) of every user and resources that are used in user equipments (UEs).

Let the average data rate per user, u is denoted by D_u and the data rate per PRB is given by $R(SINR)$ of user, u. Best load balancing can be achieved by Shannon's theorem [5]

$$R(SINR) = log_2(1 + SINR_u) \tag{1}$$

Load generated (amount of PRBs required by user, u) by each user for the required data rate D_u is

$$N_u = \frac{D_u}{R(SINR)BW} \tag{2}$$

Here, BW is the transmission rate of one PRBs per frame. The load of a cell, c can be expressed as

$$\rho_c = Min[\frac{1}{N_{tot}} \sum_{u|X(u)} Nu, 1] \tag{3}$$

Here, N_{tot} is the total number of PRBs of a cell. $X(u)$ is the connection function which gives the serving cell, c for user, u.

Due to the resource limitations, number of unsatisfied user increases. When $\rho_c \leq 1$ users will be satisfied, otherwise unsatisfied. If, $\rho_c = N$, it satisfies only $\frac{1}{N}$ users [3]. Then, the number of satisfied user of a network can be expressed as

$$Z = \sum_{\forall c} max(0, u_c(1 - \frac{1}{\rho_c})) \tag{4}$$

where number of users in a cell, c is denoted as U_c.

IV. Proposed method for load balancing

We consider a cellular coverage area, c consists of n cells where, $c = \{c_1, c_2, c_3, c_4, \ldots c_n\}$ which are participating in the coalition. For a overloaded cell, $\rho_c > 1$. The overloaded cell gets all possible information from its connected UEs about the reference signal received power (RSRP) levels not only for the serving cells but also for the neighboring cells having a strong signal level [8].

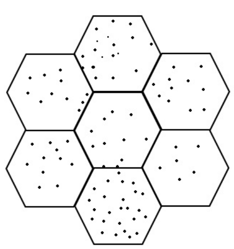


Fig. 1. Network model.

During a cell selection, cell reselection or handover an UE usually selects the cell based on RSRP which is the average power measured on UEs for cell specific reference signal [3].The link imbalance value, which is defined as the difference in the RSRP levels of overloaded cell, o and the neighboring under loaded cell, i is

$$\Delta RSRP_{u,i} = RSRP_{u,o} - RSRP_{u,i} \tag{5}$$

A good or optimum RSRP value provides good SINR value which determines a good signal strength of a cell to the UEs.The coalition that involves the under loaded cells, maximizes the satisfaction based on their marginal contribution in a fairly distribution manner. The utility function of the overloaded cell, o and the under loaded cell, i is denoted as

$$Utility_o = U_o - \Delta X_o \qquad \text{if } 0 \leq \rho \leq 1 \tag{6}$$

$$Utility_o = \frac{U_o - \Delta X_o}{\rho} \qquad \text{otherwise.} \tag{7}$$

$$Utility_i = U_i + \Delta X_i \qquad \text{if } 0 \leq \rho \leq 1 \tag{8}$$

$$Utility_i = \frac{U_i - \Delta X_i}{\rho} \qquad \text{otherwise.} \tag{9}$$

Here, ΔX_i is the load value that represents how much load will be added to the underloaded cell and ΔX_o is the load that will be given out from the overloaded cell.

For load balancing first we calculate each cell's obtained value by using

$$\Delta X_{c,i} = L_{c,i} - S_{c,i}; \forall c, i \tag{10}$$

Where, $\Delta X_{c,i}$ is the obtained value of each ith cell, $L_{c,i}$ is the load value before Shapley Value calculation. $S_{c,i}$ is the Shapley Value of each cell. We notice in simulation that some cell's obtained values are positive and some cell's obtained values are negative. Positive is for the overloaded cell and negative value is for the under loaded cell.

A. Definition of neighbor

When the signs of the obtained values, ΔX in Equation 10 are opposite for two cells and the cells meet the threshold value of SINR will be considered as neighbor cells.

B. Modeling

Now we design a Bipartite graph for those cells. Positive obtained value carrier cells are on the left partition and negative obtained value carrier cells on the right partition.

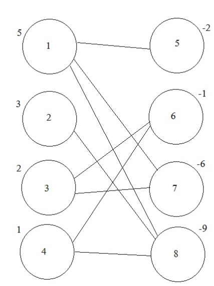


Fig. 2. Bipartite graph for cells.

In Fig. 2 a circle defines a cell and edges define neighboring cells. The value inside the circle defines the id of cells and the values outside the circles indicate the obtained value of the corresponding cells.

Initially none of the obtained values are zero, because we are only considering the cells with the positive and negative obtained values. Any cell's obtained value is zero means particular cell is satisfied in a cooperative manner.

After modeling we use our own greedy algorithm for maximizing the number of satisfied cells. If any cell fails to satisfy itself then we will try to make that cell's obtained value close to zero by losing or gaining some loads. But unfortunately, few cell's obtained value remain unchanged forever because of insufficient number of neighbor of that cell. Our first priority is to maximize the number of satisfied cell and the second priority is to maximize the number of cell's obtained value more close to zero.

C. *Proposed algorithm*

Considering the Algorithm 1 and the bipartite graph in Fig. 2, all the cells in left side are arranged in ascending order of their obtained values from bottom to top. We try to make those cells satisfied or more close to satisfied where obtained values are more close to zero. We take that particular cell's neighboring cells from right side in descending order because we also need to make those satisfied or more close to satisfied. So, for each operation, we give the maximum possible loads from left side cell to its neighboring cell then the obtained values of both cells will be closer to zero which is closer to be satisfied. Here, actually for each operation at least one of the cells must be satisfied. For each operation, we remove the satisfied cell from the bipartite graph. If both cells satisfied in each operation, we will remove both satisfied cells from the bipartite graph. If a cell in left side becomes satisfied then we go to the next left side cell instead of seeing other neighboring cells because that particular cell already becomes satisfied. So it can not make satisfy any other cells because if it gives its remaining cells to its neighbor then it becomes unsatisfied. But if that left side cell is not satisfied then we continue to see its next neighboring cells because its previous neighboring cell becomes satisfied but not itself. So, it has a chance to satisfy itself if it has any neighboring cell exist. If it has no longer any neighboring cells exist and it is not satisfied still now, it will remain unsatisfied cell forever. We will go to the next left side cell to operate that same process. In this way maximum possible number of cells will be satisfied and also maximum number of possible cell's obtained value will be more close to zero.

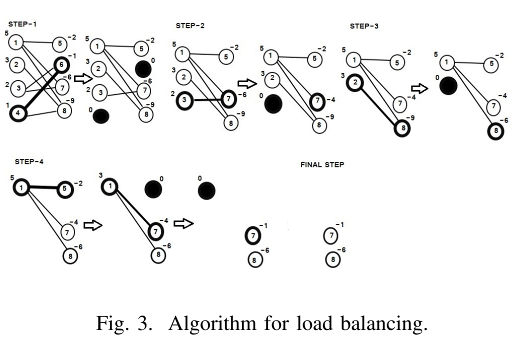


Fig. 3. Algorithm for load balancing.

Algorithm 1 Proposed algorithm for load balancing.

```
Model Bipartite Graph(G) of cells
Assign obtained values of cells in Array(O)
Sort the positive obtained value carrier cells in ascending order
for Each cell u ∈ positive obtained value cells do
    Sort all cells v ∈ G.Adj[u] in descending order
    for Each vertex v ∈ G.Adj[u] do
        if (O[u] > absolute(O[v])) then
            O[u] = O[u] − absolute(O[v])
            O[v] = 0
            Remove cell id− >v and remove all the edges connecting to that cell from Bipartite Graph(G)
        else if (O[u] < absolute(O[v])) then
            O[v] = O[v] + O[u]
            O[u] = 0
            Remove cell id− >u and remove all the edges connecting to that cell from Bipartite Graph(G)
            break
        else
            O[u] = 0
            O[v] = 0
            Remove cell id− >u,v and remove all the edges connecting to that cells from Bipartite Graph(G)
            break
        end if
    end for
end for
```

D. *Shapley value*

Shapley et al. [5] proposed a cooperative solution for the distribution of load balancing among n players based on their marginal contribution.in this concept all the players in the coalition will receive payments or shares based on their contributions to each possible coalition.The payoff function of Shapley Value for each player i is given as

$$\Phi_i(N,v) = \frac{1}{N!}\sum_{S\subseteq N\{i\}} |S|!(|N|-|S|-1)![v(s)-v(s\cup\{i\})] \tag{11}$$

For n players, the set of coalitions, 2^N; here i is the index for each player, N is the set of n players.The formula can be interpreted as follow. S is the set of a coalition then a cell i can contribute for each coalition is $[v(S\cup\{i\})v(S)]$ the marginal contribution of i to coalition S,where $v(S)$is the value including i and $v(S\cup i)$ the value of contribution to $|S|$ without i. $1/N!$ denotes that each player should take the average of this contribution for different permutation in which the coalition can form.

E. *Algorithm analysis*

Fig. 3 shows the steps about how the algorithm works. Here is a particular scenario of some cells forming a bipartite graph. Applying the algorithm 1 in Fig. 3, we find how it works.

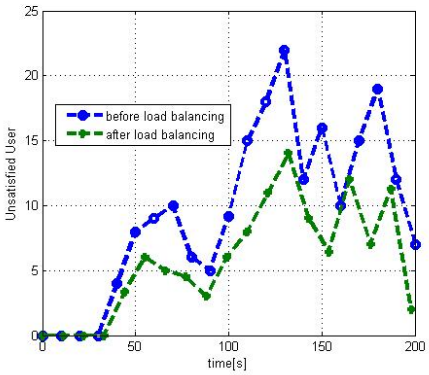


Fig. 4. Unsatisfied users vs time

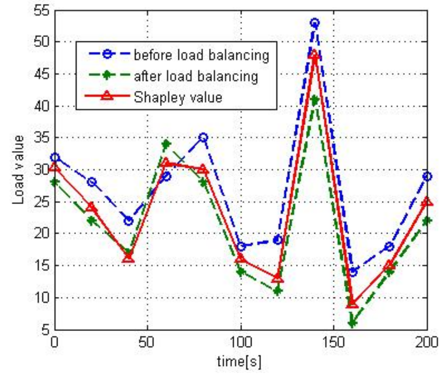


Fig. 5. Load values vs time

In each step if any cell becomes satisfied, it is marked as bold and it will be removed from the bipartite graph. All of its edges will also be removed. In the final step we can see that 2 cells are remain unsatisfied forever.

V. Experimental Results and Evaluation

A. Settings of parameters

For simulation we design a homogeneous network with 19 sites, 3 sectors per site. That means total 57 cells in the network. In all simulation we use default values for the parameters mentioned in Table I.

TABLE I
CONSIDERED PARAMETERS.

Parameter	*Value*
Network Layout	19 BS site, 3 Sectors, 57 cells
N_{tot}	50 PRBs
eNodeB TX Power	40 W
Threshold load	0.85
$\Delta RSRP_{th}$	5 dB
D_u	512 kbps
Inter site distance	500 m

B. Evaluation

For the simulation and evaluation we use Matlab. Shapley value provides the best way for load balancing in a cooperative manner where maximum number of cells become satisfied. Since all cells in the coalition get benefits, all wish to join in the coalition. In the simulation we consider a small homogeneous network. Some cells cant be satisfied from our simulation but if we consider for a grand coalition then the number of unsatisfied users decreases.

In Fig. 4 we observe that the number of unsatisfied users decreases over time after load balancing due to the availability of free PRBs. Overloaded cells that are much closer to its under loaded neighboring cell will get maximum resources. As a result, the overloaded cell can distribute this resource to its cell users.

In Fig. 5 we notice that after load balancing the Shapley value line is in middle. Since, overloaded cells offload and underloaded cells get this load. So, Shapley value falls in the middle between before load balancing and after load balancing curves.

VI. Conclusion

Load balancing is an important research issue for the increased performance in a LTE. Future technology demands a self-organizing algorithm for load balancing which is not vendor specified. In this paper, we propose a self-organized algorithm for load balancing. Our method is based on co-operative game theory which provides a way to allocate the resources efficiently in a distributed network. We compare and evaluate our proposed method before and after load balancing to get the number of satisfied users and the load values. Our simulation results show better performance in terms of satisfied users after balancing the load using our algorithm. Future work includes the consideration of a heterogeneous network and to implement a faster algorithm for load balancing.

References

[1] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. M. Mao, S. Sen, and O. Spatschek, "An In-depath Study of LTE: Effect of Network Protocol and Application Behavior on Performance," in *Proceedings of the ACM SIGCOMM*, 2013, pp. 363–374.

[2] S. Kumar, E. Hamed, D. Katabi, and L. E. Li, "LTE Radio Analytics Made Easy and Accessible," in *Proceedings of the ACM SIGCOMM*, 2014, p. 12.

[3] A. Awada, I. Viering, B. Wegmann, and A. Klein, "Application of Game Theory for Load Balancing in Long Term Evolution Networks," *Frequenz Journal of RF-Engineering and Telecommunications*, vol. 10, pp. 180–184, 2010.

[4] Cai, Jianfeng, and U. Pooch, "Allocate Fair payoff for Cooperation in Wireless Ad hoc Networks using Shapley Value," in *Proceedings of the IEEE International Conference on Parallel and Distributed Processing Symposium*, 2004, p. 219.

[5] L. S. Shapley, "A Value for n-person Games," *In Contribution to the Theory of Games*, vol. 28, pp. 307–317, 1953.

[6] C. Frank and K. Romer, "Algorithms for Generic Role Assignments in Wireless Sensor Networks," in *Proceedings of the ACM Conference on Embedded Networked Sensor Systems*, 2005.

[7] A. Awada, B. Wegmann, I. Viering, and A. Klein, "A Game-Theoretic Approach to Load Balancing in Cellular Radio Networks," in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications Conference*, 2010.

[8] O. Altrad and S. Muhaidat, "Load Balancing Based on Clustering Methods for LTE Networks," *Journal of Selected Areas in Telecommunications*, vol. 2, pp. 1–6, 2013.

Bangladeshi Road Sign Detection Based on YCbCr color model and DtBs Vector

Soumen Chakraborty
Dept. of CSE
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
Soumen0123@yahoo.com

Kaushik Deb
Dept. of CSE
Chittagong University of Engineering & Technology
Chittagong, Bangladesh
debkaushik99@cuet.ac.bd

***Abstract*— Detection of road sign from a road image is very crucial for intelligent transportation system for awareness of drivers and blind pedestrians. In this regard, a framework has been proposed in this paper for detection of the road sign from road images. For detection of the road sign, two natural properties of a Bangladeshi road sign is utilized, they are – border color rim of the road sign and the shape of the road sign. Based on these ideas initially, YCbCr color model is used to eliminate the illumination sensitiveness for segmentation. In the second step, statistical threshold value is used for color segmentation. Next, labeling and filtering is used to extract the shapes. Finally, Distance to Borders (DtBs) vector is used to verify the region of interest (ROI) for detecting the Bangladeshi road sign (BRS). Various road images are used with a variety of conditions to test the proposed framework and results are presented to prove its efficiency.**

Keywords*— *YCbCr color model, road sign detection, DtB vector

I. Introduction

During the past few years, people start to pay more and more attention on the advanced, efficient and accurate intelligent transportation systems (ITSs). In the field of computer vision or digital image processing, detection of specific object in an image is one of the most difficult topics. Detection of road sign (RS) region in a road image is such a topic. The Road Sign Detection (RSD) is widely used for driving assistance, autonomous vehicle, traffic rules awareness, blind pedestrian awareness etc. RSD task is quite challenging from road images due to the multi-shape and multi-color road sign formats, viewpoint changes, same color or shape objects in surroundings and the non-uniform outdoor illumination conditions during image acquisition. In addition, RSD system should operate fast enough (real time) to aware autonomous driving assistance, the needs of ITS and not to miss a single interest object from the road image. These systems minimize human interaction which enhances the performance to detect the road signs and follow the traffic rules.

Various road sign detection methods have been developed. This section provides a descriptive summary of some methods that have been implemented and tested for RSD. As far as detection of the road sign region is concerned, researchers have found many methods of finding RS. For example, in [1], HSV color model is used for image segmentation with heuristic color thresholding the image and Gielis curve fitting algorithm is used to filter the ROI shape with high accuracy though its more time consuming technique for filtering process.

YCbCr color space is used for the color segmentation and template matching technique is used for shape classification in [2]. In [3], RGB color space is used with chromatic and achromatic colors for color segmentation. ROI based on geometric means for detection state is used. In [5], a method based on two new criteria for analyzing different shapes with basic geometric prosperities has been proposed. The relationship between area, perimeter and the number of sides of a given shape is considered as shape classification criteria. Image segmentation is used for integration of shape analysis. Another shape filtering technique is used in [6], by Gabor Wavelet filter. One vs. rest SVM is used by this author for recognition or classification of the road sign. SVM is also used in [4] for classification section.

A color centroid matching using YCbCr has been used to detect the RS. The main thought of color centroid matching is for detecting the two color road sign like red-white, red-blue etc [7]. An efficient method is proposed in this paper for shape analysis and verification. Authors used SVM to verify the RS-shape and HSI color model is used for color segmentation [8]. In [9], HSI color model is used to segment the image. Shape classification using linear SVM is used in proposed method. And In [10], A nonlinear composite correlation filters is combined a bank to design a versatile road sign recognition system. Fuzzy logic has been applied to detect the RS. Some intuitive rules has been developed to describe the road sign colors gave some membership functions for fuzzy sets based on hue and saturation [11].

In this research, an algorithm enforcing a new method to select automatically statistical threshold value in YCbCr color model for detecting candidate region in Bangladeshi road sign. In this paper, color analysis has been based on Bangladeshi road sign that is because of there are mostly two color rim of road sign in Bangladesh. Those colors are red and blue. These candidate regions of interest may include road sign regions with some false alarm. Labeling and filtering is used to filter out the false alarm. In filtering process basic geometrical properties of road sign shape has been used. Distance to Borders Vector will verify the shape of road sign from the candidate ROI and detect the shape of the road sign.

978-1-4673-8342-4/15/$31.00 © 2015 IEEE

The rest of the paper is structured as follows: We described in Section II the system overview. Then we presented the Experimental Results in Section III. And at last Conclusion is given in Section IV.

II. SYSTEM OVERVIEW

Road sign is different from any other objects of the environment based on its color and shape. So color analysis and shape analysis are very important issues for segmentation of the road sign from any other objects. In this paper we have been emphasized on both color image processing for multi-color road sign of Bangladesh and filter out the segmented objects with basic geometrical shape analysis. The proposed algorithm for detection and verification of road sign which consists of four distinct parts is shown in Fig. 1. Initially deals with the conversion of road image from RGB color model to YCbCr color model for extracting candidate regions and to avoid the illumination sensitiveness of color. The second part allows procedures for detecting ROI by using color segmentation in YCbCr color space. In the third part the region of interest is refined using labeling and filtering and different geometrical properties such as area, aspect ratio, perimeter are used to determine whether the ROI contains road sign or not. Finally the verification of the shapes of road sign with DtBs vector is shown to detect the BRSs.

An input dataset has been created by the real time image of road sign collected from the road and high way of Bangladesh. The image has been collected by the high definition digital camera in RGB format. The dataset is collected with cluttered and non-cluttered background with different lighting condition. Occluded and shape invariant road sign image is also added in dataset which is the challenge of my proposed framework.

A. Convert to YCbCr Color Model

There is a common issue of RGB color image is illumination sensitiveness. The conversion of color space is a way to avoid this problem. In this paper the YCbCr color model has been used because most of the BRS image is based on red or blue rim. Moreover the YCbCr color space can eliminate the illumination of the image. In Y we get all the intensity of the color. And in Cb and Cr we get the red and blue colors. It is an easier way to separate the red color or blue color object simply from the Cb and Cr part of the image. Here we will get details of the blue and red objects from the image. Equation (1) shows the basic conversion of road image from RGB to YCbCr. A threshold value is needed for Segmentation to convert image into binary image.

$$[Y\ \mathrm{Cb}\ \ \mathrm{Cr}] = [\mathrm{R}\ \ \mathrm{G}\ \ \mathrm{B}]\begin{bmatrix} 0.299 & -0.168935 & 0.499813 \\ 0.587 & -0.331665 & -0.418531 \\ 0.114 & 0.50059 & -0.081282 \end{bmatrix} \quad (1)$$

B. Image Segmentation

In RGB color space, the three spectral components (e.g., red, green, and blue) are combined additively to produce the resultant color. On the other hand, in YCbCr color space, any color is represented by the values of its intensity. Many machine vision systems use the YCbCr color model for identifying the color of different objects to avoid illumination sensitiveness. We also use the YCbCr color model to obtain the candidate regions from given road images in the RGB color format. In YCbCr color space: Y represents the light intensity of all colors which represent the intensity of all three RGB components, Cb represents the blue color intensity and Cr represents the red color intensity components of the image. RGB-to-YCbCr conversion equations is used to convert the color space and then red or blue rim of road sign information is used to detect candidate regions from Cb and Cr components and geometric properties of road sign to reduce the number of BRS-like candidates.

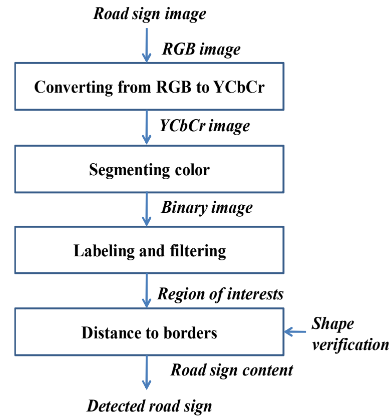


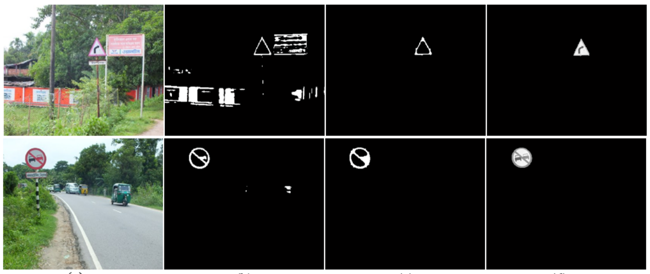
Fig. 1. The proposed framework for BRSD.

In our experiments, the color properties of mean and standard deviation values of intensity of the converted YCbCr road images are used to detect white road sign rim (e.g., red, blue) pixels. To estimate these properties, we used 25% images of road images taken under different illumination and weather conditions. The binarization process of obtaining red and blue RS pixels can be formulated as follows:

$$b_{blue} = \begin{cases} 1, & [Cb(x,y) \geq (\mu^{Cb} + \sigma^{Cb})] \\ 0, & otherwise \end{cases} \quad (2)$$

$$b_{red} = \begin{cases} 1, & [Cr(x,y) \geq (\mu^{Cr} + \sigma^{Cr})] \\ 0, & otherwise \end{cases} \quad (3)$$

Where b_{blue} and b_{red} are the blue and red candidate binary masks, *Cb(x,y)* and *Cr(x,y)* are the blue and red color intensity component of the pixel, and μ^{Cb}, μ^{Cr} are the mean intensity of and σ^{Cb}, σ^{Cr} are the standard deviation for the blue and red color rim RS pixels of the sample data, respectively shown in Eq. (2 & 3). Road images and color segmentation results are shown in Fig. 2. Color segmentation parameters are very sensitive in order to detect as many candidates as possible. All false candidates will be filtered out on the next stage.



(a) (b) (c) (d)

Fig. 2. Steps for road sign color segmentation: (a) input road images, (b) color segmentation result, (c) extracted candidate region after filtering, (d) candidate region extraction.

After color segmentation, there still exist noises such as small holes, occlusion in the target candidate regions that are not ideal. These noises are eliminated using a morphological closing operation to recover the gaps smaller than 5 consecutive pixels on the segmented objects. It plays an important role in calculating the DtBs Vector.

There are two examples of input image shown in Fig. 2(a). The first row image of the figure is with cluttered background and lots of RS-like objects and false alarms. In the second row, there is an example of circular red rim road sign. Color detection using YCbCr color model is used with statistical threshold shown in Fig. 2(b) for the input image of 2(a). The 3(b) images have lots of false alarm for same color image but due to the aspect ratio, area and geometrical shapes, false alarms have been omitted in 3(c). In filtering, challenges have been faced for scale invariant and false alarm of the surroundings. Morphological analysis is performed when any holes or gaps of objects are contained shown in Fig. 2(c).

C. Labeling and Filtering

The next step of the proposed framework is labeling the connected components. Eight connected components are used for labeling. After this operation, in order to correctly differentiate the RS regions from other RS-like regions, geometrical features such as the area and aspect ratio of each region are to be extracted.

Regions that satisfy the properties mentioned in Table I for red and blue road signs are considered as candidate regions. These parameters are used in the filtering operation to eliminate RS-like objects from the candidate list. Fig. 2 illustrate the steps for RS region extraction.

TABLE I. FILTERING PROPERTIES

Filtering Parameter	Candidate Region
Aspect Ratio	[0.8, 1.2]
Area	[40, 10000]

D. Shape Verification

Distance to Borders (DtBs) vector is a robust method for rotated and scaled objects to recognize the shape. It will be the input of verification stage. It is very efficient way of pattern recognition to recognize the shape of the road sign. To calculate the DtBs we have been calculated the distance between bounding box (BB) and Perimeter position with the Euclidian distance. A BB has 4 sides so the DtBs will generate 4 vectors for every connected candidate objects. For four sides calculation we needed to calculate vectors for 4 separate times. In Fig. 3(c1) there are two objects which are triangular and circular. Fig. 3(e1) represents the DtBs vector of triangular shape and Fig. 3(f1) represents the circular shape. In Fig. 3(c2) there are two objects which are quadrangular and triangular. Fig. 3(e2) represents the DtBs vector of quadrangular shape and Fig. 3(f2) represents the triangular shape. Here Fig. 3(d1 & d2) shows the extracted road sign with DtBs verification processes. DtBs represent those patterns to take decision about the characteristics of the BRS.

III. EXPERIMENTAL RESULTS

From the proposed system we have been implement the detection system of the multi-shape and multi-color road signs from the road image. The input dataset has been collected from the different high way road of Bangladesh in different weather and lighting condition with the high definition digital camera. The images were captured at different distances between camera and road sign. Hence our input data set is realistic. The frame size of image is resized as 448X336 pixels. Fig. 4 shows some example images taken under different illuminations and weather conditions, with complex scenes and occluded road sign with trees or any other objects. The data set is classified into three shapes of road sign which are shown in Table II. Detection result is illustrated with these three categories in the following table. Celeron Dual Core CPU 2.10GHz processor computer with 2GB ram has been used to execute the program to detect the BRS. The program has been developed under the MATLAB environment.

Considering the above conditions, the successful road sign detection rate is about 91.73%. Table II shows the results of the extracted road sign. Here all images are Bangladeshi road image, with different shapes and different colors, taken from natural scenes. These images are obtained with non-uniform outdoor illumination conditions at various angles. In our experiment, in total 121 images are used as input road image and 111 images have been detected. Some of image has one road sign and some of has more than one road sign. All the detection results include the occluded and scale invariant road sign which shown in Fig. 2 & 4.

The average computational time at different stages of the proposed framework is shown in Table IV. The average computational time of the full method is 0.58s – 0.71s.

TABLE II. DETECTED RESULTS OF PROPOSED SYSTEM

Shape of Road Sign	Total Image	Extracted RS	Success rate (%)
Triangular Road Signs	71	65	91.54
Circular Road Signs	46	42	91.30
Quadrangular Road Signs	4	4	100
Total	**121**	**111**	**91.73**

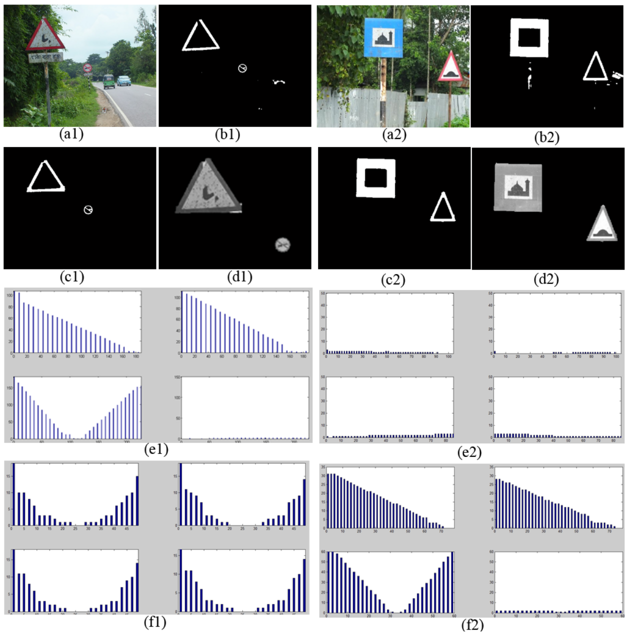


Fig. 3. DtBs analysis steps: (a1, a2) are the input image. (b1, b2) are the color segmented result, (c1, c2) are extracted candidate region after filtering, (d1, d2) are the extracted candidate region, (e1) is the DtBs of e1's triangular image, (f1) is for e1's circular image, (e2, f2) also the DtBs of c2's quadrangular and triangular image respectively.

Fig. 4. Some input of road image.

TABLE III. COMPARISON WITH RESPECT TO DETECTION ACCURACY

Method	Detection accuracy
The proposed method	91.73%
[4]	87.6%

TABLE IV. COMPUTATIONAL COST OF PROCESSES

Methods	Estimated Computational Cost
Image acquisition	0.016s
Image conversion	0.02s
Binarization technique	0.004s
Morphological analysis	0.035s / object
Labeling and filtering	0.08s / object
Verification with DtB	0.2s / object

IV. CONCLUSION

This work describes a complete detection process of Bangladeshi road sign has been considered the complication of outdoor environment. Segmentation of road sign from given road image using YCbCr color space with the statistical thresholding allows us to avoid the illumination sensitiveness and choose all road sign like color objects. Geometrical shape analysis is used to filter the false alarm from the segmented image in our work. We have been used DtBs using Euclidean distance between the bounding boxes and the perimeter of the object which is used to verify the shape of the RS. Our experimental results represent that this system is robust because it can detect multi-colors and multi-shapes road sign which were the big challenges to detect the Bangladeshi road sign. It shows 91.73% of accuracy to detect those types of realistic road signs from Bangladeshi road image.

References

[1] Manjare, Pratiksha, and S. M. Hambarde. "Image Segmentation and Shape Analysis for Road Sign Detection." ESRSA Publications on International Journal of Engineering Research and Technology. Vol. 3, No. 12, pp. 668-672, 2014

[2] Hechri, Ahmed, and Abdellatif Mtibaa. "Automatic detection and recognition of road sign for driver assistance system." 16th IEEE Mediterranean on Electrotechnical Conference (MELECON), 2012, pp. 888-891.

[3] Abukhait, Jafar, et al. "Road sign detection and shape recognition invariant to sign defects." IEEE International Conference on Electro/Information Technology (EIT), 2012, pp. 1-6.

[4] Bui-Minh, Thanh, Ovidiu Ghita, Paul F. Whelan, and Thai V. Hoang. "A robust algorithm for detection and classification of traffic signs in video data." InControl, Automation and Information Sciences (ICCAIS), 2012 International Conference on, IEEE, 2012. pp. 108-113.

[5] Khan, Jesmin F., Sharif Bhuiyan, and Reza R. Adhami. "Image segmentation and shape analysis for road-sign detection." IEEE Transactions on Intelligent Transportation Systems, Vol. 12, No. 1, 2011, pp: 83-96.

[6] Fatmehsan, Younes Rakhshani, Alireza Ghahari, and Reza Zoroofi. "Gabor wavelet for road sign detection and recognition using a hybrid classifier." IEEE International Conference on Multimedia Computing and Information Technology (MCIT), 2010, pp. 25-28.

[7] Chourasia, Jitendra N., and Preeti Bajaj. "Centroid based detection algorithm for hybrid traffic sign recognition system." IEEE International 3rd Conference on Emerging Trends in Engineering and Technology (ICETET), 2010, pp. 96-100.

[8] Kiran, C. G., et al. "Traffic sign detection and pattern recognition using support vector machine." IEEE Seventh International Conference on Advances in Pattern Recognition, ICAPR'09, 2009. pp. 87-90.

[9] Maldonado-Bascón, Saturnino, et al. "Road-sign detection and recognition based on support vector machines." IEEE Transactions on Intelligent Transportation Systems, Vol. 8, No.2, 2007, pp: 264-278.

[10] Perez, Elisabet, and Bahram Javidi. "Nonlinear distortion-tolerant filters for detection of road signs in background noise." *Vehicular Technology, IEEE Transactions on* 51.3 (2002): 567-576.

[11] Fleyeh, Hasan. "Road and traffic sign color detection and segmentation- a fuzzy approach." Red 250 (2005): 207, pp. 124-127.

GIS-based Extraction of Open Space in Rajshahi City

Fatima Jahan Sarmin
Department of Computer Science and Engineering,
Rajshahi University of Engineering and Technology
Rajshahi, Bangladesh
sarmin_ruet_cse@yahoo.com

Md. Shahid Uz Zaman
Dept. of Computer Science and Engineering,
Rajshahi University of Engineering and Technology
Rajshahi, Bangladesh
szaman22@yahoo.com

***Abstract*—In this paper we present a way to extract features and information from a digitize GIS-based map and for obtaining the open space in Rajshahi city. As we know the environment plays a most vital role for our living, the unplanned use of land can be a great threat to our environment and life. In this respect Geographic information system (GIS) gives us a way to keep track of land use through a digital map. We can capture, digitize, and store information and data in a GIS map. This paper aims to find out the useful data from the GIS based digitize map and analyze the information. After extracting the information a yearly based analysis is made and shows how terribly the number of building in Rajshahi area is increasing day by day since 1931 to 2000. And finally it is tried to make a forecast about the increasing the effect of it.**

Keywords— geographic information system (GIS); feature extraction; digital maps

I. Introduction

In Bangladesh unprecedented population growth coupled with unplanned developmental activities has resulted in rapid but skewed urbanization. This has posed serious implications on the resource base, access to infrastructure and the development of the region. The problems created by the unplanned and unrestricted growth of city aggravates irregular and chaotic development of residential, industrial and commercial areas resulting in traffic bottle necks, slums, polluted environment and others all known and felt by the residents of the city. This is reflected in almost every city in Bangladesh.

In Rajshahi, a divisional city in Bangladesh, increasing the number of land use for urbanization is becoming more important issue as the agricultural lands as well as trees are decreasing due to use the land for other purposes. Geographical information system (GIS) is a way to keep track of information about the land use and analyze the information for further processing. After processing the information, it would be helpful to show the effect of urbanization on environment.[1]

Thus the goal of the paper is to extract the features from the GIS based digital map, analyze the effect of urbanization and measure the open space available for present and future.

II. Methodology

A GIS-based map is used to extract the specific features.
Few steps are used to complete the whole process: [2]

1. Taking digitize map as input
2. Linking ArcGIS with python
3. Algorithm for extracting features
4. Analyze the data
5. Mapping processed data into GIS map

The following map in Fig. 1 is used as input (Digital GIS-map) collected from RDA “Rajshahi Development Authority”.

Here is the flowchart in Fig. 2 for extracting information from the input image. Any kind of information such as information of buildings, the owners of those buildings, year of establishment, types of the building, total area used etc can be extracted from the GIS-based map using the flowchart.

After extracting the information, following steps are done-

- Taking 10 year basis data from that digital map, geo-processing and geo-coding are done with arcGIS 10.1 which are shown in the result section.
- After geo-processing and geo-coding, a statistical analysis is made.

Finally, a forecasting is made from the extracting data to show the effect of open space in Rajshahi area.

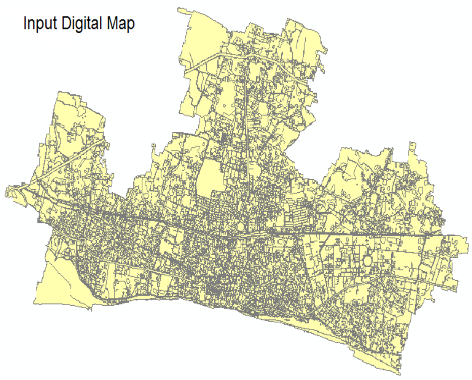


Fig. 1. Input Digital Map of Rajshahi

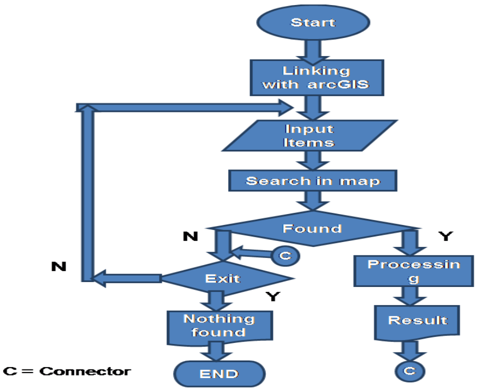


Fig. 2. Flowchart of extracting facility.

III. RESULTS

The results are processed with ArcGIS 10 and Python with Windows 7.

Fig. 3 showing polygons in the map and when a specific polygon is select and its corresponding information displayed in the Fig. 4.[4]

After extracting information, geo-processing and geo-coding are done to make further processing. [3] The following figures show the number of building in Rajshahi city during last 10 years decreasing the open space considerably. Here, red dots (actually polygon boundary by red border) are showing the increasing number of buildings from 1931 to 1940 in Fig. 5.

Fig. 6 is showing a closer view of the above map so that user can understand that every red bounded building was built in 1931 to 1940. Each polygon denotes building features.

Fig. 7 is showing the increasing number of buildings in 1941 to 1950 where green dots denote the buildings.

Fig. 8 is showing the increasing number of buildings in 1951 to 1960 where purple dots denote the buildings.

Fig. 9 is showing the increasing number of buildings in 1961 to 1970 where green dots denote the buildings.

Fig. 10 is showing the increasing number of buildings in 1971 to 1980 where yellow dots denote the buildings.

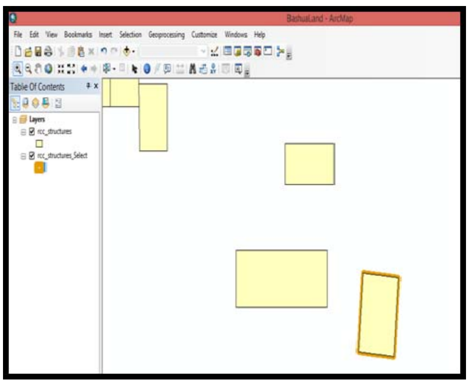


Fig. 3. Selecting a specific polygon (arcGIS view)

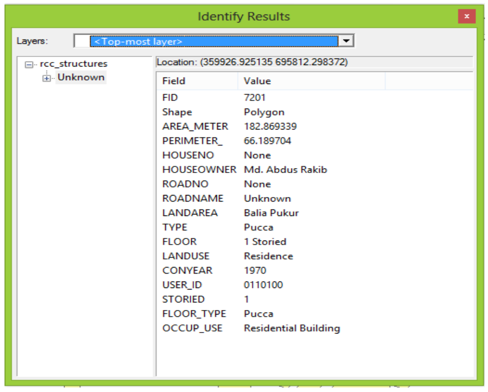


Fig. 4. Extracting information of a selected polygon

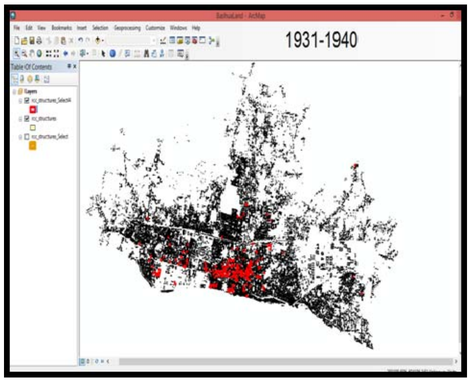


Fig. 5. Data from 1931 to 1940 (arcGIS view)

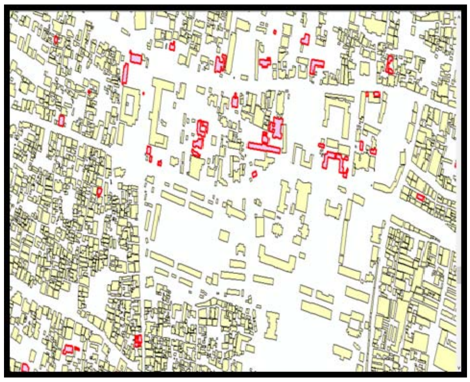


Fig. 6. Data from 1931 to 1940 (arcGISCloser view)

Fig. 11 is showing the increasing number of buildings in 1981 to 1990 where brown dots denote the buildings.

Fig. 12 is showing the increasing number of buildings in 1991 to 2000 where red dots denote the buildings. This is showing how terribly the numbers of buildings are increasing in Rajshahi area.

There is a summary of increasing number of buildings from 1931 to 2000 in table no 1. This table shows how the number of buildings is increasing in every 10 years. Fig. 13 is showing its histogram.

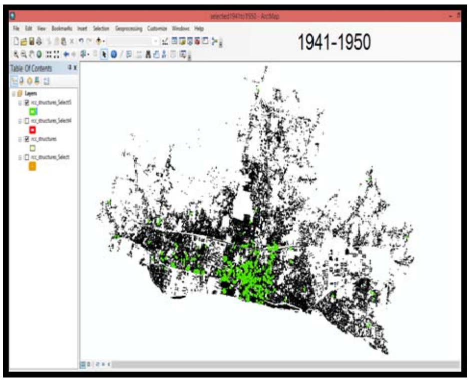


Fig. 7. Data from 1941 to 1950 (arcGIS view)

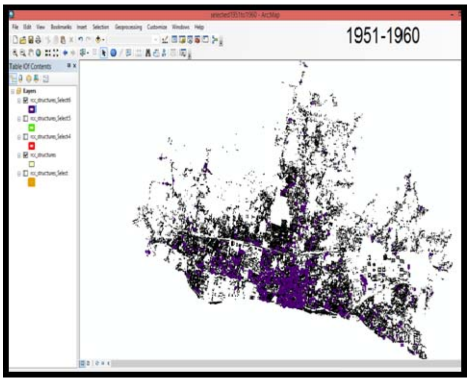


Fig. 8. Data from 1951 to 1960 (arcGIS view)

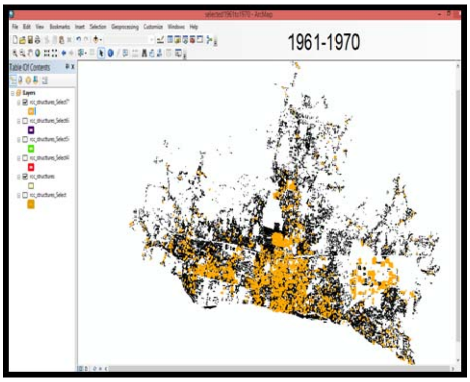


Fig. 9. Data from 1961 to 1970 (arcGIS view)

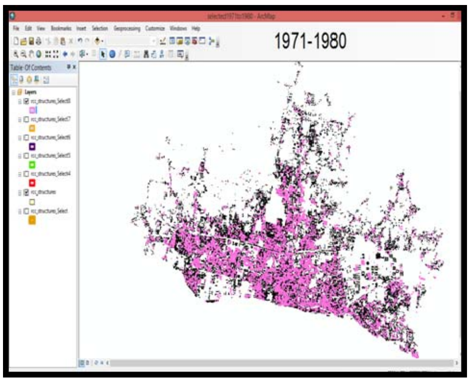


Fig. 10. Data from 1971 to 1980 (arcGIS view)

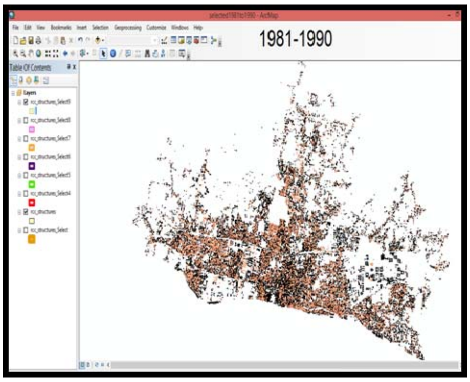


Fig. 11. Data from 1981 to 1990 (arcGIS view)

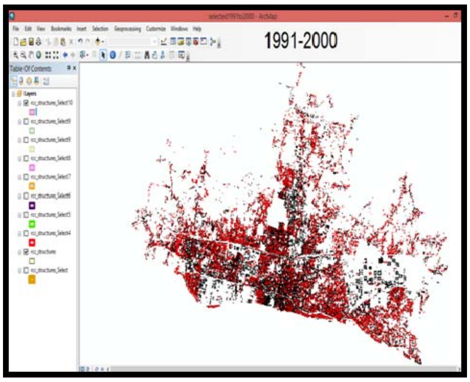


Fig. 12. Data from 1991 to 2000 (arcGIS view)

TABLE I. INCREASING NUMBER OF BUILDINGS PER 10 YEAR

Year	Number of buildings made
1931-1940	393
1941-1950	612
1951-1960	1713
1961-1970	5016
1971-1980	12461
1981-1990	23973
1991-2000	32373

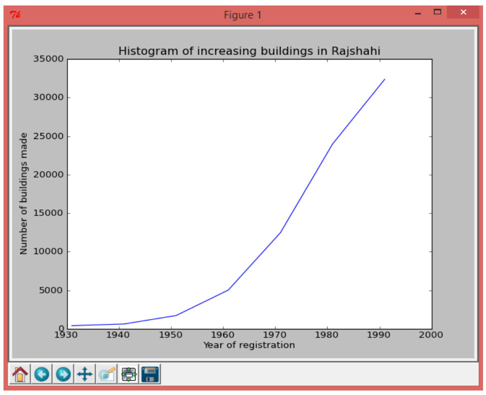


Fig. 13. Histogram of increasing buildings in Rajshahi

After analyzing the occupied land area, the bar chart in Fig. 14 showing how the land area is decreasing in every 10 years from 1931 to 2000.

If the data from 1931 to 2000 is analyze, a forecasting can be made. This forecasting in Fig. 15 is showing, if every 10 year the open space is decreasing, there will be no open space left after 2041 and it will be a threaten to environment in Rajshahi as well as for the world.

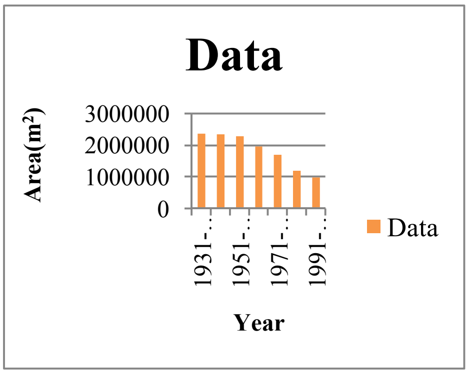


Fig. 14. Bar chart of decreasing open space in Rajshahi

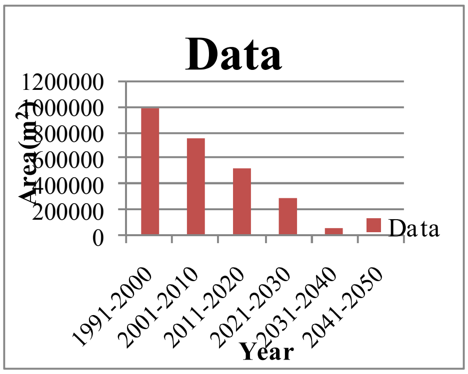


Fig. 15. Forecasting from the present data

IV. Conclusions

Safe environment is the most important for our survival. But unplanned using of land can make it unsafe and sometime dangerous for living in it. The main target of this paper is to make awareness by showing how the open space in Rajshahi city is decreasing day by day. As we take only building features to detect the open space in Rajshahi, in future this research can be extended for extracting other features like trees, water bodies, green area and other features in a GIS based map for any area and shows the effect for the whole environment.

V. Acknowledgment

Our appreciation goes to "Rajshahi Development Authority" for providing us the GIS-based digital map of Rajshahi city. The authors also want to thank the department of CSE, RUET for supporting the research in every aspect.

VI. References

[1] Qing Chang ,Xue Lia ,Xiulan Huang ,Jiansheng Wu, "A GIS-based Green Infrastructure Planning for Sustainable Urban Land Use and Spatial Development", Procedia Environmental Sciences 12 (2012) 491 – 498

[2] Norzailawati Mohd Noor, M.Zainora Asmawi, Alias Abdullah, "Sustainable Urban Regeneration: GIS and Hedonic Pricing Method in determining the value of green space in housing area", Procedia - Social and Behavioral Sciences 170 (2015) 669 – 679

[3] Salwa Farouk Elbeih, "An overview of integrated remote sensing and GIS for groundwater mapping in Egypt", Engineering Applications and Water Division, National Authority for Remote Sensing and Space Sciences, 24 August 2014

[4] N. Sofina, M. Ehlers, "*Object-Based Change Detection Using High-Resolution Remotely Sensed Data and GIS*", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B7, 2012 XXII ISPRS Congress, 25 August – 01 September 2012, Melbourne, Australia.

[5] T. Ohlhof , E.Gülch , H. Müller , C. Wiedemann , M. Torre "*Semi-Automatic Extraction of Line and Area Feature from Areal and Satellite Images*", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 34, Part XXX

[6] Zhifang Wang *"Planning for open space conservation: Using GIS to match cultural values and ecological quality of open spaces"*, thesis document, The University of Michigan, 2008.

[7] Joseph K. Berry, "*GIS TECHNOLOGY IN ENVIRONMENTAL MANAGEMENT: a Brief History, Trends and Probable Future",* Berry and Associates // Spatial Information Systems, Inc, Web: http://www.innovativegis.com/basis

[8] Barbara Huddleston, Ergin Ataman, Luca Fè d'Ostiani, "*Towards a GIS-based analysis of mountain environments and populations*", FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS, ROME 2003.

[9] Patric Wallis, "*The role of GIS technology in sustaining the built environment*", web document, ttps://www.esri.com/library / -the-built-environment.pdf ebooks/sustaining

[10] Prafull Singh *, Ankit Gupta, Madhulika Singh, "*Hydrological inferences from watershed analysis for water resource management using remote sensing and GIS techniques*", The Egyptian Journal of Remote Sensing and Space Sciences, 17 September 2014

[11] Masato Katoh 1,* and François A. Gougeon, "*Improving the Precision of Tree Counting by Combining Tree Detection with Crown Delineation and Classification on Homogeneity Guided Smoothed High Resolution (50 cm) Multispectral Airborne Digital Data*", Remote Sensing ISSN 2072-4292, . 2012, 4, 1411-1424; doi:10.3390/rs4051411

[12] Haifaa Al-Ballaa, Alexis Comber, Claire Smith, "*Distribution Pattern Analysis of Green space in Al-Madinah Using GIS*", Department of Geography, University of Leicester, Leicester LE1 7RH, UK , Web document.