ASDF

Proceedings of the 2nd International Conference on Information Engineering, Management and Security 2015

ICIEMS 2015

13-14 August 2015

Indian Institute of Technology Madras –Research Park
Chennai, India, Asia

Editors

Kokula Krishna Hari K
Daniel James, Julie Rue Bishop, Thiruvengadam B

ISBN 13 : 978-81-929742-7-9 $399
ISBN 10 : 81-929742-7-8

International Conference on Information Engineering, Management and Security 2015

ICIEMS 2015

Volume 1

By
ASDF, India

Financially Sponsored By
Association of Scientists, Developers and Faculties, India

Multiple Areas

13 – 14, August 2015
IIT-M Research Park,
Chennai, India

Editor-in-Chief
Kokula Krishna Hari K

Editors:

Kokula Krishna Hari K, Julie Rue Bishop, Thiruvengadam B

Published by

Association of Scientists, Developers and Faculties
Address: RMZ Millennia Business Park, Campus 4B, Phase II, 6th Floor, No. 143, Dr. MGR Salai, Kandanchavady, Perungudi, Chennai – 600 096, India.
Email: admin@asdf.org.in || www.asdf.org.in

International Conference on Information Engineering, Management and Security (ICIEMS 2015)

VOLUME 1

Editor-in-Chief: **Kokula Krishna Hari K**
Editors: **Kokula Krishna Hari K, Julie Bishop, Thiruvengadam B**

ISBN-13: 978-81-929742-7-9
ISBN-10: 81-929742-7-8

TECHNICAL REVIEWERS

- A Amsavalli, Paavai Engineering College, Namakkal, India
- A Ayyasamy, Annamalai University, Chidambaram, India
- A C Shagar, Sethu Institute of Technology, India
- A Kavitha, Chettinad College of Engineering & Technology, Karur, India
- A Padma, Madurai Institute of Engineering and Technology, Madurai, India
- A S N Chakravarthy, JNTU Kakinada, India
- A Tamilarasi, Kongu Engineering College, Perundurai, India
- Abdelbasset Brahim, University of Granada, Spain
- Abdelnaser Omran, Universiti Utara Malaysia, Malaysia
- Abdul Aziz Hussin, Universiti Sains Malaysia, Malaysia
- Abdul Nawfar Bin Sadagatullah, Universiti Sains Malaysia, Malaysia
- Abhishek Shukla, U.P.T.U. Lucknow, India
- Aede Hatib Musta'amal, Universiti Teknologi Malaysia, Malaysia
- Ahmed Mohammed Kamaruddeen, Universiti Utara Malaysia, Malaysia
- Ahmed Salem, Old Dominion University, United States of America
- Ali Berkol, Baskent University & Space and Defence Technologies (SDT), Turkey
- Alphin M S, SSN College of Engineering, Chennai, India
- Alwardoss Velayutham Raviprakash, Pondicherry Engineering College, Pondicherry, India
- Anand Nayyar, KCL Institute of Management and Technology, Punjab
- Anbuchezhiyan M, Valliammai Engineering College, Chennai, India
- Ang Miin Huey, Universiti Sains Malaysia, Malaysia
- Anirban Mitra, VITAM Berhampur, Odisha, India
- Ariffin Abdul Mutalib, Universiti Utara Malaysia, Malaysia
- Arniza Ghazali, Universiti Sains Malaysia, Malaysia
- Arumugam Raman, Universiti Utara Malaysia, Malaysia
- Asha Ambhaikar, Rungta College of Engineering & Technology, Bhilai, India
- Asrulnizam Bin Abd Manaf, Universiti Sains Malaysia, Malaysia
- Assem Abdel Hamied Mousa, EgyptAir, Cairo, Egypt
- Aziah Daud, Universiti Sains Malaysia, Malaysia
- B Paramasivan, National College of Engineering, Tirunelveli, India
- Badruddin A. Rahman, Universiti Utara Malaysia, Malaysia

- Balachandran Ruthramurthy, Multimedia University, Malaysia
- Balasubramanie Palanisamy, Professor & Head, Kongu Engineering College, India
- Brahim Abdelbasset, University of Granada, Spain
- C Poongodi, Bannari Amman Institute of Technology, Sathyamangalam, India
- Chandrasekaran Subramaniam, Professor & Dean, Anna University, India
- Choo Ling Suan, Universiti Utara Malaysia, Malaysia
- Cristian-Gyozo Haba, Technical University of Iasi, Romania
- D Deepa, Bannari Amman Institute of Technology, Sathyamangalam, India
- D Gracia Nirmala Rani, Thiagarajar College of Engineering, Madurai, Tamil Nadu
- D Sheela, Tagore Engineering College, Chennai, India
- Daniel James, Senior Researcher, United Kingdom
- David Rathnaraj Jebamani, Sri Ramakrishna Engineering College, India
- Deepali Sawai, Director - MCA, University of Pune (Savitribai Phule Pune University), India
- Dewi Nasien, Universiti Teknologi Malaysia, Malaysia
- Doug Witten, Oakland University, Rochester, United States of America
- Dzati Athiar Ramli, Universiti Sains Malaysia, Malaysia
- G A Sathish Kumar, Sri Venkateswara College of Engineering, India
- G Ganesan, Adikavi Nannaya University, India
- Ganesan Kanagaraj, Thiagarajar College of Engineering, Madurai, Tamil Nadu
- Geetha G, Jerusalem College of Engineering, Chennai, India
- Geetha V, Pondicherry Engineering College, Pondicherry, India
- Guobiao Yang, Tongji University, China
- Hanumantha Reddy T, RYM Engineering College, Bellary, India
- Hareesh N Ramanathan, Toc H Institute of Science and Technology, India
- Hari Mohan Pandey, Amity University, Noida, India
- Hidayani Binti Jaafar, Universiti Malaysia Kelantan, Malaysia
- Itebeddine GHORBEL, INSERM, France
- J Baskaran, Adhiparasakthi Engineering College, Melmaruvathur, India
- J Karthikeyan, Anna University, Chennai, India
- J Sadhik Basha, International Maritime College, Oman
- Jebaraj S, Universiti Teknologi PETRONAS (UTP), Malaysia
- Jia Uddin, International Islamic University Chittagong, Bangladesh
- Jinnah Sheik Mohamed M, National College of Engineering, Tirunelveli, India

- Julie Juliewatty Mohamed, Universiti Sains Malaysia, Malaysia
- K Latha, Anna University, Chennai, India
- K Mohamed Bak , GKM College of Engineering and Technology, India
- K Nirmalkumar, Kongu Engineering College, Perundurai, India
- K P Kannan, Bannari Amman Institute of Technology, Sathyamangalam, India
- K Parmasivam, K S R College of Engineering, Thiruchengode, India
- K Senthilkumar, Erode Sengunthar Engineering College, Erode, India
- K Suriyan, Bharathiyar University, India
- K Thamizhmaran, Annamalai University, Chidambaram, India
- K Vijayaraja, PB College of Engineering, Chennai, India
- Kamal Imran Mohd Sharif, Universiti Utara Malaysia, Malaysia
- Kannan G R, PSNA College of Engineering and Technology, Dindigul, India
- Kathiravan S, Kalaignar Karunanidhi Institute of Technology, Coimbatore, India
- Khairul Anuar Mohammad Shah, Universiti Sains Malaysia, Malaysia
- Kokula Krishna Hari Kunasekaran, Chief Scientist, Techno Forum Research and Development Center, India
- Krishnan J, Annamalai University, Chidambaram, India
- Kumaratharan N, Sri Venkateswara College of Engineering, India
- L Ashok Kumar, PSG College of Technology, Coimbatore, India
- Laila Khedher, University of Granada, Spain
- Lakshmanan Thangavelu, SA College of Engineering, Chennai, India
- M Ayaz Ahmad, University of Tabuk, Saudi Arabia
- M Chandrasekaran, Government College of Engineering, Bargur, India
- M K Kavitha Devi, Thiagarajar College of Engineering, Madurai, Tamil Nadu
- M Karthikeyan, Knowledge Institute of Technology, India
- M Shanmugapriya, SSN College of Engineering, Chennai, India
- M Thangamani, Kongu Engineering College, India
- M Venkatachalam, RVS Technical Campus - Coimbatore, India
- M Vimalan, Thirumalai Engineering College, Kanchipuram, India
- Malathi R, Annamalai University, Chidambaram, India
- Mansoor Zoveidavianpoor, Universiti Teknologi Malaysia, Malaysia
- Manvender Kaur Chahal, Universiti Utara Malaysia, Malaysia
- Mariem Mahfoudh, MIPS, France
- Marinah Binti Othman, Universiti Sains Islam Malaysia, Malaysia

- Mathivannan Jaganathan, Universiti Utara Malaysia, Malaysia
- Mehdi Asadi, IAU (Islamic Azad University), Iran
- Mohammad Ayaz Ahmad, University of Tabuk, Saudi Arabia
- Mohd Hanim Osman, Universiti Teknologi Malaysia, Malaysia
- Mohd Hashim Siti Z, Universiti Teknologi Malaysia, Malaysia
- Mohd Murtadha Mohamad, Universiti Teknologi Malaysia, Malaysia
- Mohd Zulkifli Bin Mohd Yunus, Universiti Teknologi Malaysia, Malaysia
- Moniruzzaman Bhuiyan, University of Northumbria, United Kingdom
- Mora Veera Madhava Rao, Osmania University, India
- Muhammad Iqbal Ahmad, Universiti Malaysia Kelantan, Malaysia
- Muhammad Javed, Wayne State University, United States of America
- N Rajesh Jesudoss Hynes, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India
- N Karthikeyan, SNS College of Engineering, Coimbatore, India
- N Malmurugan, Vidhya Mandhir Institute of Technology, India
- N Senthilnathan, Kongu Engineering College, Perundurai, India
- N Shanthi, Nandha Engineering College, Erode, India
- N Suthanthira Vanitha, Knowledge Institute of Technology, India
- Nasrul Humaimi Mahmood, Universiti Teknologi Malaysia, Malaysia
- Nida Iqbal, Universiti Teknologi Malaysia, Malaysia
- Nithya Kalyani S, K S R College of Engineering, Thiruchengode, India
- Nor Muzlifah Mahyuddin, Universiti Sains Malaysia, Malaysia
- Norma Binti Alias, Universiti Teknologi Malaysia, Malaysia
- P Dhanasekaran, Erode Sengunthar Engineering College, Erode, India
- P Ganesh Kumar, K. L. N. College of Information Technology, Madurai, India
- P Kumar, K S R College of Engineering, Thiruchengode, India
- P Ramasamy, Sri Balaji Chockalingam Engineering College, India
- P Raviraj, Kalaignar Karunanidhi Institute of Technology, Coimbatore, India
- P Sengottuvelan, Bannari Amman Institute of Technology, Sathyamangalam, India
- P Shunmuga Perumal, Anna University, Chennai, India
- P Tamizhselvan, Bharathiyar University, India
- P Thamilarasu, Paavai Engineering College, Namakkal, India
- Pasupuleti Visweswara Rao, Universiti Malaysia Kelantan, Malaysia
- Pethuru Raj, IBM Research, India

- Qais Faryadi, USIM: Universiti Sains Islam Malaysia, Malaysia
- R Ashokan, Kongunadu College of Engineering and Technology, India
- R Dhanasekaran, Syed Ammal Engineering College, Ramanathapuram, India
- R Muthukumar, Shree Venkateshwara Hi-Tech Engineering College, India
- R Nallusamy, Principal, Nandha college of Technology, Erode, India
- R Ragupathy, Kongu Engineering College, Perundurai, India
- R Sudhakar, Dr. Mahalingam College of Engineering and Technology, India
- R Suguna, SKR Engineering College, Chennai, India
- R Sundareswaran, SSN College of Engineering, Chennai, India
- Radzi Ismail, Universiti Sains Malaysia, Malaysia
- Rajesh Deshmukh, Shri Shankaracharya Institute of Professional Management and Technology, Raipur
- Rathika P, V V College of Engineering, Tirunelveli, India
- Rathinam Maheswaran, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India
- Razauden Mohamed Zulkifli, Universiti Teknologi Malaysia, Malaysia
- Reza Gharoie Ahangar, Islamic Azad University, Iran
- Roesnita Ismail, USIM: Universiti Sains Islam Malaysia, Malaysia
- Rohaizah Saad, Universiti Utara Malaysia, Malaysia
- Roselina Binti Salleh, Universiti Teknologi Malaysia, Malaysia
- Ruba Soundar K, P. S. R. Engineering College, Sivakasi, India
- S Albert Alexander, Kongu Engineering College, Perundurai, India
- S Anand, V V College of Engineering, Tirunelveli, India
- S Appavu @ Balamurugan, K. L. N. College of Information Technology, Madurai, India
- S Balaji, Jain University, India
- S Balamuralitharan, SRM University, Chennai, India
- S Balamurugan, Kalaignar Karunanidhi Institute of Technology, Coimbatore, India
- S Geetha, VIT University, Chennai, India
- S Jaganathan, Dr. N. G. P. Institute of Technology, Coimbatore, India
- S Natarajan, United Institute of Technology, Coimbatore, India
- S Poorani, Erode Sengunthar Engineering College, Erode, India
- S Prakash, Nehru Colleges, Coimbatore, India
- S Rajkumar, University College of Engineering Ariyalur, India
- S Ramesh, Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College, India
- S Selvaperumal, Syed Ammal Engineering College, Ramanathapuram, India

- S Selvi, Institute of Road and Transport Technology, India
- S Senthamarai Kannan, Kalasalingam University, India
- S Senthilkumar, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India
- S Vengataasalam, Kongu Engineering College, Perundurai, India
- Samuel Charles, Dhanalakshmi Srinivasan College of Engineering, Coimbatore, India
- Sangeetha R G, VIT University, Chennai, India
- Sanjay Singhal, Founder, 3nayan Consulting, India
- Saratha Sathasivam, Universiti Sains Malaysia, Malaysia
- Sarina Sulaiman, Universiti Teknologi Malaysia, Malaysia
- Sathish Kumar Nagarajan, Sri Ramakrishna Engineering College, Coimbatore, India
- Sathishbabu S, Annamalai University, Chidambaram, India
- Seddik Hassene, ENSIT, Tunisia
- Selvakumar Manickam, Universiti Sains Malaysia, Malaysia
- Shamshuritawati Sharif, Universiti Utara Malaysia, Malaysia
- Shankar S, Kongu Engineering College, Perundurai, India
- Shazida Jan Mohd Khan, Universiti Utara Malaysia, Malaysia
- Sheikh Abdul Rezan, Universiti Sains Malaysia, Malaysia
- Shilpa Bhalerao, Acropolis Institute of Technology and Research, Indore, India
- Singaravel G, K. S. R. College of Engineering, India
- Sivakumar Ramakrishnan, Universiti Sains Malaysia, Malaysia
- Somasundaram Sankaralingam, Coimbatore Institute of Technology, India
- Subash Chandra Bose Jeganathan, Professional Group of Institutions, India
- Subramaniam Ganesan, Oakland University, Rochester, United States of America
- Suganthi Appalasamy, Universiti Malaysia Kelantan, Malaysia
- Sunil Chowdhary, Amity University, Noida, India
- Suresh Sagadevan, Indian Institute of Science, Bangalore, India
- Syed Sahal Nazli Alhady, Universiti Sains Malaysia, Malaysia
- T Krishnakumar, Tagore Engineering College, Chennai, India
- T Ramayah, Universiti Sains Malaysia, Malaysia
- T Subbulakshmi, VIT University, Chennai, India
- T V P Sundararajan, Bannari Amman Institute of Technology, Sathyamangalam, India
- Tom Kolan, IBM Research, Israel
- Uma N Dulhare, Muffkham Jah College of Engineering & Technology, Hyderabad, India
- Uvaraja V C, Bannari Amman Institute of Technology, Sathyamangalam, India

- V Akila, Pondicherry Engineering College, Pondicherry, India
- V C Sathish Gandhi, University College of Engineering Ariyalur, India
- V Mohanasundaram, Vivekanandha Institute of Engineering and Technology for Women, India
- V Sathish, Bannari Amman Institute of Technology, Sathyamangalam, India
- V Vijayakumari, Sri Krishna College of Technology, Coimbatore, India
- Veera Jyothi Badnal, Osmania University, India
- Vijayalakshmi V, Pondicherry Engineering College, Pondicherry, India
- Vijayan Gurumurthy Iyer, Entrepreneurship Development Institute of India
- Vikrant Bhateja, Shri Ramswaroop Memorial Group of Professional Colleges (SRMGPC), India
- Wei Ping Loh, Universiti Sains Malaysia, Malaysia
- Yaty Sulaiman, Universiti Utara Malaysia, Malaysia
- Yongan Tang, Oakland University, Rochester, United States of America
- Yousef FARHAOUI, Moulay Ismail University, Morrocco
- Yudi Fernando, Universiti Sains Malaysia, Malaysia
- Yu-N Cheah, Universiti Sains Malaysia, Malaysia
- Zahurin Samad, Universiti Sains Malaysia, Malaysia
- Zailan Siri, University of Malaya, Malaysia
- Zamira Zamzuri, Universiti Kebangsaan Malaysia, Malaysia
- Zul Ariff Abdul Latiff, Universiti Malaysia Kelantan, Malaysia

PREFACE

It is a great honour to welcome for the International Conference on Systems, Science, Control, Communication, Engineering and Technology - ICIEMS 2015 at IIT-M Research Park, Chennai, India, Asia on 13 – 14 August, 2015.

ICIEMS 2015 aims to provide a chance for academic and industry professionals to share ideas on progress in the field of technology and to bring together the researchers and practitioners to discuss the problems and find solutions for the multifaceted aspects of Interdisciplinary Research Theory and Technology.

This conference provides an opportunity for various departments in the field of Engineering and Technology. It also focuses on the important aspects of advances in Systems, Science, Embedded Systems, Mobile Communications, Robotics, Engineering, Technology and Information Engineering, Management and Security. This conference highlights the new concepts and the improvements related to the research and technology.

It provides a chance for academic and industry professionals to discuss recent progress in the area of Interdisciplinary Research Theory and Technology. The proceeding of the conference consists of the information of various advancements in the field of Research and Developments globally and would act as a primary source for researchers to gain knowledge on the latest developments.

With the constant support and encouragement from the ASDF Global President Dr. S. Prithiv Rajan, ASDF International President Dr. P. Anbuoli and ASDF Governing Council Members, this conference will stay in our hearts. Without them, this proceeding could not have been completed within the shortest span.

Heartfelt Gratitude are due to the team members of Association of Scientists, Developers and Faculties – International, Family, Friends and Colleagues for their cooperation and commitment for making this conference a successful one.

Dr. K. Kokula Krishna Hari,
Chief Editor, ICIEMS.

Table of Content

Volume	01	ISBN	978-81-929742-7-9
Month	August	Year	2015

International Conference on Information Engineering, Management and Security 2015

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS001

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.001

INTEGRATING ASPECTS BASED ON OPINION MINING FOR PRODUCT REVIEWS

Suganya S[1], Sureka K[1], Vishnupriya P[1]
[1]V.S.B.Engineering College

Abstract: *It is a common practice that merchants selling products on the Web ask their customers to review the products and associated services. As e-commerce is becoming more and more popular, the number of customer reviews that a product receives grows rapidly. For a popular product, the number of reviews can be in hundreds. This makes it difficult for a potential customer to read them in order to make a decision on whether to buy the product. We aim to summarize all the customer reviews of a product. This summarization task is different from traditional text summarization because we are only interested in the specific features of the product that customers have opinions on and also whether the opinions are positive or negative. We do not summarize the reviews by selecting or rewriting a subset of the original sentences from the reviews to capture their main points as in the classic text summarization. We only focus on mining opinion/product features that the reviewers have commented on. A number of techniques are presented to mine such features. The proposed system is used to decisive the customer reviews for multiple product reviews then find the aspect from the review and classify the review whether they wrote positive or negative. We only focus on mining opinion/product features that the reviewers have commented on and compare the more product and rank the product based on the reviews automatically.*

Keywords: Aspect –based, Opinion mining, Product reviews

INTRODUCTION

With the inception of the Web 2.0 and the explosive growth of social networks, enterprises and individuals are increasingly using the content in these media to make better decisions. For instance, customers check opinions and experiences published by other customers on different Web platforms when they planning to buy the products through online. On the other hand, for organizations, the vast amount of information available publicly on the Web could make polls, focus groups and some similar techniques an unnecessary requirement in market research. However, due to the amount of available opinionated text, users are often overwhelmed with information when trying to analyze Web opinions. So far, many authors have tacked the problem of human limitation to process big amounts of information and extract consensus opinions from a large number of sources relying on data-mining-based tools. Considering a similar problem, this work is an effort to create a tool that offers a set of summarization methods and help users digest in an easy manner the vast availability of opinions.

Three main components of Opinion Mining are:
1. Opinion Holder: Person that expresses the opinion is opinion holder.
2. Opinion Object: Object on which opinion is given.
3. Opinion Orientation: Determine whether the opinion about an object is positive, negative or neutral. The core of our system is a novel extension of aspect-based opinion mining methodology, which was developed by us for online shopping of products. The core of our system is concerned with the fact that

Cite this article as: Suganya S, Sureka K, Vishnupriya P. "INTEGRATING ASPECTS BASED ON OPINION MINING FOR PRODUCT REVIEWS." *International Conference on Information Engineering, Management and Security (2015)*: 01-06. Print.

Users refer differently to different kinds of generic products when writing reviews on the Web. For instance, when a person writes a movie review, he probably comments not only movie elements, but also movie-related people. The contributions of this paper are mainly three. First, to the best of our knowledge existing approaches do not address the special issues. So we developed a model for aspect-based opinion mining that specially considers these features. Secondly, as a result of the analysis of the domain, we created a special datasets that help representing the features of the mentioned domain. The rest of this paper is structured in the following manner.

2. Related work

opinion mining or sentiment analysis comprises an area of NLP, computational linguistics and text mining, and refers to a set of techniques that deals with data about opinions and tries to obtain valuable information from them. The aspect-based approach is very popular and many authors have developed their own perspectives and models. Other related approaches are unsupervised topic-based document modeling techniques, which model an input document as a mixture of topics. In this context, our work lies on a radically different paradigm, as the former consists in identifying the aspects reviewed in a piece of text based on a bag-of-words model of the document, rather than extracting individual feature mentions and their related opinions. Therefore, our work is not directly comparable to these kinds of works. Our work acknowledges the differences between domains that is discussed in the paper, and proposing a general model that works for all the domains. Also, our system does not require any training datasets and only a small amount of human support. Finally, one last related topic is the set of so-called concept-level sentiment analysis approaches. These approaches focus on a semantic analysis of text through the use of Web ontologies or semantic networks, which allow the aggregation of conceptual and affective information associated with natural language opinions. Our approach is different from all these applications since it is aspect-based and analyzes opinions at the sentence level.

3. Background

In this section, we proposed our approach in general terms. The opinions are 5-tuples composed of the following parts.

• An entity: Proposed to denote the opinion objective other words. An entity can contain a set of components and attributes and, similarly, each entity component can have its own subcomponents and attributes.
• An aspect: Because it is difficult to study an entity at an arbitrary hierarchy level, this hierarchy is simplified to one or two levels, denoting as aspect every component or attribute of the entity.
• The Sentiment orientation, considering that opinions express a positive or negative sentiment about what they evaluate.
• The Opinion holder, which corresponds to the user that gives the opinion.
• Time: Time and date when the opinion was given. In this manner, opinions are considered to be a positive or negative view, attitude, emotion or appraisal about an entity or an aspect of that entity from an opinion holder in a specific time. The following concepts are also introduced:
• Entity expression: Corresponds to the actual word or phrase written by the user to denote or indicate an entity. As a result, entities
• are then generalizations of every entity expression used in the analyzed documents, or a particular realization of an entity expression.
• Aspect expression: As for an entity expression, the aspect expression is the actual word or phrase written by the user to denote or indicate an aspect. Thus, aspects are also general concepts that comprise every aspect expression.

3.1. Aspect identification

This stage aims to find and extract important topics in the text that will then be used to summarize. In their proposal, part-of- speech (POS) tagging and syntax tree parsing (or chunking) are used to find nouns and noun phrases or NPs. Then, using frequent item set mining, the most frequent nouns and NPs are extracted. The extracted sets of nouns and NPs are then filtered using special linguistic rules. These rules ensure that the terms inside those aspects that are composed of more than one word are likely to represent real objects together and also eliminate redundant aspects. They also extract non-frequent aspects using an approach by finding nouns or NPs that appear near to opinion words with high

frequency. This approach does not extract adjectives or any other kind of non-object aspects.

3.2. Sentiment prediction

The next phase is sentiment prediction, to determine the sentiment orientation on each aspect. This method relies on a sentiment word dictionary that contains a list of positive and negative words (called opinion words) that are used to match terms in the opinionated text. Also, since other special words might also change the orientation, special linguistic rules are proposed. Among others, these rules consider negations words ‘‘no’’ or ‘‘not’’ and also some common negation patterns. However, despite how simple these rules might appear, it is important to handle them with care, because not all occurrences of such rules or word apparitions will always have the same meaning.

Cite this article as: Suganya S, Sureka K, Vishnupriya P. “INTEGRATING ASPECTS BASED ON OPINION MINING FOR PRODUCT REVIEWS.” *International Conference on Information Engineering, Management and Security (2015)*: 01-06. Print.

3.3. Summary generation

The last step is summary generation, to present processed results in a simple manner. In this context, defined opinion quintuples are a good source of information for generating quantitative summaries. In this case, each bar above or below the x-axis can be displayed in two scales: (1) the actual number of positive or negative opinions normalized with the maximal number of opinions on any feature of any product and (2) the percent of positive or negative opinions, showing the comparison in terms of percentages of positive and negative reviews.

4. Proposed extension

Our extension, considers the same set of structured steps mentioned in Section. Here, we discuss issues on each one of the three steps and explain our own approach in the context of product reviews.

4.1. Aspect expression extraction

The aspects do not directly appear in a text but they exist in the manner of aspect expressions. Accordingly, when trying to apply opinion model to extract opinions from real data, concepts can be somewhat confusing or unclear. It is also unclear how aspects that appear more than once in a document are managed. Having noticed these issues, a model to build opinion tuples from an opinionated document has been developed here. We will not extract implicit nor not-frequent aspect expressions.

4.2. Determination of the opinion orientation

Taking the work as inspiration, a set of rules to determine the sentence orientation was developed, always considering opinion words as a basis.

4.2.1. Word orientation rules

In first place, we need to determine the orientation of each word in a sentence. In order to do so, we propose Algorithm 1.The algorithm applies a set of linguistic rules, which are explained below:

Algorithm 1.Word orientation

1: if word is in opinion- words then
2: mark (word)
3: Orientation Apply Opinion Word Rule (marked word)
4: else
5: if word is in neutral_words Then
6: mark (word)
7: orientation -0
8: end if
9: end if
10: if word is near a too_word then
11: orientation Apply Too Rules(orientation)
12: end if
13: if word is near a negation_word then
14: orientation Apply Negation Rules (orientation)
15: end if
16: return orientation

Word rules: Positive opinion words will intrinsically have a score of 1, denoting a normalized positive orientation, while negative ones will have associated a score of 1. Every noun and adjective in each sentence that is not an opinion word will have an intrinsic score of 0 and will be called neutral word.

Negation rules: A negation word or phrase usually reverses the opinion expressed in a sentence. Consequently, opinion words or Neutral words that are affected by negations need to be specially treated.

Too rules: Sentences where words ‘‘too’’, ‘‘excessively’’ or ‘‘overly’’ appear, are also handled specially. When an opinion word or a neutral word appears near one of the mentioned terms, denoted too words, its orientation will always be Negative (score=-1).

4.2.2. Aspect orientation rules

Having mentioned rules that help in determining each word orientation in a sentence, it is now explained how all these orientations should be combined to determine the final orientation of a sentence on a particular aspect. Our proposal is summarized in Algorithm 2 and it only considers words marked as opinion words or neutral words, which we call marked words, as they are the only ones that will provide the orientation for each sentence. The detailed process is explained below.

Algorithm 2: Opinion orientation

1: if but_word is in sentence then
2: orientation Opinion
Orientation (aspect,marked_words,but_clause)

```
if orientation – 0 then
return orientation
else
orientation Opinion Orientation(aspect,marked_words,not but_clause)
if orientation – 0 then
return À1 Â orientation
else
return 0
end if
end if
else
for all aspect_position in aspect do
for all aspect_word in aspect_position do
orientation += suborientation
end for
final_orientation += orientation
end for
if final_orientation > 0 then
return 1
else
if final_orientation < 0 then
return À1
else
return 0
end if
end if
end if
```

4.3 Summarization

This proposal seems fairly, simple and effective for summarizing opinions. However, it lacks a robust way of measuring the importance of each evaluated aspect. Here, we attempt to measure the importance of each aspect simultaneously using the amount of positive and negative opinions of it. We also use that measure to rank aspects. We calculate the standard deviation of these scores using:

$$AVScorei = \frac{Scorei + NScorei}{2}$$

We propose that aspect-based summaries should include bar charts and a table that shows the actual values of PScorei, NScorei and Relative Importance for each aspect expression.

5. System architecture

Two different tasks need to be performed, aspect extraction and orientation determination, for which two submodules are included:

- **Aspect extraction sub-module:** in charge of applying the aspect extraction algorithm to a set of POS-tagged sentences.
- **Orientation determination sub-module:** This sub-module applies the algorithms presented in Section 4 to determine the orientation of an opinion on a given aspect. It also extracts the set of adjectives that appeared near each aspect. Results include the following features:
- **Aspect-based summaries:** Bar charts, in which each bar measures the number of positive and negative mentions of each attribute or component of one product. Bars are initially sorted according to Relative Importance.
- **Adjective bubble charts:** Nearby adjectives in all sentences where an aspect appears are shown in a bubble chart. The size of each bubble counts the times that each adjective is used to describe the aspect.
- **Original opinions**: A list of all original sentences is also displayed in an ad-hoc manner, separating them into positive or negative.

Cite this article as: Suganya S, Sureka K, Vishnupriya P. "INTEGRATING ASPECTS BASED ON OPINION MINING FOR PRODUCT REVIEWS." *International Conference on Information Engineering, Management and Security (2015)*: 01-06. Print.

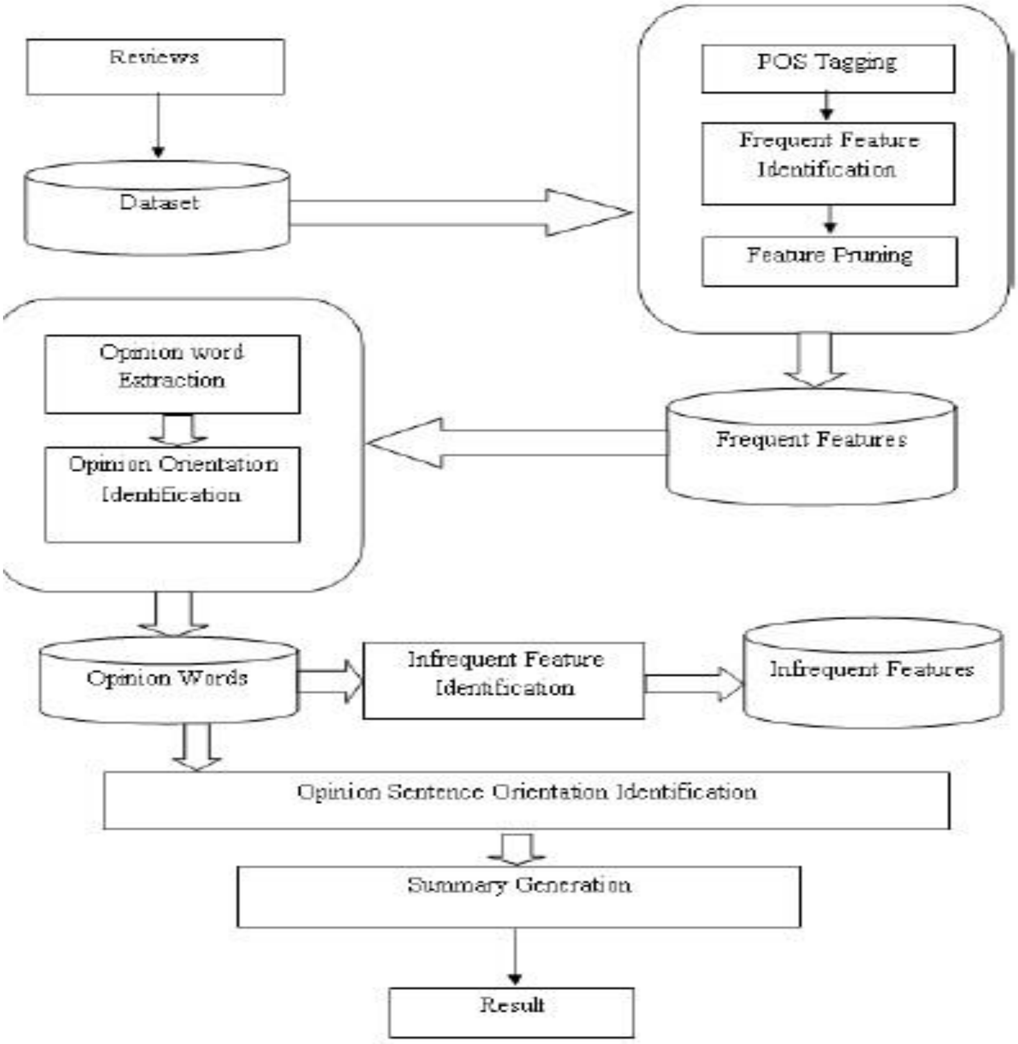


Fig.1: General Design of our system.

7. Conclusions and future work

In this study, we present a generic design of a tourism opinion mining system that aims to be useful in many industries. The core of our system is an extension of aspect-based opinion mining technique. On the one hand, the non-tailored algorithm for aspect expressions extraction, based on frequent nouns and NPs appearing in reviews, achieved a poor performance. This result shows that, in fact, multiple expressions are used to denote the same attribute or component in online product reviews. Therefore, not only the most frequent words need to be considered when extracting aspect expressions in order to achieve a better recall for this task. Our design and models for aspectbased opinion can be used in many possible applications in the online shopping domain. Benefits that may arise entail both merchants and customers.

7.1. Future work

For future work, the primary objective should be to improve Recall on the task of aspect expression extraction, finding infrequent and implicit aspect expressions. On the other hand, we have seen that tourism product reviews contain an important number of sentences that have no opinions. These sentences need to be filtered since they introduce noise to the opinion mining process. This also includes the problem of analyzing context and domaindependent opinions. New methods to determine subjectivity or sentiment orientation need to be tested on the tourism domain in order to improve the performance of these tasks. Future work should also tackle the problem of transforming aspect expressions into aspects. This is a difficult problem yet a crucial feature for any system like ours, because presenting aspect expressions to users implies redundancy and makes the analysis more complex. Here, the objective is to build or use ontologies, hierarchies or clusters of aspect expressions to make the system become easier to navigate and more intuitive for users. Finally, another extension of this work implies working with tourism products reviews written in different languages. Some of the NLP tasks that are used by our system, including sentence and word tokenizers, are generally machine leaning algorithms that need to be properly trained in order to generate good results. The vast availability of data in English to train these models contrasts with a relative scarcity for other languages. Therefore, there is an immense room for future work on this area.

Cite this article as: Suganya S, Sureka K, Vishnupriya P. "INTEGRATING ASPECTS BASED ON OPINION MINING FOR PRODUCT REVIEWS." *International Conference on Information Engineering, Management and Security (2015)*: 01-06. Print.

REFERENCES

[1] Archak, N., Ghose, A., & Ipeirotis, P. (2007). Show me the money!:

[2] Deriving the pricing power of product features by mining consumer reviews. In Proceedings of the 13th ACM

a. *SIGKDD international conference on knowledge discovery and data mining (pp. 56–65). ACM.*

[3] Bollegala, D., Matsuo, Y., Ishizuka, & M. (2007). An integrated approach to measuring semantic similarity between words using information available on the web. In HLT-NAACL (pp.340–347).

[4] Cadilhac, A., Benamara, F., & Aussenac-Gilles, N. (2010). Ontolexical resources for feature based opinion mining: A case-study. In 23rd International conference on computational linguistics (p.77).

[5] Cambria, E. (2013). An introduction to concept-level sentiment analysis. In F. Castro,

a. *Gelbukh, & M. González (Eds.), Advances in soft computing and its applications. Lecture notes in computer science (Vol. 8266,pp. 478–483). Berlin Heidelberg: Springer.*

[6] Cambria, E., Poria, S., Gelbukh, A., & Kwok, K. (2014). A common-sense based api for concept-level sentiment analysis. Making Sense of Microposts, 1(1), 1–2.

[7] Cruz, F. L., Troyano, J. A., Enríquez, F., Ortega, F. J., & Vallejo, C. G. (2013). 'Long autonomy or long delay?' The importance of domain in opinion mining. Expert Systems with Applications, 40(8), 3174–3184. Decker, R., & Trusov, M. (2010). Estimating aggregate consumer preferences from online product reviews. International Journal of Research in Marketing, 27(4), 293–307. Ding, X., Liu, B., & Yu, P.(2008). A holistic lexicon-based approach to opinion mining. In Proceedings of the international conference on Web search and web data mining (pp. 231–240). ACM. Dueñas-Fenández, R., Velásquez, J. D., & LH′ uillier, G. (2014). Detecting trends on the Q3 web: A multidisciplinary approach. Information Fusion 0 (0), in press. Fromkin, V., Rodman, R., & Hyams, N. (2010). An introduction to language. Wadsworth Publishing Company. Havasi, C., Cambria, E., Schuller, B., Liu, B., & Wang, H. (2013a). Knowledge-based approaches to conceptlevel sentiment analysis. Intelligent Systems, 28(2), 12–14. Havasi, C., Cambria, E., Schuller, B., Liu, B., & Wang, H. (2013b). Statistical approaches to concept-level sentiment analysis. Intelligent Systems, 28(3), 6–9. Hu, M., & Liu, B. (2004a). Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining (pp. 168–177). ACM. Hu, M., & Liu, B. (2004b). Mining opinion features in customer reviews. In Proceedings of the national conference on artificial intelligence (pp. 755–760). Menlo Park, CA; Cambridge, MA; London: AAAI Press. MIT Press; 1999. Hu, M., & Liu, B. (2006). Opinion extraction and summarization on the web. Proceedings of the national conference on artificial intelligence (Vol. 21, pp. 1621). Menlo Park, CA; Cambridge, MA; London: AAAI Press. MIT Press; 1999.Kim, H., Ganesan, K., Sondhi, P., & Zhai, C. (2011). Comprehensive review of opinion summarization.

Cite this article as: Suganya S, Sureka K, Vishnupriya P. "INTEGRATING ASPECTS BASED ON OPINION MINING FOR PRODUCT REVIEWS." *International Conference on Information Engineering, Management and Security (2015)*: 01-06. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS002		**eAID**	ICIEMS.2015.002

Security of sensitive data in xml or file system by using Encoding through URL

Kajal Shukla[1], S. K. Singh[1]
[1]M. Tech. Student, Department of CSE VIET, DADRI (G. B. NAGAR) UP- INDIA

Abstract: *Database services have the web applications which are interactive targeted by an SQL Injection. User gives some data as a input and at last that coded input data is being used as to form SQL statement at runtime in these applications. A person who is a n attacker can be able to input a malicious or harmful query segment when user inputs any SQL statement during SQL Injection attacks, that is the result which could be used in many more different database request. Sensitive/Confidential information can be added or modified by an attacker to form attacks of SQL Injection. SQL Injection vulnerability could be used by an attacker as an IP scanner rudimentary. There are several paper published in literature having discussed that how to secure sensitive data in xml or file system , by checking SQL dynamic query commands.SQL Injection attacks However, for secure stored procedures in the database higher level layer / application layer a very less attention is given, which surely can be too suffered from attacks of SQL Injection.*

Keywords: SQL query, SQL server, SQL Injection.

I.INTRODUCTION

One of the most demanding and challenging causes which can make impacts on the business and industry level in a Structured Query Language is that it can explore all of the sensitive information which is stored in our database, including most highly important information such as credit card details, usernames, addresses, passwords, names, phone, email id etc. To inject a Structured Query Language is the liability that when attacker gets the ability with SQL queries which can be passes to a database. The query which is passed through an attacker in to the database, an attacker can allows the query to database which is supporting element with database and our operating system. SQL Query which is able to accept the inputs from the attacker sides can harms our real web application. Attacker always try to insert harmful SQL query commands into a database so as on execution the query they can destroy or alter the database i.e. this technique is called code injection technique. So this attacker is also called attack vector for websites and this kind of attacker is used by any kind of SQL database.

According the study last year, Security Company **"Imperva"** find that the most web application attacks is done 4 times per month and other side retailers company is attacked by 2 times per month. That is not a good practice on the behalf of security.

II. TYPES OF SQL INJECTION :

- ☐ Redirection and reshaping of a query
- ☐ Based on Error message
- ☐ Blind injection

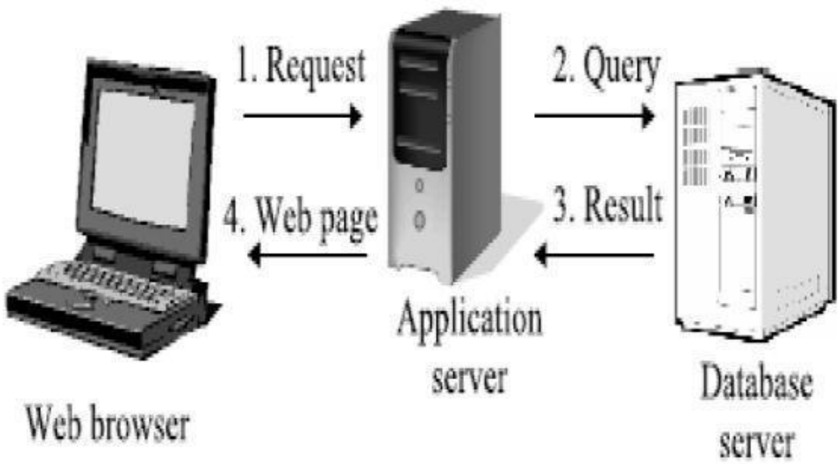
Cite this article as: Kajal Shukla, S. K. Singh. "Security of sensitive data in xml or file system by using Encoding through URL." *International Conference on Information Engineering, Management and Security (2015)*: 07-11. Print.

i. Blind SQL Injection

• Formation of queries that results in Boolean values, and interprets output of HTML Pages is provided by Blind SQL injection technique IN databse.
• Final result of SQL injection gives significant data theft and/or attacks in data modification.
• Essentially Blind attack playing 20 questions with the web server.

ii. Focus on Blind SQL Injection

• This type of SQL Injection is as common as any other type of injection.
• An incorrect or wrong sense of security on the host is provided by the Blind holes.
• Requires a larger of time investment to properly execute manual penetration against.



iii. Concepts of SQL Injection Attacks:

(a) SQL injection attack is a process to find the query which is entered through the user for execution the command.
(b) SQL attackers create crafted manually input data so that SQL interpreter has to accept the query and give the permission to execute the commands and give his desired results.
(c) SQL Injection attack breaks the security of the database layer. When attacker breaks flaws through the SQL injection then attackers can drop, modify, create, and alter our sensitive database.

III. SECURITY IN SQL INJECTIONS:

Web vulnerabilities minimum 20% of all that is being related to SQL Injection, called as the one of the most widespread type of catalog application security and as well as the subsequent most common software susceptibility which have the find and prevent capability .SQL Injection always must be on high priority for web developer and also for security basis. Generally a SQL Injection assault diminishes any web network application software which has not provided a proper validation or we can say coded by a user given input data. In the last phase that crafted input data is being used as an element of query over again whichever back-end database. Acquire an example, what time we generate any form it always asks for the ID that is called as identity. After that URL:"http://www.anywebsite.com/id/id.asp?id=anymanualdata" is created.
An invader, using the SQL Injection may perhaps go through any data or "1=1". At the particular time if the application software of the web network is not specified proper validation or incorrectly encoded the user given data that is directly send in the direction of database, and as well as input with the vulnerable query will reach there in reply that will depict every single one ids in the database ever since the condition is "1=1" is for all time true. The example given is an indispensable example however it illustrates the consequence of sanitizing client input data prior to using it in a SQL database query or SQL commands.

IV. LITERATURE REVIEW:

Our web applications allow the visitors to enter or submit or retrieve database using any web browser through the internet. These kinds of data have to centralize therefore they be capable of storing data which is needed for websites. If any Suppliers, Employees, a host of stakeholders, customers etc. want to achieve specific content from database side then he can receive it. Company statistics, User Details, financial, economical and payment information etc are stored in our database which is access through custom web applications. Our Database and Web applications allow us to run the production business frequently.

Cite this article as: Kajal Shukla, S. K. Singh. "Security of sensitive data in xml or file system by using Encoding through URL." *International Conference on Information Engineering, Management and Security (2015)*: 07-11. Print.

The Process which attempt to get ahead of commands or statements of SQL intended for implementation through the database over the web network application is called hacking technique in SQL Injection. If their attempts are right then our database allows hackers to view their desire information from the database and he can hampers our database, and be able to do the whole lot which he wants.

For example Like feedback forms , Shopping carts , Search pages , product and support request forms, figure current websites , provide businesses and login pages etc pages are very necessary to commune with customers for keep our customer in touch. These kinds of pages of websites are very to use customer. These types are pages are suspicious for SQL hackers and foremost they attempts to try on these pages .We cannot hide this category of pages on website. If we do it then our client cannot be handling with us. So hacking the website is becoming very easy task for Hackers.

For Simple Example
To access the catalog database , normal user would input their username and password to come into their profile and access his personal details and change the contents which is allow by the administrative section i.e. authenticate user are allowed to access our database. In other sides, our web network application which controls the authentication page will foremost communicate with our database through the specific planned commands as a result they be able to filter that he is authenticate user or not. In the case of valid user, database allows to access the contents.
In other sides, In case of SQL Injection, Specifically craft SQL commands with the intent of passing the login form difficulty is inputted by the hacker. In case of SQL Injection vulnerabilities, Hackers are eligible to converse with our database directly. Script languages which are Dynamic like JSP, PHP, and ASP.NET, CGI etc are the target technologies by the hacker. For publicity, our website wishes to be communal public so our safety mechanism will agree to to be communal public with our application (generally at beyond port 80/443).
SELECT count (*)
FROM person_list_table
WHERE
username=**'FIELD_USERNAME'** AND password=**'FIELD_PASSWORD"**
This SQL command is given instruction to the catalog database to compare User Id and secret code (password) filled by the current user to the combination that it has already stored in its database. Each and every web network application is hardly coded with specified SQL query so as to it will implement when executing functions and communicating with the database. If any data input of web network application is not accurately encoded, a hacker possibly will introduce extra vulnerable SQL queries which enlarge the area of SQL commands.
An attacker will therefore have a plain channel of communiqué to the web application database irrespective of the entire intrusion uncovering systems vulnerability and network based security equipments installed on the database layer.

V. SQL INJECTION IMPACT:

When a hacker feels that a organism is ready to SQL Injection attacks, he is now able to insert SQL Commands to the n/w database an input from field. This is similarly like as to say attacker comes to make changes in our catalog and allow him to do insert or delete like DROP in to database. Execution of illogical SQL queries on the susceptible structure may be done by an attacker. This may break the reliability of your secure information . It depends on the back-end database, SQL injection vulnerabilities can be lead to varying levels of data/system access to the attacker. Manipulate in any existing queries, to UNION that is used to select related information from two tables use sub-selects arbitrary data, or append additional queries.
Some of the SQL Servers like Microsoft SQL Server contains stored and extended procedures for database server functions. In certain cases, it can be possible to read in or write out in files, and can execute shell commands on the underlying operating system. Data is being stolen through the various attacks at all the time. Hackers rarely get caught which are more expert.
Any attacker that can obtain access, it could spell disaster. A SQL injection attacks involves the modification of SQL statements that are used in a web application through the use of attacker-input data. Unfortunately the harm of SQL Injection is only found when the theft is discovered. Improper validation and improper construction and incorrect input of SQL statements in web applications can lead them theft to SQL injection attacks. Thus SQL injection is a potentially destructive and prevalent attack that the Open Web Application Security Project (OWASP) listed it as the number one threat to web applications.

VI. PROPOSED SOLUTION

SQL injection can helps to retrieve sensitive information like password or credit card details, to prevent SQL injection developer should has to take some measure steps like use session in place of query string to transfer value from one page to another. Store sensitive information like password or credit card to XML or file system which is not easily accessible. If using Query
String is necessary try using URL Encoding technique. Now a day's some DBMS like MS SQL server supports
regular expression validation which protect data insertion like " ' ". All DBMS doesn't support """handle it is very
necessary replace it with some other character.

Blindfolded SQL Injection techniques:

(a) Boolean queries and WAIT FOR DELAY are used by Blind folded injection technique.
(b) By using commands such as BETWEEN, LIKE, IS NULL Comparison in queries is done.

Cite this article as: Kajal Shukla, S. K. Singh. "Security of sensitive data in xml or file system by using Encoding through URL." *International Conference on Information Engineering, Management and Security (2015)*: 07-11. Print.

IDS signature evasive SQL Injection techniques:
(a) By using CONVERT & CAST commands by masking the attack payload.

(b) By using Null bytes to break the signature patterns

(c) By using HEX encoding mixtures.

(d) By using SQL CHAR ( ) to convert ASCII values as numbers.

Example, when the attacker decided to go with a attack using: 1 = 1, at that time when it is entered as input box. The server recognizes 1 = 1 as a true statement and -- symbol is used for comment, everything after that is ignored making it possible to the attacker to access to the database. Through this SQL injection example page you can see precisely how this attack works on:

Welcome to SQL Injection Application:
Logged in as: or '1=1--'AND Password='
Other sample pages:
BadProductList- Product List that is vulnerable to SQL Injection.
BetterProductList- Product List that is still vulnerable but that uses a lower privilege account to minimize damage.
EncryptCnxString- Utility for encrypting any string: use it to encrypt cnxWindBest connection string in web config.
AddSecureUser- Add new users to Secure User table: Password will be hashed use it with BestLoginaspx.
PRODUCT LIST:
Product **Filter: 'UPDATE** Products SET Unit Price=0.0
WHERE Product=Set Filter

Product Id	Product Name	Quantity Per Unit	Unit Price
1.	pen	10 boxes*20 bags	0.0000
2.	Alcohol	24-12 oz bottles	19.0000
3.	paper	36 boxes	21.3500
4.	Aniseed Syrup	12-550 ml bottles	0.0111
5.	Seasoning	48-6 oz jars	22.0000
6.	Jelly	12-8 oz jars	25.0000
7.	Uncle Bob's Organic Pears	12-1 ib pkgs.	30.0000
8	Cranberry Sauce	12-12 oz jars	40.0000
9.	Jam	18-500 g pkgs.	97.0000
10.	Pickle	12-200 ml jars	31.0000

OUR ALGORITHM STEPS OF URL ENCODING ARE:

```
string strCnx = ConfigurationSettings.AppSettings["cnxNWindBad"];
SqlConnection cnx = new SqlConnection(strCnx);
cnx.Open();
string strQry = "SELECT Count(*) FROM Users WHERE UserName='" + txtUser.Text + "' AND Password='"
+txtPassword.Text + "'";
int intRecs;
SqlCommand cmd = new SqlCommand(strQry, cnx); cmd.CommandType= CommandType.Text;
intRecs = (int) cmd.ExecuteScalar();
```

Cite this article as: Kajal Shukla, S. K. Singh. "Security of sensitive data in xml or file system by using Encoding through URL." *International Conference on Information Engineering, Management and Security (2015)*: 07-11. Print.

```
if (intRecs>0)
{
FormsAuthentication.RedirectFromLoginPage(txtUser.T ext, false);
}
else
{
lblMsg.Text = "Login attempt failed.";
}
cnx.Close();
//Prevention string strCnx =
ConfigurationSettings.AppSettings["cnxNWindBetter"];
using(SqlConnection cnx = new SqlConnection(strCnx))
{
cnx.Open(); SqlCommand cmd = new
SqlCommand("procVerifyUser", cnx); cmd.CommandType= CommandType.StoredProcedure;
SqlParameter prm = new SqlParameter("@username",SqlDbType.VarChar,50);
prm.Direction=ParameterDirection.Input;
prm.Value = txtUser.Text;
cmd.Parameters.Add(prm);
prm = new
SqlParameter("@password",SqlDbType.VarChar,50);
prm.Direction=ParameterDirection.Input;
prm.Value = txtPassword.Text;
cmd.Parameters.Add(prm);
string strAccessLevel = (string) cmd.ExecuteScalar();
if (strAccessLevel.Length>0)
{
FormsAuthentication.RedirectFromLoginPage(txtUser.T ext, false);
}
else
{
lblMsg.Text = "Login attempt failed.";
}
}
```

VIII. CONCLUSION

SQL attackers create crafted input data so that SQL interpreter has to accept the query and give the permission to execute the commands and give his desired results. SQL Injection attack breaks the security in the Database layer and can alter, steal or destroy our database through using web application.

REFERENCES

[1] Ke Wei, M. Muthuprasanna, Suraj Kothari , Dept. of Electrical and Computer Engineering , Iowa State University Ames, IA – 50011 ,Email: {weike,muthu,kothari}@iastate.edu

[2] Cerrudo. Manipulating Microsoft sql server using sql injection.

[3] http://www.appsecinc.com/presentations/Mani pulating SQL Server Using SQL Injection.pdf, White Paper.

[4] William G.J. Halfond, Jeremy Viegas, and Alessandro Orso College of Computing Georgia Institute of Technology {whalfond|jeremyv|orso}@cc.gatech.edu

[5] Z. Su and G. Wassermann. The Essence of Command Injection Attacks in Web Applications. In The 33rd Annual Symposium on Principles of Programming Languages (POPL 2006), Jan. 2006.

[6] F. Valeur, D. Mutz, and G. Vigna. A Learning-Based Approach to the Detection of SQL Attacks. In Proceedings of the Conference on Detection of Intrusions and Malware and Vulnerability Assessment (DIMVA), Vienna, Austria, July 2005.

[7] T. M. D. Network. Request.servervariables collection. Technical report, Microsoft Corporation, 2005. http://msdn.microsoft.com/library/default. asp?url=/library/en-us/iissdk/html/9768ecfe-8280-4407-b9c0-844f75508752.asp.

[8] José Fonseca CISUC - Polithecnic Institute of Guarda, Marco Vieira, Henrique Madeira DEI/CISUC - University of Coimbra. Testing and comparing web vulnerability scanning tools for SQL injection and XSS attacks. Retrieved July 10, 2007, from http://ieeexplore.ieee.org

[9] Yuji Kosuga, Kenji Kono, Miyuki Hanaoka Department of Information and Computer Science Keio University. Sania: Syntactic and Semantic Analysis for Automated Testing against SQL Injection. Retrieved November 12, 2007, from IEEE Computer Society. http://ieeexplore.ieee.org

[10] Benjamin Livshits and U´ lfar Erlingsson. Microsoft Research. Using Web Application Construction Frameworks to Protect Against Code Injection Attacks. Retrieved June 14, 2007, from http://ieeexplore.ieee.org

Cite this article as: Kajal Shukla, S. K. Singh. “Security of sensitive data in xml or file system by using Encoding through URL.” *International Conference on Information Engineering, Management and Security (2015)*: 07-11. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS003		**eAID**	ICIEMS.2015.003

ADVANCED LOCKER SECURITY SYSTEM

Prof R.Srinivasan, T.Mettilda, D.Surendhran, K.Gopinath, P.Sathishkumar
Guided by Assistant Professor, Department of Electronics and Instrumentation Engg.
UG Students of Dept. of Electronics and Instrumentation Engineering,
K.S.Rangasamy College of Technology, (Autonomous) - Trichengode.

Abstract: *The purpose of this paper is to provide a secured locker security system based on RFID, PASSWORD, CONVEYER and GSM technology which can be organized in bank, secured offices and homes. This system allows authentic person only can be recovered money from locker. The implemented locker security system based on RFID, PASSWORD and GSM technology containing automatic movement of lockers system which can be easily activate, authenticate, and validate the user in real time for secured locker access. The RFID, PASSWORD, GSM and HEAT SENSOR provides the advantage of high security than other systems. In general terms, RFID is an object or person identifier using a radio frequency transmission. In electronic terms RFID is an electronic method of exchanging data over radio frequency waves. With RFID technology we can identify, sort, track or detect variety of objects.*

***Keywords*:** RFID, GSM, Conveyer, Microcontroller, Heat sensor.

I. INTRODUCTION

The main purpose of this paper is to implement a locker system with high security based on RFID, PASSWORD, CONVEYER, GSM and HEAT SENSOR technology which can be organised in banks, offices and other places where high security is required. In this only authorized person can open the locker. The initial security levels are RFID verification and PASSWORD. The After this security verification the details of the person will provided to the security in charge like manager, after that conformation CONVEYER setup will bring only the appropriate locker from the locker to the person. The GSM server send the random password to the customer mobile. The locker can be accessed if the password matches. Otherwise the alarm is on. In addition to this, the heat sensor can access the alarm when anyone try to open the locker by using electrical machine which produce heat.

II. EXISTING SCENARIOS:

The locker systems involve manual lock in most of the banks. Whenever the user uses the locker, user should be assisted by the bank employee. It leads to waste of time for both the customer and the employee. Lack of security and the waiting time of the customers are the major drawbacks of such manual lock systems. The person accompanying the customer can be any employee who is free at that instant of time it should be noted. Hence, time is wasted. This can be overcome with the automatic locker system. There are many techniques in which the proposed technology can be implemented. The RFID tags are used in this project which holds the user's information like locker number, username, etc, in the existing project RFID tag read by the RFID reader will automatically open and close the locker. Hence, security is guaranteed and the customers waiting time is reduced.

Cite this article as: Prof R Srinivasan, T Mettilda, D Surendhran, K Gopinath, P Sathishkumar. "ADVANCED LOCKER SECURITY SYSTEM." *International Conference on Information Engineering, Management and Security (2015)*: 12-16. Print.

III. PROPOSED METHOD

In this proposed method after the password verification for the RFID tag the details of the customer will provided to the manager. The manager authentication Selects the locker and moves it to the opening with the help of the stepper motor. The locker will have keypad for password. By GSM technology the customer receives the random password provided by the server. The locker can be accessed if the password matches otherwise the alarm rang. To avoid theft by using electrical gadgets to break the locker the heat sensor is provided to detect the heat while breaking with alarm.

IV. RFID FUNDAMENTALS

RFID is an effective automatic identification technology for variety of objects and person. The most important functionality of RFID is to track the location of the tagged item. The RFID tags can be classified into three major categories which is based on power source, active tags, passive tags, and semi-passive (semi-active) tags. An active tag contains both a radi transmitter, receiver and a battery that is used to power the transceiver. Active tags are more powerful than the passive tags/semi-passive tags. RFID tags can also be classified into two categories: tags with read/write memory, and tags with read-only memory. The tags with read/write memory are more expensive than the tags with read-only memory. RFID tags operate in three frequency ranges: low frequency (LF, 30–500 kHz), high frequency (HF, 10–15MHz), and ultrahigh frequency (UHF, 850–950MHz, 2.4–2.5GHz, 5.8GHz). The LF tags are less affected by the presence of fluids or metals when compared to the higher frequency tags. RFID reader is shown in th fig 1. The most important functionality of RFID is the ability to track the location of the tagged item. Typical applications of HF tags are access control and smart cards. RFID smart cards, working at 13.56MHz, are the most commonly used tags.

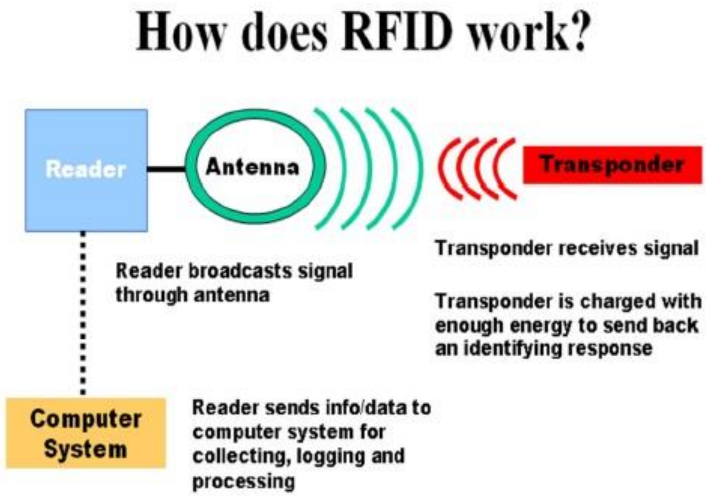


Fig 1 RFID Reader

However, UHF tags are severely affected by fluids and metals. UHF tags are more expensive than any other tag. The typical frequency of UHF tags are 868MHz (Europe), 915MHz (USA), 950MHz (Japan), and 2.45GHz The active tag enables higher signal strength and extends communication range up to 100-200m.

V. GSM

GSM (Global System for Mobile communications) is the technology that underpins most of the world's wireless mobile phone networks. GSM is a digital cellular and an open technology used for transmitting mobile voice and data services. GSM operates in the 900MHz to 1.8GHz bands. The supported data transfer speed of GSM is up to 9.6kbps. It allows the transmission of basic data services such as SMS. In the current work, GSM module SIM300 is used, it is shown in figure.2. The SIM300 module is a Triband GSM/GPRS solution in a compact plug in module featuring an industry-standard interface.

Features of GSM

• Single supply voltage 3.2v-4.5v
• Typical power consumption in SLEEP
Mode: 2.5mA.
• SIM300 tri-band
• MT, MO, CB, text and PDU mode, SMS
storage SIM card
• Supported SIM Card: 1.8V, 3V

Cite this article as: Prof R Srinivasan, T Mettilda, D Surendhran, K Gopinath, P Sathishkumar. "ADVANCED LOCKER SECURITY SYSTEM." *International Conference on Information Engineering, Management and Security (2015)*: 12-16. Print.

Figure 2 GSM Modem

VI. STEPPER MOTOR

A stepper motor (or step motor) is a brushless DC electric motor that divides a full rotation into a number of equal steps. The position of the motor can then be commanded to move and hold at one of these steps without any feedback sensor (an open-loop controller), as long as the stepper motor is carefully sized to the appropriate application. In this project the stepper motor is used to move the locker towards the opening in the room and bring it back to the original position with accuracy

Figure 3 Stepper Motor

VII. KEYPAD

The keypad is used to get the password from the customer in two different situations. Initially the RFID tag requires the password. Then the server requires the password to open the locker. In this the 4*4 matrix keypad is used. Since the passwords are four digit random numbers.

Fig 4 Keypad

Cite this article as: Prof R Srinivasan, T Mettilda, D Surendhran, K Gopinath, P Sathishkumar. "ADVANCED LOCKER SECURITY SYSTEM." *International Conference on Information Engineering, Management and Security (2015)*: 12-16. Print.

VIII. LCD DISPLAY

LCD stands for liquid crystal this is an output device with a limited viewing angle. The LCD is mostly preferred as an output device because of its cost of use and is better with alphabets when compared with a 7-segment LED display. Now a days we have so many kinds of LCD and our application requires a LCD with 2 lines, each line consist of 16 characters, the LCD receives data from the microcontroller and displays the same. It has 8 data lines and 3 control line. LCD has a supply voltage Vcc (+5v) and a GND. This low voltage supply makes the whole device user friendly by showing the balance left in the card. It also shows the card that is currently being used.

IX. MICROCONTROLLER

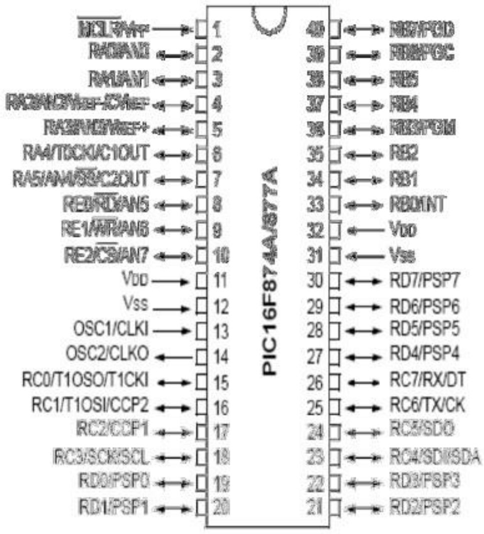


Fig 5 PIC16f874A

The security options are controlled by the microcontroller. The operating voltage is 2.0-5.5V with low power consumption. It is fully static design. The operating speed is 20MHZ. This microcontroller is 40 pins dual in line package. It has three timers with high speed. When compared to others it has high efficiency.

X. TEMPERATURE SENSOR

The LM35 series are precision integratedcircuit temperature sensors, the output voltage of the sensor is linearly proportional to the Celsius (Centigrade) temperature. The LM35 has an advantage over linear temperature sensors calibrated in § Kelvin, as the user is not required to subtract a large constant voltage from its output to obtain convenient Centigrade scaling.

Features

• Calibrated directly in § Celsius (Centigrade)
• Linear a 10.0 mV/§C scale factor
• 0.5§C accuracy guarantee able (at a25§C)
• Rated for full b55§ to a150§C range
• Suitable for remote applications
• Low cost due to wafer-level trimming
• Operates from 4 to 30 volts
• Less than 60 mA current drain
• Low self-heating, 0.08§C in still air
• Nonlinearity only g(/4§C typical
• Low impedance output, 0.1 X for 1mA load

XI. BLOCK DIAGRAM

In the given block diagram, the controller of this arrangement is microcontroller PIC16F874A. The initial security levels are controlled by the computer. The keypad reads the password entered. Then the RFID tag is swiped. The RFID reads the customer details if the password is correct. Otherwise it will not allow the process of opening. The computer verification sends the result to the microcontroller.

Cite this article as: Prof R Srinivasan, T Mettilda, D Surendhran, K Gopinath, P Sathishkumar. "ADVANCED LOCKER SECURITY SYSTEM." *International Conference on Information Engineering, Management and Security (2015)*: 12-16. Print.

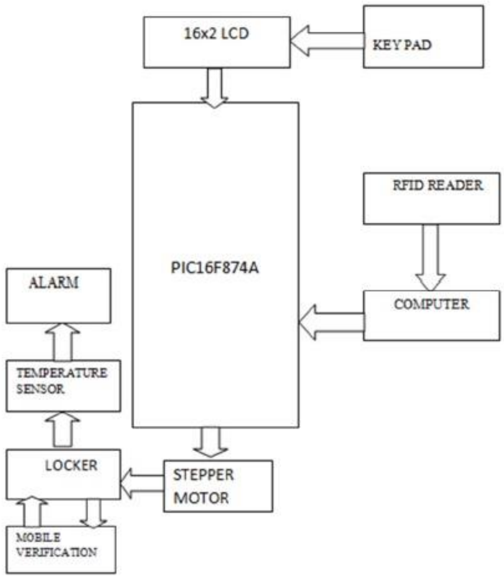


Fig 6 Block diagram

If the security proceedings is authorised by the manager the stepper motor bring the appropriate locker to the opening. Then the server will generate the random password. It received in the customer mobile phone as a message with the help of GSM technology. If the password matches the locker will open. To avoid breakages with welding equipments the heat sensor block is provided. If the heat is high enough to melt the metal then the alarm goes on.

XII. CONCLUSION

The implemented project provides a locker system with RFID, password verification, GSM technology. It provides more security facilities. In this the future extension can be made by adding the Digital Image Processing for face recognition. It will ensure high security.

REFERENCES

[1] ChetnaKoli.R, Nikita.S, Kheratkar.V, Pooja.S, Ganganalli.T &Shirsa.G, 2014 'Bank Locker System Using Iris'International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 3, Issue 3

[2] Raghu Ram.Gangi, Subhramanya Sarma & Gollapudi, 2013 'Locker Opening And Closing System Using RFID, Fingerprint, Password' International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 2, Issue

[3] BalajiVenkatesh.A.M, KarthikKalkura & Shriraam A.C, 2013 'Student Locker Protection Using RFID Tag & Reader' International Journal of Engineering and Advanced Technology (IJEAT) ,Volume- 3, Issue-2

[4] Swetha.J, 2011, 'RFID Based Automated Bank Locker System', Internationa Journal of Research in Engineering and Technology, Volume-1

[5] Bramhe, 2011 'SMS Based Secure Mobile Banking', International Journal of Engineering and Technology, Volume-3

[6] Joshua Bapu.J & sirkaziMohdArif, 2013 'Locker security system using RFID and GSM technology' International Journal on Advances in Engineering and Technology, Volume 3

[7] Ramesh.S, SoundariaHariharan & ShruthiArora, 2012 'Monitoring and Controlling of Bank Security System', International Journal of Advanced Research in Computer Science and

[8] Software Engineering, Volume-2

[9] Manoj V, Bramhe 'SMS Based Secure Mobile Banking Department' International Journal of Engineering and Technology Vol.3 (6),

[10] AbhsheksShukla, ShrutiTyagi & Shweta Gupta,2013 'An Ultimate Security to Bank Lockers Using Multi-Model Biometric Systems' International Journal of Research Review in Engineering Science &Technology (IJRREST) Volume-2, Issue-2

[11] Aruna.D.Mane& SirkaziMohdArif, 2012 'Locker Security System Using RFID and GSM Technology' International journal of advances in engineering & technology, volume-1

Cite this article as: Prof R Srinivasan, T Mettilda, D Surendhran, K Gopinath, P Sathishkumar. "ADVANCED LOCKER SECURITY SYSTEM." *International Conference on Information Engineering, Management and Security (2015)*: 12-16. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS004		**eAID**	ICIEMS.2015.004

Retrieving Information for Urgency Medical Services using Abundant Data Processing Method based on IoT

V.Balaji[1], D.Dinagaran[2]
[1]UG Scholar, Department of CSE, vbbalajir@gmail.com
[2]Assistant Professor, ddinagaran@gmail.com
IFET COLLEGE OF ENGINEERING, Villupuram-TN, India

***Abstract*:** *The Internet of Things (IoT) is the interconnection of uniquely identifiable embedded computing devices within the existing internet infrastructure. Delivering clinical information of patient at the point-of-care to physicians is critical to increase the quality of healthcare services, especially in emergency time. However, clinical data are distributed in different hospitals. It is sometimes difficult to collect clinical data of patient ubiquitously in case of urgency. In order to support the ubiquitous content accessing a resource model is first proposed to locate and get clinical data which are stored in heterogeneous hospital information systems using Hadoop Distributed File System. In the proposed method clinical data of patient is defined as resource with unique URL address. Related clinical data of one patient is collected together to form a combinational resource, and could be accessed by physician if authority is assigned to the physician, by using a mongo dB database technique efficiently in big data applications for better performance and scalability. This type of database support faster execution of queries compared to non-relational databases. By implementing the system that combines IoT with Big Data is built to provide quick and effective for different patients.*

Keywords: Big Data, Decision support system (DSS), Internet of things (IoT), Resource model.

I. INTRODUCTION

In Recent years, Healthcare faces n-number problems, including high and rising expenditures, inconsistent quality of data and gaps in care and access data. For this reason, healthcare services represent a major position of government spending more money in most countries [1]. The amount of healthcare data in the world has been increasing enormously, and to analyses these large data set referred to as Big Data becomes a key basis of competition makes an innovation for productivity growth, new ideas and consumer surplus [2]. But the Big data means the data sets whose size is vast when compared to the ability of current technology method and theory to capture, manage, and process these data within an endurable lapsed time. Today, Big Data management provides viewpoints as a challenge for all IT companies. The solution to such a challenge is shifting increasingly from providing hardware to provisioning more manageable software solutions [3]. Big data also brings new opportunities and critical challenges to industry and academia [4] [5]. In Internet of things (Iot) technologies is present enormous potential for the high-quality and more convenient healthcare servicing. By employing these technologies in the activities of healthcare servicing doctors are able to access different kinds of data resources online quickly and easily, helping to make emergency medical decisions, and reducing costs in the process[6] . Open and distributed file

Cite this article as: V.Balaji, D.Dinagaran. "Retrieving Information for Urgency Medical Services using Abundant Data Processing Method based on IoT." *International Conference on Information Engineering, Management and Security (2015)*: 17-22. Print.

system, clinical decision support system will be a technical architecture takes full advantage of Electronic Health Record (EHR), patient databases, domain expert knowledge bases like decision support system, available technologies and standards to provide efficient decision-making support for healthcare professionals [7].

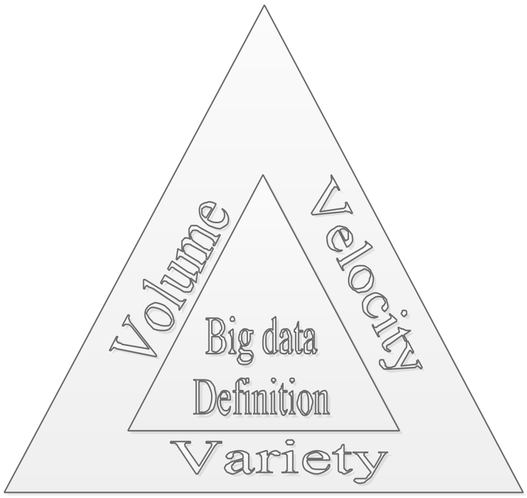


Figure 1Big data definition

II. LITRATURE SERVEY

JayavardhanaGubbi, RajkumarBuyya, [8] proposed system is to deploy large-scale, platform-independent, wireless sensor network infrastructure that includes data management and processing, actuation and analytics. It is often quite important to take advantage of the benefits of metadata for transferring the information from a database to the user via the Internet.

Marisol García-Valls, Pablo Basanta-Val, [9] proposed were the level of hardware, networked embedded systems (NES) are becoming a cloud of hundreds, and even thousands, of heterogeneous nodes connected by means of heterogeneous networks as well; they are now used in various domains such as cloud or grid.

Keling DA, Marc DALMAU, Philippe ROOSE, [10] proposed system is context collector first collects information on the operating environment from the operating system and the user context.

Li Da Xu, [11] proposed is to be properly managed, the integration of KM and ERP becomes a strategic initiative for providing competitive advantages to enterprises. ERP III enables ES applications to transform an enterprise into a knowledge-based learning organization and to capture know-how for developing business solutions**.**

Olugbara, Mathew, O. Adigun, [12] proposed system in these paper can enable improved quality of healthcare by eliminating variation and dichotomy in healthcare services and misuse of healthcare resources.

Boyi Xu, Li Da Xu, [13] proposed system is it based on emergency system.it will be collected the information about the patient and stored in a could.it will access by distributed system.It can be access the patient data in a emergency time.by using IoT it more flexibly to provide at time of emergency medical services.It can be support the data accessing in mobile computing platform. .

III. DISADVANTAGES OF EXISTING SYSTEM

1) No importance to decision making
2) Information are handled within their local system administrator.it will not support for heterogeneous formats.
3) It is on unified data model and semantic data explanation by ontology in data storage (Sql) and accessing.
4) Health Book contains patient's medical information like previous medical histories.

IV. PROPOSED SYSTEM

The proposed system, the clinical data of patient is defined as resource with unique URL address. It is access the patient information by the thumb finger authentication it useful for the emergency situations. The clinical data of one patient is collected together to form a combinational resource in cloud, and could be accessed by physician if authority is assigned to the physician, by using a mongo dB database technique efficiently in big data applications(by hdfs) for better performance and scalability. This type of database support faster execution of queries compared to non-relational databases. By implementing the system that combines IoT with Big Data is built to provide quick and effective for different patients

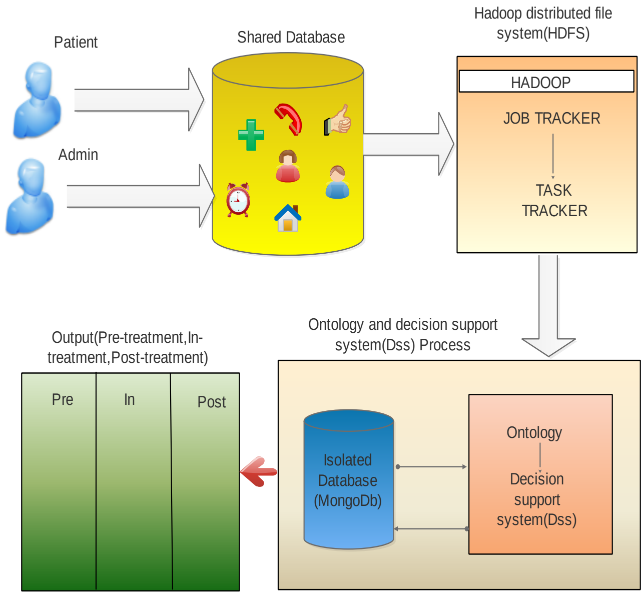


Figure 2 Proposed System Architecture

A. Authentication

Authenticity of the Patient is the main issue in now day's internet of things such as distributed file system. The Password has been the most share to provide the patients information. Hash code is generated for the patients in unique id and it is stored in the hospitals used for authentication mechanism which is subjected to online attacks (it is provides from online hacker).. To give solution for this problem one of the process using for authentication is BIOMETRIC based Cryptography scheme to address the authentication issues. This methodology proposes the finger print image which is obtained from the user is Steganographed with PIN NUMBER of the user and the Steganographed image which in turn is divided into two shares. One share is stored in the hospitals database and the database. One Time Password(OTP) is used every time to ensure the trusted submission of shares. The system not only ensures the secured transaction of process but also verifies the true identity of the person through one time password. The patients present the share during all of his/her transactions after entering the OTP. When the patients present his share the hash code is generated and compared with the database value. If it matches, the shares are databases to get the original Steganographed image. Again, the Desteganography process is carried on to obtain the original finger print image and the PIN NUMBER. The user is allowed to proceed further only after this authentication. This process ensures proper security scheme[18].

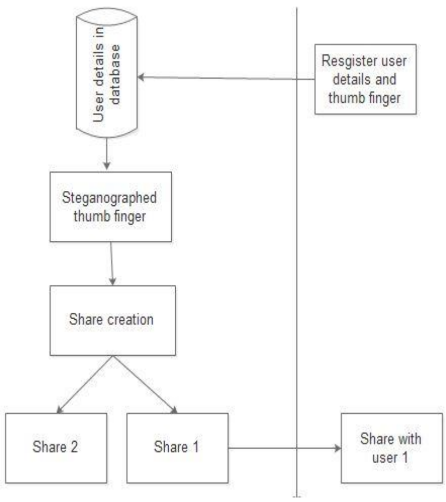


Figure 3 Authentication thumb finge

B. Hadoop distributed file system(hdfs)

The HDFS (Hadoop distributed file system) is a open-source platform. It is as some advantages more scalable and reliable. The Hadoop frameworks will allow for the distributed processing of large (million) number data sets to clusters from computers using small programming models. It is use full to scale up from one server to ten thousands of machines. Another advantages in hadoop distributed file system the hardware to deliver high-availability, The Hadoop has two components such as and MapReduce. HDFS used for data storage and Map Reduce for data processing. HDFS will "just work "under a variety of physical and systemic circumstances. By

Cite this article as: V.Balaji, D.Dinagaran. "Retrieving Information for Urgency Medical Services using Abundant Data Processing Method based on IoT." *International Conference on Information Engineering, Management and Security (2015)*: 17-22. Print.

distributing storage is useful for computation across multi servers, the combined storage resource can reduce the size of the server and will very efficiency at every size.

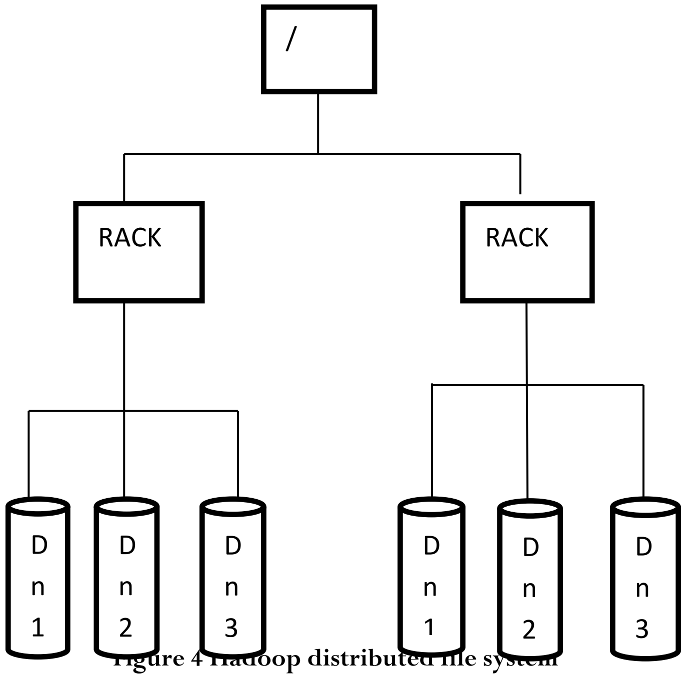


Figure 4 Hadoop distributed file system

C. Map reduce

HDFS was designed to be a scalable, fault-tolerant, distributed storage system that works closely with Map Reduce. Map Reduce is used to execute the Mongodb query and provide the parallel processing over a large number of nodes to simplify the data[17]. Finally, the Mangodb query language is created for the graph and the data is retrieved from the HDFS.

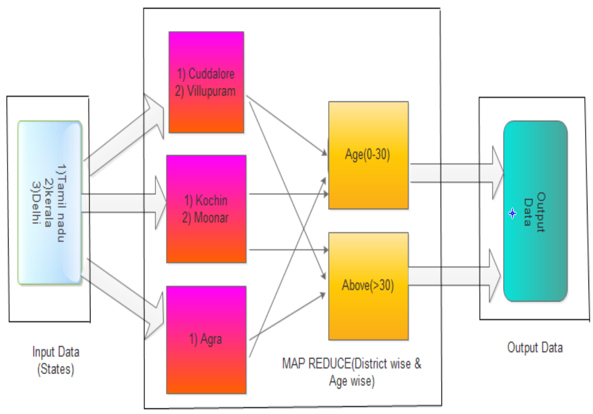


Figure 5 Map Reduce

D. Resource description framework(RDF)

Semantic Web is based on RDF, which integrates a variety of applications by using extensible markup language (XML) for syntax and universal resource identifier (URI) for naming. RDF [20] is an assertional language intended to be used to express propositions via precise formal vocabularies. An RDF data model is similar to conceptual modeling approaches, as it is based on the idea of making statements about resources.

The fundamental unit of RDF is a triple that is used to describe the relationship between two things. Its formal definition is <subject, predicate, object>, in which subject denotes a resource, and predicate denotes properties or aspects the resource and expresses in relationship between the resource and the object [19].

Cite this article as: V.Balaji, D.Dinagaran. "Retrieving Information for Urgency Medical Services using Abundant Data Processing Method based on IoT." *International Conference on Information Engineering, Management and Security (2015)*: 17-22. Print.

E. Decision support system(DSS)

Data + Analysis = Decision Support

A clinical decision-support system is any computer program designed to help health professionals make clinical decisions. In a sense, any computer system that deals with clinical data or medical knowledge is intended to provide decision support. Three types of decision-support function, ranging from generalized to patient specific

E1. Benefits of using the Dss

1) Time savings: The time savings that have been documented from using computerized decision support are often substantial [16].
2) Cost reduction: DSS cost saving from labor savings in making decisions and from lower infrastructure or technology costs.
3) Allows for faster decision-making.
4) Provides more evidence in support of a decision.

V. Advantages in proposed system

1) Full advantage it is the available Internet technology. Information is transferred from a database to the user via internet.
2) Large amount of data can be gathered, access time of data very less.

VI. Experimental Setup

Our experiments use the Windows XP operating system with Intel processor, 4-Gbyte RAM with a clock speed as 1.8 GHz. The capacity of the Hard disk drive is 1TB. The tools and database such as Hadoop 0.18.10 and Mongodb are installed in the system. Our approaches are implemented in Java language with the version JDK 1.7 and running in eclipse-SDK-3.3.1.1.

VII. Data Model For IoT Urgency Medicial Serviecs

A healthcare service is a dynamic process that includes the pre-treatment, in-treatment, post-treatment are shown in fig.6.

VIII. Conclusion

Innovative uses of IoT technology in healthcare not only bring benefits to hospitals (doctors and managers) to access wide ranges of data sources but also challenges in access heterogeneous IoT data, especially at real time IoT application systems. The big data and mongodb accumulated by IoT devices creates the easy for the IoT data accessing is fast, easy and quickly with more efficiency.it will reduce the time complexity at emergency services.

REFERENCES

[1] Shaker H. "A distributed clinical decision support system architecture support system" 2014

[2] J. Manyika et al., "Big Data: The Next Frontier for Innovation, Competition, and Productivity," 2011.

[3] C. Lynch, "Big Data: How Do Your Data Grow?" Nature, vol. 455,no. 7209, pp. 28-29, 2008.

Cite this article as: V.Balaji, D.Dinagaran. "Retrieving Information for Urgency Medical Services using Abundant Data Processing Method based on IoT." *International Conference on Information Engineering, Management and Security (2015)*: 17-22. Print.

[4] F. Chang, J. Dean, S. Ghemawat, and W.C. Hsieh, "Big table: A Distributed Storage System for Structured Data," ACM Trans. Computer Systems, vol. 26, no. 2, article 4, 2008.

[5] W. Dou, X. Zhang, J. Liu, and J. Chen, "Hire Some -II: Towards Privacy-Aware Cross-Cloud Service Composition for Big Data Applications," IEEE Trans. Parallel and Distributed Systems, 2013.

[6] R. Burke, "Hybrid Recommender Systems: Survey and Experiments," User Modeling and User-Adapted Interaction, vol. 12,no. 4, pp. 331-370, 2002.

[7] Boyi Xu, Li Da Xu . "Ubiquitous Data Accessing Method in IoT-Based Information System for Emergency Medical Services"2014

[8] Jayavardhana Gubbi, Rajkumar Buyya,b "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions" (2012).

[9] Marisol García-Valls, Pablo Basanta-Val, "A bi-dimensional QoS model for SOA and real-time middleware" (2011).

[10] Keling DA, Marc DALMAU."A Survey of adaptation systems" (2012).

Cite this article as: V.Balaji, D.Dinagaran. "Retrieving Information for Urgency Medical Services using Abundant Data Processing Method based on IoT." *International Conference on Information Engineering, Management and Security (2015)*: 17-22. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS005

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.005

FOREST FIRE PREDICTION AND ALERT SYSTEM USING BIG DATA TECHNOLOGY

Mr.T.Rajasekaran[1], J.Sruthi[2], S.Revathi[2], N.Raveena[2]
[1]Assistant Professor[1]
[2]DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
KPR INSTITUTE OF ENGINEERING AND TECHNOLOGY

Abstract: *In this paper we discuss about the forest fire prediction and alert system using big data technology. Forest fire is considered as one of the major natural disaster. Our method is to collect and analyze the data from wireless sensor using hadoop tool to predict the forest fire before it occurs. Here we use machine learning tool called as Mahout which is used for clustering and filtering the datasets and it can be able to predict the valid output. By using GSM we can give alert message to people so that they can relocate to a safety place immediately, when fire occurs. Signal and Infrare image processing is used to monitor the signals and images of the entire forest for every 30 min and those data will be stored in datasets, by using those data we can be able to predict the forest fire in advance.*

Keyword: Fire prediction, wireless sensor, hadoop, temperature sensor, Mahout, GSM, Signal proceesing, Infrared image processing.

INTRODUCTION

Several million acres of forest are destroyed every year due to forest fire. Forest fire not only destroys many valuable trees but also destroys the vegetation in that area. 90% of the forest fire occurs due to humans. **"Crown Fires"** are spread quickly by wind moving across the tops of trees. **"Running crown fires"** are more dangerous because they burnt extremely hot, travel rapidly, and can change direction quickly. Lightning strikes the earth over 100000 times a day. 10 to 20% of these lightning strikes can caus fire. Forest fire is one of the major causes of global warming as tones of greenhouse gases are emitted into the atmosphere. Nowadays the detection mechanisms are used for watching towers, satellite imaging, long distance video recording, etc. But it will not provide any quicker response which is most important in forest fire detection. Video surveillance is a low cost system but it produces false alarm due to environmental condition like fog, clouds, dust and human activities. Another method is used to take snapshot of the forest by using visual camera, and it will be placed on the towers to cover the maximum area of the forest. A motor is used to rotate the camera 360° so that we get the full view of the forest. The images obtained using these cameras are processed by using a program or a code.
These images are used to find forest fire by comparing it with the normal images. The major advantage of this method is that the system can be programmed to take into considerations of the environmental conditions and the effect of fog or clouds that can be eliminated The serious disadvantage is that it may sometime do not predict the fire considering the signals are due to environmental conditions. We also need to build towers to place the camera at a higher position so that it may increase the cost of the system. A good and effective method is the use of wireless sensor network. In this method the sensor module is deployed in the forest manually or through a helicopter. The sensor module consists of multiple sensors like temperature sensor, humidity sensor, etc. They collect the target environment information and continuously transfer it to the control center where the necessary process is carried out. Sensor nodes are less costly and even if it gets damaged in fire it won't be a great loss. WSN has the property of self configuration and hence

Cite this article as: Ast.proff. T Rajasekaran, J Sruthi, S Revathi, N Raveena. "FOREST FIRE PREDICTION AND ALERT SYSTEM USING BIG DATA TECHNOLOGY." *International Conference on Information Engineering, Management and Security (2015)*: 23-26. Print.

need not be organized manually. **GPS** can be used to track the exact location of the fire that can be easily obtained and the nearest fire service can also be easily informed by using **GSM.**

II. METHODOLOGY

Big data is used to store huge amount of data that can be analyzed later. The data will be in the form of structured, unstructured or semi structured. Big data uses **Hadoop** tool for storing huge amount of data. In forest fire prediction, the data will be collected and stored in hadoop as unstructured form. **Hadoop** ecosystem components are pig, hive, map reduce, HDFS. **Map reduce** is a programming model, and an associated implementation for processing and generating large data sets with parallel and distributed algorithm on a cluster. In forest fire prediction Map Reduce plays an important role in forest fire prediction because it is used to reduce the large datasets into simpler datasets. **Hive** is the data warehouse infrastructure built on top of hadoop for providing data, query, analysis. Apache hive supports analysis of large datasets stored in hadoop's **HDFS**. I provides an SQL like language called **HiveQL** and it is use to convert data into map reduce. To accelerate queries in hadoop it provides indexes and bit map indexes. We use hadoop tool to store the data of forest fire prediction for analyzing**. Hive** is used to analyze the datasets of forest fire that are stored in hadoop **HDFS**. We can use machine learning tool to collect the temperature, rare trees, weather, gas, etc. it is used to store the scalability of large datasets. **Mahout** tool is one of the machine learning tool used in hadoop ecosystem. It's OS is independent**. Mahout** tool is used to filter and classify the datasets based on keyword. Cameras are used for monitoring the entire forest**. Signal processing and Infrared image processing** is also used to monitor the signal and image of the entire forest. By using the signal processing and multi sensor we can get an alert message from the server and can be able to predict the fire in advance.

III. PROPOSED FIRE DETECTION MECHANISM

The proposed method consists of varieties of standalone boxes, and each box consistingof various sensors like humidity and temperature sensors. These boxes are spread around the entire forest area, so that we can be able to monitor the entire forest area.

3.1. Sensor Deployment

Sensor deployment is one of the most significant factor as it determines the efficiency of the entire system

1. The entire forest with minimum number of nodes should be covered by the sensor
2. The rate of spread of fire can be calculated easily, only if the distance between the sensor are equal
3. The sensors must be positioned such that false alarms are avoide These sensors collect the data wirelessly and transmit the data to a base station. The sensors form a cluster and are active always. They sense the parameters every 15 minutes and if there is a possibility of fire detected then the parameters will be measured every 2 minutes. This purpose is to reduce the usage of battery power. These sensors cannot be powered using electricity so they need to be deployed deep into the forest. Solar panels are used for powering the rechargeable batteries.

3.2 Topology Design

Based on the density of trees in the area, the topology of the sensor nodes must be preplanned. When the density of trees is more then there are more chances of fire as the trees more often rub each other producing heat due to friction. In such cases the number of sensors deployed must be higher. While considering the energy restriction the detection of forest fire as early as possible must not be compromised.

IV. MATERIAS USED

4.1 Temperature Sensor

One of the main changes that happen when a fire occurs is the increase in temperature of the environment. This might be considered as the cause of forest fire or due to change in temperature during summer. Due to forest fire the change in temperature can be differentiated from other environmental factors as the rate of change of temperature due to fire will be rapid. Here we use **LM35** is as fire sensor and this can be able to measure the temperature only in the range of **-55°C to 150°C**

4.2 Humidity Sensor

By measuring humidity We can detect and predict fire greatly. When a fire occurs the air becomes dry and the humidity will be low. And there is a maximum possibility of occurrence of fire when the air is dry than being moistures.

4.3 Battery

The battery used for this project must be rechargeable, small, light, cheap, environmental friendly, fast in charging and discharging, reliable, long lasting, etc. Not all these are satisfied in one battery but **Liion** battery seems to suit this purpose.

4.4 GSM

Global System For Mobile Communication is a digital mobile telephony system that is widely used in every parts of the world. GSM is used to send alert messages to the neighbor areas quickly. It is the most widely used in three digital wireless telephony technologies **(TDMA**, **GSM, and CDMA).** GSM digitizes and compresses data, then sends it down a channel with two other streams of user data, each in its own time slot. It operates at either the900 MHz or 1800 MHz frequency band. GSM is used to send status about the occurrence of fire in the forest. GSM is interfaced to the microcontroller through **RS 232 to USART** terminals.

4.5 Zig Bee

Zigbee is a specification for communication in a wireless personal area network (WPAN). Zigbee is based on an **IEEE 802.15** standard. It consumes low power with transmission distance of 10 to 100 meters line of sight. It can transmit data over long distance through intermediate devices such as by forming mesh network. Zigbee has a defined rate of 250 Kbit/s, and best suited for intermittent data transmissions from a sensor or input device. It is simple to use and much less expensive than other

WPANs such as **Bluetooth and Wi-Fi.**

V. ALGORITHM

1. All the nodes should be initialized and synchronized to same clock
2. A cluster of nodes will be connected to a base station, and all the base station are connected to the control center
3. When humidity of air is high**, LM 35** senses the temperature and transmit it to the base station every 30 minutes
4. When the humidity of the air reduces there is more possibility for the fire hence the rate of measurement will be increased to every 15 minutes
5. If the temperature is less than the threshold value then the node enters the sleep state else the sensor continuously senses the temperature and transmits the result to the base station
6. When a node senses fire it sends a danger packet to its neighboring nodes and the timer is started and it will run till it gets a fire alert. This is to calculate the rate of spread of fire and the direction of spread
7. The base station collects all the values and calculates the rate and direction of spread of fire
8. Through the GSM, alert messages is sent to nearby villages to relocate the people to a safe locality This is a simple method where we have a less overhead in the data packets, and this topology is easy to expand. The energy consumption is also less as the node senses the parameters only on certain intervals which are controlled by the base station

VI. WORKING OF FIRE DETECTOR

6.1Prediction of fire

It is necessary for us to detect the fire as early as possible and it would be better if it is predicted in advance. The fire usually occurs when the humidity of the air is lower and the temperature is higher. Thus if the humidity of the air is below a threshold value and the temperature is higher than the threshold value then an alert signal is sent to the control center After the alert signal is given to control signal we can relocate our area or else we can predict ourselves from the fire. Once the fire is predicted at a particular location then the necessary precautionary measures are carried out. The fire may occur even without being predicted. This prediction will work only when the fire arises due to increase in the relative temperature but when the fire occurs due to incidents such as lightning or manmade events or due to crown fires then the fire cannot be predicted.

6.2 Detection of fire

When the temperature in a particular node gets increased over a fixed threshold value then the alert is sent to the control center. The threshold value will always be fixed above the maximum temperature which is experienced in that particular region to avoid any false alarm due to the increase in the atmospheric temperature. As soon as the fire is detected in a particular node the alert will be sent to the control center and also to the neighboring nodes. Once the nearer nodes get the alert, timer gets started and it is run till the nearer node detects the fire. This is used to find the rate of spread of the fire in the forest. When the rate of spread is known then the necessary action can be taken immediately. All the nodes are equally spaced in order to easily find the rate of spread of fire. The rate of spread directly depends on the speed of air blowing. And also the fire usually spreads upwards in a hilly area. These are taken into considerations while designing the detection system.

6.3 Finding the direction and rate of spread of fire

The direction of spread of fire is more important to prevent further damage to the forest and wildlife. This can be obtained by using the data collected from the sensor nodes. Normally the fire spreads in all the directions hence when a fire is detected in a node then it sends danger alert packets to all the neighboring nodes and all the neighboring nodes start a timer and measure the time between the reception of the alert packets and the detection of fire. This is done for all the neighboring nodes.

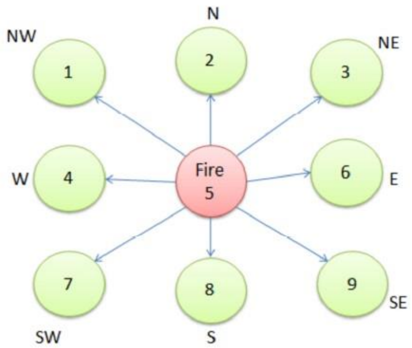


Figure 6.3: **Fire detected node sending alert packets to neighboring node**

Cite this article as: Ast.proff. T Rajasekaran, J Sruthi, S Revathi, N Raveena. "FOREST FIRE PREDICTION AND ALERT SYSTEM USING BIG DATA TECHNOLOGY." *International Conference on Information Engineering, Management and Security (2015)*: 23-26. Print.

In the above method the middle node 5 detects the fire first and it sends alert packets to all eight neighbors. If any one of the node interrupts it won't affect the remaining nodes, because we can provide secondary node for all nine nodes. Backup should be available in the secondary node and it will be analyzed for every half an hour. Hence the rate of spread of fire in all the eight directions can be found.

$$\text{Rate of spread of fire} = \frac{\text{Distance between two nodes}}{\text{Time interval between reception of alert and fire detection}}$$

VII. CONCLUSION

The objective of this paper is to reduce the damage and destruction that are caused by the forest fire to the life and property of humans and also wild animals. Apart from early detection of forest fire we have also attempted to predict the fire in advance with the help of the data obtained from the sensors that are deployed in the forest.

REFERENCE

[1] T. Akyildiz, W. Su, E. Cayirci, and Y. Sankarasubramaniam, "A survey on sensor networks," in proc. of IEEE Communications Magazine, vo1.40, no. 8, pp. 102-114, Aug.2002.

[2] C. F. Garcia Hernandez, P. H. Ibargliengoytia-Gonzalez, J. , J.A. Perez Diaz and Garcia Hernandez," Wireless Sensor Networks and Applications a Survey," International Journal of Computer Science and Network Security, vo!.7 no.3, Mar. 2007.

[3] A. Mainwaring, D. Culler, R. Szewczyk, and J. Anderson. "Wireless sensor networks for habitat monitoring", In Proc. of the First International Workshop on Wireless Sensor Networks and Applications (WSNA'02), pages 88-97, Atlanta, Georgia, Sep. 2002.

[4] Ajay V. ,Chetan B., Ganesh S., Kumaran S., "Sensor Nodes Approach to Forest fire Prediction and Prevention," Stony Brook University, 2009.

[5] Stipaničev, D., et al., "Forest Fire Protection by Advanced Video Detection System- Croatian Experiences" Split, Croatia, 2006.

[6] D. M. Doolin and N. Sitar, "Wireless Sensor Nodes for Wildfire Monitoring", SPIE Symposium on Smart Structures and Materials, San Diego, pp. 477484, 2006.

[7] D. Morvan, M. Larini, J. L. Dupuy, A. 1. Iranda, J. Andre, O. Sero-Guillaume, P. Cuinas, "Behavior Modeling of Wild land Fires: A State of the Art", EUFIRELAB: Euro-Mediterranean Wild land Fire Laboratory, 2002.

[8] Wikipedia, the free encyclopaedia, RS- 232, [Online], Available at: http://en.wikipedia.org/wiki/RS-232

Cite this article as: Ast.proff. T Rajasekaran, J Sruthi, S Revathi, N Raveena. "FOREST FIRE PREDICTION AND ALERT SYSTEM USING BIG DATA TECHNOLOGY." *International Conference on Information Engineering, Management and Security (2015)*: 23-26. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	VOL	01
Website	www.iciems.in	eMail	iciems@asdf.res.in
Received	10 - July - 2015	Accepted	31- July - 2015
Article ID	ICIEMS006	eAID	ICIEMS.2015.006

EFFICIENT ROUTING AND FALSE NODE DETECTION IN MANET

Siva Ranjani[1], Ranjani N[1], Sri Abarna[1], Ananthi K[1], Parkavvi[1]
[1]Thiagarajar College Of Engineering

INTRODUCTION

In MANET there is no fixed communication infrastructure. Each node is free to move in an arbitrary manner. Hence it is necessary for nodes to maintain updated position information with the immediate neighbor. Also there will be frequent changes in the topology of the mobile nodes in MANET. In geographic routing, the destination node and the node in the forwarding path can be mobile. In such case it is necessary to reduce the effects caused by the changing topology, which is a difficult task in geographic routing to reconstruct the network topology in presence of changing topology. To obtain the location of node's neighbor, each node exchanges its location information with its neighbor by periodic broadcasting of beacons. This periodic beaconing is not fair in terms of update cost collision of, packet delivery ratio and may lead to collision of data packet with beacon packet .To overcome this drawback, in this paper we propose an efficient beacon scheme called GPSR(Greedy Perimeter Stateless Routing Protocol) which dynamically adjust the frequency for beacon update based on nodes mobility. GPSR comprises two rules : The first rule, referred as Mobility Prediction (MP),which is used to significantly reduce the frequency of beacon overhead. The second rule, referred as On-Demand Learning (ODL), aims at improving the accuracy of local topology among the communicating nodes. Certain nodes considering their limited resources, mainly energy do not forward the data packet to its successive node although they are considered as active nodes in the neighbor list configuration. These nodes are identified as false nodes or selfish nodes and they are removed from the neighbor list and an alternate path is chosen to forward the packet. In this paper, we propose to reduce the beacon packet overhead and identify the false node in MANET.

2 LITERATURE SURVEY:

We gone through some of the literatures and acquired knowledge for choosing technique for efficient routing.

1)" False Node Detection Algorithm in Cluster Based MANET"- Mobile Ad hoc network are collection of mobile nodes that can dynamically form temporary networks, it is necessary to bring the smart technologies in the Ad hoc network environment. Huge amount of time and resources are wasted while travelling due to traffic congestion. The idea behind clustering is to group the network nodes into a number of overlapping clusters. In the clusters of MANET the resource constraints leads to a big problem as decrease in performance and the network partitioning leads to poor data accessibility due to false and selfish node. In our proposal the MANET area has been split into a number of size clusters having cluster head and storage capability according to connectivity degree, RSS (relative signal strength) as per the cluster formation algorithm given. In this cluster architecture they try to find false node inside clusters of MANET using a modified algorithm and try to remove them. Inside the cluster one node that manages the cluster activities is cluster head. Inside the cluster, there are ordinary nodes also that have direct access only to this one cluster head, and gateway. Gateways are nodes that can hear two or more cluster heads. Ordinary nodes send the packets to their cluster head that either distributes the packets inside the cluster, or (if the destination is outside the cluster) forwards them to a gateway node to be delivered

Cite this article as: Siva Ranjani, Ranjani N, Sri Abarna, Ananthi K, Parkavvi. "EFFICIENT ROUTING AND FALSE NODE DETECTION IN MANET." *International Conference on Information Engineering, Management and Security (2015)*: 27-42. Print.

to the other clusters. Several nodes will be take part in the MANET for data forwarding and data packets transmission between source and destination.

They must forward the traffic which other nodes sent to it. Among all the nodes some nodes will behave selfishly, these nodes are called selfish nodes. In our paper we called selfish node as false node. Selfish nodes only to cooperate partially , or not at all, with other nodes. These selfish nodes could then reduce the overall data accessibility in the network. Selfish nodes use the network for their own communication, but simply decline to cooperate in forwarding packets for other nodes in order to save battery power. In the clusters of MANET the false nodes leads to a big problem as increase congestion. The idea behind splitting MANET into a number of size clusters having cluster head and storage capability as per the cluster formation algorithm given .But cluster formation is very difficult in MANET.

2)" Adaptive Position Update for Geographic Routing in Mobile Ad-hoc Networks"-In geographic routing, nodes need to maintain up-to-date positions of their immediate neighbors for making effective forwarding decisions. Periodic broadcasting of beacon packets that contain the geographic location coordinates of the nodes is a popular method used by most geographic routing protocols to maintain neighbor positions. We contend and demonstrate that periodic beaconing regardless of the node mobility and traffic patterns in the network is not attractive from both update cost and routing performance points of view. We propose the *A*daptive Position Update (APU) strategy for geographic routing, which dynamically adjusts the frequency of position updates based on the mobility dynamics of the nodes and the forwarding patterns in the network. APU is based on two simple principles: (i) nodes whose movements are harder to predict update their positions more frequently (and vice versa), and (ii) nodes closer to forwarding paths update their positions more frequently (and vice versa). Our theoretical analysis, which is validated by NS2 simulations of a well known geographic routing protocol, Greedy Perimeter Stateless Routing Protocol (GPSR), shows that APU can significantly reduce the update cost and improve the routing performance in terms of packet delivery ratio and average end-to-end delay in comparison with periodic beaconing and other recently proposed updating schemes. The benefits of APU are further confirmed by undertaking evaluations in realistic network scenarios, which account for localization error, realistic radio propagation and sparse network.

3) "EAACK-A Secure Intrusion Detection System for MANET"- The migration to wireless network from wired network has been a global trend in the past few decades. The open medium and wide distribution of nodes make MANET vulnerable to malicious attackers. A new technique EAACK (Enhanced Adaptive Acknowledgement) method designed for MANET was proposed for intrusion detection. EAACK demonstrates higher malicious-behavior-detection rates in certain circumstances while does not greatly affect the network performances. MANET is vulnerable to various types of attacks because of open infrastructure, dynamic network topology, lack of central administration and limited batterybased energy of mobile nodes. But most of these schemes become worthless when the malicious nodes already entered the network or some nodes in the network are compromised by attacker. Such attacks are more dangerous as these are initiated from inside the network. Routing protocols are generally necessary for maintaining effective communication between distinct nodes. Routing protocol not only discovers network topology but also built the route for forwarding data packets and dynamically maintains routes between any pair of communicating nodes. Routing protocols are designed to adapt frequent changes in the network due to mobility of nodes. MANET is capable of creating a self-configuring and self-maintaining network without the help of a centralized infrastructure, which is often infeasible in critical mission applications like military conflict or emergency recovery.

3.1 PROBLEM DEFINITION:

The problem with AODV(Ad-hoc Ondemand Distance Vector Routing) is that there is route setup latency when a new route is needed, because AODV queues data packets while discovering new routes and the queued packets are sent out only when new routes are found. This situation causes throughput loss in high mobility scenarios, because the packets get dropped quickly due to unstable route selection. Similarly,periodic beaconing used in AODV is not suitable for all nodes. Adaptive Position Strategy(APU) can be used to overcome this.

3.2 PROBLEM DESCRIPTION:

In geographic routing, nodes need to maintain up-to-date positions of their immediate neighbors for making effective forwarding decisions. Periodic broadcasting of beacon packets that contain the geographic location coordinates of the nodes is a popular method used by most geographic routing protocols to maintain neighbor positions. To demonstrate the periodic beaconing regardless of the node mobility and traffic patterns in the network is not attractive from both update cost and routing performance points of view. Adaptive Position Update (APU) strategy for geographic routing, which dynamically adjusts the frequency of position updates based on the mobility dynamics of the nodes and the forwarding patterns in the network. APU is based on two simple principles: (i) nodes whose movements are harder to predict update their positions more frequently (and vice versa), and (ii) nodes closer to forwarding paths update their positions more frequently (and vice versa). A poorly adjusted rate of beacon transmissions may lead to vast resource usage (power and bandwidth) on one side, or may lead to poor throughput on the other side. We use a general model without assuming a particular mobility model. The model is instantiated for periodic and exponential beaconing and it is then applied to compare two-way beaconing with one-way beaconing. The disadvantage of this protocol is it is not scalable in large networks and it does not support asymmetric links. Periodic beaconing consumes network bandwidth,increase update cost,end-to-end delay. Thus Packet delivery ratio will get decreased.Beacon packets traffic will be overhead for data packets and most of the data packets will be dropped. Average end-to-end delay is more in periodic beaconing,because neighbor list is updated periodically not based on mobility of nodes.False nodes in the routing path affects routing performance. These nodes do not forward data packets in order to save their energy. Alternate path for forwarding should be chosen. The unreachability of even a small fraction of destinations on static networks because of the failure of the no-crossing heuristic is also problematic; such routing failures are permanent, not transitory. The power of greedy forwarding to route using only neighbor nodes' positions comes with one attendant drawback: there are topologies in which the only route to a destination requires a packet move temporarily farther in geometric distance from the destination. In Distance Source Vector (DSV) routing, by caching the negative information, the link may get broken, this cause the problem in the system. Source routes in use may be automatically shortened if one or more intermediate hops in the route become no longer necessary.

4.1 GPSR PROTOCOL:

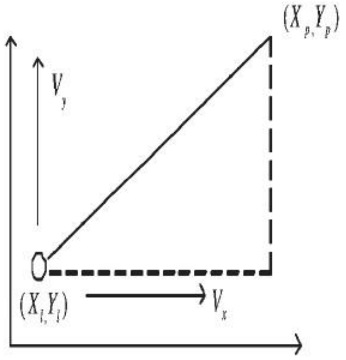
Greedy Perimeter Stateless Routing (GPSR), a novel routing protocol for wireless datagram networks that uses the positions of routers and a packet's destination to make packet forwarding decisions. Geographic routing is also called georouting or position based routing which is a routing principle that relies on geographic position information. It is mainly proposed for wireless networks and based on the idea that the source sends a message to the geographic location of the destination instead of using the network address. The idea of using position information in the area of packet radio networks and interconnection networks. Geographic routing requires that each node can determine its own location and that the source is aware of the location of the destination. With this information a message can be routed to the destination without knowledge of the network topology or a prior route discovery. GPSR makes greedy forwarding decisions using only information about a router's immediate neighbors in the network topology. When a packet reaches a region where greedy forwarding is impossible, the algorithm recovers by routing around the perimeter of the region. GPSR scales better in per router state than shortest path and adhoc routing protocols as the number of network destinations increases. GPSR can use local topology information to find correct new routes quickly. However, in situations where nodes are mobiles or when nodes often switch off or on ,the local topology rarely remain static. Hence, its necessary that each node broadcasts its updated location information to all of its neighbors.These location updated packets are usually referred as a beacons.Greedy perimeter stateless routing protocol shows that APU can significantly reduce the update cost and improve the routing performance in terms of packet delivery ratio and average end to end delay in comparison with periodic beaconing and other recently proposed updating schemes.GPSR protocol use extensive simulation of mobile wireless networks to compare its performance with Dynamic Source Routing. In networks of wireless stations, communication between source and destination nodes may require traversal of multiple hops, as radio ranges are finite. A community of adhoc network researchers has proposed, implemented, and measured a variety of routing algorithms for such networks. The observation that topology changes more rapidly on a mobile, wireless network than on wired networks, where Link State Protocol is used. In a linkstate protocol, the only information passed between the nodes is information used to construct the connectivity maps. GPSR benefits from geographic routing use of only immediateneighbor information in forwarding decision. GPSR allows nodes to figure out who its closest neighbors are also close to the destination the information is supposed to travel.

4.2 MOBILITY PREDICTION RULE :

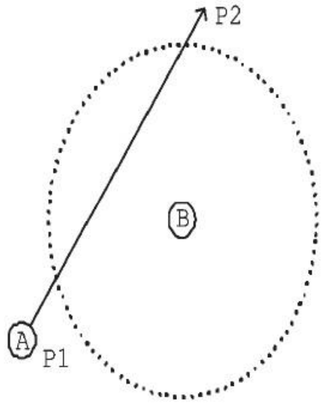
A Mobile Ad hoc NETwork (MANET) is a collection of wireless mobile nodes forming a network without using any existing infrastructure. All mobile nodes function as mobile routers that discover and maintain routes to other mobile nodes of the network and therefore, can be connected dynamically in an arbitrary manner. The mobility attribute of MANETs is a very significant one.

The mobile nodes may follow different mobility patterns that may affect connectivity, and in turn protocol mechanisms and performance. Mobility prediction may positively affect the service-oriented aspects as well as the application-oriented aspects of ad hoc networking. At the network level, accurate node mobility prediction may be critical to tasks such as call admission control, reservation of network resources, pre-configuration of services and QoS provisioning. At the application level, user mobility prediction in combination with user's profile may provide the user with enhanced location-based wireless services, such as route guidance, local traffic information and on-line advertising. In this chapter we present the most important mobility prediction schemes for MANETs in the literature, focusing on their main design principles and characteristics. This rule adapts the beacon generation rate to the frequency with which the nodes change the characteristics that govern their motion (velocity and heading).The motion characteristics are included in the beacons broadcast to a node's neighbors. The neighbors can then track the node's motion using simple linear motion equations. Nodes that frequently change their motion need to frequently update their neighbors, since their locations are changing dynamically. On the contrary, nodes which move slowly do not need to send frequent updates. A periodic beacon update policy cannot satisfy both these requirements simultaneously, since a small update interval will be wasteful for slow nodes, whereas a larger update interval will lead to inaccurate position information for the highly mobile nodes. The MP rule, thus, tries to maximize the effective duration of each beacon, by broadcasting a beacon only when the predicted position information based on the previous beacon becomes inaccurate. This extends the effective duration of the beacon for nodes with low mobility, thus reducing the number of beacons. Further,highly mobile nodes can broadcast frequent beacons to ensure that their neighbors are aware of the rapidly changing topology.

This rule adapts the beacon generation rate to the mobility of nodes. Nodes which contains highly mobile need to frequently update their neighbors since their locations are changing dynamically. At the same time, nodes which move slowly do not need to send frequent updates. This MP rule adapts the beacon generation rate to the frequency with which the nodes change the characteristics that govern their motion (velocity). The motion characteristics are included in the beacons broadcast to a node's neighbors. The neighbors can then track the node's motion using simple linear motion equations. Nodes that frequently change their motion need to frequently update their neighbors, since their locations are changing dynamically. Nodes which move slowly do not need to send frequent updates. A periodic beacon update policy cannot satisfy both these requirements simultaneously, since a small update interval will be wasteful for slow nodes, whereas a larger update interval will lead to inaccurate position information for the highly mobile nodes.



4.2.1 Example of MP Rule

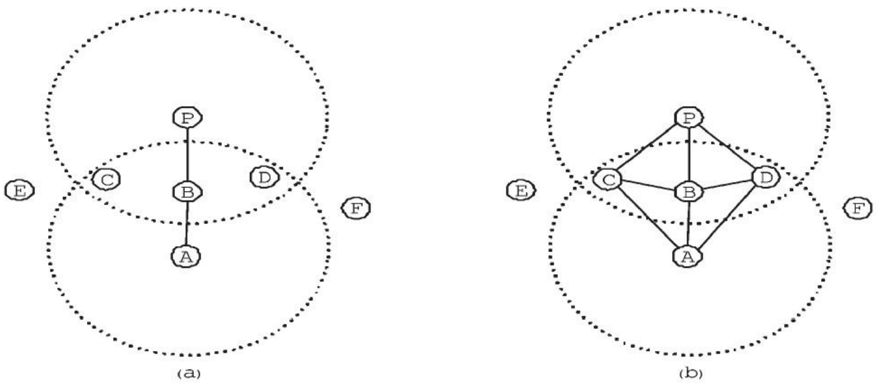


4.2.2 Drawback of MP Rule

Variables	Definition
(Xl,Yl)	The coordinate of node i at time Tl (included in the previous beacon)
(Vx,Vy)	The velocity of node i along the direction of the x and y axes at time Tl (included in the previous beacon)
Tl,Tc	The time of the last beacon broadcast and current time
(Xp,Yp)	The predicted position of node i at the current time

4.3 ON DEMAND LEARNING RULE:

A node broadcasts beacons response to data forwarding activities that occur in the vicinity of that node. Whenever a node overhears a data transmission from a new neighbor, it broadcasts a beacon as a response, it implies a neighbour who is not contained in the neighbor list of this node. A node waits for a small random time interval before responding with the beacon to prevent collisions with other beacons. The location updates are piggybacked on the data packets and that all nodes operate in the promiscuous mode, which allows them to overhear all data packets transmitted in their vicinity. Since the data packet contains the location of the final destination, any node that overhears a data packet also checks its current location and determines if the destination is within its transmission range. According to this rule,whenever a node overhears a data transmission from a new neighbor, it broadcasts a beacon as a response. By a new neighbor, we imply a neighbor who is not contained in the neighbor list of this node. In reality, a node waits for a small random time interval before responding with the beacon to prevent collisions with other beacons.Recall that, we have assumed that the location updates are piggybacked on the data packets and that all nodes operate in the promiscuous mode, which allows them to overhear all data packets transmitte in their vicinity.In addition, since the data packet contains the location of the final destination, any node that overhears a data packet also checks its current location and determines if the destination is within its transmission range. If so,the destination node is added to the list of neighboring nodes, if it is not already present. Note that, this particular check incurs zero cost, i.e. no beacons need to be transmitted. The MP rule solely may not be sufficient for maintaining an accurate local topology. In the worstcase, assuming no other nodes were in the nearby range, the data packets would not be transmitted at all here To maintain a more accurate local topology devise a mechanism in those regions of the network . This is precisely On-Demand Learning (ODL) rule aims to achieves this. As the name suggests, a node broadcasts beacons packet on-demand, i.e. in response to data forwarding node that occur in activities involve the vicinity of that node According to this rule, whenever a node overhears a data transmission from a new neighbor, it broadcasts a beacon as a response. Node waits for a small random time interval before responding with the beacon to prevent collisions with other beacon .In addition, since the data packet contains the location of the final destination, any node that overhears data packet also checks its current location and determines if the destination is within its transmission range. If so, the destination node is added to the list of nodes neighbor if it is not added. Note that, this particular check incurs turns to zero cost, i.e., no beacons need to be transmitted.

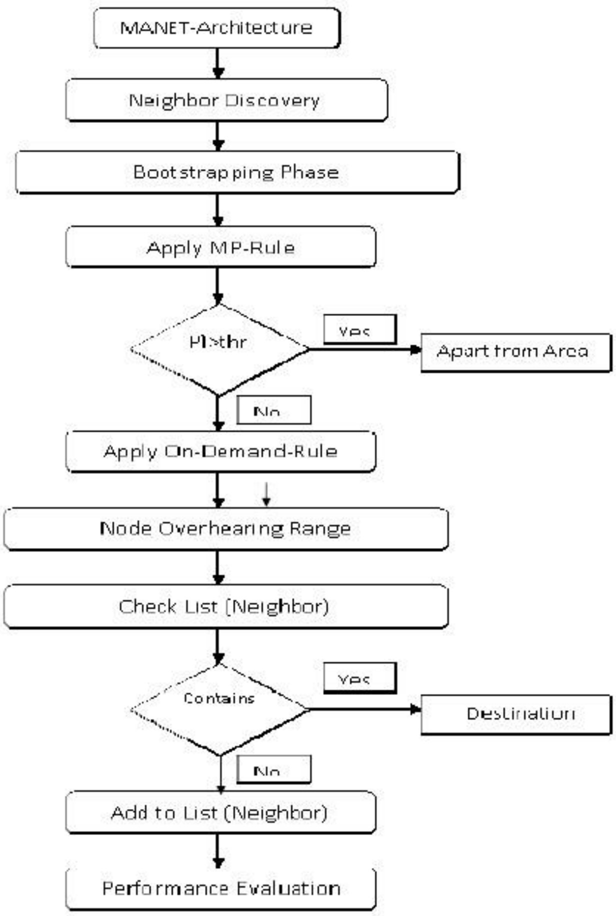


4.3.1 Example for ODL rule

4.4 FALSE NODE DETECTION:

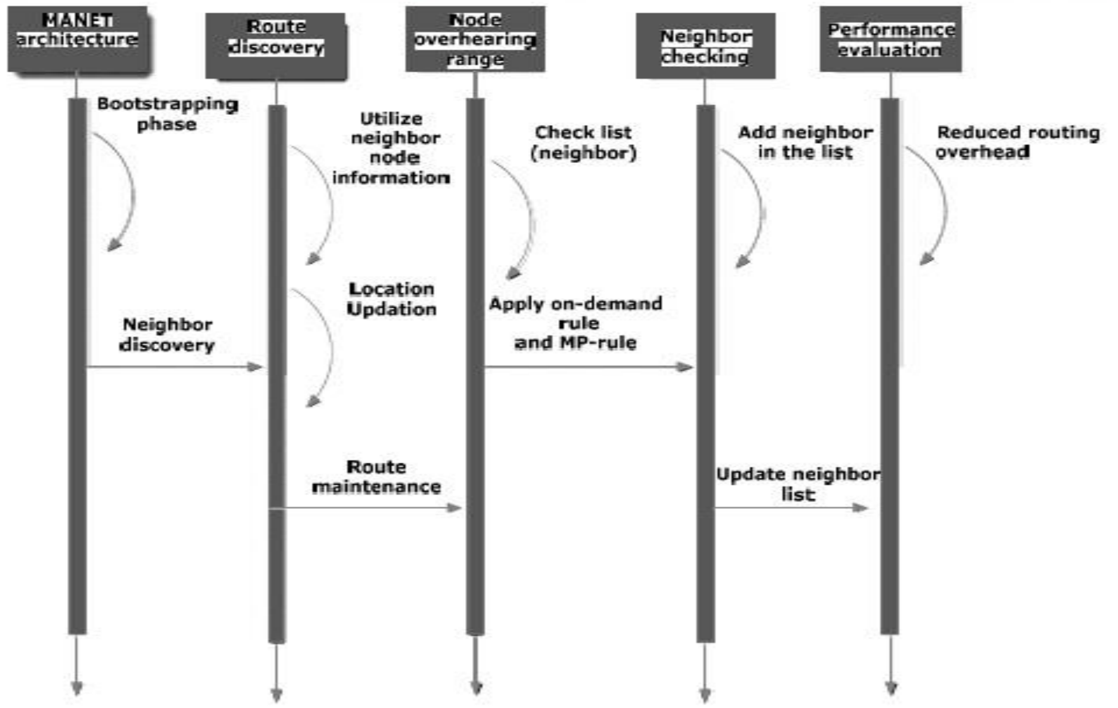
The nodes participating in the packet forwarding should co-operate, if these nodes are not forwarding the packets to the destination then these nodes are considered as the selfish nodes. These selfish nodes detection is an important factor in the network performance. The detected selfish nodes are avoided from the routing path to avoid the lost of the packets. The amount of packets can be saved from these selfish nodes and thus can enhance the network performance through the detection of these nodes. Selfish nodes are inclined to get the greatest profits from the networks and at the same time these nodes trying to conserve their own resources like bandwidth, batterylife or hardware. A selfish node only communicates to other nodes if its data packet is required to send to some other node and refuses to cooperate other nodes whenever it some data packets or routing packets are received by it that it has no interest in. Hence data packets are either refused to retransmit or are dropped for being received by a selfish node. The nodes which don't send RREQ packets don't impact the network, this sort of selfish nodes can increase end to end delay because the number of nodes in the transmission path will increase. If a hello message is not accepted from a neighbour inside two seconds of the last message, connectivity lost is determined to that neighbor node.

5.1 FLOWCHART:



Cite this article as: Siva Ranjani, Ranjani N, Sri Abarna, Ananthi K, Parkavvi. “EFFICIENT ROUTING AND FALSE NODE DETECTION IN MANET.” *International Conference on Information Engineering, Management and Security (2015)*: 27-42. Print.

5.2 SEQUENCE DIAGRAM:



5.3 SYSTEM CONFIGURATION

5.3.1 HARDWARE CONFIGURATION:

Processor : Intel Pentium dual core

RAM : 2 GB

Clock speed : 1.6 GHz

Hard disk : 40 GB

5.3.2 SOFTWARE CONFIGURATION:

Operating System : Windows XP /Red Hat Linux 9.0

Tools : NS2

Languages : TCL/Tk,awk,GCC

6.1 ADAPTIVE POSITION UPDATE:

In this paper, we propose a novel beaconing strategy for geographic routing protocols called Adaptive Position Updates strategy (APU). Our scheme eliminates the drawbacks of periodic beaconing by adapting to the system variations. APU incorporates two rules for triggering the beacon update process. The first rule uses a simple mobility prediction scheme to estimate when the location information broadcast in the previous beacon becomes inaccurate. The next beacon is broadcast only if the predicted error in the location estimate is greater than a certain threshold, thus tuning the update frequency to the mobility of the nodes. The second rule proposes an on-demand learning strategy, whereby beacons are exchanged in response to data packets from new neighbors in a node's vicinity. This ensures that nodes involved in forwarding data packets maintain a fresh view of the local topology. On the contrary, nodes that are not in the vicinity of the forwarding path are unaffected by this rule and do not broadcast beacons. By reducing the beacon updates, APU reduces the power and bandwidth utilization, resources which are scarce in MANETs. It also decreases the chance of link-layer collisions with the data packets and consequently reduces the end-to-end delay. Note that, APU simply governs the beacon update strategy and is hence compatible with any geographic routing protocol. In this work, we have incorporated the APU strategy within GPSR (Greedy Perimeter Stateless Routing) [2] as a representative example. We have carried out simulations to evaluate the performance improvement achieved by APU with randomly generated network topologies and mobility patterns. We have also performed some initial experiments with realistic movement patterns of buses in a metropolitan city. Our initial results indicate that APU significantly reduces beacon overhead without having any noticeable impact on the data delivery rate.

6.2 GPSR(GREEDY PERIMETER STATELESS ROUTING PROTOCOL):

Greedy Perimeter Stateless Routing, GPSR, is a responsive and efficient algorithms before it, which use graph-theoretic notions of shortest paths and transitive reachability to find routes, GPSR exploits the correspondence between node and connectivity in a wireless network, by using the positions of nodes to make packet forwarding decisions. In this paper,we aim at reducing the beacon overhead.In case of MANET Upon initialization, each node broadcasts a beacon informing its neighbors about its presence and its current location and velocity. Following this, in most geographic routing protocols such as GPSR, each node periodically broadcasts its current location information. The position information received from neighboring beacons is stored at each node. Based on the position updates received from its neighbors, each node continuously updates its local topology, which is represented as a neighbor list. Only those nodes from the neighbor list are considered as possible candidates for data forwarding. Thus, the beacons play an important part in maintaining an accurate representation of the local topology.GPSR uses greedy forwarding to forward packets to nodes that are always progressively closer to the destination. In regions of the network where such a greedy path does not exist (*i.e.,* the only path requires that one move temporarily farther away from the destination), GPSR recovers by forwarding in *perimeter mode*, in which a packet traverses successively closer *faces* of a planar sub graph of the full radio network connectivity graph, until reaching a node closer to the destination, where greedy forwarding resumes. GPSR makes greedy forwarding decisions using only information about a router's immediate neighbors in the network topology. When a packet reaches a region where greedy forwarding is impossible, the algorithm recovers by routing around the perimeter of the region. By keeping state only about the local topology, GPSR scales better in per-router state than shortest-path and ad-hoc routing protocols as the number of network destinations increases. Under mobility's frequent topology changes, GPSR can use local topology information to find correct new routes quickly.

Greedy Forwarding: As mentioned in the introduction, under GPSR, packets are marked by their originator with their destinations' locations. As a result, a forwarding node can make a locally optimal, greedy choice in choosing a packet's next hop. Specifically, if a node knows its radio neighbors' positions, the locally optimal choice of next hop is the neighbor geographically closest to the packet's destination. Forwarding in this regime follows successively closer geographic hops, until the destination is reached. An example of greedy nexthop choice appears in Figure 1. Here, x receives a packet destined for D. x's radio range is denoted by the dotted circle about x, and the arc with radius equal to the distance between y and D is shown as the dashed arc about D. x forwards the packet to y, as the distance between y and D is less than that between D and any of x's other neighbors. This greedy forwarding process repeats, until the packet reaches D. A simple beaconing algorithm provides all nodes with their neighbors' positions: periodically, each node transmits a beacon to the broadcast MAC address, containing only its own identifier (e.g., IP address) and position. We encode position as two four-byte floatingpoint quantities, for x and y coordinate values. To avoid synchronization of neighbors' beacons, as observed by Floyd and Jacobson , we jitter each beacon's transmission by 50% of the interval B between beacons, such that the mean inter-beacon transmission interval is B, uniformly distributed in [0:5B; 1:5B]. Upon not receiving a beacon from a neighbor for longer than timeout interval T, a GPSR router assumes that the neighbor has failed or gone out-ofrange, and deletes the neighbor from its table. The 802.11 MAC layer also gives direct indications of link-level retransmission failures to neighbors; we interpret these indications identically. We have used T = 4:5B, three times the maximum jittered beacon interval, in this work. Greedy forwarding's great advantage is its reliance only on knowledge of the forwarding node's immediate neighbors. The state required is negligible, and dependent on the density of nodes in the wireless network, not the total number of destinations in the network.1 On networks where multi-hop routing is useful, the number of neighbors within a node's radio range must be substantially less than the total number of nodes in the network. The position a node associates with a neighbor becomes less current between beacons as that neighbor moves. The accuracy of the set of neighbors also decreases; old neighbors may leave and new neighbors may enter radio range. For these reasons, the correct choice of beaconing interval to keep nodes' neighbor tables current depends on the rate of mobility in the network and range of nodes' radios. We show the effect of this interval on GPSR's performance in our simulation results. We note that keeping current topological state for a one-hop radius about a router is the minimum required to do any routing; no useful forwarding decision can be made without knowledge of the topology one or more hops away. This beaconing mechanism does represent pro-active routing protocol traffic, avoided by DSR and AODV. To minimize the cost of beaconing, GPSR piggybacks the local sending node's position on all data packets it forwards, and runs all nodes' network interfaces in promiscuous mode, so that each station receives a copy of all packets for all stations within radio range. At a small cost in bytes (twelve bytes per packet), this scheme allows all packets to serve as beacons. When any node sends a data packet, it can then reset its inter-beacon timer. This optimization reduces beacon traffic in regions of the network actively forwarding data packets. In fact, we could make GPSR's beacon mechanism fully reactive by having nodes solicit beacons with a broadcast "neighbor request" only when they have data traffic to forward. We have not felt it necessary to take this step, however, as the one-hop beacon overhead does not congest our simulated networks. The power of greedy forwarding to route using only neighbor nodes' positions comes with one attendant drawback: there are topologies in which the only route to a destination requires a packet move temporarily farther in geometric distance from the destination . A simple example of such a topology is shown in . Here, x is closer to D than its neighbors w and y. Again, the dashed arc about D has a radius equal to the distance between x and D. Although two paths, (x ! y ! z ! D) and (x ! w ! v ! D), exist to D, x will not choose to forward to w or y using greedy forwarding. x is a local maximum in its proximity to D. Some other mechanism must be used to forward packets in these situations.

6.3 MOBILITY PREDICTION RULE:

To avoid periodic beaconing in the routing strategy,APU adapts the beacon update intervals to the mobility dynamics of the nodes and the amount of data being forwarded in the neighborhood of the nodes. To achieve this APU employs MP rule. The beacons transmitted by the nodes contain their current position and speed. Nodes estimate their positions periodically by employing linear kinematic equations based on the parameters announced in the last announced beacon. If the predicted location is different from the actual location, a new beacon is broadcast to inform the neighbors about changes in the node's mobility characteristics The Mobility Prediction rule is triggered when there is change in the location of the node. The change in the location of the node cannot be predicated feasibly because the nodes move in the random fashion. This rule adapts the beacon generation rate to the frequency with which the nodes change the characteristics that govern their motion (velocity and heading). The motion characteristics are included in the beacons broadcast to a node's neighbors. The neighbors can then track the node's motion using simple linear motion equations.

Nodes that frequently change their motion need to frequently update their neighbors, since their locations are changing dynamically. On the contrary, nodes which move slowly do not need to send frequent updates. A periodic beacon update policy cannot satisfy both these requirements simultaneously, since a small update interval will be wasteful for slow nodes, whereas a larger update interval will lead to inaccurate position information for the highly mobile nodes. In our scheme, upon receiving a beacon update from a node *i*, each of its neighbors records node *i'*s current position and velocity and periodically track node *i's* location using a simple prediction scheme based on linear kinematics (discussed below). Based on this position estimate, the neighbors can check whether node *i* is still within their transmission range and update their neighbor list accordingly. The goal of the MP rule is to send the next beacon update from node *i* when the error between the predicted location in the neighbors of *i* and node *i's* actual location is greater than an acceptable threshold.The neighbours estimate the current position of node I by using linear kinematics equation.On the contrary node *i* uses the same prediction scheme to keep track of its predicted location among its neighbors.Node *i* then computes the deviation with this information.If the deviation is greater than a certain threshold, known as the Acceptable Error Range (AER), it acts as a trigger for node *i* to broadcast its current location and velocity as a new beacon. The MP rule, thus, tries to maximize the effective duration of each beacon, by broadcasting a beacon only when the predicted position information based on the previous beacon becomes inaccurate. This extends the effective duration of the beacon for nodes with low mobility, thus reducing the number of beacons. Further, highly mobile nodes can broadcast frequent beacons to ensure that their neighbors are aware of the rapidly changing topology. In this method, the mobility prediction rule (MP rule) helps in reducing the amount of beacon packets transmitted in the MANET. Mobility prediction rule help in reducing the traffic of beacon overhead and enabling the increase of packet delivery ratio. Mobility Prediction rule also helps in reducing update cost,bandwidth,end-to-end delay. Mobility Prediction (MP) uses a simple mobility prediction scheme to estimate when the location information broadcast in the previous beacon becomes inaccurate. The next beacon is broadcast only if the predicted error in the location estimate is greater than a certain threshold, thus tuning the update frequency to the dynamism inherent in the node's motion.

6.4 ON-Demand Learning Rule :

The MP rule solely may not be sufficient for maintaining an accurate local topology. Consider the example illustrated in which node A moves from P1 to P2 at a constant velocity. Now, assume that node A has just sent a beacon while at P1. Since node B did not receive this packet, it is unaware of the existence of node A. Further, assume that the AER is sufficiently large such that when node A moves from P1 to P2, the MP rule is never triggered. However, node A is within

the communication range of B for a significant portion of its motion. Even then, neither A nor B will be aware of each other. Now, in situations where neither of these nodes are transmitting data packets, this is perfectly fine since they are not within communicating range once A reaches P2. However, if either A or B was transmitting data packets, then their local topology will not be updated and they will exclude each other while selecting the next hop node. In the worst case,

assuming no other nodes were in the vicinity, the data packets would not be transmitted at all. Hence, it is necessary to devise a mechanism, which will maintain a more accurate local topology in those regions of the network where significant data forwarding activities are on-going. This is precisely what the On- Demand Learning rule aims to achieve. As the name suggests, a node broadcasts beacons on-demand, i.e., in response to data forwarding activities that occur in the vicinity of that node. According to this rule, whenever a node overhears a data transmission from a new neighbor, it broadcasts a beacon as a response. By a new neighbor, we imply a neighbor who is not contained in the neighbor list of this node. In reality, a node waits for a small random time interval before responding with the beacon to prevent collisions with other beacons. Recall that, we have assumed that the location updates are piggybacked on the data packets and that all nodes operate in the promiscuous mode, which allows them to overhear all data packets transmitted in their vicinity. In addition, since the data packet contains the location of the final destination, any node that overhears a data packet also checks its current location and determines if the destination is within its transmission range. If so, the destination node is added to the list of neighboring nodes, if it is not already present. Note that, this particular check incurs zero cost, i.e., no beacons need to be transmitted. We refer to the neighbor list developed at a node by virtue of the initialization phase and the MP rule as the basic list. This list is mainly updated in response to the mobility of the node and its neighbors. The ODL rule allows active nodes that are involved in data forwarding to enrich their local topology beyond this basic set. In other words, a rich neighbor list is maintained at the nodes located in the regions of high traffic load. Thus, the rich list is maintained only at the active nodes and is built reactively in response to the network traffic. All inactive nodes simply maintain the basic neighbor list. By maintaining a rich neighbor list along the forwarding path, ODL ensures that in situations where the nodes involved in data forwarding are highly mobile, alternate routes can be easily established without incurring additional delays. ODL diagram illustrates the network topology before node A starts sending data to node P. The solid lines in the figure denote that both ends of the link are aware of each other.

6.5 PERFOMANCE EVALUATION:

6.5.1 PACKET DELIVERY RATIO:

The ratio of the number of delivered data packet to the destination. This illustrates the level of delivered data to the destination. The greater value of packet delivery ratio means the better performance of the protocol.

PDR=Σ Number of packet receive / Σ Number of packet send

6.5.2 END-TO-END DELAY:

The average time taken by a data packet to arrive in the destination. It also includes the delay caused by route discovery process and the queue in data packet transmission. Only the data packets that successfully delivered to destinations that counted. The lower value of end to end delay means the better performance of the protocol.

End-to-End delay=Σ (arrive time – send time) / Σ Number of connections

6.5.3 PACKET LOSS:

The total number of packets dropped during the simulation. The lower value of the packet lost means the better performance of the protocol.

Packet lost = Number of packet send – Number of packet received .

IMPLEMENTATION:

WIRELESS-GPSR.TCL:

```
set opt(chan) Channel/WirelessChannel
set opt(prop) Propagation/TwoRayGround
set opt(netif) Phy/WirelessPhy
set opt(mac) Mac/802_11
set opt(ifq) Queue/DropTail/PriQueue ;# for dsdv
set opt(ll) LL
set opt(ant) Antenna/OmniAntenna
set opt(x) 670 ;# X dimension of the topography
set opt(y) 670 ;# Y dimension of the topography
set opt(cp) "./cbr100.tcl"
set opt(sc) "./grid-deploy10x10.tcl"
set opt(ifqlen) 50 ;# max packet in ifq
set opt(nn) 100 ;# number of nodes
set opt(seed) 0.0
set opt(stop) 250.0 ;# simulation time
set opt(tr) trace.tr ;# trace file
set opt(nam) out.nam
set opt(rp) gpsr ;# routing protocol script (dsr
or dsdv)
set opt(lm) "off" ;# log movement
LL set mindelay_ 50us
LL set delay_ 25us
LL set bandwidth_ 0 ;# not used
Agent/Null set sport_ 0
Agent/Null set dport_ 0
Agent/CBR set sport_ 0
```

Cite this article as: Siva Ranjani, Ranjani N, Sri Abarna, Ananthi K, Parkavvi. "EFFICIENT ROUTING AND FALSE NODE DETECTION IN MANET." *International Conference on Information Engineering, Management and Security (2015)*: 27-42. Print.

```
Agent/CBR set dport_ 0
Agent/TCPSink set sport_ 0
Agent/TCPSink set dport_ 0
Agent/TCP set sport_ 0
Agent/TCP set dport_ 0
Agent/TCP set packetSize_ 1460
Queue/DropTail/PriQueue set Prefer_Routing_Protocols 1
# unity gain, omni-directional antennas
# set up the antennas to be centered in the node and 1.5 meters above
it
Antenna/OmniAntenna set X_ 0
28
Antenna/OmniAntenna set Y_ 0
Antenna/OmniAntenna set Z_ 1.5
Antenna/OmniAntenna set Gt_ 1.0
Antenna/OmniAntenna set Gr_ 1.0
# Initialize the SharedMedia interface with parameters to make
# it work like the 914MHz Lucent WaveLAN DSSS radio interface
Phy/WirelessPhy set CPThresh_ 10.0
Phy/WirelessPhy set CSThresh_ 1.559e-11
Phy/WirelessPhy set RXThresh_ 3.652e-10
Phy/WirelessPhy set Rb_ 2*1e6
Phy/WirelessPhy set freq_ 914e+6
Phy/WirelessPhy set L_ 1.0
# The transimssion radio range
#Phy/WirelessPhy set Pt_ 6.9872e-4 ;# ?m
Phy/WirelessPhy set Pt_ 8.5872e-4 ;# 40m
#Phy/WirelessPhy set Pt_ 1.33826e-3 ;# 50m
#Phy/WirelessPhy set Pt_ 7.214e-3 ;# 100m
#Phy/WirelessPhy set Pt_ 0.2818 ;# 250m
#
proc usage { argv0 } {
puts "Usage: $argv0"
puts "\tmandatory arguments:"
puts "\t\t\[-x MAXX\] \[-y MAXY\]"
puts "\toptional arguments:"
puts "\t\t\[-cp conn pattern\] \[-sc scenario\] \[-nn nodes\]"
puts "\t\t\[-seed seed\] \[-stop sec\] \[-tr tracefile\]\n"
}
proc getopt {argc argv} {
global opt
lappend optlist cp nn seed sc stop tr x y
for {set i 0} {$i < $argc} {incr i} {
set arg [lindex $argv $i]
if {[string range $arg 0 0] != "-"} continue
set name [string range $arg 1 end]
set opt($name) [lindex $argv [expr $i+1]]
}
}
proc log-movement {} {
global logtimer ns_ ns
set ns $ns_
# source ../tcl/mobility/timer.tcl
Class LogTimer -superclass Timer
29
LogTimer instproc timeout {} {
global opt node_;
for {set i 0} {$i < $opt(nn)} {incr i} {
$node_($i) log-movement
}
```

```
$self sched 0.1
}
set logtimer [new LogTimer]
$logtimer sched 0.1
}
getopt $argc $argv
if { $opt(x) == 0 || $opt(y) == 0 } {
usage $argv0
exit 1
}
if {$opt(seed) > 0} {
puts "Seeding Random number generator with $opt(seed)\n"
ns-random $opt(seed)
}
#
# Initialize Global Variables
#
set ns_ [new Simulator]
set chan [new $opt(chan)]
set prop [new $opt(prop)]
set topo [new Topography]
set tracefd [open $opt(tr) w]
$ns_ trace-all $tracefd
#set namfile [open $opt(nam) w]
#$ns_ namtrace-all $namfile
#modified
set namfile [open $opt(nam) w]
$ns_ namtrace-all-wireless $namfile $opt(x) $opt(y)
$topo load_flatgrid $opt(x) $opt(y)
$prop topography $topo
#
# Create God
#
set god_ [create-god $opt(nn)]
30
$ns_ node-config -adhocRouting gpsr \
-llType $opt(ll) \
-macType $opt(mac) \
-ifqType $opt(ifq) \
-ifqLen $opt(ifqlen) \
-antType $opt(ant) \
-propType $opt(prop) \
-phyType $opt(netif) \
-channelType $opt(chan) \
-topoInstance $topo \
-agentTrace ON \
-routerTrace ON \
-macTrace OFF \
-movementTrace OFF
source ./gpsr.tcl
for {set i 0} {$i < $opt(nn) } {incr i} {
gpsr-create-mobile-node $i
$node_($i) namattach $namfile
$ns_ at 0.0 "$node_($i) setdest [ expr { rand() * 670 } ] [ expr {
rand() * 670 } ] 10.0"
}
#
# Source the Connection and Movement scripts
#
if { $opt(cp) == "" } {
```

```
puts "*** NOTE: no connection pattern specified."
set opt(cp) "none"
} else {
puts "Loading connection pattern..."
$ns_ at 10.0 "$ns_ trace-annotate \"Loadin connection pattern
...........\""
source $opt(cp)
}
#
# Tell all the nodes when the simulation ends
#
for {set i 0} {$i < $opt(nn) } {incr i} {
$ns_ at $opt(stop).000000001 "$node_($i) reset";
}
$ns_ at $opt(stop).00000001 "puts \"NS EXITING...\" ; $ns_ halt"
if { $opt(sc) == "" } {
31
puts "*** NOTE: no scenario file specified."
set opt(sc) "none"
} else {
puts "Loading scenario file..."
$ns_ at 0.1 "$ns_ trace-annotate \"Loading Scenario
File...........\""
source $opt(sc)
puts "Load complete..."
$ns_ at 0.15 "$ns_ trace-annotate \"Load complete...........\""
}
#added by zhou
for {set i 0} {$i < $opt(nn)} {incr i} {
$ns_ initial_node_pos $node_($i) 10
}
##
puts $tracefd "M 0.0 nn $opt(nn) x $opt(x) y $opt(y) rp $opt(rp)"
puts $tracefd "M 0.0 sc $opt(sc) cp $opt(cp) seed $opt(seed)"
puts $tracefd "M 0.0 prop $opt(prop) ant $opt(ant)"
puts "Starting Simulation..."
proc finish {} {
global ns_ tracefd namfile
$ns_ flush-trace
close $tracefd
close $namfile
exec nam out.nam &
exit 0
}
$ns_ at $opt(stop) "finish"
$ns_ run
DATASHORT.PL:
#!/usr/bin/perl
$ofile="simresult.txt";
$nNodes=10;
$inEnergy=100;
open OUT, ">$ofile" or die "$0 cannot open output file $ofile: $!";
print "Please Stand By. Analyzing File: simple.tr in "; print `pwd`;
print "\n";
open OUT, ">$ofile" or die "$0 cannot open output file $ofile: $!";
print OUT "=================== Simulation Result
===========================\n";
print OUT " Date:"; print OUT `date`;
print OUT "\n Analyzed File: simple.tr in "; print OUT `pwd`;
32
```

```
printQQ22\A\\QQ2 OUT
"\n========================================================================\
n";
while(<>){
@mline = split(':', $_);
@mline2 = split('\[', $mline[0]);
@word = split('\]',$mline2[2]);
@eng = split(" ",$word[0]);
@tline = split('_', $_);
$src=$tline[1];
$Emin[$src] = $eng[1];
}
for ($i=0;$i < $nNodes; $i++) {
# print "Node($i) : $Emin[$i]\n";
# print OUT "Node($i) : $Emin[$i]\n";
$total = $total + $Emin[$i];
}
$consume=($total/($nNodes*$inEnergy))*100;
$average=$total/$nNodes;
for ($i=0;$i<$nNodes;$i++) {
$sub = $average - $Emin[$i];
if($sub < 0) {
$sub = $sub * -1;
}
$sub_total = $sub_total + $sub;
}
$pyuncha=$sub_total/$nNodes;
#Primary Information
print OUT " Total Remained Energy : $total\n";
print OUT " Average Remained Energy : $average\n";
print OUT " Energy Difference : $pyuncha\n";
close OUT;
```

SCREEN SHOTS:

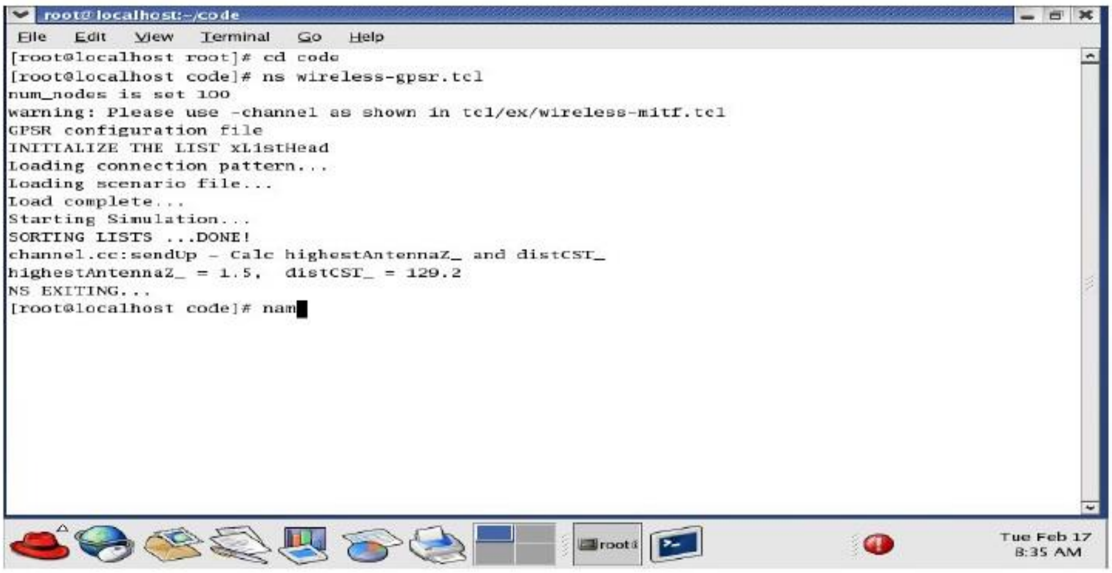


FIG 8.1 RUNNING OF PROGRAM

Cite this article as: Siva Ranjani, Ranjani N, Sri Abarna, Ananthi K, Parkavvi. "EFFICIENT ROUTING AND FALSE NODE DETECTION IN MANET." *International Conference on Information Engineering, Management and Security (2015)*: 27-42. Print.

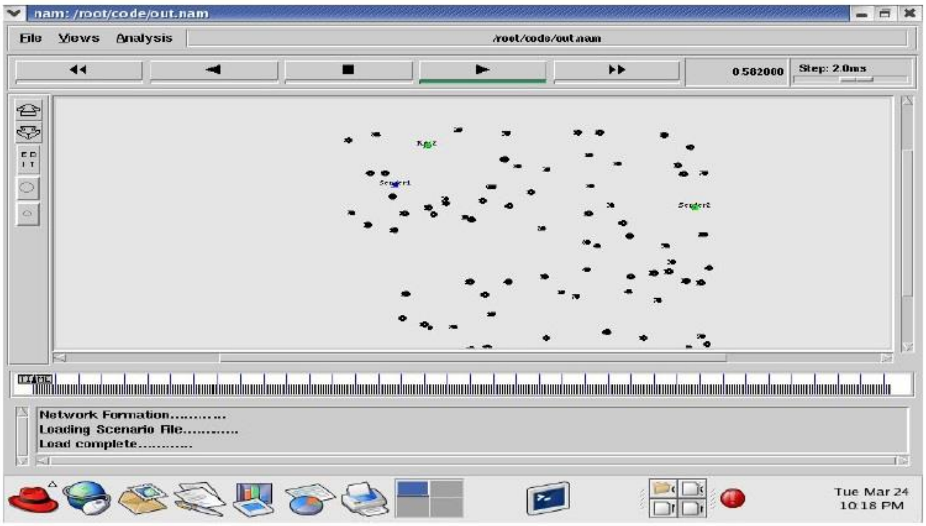


FIG 8.2 APPLYING MP AND ODL RULE

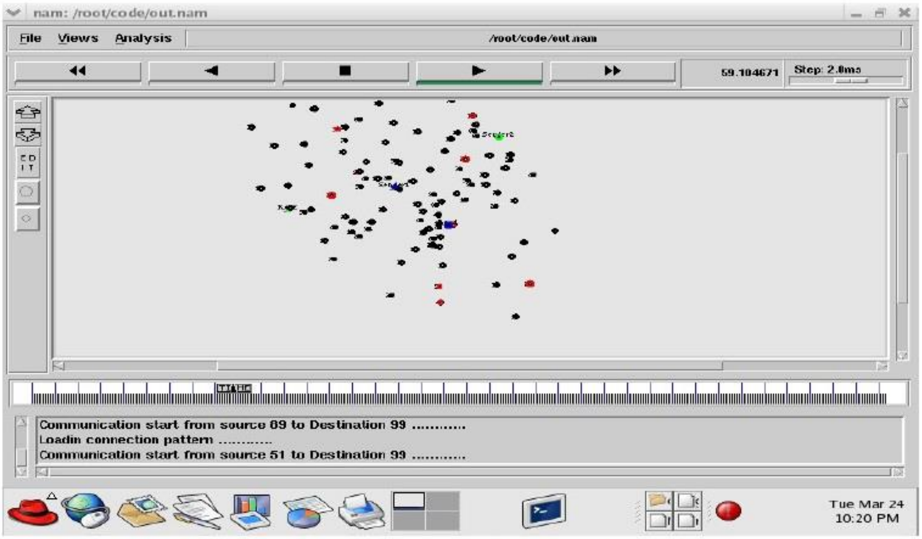
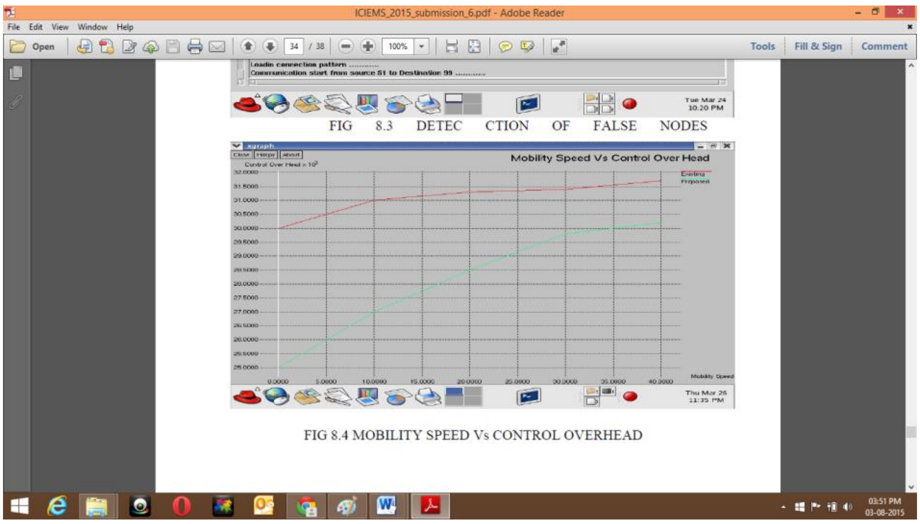


FIG 8.3 DETEC CTION OF FALSE NODES



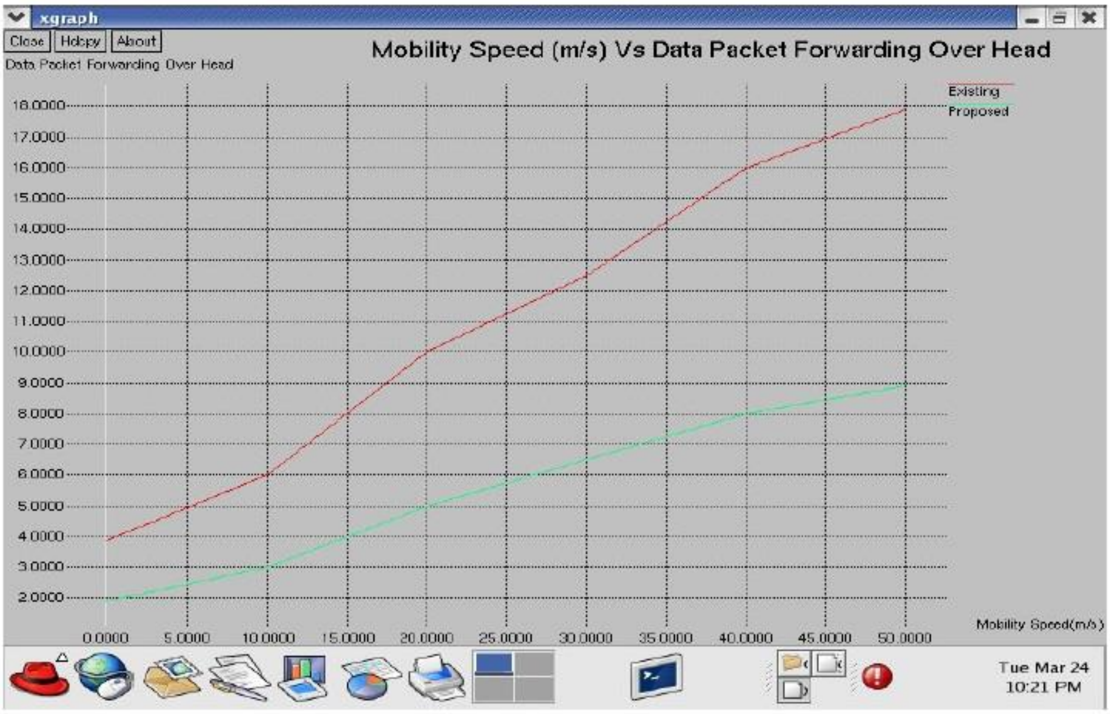


FIG 8.4 MOBILITY SPEED Vs DATA PACKET OVERHEAD

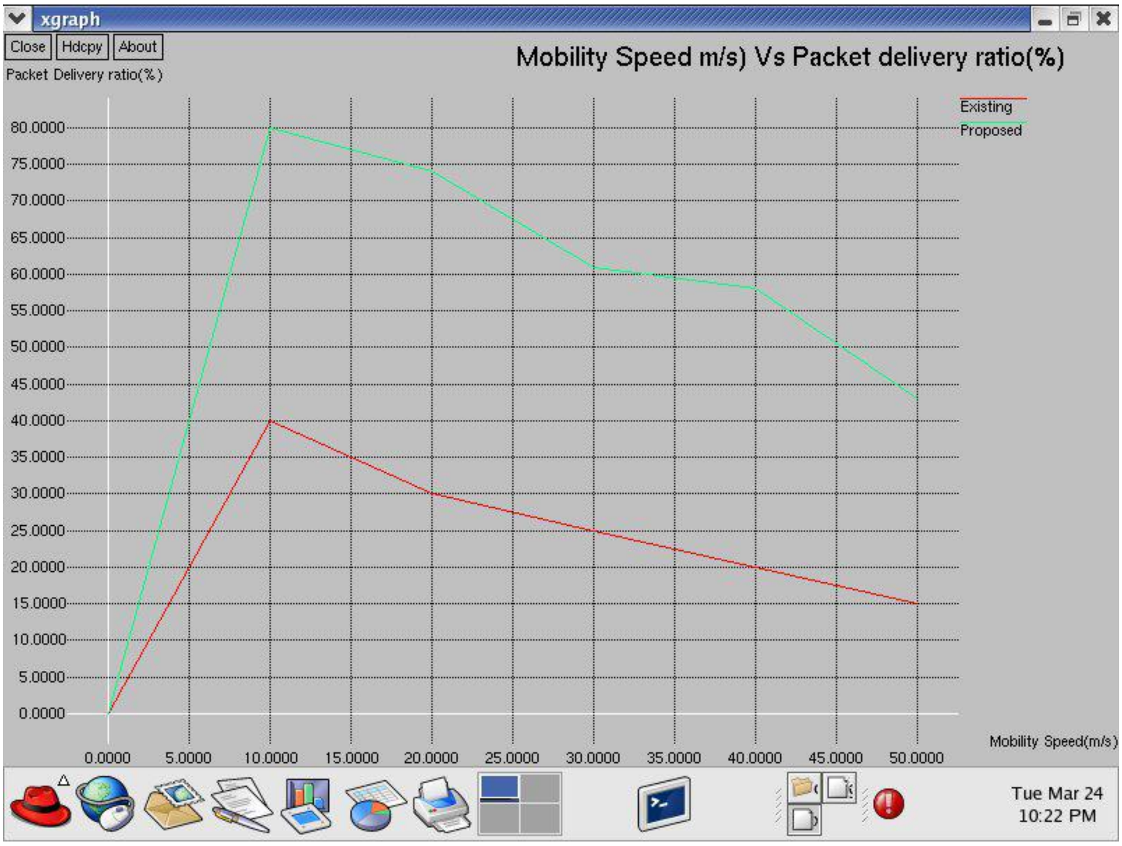


FIG 8.4 MOBILITY SPEED Vs PACKET DELIVERY RATIO

Cite this article as: Siva Ranjani, Ranjani N, Sri Abarna, Ananthi K, Parkavvi. "EFFICIENT ROUTING AND FALSE NODE DETECTION IN MANET." *International Conference on Information Engineering, Management and Security (2015)*: 27-42. Print.

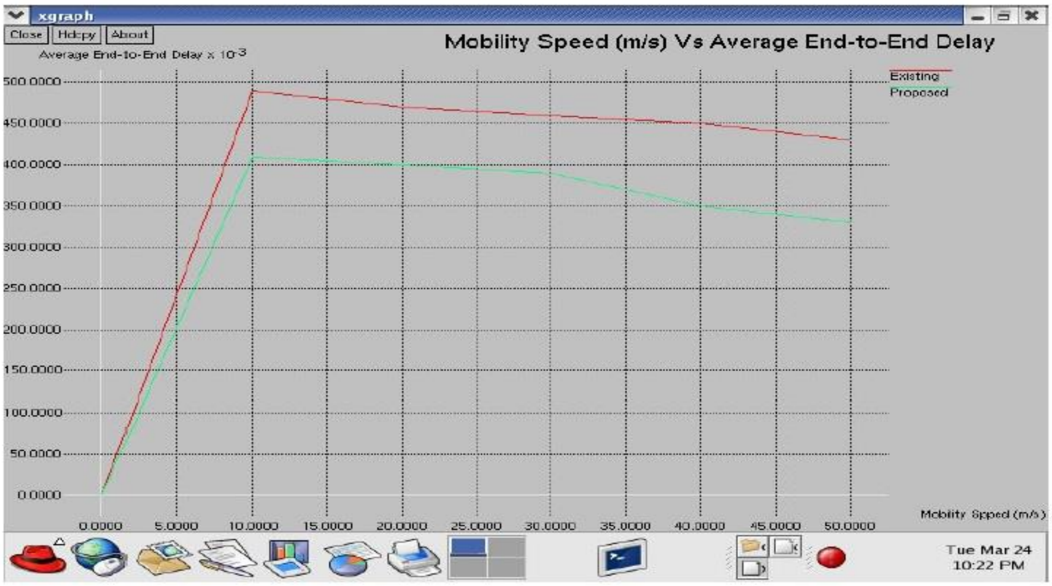


FIG 8.4 MOBILITY SPEED Vs END-TO-END DELAY

CONCLUSION

Each update consumes node energy, wireless bandwidth, and increases the risk of packet collision at the medium access control (MAC) layer. Packet collisions cause packet loss which in turn affects the routing performance due to decreased accuracy in determining the correct local topology (a lost beacon broadcast is not retransmitted). A lost data packet does get retransmitted, but at the expense of increased end-to-end delay. Clearly, given the cost associated with transmitting beacons, it makes sense to adapt the frequency of beacon updates to the node mobility and the traffic conditions within the network, rather than employing a static periodic update policy. In our system, we use Adaptive Position Update strategy.Using this strategy, we can update the node position and velocity dynamically. The system use Periodic beaconing scheme, node can broadcast the beacon for fixed interval because this research based on proactive model.We studied the different recovery delays consecutive to a link failure and observed that this delay, under several topologies and mobility scenarios, was significant and incompatible with delay constrained applications.The simulation studies demonstrate that the proposed routing protocols are more robust and outperform the existing geographic routing protocol and conventional on demand routing protocols under various conditions including different motilities, node densities and traffic loads.

REFERENCES

[1] J. Hightower and G. Borriello, "Location Systems for Ubiquitous Computing," Computer, vol. 34, no. 8, pp. 57-66, Aug. 2001.
[2] B. Karp and H.T. Kung, "GPSR: Greedy Perimeter Stateless Routing for Wireless Networks," Proc. ACM MobiCom, pp. 243- 254, Aug. 2000.
[3] L. Blazevic, S. Giordano, and J.-Y. LeBoudec, "A Location Based Routing Method for Mobile Ad Hoc Networks," IEEE Trans. Mobile Computing, vol. 4, no. 2, pp. 97-110, Mar. 2005.
[4] Y. Ko and N.H. Vaidya, "Location-Aided Routing (LAR) in Mobile Ad Hoc Networks," ACM/Baltzer Wireless Networks, vol. 6, no. 4, pp. 307-321, Sept. 2002.
[5] T. Camp, J. Boleng, B. Williams, L. Wilcox, and W. Navidi, "Performance Comparison of Two Location Based Routing Protocols for Ad Hoc Networks," Proc. IEEE INFOCOM, pp. 1678-1687, June 2002.
[6] D. Johnson, Y. Hu, and D. Maltz, The Dynamic Source Routing Protocol (DSR) for Mobile Ad Hoc Networks for IPv4, IETF RFC 4728, vol. 15, pp. 153-181, Feb. 2007.
[7] C. Perkins, E. Belding-Royer, and S. Das, Ad Hoc On-Demand Distance Vector (AODV) Routing, IETF RFC 3561, July 2003.
[8] J. Li, J. Jannotti, D.S.J.D. Couto, D.R. Karger, and R. Morris, "A Scalable Location Service for Geographic Ad Hoc Routing," Proc. ACM MobiCom, pp. 120-130, Aug. 2000.
[9] Z.J. Haas and B. Liang, "Ad Hoc Mobility Management with Uniform Quorum Systems," IEEE/ACM Trans. Networking, vol. 7, no. 2, pp. 228-240, Apr. 1999.

Cite this article as: Siva Ranjani, Ranjani N, Sri Abarna, Ananthi K, Parkavvi. "EFFICIENT ROUTING AND FALSE NODE DETECTION IN MANET." *International Conference on Information Engineering, Management and Security (2015)*: 27-42. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS007

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.007

Colored Noise Reduction By Kalman Filter

Maitry Chakraborty[1], Depanwita Sarkar[2]
[1]M.Tech,Microelectronics & VLSI Design, [2]Assitant Professor,Electronics & CoEngineering Department
Techno India,Salt Lake

Abstract: *This paper deals with reduction of colored noise from audio signal to improve the audibility of the signal.In this design an audio signal is used as an input signal whose Auto regressive coefficients is calculated by AR-Yule method and is given to the system in a matrix form. Brown colored is used as a noisy signal. In this work Kalman filter is used to reduce colored noise from the audio signal.Kalman filter has the advantage of zero Gaussian noise. It works in a two-step algorithm. Prediction and estimation are the essential steps of it Aim of these designs is to reduce the noise from the signal. This system is designed and simulated using Matlab Simulink version R2013a.*

Keywords: Kalman filter, colored noise,auto regressive co efficients.

I. INTRODUCTION

The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) solution of the least-squares method. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modeled system is unknown. It is shown that the Kalman filter is a linear, discrete time, finite dimensional time-varying system that evaluates the state estimate that minimizes the mean-square error. The Kalman filter dynamics results from the consecutive cycles of prediction and filtering. The dynamics of these cycles is derived and interpreted in the framework of Gaussian probability density functions. Under additional conditions on the system dynamics, the Kalman filter dynamics converges to a steady-state filter and the steady-state gain is derived. The innovation process associated with the filter, that represents the novel information conveyed to the state estimate by the last system measurement, is introduced. The Kalman filter is a recursive estimator. This means that only the estimated state from the previous time step and the current measurement are needed to compute the estimate for the current state.

The Kalman filter addresses the general problem of trying to estimate the state of a $x \in R^n$ discrete-time controlled process that is governed by the linear stochastic difference equation

$x_{k+1} = A_k x_k + Bu_k + w_k$ ...... eq(1)

with a measurement $z \in R^n$

$z_k = H_k x_k + v_k$ ....eq(2)

The random variables w_k and v_k represent the process and measurement noise respectively They are assumed to be independent (of each other), white, and with normal probability distributions.

P(w)=N(0,Q)

P(v)=N(0,R)

In practice, the *process noise covariance Q* and *measurement noise covariance R* matrices might change with each time step or measurement, however here we assume they are constant The n*n matrix A in the difference equation relates the state at the previous time step k-1

Cite this article as: Maitry Chakraborty, Depanwita Sarkar. "Colored Noise Reduction By Kalman Filter." *International Conference on Information Engineering, Management and Security (2015)*: 43-48. Print.

to the state at the current step k , in the absence of either a driving function or process noise. Note that in practice A might change with each time step, but here we assume it is constant. The n*l matrix *B* relates the optional control input to the state *x*. The matrix in the measurement equation relates the state to the measurement *zk*. In practice H might change with each time step or measurement, but here we assume it is constant.

1.3 Computational Origins of The filter :

We define x_k^- € R^n to be the *priori* state estimate at step *k* given knowledge of the process prior to step *k*, and x_k € R^n to be our *a posteriori* state estimate at step *k* given measurement z_k . We can then define *a priori* and *a posteriori* estimate errors as

$$e_k^- \Xi \; x_k - x_k^- \quad \text{and} \quad e_k \;\Xi\; x_k - x_k \qquad \text{.........eq (3)}$$

The *a priori* estimate error covariance is then

$$P_k^- = E\,[e_k^-\, e_k^{-T}] \qquad \text{...........eq (4)}$$

And the a posteriori error covariance is

$$P_k^- = E\,[e_k^-\, e_k^{T}] \qquad \text{.............eq (5)}$$

In deriving the equations for the Kalman filter, we begin with the goal of finding an equation that computes an *a posteriori* state estimate x_k as a linear combination of an *a priori* estimate x^-_k and a weighted difference between an actual measurement z_k and a measurement prediction $H\,x^-_k$.

$$x_k = x_k^- + k\,(z_k - Hx_k^-) \qquad \text{...............eq(6)}$$

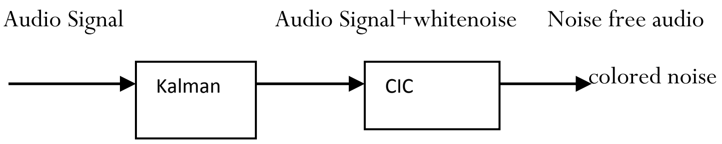
the difference (z_k – Hx_k^-) is called measurement innovation or the residual. The residual reflects the discrepancy between the predicted measurement and the actual measurement z_k A residual of zero means that the two are in complete agreement.

The n*n matrix K is chosen to be the gain or blending factor that minimizes the *a posteriori* error covariance equation

$$K_k = P_k^-\, H^T\,(\,H\,P_k^-\,H^T + R\,)^{-1}$$

$$= (P_k^-\,H^T) \,/\, (\,H\,P_k^-\,H^T + R\,) \qquad \text{....... eq (7)}$$

as the measurement error covariance R approaches zero,the gain K weights the residual more heavily.On the other hand, as the *a priori* estimate error covariance P_k^- approaches zero, the gain*K* weights the residual less heavily as the measurement error covariance R approaches zero,the gain K weights the residual more heavily.



Noise suppression by kalman filter[3] : There are several noise reduction algorithms base on linear prediction have been proposed in case that noise signal is AWGN(additive white Gaussian noise).There are some advantages of using kalman filter.

Assuming that the speech signal d(n) is degraded by an additive observation noise v(n), a noisy speech signal r(n) is given by

r(n) = d(n)+v(n)

A noise suppression procedure consists in estimating the speech signal d(n) from the sole noisy speech signal r(n).Here it is assumed that the noise v(n) is additive noise with known variance σ_v^2.The noise variance may not be known in practice but it can be estimated in many methods. Our purpose of it is to achieve high performance noise suppression without sacrificing quality of the speech signal from the only noisy speech signal r(t) for the additive white and colored noise.

The method utilizes the canonical state space models with a state equation considered of the speech signal, and an observation equation consisted of the speech signal and additive noise.The proposed algorithm is performed without the conception of the generated model of speech signal.The main tool in our algorithm is Kalman filter theory only.

For the $L_p * 1$ state vector $x_p\ (n+1)$ of the proposed method

Define

$X_c(n+1)=d(n+1),d(n),\ldots\ldots\ldots\ldots\ldots,d(n-L_p+2)]^T$ ....eq(8)

Nothing the signal d(n) ,we give state equation :

State equation

$X_c\ (n+1) = \varphi_p\, x_p\ (n) + \delta_p\ (n+1)$ ....eq(9)

Where the $L_p * L_p$ transition matrix φ_p and the $L_p * 1$ driving source vector $\delta_p\ (n+1)$ are expressed as

$$\varphi_p = \begin{bmatrix} 0 \ldots\ldots\ldots\ldots\ldots\ldots \ldots & 0 \\ 0 \quad 1 \ldots\ldots\ldots\ldots\ldots & 0 \\ 0 \quad 0 \quad 1 \ldots\ldots\ldots\ldots\ldots & 0 \end{bmatrix}$$

$\delta_c\ (n+1) = [d(n+1),0,\ldots\ldots.0]^T$ ....eq(10)

furthermore define the $L_p * 1$ extended observation vector $y_p\ (n+1)=[1,r(n+1),\ldots\ldots\ldots.r(n-L_p+3)]^T$
the observation equation may be also expressed as
observation equation
$y_p\ (n+1) = M_p\ x_p\ (n+1) + \epsilon_p\ (n+1)$ ....eq(11)

where the $L_p * L_p$ observation matrix M_p and the $L_p * 1$ extended vector $\epsilon_p\ (n+1)$ are expressed as

$M_p = \varphi_p$,

$\epsilon_p\ (n+1) = [1,v(n+1),\ldots\ldots v(n - L_p + 3)]^T$ .....eq(12)

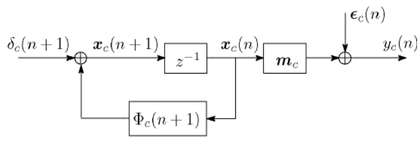


Fig: Block Diagram of the Overview of Noise

It is easily guessed that the proposed algorithm does not depends on both the estimation accuracy of the parameters of AR system and the order L_c of it, since the proposed algorithm does not include the conception of the generated model of signal,
CIC (Cascaded Integrated Comb Filter)[5]:

CIC filters were invented by Eugene B. Hogenauer, and are a class of FIR filters used in multi-rate digital signal processing. The CIC filter finds applications in interpolation and decimation. Unlike most FIR filters, it has a decimator or interpolator built into the architecture. The figure at the right shows the Hogenauer architecture for a CIC interpolator.

CIC filter is one of the simplest filter because the logic requirement of this filter is less as this filter does not require a multiplier. Since the filter co efficients are all unity multiplier is not required in that kind of filter. However the comb filter is not very efficient for removing noise in wide frequency range.For many applications the comb filter must be used in connection with one or more additional design. A cascaded Integrator Comb filter is a special class of linear phase, finite impulse response filter. For many applications which cannot tolerate this distortion.the comb filter must be used in conjunction with one or more additional digital filter stages.CIC filters generally used in multirate systems for better performance.The system is completely designed to reduce noise while noise is mixing continuously with the source signal.For developing such system a CIC filter is crude to eliminate noise from sine wave.

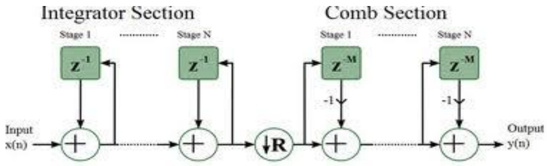


Fig: Block diagram of Cascaded Integrator Comb Filter

$H(z) = [\sum_{k=0}^{RM-1} Z^{-k}]^N$

$= [(1-Z^{-RM}) / (1-Z^{-1})]^N$

Where R=Decimation or interpolation ratio

M=number of samples per stage

N=number of stages in filter

Characteristics of CIC filters are that it gives linear phase response and utilizes only delay and addition and subtraction.CIC filter structure The basic elements of a CIC filter are integrator filters and comb filters.

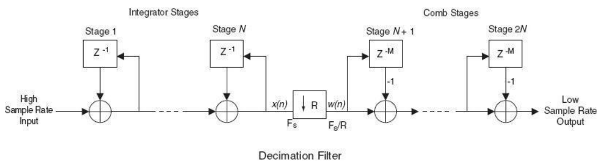
1.7Frequency characteristics :

The transfer function for a CIC filter is at f_s is

$H(z) = H_I(z)\, H_C(z) = = [(1-Z^{-RM}) / (1-Z^{-1})]^N$

This equation shows that even though a CIC has integrators in it, which by themselves have an infinite impulse responses , a CIC filter is equivalent to N FIR filters. Each having a rectangular impulse responses. Since all of the coefficients of these FIR filters are unity and therefore symmetric,a CIC filter also has a linear phase response and constant group delay.

$|H(f)| = |[(\sin \pi Mf)/(\sin \pi f/R)]^N|$



Block Diagram of CIC Filter

Here 10*10 matrix is being considered.By using AR-yule method the auto regressive co efficient of the input signal and the brown noise is calculated.

The matlab code for it is simulated to find out the AR coefficient.

$\delta_c(n+1) =$

$$\begin{pmatrix}
1.0 & 0.8237 & 0.1114 & -0.0471 & -0.0079 & 0.0151 & 0.0799 & -0.1224 & 0.0271 & 0.04231 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{pmatrix}$$

Cite this article as: Maitry Chakraborty, Depanwita Sarkar. “Colored Noise Reduction By Kalman Filter.” *International Conference on Information Engineering, Management and Security (2015)*: 43-48. Print.

$\Phi_c(n+1) =$

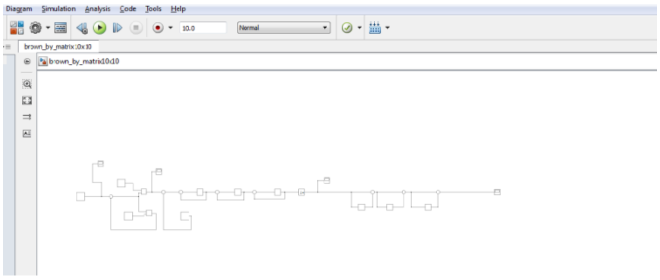
$$\begin{pmatrix}
1.0000 & 0.0943 & 0.9264 & -0.0745 & 0.0252 & -0.1387 & 0.0822 & 0.0252 & -0.0974 & 0.1850 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}$$

$X_c(n+1)=$

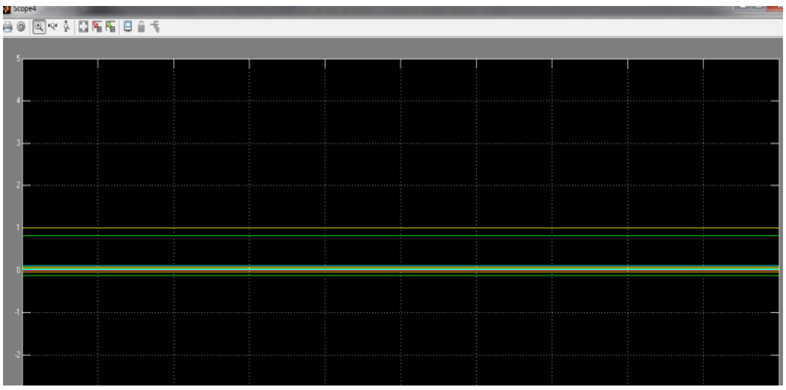
$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

Results and Discussion :

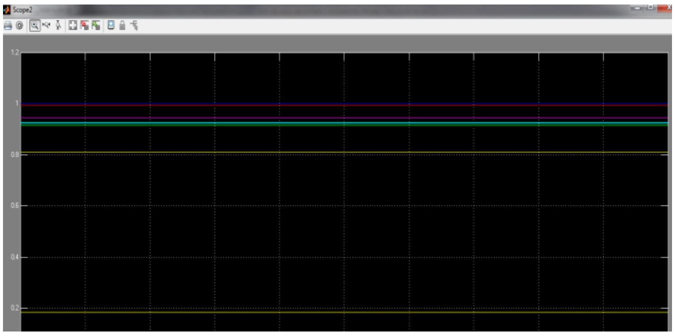
These values are given and then simulated by Matlab R2013a.The circuit which is simulated after giving the AR values of audio sign. and brown noise is given below



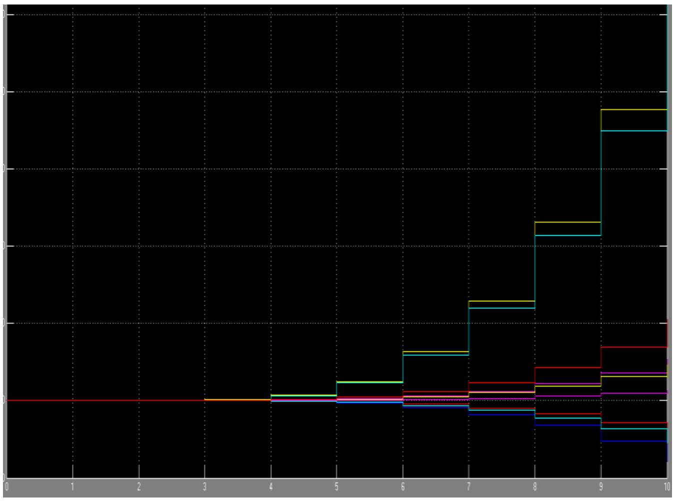
After the simulation of the above circuit the scope block of the audio signal give the output like,



Output of the scope block after mixing brown noise with audio signal looks like,



Output of the final scope shown below where the noise is down to zero.The audio signal is amplified.



CONCLUSION

In this paper the auto regressive coefficient form of the input audio signal is used as input in the Kalman filter algorithm.By using the Kalman filter the brown noise is removed successfully.the advantage of Kalman filter is zero Gaussian noise.Cascaded Integrator Comb Filter is used as a linear FIR filter to remove the colored noise.

REFERENCES

[1] Greg Welch, Garry Bishop, "An Introduction to the Kalman Filter ", Department of Computer Science, University of North Carolina at Chapel Hill.

[2] Maria Isabel Ribeiro, "Kalman and Extended Kalman Filters: Concept, Derivation and Properties " Institute of Systems and Robotics.

[3] Nari Tanabe , Toshirio Furukawa , Hideaki Matsue , Shigeo Tsujii ,"Noise suppression with High Speech Quality Based on Kalman filter" ,IEEE,2006 International Symposium on Intelligent Signal Processing and Communication Systems(ISPAC 2006), Yonago Convention Center, Tottori, Japan

[4] Nari Tanabe , Toshirio Furukawa , Shigeo Tsujii, "A Kalman Filter based Fast Noise Suppression Algorithm" ,IEEE,Tokyo University of Science , Sua ,Japan.

[5] Anil Singh, Poonam Singhal and Rajeev Ratan "Multistage implementation of multirate CIC filters" ECE,DCRUST , Campus Murthal , Sonepat

Cite this article as: Maitry Chakraborty, Depanwita Sarkar. "Colored Noise Reduction By Kalman Filter." *International Conference on Information Engineering, Management and Security (2015)*: 43-48. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS008		**eAID**	ICIEMS.2015.008

Mining URL based feedback comments using Multi – Dimensional Trust Algorithm

Ms. A. Divya
ME Software Engineering, St. Joseph's College of Engineering Chennai, India

ABSTRACT: *A product aspect ranking framework is proposed which automatically identifies the important aspect of the product based on the customer review that is to be used for further numerous reviews to make the customer understand the product and its aspects more easily. The numerous reviews of products are available on the Internet. The reviews are based on the product. The customer reviews contain rich and valuable knowledge for users. The aspect ranking is proposed, which automatically identifies the important aspects of products from online customer reviews, to improving the usability of the numerous reviews. The main contributions include proposing a multidimensional trust model for computing reputation scores from user feedback comments and propose an algorithm for mining feedback comments for dimension ratings and weights. The product based on the customer reviews. The important product aspects are identified based on two observations. The first observation is important aspects are usually commented on by a large number of consumers and the second observation is the important aspects of the consumer opinions are greatly achieve their overall opinions on theproduct. The first is identifying product aspects and develop a probabilistic aspect ranking algorithm. There are document-level sentiment classification and extractive review summarization. To achieve the significant performance improvements, that demonstrates the product aspect ranking capacity in facilitating real-world applications. The product reviews are taken in the Internet based on the URL.*

***Keywords*:** Product aspects, product aspect ranking, Product aspect identification, sentiment classification, consumer feedback, review extractive summarization, Multi Dimensional trust algorithm.

1. INTRODUCTION

Data mining is the extraction of hidden predictive information from large databases. It means extracting or "mining" knowledge of large amount of data. It is a new technology powerful with great potential to help companies focus on the most important information in their data warehouses. The Product aspect ranking is proposed, which automatically identifies the important aspects of products from online consumer reviews and it is aiming at to improve the usability of the numerous reviews. Most companies already collect and refine quantities of data. Data Mining techniques can be implemented rapidly on existing software and hardware platforms to enhance the value of existing information resources and can be integrated with new products and systems as they are bought on-line.

1.1 Definition of Problem

The Previous system gets the reviews from the single user; it does not allow multiple users to give review on the product. The user will be restricted to minimum. To overcome this problem, the reviews from the multiple users are allow registering in a specified URL, and then analyzed using sentiment classification, to identify the reviews of the product. The performance will be improved while comparing with the users involved based on the continuous reviews.

1.2 Need for Sentiment Rating

The Sentiment Rating is used for rating the customer reviews based on the product. It is used for Sentiment classification. The Product Aspect Ranking is used for rating the product. It consists of three main components. There are aspect identification, sentiment classification on aspects and probabilistic aspect ranking. To given the product with consumer reviews, to identify first the aspects in the reviews and then analyse consumer opinions on the aspects via a sentiment classifier. The Sentiment Classification is about assigning a positive, negative and neutral label to a piece of text based on its overall opinion. In this project, it is used for analysing the reviews and overall rating the reviews based on the product. The task of analyzing the sentiments expressed on aspects is called aspect-

Cite this article as: A. Divya. "Mining URL based feedback comments using Multi – Dimensional Trust Algorithm." *International Conference on Information Engineering, Management and Security (2015)*: 49-55. Print.

level sentiment classification in literature [12]. Exiting techniques include the supervised learning approaches and the lexicon-based approaches, which are typically unsupervised. The lexicon-based methods utilize a sentiment lexicon consisting of a list of sentiment texts, words and idioms, are determine the sentiment orientation on different aspect [23]. The sentiment classifier is then leveraged to determine the opinion of the opinionated expression, i.e. the opinion on the aspect.

1.3 Objectives of Project

The Product Ranking framework, that automatically generates the aspect of the product only based on the individual user's opinion in URL using Multi Dimensional Trust algorithm. The further work conducts the product ranking that is generated in source of a URL in large scale. Executing the review in real time environment based on the customer feedback. The Product aspect ranking is used for analysing feedback from the online marketing websites and rating the product.

1.4 Structure of the Project

Product aspect ranking is a wide range of real-world applications. It is used for business improvement. In this project, the product can be analyzed by the customer feedbacks. In the customer feedbacks will be taken from Amazon web services based on URL. The word extraction is used for pre-processing and opinion extraction methods. The pre-processing has stop word removal, stemming and part of speech. The opinion word extraction has unigram and bigram method. Stop words are words which are filtered out before or after, processing of natural language data (text). The stemming is process for reducing inflected or something derived words to their stem, base or root form. A part of speech is a category to which a word is assigned in accordance with its syntactic function. It is noun, pronoun, adjective, adverb, etc. This module will classify the word as dependent and independent. The word that is classified will be given as input and we will find the polarity of the given opinion word. Using the opinion word that is tagged first will form unigram and bigram. Unigram is a single word and bigram is a combination of unigram. In the Formation of unigram it will consider each word as a unigram and in bigram we will combine two words to form a word of bigram. The bigram formation is used classify the polarity of word correctly. The classification is the process in which ideas and objects are recognized, differentiated and understood. It is the process of finding set of models that describe and distinguish data classes or concepts for the purpose of being able to use the model to predict the class objects whose the unknown class label is found. The derived model is used for the analysis of a set of training data. The clustering is the process of grouping a set of physical or abstract objects into classes of similar objects. Clustering techniques consider tuples of data as objects. They partition the data objects into groups or clusters so that objects within a cluster are similar to one another and dissimilar to objects in other clusters. Clustering is a process of grouping the data or words belonging to a same polarity (i.e.) clustering the positive and negative polarity separately. The K-means is a portioning method. It is also K-nearest classifier. Main thing of the project is update and add URL, the URL will be specified the product. It will update the details of produc, product reviews, etc. Then, the aspect raking will be calculated the customer reviews in sentiment classification. The Sentiment classification means analyzing the customer reviews and rating the reviews in positive, negative and neutral. Finally, the overall rating will be calculated and it will be displayed in 3D view.

2. REVIEWS OF LITERATURE

The Survey papers are mainly described by rating of the customer reviews. Product aspect ranking is a wide range of real-world applications. It is used for business improvement. In this project, the product can be analyzed by the customer feedbacks. In this paper [1] Author said, the cross domain algorithm is used for analysing the customer reviews and the sentiment is expressed differently in different domains. Its interest is costly. After applying the Sentiment classifier it uses labeled and unlabelled data. The empirical analysis can be conducted in this method. It evaluate single and multisource domain adaptation supervised and unsupervised domain adaptation, these are used for creating the sentiment sensitive thesaurus. These papers are mainly analyzed by sentiment classification with any other concepts and algorithms. In this paper, the author [2] says sentiment analysis based on the conditions. For example Affect, Judgment and Appreciation. The customer reviews are based on these conditions. Some papers are analyzed by the Cross domain sentiment classification analysis and some are supervised and unsupervised algorithm. The paper [3] author says, this sentiment classification is based on the text to speech. The comments based on the text. It will execute speech. This execution is related to emotion. Text classification is main technique is used. The paper [4] describes, probabilistic aspect mining will be generated. It will identify aspects related to class. The Efficient EM algorithm is developed and it is used. In case, the mining is used for identify the customer reviews based on the Intrinsic and Extrinsic Domain relevance. A novel based method is used to identify the opinion features from online reviews by exploiting the difference in opinion features statistics across two corpora, the author said [5]. In this literature survey papers, the efficiency and accuracy will be very low. In this project, it will overcome all these things.

2.1 Approaches to Sentiment Classification

The Sentiment Classification is used for analyzing the customer reviews based on the product. The sentiment classification is about assigning a positive, negative or neutral label to apiece of text based on its overall opinion. The document-level sentiment classification is to determine the overall opinion of a given review document. A review document often expresses various opinions on multiple aspects of a each product. The opinions on many aspects might be in contrast, and have other degree of impacts on the overall opinion of the review document. The supervised learning methods train a sentiment classifier based on training corpus. The classifier is then used to predict the sentiment on each aspect. Many learning-based classification models are applicable. Sentiment Classifier (SC) will be classifying reviews as Positive or Negative based on the sentiment expressed in customer review. The Sentiment Classification is used in many other approaches. They are supervised and unsupervised, Intrinsic and Extrinsic domain, etc.

Cite this article as: A. Divya. "Mining URL based feedback comments using Multi – Dimensional Trust Algorithm." *International Conference on Information Engineering, Management and Security (2015)*: 49-55. Print.

2.1.1 Supervised Learning

The Supervised Learning Approaches is dependent on the training data and cannot perform well without sufficient training samples. Automatic classification of sentiment is important for numerous applications such as opinion mining, opinion summarization and market analysis. To apply a sentiment classifier and it is trained using labeled data for a particular domain. It is used to classify sentiment of user reviews on a different domain.

2.2 Approaches to Sentiment Ranking

Product Aspect ranking framework is used for analysing the reviews. We start with an overview of its pipeline. It consisting of three main components: aspect identification, sentiment classification on aspects and probabilistic aspect ranking. To given the product with consumer reviews, to identify first the aspects in the reviews and then analyze consumer opinions on the aspects via a sentiment classifier.

Fig.1. Overall Rating in 3D view

Probabilistic aspect ranking algorithm is used to infer the importance of the aspects by simultaneously taking into account aspect frequency and the influence of consumers' opinions given to each aspect over their overall opinions. The Overall Rating is analyzed by the reviews; it will be shown in the Figure 1.

2.2.1 Multi-Dimensional Trust Algorithm

A Multi-Dimensional trust algorithm is used for computation of reputing scores from customer feedback comments. It is used for a fine-grained multidimensional trust evaluation model by mining ecommerce feedback comments. Dimension trust scores together with their weights are further computed by clustering aspect expressions into dimensions and aggregating the dimension rating. This algorithm is used for mining the customer feedback comments for rating the dimensions and the computing dimension weights will be described. This algorithm based on clustering dimension expressions into dimensions and computing dimension weights. A sentence in feedback is represented as a set of dependency relations between pairs of words in the form of (head, dependent), where content words are chosen as heads, and other related words depend on the heads. Aspect opinion expressions and their associated ratings (positive, negative) are first extracted from feedback comments shown in the Fig. 2. The feedback comments are shown in sources. Their opinions more honestly and openly. The analysis of feedback comments on Amazon reveals that even if a buyer gives a positive rating for transaction. The comments regarding different aspects of transactions in feedback comments.

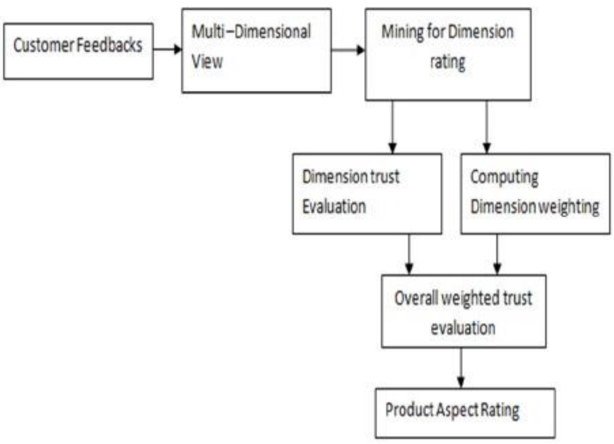


Fig.2 Multi-Dimensional trust Algorithm

Cite this article as: A. Divya. "Mining URL based feedback comments using Multi – Dimensional Trust Algorithm." *International Conference on Information Engineering, Management and Security (2015)*: 49-55. Print.

2.2.2 K-mean Clustering and hierarchal Clustering

The K-Nearest Neighbour classifiers are based on learning the analogy. The samples are described in trained by n - dimensional numeric attributes. Each Sample represents a point in an n – dimensional space. This way of the training samples are stored in an n - dimensional pattern space. When given an unknown sample a K – nearest neighbour classifier searches the pattern space for the K – training samples that are closes to the unknown sample. These K training samples are the K – nearest neighbours of the unknown sample. It will group the data based on the k-nearest neighbours. In this we Insert edges between a node and its k-nearest neighbours. Each node will be connected to (at least) k nodes. The k-means clustering is a method of cluster analysis which aims to partition n observations into kclusters in which each observation belongs to the cluster with the nearest mean.

2.3 Extraction and Summarization of Aspect opinion

This project is related to mining or sentiment analysis on free text documents or customer feedbacks. This field is presented in [6]-[7]. There has been existing work on product aspect ranking in reviews of product. The product reviews are considered by frequent and noun phrases. The dependency relation parsing [30] is used to mining the aspect opinions for product reviews. The aspects are grouped into clusters and group them into meaningful clusters. Unsupervised techniques have been developed to jointly model opinions and aspects based on LDA [9]. The aspect rating has been computed from overall ratings in e-commerce feedback comments or reviews [10]-[12]. Their aspect ratings weights are computed based on regression from overall ratings and the positive in overall ratings. The approach based on the typed dependency analysis to extracting aspect opinion expressions and identifying their associated ratings. This algorithm based for clustering dimension expressions into dimensions and computing dimension weights.

Table 1. Sample comments on Amazon

No	Comment	Amazon rating
C1	Quick response	1
C2	Bad communication, will not buy from again. Super slow ship. Item as described	1
C3	Top seller, many thanks, A+	1
C4	Great price and awesome service! Thank you!	1
C5	Product arrived swiftly! great seller	1

2.4 Limitations of earlier approaches to sentiment Ranking

The sentiment ranking is used for analyzed the customer feedbacks based on the product. It is used for ranking the words. The words are analyzed in positive, neutral and negative. It consists of three main components: aspect identification, sentiment classification on aspects and probabilistic aspect ranking. In this sentiment rating has many other approaches. They are intrinsic and extrinsic domain dependent, supervised and unsupervised, support vector machine, back propogation, neural network, lexicon, Maximum Entrophy, Naïve bayes, etc.. The product reviews can be compared the following methods of sentiment classification. One is unsupervised method; it is determined by referring to the sentiment lexicon WordNet. The lexicon contains a list of positive and negative sentiment words. The method is supervised method, it determines naïve bayes, maximum entropy and support vector machine. The sentiment classifiers are well trained on the reviews based on producers and consumers. The maximum entropy was implemented with parameter estimation. The support vector machine was implemented by using library svm with linear kernel. The naïve bayes was implemented with Laplace smoothing.

3. RESULT AND DISCUSSION

The Multi-Dimensional trust algorithm collects the reviews from the specified input URL, and the rating is performed. The rating distinguishes the positive, negative and neutral feedbacks of the customers on the product. The sentiment value is calculated for each word in the rating, so that the overall rating of the product is analysed.

Cite this article as: A. Divya. “Mining URL based feedback comments using Multi – Dimensional Trust Algorithm.” *International Conference on Information Engineering, Management and Security (2015)*: 49-55. Print.

Fig 3. URL based Product from Amazon

Fig 4. Overall Rating of the Product

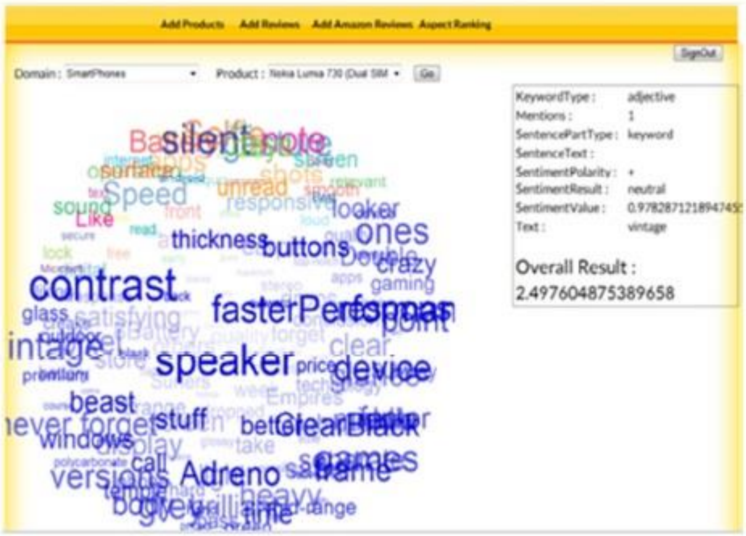
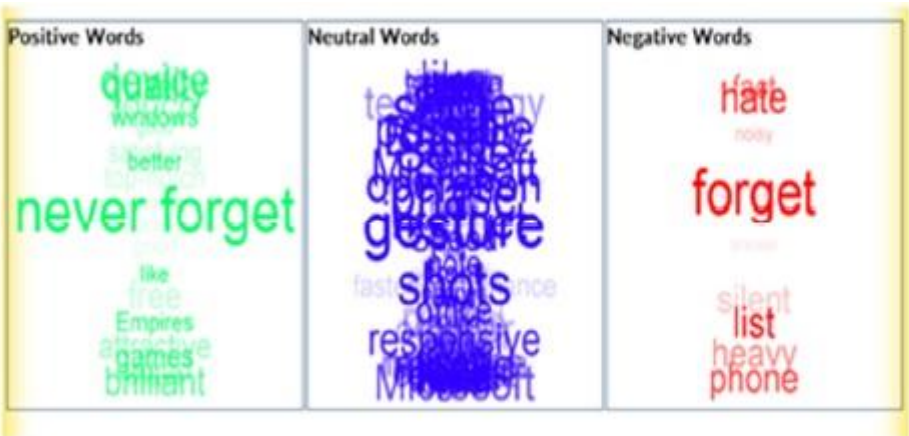


Fig 5. Positive, Negative, Neutral Rating of the Product



Cite this article as: A. Divya. "Mining URL based feedback comments using Multi – Dimensional Trust Algorithm." *International Conference on Information Engineering, Management and Security (2015)*: 49-55. Print.

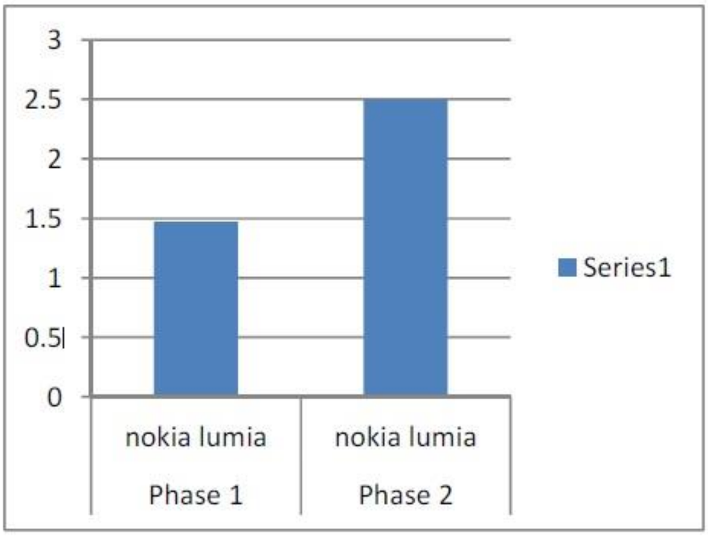
4.2 Sentiment ranking with Multi-Dimensional Trust Algorithm and Overall Rating Estimation

Product Aspect Ranking framework start with an overview of its pipeline. It is used for ranking the words. The words are analyzed in positive, neutral and negative. It consisting of three main components: aspect identification, sentiment classification on aspects and probabilistic aspect ranking. Given the consumer reviews of a product, it identifies the aspects in the reviews and then analyzes consumer opinions on the aspects via a sentiment classifier.

Rank-diff = Σirank(i) – rank'(i) / N

The rank difference between two ranking vectors is defined as: where rank(i) and rank'(i) are respectively the rank for seller *i* by two ranking methods, and *N*=10.

Fig 6. Improved rating of phase 1 to Phase 2



The aspect ranking algorithm is used to infer the importance of the aspects by simultaneously taking into account aspect frequency and the influence of consumers' opinions given to each aspect over their overall opinions. The words are collected in the 3D view. The words are rotating and analyzed. The clustering is used for grouping the words. Finally, the sentiment value is calculated and overall value is described. In the Phase 1, the user is created own reviews. The user level is very low and the user knows the review level and overall rating. It is the disadvantage of the Phase 1. But, the phase 2 is the user has unlimited, so the reviews also unlimited. The Developer does not know the rating. It only calculated be the Sentiment Classification.

4. CONCLUSION

The product aspect ranking has been proposed to identify the important aspects of products from numerous customer feedbacks. This process contains three components. They are product aspect identification, aspect sentiment classification and product aspect ranking. It is used to improve aspect identification and sentiment classification on text reviews. The high reputation scores for sellers cannot effectively rank sellers and therefore cannot guide potential buyers to select trustworthy sellers to transact with. In this paper a multi-dimensional trust algorithm has been proposed for sellers by uncovering dimension ratings embedded in feedback comments. Extensive experiments on feedback comments for Amazon sellers demonstrate that our approach computes trust scores highly effective to distinguish and rank sellers. The effective algorithms is proposed to compute dimension trust scores and dimension weights automatically via extracting aspect opinion expressions from feedback comments and clustering them into dimensions. Our approach demonstrates the novel application of combining natural language processing with opinion mining and summarization techniques in trust evaluation for e0- commerce applications.

REFERENCES

[1] Danushka Bollegala, Member, IEEE, David Weir, and John Carroll, "Cross-Domain Sentiment Classification Using a Sentiment Sensitive Thesaurus" IEEE Transactions on Knowledge and Data Engineering, vol. 25, NO. 8, aug. 2013

[2] Alena Neviarouskaya and Masaki Aono, "Sentiment Word Relations with Affect, Judgement & Appreciation", IEEE transactions on affective computing, vol. 4, no. 4, oct. – dec. 2013,

[3] Alexandre Trilla and Francesc Alías, "Sentence-Based Sentiment Analysis for Expressive Text-to-Speech", IEEE Transactions On Audio, Speech, And Language Processing, vol. 21, no. 2, 2013

[4] Victor C. Cheng, C.H.C. Leung, Jiming Liu, Fellow, IEEE, and Alfredo Milani, "Probabilistic Aspect Mining Model for Drug Reviews", IEEE transactions on knowledge and data mining, vol. 26, no. 8, aug. 2014

Cite this article as: A. Divya. "Mining URL based feedback comments using Multi – Dimensional Trust Algorithm." *International Conference on Information Engineering, Management and Security (2015)*: 49-55. Print.

[5] Zhen Hai, Kuiyu Chang, Jung-Jae Kim, and Christopher C. Yang, “Identifying Features in Opinion Mining via Intrinsic and Extrinsic Domain Relevance”, IEEE transactions on knowledge and data mining, vol. 26, no. 3, mar. 2014.

[6] Shulong Tan, Yang Li, Huan Sun, Ziyu Guan, Xifeng Yan, Jiajun Bu, Chun Chen, and Xiaofei He, “Interpreting the public sentiment variations on twitter”, IEEE transactions on knowledge and data mining , vol. 26, no. 5, may 2014.

[7] Nikolaos Malandrakis, Student Member, Alexandros Potamianos, Senior Member, Elias Iosif, Student Member and Shrikanth Narayanan, Fellow, “Distributional Semantic Models for Affective Text Analysis”, IEEE Transactions on Audio, Speech and Language Processing, vol. 21, no. 11, nov. 2013.

[8] Georgios Paltoglou and Michael Thelwall, “Seeing Stars of Valence and Arousal in Blog Posts”, IEEE Transactions on Affective Computing, vol. 4, no. 1, jan.- mar. 2013.

[9] Uros Krcadinac, Philippe Pasquier, Jelena Jovanovic, and Vladan Devedzic, “An Open Source Library for Sentence-Based Emotion Recognition”, IEEE Transaction on Affective Computing, vol. 4, no. 3, julsep 2013.

[10] Chun-Wei Seah, Ivor W. Tsang, and Yew-Soon Ong, “Transfer Ordinal Label Learning”, IEEE Transactions on Neutral Networks and Learning Systems, vol. 24, no. 11,nov 2013.

[11] [LIU Lizhen1, SONG Wei1, WANG Hanshi1, LI Chuchu1, LU Jingli2,”A Novel Feature-based Method for Sentiment Analysis of Chinese Product Reviews”, conf. Information and Engineering College, mar. 2014

[12] XU Xueke, CHENG Xueqi, TAN Songbo, LIU Yue, SHEN Huawei, “Aspect-Level Opinion Mining of Online Customer Reviews”, Key Laboratory of Web Data Science and Technology,Conf. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China, mar 2013

[13] ZHAI Zhongwu, XU Hua, JIA Peifa, “An Empirical Study of Unsupervised Sentiment Classification of Chinese Reviews”, tsinghua science and technology, issn 1007-14 11/14 pp702-708, Volume 15, Number 6, December 2010

[14] Dmitriy Bespalov, Bing Bai, Yanjun Qi, Ali Shokoufandeh, “Sentiment Classification Based on Supervised Latent n-gram Analysis”, CIKM’11, October 24–28, 2011.

[15] Vivek Narayanan1, Ishan Arora2, Arjun Bhatia3, “Fast and accurate sentiment classification using an enhanced Naive Bayes model”, Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10.

[16] [16] Valentina Presutti and Diego Reforgiato Recupero, “Frame-Based Detection of Opinion Holders and Topics: A Model and a Tool”, IEEE Computational intelligence magazine, Digital Object Identifier 10.1109/MCI.2013.2291688 feb. 2014

[17] [17] E.Divya1, “Real Time Sentiment Classification Using Unsupervised Reviews”, International Journal of Innovative Research in Computer and Communication Engineering, vol.2, Special Issue 1, mar 2014.

[18] Ahmed Abbasi, Member, Stephen France, Member, Zhu Zhang, and Hsinchun Chen, Fellow, “Selecting Attributes for Sentiment Classification Using Feature Relation Networks”, IEEE Transactions on Knowledge and Data Engineering, vol. 23, no. 3, mar 2011.

[19] YE Qiang, LI Yijun, ZHANG Yiwen, “Semantic-Oriented Sentiment Classification for Chinese Product Reviews: An Experimental Study of Book and Cell Phone Reviews”, School of Management, Harbin Institute of Technology, Harbin 150001, China;f Department of Management Information Systems, University of Arizona, Tucson 85705, U. S. A., TSINGHUA SCIENCE AND TECHNOLOGY , ISSN 1007-0214 07/15 pp 7 9 7 - 8 0 2 Vol. 10 , no. S1 , dec. 2005.

[20] Erik Cambria, Nanyang Technological University,Yangqiu Song, University of Illinois at Urbana Champaign, Haixun Wang, Microsoft Research Asia, Newton Howard, Massachusetts Institute of Technology, “Semantic Multidimensional Scaling for Open-Domain Sentiment Analysis”, IEEE Intelligent Systems, Published by IEEE Computer Society, mar. – apr. 2014.

[21] Chihli Hung and Hao-Kai Lin, “Using Objective Words in SentiWordNet to Improve Word-of-Mouth Sentiment Classification”, 1541-1672/13/$31.00 © 2013 IEEE, Published by the IEEE Computer Society, March/April 2013

[22] Rui Xia, Chengqing Zong, Xuelei Hu, Erik Cambria, “Feature Ensemble Plus Sample Selection: Domain Adaptation for Sentiment Classification”, Published by the IEEE Computer Society, 1541-1672/13/$31.00 © IEEE 2013.

[23] Anindya Ghose and Panagiotis G. Ipeirotis, “Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics”, IEEE Transactions on Knowledge and Data Engineering, vol. 23, no. 10, oct. 2011.

[24] Giuseppe Di Fabbrizio, Ahmet Aker, and Robert Gaizauskas, “Summarizing Online Reviews Using Aspect Rating Distributions andLanguage Modeling”, Published by the IEEE Computer Society, 1541- 1672/13/$31.00 © 2013.

[25] Cristina Bosco, Viviana Patti, Andrea Bolioli,“Developing Corpora for Sentiment Analysis: The Case of Irony and Senti-TUT”, Published by the IEEE Computer Society, mar/apr 2013.

[26] Lisette García-Moya, Henry Anaya-Sánchez, and Rafael Berlanga-Llavori , “Retrieving Product Features and Opinions from Customer Reviews”, Published by the IEEE Computer Society, may-jun. 2013.

[27] Magdalini Eirinaki, Shamita Pisal1, Japinder Singh2 , “Feature-based Opinion Mining and Ranking”, Preprint submitted to Journal of Computer and System Sciences, jul 18, 2011.

[28] Erik Cambria, Björn Schuller, Yunqing Xia, Catherine Havasi, “New Avenues in Opinion Mining and Sentiment Analysis”, Published by the IEEE Computer Society, mar-apr. 2013.

[29] Erik Cambria, Yangqiu Song, Haixun Wang and Newton Howard, “Semantic Multi-Dimensional Scaling for Open-Domain Sentiment Analysis”, IEEE Intelligent Systems, 2012.

[30] Soujanya Poria, Alexander Gelbukh, Amir Hussain, Newton Howard, Dipankar Das, Sivaji Bandyopadhyay, “Enhanced SenticNet with Affective Labels for Concept- Based Opinion Mining”, Published by the IEEE Computer Society, March/April 2013.

Cite this article as: A. Divya. “Mining URL based feedback comments using Multi – Dimensional Trust Algorithm.” *International Conference on Information Engineering, Management and Security (2015)*: 49-55. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS009		**eAID**	ICIEMS.2015.009

Accomplishment of encryption technique to secure file

Mrs. Parul Rathor
Department of Computer Science & Engineering,
Institute of Engineering & Technology, Alwar (R.J), India

ABSTRACT: *The aim of this paper is to implement the new concept of cryptography. To protect information of files in digital form and how to get a security services by network security and cryptography. Though, a general summary of such algorithms like RSA, DES and AES of network security and cryptography is provided first. A complete review of the purposed system of network security and cryptography is then presented by using transmitter. The general attacks of security were reviewed. The purpose of this implementation is to secure a huge amount of files. So that others will unable to know the original data still they know about the procedure of encryption and decryption. This implementation has many applications to secure information including authentication. Here we create new technology by using such transmitter and receiver for decrypting data which is highly secure and accurate.*

Keywords: Master-file, various keys, transmitter and receiver.

I. INTRODUCTION

As we have learn about such cryptographic algorithms like Rivest-Shamir-Adleman [RSA], Data Encryption Standard [DES] and Advanced Encryption Standard [AES]. Also in previous paper we introduce concept of security in cryptography. Here in our data we encapsulate above three algorithm schemes. The very first one is RSA which is also known for asymmetric key algorithm. RSA is a reproduction of two big prime numbers and the secret key and public key are based on this numbers. This RSA algorithm is very easy to understand.

The second one is DES which is of two types double DES and triple DES. Also substitution known as confusion and transposition known as diffusion are the two attributes of cryptography in DES. In substitution characters are altered to numbers or symbols or any other characters and in transposition, it perform permutation over original data. We used DES generally to encrypt data in blocks which have some particular size that is 64 bits. An algorithm and key are also used for encoding and decoding. And last is AES, there are various steps mainly four is used and that are substitution of bytes, shifting of rows, mixing of columns and addition of keys. All procedure is to be process in matrixes which is of four by four. Because of actual weakness in DES, AES is invented. In DES, 56 bit keys were not safe which is based on complete key searches and also 64 bit blocks were measured as weak. So AES as developed which was based on 128 bit blocks with 128 bit keys. There are three major features of AES like Symmetric and parallel structures, Adapted to modern processors and last is suited to smart cards. So these are the basic review of our three cryptographic schemes that is RSA, DES and AES.

Cite this article as: Mrs. Parul Rathor. "Accomplishment of encryption technique to secure file." *International Conference on Information Engineering, Management and Security (2015)*: 56-59. Print.

II. PURPOSED METHODOLOGY

Now these three things we used in main file as a data. Firstly here we create one master file then breaking it into four parts. In all parts we used same data of above three schemes but sequences are different. Likewise for part one we encapsulate RSA on first place, DES on second place and AES on third place. Then for part two it is DES on first place, RSA on second place and AES on third place. In third part we put RSA on first, AES on second and DES on third.. In this way we merge our data in all four parts of master file.

Suppose part 1 is encrypted by RSA, part 2 is encrypted by DES, part 3 is encrypted by AES and so on. At last part 12 is encrypted by DES according to procedure of above so we get twelve encrypted parts e1 to e12 and master file contains all keys in the form. Now we used key for each algorithm but type is different. Also we have separators for each key. If there are four parts of master file then we split that four parts again into three and name them as begin with e1, e2, e3 up to e12. Means for twelve encrypted data we used here twelve separators as shown in figure below.

Fig 1. Keys with their separators.

Master file = RSA key _1_1_ DES key _1_1_ AES key _1_1_ DES key _1_1_ RSA key _1_1_ AES key _1_1_ RSA key _1_1_ AES key _1_1_ DES key _1_1_

The master file will now be separated and twelve keys which we used for encryption, the same keys obtain. Again the same twelve keys will be used for decryption. At last the decrypted file will be obtained After this we used transmitter for keeping original data that is encrypted and also with its master file. The work of this transmitter is used to transmit all above twelve parts along with its separators.

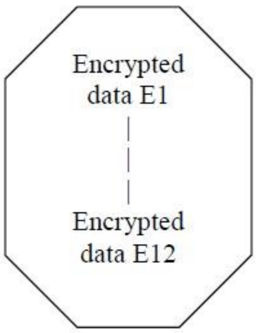


Fig 2. Transmitter

So here we have original data in encrypted form which is kept in transmitter. We can say data of transmitter are kept in input file. Now what about output file?

III. SYSTEM ARCHITECHTURE

After transmission of data encryption are converted into decryption and these decrypted data are to be found in output file. This file is kept in receiver, just like we used transmitter for keeping input file after transmission output file is kept in receiver. The receiver receives all twelve encrypted parts and separates them with its separators. This is only because to get master file and all its parts. The transmitter contains codes of encryption only. While at receiver codes of decryption are kept.

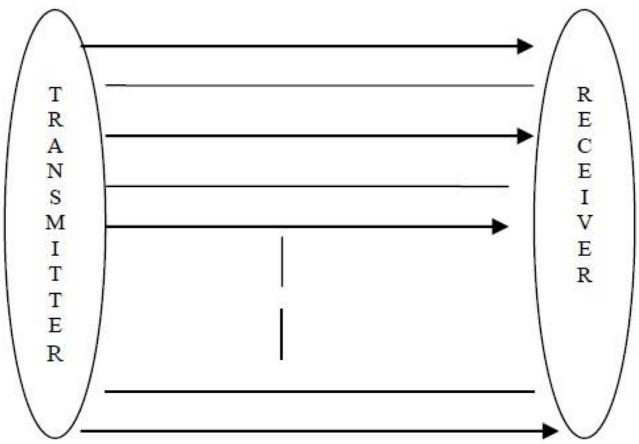


Fig 3. Conversion of original data to decrypted data.

Cite this article as: Mrs. Parul Rathor. "Accomplishment of encryption technique to secure file." *International Conference on Information Engineering, Management and Security (2015)*: 56-59. Print.

In above figure, there are two sides one side we kept transmitter and on the other side there is receiver. In between this figure shows arrows and lines. An arrow is known for file and lines known for separators. The arrow 1 is for master file then kept one separator below this file that shows like single line. In the same way we arrow 2 is known for encrypted data part E1 and kept separator below it. Likewise we arrange all twelve encrypted parts from E1 to encrypted part E12 with its separators.
In other side receiver obtained output file which contains decrypted data.

IV. EXPERIMENTAL RESULTS
We can observe it with the help of results.

Size	Single delay	Multiple delay
20kb	672	531
34kb	2875	578

Fig 4. Single delay versus multiple delays

Size	Single throughput	Multiple throughput
20kb	30236.61	38265.54
34kb	12087.65	61015.57

Fig 5. Single throughput versus multiple throughputs

Here at receiver mode we have transmitter for single connection and for multi connection. Its depend on our choice. According to size of file we have single delay with its throughput and multiple delays corresponding to its throughput. So we find here what will be range between single delay versus multiple delays in figure 4 and in figure 5 single throughput versus multiple throughputs. As well for file size 20 kilobytes there are 672 observe in single delay and 531 observe in multiple delay. And 30236.61 find in single throughput and 38265.54 obtain in multiple throughput. In the same way we can find for file size 34 kilobytes and so on to get desire output. Also we can express more deeply these output with the help of graph as shown below figure 6. There are some advantage and disadvantage to every research, lets discuss about first advantage. An advantage of transmitter (input file) and receiver (output file) for encryption and decryption is make data highly secure. Also it is accurate and gives high throughput.
On the other side if there are good points so it should have bad points also. The disadvantage is here like multiple types of encryption is used, so because of this computation complexity is high. Also as there are twelve parts and for these parts each having its own separators so large amount of space is required which mean memory requirement is high.

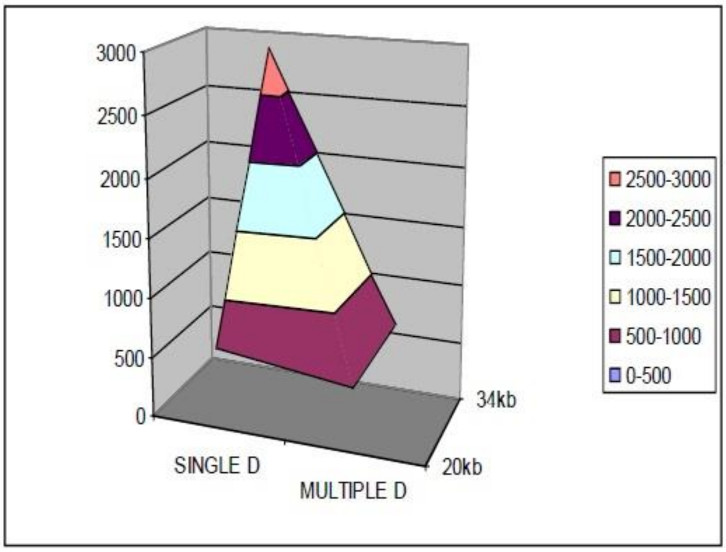


Fig 6. Graph for single delay versus multiple delay.

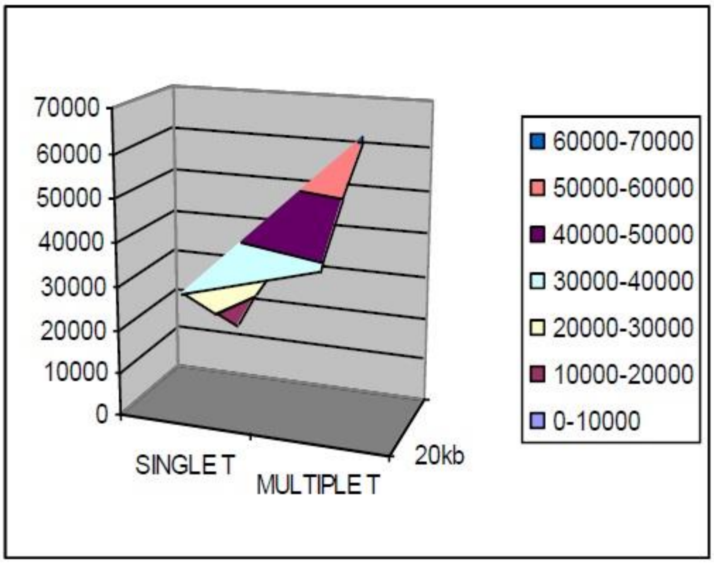


Fig7. Graph for single throughput versus multiple throughput.

V. CONCLUSION

The main purpose for this is today any one can know about cryptographic techniques and about encryption/decryption. They can easily encrypt data but will make difficult for them to decrypt it. This above procedure is somewhat complicated for others to hack it. Hence we discuss by taking shortest review of all algorithms like RSA, DES and AES which were used here. Also try to implement the procedure of conversion data from encryption to decryption in different format. We were found some results which were shown with respect to graph.

ACKNOWLEDGEMENT

It is with deep sense of appreciation and veneration that I express my sincere thanks to my highly respectable supervisor Prof. Rohit sehgal. He has played a pivotal role for my guidance, encouragement, help and useful suggestion throughout. His untiring and painstaking efforts, methodological approach and individual help made it possible to complete this work in time. I like to thank our Principal Prof.(Dr.) Anil Kumar Sharma , H.O.D(CS/IT) Dr. B.K.Verma for providing all the facilities and working environment in the institute.

REFERENCES

[1] Parul Rathor, "Implementation of split based encryption technique for securing file transfer over a network", in International Conference on Industrial Automation and computing [ICIAC-2014].
[2] Parul Rathor, "SPLIT BASED ENCRYPTION IN SECURE FILE TRANSFER",
[3] [3]Rajani Devi.T, "Importance of cryptography in network security", in 2013 International conference on communication systems and network technologies.
[4] Atul kahate, "Cryptography and Network Security" book, tata McGraw-Hill publishing company limited, 2003.
[5] "Cryptography And Network Security" principles and practice, Fifth edition, pearson by William Stallings.
[6] Trappe, W., & Washington, L.C. (2006). *Introduction to Cryptography with Codin Theory*, 2nd ed. Upper Saddle River, NJ: Pearson Prentice Hall.
[7] Denning, D.E. (1982). *Cryptography and Data Security.* Reading, MA: Addison-Wesley.
[8] Ferguson, N., & Schneier, B. (2003). *Practical Cryptography.* New York: John Wiley & Sons.
[9] Electronic Frontier Foundation. (1998). *Cracking DES: Secrets of Encryption Research, Wiretap Politics & Chip Design*. Sebastopol, CA: O'Reilly & Associates.
[10] Schneier, B. (2000). Secrets & Lies: Digital Security in a Networked World. New York: John Wiley & Sons.
[11] Network Security Essential, William Staling, Pearson Publications Ltd

Cite this article as: Mrs. Parul Rathor. "Accomplishment of encryption technique to secure file." *International Conference on Information Engineering, Management and Security (2015)*: 56-59. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS010

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.010

Efficient Method for Identifying Shortest Path in Duty Cycled Wireless Sensor Networks

P.Raja
Student, M.E Multimedia Technology, CEG campus, Anna University Chennai, India.

Abstract: *A Wireless Sensor Network contains a large amount of Sensor Nodes that are mainly data-centric. The sensor nodes operate on battery they cooperate among themselves and transfer data to the sink which processes the sensed information. The end user can access the information collected by these sensor nodes through the Internet. The proposed algorithm deals to find the optimistic path between the sensor nodes and also between the sink and the sensor nodes. The minimum spanning tree algorithm is used in finding the shortest path and a comparison is performed between the enhanced Kruskal's and Prim's algorithm.*

Key Words: Wireless Sensor Networks, Duty Cycled, Routing Protocols, Shortest Path

I. INTRODUCTION

Wireless Sensor Networks consists of battery powered small nodes which are distributed across the sensing areas. The sensor nodes are usually scattered in a sensor field as shown in Fig 1.1 Each of these scattered sensor nodes has the capabilities to collect data and route data back to the sink. Data are routed back to the sink by multi-hop infra structureless architecture through the sink. The sink may communicate with the task manager node via Internet and satellite.

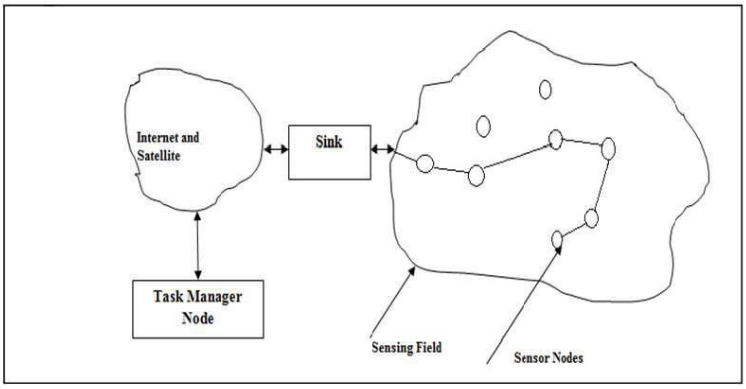


Fig 1.1 Sensor nodes scattered in a sensor field

Routing is the mechanism to find the optimistic path between source to destination. There may be multiple paths from the source to the destination. Therefore, message routing is an important technology. The main performance measures affected by the routing scheme are throughput (quantity of service) and average packet delay (quality of service). Routing schemes should also avoid deadlock. Routing methods can be fixed (i.e. pre-planned), adaptive, centralized, distributed, broadcast.

Cite this article as: P.Raja. "Efficient Method for Identifying Shortest Path in Duty Cycled Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 60-69. Print.

Fixed Routing scheme uses Routing Tables that dictate the next node to be routed to, given the current message location and the destination node. Routing tables can be very large for large networks, and cannot take into account real-time effects such as failed links, nodes with backed up queues, or congested links. There is another method called Adaptive Routing scheme which depends on the current network status and can take into account various performance measures, including cost of transmission over a given link, congestion of a given link, reliability of a path, and time of transmission. They can also account for link or node failures. Some of the routing challenges and design issues are

- Data routing methods
- Fault tolerance
- Transmission media
- Coverage
- Data aggregation
- Quality of Service

Also many energy aware routing protocol are proposed to reduce the energy consumption. This protocol maintains set of path instead of maintains or reinforce one optimal path. Maintenance and selection depends on a certain probability which relays on how low energy consumption of each path can be achieved. The possible goal of energy aware routing are

- Shortest-hop (fewest nodes involved)
- Lowest energy route
- Route via highest available energy
- Distribute energy burden evenly
- Lowest routing overhead

This is very much needed in Wireless Sensor Networks since sensor nodes are battery powered and energy consumption is absolutely necessary. Energy consumption can be reduced by following least latency routing in Time Dependent Duty Cycled Wireless Sensor Networks.Least Latency can be achieved by following a shortest path in order to transfer the messages efficiently from the source to the destination.

II. LITERATURE SURVEY

Low Power Listening(LPL) was proposed by Shouwen Lai,Binoy Ravindran. In LPL the Node only wakes up and listens the channel state for a short time period. B-MAC is a Carrier Sense Media Access(CSMA) protocol for Wireless Sensor Networks.It provides Flexible interface to obtain ultralow power operation,Effective collision avoidance and High channel utilization.The disadvantages of B-MAC are it is difficult to reconfigure the protocols after deployment, thus lacking in flexibility.They do not explicitly support adaptive duty cycling, where nodes choose their duty cycle depending on their residual energy.Showen Lai ,Binoy Ravindran also proposed the Adaptive Low Power Listening (ALPL).In ALPL, since nodes have heterogenous duty cycle setting, it is more difficult for neighbor discovery since a node cannot differentiate whether a neighbor is sleeping or failing when it does not receive feedback from the neighbor.Delay-efficient routing over adaptively duty-cycled Wireless Sensor Networks stated by Shouwen Lai and Binoy Ravindran Routing becomes more difficult due to two reasons:

i. Intermittent connection between two neighbour nodes
ii. Changes in the transmission latency at different times.

Two methods to solve routing over intermittently connected Wireless Sensor Networks due to duty cycling are On-Demand Approach which uses probe messages to determine the least-latency route and Proactive Method where all least-latency routes at different departure times are computed at the beginning. The Limitations of On-Demand Approach are that the method does not work well for frequent data deliveries. The Proactive Approach is Centralized approach and is not flexible for distributed construction.Thus both of these approaches suffer from limitations.Routing must be done with least latency inorder to transmit data efficiently. Hence we first model the Wireless Sensor Networks as Time Dependent which can operate for any particular time interval. The sensor nodes are either in a wake mode or sleep state.Solutions for this problem using a centralized approach and has been studied by Time Dependent Graphs .Limitations are it cannot be applied to WSNs where the global network topology is not known by a centralized node. If the whole time period has M discrete intervals we have to execute the algorithm M times, which is inefficient.Multiple executions suffers from high message cost, which is undesirable for resource-limited Wireless Sensor Networks.It computes all pair shortest path which includes node insertions and deletions.It is an efficient incremental solution. Some of the limitations are it needs O(n) space at each node which is impractical for sensor nodes with limited memory capacity.In the Fast Time Dependent Shortest Path(FTSP)algorithm proposed by Binoy Ravindran Initial Route Construction from the source to destination is computed by this algorithm.It exchanges vectors among sensor nodes rather than single values of static link cost and distance.FTSP computes the shortest path for N discrete Time intervals in one execution rather than N different executions.In first Iteration FTSP computes the shortest path for nodes in layers nearest to sink .In next Iteration it goes beyond every layer until last layer is reached.FTSP-M is used for distributed route maintenance with node insertion, updating,and deletion of sensor nodes in a sensing Network. In Dynamic route maintenance each node stores the route information only to sink and it is memory efficient in sensor node with limited memory.Single sink exists in the sensing areas. Limitations include Active Neighbour Discovery and Synchronization . In Active Neighbour Discovery sensor nodes needs to probe the schedule of the neighbours actively to find out it's state.Synchronization must be ensured at all neighbouring nodes so that they wake up at the same time when an activity is detected.

A distributed algorithm was proposed for efficient routing in duty cycled Wireless Sensor Networks Shouwen Lai and Binoy Ravindran.It uses a synchronizer which provides synchronization among the Sensor Nodes in asynchronous Duty Cycled Wireless Sensor Networks.The nodes are modeled as FIFO(First In First Out.) where first arriving nodes are served first.The main disadvantage is that the algorithm is based on the observation that the time-varying link cost function is periodic, and hence by derivation, the time-varying distance function for each node is also periodic.

A technique for Power Efficient Routing was proposed by Raja Jurdak, Pierre Baldi, and Cristina Videira Lopes which contains a framework for local optimizations in sensor networks that reduces overall power consumption and provides Load Balancing across sensor nodes.Thus the duties of a sensor node are equally shared among the available nodes. The performance decreases drastically when presented with data sets exhibiting higher variability, spending significant portions of time with an either empty or full battery.

A. Woo, T. Tong, and D. Culler provided a method for Reliable Multihop Routing . In this methodology Link status and routing information are maintained in a neighbourhood table with constant space and reliable routing protocols are also discussed. This method stores more space since every node maintains the routing information of its neighbor. When conjestion occurs it results in unstable network performance due to inteference from collisions and estimations over the same pair of nodes behave differently under different channel utilization The Clustered nodes Routing methodology was initiated by

C. Schurgers and M.B. Srivastava which contains implementation of Bitmap-Assisted (BMA)B-MAC routing protocol for large-scale cluster-based Wireless Sensor Networks. Performance decreases when there is high traffic loads and the number of nodes per cluster increases.An adaptive bidirectional interface for wireless sensor network applications uses this B-MAC protocol. It provides an interface allowing middleware services to reconfigure the MAC protocol based on the present workload of the system.A node's lifetime can be extended by changing the configuration of the B-MAC protocol. A set of benchmark standards are used to assess the performance of Wireless Sensor Networks. Clustered nodes routing using B-MAC protocol extends the lifetime of the system by 50%. Data Routing in Extremely Low DutyCycle Wireless Sensor Network was proposed by Y. Gu and T. He where end-to-end routing does not offer to maintain an always-awake communication backbone. Low duty-cycle always causes unreliable nature of wireless communication. It forces to develop a new routing strategy for such networks, so as to achieve network energy efficiency, reliability. The concept of Dynamic Switch-based Forwarding (DSF) that optimizes the expected data delivery ratio, improves energy consumption and reduces the communication latency DSF is designed for networks with already determined node communication schedules.

H. Chon, D. Agrawa, and A. Abbadi developed a Route Planning based on Time dependency. There are certain methods for speeding up the routing process in Time Dependent Networks called Contraction Hierarchy algorithm which is an modified version of Dijkstra shortest path algorithm is discussed in this paper.It is suitable for time dependent Wireless Sensor Networks.A query is sent to all nodes and based on the response a route plan is proposed. In the algorithm for Initial Route Construction the distances from all nodes to the sink node are initially infinite. The time axis is infinite, the time-varying link costs and distance are implemented by vectors. The algorithm is similar to the distributed Bellman-Ford algorithm. The main difference is that algorithm is exchanges vectors for time-varying link costs and distances, rather than single values of static link cost and distance.However, due to the diagram of Time Dependent Wireless Sensor Networks. Routing in Duty Cycled limited resource in WSNs, it is more important to avoid the exponential message complexity of the traditional Bellman-Ford algorithm. This algorithm is equipped with the vector implementation, this algorithm computes the shortest path in infinite time intervals.The adaptively duty cycled Wireless Sensor networks are modeled as time dependent networks which satisfy the FIFO condition.Then a distributed algorithm is computed which can be executed for a single execution across the entire timeperiod.Then the shortest path found is maintained so that it can be used with suboptimal implementation.The FIFO condition referes to the fact that packet which was delivered earlier will always arrive at a direct neighbor earlier.There are two steps in construction of
shortest path.

i.)A shortest path is created by Minimum Spanning Tree Algorithm.

ii.)A shortest path is calculated and an acknowledgement message is sent to the sink

When comparing to static networks, link changes and node changes are more frequent in duty-cycled WSNs. If a node changes its duty-cycle configuration, or dynamically joins or leaves the network, the links connecting with all its neighbors will be changed at different time intervals. During such conditions occurring a single node update usually causes multiple link updates.Some previous works in static networks have proposed solutions that efficiently deal with single link updates. They are inefficient for multiple link updates caused by a single node update. The algorithms in are also memory-inefficient, since each node stores the route entries for all other nodes, increasing the space complexity. In dynamic Route Maintenance algorithm a node only stores the route to the sink, which is more efficient in WSNs due to their memory constraints.The updates of the shortest path are updated only from the sink to the destination and not on the entire collection of nodes.The node which is to be monitored is first identified . The updating involves creation of initial route and sending an acknowledgement to the sink.The algorithm runs on multiple node when there are multiple node updates.

Cite this article as: P.Raja. "Efficient Method for Identifying Shortest Path in Duty Cycled Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 60-69. Print.

III. SYSTEM DESIGN

A collection of Sensor nodes in a Sensing area can communicate among themselves and the shortest path is found out by the minimum spanning tree algorithm.The diagram shown below in Fig 3.1 gives a high level block diagram of Time Dependent Wireless Sensor Networks.

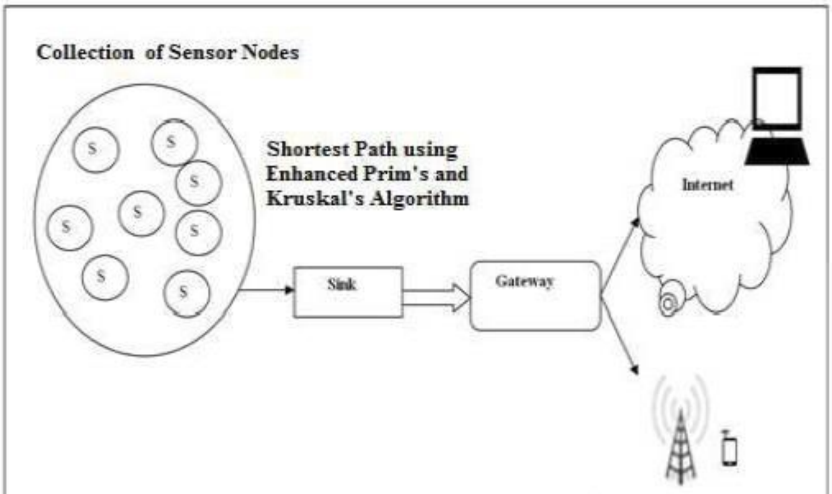


Fig 3.1 Overview of Fast Distributed Routing in WSN

Every node maintains the shortest path to the sink node and does not contain the routing information of other nodes. The sink node interacts through a gateway by which it can be connected with the Internet or transmit information through the Base Station which can be received by the mobile.The Dynamic addition and deletion of Nodes can be easily maintained by changing the initial route maintained.The sensor nodes can represent any of the states as shown in Fig 3.2

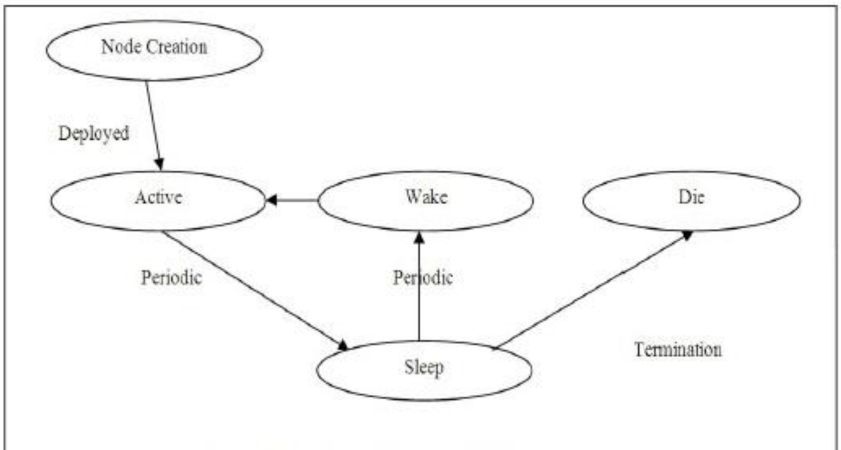


Fig 3.2 States of Sensor Node

The node is initially introduced in to the sensing system which begins to operate in datacentric manner depending on the application. It is active when a particular condition is to be sensed. Then all the nodes either enter in to sleep state or wake up state based on S-MAC protocol. The S-MAC protocol uniformly awakes all the nodes for a certain time period and also the nodes sleeps uniformly for a certain time period. The awake time can be chosen as 120 ms and sleep time can be taken as 2ms as shown in Fig 3.3.The sleep and wake up time period can be changed as per the type of application environment where the sensor node is deployed.A timer is set by the sleeping node and wakes up when it expires.It requires periodic synchronization among nodes to take care of clock drift

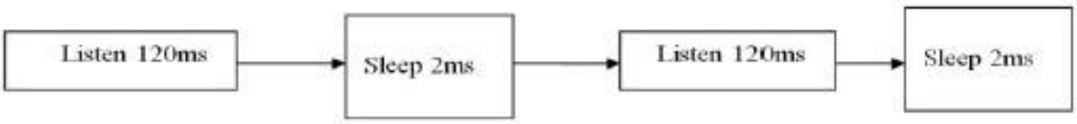


Fig 3.3 S-MAC Protocol

The sending of message can be by the S-MAC protocol can be represented by maintaining a vector Xi(t).When the value of Xi(t)=1 then the sensor node is active and message can be sent. When the sensor node enters in to sleep state it sets the vector as Xi(t)=0.So when a node wants to send a message to another node in a sensing area then it checks whether Xi(t) is One or Zero and sends message only when it is in awake state ie.) Xi(t)=1.If Xi(t)=0 then it waits for a time period say 120 ms as described in the above S-MAC protocol .Then using the Intial shortest path construction it sends the message. The below flowchart in Fig 3.4 explains the procedure more elaborately

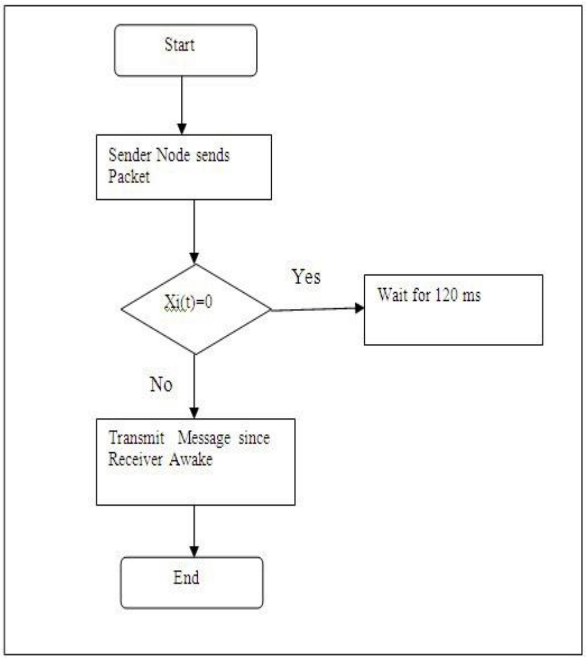
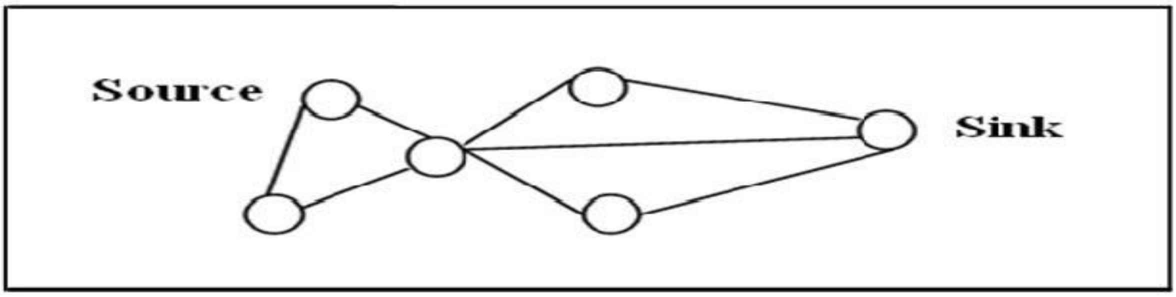


Fig 3.4 Flowchart for Transmitting Messages

The Sensor nodes duty cycle refers the ratio of the active period of the nodes to the full active or dormant period of the sensor node over a predefined time interval.The sensor nodes in a Sensing System can be modeled as directed graph and message can be transferred over the time dependent duty cycled sensor networks.When the message is transferred over the shortest path calculated there is least latency since the optimal path is used to transfer message across the nodes and also from nodes to the sink.Thus the optimal performance is achieved in a sensing system under consideration.The estimation of shortest path for routing in the architecture mentioned above leads to the refinement of Minimum Spanning Tree algorithm namely Prim's Algorithm ,Kruskal's Algorithm.Both of the modules are implemented and a comparison is made.Collision avoidance among the sensor nodes is also implemented.The type of routing followed in the above mentioned architecture is multicast routing.

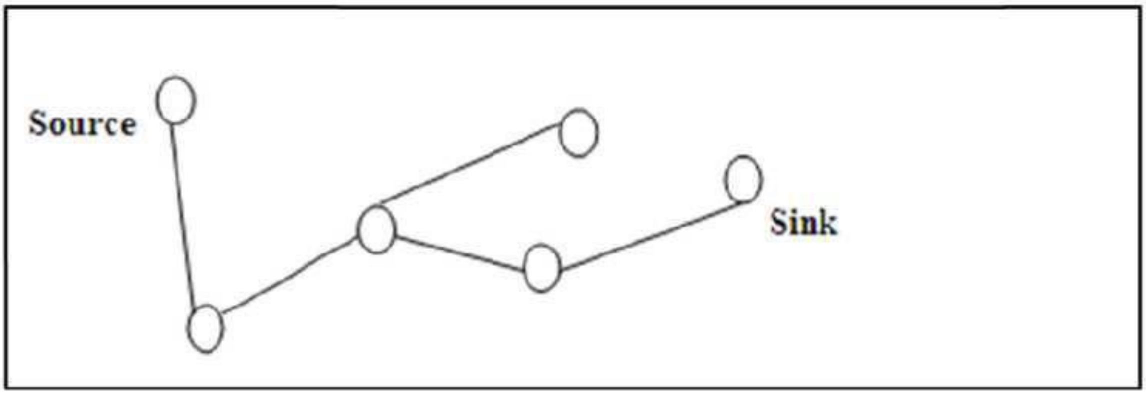
3.1 Enhanced Prim's Algorithm:

Prim's algorithm is used for finding the shortest path as the sensor nodes are added to the sensing area. In this algorithm the the array holds the node that have been added.Instead of manipulating the traditional Prim's Algorithm with weights in each edge the enhanced Prim's Algorithm works only by the addition of only the active nodes in a sensing area.The distance of every node from the sink is calculated by means of hopcount.During every iteration all of the sensor nodes that are part of the sensing area are checked is selected and then added to the array. It is quite clear that after each round the number of sensor nodes to be checked increases and since each node has a number of edges, we have a loop within a loop.For the last iteration the function must go through almost the entire sensing area.



Cite this article as: P.Raja. "Efficient Method for Identifying Shortest Path in Duty Cycled Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 60-69. Print.

The algorithm repeats until a minimum spanning tree is formed without loop and this gives the shortest path from the source node to the sink.The Fig 3.1.2 gives the structure of minimum spanning tree formed by enhanced Prim's algorithm



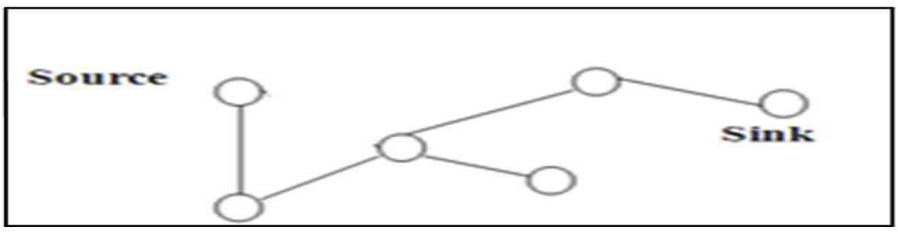
Prim's Algorithm

The time complexity for enhanced Prim's algorithm is O(E +N logN) depending on network structure.The number of edges(Communication Links)are represented by E and Sensor Nodes are represented by N.

A.Algorithm for Source Node in Enhanced Prim's Algorithm:
Initialization:
For any ni which belongs to Nsrc
Curr_Src=ni
ACK_SINK_SRC(i)=0;
Data_Value (i)=0
Data Transfer
If(Data_Value (i) = Sensing_Threshold)
{
then send MSG(src,dest,Hopcount,data) to sink along shortest path by Enhanced Prim Set_Timer(25);
}
If ACK_SINK not received after 25ms
then resend the MSG;
End Data Transfer

B.Algorithm for Sink Node in Enhanced Prim's
Algorithm: Initialization:
For every Nsink which belongs to a Sensing area of N nodes
ACK_SINK=0;
Response Message
If (Data_value(i) = valid_data)
Then send ACK_SINK =1;
Else
Send "Error_Message"

3.2 Enhanced Kruskal's Algorithm:
In this algorithm the the nodes may be in a active state or sleep state.Instead of considering the weights among the edges and sorting them only the active edges along two sensor nodes are identified.Then the number of hops needed to reach the sink is taken as the weight among the edges.The active link among the nodes and the edge with the minimum hop count to the sink is considered first and added to the sensor nodes.The edges are added among the sensor nodes until a minimum spanning tree is formed without a cycle.



Kruskal's Algorithm

The Fig 3.2.1 gives the Minimum Spanning Tree by
Enhanced Kruskal's algorithm for the sample WSN in Fig
3.1.1. The Time complexity of Kruskal's Algorithm is O(E+log N)
A.Algorithm for Source Node in Enhanced Kruskal's
Algorithm:
Initialization:
For all link Lis which belongs to Nsrc and Nneighbor along the shortest path
Lis=active;
ACK_SINK_SRC(i)=0;
Data_Value (i)=0
Data Transfer
If(Data_Value (i) = Sensing_Threshold and Lis=active)
{
then send MSG(src,dest,Hopcount,data)to sink along
shortest path by Enhanced Kruskal;
Set_Timer(25);
}
If ACK_SINK not received after 25ms
then resend the MSG;
End Data Transfer
B.Algorithm for Sink Node in Enhanced
Kruskal's Algorithm:
Initialization:
For every Nsink which belongs to a Sensing area of N nodes
ACK_SINK=0;
Response Message
If (Data_value(i) = valid_data)
Then send ACK_SINK =1;
Else
Send "Error_Message"

3.3 Collision Avoidance:

When more than one node tries to transfer data at the same time through the same channel then there is a possibility of collision and this leads to corruption of packets. It also leads to energy wastage among sensor nodes.This can be avoided by RTS/CTS mechanism.When a sensor node wants to transfer message the RTS message is sent to the Sink .When a RTS message is sent by a node it is overheard by all the nodes along the shortest path identified.So the overhearing nodes along the shortest path constructed will not send RTS message since there is some node already waiting to transfer the message.The Sink on receiving the RTS message transfer to the sink at the time of RTS being sent then the sink replies with CTS.Then the Source node transfers the message.After the entire message is transferred the next node sends the CTS message

A.Algorithm for Collision Avoidance in Source Node
: Initialization:
For every Ni in N nodes of sensing area
Send Request_To_Send_MSG(src,dest,hopcount,data) to
Sink
RTS[i]=1
Data Transfer
Multicast RTS[i]=1 from Source node to nodes along the
Shortest Path
Algorithm for Collision Avoidance in Sink Node :
Initialization:
For all i=1 toN in Sensing Area
{
CTS[i]=0
RTS[i]=0
}
Response Message
For all i=1 to N in Sensing Area
For all j=i+1 to N in Sensing Area

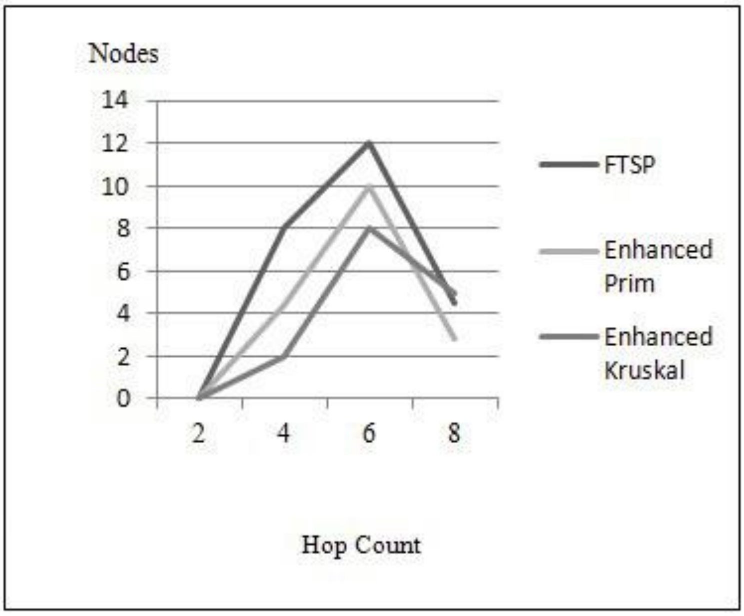
If(RTS[i] = 1 and Time_of_RTS[i] !=Time_of_RTS[j])
Then send CTS[i]=1
Else
Send "PKT_COLLSION_MSG"

IV. IMPLEMENTATION

The Prim's and Kruskal's Algorithm were implemented in NS-2 .The sensor nodes always function according to S-MAC protocol and the time taken to route according to the shortest path is noted.The number of sensor nodes taken in to consideration is 12. The algorithm is simulated using random placement of sensor nodes since the topology of sensor network always vary according to the need of the user.The distance between every sensor node is calculated by the Euclidean's distance formula|(x1-x2)2-(y1-y2)2|.The distance between every sensor node and the sink is displayed along with the vizualization of packet transmission along the shortest path.A distance of 150 meters is taken as the simulation area for Wireless Sensor Network.When the distance calculated between the sensor nodes exceeds 150 meters then an error message is displayed.This criteria is achieved by setting the threshold value as 150 meters.Destination Sequence Distance Vector Routing (DSDV) algorithm is used .The files for implementing DSDV is correspondingly modified in the simulation tool. The shortest path for a specific discrete time interval is simulated and the trace file values are gathered.

V. EVALUATION

A comparison is made between the Enhanced Prim's,Enhanced Kruskal's and Fast Time Dependent Shortest Path(FTSP)Algorithm.The time taken by enhanced Prim algorithm to transfer a message is the least since it works well when the number of nodes are added dynamically to a Wireless Sensor Network.In the enhanced Kruskal algorithm the efficiency increases as the number of links between the sensor nodes increases.This algorithm is best examines if it can receive the message.If there are no data suited when all the messages are broadcasted among themselves.FTSP takes more time to route compared with the Prim and Kruskal's algorithm.The metrics used in evaluation are as follows
A. Hop Value corresponds the routing distance of nodes throughout the network which considers the end to end latency among the sensor nodes and the energy consumption among them.
B. Data Reliability measures the link quality along the shortest path discovered from each node in the network.It accounts for the end to end message transfer in a Wireless Sensor Network without retransmission.
C. Transmission Rate refers to the ratio of number of packets received at the sink for a particular node to the number of packets actually originated at that node.A value of one indicates that the network is highly reliable since all the messages are received at the processing unit without any loss.
5.1 Analysis of Network by Graph The Wireless Sensor Network of 12 nodes organized as a grid of 100x100 with a random placement of nodes is analysed by means of Graph.The Hop count value is tabulated for the three algorithms as shown in the Fig 5.1.1 which conveys the information that enhanced Kruskal's algorithm is 50 % more efficient compared to FTSP .The enhanced Prim's algorithm is 75% more efficient compared to the existing FTSP as the number of nodes increases



The next metric taken for analysis is Data Reliability.The values of data transferred from the source to the sink can be tabulated in a graph as shown in the Fig 5.1.2 .It can be concluded from the graph that enhaced Prim's and Kruskal's algorithm are 50% more reliable than the existing FTSP algorithm.

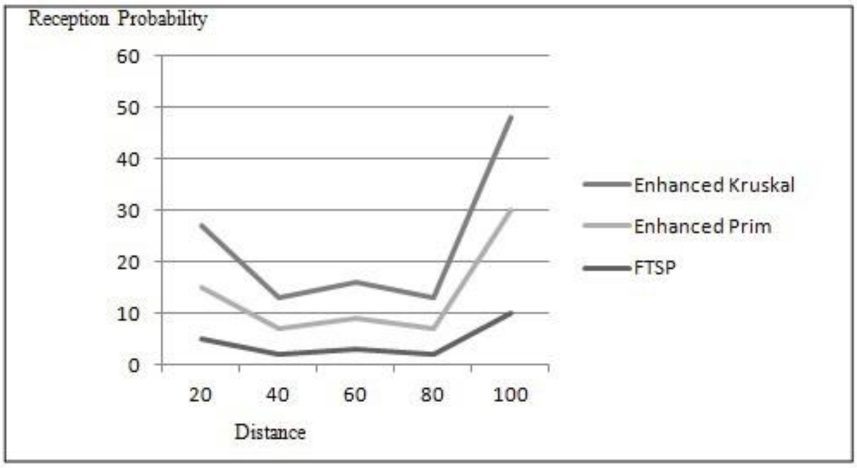
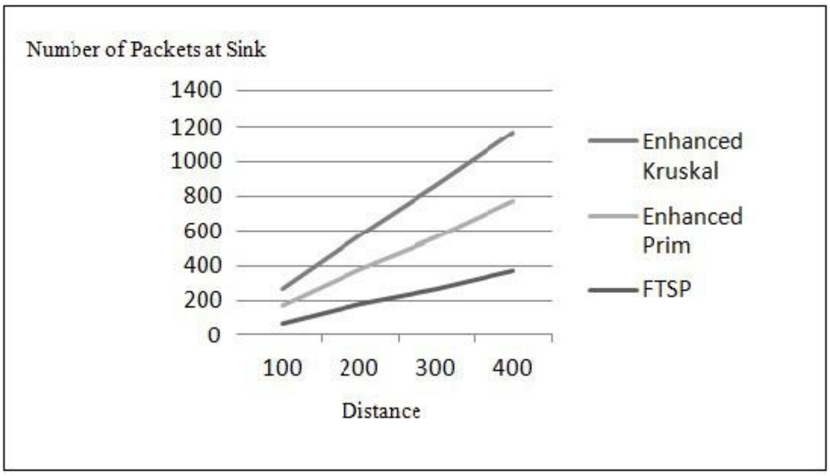


Fig 5.1.2 Data Reliability Analysis

The Transmission Rate can be tabulated in a graph by measuring the total number of packets received at the sink.



It depends on the congestion of the packets in a network and the time period of transfer.Based on the analysis from the Fig 5.1.3 it can be concluded that the Transmission Rate is maximum for Kruskal followed by Prim and finally the FTSP algorithm.Therefore the enhanced Prim's and Kruskal's algorithm are almost 75% more efficient than the existing FTSP algorithm.

VI. CONCLUSION AND FUTURE WORK

In this paper we addressed the various types of routing in Duty Cycled Wireless Sensor Networks.The strengths and limitations of various methods are surveyed.The best method to achieve least latency is to construct the shortest path using minimum spanning tree algorithm.The enhanced Prim's and Kruskal's algorithm can be used and a comparison can be done based on time and number of sensor nodes in the sensing area.The security of data transferred by multicast routing can be a problem and it is to be addressed in future.The failure of sink posses a threat since the data cannot be processed and some alternate solution needs to be proposed.

REFERENCES

[1] Shouwen Lai,Binoy Ravindran," Least-Latency Routing over Time-Dependent Wireless Sensor Networks, IEEE TRANSACTIONS ON COMPUTERS, VOL. 62, NO. 5, MAY 2013.

[2] Shouwen Lai, Binoy Ravindran "On Distributed Time-Dependent Shortest Paths over Duty-Cycled Wireless Sensor Networks", Department of Electrical and Computer EngineeringVirginia Tech,Blacksburg Technical Program at IEEE INFOCOM 2010.

[3] Raja Jurdak, Pierre Baldi, and Cristina Videira Lopes," Adaptive Low Power Listening for Wireless Sensor Networks" , IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 6, NO. 8, AUGUST 2007

Cite this article as: P.Raja. "Efficient Method for Identifying Shortest Path in Duty Cycled Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 60-69. Print.

[4] A. Woo, T. Tong, and D. Culler, "Taming the Underlying Challenges of Reliable Multihop Routing in Sensor Networks,"Proc. First Int'l Conf. Embedded Networked Sensor Systems (Sensys),pp. 14-27, 2003.

[5] C. Schurgers and M.B. Srivastava, "Energy Efficient Routing in Wireless Sensor Networks," Proc. Military Comm. Conf. (MILCOM'01), vol.1, pp. 357-361, 2001.

[6] Y. Gu and T. He, "Data Forwarding in Extremely Low Duty-Cycle Sensor Networks with Unreliable Communication Links," Proc.Sixth ACM Conf. Embedded Network Sensor Systems (Sensys), pp. 32-38, 2007.

[7] H. Chon, D. Agrawa, and A. Abbadi, "Fates: Finding a Time Dependent Shortest Path," Mobile Data Management, vol. 2574,pp. 165-180, 2003.

[8] J. Hui and D. Culler. The Dynamic Behavior of a Data Dissemination Protocol for Network Programming at Scale. In SenSys'04, 2004.

[9] A. Keshavarzian, H. Lee, and L. Venkatraman. Wakeup Scheduling in Wireless Sensor Networks. In MobiHoc '06, 2006.

Cite this article as: P.Raja. "Efficient Method for Identifying Shortest Path in Duty Cycled Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 60-69. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS011		**eAID**	ICIEMS.2015.011

CLOUD DATA PROTECTION FOR THE MASSESS

V.Prasanth[1], B.Ajay[2], R.Nijanthan[3]

[1]Dept. of computer science and engineering, Velammal Institute of technology,Chennai

[2]Dept. of computer science and engineering, S.R.M. University,Chennai

[3]Dept. of computer science and engineering, Velammal Institute of technology,Chennai,

Abstract: *Offering strong data protection to cloud users while enabling rich applications is a challenging task. We explore a new cloud platform architecture called Data Protection as a Service, which dramatically reduces the per-application development effort required to offer data protection, while still allowing rapid development and maintenance.*

Intoduction

Although cloud computing promises lower costs, rapid scaling, easier maintenance, and service availability anywhere, anytime, a key challenge is how to ensure and build confidence that the cloud can handle user data securely. A recent Microsoft survey found that ―58 percent of the public and 86 per-cent of business leaders are excited about the possibilities of cloud computing. But more than 90 percent of them are worried about security, availability, and privacy of their data as it rests in the cloud.‖1 This tension makes sense: users want to maintain con-trol of their data, but they also want to benefit from the rich services that application developers can provide using that data. So far, the cloud offers little platform-level support or standardization for user data protection beyond data en-cryption at rest, most likely because doing so is nontrivial.

Protecting user data while enabling rich computation re-quires both specialized expertise and resources that might not be readily available to most application developers.

Building in data-protection solutions at the platform layer is an attractive option: the platform can achieve economies of scale by amortizing expertise costs and dis-tributing sophisticated security solutions across different applications and their developers.

WHAT ABOUT ENCRYPTION?

In the realm of data protection, *full-disk encryption* (FDE) and computing on encrypted data have recently gained attention, these techniques have fallen short of answering all of the security and maintenance challenges mentioned earlier.FDE encrypts entire physical disks with a symmetric key, often in disk firmware, for simplicity and speed. At the other end of the spectrum, Craig Gentry re- cently proposed the first realization of *fully homomorphic encryption* (FHE),[2] which offers the promise of general computation on ciphertexts. Basically, any function in plaintext can be transformed into an equivalent function in ciphertext: the server does the real work, but it doesŶ't kŶow the data it's ĐoŵputiŶg. Naturally, this property gives strong privacy guarantees when computing on private data, but the question of its practicality for general cloud applications still remains.

Cite this article as: V.Prasanth, B.Ajay, R.Nijanthan. "CLOUD DATA PROTECTION FOR THE MASSESS." *International Conference on Information Engineering, Management and Security (2015)*: 70-73. Print.

1.FDE versus FHE

A comparison of FDE and FHE in the cloud comput-ing setting reveals how these encryption techniques fall short of addressing the aforementioned security and maintenance challenges simultaneously.

1.1 Key management and trust.

With FDE, the keys reside with the cloud platform, generally on or close to the physic cal drive: the cloud application user isn't involved in key management. While user data is encrypted on the physical disk, it is always accessible in the clear to any layer above it. Consequently, FDE doesn't prevent online attacks from leaking the data to an unauthorized party, which is far more common in the cloud setting than physical attacks. With FHE, untrusted applications can't easily learn or leak data. Users typically own and manage FHE encryp-tion keys, while applications compute on encrypted forms of user data without actually seeing‖ the data. This raises questions about how users can store their keys securely and reliably, especially in the presence of sharing. After all, the point of the cloud is to avoid maintaining local state.

1.2 Sharing.

Collaboration is often cited as a ―killer feature‖ for cloud applications. Fine-grained access control is nec-essary to let a data owner selectively share one or more data objects with other users.

With FDE, users must fully trust the cloud provider to enforce correct access control because the key granularity (the whole disk) doesn't line up with access control granu-larity (a single data unit).

With FHE, because the user—or a third-party cloud pro-vider employed by the user—manages the encryption keys, the best way of providing access control isn't clear yet. To offer fine-grained encryption-based access control, we might need to define key management on a per data object granularity basis or over collections of data objects. How-ever, to support homomorphic operations across multiple encrypted objects, those objects must still be encrypted under the same public key.

1.3 Aggregation.

Many cloud applications require per-forming data mining over multiple users' data for tasks such as spam filtering or computing aggregate statistics. Because users fully trust the cloud provider, performing such data aggregation is relatively easy with FDE.

Current FHE techniques don't readily allow computing on multiple users' data encrypted under different keys. Therefore, it isn't clear yet how to support such data aggre-gation applications with FHE; similarly, offline aggregation across users' data isn't possible. One solution might be to escrow keys to the cloud provider, but that would eliminate many of FHE's benefits, making its cost harder to justify.

1.4 Performance.

According to a recent survey, 49 percent of users abandon a site or switch to a competitor after experiencing performance issues.[3] And the need for speed is only increasing: in 2000, a typical user was willing to wait 8 seconds for a webpage to load before navigating away; by 2009, that number dropped to 3 seconds.

When FDE is implemented in disk firmware, its sym-metric encryption can run at the disk's full bandwidth, effectively avoiding a slowdown. Although researchers have made significant advances in improving FHE's per-formance since Gentry's original proposal, it has a long way to go before becoming efficient enough to deploy at scale. In Gentry's estimation, implementing something like a Google search with FHE would require roughly 1 trillion times more computation than the one without FHE.[4]

1.5 Ease of development.

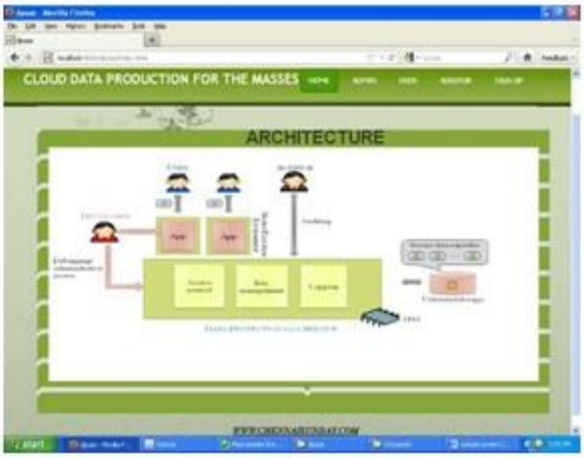
Because FDE is hidden behind an abstraction of the physical disk, it typically has no impact on application development. In theory, FHE could also be relatively automatic: it works on an abstraction of the program as a circuit and transforms that circuit. In practice, however, performing this translation for arbitrary programs—especially when marshaling data could be quite complex. At a minimum, programming tools would need to evolve dramatically.

2. Splitting the difference

Although FDE offers excellent performance and ease of development, it does little to protect privacy at the required granularity. FHE, on the other hand, pushes the privacy envelope in the other direction by removing data visibility entirely from both the server and application developer. However, having a remote machine see and compute on sensitive data isn't automatically a privacy violation. FHE's guarantees go beyond what's necessary to protect data, and in so doing, it incurs significant performance and de-velopment costs. We believe the DPaaS approach is better suited for the target applications because it falls between the two. It keeps the ―natural‖ granularity of FHE by keying on units of sharable data and maintains the performance of FDE by using symmetric encryption. It moves key manage-ment and access control to a middle tier—the computing platform—to balance rapid development and easy main-tenance with user- side verifiability.

A WAY FORWARD

In an OS, processes and files are the primary units of access control, and the OS provides suitable isolation for these boundaries.



In a cloud setting, the unit of access control is typically a sharable piece of user data—for example, a document in a collaborative editor. Ideally, the system offers some analogous confinement of that data, restricting its visibility only to authorized users and applications while allowing broad latitude for what operations are done on it. This can make writing secure systems easier for programmers because confinement makes it more difficult for buggy code to leak data or for compromised code to grant unau- thorized access to data. A malicious program might find different ways to exfiltrate data, such as employing a side channel or covert channel, but the priority here is to sup-port benign developers, while making all applications and their actions on users' sensitive data more easily auditable to catch improper usage.

One of the main concerns people and organizations have about putting data in the cloud is that they don't know what happens to it. Having a clear audit trail of when data is accessed—and by whom or what— bolsters confidence that data is being handled appropriately. Con-finement can be effective for most normal user accesses, but administrative access that's outside the normal flow of user access and involves human administrators (for example, for debugging and analysis) can especially ben-efit from auditing.

3. Verifiable platform support

Bugs need to be fixed. Data needs to be updated and amigrated as schemas change. Offline computation is valuable for data aggregation across users or for precom-putation of expensive functions. To reduce the risk of unaudited backdoor access, all these functions should be subject to the same authorization flows and platform-level checks as normal requests, albeit with a separate, appropriate policy.

Platform providers should build support for confine-ment and auditing into the platform in a verifiable way. This approval has many advantages:

- application developers don't have to reinvent the wheel;
- application code is independent of ACL enforcement;
- third-party auditing and standards compliance are easier; and
- the verifiable platform extends to virtualized environ-ments built atop it.
- Finally, the cost of examining the platform is amortized across all its users, which means significant economies of scale.

4. Design space and a sample architecture

Figure 1 illustrates an example architecture for explor-ing the DPaaS design space.[5] Here, each server contains a *trusted platform module* (TPM) to provide secure and verifiable boot and dynamic root of trust. This example ar-chitecture demonstrates at a high level how it's potentially possible to combine various technologies such as applica-tion confinement, encryption, logging, code attestation, and information flow checking to realize DPaaS.

4.1 Confinement.

A *secure data capsule* (SDC) is an encrypted data unit packaged with its security policy. For example, an SDC might encompass a sharable docu-ment or a photo album along with its ACL. The platform can use confinement and information-flow controls to enforce capsules' ACLs.

4.2 Audit trails.

Because the platform mediates all data access, authenticates users, and runs binaries, it knows what data is accessed by what user, and with which appli-cation. It can generate meaningful audit logs containing all these parameters and optionally incorporate additional information from the application layer.

Cite this article as: V.Prasanth, B.Ajay, R.Nijanthan. "CLOUD DATA PROTECTION FOR THE MASSESS." *International Conference on Information Engineering, Management and Security (2015)*: 70-73. Print.

4.3 Platform verifiability.

The DPaaS approach provides logging and auditing at the platform level, sharing the benefits with all applications running on top. Offline, the auditor can verify that the platform implements each data protection feature as promised. At runtime, the platform provider can use *trusted computing* (TC) technologies to attest to the particular software that's running. TC uses the tamperproof TPM as well as the virtualization and isolation features of modern processors, such as Intel VT or AMDV.

5.Achieving data protection goals

We assume in the analysis that the platform behaves correctly with respect to code loading, authorization, and key management, and that the TPM facilitates a runtime attestation to this effect. DPaas uses a combination of encryption at rest, ap-plication confinement, information flow checking, and auditing to ensure the security and privacy of users' data. Application confinement isolates faults and compromises within each SEE, while information flow checking ensures that any information flowing among SEEs, data capsules, and users satisfies access-control policies. Controlling and auditing administrative accesses to data provides account-ability. DPaaS can guarantee the integrity of the data at rest via cryptographic authentication of the data in storage and by auditing the application code at runtime.

Access controls, authorization, and auditing capability are common challenges for application developers. Incor-porating these features within the platform is a significant improvement in terms of ease of use, and it doesn't con-strain the types of computation that can be performed within a SEE. The platform logs common maintenance and batch processing tasks to provide accountability. These tasks too often require one-off work in the development process and can benefit from standardization.

CONCLUSION

As private data moves online, the need to secure it properly becomes increasingly urgent. The good news is that the same forces concentrating data in enormous datacenters will also aid in using collective security expertise more effectively. Adding protections to a single cloud platform can immediately benefit hundreds of thousands of applications and, by extension, hundreds of millions of users. While we have focused here on a particular, albeit popular and privacy-sensitive, classes of applications, many other applications also need solutions.

REFERENCES

[1] C. Dwork,―The Differential Privacy Frontier Extended Abstract,‖ *Proc. 6th Theory of Cryptography Conf.* (TCC 09), LNCS 5444, Springer, 2009, pp. 496-502.

[2] C. Gentry,―Fully Homomorphic Encryption Using Ideal Lattices,‖ *Proc. 41st Ann. ACM Symp. Theory Computing* (STOC 09), ACM, 2009, pp. 169-178.

[3] E. Naone,―The Slow-Motion Internet,‖ *Technology Rev.*,Mar./Apr. 2011; www.technologyreview.com/files/54902/Google Speed_charts.pdf.

Cite this article as: V.Prasanth, B.Ajay, R.Nijanthan. "CLOUD DATA PROTECTION FOR THE MASSESS." *International Conference on Information Engineering, Management and Security (2015)*: 70-73. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	VOL	01
Website	www.iciems.in	eMail	iciems@asdf.res.in
Received	10 - July - 2015	Accepted	31- July - 2015
Article ID	ICIEMS012	eAID	ICIEMS.2015.012

BRAIN WAVE CONTROLLER FOR STRESS REMOVAL AND AUTOMATION OF AUTOMOBILE IGNITION TO PREVENT DRIVING UNDER INFLUENCE

[CONTROLLING BRAINWAVES USING EMBEDDED SYSTEMS]

K.KABILAN
Department of Electronics and Communication Engineering
Nandha College of Technology, Erode

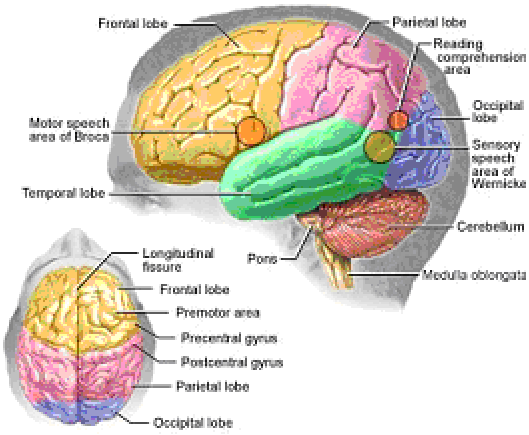
Abstract: *Stress is a prevalent and costly problem in today's workplace. It is the harmful physical and emotional response that occurs when there is a poor match between job demands and the capabilities, resources, or needs of the worker. Persistence of stress results in cardiovascular disease such as depression, concentration and memory loss. Addiction is one of the chronic disorders that are characterized by the repeated use of substances or behaviors despite clear evidence of morbidity secondary to such use. It is a combination of genetic, biological / pharmacological and social factors. Example: gambling, alcohol drinking, taking narcotic drugs and certain mannerisms. The therapies at present consume time. About 24% of the accidents taking place are due to drunken drive. A driver subjected to long drive falls sleepy and ends up in accident. In this paper we briefly discuss about the brain wave and brains reaction during stress, addiction and drunk. This paper also explains you the basic task of Brainwave Controller, that how stress, addiction is identified with the help of brainwave and how these are controlled using the principle binaural beats. Also we have designed a device to detect the brainwaves and process it to determine whether it is addiction or stress. In addition to controlling of brainwaves, it also has a feature to avoid an individual who consumes alcohol to drive a vehicle. This paper promises to be an economical solution for the people who suffer from stress, addiction and to prevent accidents.*

I. Introduction

Driving Under Influence (DUI):

Driving under the influence of alcohol (operating under the influence, drinking and driving, impaired driving) or other drugs is the act of operating a vehicle (including boat, airplane, or tractor) after consuming alcohol or other drugs. DUI or DWI are synonymous terms that represent the criminal offense of operating (or in some jurisdictions merely being in physical control of) a motor vehicle while being under the influence of alcohol or drugs or a combination of both. It is a criminal offense in most countries as it contributes to some of the major accidents. Now let us consider few basic concepts upon which this project is largely based upon.

Cite this article as: K.KABILAN. "BRAIN WAVE CONTROLLER FOR STRESS REMOVAL AND AUTOMATION OF AUTOMOBILE IGNITION TO PREVENT DRIVING UNDER INFLUENCE." *International Conference on Information Engineering, Management and Security (2015)*: 74-80. Print.



Blood alcohol content or blood alcohol concentration (abbreviated BAC) is the concentration of alcohol in a person's blood. BAC is most commonly used as a metric of intoxication for legal or medical purposes. It is usually expressed in terms of volume of alcohol per volume of blood in the body. That is a unit-less ratio commonly expressed as parts per million (PPM) or as a fractional percentage. That is a decimal with 2-3 significant digits followed by a percentage sign, which means 1/100 of the previous number (E.g., 0.0008 expressed as a percentage as 0.08%). Since measurement must be accurate and inexpensive, several measurement techniques are used as proxies to approximate the true parts per million measure.
Some of the most common are listed here: (1) Volume of alcohol per volume of exhaled breath (E.g. 0.08 mL/L), (2) Mass per volume of blood in the body (E.g.: 0.08 g/L), and (3) Mass of alcohol per mass of the body (E.g.: 0.08 g/Kg). After one drink you reach your peak after 30 minutes and you should wait a few hours before you drive. The number of drinks consumed is often a poor measure of BAC, largely because of variations in weight, sex, and body fat.

Ignition Interlock System:

An ignition interlock device or breath alcohol ignition interlock device (IID and BIID) is a mechanism, like a breathalyzer, installed to a motor vehicle's dashboard. Before the vehicle's motor can be started, the driver first must exhale into the device, if the resultant breath-alcohol concentration analyzed result is greater than the programmed blood alcohol concentration, usually 0.02% or 0.04%, the device prevents the engine from being started. At random times after the engine has been started, the IID will require another breath sample. The purpose of this is to prevent a friend from breathing into the device, enabling the intoxicated person to get behind the wheel and drive away. If the breath sample isn't provided, or the sample exceeds the ignition interlock's preset blood alcohol level, the device will log the event, warn the driver and then start up an alarm (e.g., lights flashing, horn honking, etc.) until the ignition is turned off, or a clean breath sample has been provided. A common misconception is that interlock devices will simply turn off the engine if alcohol is detected; this would, however, create an unsafe driving situation and expose interlock manufacturers to considerable lia
bility. An interlock device cannot turn off a running vehicle, all that an Interlock device can do is interrupt the starter circuit and prevent the engine from starting…

II. PRINCIPLE

The principle behind this device is *Binaural Beats*. Binaural beats or binaural tones are auditory processing artifacts, which are apparent sounds, the perception of which arises in the brain independent of physical stimuli. The brain produces a similar phenomenon internally, resulting in low-frequency pulsations in the loudness of a perceived sound when two tones at slightly different frequencies are presented separately, one to each of a subject's ears, using stereo headphones. A beating tone will be perceived, as if the two tones mixed naturally, out of the brain. The frequency of the tones must be below about 1,000 to 1,500 hertz. The difference between the two frequencies must be small (below about 30 Hz) for the effect to occur; otherwise the two tones will be distinguishable and no beat will be perceived.

Binaural beats can influence functions of the brain besides those related to hearing. This phenomenon is called frequency following response (FFR). The concept is that if one receives a stimulus with a frequency in the range of brain waves, the predominant brain wave frequency is said to be likely to move towards the frequency of the stimulus (a process called entrainment). Human hearing is limited to the range of frequencies from 20 Hz to 20,000 Hz, while the frequencies of human brain waves are below about 40 Hz. To account for this, binaural beat frequencies must be used.

According to this view, when the perceived beat frequency corresponds to any of the brainwave frequencies, the brain entrains to or moves towards the beat frequency. For example, if a 315 Hz sine wave is played into the right ear and a 325 Hz one into the left ear,

Cite this article as: K.KABILAN. "BRAIN WAVE CONTROLLER FOR STRESS REMOVAL AND AUTOMATION OF AUTOMOBILE IGNITION TO PREVENT DRIVING UNDER INFLUENCE." *International Conference on Information Engineering, Management and Security (2015)*: 74-80. Print.

the brain is supposed to be entrained towards the beat frequency 10 Hz. Alpha range is usually associated with relaxation, this is supposed to have a relaxing effect. Some people find pure sine waves or *pink noise* unpleasant, so background music (e.g. natural sounds such as river noises) can also be mixed with them.

III.CONSTRUCTION AND WORKING

A. *Block diagram*

The General block diagram of controlling addiction/stress is shown in the figure 4.

Fig. 4 General block diagram to control addiction / stress

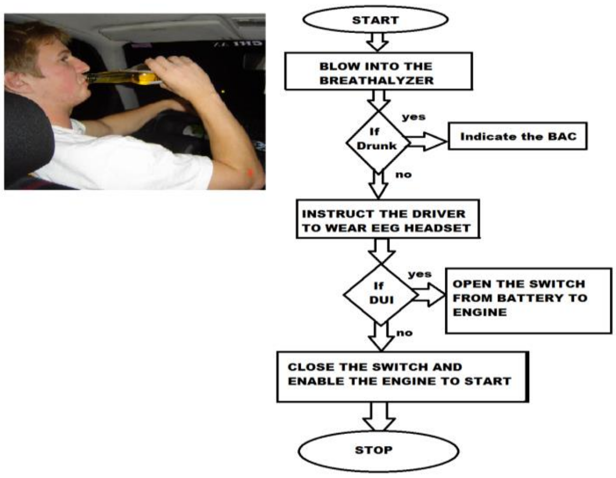


Fig. 5 General block diagram to avoid DUI

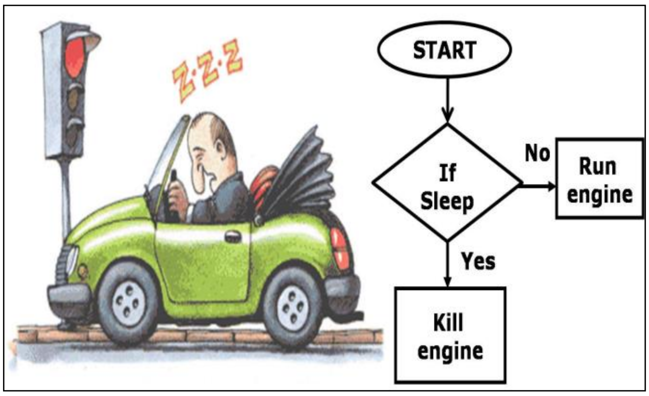
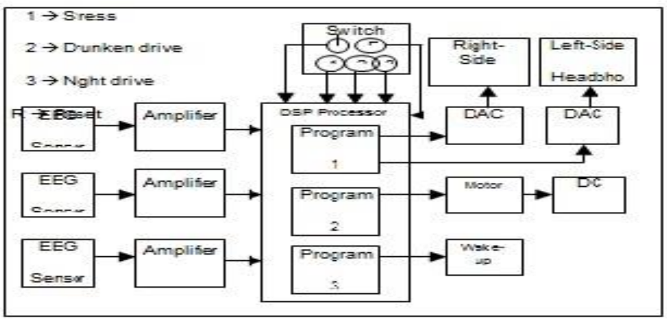


Fig. 6 General block diagram to avoid accidents

The block diagram used in implementation of brainwave controller with all its modes is given in figure 7.



Cite this article as: K.KABILAN. "BRAIN WAVE CONTROLLER FOR STRESS REMOVAL AND AUTOMATION OF AUTOMOBILE IGNITION TO PREVENT DRIVING UNDER INFLUENCE." *International Conference on Information Engineering, Management and Security (2015)*: 74-80. Print.

Fig. 7 General block diagram of Brainwave controller

B.EEG Sensors

EEG sensors is used to measure the electrical equivalent signal of brain wave.It consist of a 0.7 inch diameter hard plastic outer disc housing with a pre-jelled Silver chloride snap style post pellet insert. These sensors do not contain any latex. Figure 8 shows the representation of Ag/Agcl EEG sensor.

Fig. 8 Electroencephalography (EEG) sensors

Fig.9 EEG Fabricated Headset

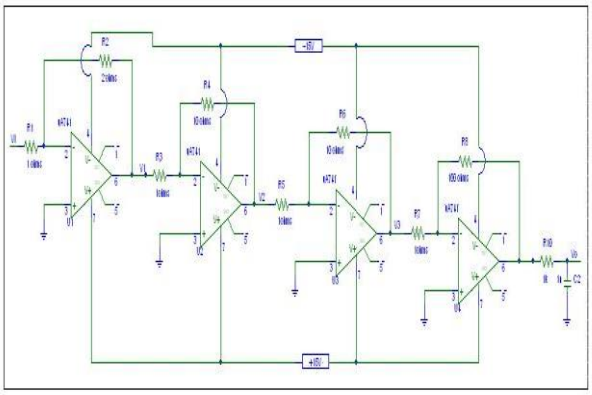
The sensor sends the analog brainwave signal into the instrumentation amplifier circuit.

B. Instrumentation Amplifiers

The amplitude of analog brainwaves is in between the range of 150 – 250 micro volts (µV). This is very low. For processing, at least amplitude above 2 volt is needed. For this a high gain and low noise amplifier is needed. For that instrumentation amplifier with high gain and high CMRR ratio is employed.

Here one operational is not enough to produce this much high gain. So a series of amplifier is cascaded to give required gain. The gain of an individual inverting operational amplifier is given by.

$$\text{Gain (A)} = -R_2/R_1$$



Cite this article as: K.KABILAN. "BRAIN WAVE CONTROLLER FOR STRESS REMOVAL AND AUTOMATION OF AUTOMOBILE IGNITION TO PREVENT DRIVING UNDER INFLUENCE." *International Conference on Information Engineering, Management and Security (2015)*: 74-80. Print.

Fig.9 Circuit diagram of cascaded inverting amplifier with a gain of 20,000. (Designed in Pspice)

Here we are using four inverting amplifier cascaded as shown in figure 9. Let the gain of each inverting amplifier from left to right be A_1, A_2, A_3 and A_4. And let V_i and V_o be the input and output voltages of the amplifier.

Now,

$A_1 = (-R_2/R_1)$
$= (-2/1)$
$= -2$

$A_2 = (-R_4/R_3)$
$= (-10/1)$
$= -10$

$A_3 = (-R_6/R_5)$
$= (-10/1)$
$= -10$

$A_4 = (-R_8/R_7)$
$= (-100/1)$
$= -100$

*all resistors are in kilo ohm

Now Total Gain of the amplifier (A_{eff}),

$A_{eff} = A_1 * A_2 * A_3 * A_4$
$A_{eff} = (-2)*(-10)*(-10)*(-100)$
$A_{eff} = 20,000$

Therefore,

$V_o = Vi * A_{eff} * V_i$
$= 15 * 10^{-5} * 20,000$ V
$V_o = 3\ V_i$

Hence an amplifier with gain 20,000 is designed using basic operational amplifier.

B. DSP Processor

The Amplified EEG signal is given to TMS320C6713DSP Processor where it has a pre-defined program to select a range of frequency from 1 – 50 Hertz. Then that selected range is converted from analog to digital samples. The DSP processor has a definite program according to different modes selected from the switch. There are 3 modes with which we can use this device.

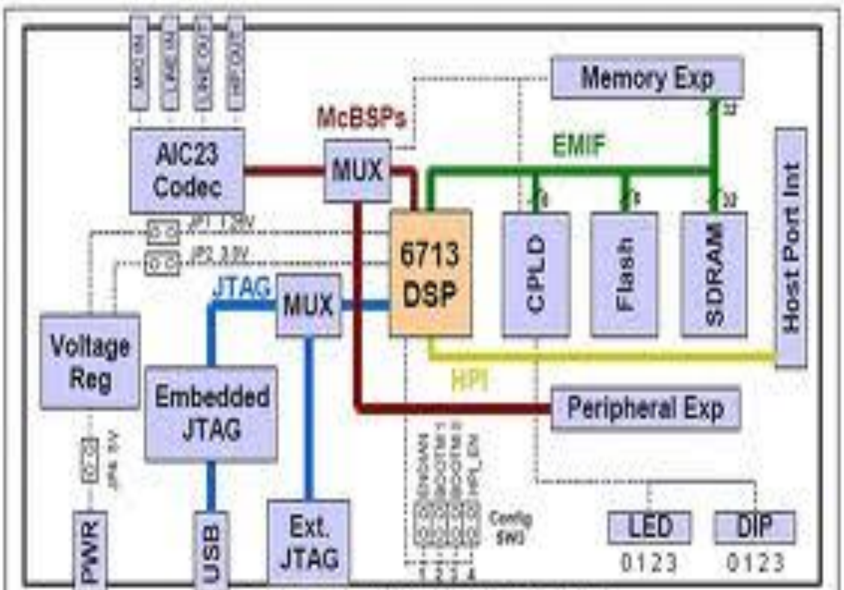
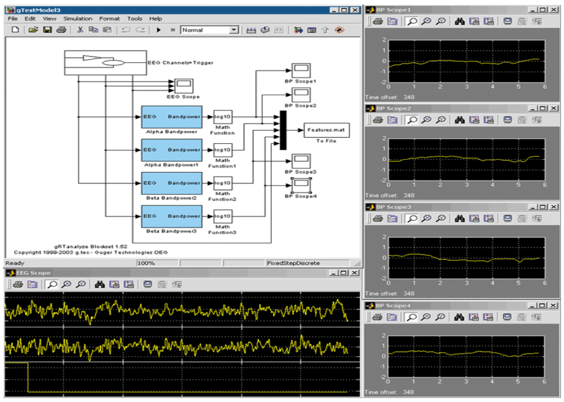
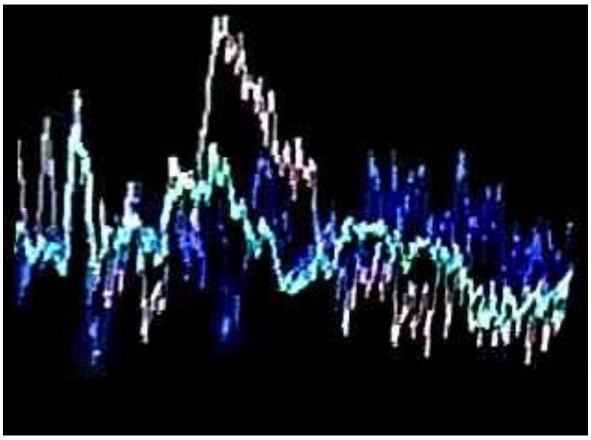


Figure 1-1, Block Diagram C6713 DSK

E. Simulation:

The following simulation images indicate the recorded brainwave using 2 EEG electrodes:





1. ***Mode - 1: Controlling addiction / stress:***

In mode 1 the DSP processor checks for the frequency range between 32 – 40 Hertz. If the range is between 32 – 40 Hertz, it is considered that the person is under addiction/stress and the processor runs a look-up table which contains the digital samples of binaural beats. The samples produce sine wave with a difference of 10 Hertz. These two waves are sent to each side of headphone.

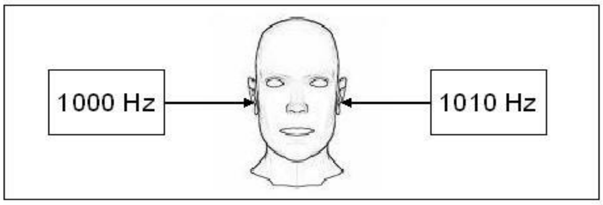


Fig.10 Block Diagram – Sending two similar tones with difference in frequency.

This generates the binaural beats. This is given to the ears . As the difference in these two waves is 10 Hz which is below 20 Hz, it cannot be detected by the human ear. But there is a neuron called afferon neuron inside the ear which senses this 10 Hz and sends it to brain as a stimulus. Now this stimulus entrains the brain to generate a stimulus of brainwaves similar to the supplied stimulus thereby reducing the brainwaves from 32 – 40 Hz to 9 – 14 Hz making the mind relaxed.

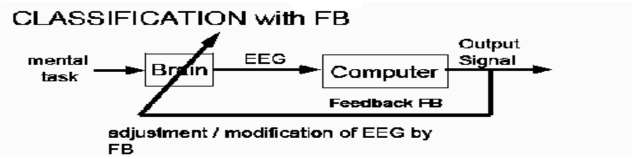
2. ***Mode – 2: Avoiding DUI*** :

When the switch is turned to mode 2, the device is connected to internal circuitry of a vehicle. Firstly, an Ignition Interlock Device (IID) is placed in an automobile and it is made mandatory for the driver to exhale in to the breathalyzer to switch on the automobile's engine. Here, the Blood Alcohol Level (BAC) of the driver is analyzed and in case of high BAC the engine does not start. If BAC is low, then a specially fabricated Electroencephalograph (EEG) headset (which contains EEG sensors according to international 10 -20 system) should be placed in the driver's head to analyze the driver's brainwaves. By analyzing the driver's brainwaves, the risk of driving under the influence of drugs is reduced. Here the BAC is measured as per the international threshold value of 0.04ml/L. Once the BAC is declared low, then the driver has to take up the drug test, which involves the usage of EEG headset in order to detect the brainwave activity of the driver. Unless and otherwise the driver has his brain wave levels at alpha or beta mode, the engine will not start.

Cite this article as: K.KABILAN. "BRAIN WAVE CONTROLLER FOR STRESS REMOVAL AND AUTOMATION OF AUTOMOBILE IGNITION TO PREVENT DRIVING UNDER INFLUENCE." *International Conference on Information Engineering, Management and Security (2015)*: 74-80. Print.

*3. **Mode – 3: Night drive:***

When the switch is turned to mode 3, the device is used to avoid sleep while long night drive. Here the individual must wear a cap at all times. This cap is embedded with EEG sensors as explained in mode 2. Here too DSP processor checks whether the signal has frequency below 7 Hertz which means the individual is nearly asleep. If such is the case, the processor triggers an alarm, so that accidents can be avoided. This alarm could be of any form. For example it could be horns honking or the audio system playing loud music or enabling specific alarm device to perform the waking up operation.



IV. Conclusion

Firstly, the brainwaves are controlled by the principle of binaural beats and frequency following response thereby controlling addiction or stress by making the mind relaxed temporarily. Secondly, the brainwaves are continuously monitored to avoid drunken drive in a vehicle. Thirdly, the brainwaves are continuously monitored to falling asleep while driving long distance in a vehicle. Though all the above applications are discreet to each other, it is absolutely useful to use a same device for all the three purposes.

V. Merits

1. The whole device is light weight and can be carried anywhere.
2. The whole device including sensors and headphone is cheap and costs only about Rs. 1500 and slightly above.
3.

VI. Demerits

Those meeting any of the following criteria should not use binaural beats

1. Epileptics
2. Pregnant women
3. People susceptible to seizures
4. Pacemaker users
5. Photosensitive people.

VII. Future enhancements

1. The concept of *frequency following response* can be further researched to ease communication with deaf and dumb individuals.
2. The concept of *binaural beats* can be further used to study the resonance of brain during brain diseases.

References

[1] 'Detection of seizures in epileptic and non-epileptic patients using GPS and Embedded Systems' by Abhiram Chakraborty – Ukranian Journal of Telemedicine and medical Telematics (TOM 3 No.2 pp 211)
[2] 'Automated realtime interpretation of brainwaves' by Kenneth George Jordan US patent
[3] 'Method and apparatus for changing brain wave frequency' by John L Carter US patent 5036858
[4] en.wikipedia.org/wiki/ binaural beats
[5] en.wikipedia.org/wiki/Addiction

Cite this article as: K.KABILAN. "BRAIN WAVE CONTROLLER FOR STRESS REMOVAL AND AUTOMATION OF AUTOMOBILE IGNITION TO PREVENT DRIVING UNDER INFLUENCE." *International Conference on Information Engineering, Management and Security (2015)*: 74-80. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS013		**eAID**	ICIEMS.2015.013

Portfolio Graph: Risk Vs. Return Trade Off

Tuhin Mukherjee[1], Gautam Mitra[2]
[1]Department of Business Administration, University of Kalyani, WB, Nadia, India
[2]Department of Business Administration, University of Burdwan, WB, Burdwan, India.

Abstract: Due to possibility of high return, investment in stock is a challenging task for research community not only in finance but also in other disciplines of applied science and engineering. In this paper graph theory is used for portfolio optimization. All the constraints like upper bound of investment risk, fund constraint, time horizon etc. is being satisfied within this technique. The newly proposed technique outperforms the traditional techniques. However, due to paucity of time and non-availability of data set, the entire study is based on empirical in Indian context only. Future researchers can extend and verify its applicability for stock markets of other countries.

Keywords: Portfolio, Optimization, Investment Risk, Constraint Satisfaction Problem.

I. Introduction

A portfolio refers to a collection of securities held by an investor. Selection of securities within a portfolio obviously depends on risk aversion of the investor. So a good mix of securities implies different meaning to different people. Risk analysts usually consider return appreciation of securities pairwise within a portfolio. Types of securities included in a portfolio, may not necessarily be those which are positively correlated to each other [3]. The rational of such selection is to protect against adverse performance by limiting the risk. For example, to hedge against adverse downside risk of stock market, investors can simultaneously hold a put option to offset potential losses. This paper will address such issues using a graph theoretical framework. To the best of the knowledge, authors claim that sufficient literature is presently not available to link graph theory and portfolio management.

The next section describes graph theoretical concept followed by graph modeling of portfolio. Subsequently proposed methodology is illustrated followed by empirical evidence in Indian context. The paper terminates with vital conclusions along with scope future expansion.

II. Graph terminologies

A graph G is mathematically defined as tuple (V, E, φ) where V is a non-empty set whose elements are called vertices of G, E is a set whose elements are called edges of the graphs and φ is a function from E to V x V and called incidence function of the graph [2]. Geometrically the graph G = ({v1, v2, v3}, {e1, e2}, {(e1, v1, v2), (e2, v1, v2)} can be represented as in Fig. 1. Extending the notion of a graph, a signed graph is simply a graph where each edge corresponds to either positive or negative sign. For example Fig. 2 is a signed graph. Signed graph is relevant in describing the concept of structural balance [4]. According to the principle extolled by balance theory, a structure is balanced or stable iff all the cycles (closed walk between vertices and edges) of the signed graph are positive. Otherwise, it will be unbalanced. A cycle is said to be positive iff it has even number of negative edges in the cycle. Otherwise it will be negative. According to balance theory, any real life scenario with unbalanced structure has a tendency always to

Cite this article as: Tuhin Mukherjee, Gautam Mitra. "Portfolio Graph: Risk Vs. Return Trade Off." *International Conference on Information Engineering, Management and Security (2015)*: 81-85. Print.

undergo changes, (i.e. correction in bias) in order to restore its state of balance. Several approaches for determining the degree of balance associated with signed graph has been reported [5]. Among them, the triangular degree of balance index is perhaps the most elegant. The index is calculated by taking the ratio of the number of positive triangles (i.e. 3 vertices cycle) to the total number of triangles in the signed graph.

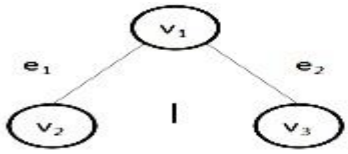


Fig. 1. A simple graph

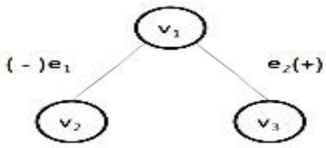


Fig. 2. A signed graph

This paper requires the concept of complete graph also. G is said to be complete iff it includes all possible edges between its vertices. A complete graph with n vertices is denoted by K_n as in Fig. 3.

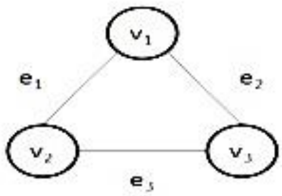


Fig. 3. A complete graph with 3 vertices (k_3)

If we denote negative edges by dotted line then some possibilities of complete graphs with 3, 4 and 5 vertices are shown in Fig. 4, Fig. 5 and Fig. 6 respectively.

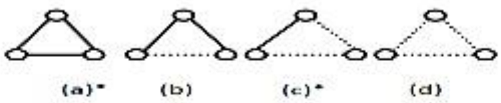


Fig. 4. Possible K_3

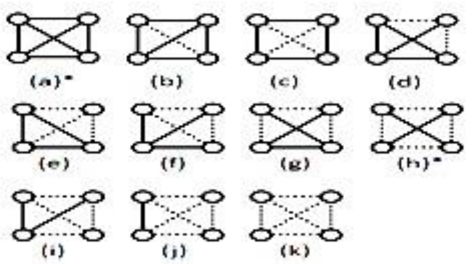


Fig. 5. Possible K_4

Cite this article as: Tuhin Mukherjee, Gautam Mitra. "Portfolio Graph: Risk Vs. Return Trade Off." *International Conference on Information Engineering, Management and Security (2015)*: 81-85. Print.

Fig. 5. Possible K_4

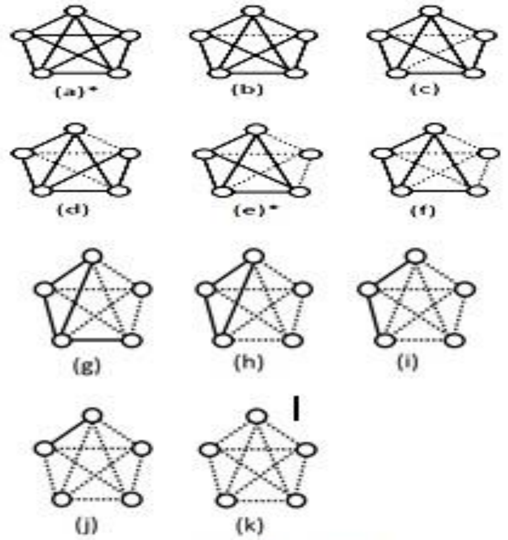


Fig. 6. Possible K_5

With these concepts of graph theory, now we are in a position to move in the next section describing link between graph theory and portfolio management.

III METHODOLOGY AND ANALYSIS

A portfolio of n securities can be represented as a complete graph whose vertices are securities and edges are sign of correlation between vertices. Hence a negative edge implies purpose of hedging. By analysing such portfolio graph, this paper recommends following benefits.

- The nature of portfolio graph can be used to judge the motive of investor. If protection against market shocks is the intention then graph should be balanced with at least one negative edge. On the other hand a balanced graph with all positive edges offers tendency to move either upside or downside. Hence it will be speculative in nature.
- A given portfolio graph is able to suggest possible course of actions for investors. If the portfolio graph is not balanced then suggestive restructuring can enhance the state of balance. On the other hand, for a given balanced graph, we can suggest structural transformations to move towards either hedging or speculation subject to risk.

Let us consider some possibilities of the portfolio graphs of 3, 4 and 5 vertices respectively as shown earlier in Fig. 4, Fig. 5 and Fig. 6. Only the balanced graphs are marked asterisk. Now we will discuss two possible situations. Case -I depicts the portfolio is balanced initially and case – II considers when it is unbalanced initially.
Case –I: Evidently 5(a) is balanced and has all positive edges and hence unprotected against downside risk. In order to hedge, the paper finds two possible courses of actions (recommended for an investor):

- Replacement of one or more vertices within Portfolio graph.
- Removal or Incorporation of one or more vertices within portfolio graph.

Evidently 5(e) and 5(h) both graphs are balanced. A transformation from 5(a) to 5(h) will involve replacing more than one security while 5(e) will need only one replacement.
A balanced K_n can be transformed to either balanced K_{n+m} or balanced K_{n-m} for some suitable natural number m. For example hedging can be done by incorporating stock option or future. It will imply transformation like 5(a) to 6(e).
Case –II: In this case, also we have two courses of actions as like as first case. To illustrate restructuring of unbalanced portfolio, we consider 6(a) which is unbalanced which can be transformed to balanced 6(e) without changing number of vertices. Even if a stable structure like 6(e) is not achievable, in order to reduce level of risk, investors may seek to restructure to one that has a higher degree of balance such as 6(h) or 6(g). Similarly to case–I, we can also restructure unbalanced portfolio by removal or incorporation of securities.
This section finds the evidence that goodness of a portfolio can be determined by its degree of balance. To arrive such empirical findings, we have considered 4 portfolios, each of which consists of 10 randomly selected stocks from National Stock Exchange of India, such that no two within the same portfolio, belong to same sector. The list of stocks is given in sorted order in Table 1.

Table 1. Four Portfolios (Ten stocks each)

No	Stock Name (Industry Affiliation)-Portfolio1	Stock Name (Industry Affiliation)-Portfolio2	Stock Name (Industry Affiliation)-Portfolio3	Stock Name (Industry Affiliation)-Portfolio4
1	Axis Bank Ltd (Finance)	ABB LTD (Capital Goods)	ACC LTD (Housing Related)	Axis Bank Ltd (Finance)
2	BPCL (Refineries)	Adani Power Ltd (Power)	Axis Bank Ltd (Finance)	Bajaj Hindustan Ltd. (Agriculture)
3	CESC Ltd (Power)	Berger Paints India Ltd (Chemical & Petrochemical)	Bilcare Ltd (Health care)	Berger Paints India Ltd (Chemical & Petrochemical)
4	CMC Ltd (Software)	CIPLA (Pharmaceutical)	CIPLA (Pharmaceutical)	Bharat Forge Ltd (Transport Equipments)
5	DLF (Housing Related)	Dish TV India Ltd (Media& Publishing)	DLF (Housing Related)	DLF (Housing Related)
6	EIH Ltd (Tourism)	EIH Ltd (Tourism)	Escorts Ltd (Transport Equipments)	EIH Ltd (Tourism)
7	FDC Ltd (Healthcare)	Emami Ltd (FMCG)	Finolex Industries Ltd (Chemical & Petrochemical)	Entertainment Network(India) Ltd (Media & Publishing)
8	Grasim Industries Ltd (Textile)	Grasim Industries Ltd (Textile)	Grasim Industries Ltd (Textile)	Essar Oil Ltd (Refineries)
9	Havel's India Ltd (Capital Goods)	HCL Technologies Ltd (Software)	GTL Infrastructure Ltd (Telecom)	Havel's India Ltd (Capital Goods)
10	ITC Ltd (FMCG)	Hindustan Unilever Ltd (Diversified)	Hindustan Unilever Ltd (Diversified)	Hindalco Industries Ltd (Aluminum)

Source of Industry Affiliation: NSE India

The sample is drawn with replacement between portfolios but without replacement within a portfolio. Naturally, the same stock may participate in different portfolio and restrict our choice in such a way that no two portfolios have more than 5 stocks (50% of 10 stocks) in common. The merit of such selection is not to be biased with predefined sectorial correlation between stocks (as nicely pointed out by [1]. Depending on their average prices (i.e. (closing prices + opening prices)/2) from 01.04.10 to 30.11.14, we have completed pair wise correlation$\rho = \frac{Cov(x,y)}{\sigma x \, \sigma y}$ using statistical package SPSS and use the sign of ρ in our empirical study. Now for each of these portfolios, we constructed their portfolio graphs and recorded their balance index shown in Table 2.

Table 2. Rank of four Portfolios (Ten stocks each) vs. Balanced index of their graph.

Portfolio	Balance Index	Sharpe	Treyner	Jenson	Rank
1	0.314	-0.0089	-0.1370	-0.0640	**4**
2	0.726	0.0041	0.2056	0.2957	**1**
3	0.568	0.0015	0.0527	0.1492	**3**
4	0.621	0.0023	0.1663	0.2141	**2**

Result: Computed.

It is evident from finance literature that only return cannot be the proper tool for measuring performance of any portfolio. One must consider the risk undertaken by each portfolio also. For measuring effectiveness of our portfolios, we have used Sharpe, Treyner, and Jenson which are the proven tools in financial management.
With these measures, we computed our 4 portfolios to rank them and results are presented in Table 2. From those results, it is clear that portfolio graph with highest degree of balance is proved to be rank first and similarly remaining portfolios are ranked according as decreasing balance of index. Though the author agree personally that to keep the degree of balance as the sole criteria to invest may not work fully but to significant extent this paper proved it viable for randomly selected portfolios from Indian stock market. It opens a new area of research towards stock market using a tool like graph theory which has almost never been used previously.

Cite this article as: Tuhin Mukherjee, Gautam Mitra. "Portfolio Graph: Risk Vs. Return Trade Off." *International Conference on Information Engineering, Management and Security (2015)*: 81-85. Print.

I V CONCLUSION

Main objective of this paper is to illustrate a methodology describing how complete signed balanced graph representation of portfolio can be used for risk analysis. It also provides a surprisingly elegant method to describe state of balance for a given portfolio graph. The paper recommends a portfolio graph which is complete signed and balanced for effective hedging but market dynamics can anytime change these characteristics of an existing portfolio. Proactive measures to restructure the portfolio graph may therefore be a necessary challenge in finance literature (future scope). In the empirical part, this paper attempts to study the goodness of a portfolio in terms of its balance index. It has been found that more the balance index, more suitable the portfolio is. Hence the paper recommends the investors to go for that portfolio which has more degree of balance. Although our illustrations are based on simple portfolio but the idea is generally applicable for portfolio involving larger cluster of assets. As the number of assets increases, the geometrical representation of portfolio graph should provide attractive alternatives for exposing structurally weak portfolios. Hence finding optimal transformation (of portfolio graph structure) subject to the available investment capital constraint is another challenge in future. Such optimality of a restructuring can be perceived in terms of number of vertices (securities) involved, changing amount of investment capital etc.

REFERENCES

[1] Basu, S. : The Investment Performance of Common Stocks in relation to their Price Earning Ratios : Journal of finance, Vol. 2, 1977, pp. 663-82.
[2] Donald, C.W. : Signed graphs for portfolio analysis in risk management : Journal of Management Mathematics, Vol. 13, 2001, pp. 201-210.
[3] Hull, J.C.: Options features and other Derivative Securities . 2nd Edition. Prentice Hall.
[4] Harary, F. : A structural analysis of the situation in the middle east in 1956 : Journal of conflict resolution,Vol. 5,1961, pp.167-178
[5] Harary, F.: A graph-theoretic approach to analysis. Research Note, 1961.

Cite this article as: Tuhin Mukherjee, Gautam Mitra. “Portfolio Graph: Risk Vs. Return Trade Off.” *International Conference on Information Engineering, Management and Security (2015)*: 81-85. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS014		**eAID**	ICIEMS.2015.014

Wireless Sensor Network Based Envoronmental Temparature Monitoring System

Ravi P. Athawale[1], Prof. J. G. Rana[2]
[1]ME(Appearing) Electronics Engineering
[2]Head of Department, Electronics & Telecommunication Engineering,
MGM's Jawaharlal Nehru Engineering College, Aurangabad, India

Abstract: *A sensor network consists of multiple detection stations called sensor nodes, each of which is small, lightweight and portable. Every sensor node is compact and doing task of the sensing, processing, etc. The implementation of monitoring of environment is hard task. This Sensor applications are to be implemented in multiple fields such as smart power grids, smart buildings and smart industrial process control significantly contribute to more efficient use of resources, Each node has sensor interfacing with signal conditioning and processing unit such as MSP430G2553 ultra low power device. The interfacing of the temperature sensors to MSP430 ICs on all nodes and light dependent resister to source and destination node. Destination node is assembled with 16*2 LCD display to get the values continuously on the display module. Continuous monitoring of all the nodes and observation is done on Master Computer. This operating system is designed to work on Windows 7.*

Keywords: LDR, wireless sensor network, environment***, monitoring, detection.***

I. INTRODUCTION

A wireless sensor network consists of a set of nodes powered by batteries and collaborating to perform sensing tasks in a given environment. It may contain one or more sink nodes (base stations) to collect sensed data and relay it to a central processing and storage system. A sensor node can be divided into three main functional units: a sensing unit, a communication unit and a computing unit[1]. My project aim is to make generalized platform for monitoring of the environment in suitable and non suitable conditions. As the emulation includes five nodes (slaves) are used out of which one is connected to PC/master computer called master slave and other four nodes are kept at some distance. Communication is done by RF2500[2] transceivers i.e. 2.5GHz ISM license free band (zigbee). This module is CC2500 Serial Transreceiver Wireless Module made by Vegarobokits. This module is interfaced with each node. Each node has sensor interfacing with signal conditioning and processing unit such as MSP430G2553 ultra low power device[1]. We have interfaced the temperature sensors on all nodes and light dependent resister to source and destination node. Destination node is assembled with 16*2 LCD[5] display to get the values continuously on the display module. Continuous monitoring of all the nodes and observation is done on Master Computer. This operating system is designed to work on XP, or Windows 7, platforms. The sensor networks are often deployed in unattended remote geographic areas, such as inside a large machinery at the bottom of an ocean, in contaminated field, battlefield, in home or large building. In a multi-hop sensor network, communicating nodes are linked by a wireless medium such as radio, infrared, or optical media.

Cite this article as: Ravi P. Athawale, Prof. J. G. Rana. "Wireless Sensor Network Based Envoronmental Temparature Monitoring System." *International Conference on Information Engineering, Management and Security (2015)*: 86-91. Print.

II. SENSOR NODE STRUCTURE

A. Processing Unit

In order to control the components of the sensor nodes and perform the required computations this subsystem is responsible for it. There are two sub-units, storage unit and processor unit[7]. Microcontroller MSP430G2553: The device features a powerful 16-bit RISC CPU, 16-bit registers, and constant generators that contribute to maximum code efficiency. The digitally controlled oscillator (DCO) allows wake-up from lowpower modes to active mode in less than 1 µs. The MSP430G2x13 and MSP430G2553[6] series are ultra-lowpower mixed signal microcontrollers with built-in 16- bit timers, up to 24 I/O capacitive-touch enabled pins, a versatile analog comparator, and built-in communication capability using the universal serial communication interface. In addition the MSP430G2553[7] family members have a 10-bit analogto- digital (A/D) converter.

B. Communication Unit

The sensor nodes due to this component interact with the base station and to the other nodes. Usually this subsystem is radio of short range but other fields has also been explored like ultrasound, infrared communication and inductive fields, communication by using CC2500 modules made by Vegarobokits[11] and these are able to interface with microcontroller IC MSP 430G2553[7] as both theses device have interfaced easily. As modules has capability of analog to digital conversion (ADC). So user has to just configure the module once for ADC, & the module will send the data to the respective receiver, at the given interval of time. Interfacing includes the communication of the microcontroller MSP 430G2553[7][8] IC of every node with other node. As we are implementing the sensor network for 5 nodes, four nodes are communicated with each other. Diagram for communication with microcontroller is as shown in below figure. It is only the communication of pins of RX of CC2500 with TX of Microcontroller and TX of CC2500[11] with RX of microcontroller with voltage of 5.12V DC as Vcc and ground is connected same of Launchpad[12].

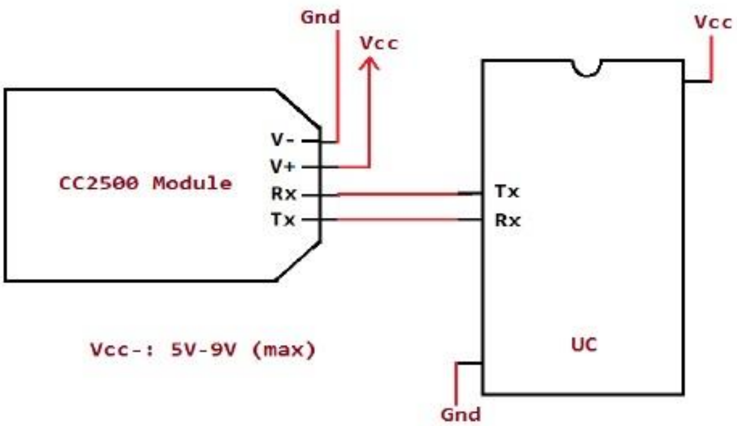


Fig. 1. Interfacing of CC2500 module with MSP430IC

C. Power Unit

Here power supply is provided to all the sections of the Node. Each node has sensing unit, processing unit, communication unit. Communication unit is provided with 5.12 DC V from power unit, as conversion of 12.52V DC to 5.12V by using 7805 voltage regulator. The processing unit includes the launch pad kit so it is provided with 3.3V. The range of the supplied voltage is 2.2 V to 3.6 V DC. We are providing it 3.3 V DC form power supply unit. Here 3.3 V is achieved by converting 5.12V into 3.3 V by using LM1117 IC it is the linear regulator having 800 mA low dropout IC, it gives continuous output of 3.3 VDC. We have obtained these 12.52 V DC from Adaptor and it is connected to every node. And then this voltage is converted as per our requirement. Adaptor converts 100-230V AC to 12.52 V DC with 1A rating at 50 Hz frequency. We are applying this to all five nodes.

D. Sensing Unit

Sensors are converting the temperature into electrical signals or light detection into electrical signals. In our project we are using only temperature sensor[14] (LM 35 DZ) and Light dependant resistor LDR only for Source node. Temperature sensor on all other four nodes these transducers are provided with voltage supply form voltage regulator. Signal Conditioning Section In the sensor node we have developed, LM35 DZ[14] serves as the temperature sensor. It is TO 92 Package. It generates an output voltage proportional to the temperature of the surroundings. Hence if we observe the output voltage to be 200mV at some time, we can say that the temperature at that moment is 20 degree Celsius. The output of the sensor is fed as the input of the ADC available in MSP 430 IC. In our configuration, the interfacing with the microcontroller, would give a digital output corresponding to the temperature.

E. Source Node

Source node is master node it is connected to personal computer through USB as it is made up of MSP430G2553 launch pad it is provided with DC power supply form adaptor as it is already mentioned in the power unit. As source node is beginner of the communication. When we starts the power supply of all the nodes, then we are initializing the node and begins the serial

Cite this article as: Ravi P. Athawale, Prof. J. G. Rana. "Wireless Sensor Network Based Envoronmental Temparature Monitoring System." *International Conference on Information Engineering, Management and Security (2015)*: 86-91. Print.

communication through UART port by pressing the reset (S1) and S2 switch on the launchpad kit. This node is interfaced with LDR also called as photo resistor.

F. Destination Node

Destination node is last node of communication; this node is interfaced with temperature sensor of LM35DZ in TO-92 Package and it is provided with 12 V DC adaptor and CC2500 communication module. Work of destination node starts after routing done and it is shown, the sensor value is kept at P1_5 pin, the data is shown on LCD 16*2 character display. Incoming bytes to the source node is noted by Val character. After route path is identified by source node, it will initialize the destination node and serial communication begins, then LCD[13] begins as shown in program then, data of character ‗incoming bytes' is defined and it is of 6 bytes.

G. Intermediate Nodes

As the intermediate node includes three nodes, Node-1, Node-2, and Node-3, these nodes are interfaced with temperature sensors of LM35 DZ with TO-92 package, and processor MSP430G2553 of launchpad kit and CC2500 communication module poweed with 12 V DC adaptor and in programming section path routing is done with after serial communication its main work is to check the data and transfer it to next node.

III. IDENTIFY THE ROUTE

As the source node is doing the important task of identifying the route it is the AODV protocol implementation. This node is the starting node of communication in the protocol, as this node starts serial communication it will get into the route setup of the communication, then it will check each nodes, first it will check node-1, as there is PC Frame is formed before strts of communication, in PC Frame it is noted with $1 sign for Node-1, and then it will check sensor value of node-1 and stores response time of the node-1. After thet the source node will similarly check node-2 and node-3and its notations are $2 for node-2 and $3 for node-3, and stores the interval of response and sensor value of both nodes. According to the node intervals stores the source node get the position of the nodes and assign it as the assign the node positions.

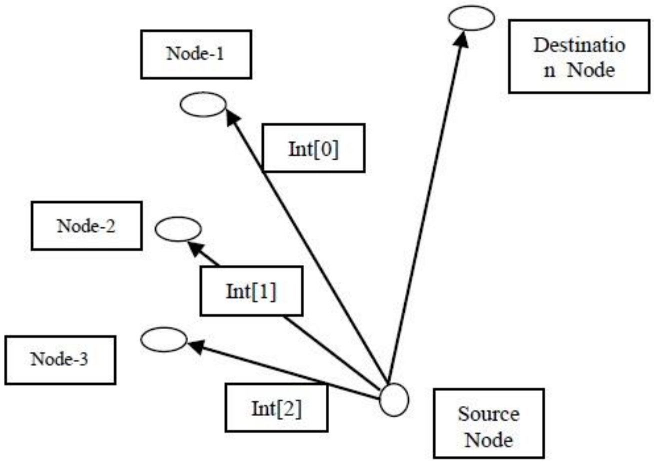


Fig. 2. Condition-1 for Node Position Assigning

Here interval[0] is time required to get acknowledgement to source node from node-1 in milliseconds. Here interval[1] is time required to get acknowledgement to source node from node-2 in milliseconds. Again here interval[2] is time required to get acknowledgement to source node from node-3 in milliseconds. This figure is made from our condition-1 in which interval[0]> interval[1] and interval[1]>interval[2] then we can say that node three is very close to source node then node-3 is given position at Data Frame[1], and node -2 is longer than node-3 but nearer than node-1 hence it is given position at Data Frame[2]. And node-1 is longer than node-2 and node-3 hence it ig given position at Data Frame[3]

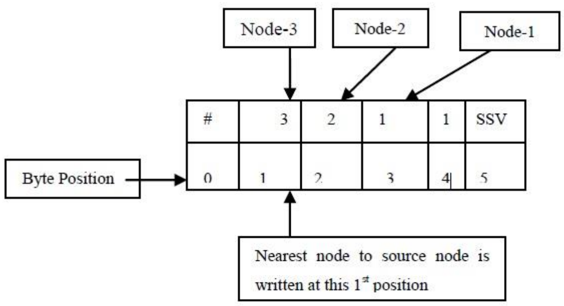


Fig. 3. Data Frame Format

Cite this article as: Ravi P. Athawale, Prof. J. G. Rana. "Wireless Sensor Network Based Envoronmental Temparature Monitoring System." *International Conference on Information Engineering, Management and Security (2015)*: 86-91. Print.

Notations of the dataframe format
Data Frame[0] = #, means it is data frame for source node
Data Frame[1] =3, node-3 is nearest to source node
Data Frame[2]= 2, Node-2 is longer than node-3 but nearer than node-1
Data Frame[3]= 1, Node-1 is longer than node-2, and node-3
Data Frame[4]= counter
Data Frame[5]= source sensor value
Notations of the PC frame format
& indicates this data is for PC
SV is sensor value for 1st intermediate node
SV is sensor value for 2nd intermediate node
SV is sensor value for 3rd intermediate node In this way the node position is assigned to the all three center nodes for first condition
Case -2: If interval[0]>interval[1] this condition is yes and interval[1]>interval[2] this condition is false, then source node will again check for interval[0]> interval[2] and this condition is true then data frame format becomes, Dataframe[1]=3, dataframe[2]=1, and dataframe[3]=2.
Case-3: If interval[0]>interval[1] this condition is yes and interval[1]>interval[2] this condition is false, then source node will again check for interval[0]> interval[2] and this condition is false then data frame format becomes, Dataframe[1]=2,
Dataframe[2]=1, Dataframe[3]=3.
Case-4: If interval[1]>interval[2] it is ture , interval[0]>interval[2] it is true, then data frame format is Dataframe[1]=2, Dataframe[2]=3, Dataframe[3]=1.
Case-5; If interval[1]>interval[2] it is ture , interval[0]>interval[2] it is false then Dataframe becomes DataFrame[1]=1, Dataframe[2]=3, Dataframe[3]=2.
Case-6: If interval[1]>interval[2] and it is false then Data frame becomes Dataframe[1]=1, Dataframe[2]=2, Dataframe[3]=3

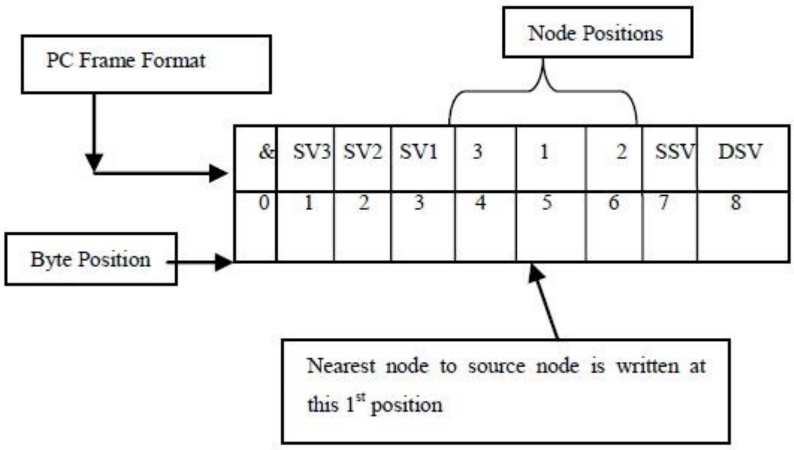


Fig. 4. PC Frame Format

All these are possibilities of the Node positions, on every identification of route discovery theses path is identified and come in effective for routing purpose.In this way the node positions are assigned by using the above data frame format, and PC frame format will this data is sent to the PC which will be monitored on PC.

A. Implementation of Source Node
Source node is master node it is connected to personal computer through USB as it is made up of MSP430G2553 launch pad it is provided with DC power supply form adaptor as it is already mentioned in the power unit. As source node is beginner of the communication. When we starts the power supply of all the nodes, then we are initializing the node and begins the serial communication through UART port by pressing the reset (S1) and S2 switch on the launchpad kit. This node is interfaced with LDR also called as photo resistor. As the source node is doing another task of identifying the route it is the AODV protocol implementation and it is described in the software development section. This node is the starting node of communication in the protocol, as this node starts serial communication it will get into the route setup of the communication, then it will check each nodes, first it will check node-1, as there is PC Frame is formed before strts of communication, in PC Frame it is noted with $1 sign for Node-1, and then it will check sensor value of node-1 and stores response time of the node-1. After thet the source node will similarly check node-2 and node-3and its notations are $2 for node-2 and $3 for node-3, and stores the interval of response and sensor value of both nodes. According to the node intervals stores the source node get the position of the nodes and assign it as the assign the node positions. This assigning node positions is the main work of the source node and after getting the node positions, another task of the source node begins after routing is completed and data start flowing form destination node to source node, and incoming bytes form desination

node are denoted by @, so its PC frame starts form @ sign then it is considered as PC frame. This data is transferred to PC and monitored on the PC.

B. Monitoring of Nodes on PC

While at first power on then USB of source node is connected to the PC, through this USB the communication between them is started. To achieve this task we have prepared .Net based platform in which the program for monitoring of all the nodes is done. The program is explained in program section some basic steps are as follows. When program is started we will load the TMS file saved in our PC. Then we check the communication port where we are connecting the USB of source node. Then next stage is debug the file, then we enter the number of communication port in to the tab shown. Then click the _Start System‘ on system as tab is shown, then message of _working‘ is observed on the screen. Then we wait for communication and collection of data on the source node and lastly all the values of nodes are shown on PC. When we want to stop the monitoring of node then press _Stop System‘. As we have made this operating system for 5 nodes[14], we are showing 8 different values on the PC, we have defined the 8 characters as per the code mentioned in the program of all the nodes. It is the screen shot of monitoring of nodes on PC

1. It shows node path for that routing in this 2->3->1 is the routing path
2. In first line 2(23)->3(23)->1(21) here these three are intermediate nodes and 2(23) means 23 Degree Centigrade is the temperature of 2nd intermediate node. in the same way temperature of other two nodes is also mentioned on the screen.
3. Then Self Value means we have attached photo resistor to the source node its rating given as 0 to 255 nos, as the Self Value shown
4. Destination Value shown is the Temperature of Destination node, here in first line it shows 22 Degree centigrade. At that time mentioned at 9:06:30 AM, mentioned
5. All this reading line come after every 35 to 39 seconds in three lines

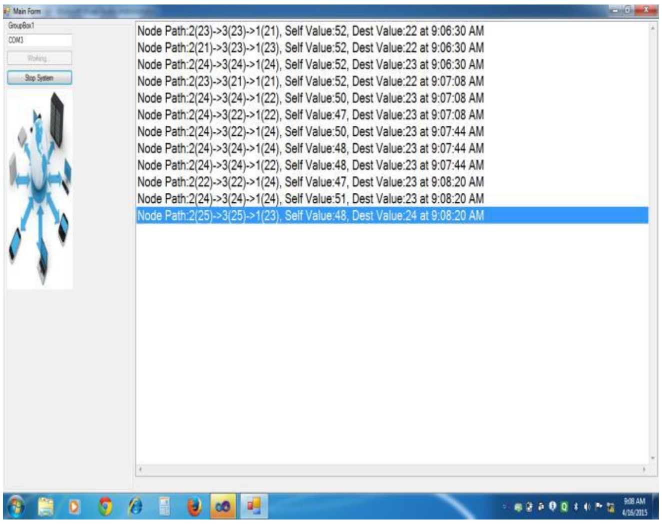


Fig. 5. Monitoring of Nodes on PC

IV. RESULTS & DISCUSSIONS

Result of this project is the monitoring of the all nodes on the PC, as this is MSP 430 based project so it is ultra low power device and its maximum power consumption is only 60 miliWats.

V. SCOPE OF PROJECT

☐ Implementation in Industries As the development of the sensor network is continuously in progress in the field of wireless communication, power efficiency, extreme miniaturization every section of node is in development and embedded computingtechnologies have led to the rise of viable wireless sensor networks for demanding Industrial environments. We can put lot of parameters on a single node and concurrently it is kept monitored on the control room of the industry.

☐ Industrial Monitoring on Internet As the advancement leads to the development of nodes and its applications to the higher level. There are thousands of the applications where these nodes are utilized hence by using the cloud computing and taking IP address to the each node, and by putting that IP on the internet address we can monitor these nodes form anywhere from the any country of the world.

REFERENCES

[1] Gang Zhao, *Wireless Sensor Networks for Industrial Process Monitoring and Control: A Survey,* Microthink Institutes, Network Protocol Algorithms,ISSN 1943-3581/2011 vol 3 No.1

[2] P. Bonnet, J. Gehrke, and P. Seshadri, ―Querying the physical world,‖ *IEEE Personal Communications*, 7(5):10–15, October 2000

Cite this article as: Ravi P. Athawale, Prof. J. G. Rana. "Wireless Sensor Network Based Envoronmental Temparature Monitoring System." *International Conference on Information Engineering, Management and Security (2015)*: 86-91. Print.

[3] Ian D Chakeres, Eleizabeth M.Blending Royer, AODV-UCSB Implementation from University of California Santa Barbara. http://moment.cs.ucsb.edu/AODV/aodv.html
[4] Perkins, C.; Belding-Royer, E.; Das, S. (July 2003). Ad hoc On-Demand Distance Vector (AODV) Routing. IETF. RFC 3561. https://tools.ietf.org/html/rfc3561. Retrieved 2010- 06-18.
[5] V. Kawadia, Y. Zhang, and B. Gupta. System Services for Implementing Ad-Hoc Routing: Architecture, Implementation and Experiences. In *Proceedings of the* 1st *International Conference on Mobile Systems, Applications, and Services (MobiSys)*, pages 99.112, San Francisco, CA, June 2003.
[6] Texas Instruments, *MSP430G Evaluation Kit,* Launchpad, Literature Number: SLAU318F, www.ti.com/tool/msp-exp430g2.
[7] Texas Instruments, MSP430G2553 , SLAS735J, http://www.ti.com/lit /ds/symlin /msp430g2453.pdf
[8] Luis Javier García Villalba *, Ana Lucila Sandoval Orozco, Routing Protocols in Wireless Sensor Networks, *sensor,*ISSN no 1424-8220, www.mdpi.com/journal/sensors
[9] http://iitkgp.vlab.co.in/?sub=38&brch=121&sim=561 &cnt=1
[10] Bruno Sinopoli, Courtney Sharp, Luca Schenato, Shawn Schaffert, Distributed Control Applications within sensor Networks, Procedings of the IEEE Vol 91 No 8 Augest 2003, invited paper
[11] Vegarobokits, CC2500 serial Transreciver Wireless Module
[12] Vegarobokits, CC2500 serial Transreciver Wireless Module/interfacing CC2500 module
[13] National Semicaonductor/LM1117/800mA low dropout linear regulator.pdf
[14] National Semiconductor, LM 35 Pricision Centigrade Temparature Sensor http://www.alldatasheet.com/datasheetpdf/ pdf/8866/NSC/LM35.html

Cite this article as: Ravi P. Athawale, Prof. J. G. Rana. "Wireless Sensor Network Based Envoronmental Temparature Monitoring System." *International Conference on Information Engineering, Management and Security (2015)*: 86-91. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS015		**eAID**	ICIEMS.2015.015

Bluetooth to Bluetooth RSSI estimation Using Smart Phones

Kumareson.P[1], Rajasekar.R[2], Prakasam.P[3]
[1]PG scholar/CSE, Tagore Institute of Engineering and Technology, Salem, India.
[2]Associate Professor/CSE, Tagore Institute of Engineering and Technology, Salem, India.
[3]Professor/ECE, United Institute of Technology, Coimbatore, India.

Abstract: *Mobile computing is one of the most advanced computing models used in scientific research applications. Although the foundation of the proximity estimation model was laid by past generations only the recently advances opened on expanding proximity estimation application range and its research implementation. Existing approaches used such as GPS and WIFI triangulation are more complicated to find the accurate extraction of proximate location and it insufficient to meet the requirements of flexibility and accuracy. In nowadays the Bluetooth which is commonly available on most modern Smartphone's. And it finds the exact location of any Bluetooth users only in a certain limit. By pairing the key in their mobile of one Bluetooth users to any another Bluetooth users. This paper proposes a proximity estimation model to identify the distance based on the RSSI values of Bluetooth and light Sensor data in different environments. And also state Bluetooth proximity estimation model on Android with respect to accuracy and power consumption with a several real world scenarios.*

Keywords: Bluetooth, RSSI, proximity estimation model, smartphone, face-to-face proximity

I. Introduction

In recent years, the mobile phone market has increasingly used Bluetooth as the preferred method of device communication, data exchange, and accessory pairing. Many PC accessories including mice, keyboards, headsets, and printers also employ the Bluetooth standard for wireless communication.Bluetooth is an industrial standard for wireless personal area networks. It is primarily designed for low power consumption and short range operations among several mobile and embedded devices.. Bluetooth provides connection management and data exchange among devices that are within close proximity and do not require high bandwidth data links. The technical challenge is how to measure face-to-face interactions.In Bluetooth the rssi signals range between two or more individuals within a certain distance that could afford those interactions.The previously mentioned schemes used as to determine the proximity estimation is GPS and WIFI are not such efficient,because its suffers from accuracy shortcoming and lack of viability indoors.

With the important shift of the problem statement, Bluetooth emerges as a straightforward and Alternative approach used as offering both accuracy and ubiquity (most modern smartphones come with Bluetooth) Although some prior work has attempted to use the detection of Bluetooth to indicate proximity nearness, it is not enough for the face-to-face proximity estimation. This paper describes to extent the range of Bluetooth and it can be an accurate estimator of such proximity.

Cite this article as: Kumareson.P, Rajasekar.R, Prakasam.P. "Bluetooth to Bluetooth RSSI estimation Using Smart Phones." *International Conference on Information Engineering, Management and Security (2015)*: 92-97. Print.

To summarize, our work it makes the following contributions:

- To explore the viability of using Bluetooth for the purposes of face-to-face proximity estimation and propose a proximity estimation model with appropriate smoothing and consideration of a wide variety of typical environments.
- To Identifying the relationship between the value of Bluetooth RSSI and distance based on empirical measurements and compare the results with the theoretical results using the radio propagation model.
- To explore the energy efficiency and accuracy of Bluetooth.

The remainder of the paper is organized as follows.In Section II the problem identification is described. In Section III, the introduction about related approaches to get relative distance determination in proximity estimation. In Section IV the data collecting system built on smartphones is documented.In Section V the proximity estimation model with smoothing and environment differentiation is proposed. Finally, we suggest ways to extend this work to future communication research in Section VI.

II. PROBLEM IDENTIFICATION

Bluetooth -to- Bluetooth interaction does not demand an absolute position as offered by the previously mentioned schemes like GPS and WI-FI but rather it requires a determination of proximity.

With that above important shift of the problem staement .Bluetooth emerges as a straightforward and alternative approach to offering both accuracy 1-1.2 m and ubiquity (most latest smart phones come with Bluetooth). Although the prior work has attempted to use the detection of Bluetooth to indicate nearness, it is not enough for the Bluetooth -to- Bluetooth proximity estimation.

Data values reported by light sensor is not reliable. Determination of proximity within a limit(coverage).Each RSSI value was not allow for environmental fluctuations.

The critical challenge is how to measure face-to-face interactions Two or more individuals within a certain distance that could afford such interactions .Interactions are not limited to any particular area and can take place at a wide variety of locations, ranging from sitting and chatting in a Starbucks coffee shop to walking and chatting across a college campus

III.RELATED WORK

Over the past years, there has been a number of technologies proposed for proximity detection. The approach used such as meme tags,active badge,place lab,zigbee techonology,location based services,3-D optical services…etc

A)Meme Tags

The Meme Tag event took place over a period of October 1997. The event was designed by MIT Media Lab's *Digital Life* (DL), *Thing That Think* (TTT), and *News In the Future* (NIF) consortia sponsor meetings

Meme Tags [12] provide good accuracy but require line of sight.

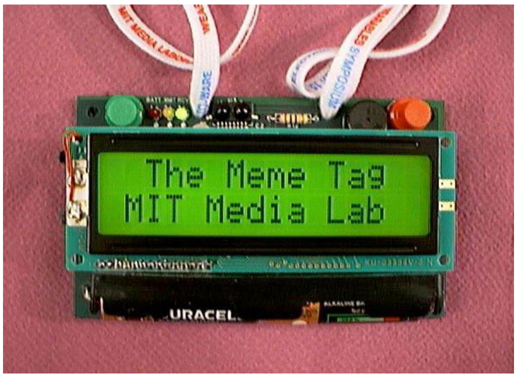


figure 1: The Meme Tag. Worn around the neck, the

Meme Tag includes a large, bright LCD screen, green

and red pushbuttons (for accepting or deleting memes), a knob (not visible) for reviewing and choosing memes to offer, and a bidirectional infrared communications device

B)Active Badge

Cite this article as: Kumareson.P, Rajasekar.R, Prakasam.P. "Bluetooth to Bluetooth RSSI estimation Using Smart Phones." *International Conference on Information Engineering, Management and Security (2015)*: 92-97. Print.

Ultrasound approaches used such as Activebadge also provide good accuracy but they require infrastructure support. Goal of active badge is to find efficient location and coordination of workers in a large organization. Existing Solutions used as Broadcasting a phone to call several possible numbers a beeper with audible signal or call-back number

Ex: location of doctors, staffs, and patients in a hospital.

C)Zig Bee technology

ZigBee technology is widely used in wireles

sensor network to provide radio proximity estimation in the environment where GPS is inoperative. Proximity can also be reported by sounds, and past work has shown audio to be effective for delivering peripheral cue. However,it is untenable to expect the use of smartphones to reduce the unobtrusiveness of cues or increase comprehension However,it is untenable to expect the use of smartphones to reduce the unobtrusiveness of cues or increase comprehension

For the purposes of this paper, we are interested in techniques that are based on commonly available technologies in smartphones,i.e., GPS, Cell, Wi-Fi and Bluetooth. Particularly, we are interested in techniques that can be applied at the smartphone itself without significant changes to the infrastructure.There are some proximity detection works using Bluetooth signal. From a specific work perspective, the works of Eagle etal. are highly relevant to this paper. In those studies, the authors use the ability to detect Bluetooth signals as indicators for people nearby within the Bluetooth range However, such indication does not meet the requirement of face-to-face proximity detection. In class, a student may discuss with others sitting beside him/her, butface-to-face talk is difficult with the students on the other side of the classroom even they are still in the Bluetooth range. Different from the above proximity detection method, our method is a fine grain Bluetooth-based proximity detection method which can provide adequate accuracy for face-to face proximity estimation without any environment limitations.

D)location based systems

Proximity detection is one of the advanced Location based Service (LBS) functions to automatically detect when a pair of mobile targets approach each other closer than a predefined proximity distance (as in Location Alerts of Google Latitude and longitude). For realizing this function these targets are equipped with a cellular mobile device with an integrated GPS receiver, which passed position fixes obtained by GPS to a centralized location server. Most proposals for such services give low accuracy and its guarantees to incur high communication costs

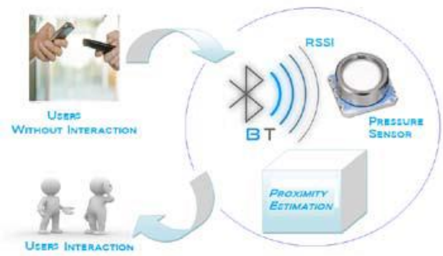
E) 3-D optical approach

3-D optical wireless based location approach is proposed which it is based on both GPS and triangulation technologies. It is another feasible way of utilizing GPS to get relative distance among objects.Some proximity estimation methods are based in Cell or WiFi signal. Using Place Lab cell phones listen for the MAC address of fixed radio beacons such as cell tower ,wireless access points, and reference the beacons positions in a cached database. It provides adequate accuracy for detecting something like buddy proximity (e.g., median accuracy of 20-30 meter)

IV.SOFTWARE DESING FOR BLUETOOTH SMARTPHONE’S

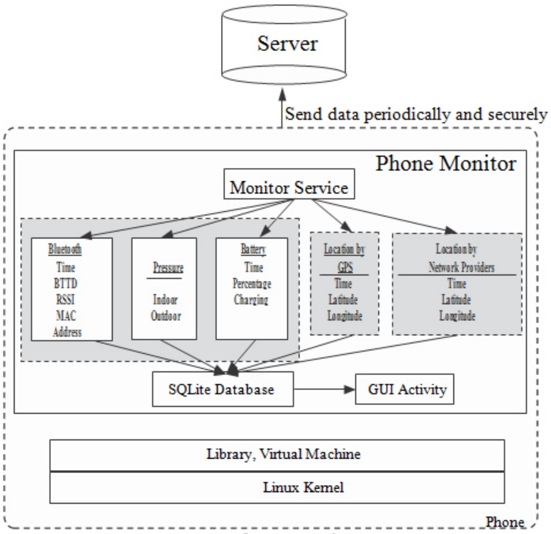
A)System Architecture

The smart phone is taken and the application for it is modeled.The first one is to enable the application and it will turn on the Bluetooth application then it asks whether to display the listed pair device. If list paired device button is pressed then the list of paired device is displayed. Then select one device and if the device is near the coverage area then it displaces the RSSI value i.e., distance between the devices. The obtained RSSI value is calculated by using the propagation formula. After that in future the pressure sensor is used to detect whether the smart phone is in indoor or outdoor location



B)Data Collection System

The application named Phone Monitor collects Bluetooth data including the detailed values of RSSI, MAC address, and Bluetooth identifier (BTID). The data is recorded in SD card once the phone detects other Bluetooth devices around. In addition to Bluetooth, data points from a variety of other subsystems (light sensor, battery level and etc.) are gathered in order to compare and improve the proximity estimation. Separate threads are employed to compensate for the variety of speeds at which the respective subsystems offer relevant data. It also record the location data reported by both GPS and network providers (either WiFi or cell network). In order to determine whether the phone is sheltered (e.g. inside a backpack or in hand) and the surroundings (e.g.inside or outside buildings) during the daytime, we keep track of the light sensor data values. The battery usage of the percentage is recorded for the energy consumption comparison.



V.BLUETOOTH PROXIMITY ESTIMATION MODEL

In this section, we explore the relationship between Bluetooth RSSI and distance in real world scenarios. The first method is using RSSI value threshold to determine whether two phones are in proximity or not. The second method introduces the light sensor data to determine whether the phone is indoors or outdoors, inside the backpack or in hand. By differentiating environments and smoothing data, a face-to-face proximity estimation model is outlined to improve the estimation accuracy in general scenarios. At the end of this section the proximity accuracy of Bluetooth, WiFi and GPS are analyzed and compared.

A) Bluetooth RSSI vs. Distance

Anti et al. presented the design and implementation of a Bluetooth Local Positioning Application (BLPA) in which the Bluetooth received signal power level is converted to distance estimate according to a simple propagation model as follows:

$$RSSI = PTX + GTX + GRX + 20 \log (c\, 4\pi f) \;\; - 10n \log (d)$$
$$= PTX + G - 40.2 - 10n \log (d)$$

where PTX is the transmit power; GTX and GRX are the antenna gains and G is the total antenna gain:
$G = GTX + GRX$, c is the speed of light ($3.0*108m/s$), f is the central frequency (2.44 GHz), n is the attenuation factor (2 in free space), and d is the distance between transmitter and receiver (in m). d is therefore:
$d = 10[(PTX-40.2-RSSI+G)/10n]$
However, such a model can only be utilized as a theoretical reference. Due to reflection, obstacles, noise and antenna orientation, the relationship between RSSI and distance becomes more complicated. Our challenge was to assess how much impact these environmental factors have on Bluetooth RSSI values. Therefore, we carried out several experiments to understand how the Bluetooth indicators fade with distance under these environmental influences.

B)Single Threshold

RSSI value (-52dBm) of direct communication distance (152cm) based on the indoor measurements was used as a threshold to estimate whether the individuals were in proximity. Accordingly,values less then -52dBm were considered as not in face-to-face proximity and labeled as a wrong estimation. It was found that both of the outdoor and backpack parts have extremely high error rates.After switching the threshold value to -58dBm which is the outdoor RSSI values with 152cm distance, the error rate was improved but still high. To reduce the error rate we go multiple thresholds with data smoothing and different environmental effect.

C)Multiple Thresholds

According to the reasons for high error rate then we introduce the proximity estimation model which is a multiple threshold-based method with the consideration of data smoothing and different environmental effects.

i)Data Smoothing

. Since there is time delay during the data collection,then do smoothing on the data collection to avoid environmental fluctuation effects and there are several ways to achieve it. using simple window function and each value RSSI(i) at time (i) is modified using the following function:

$RSSI(i) = a * RSSI(i-1) + b * RSSI(i) + c * RSSI(i+1)$

Another one smoothing method is to utilize EWMA (exponentially weighted moving average) to analyze the dataset. Let the Ei be the EWMA value at time (i) and (s) be the smoothing factor.
The EWMA calculation is as follows:

$Ei = s * RSSIi + (1 - s)Ei-1.$

Measureing the possible face to- face interaction distances across the campus (such as diagonal of desk in dinning hall and distance between desks in classrooms and etc.) and the average value is equal to1.52 (m.)Base assessment: the whole process took 30 minutes and individuals were always within the distance for face-to-face communication. After the data collection, the corresponding RSSI value (-50dBm) of direct communication distance (152cm) was used as a threshold to estimate whether the individualswere in face-to-face proximity or not

ii)Light Sensor Data

The Bluetooth RSSI values are much smaller than the indoor ones when the phone is in the backpack or outdoors. One of our observations is that it is possible to treat the light sensor data as an indicator of the environment

VI.ACKNOWLEDGEMNT

This work was funded in part by the National Science
Foundation through grant IIS-0968529.

VII.CONCLUSION AND FUTURE WORK

I have analyzed several proximity estimation model by combining Bluetooth RSSI value, light sensor data as well as data smoothing together and understanding with the method of collecting all devices around, the accuracy of utilizing proximity estimation model to estimate whether two devices are in a Direct communication distance is improved dramatically. I also analyzed and studied the battery usage and accuracy of Bluetooth method with other different location methods such as WiFi triangulation and GPS. Finally it demonstrates that Bluetooth offers an effective mechanism that is accurate and power-efficient for measuring face-to-face proximity to increase Bluetooth signal Strength level and its coverage range.Another promising method for Improving the threshold algorithms w ith data mining. The thresholds used in the proximity estimation model are based on the experiment results on android phones. For different phones, such thresholds may be different. Therefore, a more general method is necessary to determine the relationship between Bluetooth RSSI values and the face-to-face proximity. With more data reported in the next following years, a more efficient data mining algorithm is needed to analyze the data. During the nighttime, only the data reported by light sensor is not efficient. The possible method to solve this problem is to taking atmospheric pressure into consideration to determine whether the phone is indoor or outdoor.

REFERENCE

[1] Yang-Hang Lee , Kuo-Wei Ho, Hsien-Wei Tseng," Accurate Bluetooth Positioning Using Large Number of Devices Measurements", IMCES ,Vol.2,No 4, PP.381-394 March 12 - 14, 2014,

[2] M. R. Friesen & R. D. McLeod," Bluetooth in Intelligent Transportation Systems: A Survey",Springer,vol 10,No 15,pp.107-112, May .2014

[3] Dae Wook Kim and Mehmet M. Dalkilic," Analysis of Proximity Networks from Multiple Mobile Sensor Data", International Journal of Future Computer and Communication, Vol. 2, No. 3,pp 134-139, June 2013

[4] Avik Ghose,Chirabrata Bhaumik," BlueEye – A System for Proximity Detection Using Bluetooth on Mobile Phones", UbiComp'13, Vol 5,No 7,pp.8-12,September 2013

[5] Julian Benavides,"Smartphone techonologies for social network data generation and infectious disease modeling", Journal of Medical and Biological Engineerin,,Vol 32 ,No 4,pp. 235-244,jan 2012.

[6] Trinh Minh Tri Do,Daniel Gatica-Perez," GroupUs: Smartphone Proximity Data and Human Interaction Type Mining",Vol 5,No 2,pp 8-5, feb 2011

[7] M. N. Juuso Karikoski, "Measuring social relations with multiple datasets," International Journal of Social Computing and Cyber- Physical Systems, vol. 1, no. 1, pp. 98–113, November 2011.

[8] Ling Pei, Ruizhi Chen," Using Inquiry-based Bluetooth RSSI Probability Distributions for Indoor Positioning", Journal of Global Positioning Systems ,Vol.9, No.2,pp.122-130,march 2010

[9] H.Falaki,R.Mahajan,S.Kandula,D.Lymberopoulos,R.Govindan,and Estrin,"Diversity in smartphoneusage,"inProceedingofthe8thinternationalconferenceonMobilesystems,applications,andservices. ACM,2010, pp. 179–194

[10] Chakib Baouche, Antonio Freitas," Radio Proximity Detection in a WSN to Localize Mobile Entities Within a Confined Area", JOURNAL OF COMMUNICATIONS, VOL. 4, NO. 4,pp.132-138, MAY 2009

[11] Mika Raento, Antti Oulasvirta," Smartphones: An Emerging Tool for Social Scientists", Sociological Methods & Research,Volume 37 Number 3, pp.426-431,February 2009

[12] A. P. Nathan Eagle and D. Lazer, "Inferring social network structure using mobile phone data," Proc. of the National Academy of Sciences (PNAS), vol. 106, no. 36, pp. 15 274–15 278, September 2009

[13] Kevin A. Li, Timothy Y. Sohn," PeopleTones: A System for the Detection and Notification of Buddy Proximity on Mobile Phones", ACM,vol 2,No 8,pp-978-982,June 2008

[14] Aswin N Raghavan1 Harini Ananthapadmanaban,"IIT,Vol 10,No 3,pp 123-128,feb 2007

[15] Fernán Izquierdo, Marc Ciurana," Performance evaluation of a TOA-based trilateration method to locate terminals in WLAN",IEEE,Vol 5,No 3,pp. 132-140,jan2006.

[16] Georg Treu and Axel K¨upper," Efficient Proximity Detection for Location Based Services," (WPNC'05)& (UET'05),Vol 12,No 24,pp.567-572,june 2005.

[17] N. Eagle and A. Pentland, "Social serendipity: Mobilizing social software," IEEE Pervasive Computing, vol. 4, no. 2, pp. 28–34, 2005.

[18] V. Otsason, A. Varshavsky, A. LaMarca, and E. De Lara, "Accurate GSM indoor localization," UbiComp 2005: Ubiquitous Computing, pp. 903–921, 2005

[19] A. Kotanen, M. Hannikainen, H. Leppakoski, and T. Hamalainen,"Experiments on local positioning with Bluetooth," in ITCC 2003:Information Technology: Coding and Computing, april 2003, pp. 297–303

[20] V. Zeimpekis, G. M. Giaglis, and G. Lekakos, "A taxonomy of indoor and outdoor positioning techniques for mobile location services," SIGecom Exch., vol. 3, pp. 19–27, December 2002.

Cite this article as: Kumareson.P, Rajasekar.R, Prakasam.P. "Bluetooth to Bluetooth RSSI estimation Using Smart Phones." *International Conference on Information Engineering, Management and Security (2015)*: 92-97. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	VOL	01
Website	www.iciems.in	eMail	iciems@asdf.res.in
Received	10 - July - 2015	Accepted	31- July - 2015
Article ID	ICIEMS016	eAID	ICIEMS.2015.016

Implementations of Reconfigurable Cryptoprocessor A Survey

N Rajitha[1], R Sridevi[2]
[1]Research Scholar, JNTUH, Hyderabad
[2]Professor in CSE JNTUH, Hyderabad

ABSTRACT: *One among the several challenges in the area of applied cryptography is not just devising a secure cryptographic algorithm but also to manage with its secure and efficient implementation in the hardware and software platforms. Cryptographic algorithms have widespread use for every conceivable purpose. Hence, secure implementation of the algorithm is essential in order to thwart the side channel attacks. Also, most of the cryptographic algorithms rely on modular arithmetic, algebraic operations and mathematical functions and hence are computation intensive. Consequently, these algorithms may be isolated to be implemented on a secure and separate cryptographic unit.*

Keywords: Trust, FPGA security, Cryptographic processor, reconfigurable cryptosystems.

I. INTRODUCTION

There is an alarming need for securing wide area of applications of cryptography that we use in our daily life besides military, defense, banking, finance sectors and many more. To cater to this need innumerable products/services have been developed which are predominantly based on encryption. Encryption in turn relies on the security of the algorithm and the key used. The different encryption algorithms proposed so far have been subjected to various forms of attacks. While it is not possible to devise an algorithm that works perfectly well and sustains all forms of attacks, cryptographers strive to develop one that is resistant to attacks and that performs well. The task is not just to propose a new algorithm but to create an environment that improves the performance of the algorithm and that protects the keys from attacks. A cryptoprocessor is a specialized processor that executes cryptographic algorithms within the hardware to accelerate encryption algorithms, to offer better data, key protection. Commercial examples of cryptoprocessors include IBM 4758, SafeNet security processor, Atmel Crypto Authentication devices. The following are the different architectures of cryptographic computing[1].

A.Cryptoprocessor Types

· Customized General Purpose Processor: The processor is extended or customized to implement the cryptographic algorithms efficiently. Typical commercially available solutions are CryptoBlaze from Xilinx or the AES New Instructions (AES-NI) incorporated in the new Intel processors.
· Cryptographic processor (cryptoprocessor): It is a programmable device with a dedicated instruction set to implement the cryptographic algorithm efficiently.
· Cryptographic coprocessor (crypto coprocessor): It is a logic device dedicated to the execution of cryptographic functions. Unlike the cryptoprocessor it cannot be programmed, but can be configured, controlled and parameterized.
· Cryptographic array (crypto-array): It is a coarse grained reconfigurable architecture for cryptographic computing.

Cite this article as: N Rajitha, R Sridevi. "Implementations of Reconfigurable Cryptoprocessor A Survey." *International Conference on Information Engineering, Management and Security (2015)*: 98-103. Print.

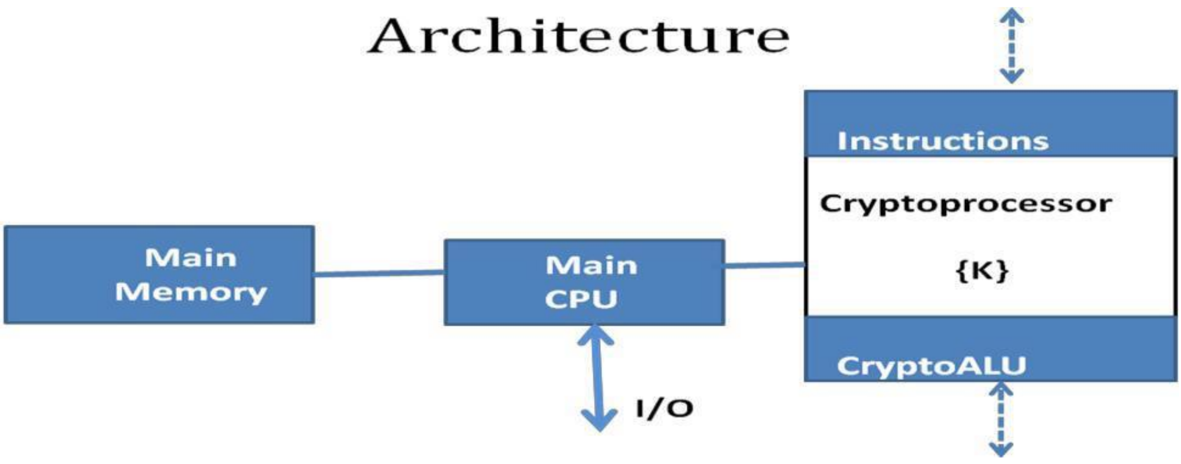


Fig 1 Architecture of Cryptoprocessor [1]

B Cryptoprocessor Implementations

i) Cryptoprocessor implemented in (field programmable gate array) FPGA are fast in terms of cryptographic processing. The complex mathematical operations can be run **quickly and efficiently**. IP blocks can be **modified** if desired as the name suggests. FPGA based cryptoprocessors are used in ATMs, automobiles, robotics etc.

ii) **ASIC based cryptoprocessors** have small footprint and offer high speed. They cannot be changed once produced. They use less power and are used in applications such as RFID, network routers, cameras, cell phones etc.

iii) **Hardware Security Module** (HSM) contains one or more secure cryptoprocessor chips to prevent tampering and bus probing. They come in the form of a plug-in card or an external device that attaches directly to a computer of some sort. HSM can be made to provide backup to computer to which it is attached, NAS, cloud server and can be used as external security token.

iv) **Trusted Platform Module** (TPM) is a cryptoprocessor integrated in software microkernel. The kernel generates and stores keys, passwords and certificates. They can be found in Digital Rights Management to ensure that audio/video file is original and not a copy.

ii) CRYPTOPROCESSOR ATTACKS

The different forms of hardware attacks on algorithmic implementations on cryptographic devices in literature have been identified as given below

i) **Side Channel Attack**: A study of the literature reveals that a major amount of research has been expended during the last decade on side channel attacks and countermeasures. Side channel attacks and can happen in one of the following ways:

a) **Timing Analysis**: Time required by the device to perform encryption/decryption can be used to get additional data to perform an attack.

b) **Electromagnetic analysis**: It is based on the electromagnetic radiation from the circuit that executes the encryption/ decryption algorithm

c) **Power Analysis**: Power consumed by the device implementing the algorithms can be used to perform the attack. It can be of the form Simple Power Analysis or Differential Power Analysis. Side channel attacks and countermeasures can be found in [25], [42],[43], [44]. Pawel Swierczynki et al[25] discuss side channel attack on bitstream encryption of Altera Stratix II and Stratix III FPGA family in the form of black box attack. To combat IP theft and physical cloning bitstream encryption is used.

ii) Fault Injection Attacks: involves inserting fault deliberately into the device and to observe erroneous output.

iii) Counterfeiting: to your name illegally on a clone.

iv) Steal bitstreams

v) Insert Trojan Horse: a common method used to capture passwords.

vi) Overbuilding

vii) Cold boot attack: is a technique to extract disk encryption keys [12].

viii) Cloning: in which your design is copied without knowing how it works

ix) Reverse Engineering: Finding out how the design works

x) Steal IP: IP is stolen either with the intention to sell it to others or to reverse engineer. Another classification of attacks on cryptoprocessor as mentioned in [26] is as follows:

A. *Invasive: Invasive attack give direct access to internal components of the cryptographic device. The attack can be performed by manual micro probing, glitch, laser cutting, ion beam manipulation etc.*

B. *Local Non Invasive: This form of attack involves close observation to operation on the device. The side channel attacks listed above may be considered as an example of such an attack.*

C. *Remote Attacks: Remote attacks involve manipulation of device interfaces. Unlike the previous attacks these attacks do not need physical access. API analysis, protocol analysis, cryptanalysis are examples of such an attack. While API analysis is concerned with cryptographic processor cryptanalysis involves finding out the flaws in the algorithms primitives.*

III. IMPLEMENTATIONS OF CRYPTOGRAPHIC ALGORITHMS

Security in the digital world is primarily fulfilled by using cryptography. Numerous optimizations have been proposed and implemented for enhancing the performance and efficiency of the cryptographic algorithms that serve the innumerable applications in various fields. We present few such algorithms which have been implemented on FPGA. The significant consideration of most of them is time area product, besides analysis related to side channel resistance, amount of hardware resources utilized etc.

A. **Symmetric key algorithm implementations** We now discuss few implementations of symmetric key cryptographic algorithms on FPGA. Cryptoraptor [45] considers high performance implementation of set of symmetric key algorithm. The architecture comprises of processing elements(PE) linked by connection row (CR). The PE have independent functional units for arithmetic, shift, logical, table look permutation and operations. Multiplication is limitation due to the limited addressing structure of TLU. It also lacks support for varying modulo in modular arithmetic operations. Rajesh Kannan et al in [46] implement AES, RC5 and RC6 block cipher algorithms in which they discuss on area analysis and power consumptions.

B. Implementations of asymmetric cryptographic algorithms

Many implementations of the asymmetric cryptographic algorithms exist with optimizations to address the needs of embedded system applications. Few of the implementations are as described below.

Base Ext. A→B B→A	Throughput (1024 bits) [Enc/s] (rel.)	Throughput (2048 bits) [Enc/s] (rel.)
M M	194 (46%)	28 (50%)
B M	267 (63%)	38 (67%)
B K	408 (97%)	55 (98%)
B S	419 (100%)	56 (100%)

Table 1 Asymmetric Cryptography with Graphic cards Base Extension Techniques RNS Method [33].

Tim Erhan Guneysu in [33] investigates High Performance Computing implementation of symmetric AES block cipher, ECC and RSA on FPGA.

Feature	ECC(146bits)	RSA(1024bits)
Frequency (MHz)	50	28
Logic size (Slices)	3,036	4,595
Execution time	7.28msec (scalar multiplication)	58.9msec (decryption with 1024-bit sized key)

Table 2 Characteristics of ECC and RSA Crypto Blocks [39]

C. Implementations of hash functions

Hash functions are used for authentication, for providing data integrity and along with public key algorithms as digital signatures. MD5, SHA1, SHA-512 are prominent hash digest algoritms. BLAKE is one of the candidate of SHA3 and Keccak is SHA3 finalist which are based on sponge structure.

Cite this article as: N Rajitha, R Sridevi. "Implementations of Reconfigurable Cryptoprocessor A Survey." *International Conference on Information Engineering, Management and Security (2015)*: 98-103. Print.

Algorithm	Technolog y	Area	Frequenc y	Through - put (Gbps)
Blake-512 [36]	FPGA Virtex 5	108 slice s	358 MHz	0.3
Keccak-1600 [34]	FPGA Stratix III	4684 LUT	206 MHz	8.5

Table 3 Comparison of hardware implementation of Hash functions [38]

D. Implementations of lightweight cryptography

For the fast growing applications of ubiquitous computing, new lightweight cryptographic design approaches are emerging which are investigated in [40]. The implementation of PRESENT-128 lightweight cryptographic algorithm on Spartan III XCS400-5 with a frequency of 254MHz achieves a throughput of 508Mbps

Algorithm	key size	block size	datapath width	cycles / block	T'put [Kbps]	Tech. [μm]	Area [GE]	Eff. [bps/GE]	Cur. [μA]
Serialized Architecture									
DES	56	64	4	144	44.44	0.18	2,309	19.25	1.19
DESL	56	64	4	144	44.44	0.18	1,848	24.05	0.89
DESX	184	64	4	144	44.44	0.18	2,629	16.9	–
DESXL	184	64	4	144	44.44	0.18	2,168	20.5	–

Table 4 Hardware implementation results of DES, DESX, DESL and DESXL. All figures are obtained at or calculated for a frequency of100KHz. [40]

FPGA implementation on low cost Spartan III of ultra light weight cryptographic algorithm Hummingbird is considered in [31]. Hummingbird has its application in RFID tags, wireless control and communication devices and resource constraint devices.

E. A glance on code based cryptography and its implementations

Encryption with Coding Theory by Claude Shannon as basis is used in McEliece and Niederreiter which are considered as candidates for post quantum cryptosystems. McEliece is based on binary Goppa Codes which are fast to decode. McEliece and Niederreiter differ in the description of the codes. While the former cannot be used to generate signatures the later can be used for digital signatures.

Property	Spartan-3an	Virtex-5
Slices	2979	1385
BRAMs	5	5
Clock Frequency	92 MHz	190 MHz
Clock cycles	94,249	94,249
Decryption Latency	1.02 ms	0.50 ms
Security	80 bits	80 bits

Table 5 Mceliece Decryption Implementations [37]

Cite this article as: N Rajitha, R Sridevi. "Implementations of Reconfigurable Cryptoprocessor A Survey." *International Conference on Information Engineering, Management and Security (2015)*: 98-103. Print.

IV. OBSERVATIONS & OPEN QUESTIONS

A. Applications of Cryptoprocessors Numerous applications of cryptoprocessor exist. They can be used in Automated Teller Machine Security, E-commerce applications, smart cards, wireless communication devices, resource constrained devices such as sensors, RFID tags, smart phones, smart cameras, digital rights management, trusted computing, prepayment metering systems, pay per use, banking, military and defense applications.

B. Open Problems

One of the open problems is the remote attacks (in the form of API attack) on cryptoprocessor which may be passive or active and which unlike the physical or invasive attacks doesn't need any contact with the implementation unit. Wollinger et al [47] discuss on the architectures of programmable routing in FPGA in the form of hierarchical and island style. FPGA security resistance to invasive and non-invasive attacks is still under experimentation as new attacks are devised before existing attacks are solved. Much of the work on cryptoprocessors is specific to the application domain or to address a particular form of attack and is not generic to cater to many applications unless customized. Key management in general is not considered as part of th cyrptoprocessor implementation. Several designs of cryptoprocessors are proposed and implemented but still fully functional cryptoprocessor designs addressing integrity, key generation, key management, privacy of both symmetric and asymmetric cryptosystems is still a challenge.

V. ACKNOWLEDGEMENT

The first author would like to express gratitude to TEQIP II. This work has been carried out as a part of Ph D under TEQIP II.

REFERENCES

[1] 1. LILIAN BOSSUET et al Architectures of Flexible Symmetric Key CryptoEngines—A Survey: From Hardware Coprocessor to Multi- Crypto-Processor System on Chip ACM Computing Surveys, Vol. 45,No. 4, Article 41, August 2013.
[2] Crypto-processor - architecture, programming and evaluation of the security Lubos Gaspar Ph D Thesis, November 2012.
[3] Sandro Bartolini, Instruction Set Extensions for Cryptographic Applications, Springer Cryptographic Engineering, 2009.
[4] N. Sklavos , On the Hardware Implementation Cost of Crypto- Processors Architectures, Information Security Journal: A Global Perspective, Taylor& Francis, Vol 19 2010.
[5] Santosh Ghosh et al, BLAKE-512-Based 128-Bit CCA2 Secure Timing Attack Resistant McEliece Cryptoprocessor, IEEE Transactions on Computers 2014.
[6] Siddhartha Chhabra et al An Analysis of Secure Processor Architectures,Springer LNCS 2010.
[7] Sujoy Sinha Roy et al, Compact Ring-LWE Cryptoprocessor, Springer LNCS Vol 8731, 2014
[8] Hans Eberle et al, A Public-key Cryptographic Processor for RSA and ECC, IEEE proceeding 2004
[9] Trimberger and Moore: FPGA Security: Motivations,Features,and Applications, Invited Paper IEEE Proceedings Aug 2014
[10] [Stepahnie Kerchof et al, Towards Green Cryptography: A Comparison of Lightweight Ciphers from the Energy Viewpoint, Springer LNCS Vol 7428, 2012
[11] [Kotaro Okamoto et al , A Hierarchical Formal Approach to Verifying Side-channel Resistant Cryptographic Processors in Hardware-Oriented Security and Trust (HOST), 2014 IEEE International Symposium.
[12] J.Alex Halderman et al Lest We Remember: ColdBoot Attacks on Encryption Keys, Proc.2008 USENIX Security Symposium
[13] Stefan Tillich, Instruction set extensions for support of cryptography on Embedded Systems, Ph D thesis, Graz University of Technology Nov 2008
[14] Michael Grand et al, Design and Implementation of a Multi-Core Crypto-Processor for Software Defined Radios, Springer LNCS Vol 6578 2011
[15] Joel Reardon et al, On secure data deletion, IEEE S&P Symposium, May 2014.
[16] Masoud Rostami et al, A Primer on Hardware Security , Models, Metrics, Vol 102, Proceeedings of IEEE, August 2014.
[17] Hao Zhang et al, In-Memory Big Data Management & Processiing: A Survey, 2014
[18] Peter A. H. Peterson, Cryptkeeper: Improving Security With Encrypted RAM, IEEE 2010.
[19] J.Alex Halderman et al, Lest We remember Cold boot attacks on encryption keys, Usenix 2008.
[20] S Subha, An algorithm for deletion in Flash Memories, IEEE 2009.
[21] Peter Gutmann, Data Remanance in Semiconductor Devices, 2000.
[22] Peter Gutmann, Secure Deletion of Data from Magnetic & Solid-State Memory, Sixth USENIX security Symposium , 1996
[23] Siddhartha Chhabra et al, An analysis of Secure Processor Architecture, AES Key Wrap Specification 2001. Lubos Gaspar et al, Secure extension for soft general purpose processor **Property Spartan-3an Virtex-5** Slices 2979 1385 BRAMs 5 5 Clock Frequency 92 MHz 190 MHz Clock cycles 94,249 94,249 Decryption Latency 1.02 ms 0.50 ms Security 80 bits 80 bits
[24] Lubos Gaspar et al, Secure extension for soft general purpose processor with secure key management, IEEE 2011.
[25] PAWEL SWIERCZYNSKI et al, Physical Security Evaluation of the Bitstream Encryption Mechanism of Altera Stratix II and Stratix III FPGAs, ACM Transactions on Reconfigurable Technology and Systems, Vol. 7, No.4, Article 7,Publication date: December 2014.
[26] MoezBen MBarka Cryptoprocessor application & attacks survey, May 2008
[27] [Mehran Mozaffaci kenani et al, Fault Resilient lightweight cryptographic block cipher for secure embedded system, IEEE Embedded Systems letter, Vol 16, Dec 2014.
[28] Gaurav Bansod et al, Implementation of a new light weight encryption design for embedded security, IEEE Transactions on Information Forensics & Security, 2013.
[29] Majzoobi & Koushnfar, Time bounded Authentications of FPGAs, IEEE Transactions on Information Forensics & Security, Sept 2011
[30] Hero Maderis, Theisis , A Scalar Multiplication in Elliptic Curve Cryptography with Binary Polynomial operation in Galois Field Oct 2009
[31] Xin Xin Fan et al, FPGA Implementation of Humming bird cryptographic algorithm, IEEE International Symposium on Hardware-Oriented Security and Trust (HOST), 2010
[32] Lejla Batina et al, Hardware Architectures for Public Key Cryptography, 2002
[33] Tim Erhan Guneysu, Thesis, Cryptography and cryptanalysis of reconfigurable devices Bochum , 2009
[34] Beuchat et al, Compact Implementation of BLAKE on FPGA, 2010

Cite this article as: N Rajitha, R Sridevi. "Implementations of Reconfigurable Cryptoprocessor A Survey." *International Conference on Information Engineering, Management and Security (2015)*: 98-103. Print.

[35] Baldwin B eta al, FPGA Implementation of SHA3 candidates July 2011
[36] Bertoni, The Keccak sponge function family: Hardware performance 2010
[37] Santosh Gosh et al, A speed area optimized embedded Coprocessor for Mc Eliece Cryptosystem, IEEE Conference 2012
[38] Zhije Shi et al, Hardware Implementation of Hash Function, Springer LLC 2012
[39] HoWon Kim et al, Design and Implementation of public key cryptoprocessor and its application to a security system
[40] Axer York Poschmann, Ph D Thesis, Lightweight cryptography Feb 2009
[41] Ricardo Chaves, Ph D Thesis, Secure Computing in reconfigurable devices, 2007
[42] PowerKotaro Okamoto, A Hierarchical Formal Approach to Verifying Side-channel Resistant Cryptographic Processors, IEEE, 2014
[43]] Amir Moradi ,Side-Channel Leakage through Static Power Should we care in practice
[44] Jen-Wei Lee, Efficient Power-Analysis-Resistant Dual-Field Elliptic Curve Cryptographic Processor Using Heterogeneous Dual-Processing- Element Architecture, IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, VOL. 22, NO. 1, JANUARY 2014
[45] Gokhan Sayiler, Cryptoraptor: High Throughput Reconfigurable Cryptographic Processor, IEEE 2014.
[46] Rajesh Kannan et al Recongfigurable Cryptoprocessor for multiple crypto Algorithm, IEEE Symposium 2011.

Cite this article as: N Rajitha, R Sridevi. “Implementations of Reconfigurable Cryptoprocessor A Survey.” *International Conference on Information Engineering, Management and Security (2015)*: 98-103. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS017

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.017

BER and PAPR Performance Analysis of MIMO System for WIMAX (IEEE 802.16) Systems

Jitendra Jain[1], Lawkush Dwivedi[2]
[1]Assistant professor, Bansal Institute of Science Technology, Bhopal, India
[2]Mtech Scholar, Bansal Institute of Science Technology, Bhopal, India

Abstract: *This paper investigates the multiple input multiple output (MIMO) space-time coded wireless systems. MIMO-OFDM system to improve the reliability of the WiMAX (IEEE 802.16) system. This paper discusses the model building of MIMO-OFDM using MATLAB R2012b version. This model is a using tool for BER (Bit Error Rate), PAPR (Peak Average Peak Ratio) and transmits spectrum performance evaluation for signal & multiple input output port by the WiMAX (IEEE 802.16) system. In this paper, transmitter and receiver model are analysis according to the parameters established by the standards, to evaluate the performance parameter.*

Keywords: WiMAX, OFDM, RAYLEIGH CHANNEL, MIMO-OFDM, BER, PAPR

I. INTRODUCTION

Wireless communications is a rapidly growing part of the communications field, with the believable to provide high-speed and high-quality information swap between portable devices located anywhere in the world. It has been the topic of study since last two decades the terrific development of wireless communication technology is due to several factors. The demand of wireless connectivity is exponentially increased. Second, the dramatic progress of VISL technology has enabled small-area and low-power implementation of sophisticated signal processing algorithm and coding algorism. Third, wireless communication standards, like CDMA, GSM, TDMA, make it possible to transmit voice and low volume digital data. Further, third generation of wireless communications can offer users more advanced service that achieves greater capacity through improved spectral efficiency [1]. Potential applications enabled by this technology include multimedia cell phones, smart homes and appliances, automated systems, video teleconferencing and distance learning, and autonomous sensor networks. However, there are two significant technical challenges in supporting these applications first is the phenomenon of fading the time variation of the channel due to small-scale effect of multi-path fading, as well as large-scale effect like pass loss by distance attenuation and shadowing by obstacles. Second,since wireless transmitter and receiver need communicate over air, there is significant interference between them [2]. Overall the challenges are mostly because of limited availability of radio frequency spectrum and a complex time-varying wireless environment (fading and multipath). The OFDM system of the WiMAX adopts abruptly deliver mode, reliability, good efficiency and High data rate is achieved between the transmitter and the receiver if they are ideally synchronized [3-4].
However, there usually exists a small timing and frequency offset whose exists will dramatically degrade the performance of the whole OFDM systems. Hence, before signals can be demodulated, OFDM symbols have to be time-synchronized and carrier frequency offset

Cite this article as: Jitendra Jain, Lawkush Dwivedi. "BER and PAPR Performance Analysis of MIMO System for WIMAX (IEEE 802.16) Systems." *International Conference on Information Engineering, Management and Security (2015)*: 104-109. Print.

compensated. This puts forward very high request to the mode piece of the synchronization system. In order to realize the synchronization, it must adopt synchronization algorithm of smaller calculation quantity. In the meantime, it should have higher examination of the first moment [5]. In nowadays, the key goal in wireless communication is to increase data rate and improve transmission reliability. In other words, because of the increasing demand for higher data rates, better quality of service, fewer dropped calls, and higher network capacity that improve spectral efficiency and link reliability, more technologies in wireless communication are introduced, like OFDM, MIMO and MIMO-OFDM [6]. This paper is organized as follows: In section II, the orthogonal frequency division multiplexing (OFDM) system and multiple input multiple output OFDM (OFDM-MIMO) system is formulated. Space time block code is introduced in section III. In section IV discussed about previous and proposed model and simulation result. Finally, the conclusions are given in section V.

II. OVERVIEW OF OFDM AND MIMO SYSTEM

o OFDM

Orthogonal frequency-division multiplexing (OFDM) is a method of digital modulation in which the data stream is split into N parallel streams of reduced data rate with each of them transmitted on separate subcarriers. In short, it is a kind of multicarrier digital communication method. OFDM has been around for about 40 years and it was first conceived in the 1960s and 1970s during research into minimizing interference among channels near each other in frequency [2] .OFDM has shown up in such disparate places as asymmetric DSL (ADSL) broadband and digital audio and video broadcasts. OFDM is also successfully applied to a wide variety of wireless communication due to its high data rate transmission capability with high bandwidth efficiency and its robustness to multi-path delay [7-8]. The basic principle of OFDM is to split a high data rate streams into a number of lower data rate streams and then transmitted these streams in parallel using several orthogonal sub-carriers (parallel transmission). Due to this parallel transmission, the symbol duration increases thus decreases the relative amount of dispersion in time caused by multipath delay spread. OFDM can be seen as either a modulation technique or a multiplexing technique.

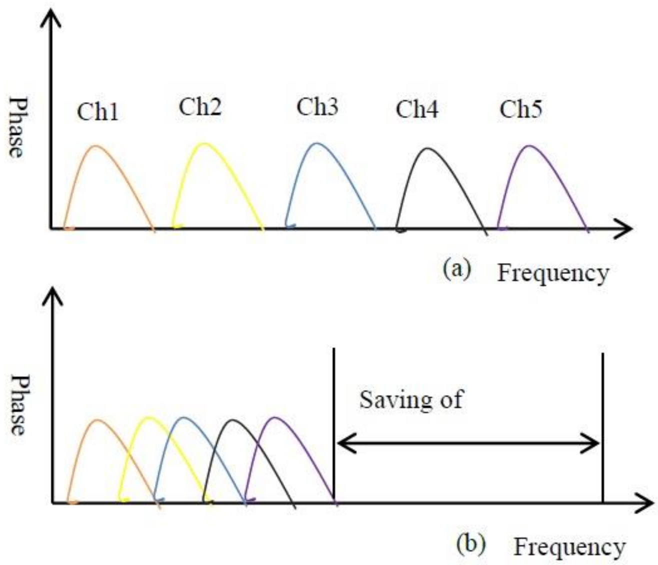


Figure 1: Comparison between conventional FDM (a) and OFDM (b)

o MIMO

MIMO has been developed for many years for wireless systems. One of the earliest MIMO to wireless communications applications came in mid-1980 with the breakthrough developments by Jack Winters and Jack Saltz of Bell Laboratories [9]. They tried to send data from multiple users on the same frequency/time channel using multiple antennas both at the transmitter and receiver. Since then, several academics and engineers have made significant contributions in the field of MIMO. Now MIMO technology has aroused interest because of its possible applications in digital television, wireless local area networks, metropolitan area networks and mobile communication. Comparing to the Single-input-single-output (SISO) system MIMO provides enhanced system performance under the same transmission conditions. First, MIMO system greatly increases the channel capacity, which is in proportional to the total number of transmitter and receiver arrays. Second, MIMO system provides the advantage of spatial variety: each one transmitting signal is detected by the whole detector array, which not only improved system robustness and reliability, but also reduces the impact of ISI (inter symbol interference) and the channel fading since each signal determination is based on N detected results. In other words, spatial diversity offers N independent replicas of transmitted signal. Third, the Array gain is also increased, which means SNR gain achieved by focusing energy in desired direction is increased.

Cite this article as: Jitendra Jain, Lawkush Dwivedi. "BER and PAPR Performance Analysis of MIMO System for WIMAX (IEEE 802.16) Systems." *International Conference on Information Engineering, Management and Security (2015)*: 104-109. Print.

o MIMO-OFDM

OFDM reduces BER performance and ISI with using multiplexing and modulation techniques to get higher data rate over wireless channels, the use of multiple antennas at both ends of the wireless link provide better performance. The MIMO technique does not require any extra transmission power and bandwidth. Therefore, the promising way to increase the spectral efficiency of a system, the combination of MIMO and OFDM is used over fading channels [10-11].

III. SPACE TIME BLOCK CODE

Multiple-Input Multiple-Output uses multiple antennas at both sides which provides transmit diversity and receiver diversity. It's applicable in every kind of networks like PAN, LAN, WLAN, WAN, MAN. MIMO system can be applied in different ways to receive either a diversity gain, capacity gain or to overcome signal fading.

Space-frequency coding basically extends the theory of space-time coding for narrowband flat fading channels to broadband time-variant and frequency-selective channels. The application of classical space-time coding techniques for narrowband flat fading channels to OFDM seems straightforward, since the individual subcarriers can be seen as independently flat fading channels. However, it was shown that the design criteria for space-frequency codes operating in the space-time and frequency domain are different from those for classical space-time codes for narrowband fading channels as introduced in. When operating in frequency selective fading channels, the application of conventional decoding algorithms results in a significant performance decrease [12]. This is due to the fact that the equivalent channel matrix is no longer orthogonal. Consequently, independent decoding of the two transmitted symbols, as in conventional decoding algorithms, is no longer appropriate.

IV. SIMULATION RESULT

Simulation experiments are conducted to evaluated the transmit spectrum, BER, PAPR reduction performance of the proposed scheme and the OFDM scheme. In addition, it is assumed that the data are QPSK, BPSK, 16-QAM modulated and are transmitted using N=256 sub-carrier.

The following subsection presents the simulation results using the OFDM and MIMO-OFDM model in figure 2, 3, 4, 5, 6 and 7 for WiMAX IEEE 802.16.

Figure 2 & 3, shows the transmit spectrum WiMAX OFDM with QPSK, QAM-16 and BPSK and transmit spectrum WiMAX MIMO-OFDM 2*2 respectively.

In our simulation binary phase-shift keying (BPSK) modulation, quadrature phase-shift keying (QPSK) modulation and quadrature amplitude modulation (QAM) will be used; the impairments of the channel include Rayleigh fading.

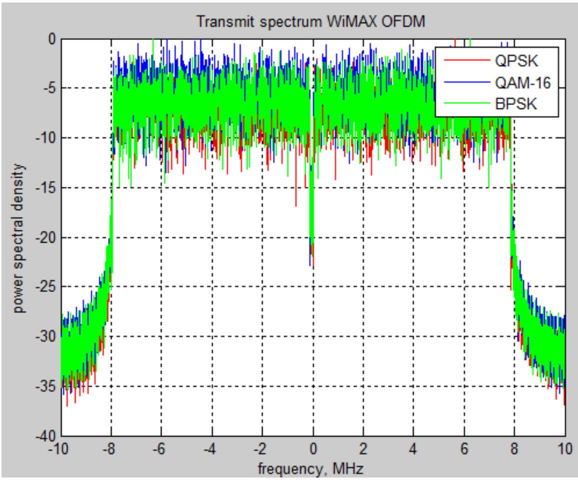


Figure 2: Simulation result of QPSK, QAM-16 and BPSK modulation in transmit spectrum WiMAX OFDM

In figure 4 the CCDF plot for SISO OFDM system of WiMAX (IEEE 802.16e) is shown. This research performs a series of simulations to evaluate PAPR performances of the OFDM system. The simulations assume the data were QPSK, QAM-16 and BPSK modulated and the system contained N-256 sub-carriers. The BPSK modulation technique is better than other technique, because the error performance of BPSK is better than other technique as we can see in figure 4.

Cite this article as: Jitendra Jain, Lawkush Dwivedi. "BER and PAPR Performance Analysis of MIMO System for WIMAX (IEEE 802.16) Systems." *International Conference on Information Engineering, Management and Security (2015)*: 104-109. Print.

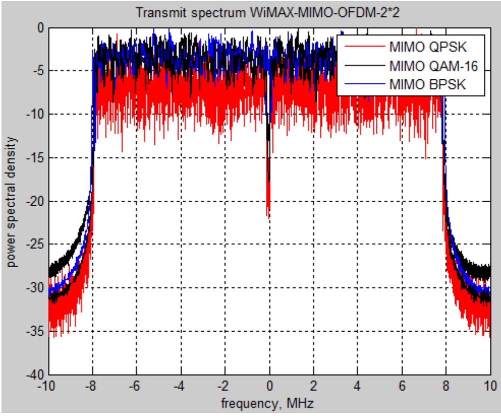


Figure 3: Simulation result of QPSK and QAM-16 modulation in transmit spectrum WiMAX MIMO-OFDM 2*2.

The CCDF is generally used to evaluate the performance of PAPR reduction on MIMO-OFDM system (IEEE 802.16e) signals for a statistical pair of view. The CCDF is defined as the probability that the PAPR as in equation (1) and $PAPR_0$ as shown in the following:

$$PAPR\{Y\} = \arg \underbrace{\max}_{k=1,2,3,\ldots\ldots N_T} (PAPR\{Y_k\}) \qquad (1)$$

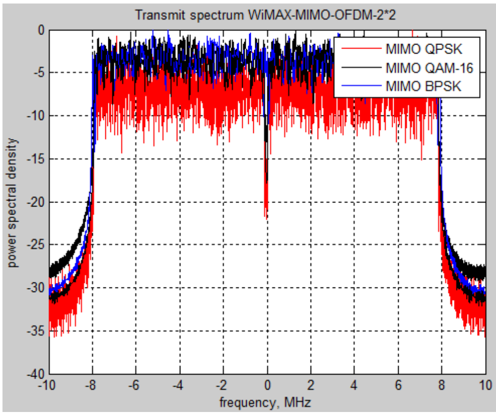


Figure 3: Simulation result of QPSK and QAM-16 modulation in transmit spectrum WiMAX MIMO-OFDM 2*2.

Where , represents the time-domain transmitted signal of the k-th antenna

$$CCDF(PAPR_0) = \Pr(PAPR\{Y\} > \{PAPR_0\}) \qquad (2)$$

Figure 5 presents the CCDF graph of PAPR for the MIMO-OFDM 2×1& MIMO-OFDM 2*2 system of WiMAX (IEEE 8002.16e). Figure 5 present the CCDF of PAPR for STBC algorithm in the condition described above.

The green line curve corresponds to the MIMO-OFDM 2*1 BPSK signal and blue line curve corresponding to the MIMO-OFDM 2*2 BPSK signal. MIMO-OFDM 2×1 system is better result compare to the MIMO-OFDM 2×2 system.

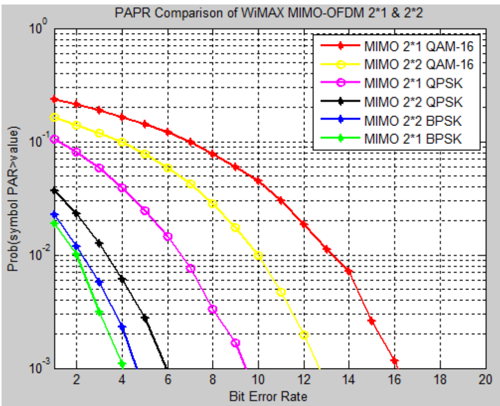


Figure 5: Simulation result of QPSK, QAM-16 and BPSK modulation in PAPR Performance of WiMAX MIMO-OFDM 2*1 and 2*2 systems.

Cite this article as: Jitendra Jain, Lawkush Dwivedi. "BER and PAPR Performance Analysis of MIMO System for WIMAX (IEEE 802.16) Systems." *International Conference on Information Engineering, Management and Security (2015)*: 104-109. Print.

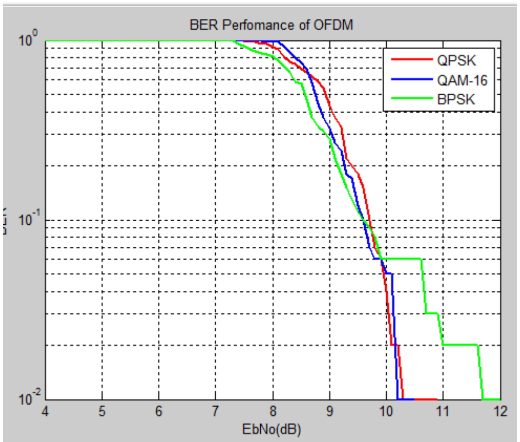


Figure 6: Simulation result of QPSK, QAM-16 and BPSK modulation in BER Performance of WiMAX OFDM

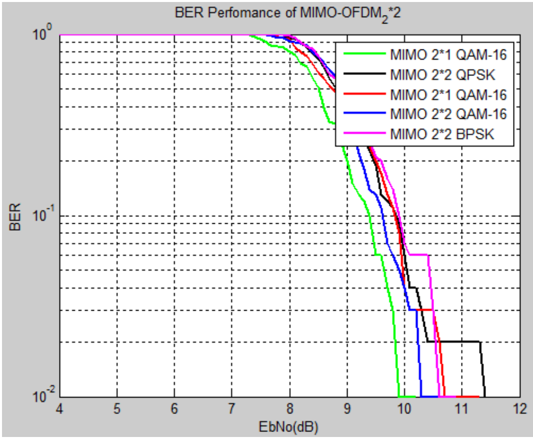


Figure 7: Simulation result of QPSK and QAM-16 modulation in BER Performance of WiMAX MIMO-OFDM 2*1 and 2*2 systems.

V. CONCLUSION

We know that a tradeoff between peak power peak ratio (PAPR) and bit error rate for WiMAX IEEE 802.16. In this paper presented low-complexity transmitter architecture for STBC MIMO-OFDM system. The proposed SBTC MIMO-OFDM 2*1 and MIMO-OFDM 2*2 scheme could offer good PAPR reduction, which is almost the same as that of OFDM system. The previous scheme used only single input single output. However, the proposed scheme designs for multiple inputs and multiple outputs. Therefore, the proposed SBTC MIMO-OFDM scheme has better bandwidth efficiency and BER performance compared with the previous scheme.

VI. ACKNOWLEDGMENT

I would like to say thanks to my guide “Dr. Ashutosh Sharma”, Director “Dr. A.K. Singh and who gives their knowledge and time in order to complete this paper. This paper will never complete without the support faculty members of ECE department of Bhopal Institute of Technology, Bhopal

REFRENCES

[1] Kumar, “Introduction to Broadband Wireless Networks” in Mobile Broadcasting with WiMAX: Principles, Technology and Applications, New York, USA: Focal Press, 2008, pp. 24-50.

[2] C. Eklund, R. B. Marks, K. L. Stanwood and S. Wang, “IEEE Standard 802.16: a Technical Overview of the WirelessMAN Air Interface for Broadband Wireless Access,” *IEEE Commun. Mag.*, vol. 40, pp. 98-100, Jun 2002.

[3] S. J. Vaughan-Nichols, “Mobile WiMAX: The Next Wireless Battleground,” in *IEEE Comp. Soc. Mag.*

[4] L. Koffman and V Roman, “Bradband wireless access solution based in IEEE 802.16,” IEEE communication Magazine, vol. 40, pp 4, Apr. 2004pp 96-103.

[5] K. Y. Cho, B. S. Choi, Y. Takushima, and Y. C. Chung, B25.78-Gb/s operation of RSOA for next-generation opticalaccess networks,[IEEE Photon. Technol. Lett., vol. 23, no. 8, pp. 495–497, Apr. 2011.

[6] J. Zhang and N. Ansari, BToward energy-efficient 1G-EPON and 10G-EPON with sleep-aware MAC control and scheduling,[IEEE Commun. Mag., vol. 49, no. 2, pp. s33–s38, Feb. 2011.

[7] A. Islam, M. Bakaul, A. Nirmalathas, and G. E. Town, BMillimeter-wave radio-over-fiber system based on heterodynedunlocked light sources and self-homodyne RF receiver,[IEEE Photon. Technol. Lett., vol. 23, no. 8, pp. 459–461,Apr. 2011.

Cite this article as: Jitendra Jain, Lawkush Dwivedi. “BER and PAPR Performance Analysis of MIMO System for WIMAX (IEEE 802.16) Systems.” *International Conference on Information Engineering, Management and Security (2015)*: 104-109. Print.

[8] M. Daneshmand, C. Wang, and W. Wei, BAdvances in passive optical networks,[IEEE Commun. Mag., vol. 49, no. 2,pp. s12–s14, Feb. 2011.

[9] Different Modulation Techniques used in WiMAX , International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 4,April 2013.

[10] PrabhakarTelagarapu,PrabhakarTelagarapu,K.Chiranjeevi, "Analysis of Coding Techniques in WiMAX", International Journal of Computer Applications ,Volume 22–No.3,May 2011.

[11] Performance of Coding Techniques in Mobile WiMAX Based System" International Journal on Recent and Innovation Trends in Computing and Communication,Volume: 1 Issue: 1,2009.

[12] [Mukesh patidar, Rupesh Dubey and Nitin kumar jain, "Performance Analysis of WimAX 802.16e Physical Layer Moodel", 978-1-4673-1989-8/12/$31.00@2012 IEEE.

Cite this article as: Jitendra Jain, Lawkush Dwivedi. "BER and PAPR Performance Analysis of MIMO System for WIMAX (IEEE 802.16) Systems." *International Conference on Information Engineering, Management and Security (2015)*: 104-109. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS018		**eAID**	ICIEMS.2015.018

A Review on Feature Extraction Techniques for CBIR System

Kavita Chauhan[1], Shanu Sharma[1]
[1]Department of Computer Science & Engineering,
ASET, Amity University, Uttar Pradesh, Noida, India

ABSTRACT: *Ongoing expansion of digital images requires improved and efficient methods for sorting, browsing and searching operations through ever-growing image databases. Content Based Image Retrieval (CBIR) systems are search engines for image databases, which perform indexing on images according to their content and features. This paper presents the systematic review of various existing CBIR systems and their feature extraction techniques. Further the performance analysis and limitations of these systems have been discussed.*

KEYWORDS: CBIR, Image Feature Extraction, Similarity Measurement, Neural Network, Support Vector Machine.

I. INTRODUCTION

The advancement in computer technologies produces huge volume of multimedia data, specifically on image data. The greatest challenge of the World Wide Web is that the more information available about a given topic, the more difficult it is to locate accurate and relevant information. Generally users know that which information they need, but are unsure where to find it. Search engines can facilitate the ability of the users to locate such relevant information.

Content Based Image Retrieval is a technique which uses visual content to search and compares images from large scale image databases according to the interest of the users [1]. In this process firstly, the user submits a query image or a series of images and the system is required to retrieve images from the database as similar as possible. It also includes another task which is a support for browsing through large image databases, where the images are supported to be grouped or organised in accordance with similar properties [2]. During the past few years, CBIR has gained much devotion for its potential application in multimedia management. The term 'content' in this context might refer to colors, shapes, texture, or any other information that can be derived from the image itself. Basically, there are two ways of image retrieval in CBIR- Query by tag and Query by example. In the former method, the query is submitted in the form of tag like square is used for searching square shape in image, and in the later method the query is given with the image example such that the results resemble the given query [2]. CBIR is also known as Query by Image Content (QBIC) and Content-Based Visual Information Retrieval (CBIR).

This poses two main challenges for the Image Retrieval researchers and practitioners:

a. Low level features extracted and their semantic meanings may differ thus forming a gap known as the "Semantic Gap"[3]

b. Granularity of classification, this granularity is closely related to the level of invariant that the CBIR system should guarantee [3].

This paper is divided into four sections. Section I contains the introduction of the CBIR system and their challenges. Section II discusses the result of the various feature extraction techniques in the form of a table defining the various attributes and how they differ from each other. Section III concludes this paper followed by future scope.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.

II. RELATED WORK

A thorough and systematic literature has been done on various CBIR systems. The origin of major studies vary from listed repositories (IEEE, ACM Library, IJCA, Science Direct etc.) to common purpose search engines like Google. Various research papers with the search string 'Feature Extraction Techniques and CBIR System' has been searched and finally 65 papers have been downloaded out of which 17 papers were considered of most relevant to the objective. Review for the selected paper is further presented: Ligade *et al.* [3] provides a review on techniques of Neural Network, Interactive Genetic Algorithm and Relevance Feedback where the characteristics of the above 3 techniques has been described along with their current achievement, uses and advantages in image retrieval. The experimental evaluation is done by them using the convergence ratio, precision and recall parameters. Walia *et al.* [4] proposed a new similarity measure that improves the efficiency of the CBIR system by using the dominant color descriptor. Their work compares two images and consider their no. of dominant colors and their distance and thus improves the performance using the colors and also the results were verified on two different image databases. Sarangi *et al.* [5] proposed an automatic contrast enhancement technique using differential evolution for gray-scale images. The technique attempts to demonstrate the methods adaptability and effectiveness for searching global optima solutions to enhance the contrast and detail in gray scale images. Agarwal *et al.* [6] proposed a novel feature descriptor for CBIR system by integrating Cooccurence of Haar like Wavelet Filter (CHLWF) with Color Histogram. It extracts the image properties from different visual perspectives to give the image representation almost similar to the human interpretation and hence improves the effectiveness of retrieval system. Jeyabharathi *et al.* [7] analyzed the performance of the most useful feature extraction techniques which includes Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Independent Component Analysis (ICA) and also the performance of the most renowned classification techniques, i.e., Support Vector Machine (SVM) and Nearest Neighbour (NN). The performance metric used by them were Recognition Rate and F-Score. Based on the performance evaluation models, they concluded that PCA with SVM provide more recognition accuracy than others. Ezekiel *et al.* [8] proposed a CBIR technique based on multi-scale feature extraction scheme. They designed a Pulse Coupled Neural Network (PCNN) based fusion of a fast wavelet transformation and Contourlet transformation coefficients applied on Rescaled Range(R/S) analysis techniques. The method highlights the edges, segment edges and finds control points to answer the image retrieval query. Zhang *et al.* [9] introduced a method of color principal feature extraction called ColorPCA which works in color image space extracting the principal features directly from the color images. It considers only one parameter known as reduced dimension to estimate the projection axes. Syam *et al.* [10] proposed a CBIR system based on GA for Medical image retrieval using the feature extraction of color, texture and shape. They used Squared Euclidean Distance (SED) for the computation of similarity measure for efficient retrieval of images. Their work assures the benefit of the shape feature in addition to the other features. Ligade *et al.* [11] proposed an image extractor method on multi-feature similarity synthesis using the GA for efficient image retrieval in CBIR systems. The method extends the GA algorithm by using the methods of relevance feedback for magnified retrieval performance, also the method uses both the implicit and explicit feedback technique. Ho *et al.* [12] proposed a novel system architecture for CBIR system in which the well-known techniques are combined like content-based image, color analysis and data mining techniques for better performance and efficiency. It combines the segmentation and grid module, feature extraction module, K-means clustering and neighbourhood module to build the CBIR system. Chadha *et al.* [13] proposed an improved technique of image retrieval by incorporating the Query modification through image cropping. This feature identifies the user's interest region in a particular image and thus resulted into more precise and personalized search results. This technique results into 28% improvement in accuracy. Rashedi *et al.* [14] proposed a method for the CBIR system improving its precision by Feature selection using the Binary Gravitational Search Algorithm. It selects the most relevant features from the query image thus leading to more accurate results by reducing the semantic gap. The results are examined in the Corel database. Also they compared the work of GA and BPSO and found BGSA to be the best among them. Pighetti *et al.* [15] proposed a new architecture combining the multi-objective interactive genetic algorithm and Support Vector Machine. The multi-objective IGA is used for its capability to converge towards global optima and SVM for its capabilities to learn user evaluations required by IGA. Madugunki *et al.* [16] described the detail classification of the CBIR system and also about their efficiency. The work compares the CBIR and TBIR technique and discusses about the effect of different matching techniques on retrieval process. Euclidean Distance Method, City Block Distance and Canberra Distance Method are used to calculate the matching distance and found Canberra Distance Method best among others. Selvarajah *et al.* [17] introduced a descriptor called Combined Feature Descriptor for the CBIR system to enhance the retrieval performance. It uses the concept of Haar Wavelet and color histogram moment. The descriptor works for the application which includes traditional color moments, 2D-Discrete Wavelet Transform. Abubacker *et al.* [18] proposed a CBIR technique based on query and extracts the most vital attributes, i.e., color, shape and texture. It includes the automatic extraction of spatial based color feature using invariant Fourier descriptor and texture feature using the Gabor filter. The distance metrics were used for distance calculation and their weightage were determined. Based on the data, the output images are sorted and ranked so that most similar images can be displayed to the user. Omar *et al.* [19] proposed a WhatAreYouLOOKing4(WAY-LOOK4) system using the Local descriptor and Image Signatures. It contains 3 system components: feature extraction, image database indexing and similarity retrieval. The system maintains reasonable storage and computational costs. The system is simple because of no iterations for clustering or complex wavelet transformation.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.

The analysis of different techniques is presented in the following table.

Table 1: Analysis of Different Feature Extraction Techniques

S.No.	Author	Year	Source	Major Findings/Limitations/Future Scope
1	Ligade *et al.*	2014	IJCSIT	Reviews NN, IGA and RF using the convergence ratio, precision and recall and provides the various CBIR system based RF techniques.
2	Walia *et al.*	2014	IEEE	1. Proposed a new similarity measure that improves the efficiency of the CBIR system by using the dominant color descriptor. 2. In future, it can be used for color based retrieval of videos.
3	Sarangi *et al.*	2014	IEEE	Proposed an automatic contrast enhancement technique using differential evolution for gray-scale images
4	Agarwal *et al.*	2014	IEEE	Proposed a novel feature descriptor for CBIR system by integrating Cooccurence of Haar like Wavelet Filter (CHLWF) with Color Histogram
5	Jeyabharathi *et al.*	2013	IEEE	Analyses the performance of feature extraction techniques (PCA, LDA & ICA) and classification techniques (SVM & NN) using the Recognition Rate and F-Score and finds PCA with SVM to the best.
6	Ezekiel *et al.*	2013	IEEE	Proposed a multi-scale feature extraction scheme based on wavelet and Contourlet transform
7	Zhang *et al.*	2013	IEEE	1. Proposed a method of color principal feature extraction called ColorPCA 2. This method contains three main problems: a. Nonlinear structures remains unclear on how to extend the ColorPCA algorithm. b. To extend the ColorPCA idea in manifold structures or discriminant information of samples for locality preservation c. To determine the optimal reduced dimensions for dimension reduction algorithms using ColorPCA.
8	Syam *et al.*	2013	IEEE	Proposed a CBIR System using the GA and SED with feature extraction process.
9	Ligade *et al.*	2013	TJPRC	1. Proposed an image extractor method on multi-feature similarity synthesis using the Genetic Algorithm fpr efficient image retrieval in CBIR systems 2. They consider only the occurrence frequencies of image in result and not the location of the retrieval result which directly reflects the similarity of it and the query image.
10	Ho *et al.*	2012	IEEE	Proposed an architecture for CBIR system combining content-based image, color analysis and data mining.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.

11	Chadha *et al.*	2012	IJCA	Proposes an improved technique for image retrieval using the Query modification through image cropping.
12	Rashedi *et al.*	2012	IEEE	Proposed a method to improve the precision of CBIR system by Feature selection using the Binary Gravitational Search Algorithm
13	Pighetti *et al.*	2012	ICPR	Proposed a new architecture combining the multi-objective IGA and SVM.
14	Madugunki *et al.*	2011	IEEE	Compares the CBIR and TBIR techniques of image retrieval and also the matching technique and found Canberra Distance Method best among other matching technique.
15	Selvarajah *et al.*	2011	IEEE	Introduced Combined Feature Descriptor for the CBIR system to enhance the retrieval performance using Wavelet and Color Histogram concepts.
16	Abubacker *et al.*	2010	IEEE	1.Proposes a CBIR technique based on query and vital attributes. 2.Scanner& camera support can be aimed. 3. Different query formats support incorporating filtering & indexing mechanism can be used for quick image access from the database.
17	Omar *et al.*	2009	IEEE	1. Proposes a WhatAreYouLOOKing4(WAY-LOOK4) system using the Local descriptor and Image Signatures. 2. System performance degrades in case of fine details like buildings and food.

III. CONCLUSION

In this work, overview of various feature extraction techniques for Content Based Image Retrieval System is explained briefly. This paper classifies the current methods and summarizes their features. After having a review, it can be considered that no single method can be resulted as best or very good for all types of images or all the methods uniformly good for a specific image type. Considering all these limitations and major findings, CBIR system remains to be a challenging problem in image processing. Feature extraction techniques for CBIR system is still a pending problem in the world and more research need to be carried out for better estimation.

REFERENCES

[1] R. Choudhary, N. Raina, N. Chaudhary, R. Chauhan and R H Goudar," - An Integrated Approach to Content Based Image Retireval", International Conference on Advances in Computing, Communications and informatics (ICACC1), pp. 2402-2410, IEEE, 2014.

[2] B. Syam and Y. Rao, " An Effective Similarity Measure via Genetic Algorithm for Content Based Image Retrieval with Extensive Features", The International Arab Journal of Information Technology, Volume 10, pp. 143-151, 2013.

[3] A. N. Ligade and M. R. Patil, " Content Based Image Retrieval Using Interactive Genetic Algorithm with Relevance Feedback Technique – Survey", International Journal of Computer Science and Information Technologies, Vol. 5(4), pp. 5610-5613, IJCSIT, 2014.

[4] E. Walia, P. Saigal and A. Pal," Enhanced Linear Block Algorithm with Improved Similarity Measure", Canadian Conference on Electrical and Computer Engineering, pp. 1-7 , IEEE, 2014.

[5] P.P. Sarangi, B.S.P. Mishra, B. Majhi and S. Dehuri," Gray-level Image Enhancement Using Differential Evolution Optimization Algorithm", International Conference on Signal Processing and Integrated Networks (SPIN), IEEE, 2014.

[6] M. Agarwal," Integrated Features of Haar-like Wavelet Filters," 7th International Conference on Contemporary Computing, pp. 370-375, IEEE, 2014.

[7] D. Jeyabharathi and A. Suruliandi," Performance Analysis of Feature Extraction and Classification Techniques in CBIR", International Conference on Circuits, Power and Computing Technologies, pp. 1211-1214, IEEE, 2013.

[8] S. Ezekiel, M.G. Alford, D. Ferris and E. Jones," Multi-Scale Decomposition Tool for Content Based Image Retrieval", Applied Imagery Pattern Recognition Workshop: Sensing of Control and Augmentation (AIPR), pp. 1-5, IEEE, 2013.

[9] Z. Zhang, M. Zhao, B. Li and P. Tang," ColorPCA: Color Principal Feature Extraction Technique for COlor Image Reconstruction and Recognition", International Joint Conference on Neural Networks (IJCNN), pp. 1-7, IEEE, 2013.

[10] B. Syam, S. R. Victor J and Y.S. Rao," Efficient Similarity Measure via Genetic Algorithm for Content Based Medical Image Retrieval with Extensive Features", International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing, pp. 704-711, IEEE, 2013.

[11] A. N. Ligade and M. R. Patil," Optimized Content Based Image Retrieval Using Genetic Algorithm with Relevance Feedback Technique", International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR), pp. 49-54, TJPRC, 2013

[12] J.M. Ho, S.Y. Lin, C.W. Fann, Y.C. Wang and R.I Chang," A Novel Content Based Image Retrieval System using K-means with Feature Extraction", International Conference on Systems and Informatics (ICSAI), pp. 785-790, IEEE, 2012.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.

[13] A. Chadha, S. Mallik and R. Johar," Comparative Study and Optimization of Feature Extraction Techniques for Content Based Image Retrieval", International Journal of Computer Applications, Vol. 52-No. 20, pp. 35-42, IJCA, 2012.

[14] E. Rashedi and H.Nezamabadi-pour," Improving the Precision of CBIR Systems by Feature Selection Using Binary Gravitational Search Algorithm", International Symposium on Artificial Intelligence and Signal Processing (AISP), pp. 039-042, IEEE, 2012.

[15] R. Pighetti, D. Pallez and F. Precioso," Hybrid Content Based Image Retrieval combining Multi-objective Interactive Genetic Algorithm and SVM", 21st International Conference on Pattern Recognition (ICPR), pp. 2849-2852. ICPR, 2012.

[16] M. Madugunki, D.S. Bormane, S. Bhadoria and C.G. Dethe," Comparison of Different CBIR Techniques", International Conference on Electronics Computer Technology, pp. 372-375, IEEE, 2011.

[17] S. Selvarajah and S. R. Kodithuwakku," Combined Feature Descriptor for Content Based Image Retrieval", 6th International Conferencec on Industrial and Information Systems (ICIIS), pp. 164-168, IEEE, 2011.

[18] [18] K.A.S. Abubacker and L.K. Indumathi," Atrribute Associated Image Retrieval and Similarity Reranking", Proceedins of the International Conference on Communication and Computational Intelligence, pp. 235-240, IEEE, 2010.

[19] S.G. Omar, M.A. Ismail and S.M. Ghanem," WAY-LOOK4: A CBIR System Based on Class Signature of the Images' Color and Texture Features", International Conference on Computer Systems and Applications, pp. 464-471, IEEE, 2009.

Cite this article as: Kavita Chauhan, Shanu Sharma. "A Review on Feature Extraction Techniques for CBIR System." *International Conference on Information Engineering, Management and Security (2015)*: 110-114. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS019		**eAID**	ICIEMS.2015.019

Top K Sequential Pattern Mining Algorithm

Karishma B Hathi[1], Jatin R Ambasana[2]
[1]Student, M.E.(CSE), Gardi Vidyapith, Gujarat, India
[2]Assistant Professor, CSE Department, Gardi Vidyapith, Gujarat, India

ABSTRACT: *Sequential pattern mining is a very chief mining technique with wide applications. Still, tune up the minsup parameter of sequential pattern mining algorithms to produce enough patterns is complex and time-consuming. To solve this problem, the assignment of top-k sequential pattern mining has been described, here k is the number of sequential patterns to be discovered, and is set by the user. In this paper, we present proposed approach for improving parameters in TKS Algorithm.*

***KEYWORDS*:** Sequential Patterns, Top K, Sequence Database, Pattern mining.

I. INTRODUCTION

The sequential pattern mining is a very important concept of data mining, a further extension to the concept of association rule mining [1]. That has a huge range of real-life application. This mining algorithm solves the problem of discovering the presence of frequent sequences in the given database [2]. Sequential Pattern Mining finds interesting sequential patterns among the huge database. It discovers frequent subsequences as patterns from a given sequence database. It is a well-understood data mining problem with broad applications such as the analysis of web clickstreams, program executions, medical data, biological data and e-learning data [1, 5]. Although many studies have been done on constructing sequential pattern mining algorithms [1, 2, 3, 4], the main problem is how the user should choose the *minsup* threshold to produce a desired amount of patterns. This problem is important because in practice, users have limited resources (time and storage space) for discovering the results and thus are often only interested in analyzing a certain amount of patterns, and fine-tuning the *minsup* parameter is very time-consuming. Depending on the choice of the *minsup* threshold, algorithms can become very slow and produce an extremely huge amount of results or generate none or too few results, getting valuable information. To address this difficulty, it was proposed to redefine the problem of mining sequential patterns as the problem of mining the top-*k* sequential patterns, where k is the number of sequential patterns to be discovered and is set by the user.

II. RELATED WORK

The problem of sequential pattern mining was proposed by Agrawal and Srikant [2] and is defined as follows. A *sequence database SDB* is a set of sequences $S = \{s1, s2\ldots, ss\}$ and a set of items $I = \{i1, i2, \ldots, im\}$ happening in these sequences. An *item* is a symbolic value. An *itemset* $I = \{i1, i2, \ldots, im\}$ is an unordered set of different items. For example, the itemset $\{a, b, c\}$ shows the sets of items a, b and c. A *sequence* is an ordered list of itemsets S=< $I1, I2, I3, \ldots, In,$ > such that $Ik \subseteq I$ for all 1≤k≤n .For example having a sequence the sequence database *SDB* depicted in Figure 1. It contains mainly four sequences having accordingly the *sequences ids* (SIDs) 1, 2, 3 and 4. In this example, each solo letter represents an item. Items between curly brackets describes an itemset. For in-stance, the first

Cite this article as: Karishma B Hathi , Jatin R Ambasana. "Top K Sequential Pattern Mining Algorithm." *International Conference on Information Engineering, Management and Security (2015)*: 115-120. Print.

sequence ⟨{*a*, *b*},{*c*},{*f*},{*g*},{*e*}⟩ shows that items *a* and *b* happened at the same time, were followed successively by *c, f, g* and lastly *e*. A sequence $sa=< A1, A2, A3, \ldots, An>$ is called to be included in another sequence $sb=< B1, B2, B3, \ldots, Bm>$ if and only if there exists integers $1\leq I1< I2 <\ldots\ldots< In\leq Im$ such that $A1\subseteq Bi1$, $A2\subseteq Bi2$, $A3\subseteq Bi3,\ldots., An\subseteq Bin$. The support of a subsequence sa in a sequence database SDB is described as the number of sequences s ∈ S such that $sa \sqsubseteq$ s and is denoted by sup(sa). The problem of mining sequential patterns in a sequence database SDB is to locate all frequent sequential patterns, i.e. each subsequence sa such that sup(sa)≥minsup for a threshold minsup set by the user. For example, Figure 2 displays five of the 29 sequential patterns found in the database of table Figure 1 for minsup = 2. Many algorithms have been proposed for the problem sequential pattern mining such as PrefixSpan [3], SPAM [4], GSP and SPADE [6].

SID	Sequences
1	⟨{a, b},{c},{f, g},{g},{e}⟩
2	⟨{a, d},{c},{b},{a, b, e, f}⟩
3	⟨{a},{b},{f},{e}⟩
4	⟨{b},{f, g}⟩

Figure 1: Sequence Database

ID	Pattern	Support
P1	⟨{a},{f}⟩	3
p2	⟨{a},{c}{f}⟩	2
p3	⟨{b},{f,g}⟩	2
p4	⟨{g},{e}⟩	1
p5	⟨{b}⟩	4

Figure 2: Some Sequential Patterns

To address the problem of setting minsup, the problem of sequential pattern mining was reconsidered as the problem of top-k sequential pattern mining [7]. The current state-of-the-art algorithm for top-k sequential pattern mining is TSP [7]. There are two versions of TSP have been proposed for correspondingly mining (1) top-k sequential patterns and (2) top-k closed sequential patterns. Here we are addressing the first case. Extending algorithm to the second case will be considered in future work. The TSP algorithm is based on PrefixSpan [3]. TSP first generates frequent sequential patterns holding a single item. Then it recursively extends each pattern s by (1) it projecting the database by s, (2) it scanning the resulting projected database to identify items that appear more than minsup times after s, and (3) it append these items to s. The main benefit of this projection-based approach is that it only considers patterns appearing in the database unlike "generate-and-test" algorithms [2, 7]. However, the drawback of this approach is that projecting/scanning databases repeatedly is costly, and that cost becomes huge for dense databases where multiples projections have to be performed. Given this limitation, a chief research challenge is to define an algorithm that would be more efficient than TSP and that would perform well on dense datasets.

III. THE BASIC TKS ALGORITHM

TKS, an algorithm to find the top-k sequential patterns having the highest support, where k is set by the user. TKS employs the vertical database representation and basic candidate-generation procedure of SPAM [8]. Furthermore, it also includes various efficient strategies to find top-k sequential pattern efficient Fine-tuning the minsup parameter of sequential pattern mining algorithms to generate enough patterns is hard and time-consuming. To address this problem, the task of top-k sequential pattern mining has been defined, where k is the number of sequential patterns to be found, and is set by the user. So here an efficient algorithm for this problem named TKS (Top-K Sequential pattern mining) is present. TKS utilizes a vertical bitmap database representation, a new data structure named PMAP (Precedence Map) and various efficient strategies to prune the search space. The experimental study on real datasets shows that TKS outperforms TSP, the current state-of-the-art algorithm for top-k sequential pattern mining by more than an order of magnitude in execution time and memory.

TKS Algorithm [9]

It takes as parameters a sequence database SDB and k.

1) It first scans SDB once to construct V(SDB).

2) Let $Sinit$ be the list of items in V(SDB)

3) Then, for each item s∈$Sinit$, if s is frequent according to bv(s) it calls the procedure "SAVE".

4) R=RU{s, $Sinit$,items from $Sinit$ that are lexically larger than s}

5) WHILE ∃<r, S1,S2>∈ R AND sup(r)≥ minsup DO

6) Select the tuple <r,S1,S2> having the pattern r with the highest support in R

7) Then calls "SEARCH" find tuple.

8) Finally calls "REMOVE" and delete infrequent patterns from database.

IV. THE PROPOSED ALGORITHM

Limitation of basic TKS algorithm is number of database scan are higher so execution time grows higher due to this limitation. For improving the efficiency of TKS algorithm and overcome the drawbacks of TKS we propose an efficient approach for mining top k sequential patterns. We can improve the efficiency of TKS algorithm by using tree structure in TKS. By using this we can improve the efficiency of TKS algorithm in the terms of execution time.

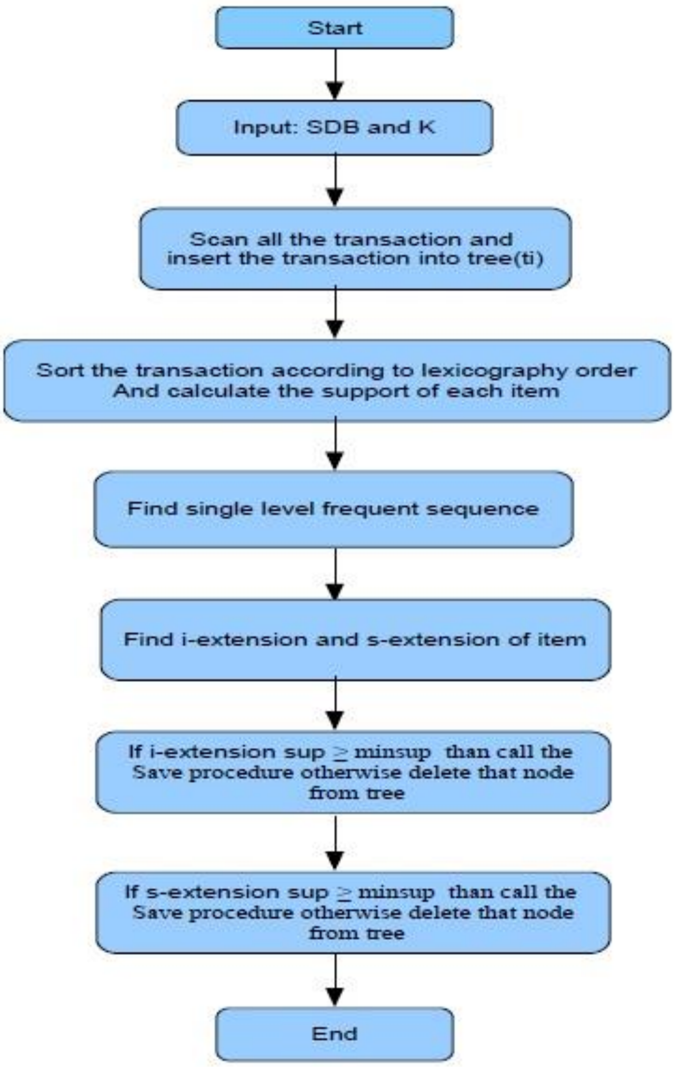


Figure 3: Propose Flow

Proposed Algorithm

Input: SDB, K

Output: Top K Sequential Patterns

Let $qtemp$ be the list of items in tree

For each q∈$qtemp$

Save(q,L,k,minsup)

Cite this article as: Karishma B Hathi , Jatin R Ambasana. "Top K Sequential Pattern Mining Algorithm." *International Conference on Information Engineering, Management and Security (2015)*: 115-120. Print.

☐ If i-extension sup(q)≥$minsup$
☐ *Save(q, all the items in* $qtemp$ that are lexically larger than q,L,k)
☐ And
☐ If s-extension sup(q)≥$minsup$
☐ *Save(q, all the items in* $qtemp$ that are lexically larger than q,L,k)
☐ End if
☐ Remove q∈$qtemp$ $when$ sup(q)<$minsup$

End for
Return L

V. TOOL STUDY

Java Technology:
JAVA is an object oriented platform independent and middle level language. It contains JVM (Java Virtual Machine) which is able to execute any program more efficiently. The feature of Platform Independence makes it different from the other Technologies available today.

Eclipse Tool:

Eclipse is an integrated development environment (IDE).It contains a base workspace and an extensible plug-in system for customizing the environment. Eclipse is written mostly in Java and thus can be used to develop applications. A vendor-neutral open-source workbench for multi-language development. An extensible platform for tool integration. Plug-in based framework to create, integrate and utilize software tools.

Sequential Pattern Mining Framework:

SPMF is an open-source data mining library written in Java, specialized in pattern mining. The source code of each algorithm can be integrated in other Java software. Moreover, SPMF can be used as a standalone program with a simple user interface or from the command line. The current version is **v0.96r16** and was released the **28th April 2015.**

VI. EXPERIMENTAL RESULTS

Various parameters are used for sequential pattern mining. Here this proposed approach conclude parameter execution time. The execution time is depend on the number of patterns are found and the number of passes that requires for database scan. We compared the performance of Proposed Algorithm with TKS Algorithm. All algorithms were implemented in Java. Experiments were carried on three real-life datasets having varied characteristics and representing three different types of data (web click stream, text from books, sign language utterances). Those datasets are accordingly FIFA, Bible and Sign. The comparison between TKS and Proposed algorithm is shown in following figures.

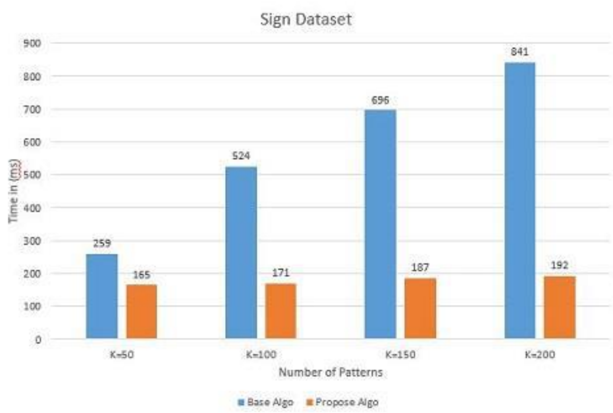


Figure 4: Sign Dataset

Cite this article as: Karishma B Hathi , Jatin R Ambasana. "Top K Sequential Pattern Mining Algorithm." *International Conference on Information Engineering, Management and Security (2015)*: 115-120. Print.

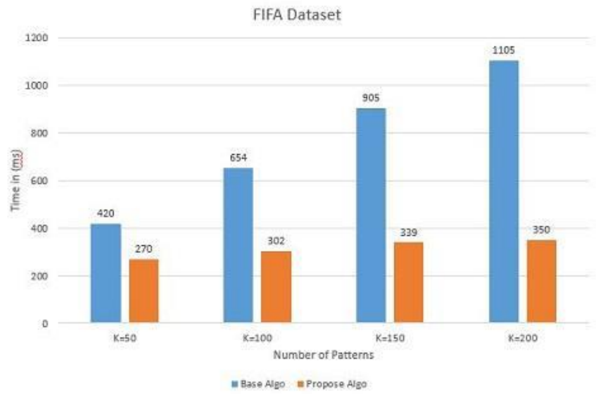


Figure 5: FIFA Dataset

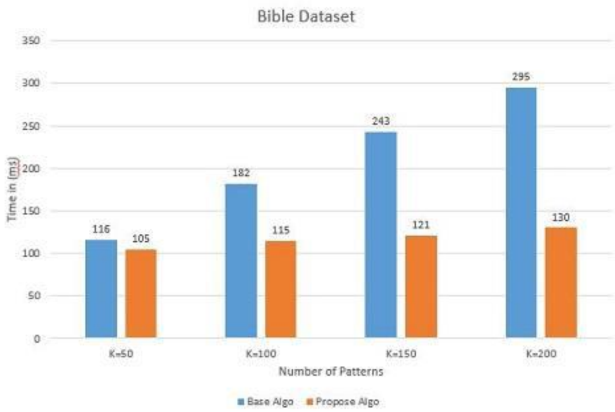


Figure 6: Bible Dataset

VII. CONCLUSION AND FUTURE WORK

The Proposed System has improves the performance of TKS algorithm by using tree Structure. From the result analysis it is clearly seen that the execution time of proposed algorithm reduced. There is less number of database scan requiring so the efficiency of the TKS algorithm improves. Thus, we conclude that the proposed system has better performance than TKS algorithm. Extending TKS algorithm for finding Top K Closed Sequential Pattern can be considered as future work.

ACKNOWLEDGEMENT

We are deeply indebted & would like to express gratitude to our thesis guide Prof. Jatin Ambasana , B. H. Gardi College of Engineering & Technology for his great efforts and instructive comments in the dissertation work.
We would also like to extend our gratitude to Prof.Hemal Rajyaguru, Head of the Computer Science & Engineering Department, B. H. Gardi College of Engineering & Technology for his continuous encouragement and motivation.
We would also like to extend our gratitude to Prof. Vaseem Ghada, PG Coordinator, B. H. Gardi College of Engineering & Technology for his continuous support and cooperation.
We should express our thanks to our dear friends & our classmates for their help in this research; for their company during the research, for their help in developing the simulation environment.
We would like to express our special thanks to our family for their endless love and support throughout our life. Without them, life would not be that easy and beautiful.

REFERENCES

[1] J. Han and M. Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufman publishers, 2001.
[2] R. Agrawal and R. Srikant, "Mining Sequential Patterns", In Proceedings of the 11th International Conference on Data Engineering, pp. 3-14, Taipei, Taiwan, 1995.
[3] Pei, J., Han, J., Mortazavi-Asl, B., Pinto, H., Chen, Q., Dayal, U., Hsu, M.: Mining Se-quential Patterns by Pattern-Growth: The PrefixSpan Approach, IEEE Trans. Knowledge and Data Engineering, vol. 16, no. 10, pp. 1-17 (2001)
[4] Ayres, J., Flannick, J., Gehrke, J. and Yiu, T.: Sequential PAttern mining using a bitmap representation, Proc. 8th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD 2002), July 23-26, 2002, Edmonton, Alberta, pp. 429-435 (2002)

Cite this article as: Karishma B Hathi , Jatin R Ambasana. "Top K Sequential Pattern Mining Algorithm." *International Conference on Information Engineering, Management and Security (2015)*: 115-120. Print.

[5] Mabroukeh, N. R. and Ezeife, C. I.: A taxonomy of sequential pattern mining algorithms, ACM Computing Surveys, vol. 43, no. 1, pp. 1-41 (2010)
[6] Cormen, T. H., Leiserson, C. E., Rivest, R. and Stein, C.: Introduction to Algorithms, 3rd ed., Cambridge:MIT Press (2009)
[7] Tzvetkov, P. Yan, X. and Han, J.: TSP: Mining Top-k Closed Sequential Patterns, Knowledge and Information Systems, vol. 7, no. 4, pp. 438-457 (2005)
[8] Jian Pei, “Mining Sequential Patterns by Pattern-Growth:The PrefixSpan Approach”, IEEE Transactions on Knowledge and Data Engineering, VOL. 16, NO. 10, OCTOBER 2004.
[9] Philippe Fournier-Viger,“TKS: Efficient Mining of Top-K Sequential Patterns”, Springer Advanced Data Mining and Application, vol. 8346, pp. 109-120, 2013.
[10] Vishal S. Motegaonkar, Prof. Madhav V. Vaidya,”A survey on sequential pattern mining algorithms”,International Journal of Computer Science and Information Technologies (IJCSIT),Vol.5 (2),2014.
[11] C.-C. Yu and Y.-L. Chen, “Mining Sequential Patterns from Multi-Dimensional Sequence Data”, IEEE Trans. Knowledge and Data Eng., Vol. 17, No. 1, pp. 136-140, Jan. 2005.
[12] Karishma B Hathi, Jalpa A Varsur, Sonali P Desai, Sagar R Manvar,” A Performance Analysis of Sequential Pattern Mining Algorithms’, ournal of Emerging Technologies and Innovative Research (JETIR), February 2015, Volume 2, Issue 2.
[13] [Jian Pei, Jiawei Han and Wei Wang, "Constraint-based sequential pattern mining: the pattern-growth methods", Journal of Intelligent Information Systems, Vol:28, No: 2 ,pp:133-160, 2007.
[14] Jian Pei, Jiawei Han and Wei Wang, "Constraint-based sequential pattern mining: the pattern-growth methods", Journal of Intelligent Information Systems, Vol:28, No: 2 ,pp:133-160, 2007.

Cite this article as: Karishma B Hathi , Jatin R Ambasana. “Top K Sequential Pattern Mining Algorithm.” *International Conference on Information Engineering, Management and Security (2015)*: 115-120. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS020

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.020

MINING RARE ITEMSET BASED ON FP-GROWTH ALGORITHM

A. Jalpa A Varsur[1], Nikul G Virpariya[2]
[1]Student, M.E.(CSE), Gardi Vidyapith, Gujarat, India
[2]Assistant Professor, CSE Department, Gardi Vidyapith, Gujarat, India

ABSTRACT: *Frequent weighted itemset represent correlation frequently holding in data in which items may weight differently.However,in some contexts,e.g., when the need is to minimize a certain cost function,discovering rare data correlations is more interesting than mining frequent ones. Now a days, some uninteresting items are consider as important.which are appear very less in the database.It is known as Rare or infrequent item.Finding infrequent itemset is useful in many recent field like medical,web,cloud,market basket analysis etc,. for i.e.,in market basket analysis some set of items such as milk and bread customer buy frequently compared to milk and bread gold chain and ring are infrequent items.This paper takes the issue of discovering rare itemset in transaction dataset.*

KEYWORDS: Data Mining,Frequent item,Infrequent item,Threshold,Support,Item weight.

I. INTRODUCTION

Itemset is a set of items.Frequent items are appear very frequently in database,with high support and high confidence. Rare items are reverse from of frequent ones,it has low support and high confidence. [1]That has a vast range of real life application like,
In **Medical** – if we identify the solution of rare disease then we can prevent the person to get affected by it.Here,we don't need this rare disease to become frequent. In **Marketing** strategy knowing the rare items can help us to make them the frequent ones.So,the businessmen gains profit. In **Market basket analysis** for finding which items tend to be purchased together,milk and bread occurs frequently and can be considered as regular case,some items like bad and pillow are infrequently associated itemsets.[1] Recently,the importance is being given for the discovery of infrequent or exceptional patterns[1].threshold plays key role for finding of frequent and infrequent itemsets.The things which are done together for eg.Buying groceries known as association,occurs between items.Association is an implication of the form X→Y.Infrequent itemsets are produced from very vast or enormous datasets.Extraction of frequent itemset is a necessary step in many association techniques[2]. Association rule mining extracts interesting correlation between transactions.In many application some items are appear very frequently in the data,while others rarely appear.if minsup is set too high,those rules that involve both frequent and infrequent items.To find the rule that contain both frequent and rare minsup is set to be very low.This may cause combinatorial explosion for those frequent items will be associated with one another in all possible ways.This problem is called the rare item problem[3]. infrequent itemset do not comprise any infrequent subset. it appears only when threshold is set to very low.The mining algorithm solves the problem of discovering the infrequent itemsets in the given database[2].Infrequent Itemset Mining finds uninteresting items among the huge database.

Cite this article as: A. Jalpa A Varsur, Nikul G Virpariya. "MINING RARE ITEMSET BASED ON FP-GROWTH ALGORITHM." *International Conference on Information Engineering, Management and Security (2015)*: 121-128. Print.

The Pattern Growth algorithm comes in the early 2000s, for the answer to the problem of generates and test.The main idea is for to avoid the candidate generation step altogether, and to concentrate the search on a specified portion of the initial database.

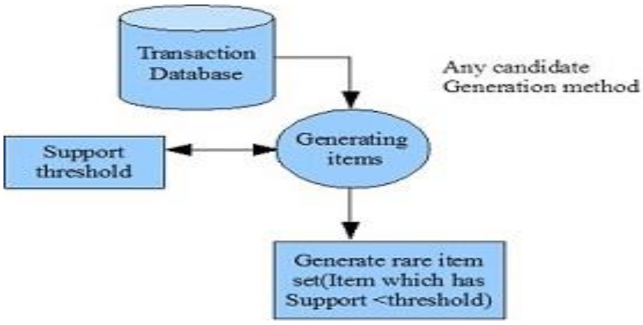


Figure 1: Rare itemset mining architecture [9]

II. RELATED WORK

For example, the rule {Bread, Butter} =>{Milk} found in the sales data of a shop would indicate that if a customer buys bread and butter together, he or she is likely to also buy milk. Such information can be used in decision making about marketing policies such as, e.g., product offers, product sales and discount schemes. In addition to the above mentioned example association rules are used today in many application areas including Web usage mining, Intrusion detection, Continuous production, and Bioinformatics [3]. As opposed to sequence mining, association rule learning typically does not consider the order of items either within a transaction or across transactions.

The problem of association rule mining [3] is defined as: Let I= {i1, i2,…, in} be a set of n binary attributes called *items*. Let D={t1,t2,…,tm} be a set of transactions called the *database*. Each transaction in database D has a unique transaction identity ID and contains a subset of the items in I [3]. A *rule* is defined as an implication of the form X=>Y where X,Y is subset of I and X intersection Y = Null Set. The sets of items (for short *itemsets*) X and Y are called *antecedent* (if) and *consequent* (then) of the rule respectively.[6]

III. PROBLEM DEFINITION

To understand the background of itemset mining, we present different techniques and algorithm in the following subsections.

Goal: Mining infrequent itemset from transaction datasets.

I={i1,i2,..,im}be a set of data items.A transactional dataset T={t1,t2,…,tn}is a set of transactions,where each transaction tq (q € [1,n])is a set of items in I and is characterized by a transaction ID (tid).An itemset I is a set of data items[6].Specifically we denote as k-itemset a set of k-items in I.The support of an itemset is the number of transactions containing I in T.

An itemset I is infrequent if its support is less than or equal to a predefined maximum support threshold ξ.Otherwise,it is said to be frequent[1].

Weighted transactional data set

Let I ={i1,i2,…,im} be a set of items.A weighted transactional data set T is a set of weighted transactions ,where each weighted transaction tq is a set of weighted items <ik,wk>

Weights could be either positive,null or negative numbers.itemsets mined from weighted transactional data sets are called weighted itemsets.

Their expression is similar to the one used for traditional itemsets,i.e., a weighted itemset is a subset of the data items occurring in a weighted transactional data set.The problem of mining itemsets by considering weights associated with each item is known as the weighted itemset mining problem[4].

This approach is focuses on considering item weights in the discovery of infrequent itemsets.To this aim,the problem of evaluating

Cite this article as: A. Jalpa A Varsur, Nikul G Virpariya. "MINING RARE ITEMSET BASED ON FP-GROWTH ALGORITHM." *International Conference on Information Engineering, Management and Security (2015)*: 121-128. Print.

itemset significance in a given weighted transactional data set is addressed by means of a two-step process.

Firstly,the weight of an itemset I associated with a weighted transaction tq € T is defined as an aggregation of its item weights in tq.Secondly,the significance of I with respect to the whole data set T is estimated by combining the itemset significance weights associated with each transaction.

TID	CPU Usage Readings
1	<a,0> <b,100> <c,57> <d,71>
2	<a,0> <b,43> <c,29> <d,71>
3	<a,43> <b,0> <c,43> <d,43>
4	<a,100> <b,0> <c,43> <d,100>
5	<a,86> <b,71> <c,0> <d,71>
6	<a,57> <b,71> <c,0> <d,71>

Figure 3: TABLE Weighted Transactional Data Set

The significance of a weighted transaction, i.e., a set of weighted items, is commonly evaluated in terms of the corresponding item weights. For instance, when evaluating the support of {a,b} in the example data set reported in Table 1, the occurrence of b in tid 1, which represents a highly utilized CPU, should be treated differently from the one of a, which represents an idle CPU at the same instant.

Task (A) entails discovering IWIs and minimal IWIs (MIWIs) which include the item with the least local interest within each transaction. Table 2 reports the IWIs mined from Table 1 by enforcing a maximum IWI-support-min threshold equal to 180 and their corresponding IWI-support-min values. For instance, {a,b} covers the transactions with tids 1, 2, 3, and 4 with a minimal weight 0 (associated with a in tids 1 and 2 and b in tids 3 and 4), while it covers the transactions with tids 5 and 6 with minimal weights 71 and

57, respectively.Hence, its IWI-support-min value is 128.

IWI	IWI-Support-min	IWI	IWI-Support-min
{c}	172 (Minimal)	{a,b,c}	0 (Not Minimal)
{a,b}	128 (Minimal)	{a,b,d}	128 (Not Minimal)
{a,c}	86 (Not Minimal)	{a,c,d}	86 (Not Minimal)
{b,c}	86 (Not Minimal)	{b,c,d}	86 (Not Minimal)
{c,d}	172 (Not Minimal)	{a,b,c,d}	0 (Not Minimal)

TABLE IWIs Extracted from the Data Set from above Table

Maximum IWI-support-max threshold = 390

IWI	IWI-Support-max	IWI	IWI-Support-max
{a}	286 (Minimal)	{a,c}	0 (Not Minimal)
{b}	285 (Minimal)	{b,c}	128 (Not Minimal)
{c}	172 (Not Minimal)		

TABLE IWIs Extracted from the Data Set from above Table

Base Algorithm
Input: T,a weighted transaction dataset ,

Cite this article as: A. Jalpa A Varsur, Nikul G Virpariya. "MINING RARE ITEMSET BASED ON FP-GROWTH ALGORITHM." *International Conference on Information Engineering, Management and Security (2015)*: 121-128. Print.

maximum IWI support threshold
Output: set of IWIs
1. initialization of items.
2. count IWI-support
3. tree -a new empty tree
4. for all weighted transaction tq in T do,
5. TEq -equivalence transaction
6. for all transaction insert tej in tree
7. end for
8. end for
9. IWI Mining
10. return set of IWIs

IV. PROPOSED ALGORITHM

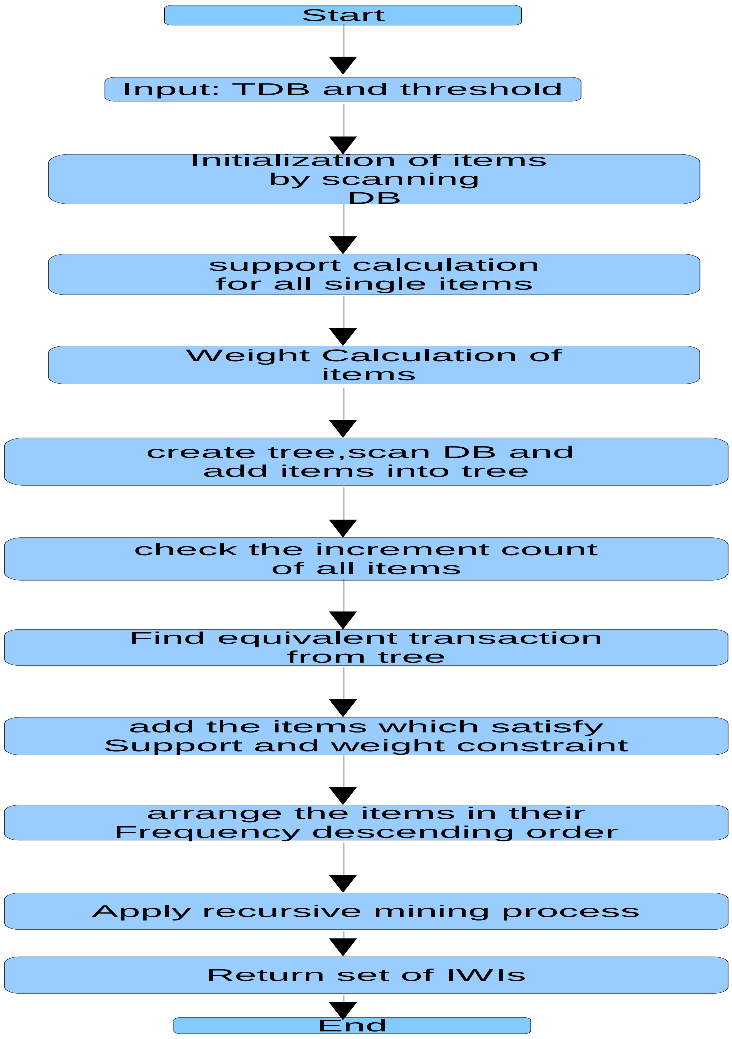


Figure 4: Proposed Algorithm Design

Proposed Algorithm Steps:

Input : - T (transaction database TDB) ,
ξ (maximum IWI-support threshold)
Output :- Ɉ (set of IWIs)

Step 1: Ɉ = ∅ /* initialization of items by scanning DB */
Step 2: count the support for all single items

$$\text{Support} = \frac{(X \cup Y).\text{count}}{n}$$

Cite this article as: A. Jalpa A Varsur, Nikul G Virpariya. "MINING RARE ITEMSET BASED ON FP-GROWTH ALGORITHM." *International Conference on Information Engineering, Management and Security (2015)*: 121-128. Print.

Confidence = $\frac{(X \cup Y).count}{c.count}$

Step 3: weight calculation of all items.

- Generation of MIS equation will be used for weight calculation,we take MIW instead of MIS
- $MIW(i) \equiv \begin{cases} M(i) & \text{where } M(i) > LS \\ LS & \text{LS otherwise} \end{cases}$
- $M(i) \equiv \beta f(i)$
- here, f (i) is the actual frequency of item i in data or the support expressed in percentage of the data set size
- LS = user specify lowest minimum weight
- β = a parameter to control MIW value for items
- If $\beta = 0$ we have only one minimum weight.
- If $\beta = 1$ and $f(i) \geq LS$
- $f(i)$ is the MIW value for i.

Step 4: create initial fp-tree

- add items into tree
- for all transaction ti ∈ T
- for all tej ∈ ti
- Insert tej in tree

Step 5: If tej.sup ≤ tsup && tej.weight ≥ tweight

- If item does not satisfy the Weight Constrain and support constrain then remove it from transaction

Step 6: Order items in their frequency. descending order.
Step 7: Ꞙ ← Mining process
Step 8: return set of infrequent items
Step 9: end

Advantages

It will Reduce time.
It Require less memory.
It removes the frequent items from tree.

Limitation

Number of database scan are higher
It performs save procedure each time when particular item is found.
So execution time grows higher due to this limitation.

V. STUDY OF TOOL

Java Technology
JAVA is an object oriented, platform independent and middle level language.It contains JVM (Java Virtual Machine) which is able to execute any programmore effciently.The feature of Platform Independence makes it different from the other Technologiesavailable today.

Eclipse Tool
Eclipse is an integrated development environment (IDE).It contains a base workspace and an extensible plug-in system for customizingthe environment.Eclipse is written mostly in Java and thus can be used to develop applications.
Eclipse started as a proprietary IBM product (IBM Visual age for Smalltalk/Java)
SPMF
SPMF is an **open-source data mining mining library** written in **Java**, specialized in **pattern mining.**It is distributed under the **GPL v3 license.**It offers implementations of **78 data mining algorithms** for:

- **sequential pattern mining,**
- **association rule mining,**
- **frequent itemset mining,**

- **high-utility pattern mining,**
- **sequential rule mining,**
- **clustering.**

The source code of each algorithm can be integrated in other Java software.Moreover, SPMF can be used as a standalone program with a simple user interface or from the command line.The current version is **v0.96r16** and was released the **28th April 2015.**

V. RESULT ANALYSIS

Aggregate function

Aggregate function is a function where the values of multiple rows are grouped together as input or certain criteria to from a single value or more significant meaning or measurement such as a set,a bag or a list.for eg. Function like average( ),count ( ),maximum ( ). It will returns a single value.

Reducing the execution time

First scan of data will remove the frequent items and reduce the no. of scans .by tree pruning strategy it will find the prunable items and reduce the time

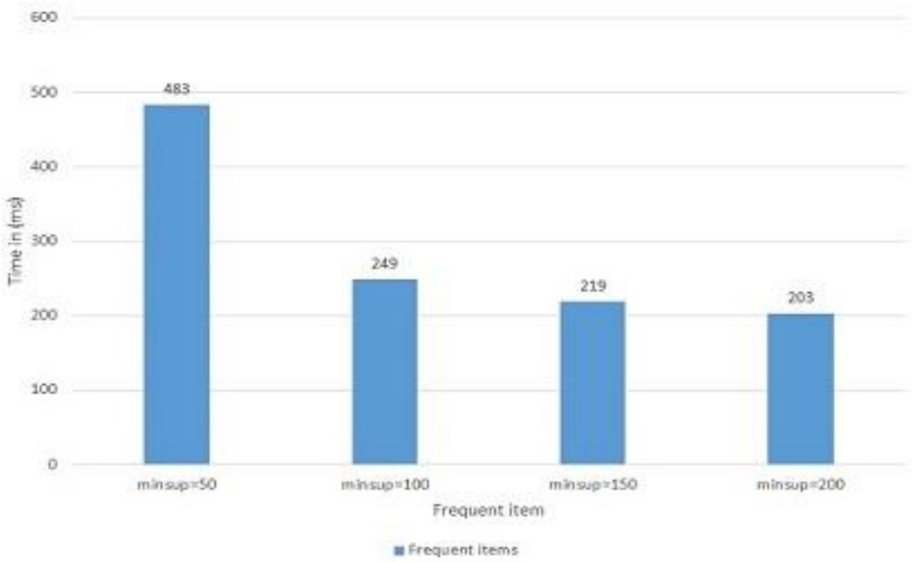


Figure 5: Performance of different threshold values

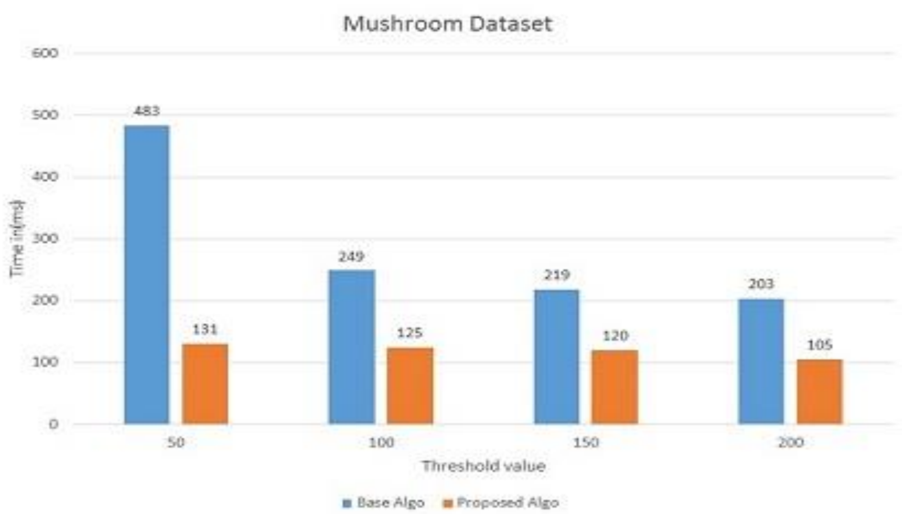


Figure 6: Execution time comparison over mushroom dataset

Cite this article as: A. Jalpa A Varsur, Nikul G Virpariya. “MINING RARE ITEMSET BASED ON FP-GROWTH ALGORITHM.” *International Conference on Information Engineering, Management and Security (2015)*: 121-128. Print.

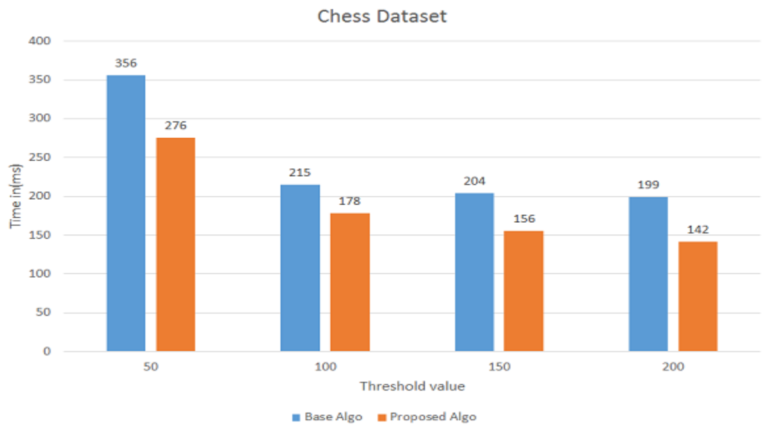


Figure 7: Execution time comparison over chess dataset

The algorithm first constructs the tree out of the original data set and then grows the frequent patterns.For a faster execution,the data should be preprocessed before applying the algorithm.

VI. CONCLUSION AND FUTURE WORK

The Propose System has improves the performance of IWI mining algorithm by using FP-Growth Structure.
It reduce the execution time.
at mining time tree will remove the frequent items and we get only rare items.
Thus,we conclude that the proposed system has better performance and it will require less memory.

Future Work:

As future,we plan for discovering rare itemset ,weight calculation of items will be done by user.IWI algorithm can also be implemented in advanced decision making system and business intelligence.

ACKNOWLEDGEMENT

We are deeply indebted & would like to express gratitude to our thesis guide Prof. Nikul Virpariya , B. H. Gardi College of Engineering & Technology for his great efforts and instructive comments in the dissertation work.

We would also like to extend our gratitude to Prof.Hemal Rajyaguru, Head of the Computer Science & Engineering Department, B. H. Gardi College of Engineering & Technology for his continuous encouragement and motivation.

We would also like to extend our gratitude to Prof. Vaseem Ghada, PG Coordinator, B. H. Gardi College of Engineering & Technology for his continuous support and cooperation.

We should express our thanks to our dear friends & our classmates for their help in this research; for their company during the research, for their help in developing the simulation environment.

We would like to express our special thanks to our family for their endless love and support throughout our life. Without them, life would not be that easy and beautiful.

REFERENCES

[1] K.S.Sadhasivam,Tamilarasi,"Mining Rare ItemSet with Automated Support Thresholds",*Journal of Computer Science*,pp.394-399,2011.
[2] Lugi Troiano,Cosimo Birtolo,"A Fast Algorithm for Mining Rare ItemSets",IEEE Ninth International Conference on Intelligent Systems Design and Applications,2009.
[3] Mehdi Adda,Lei Wu,"Rare ItemSet Mining",IEEE Sixth International Conference on Machine Learning and Applications,2007.
[4] Petko Valtchev,Amedeo Napoli,"Towards Rare ItemSet Mining",19th IEEE International Conference on Tools with Artificial Intelligence,2007.
[5] K.Sun,Fengshan Bai,"Mining weighted Association Rules without Preassigned Rules",*IEEE Transactions On Knowledge and Data Engineering*,vol.20,No.4,April 2008.

Cite this article as: A. Jalpa A Varsur, Nikul G Virpariya. "MINING RARE ITEMSET BASED ON FP-GROWTH ALGORITHM." *International Conference on Information Engineering, Management and Security (2015)*: 121-128. Print.

[6] Luca Cagliero,Paolo Garza,"Infrequent Weighted ItemSet Mining Using Frequent Pattern Growth",*IEEE Transactions On Knowledge and Data Engineering*,vol.26,No.4,April 2014.
[7] Gou Masuda,Norihiro Sakamoto,"A Framework for Dynamic evidence based medicine using data mining",*CBMS'02:Proceedings of the 15th IEEE Symposium on Computer-Based Medical Systems*,2002.
[8] H.Yun,D.Ha,"Mining association rules on significant rare data using relative support",*The Journal of Systems and Software*,pp.181-191,2003.
[9] Nidhi Sethi,Pradeep Sharma,"Efficient Algorithm for Mining Rare Itemsets over Time Variant Transactional Database",*International Journal of Computer Science and Information Technologies*,vol.5,2014.
[10] J.Jenifa,Dr.V.Sampath Kumar,"Study on predicting various Mining Techniques Using weighted Itemsets",*IOSR*,vol.9,pp.30-39,Mar-Apr 2014.

Cite this article as: A. Jalpa A Varsur, Nikul G Virpariya. "MINING RARE ITEMSET BASED ON FP-GROWTH ALGORITHM." *International Conference on Information Engineering, Management and Security (2015)*: 121-128. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS021		**eAID**	ICIEMS.2015.021

BIGDATA ANALYTICS WITH SPARK

SUBHASH KUMAR
IT Department, St Xavier's College, Mahapalika Marg, New Marine Lines,
Mumbai, PIN-400001, Maharashtra, India

Abstract: *Current generation is witnessing data explosion most of it is unstructured and is called Big Data. This data has characteristics of high volume, velocity, variety and veracity. HDFS, GFS, Ceph, Lustre, PVFS etc are used as file system for storing Big Data. MapReduce processes program in parallel across clusters and generates output. Spark framework improves performance by 10x when datasets are stored in hard disk and performance improves by 100x when data is stored in memory. This paper proposes optimization of Big Data processing using Spark framework.*

Keywords: Volume, velocity, veracity, Big Data, HDFS, Spark.

I. INTRODUCTION

Huge amount of data is being generated every second. It puts a new challenge for managing this data. Social media like Facebook generates about 600 TB of data every day, whereas Twitter generates about 120 TB each day and Google generates about 20 PB of data each day. It is evident that data is collected in an exponential rate and we have already reached Terabyte and PetaByte stage (see Figure 1). 1PB=1024TB. 1EB=1024PB.

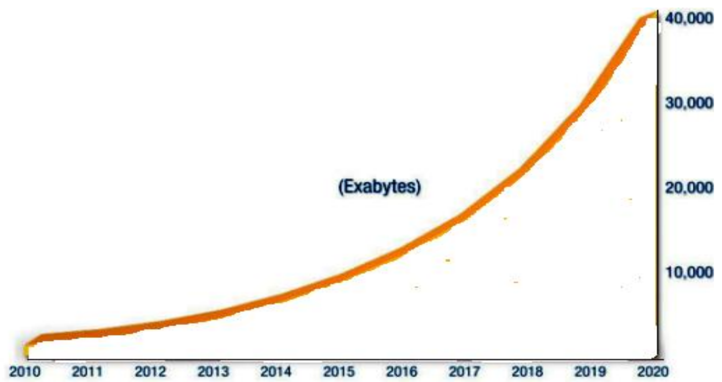


Figure 1: Digital Universe expansion

As per IDC [6], digital universe in 2010 was 1227 ExaBytes and by end of 2020 data would reach to 40ZB.

Cite this article as: SUBHASH KUMAR. "BIGDATA ANALYTICS WITH SPARK." *International Conference on Information Engineering, Management and Security (2015)*: 129-133. Print.

This Big Data needs to be stored in multiple machines in commodity hardware as clusters since single machine cannot store it. This is managed with Hadoop Framework. HDFS (Hadoop Distributed File System), GFS (Google File System), PVFS, Ceph etc is used to store Big Data. Default data size in HDFS is 64MB. These chunks (64MB) are stored in commodity hardware. In hadoop MapReduce, the tuples generated in Map and Reduce task are stored in disk (see Figure 2). This takes time. This is improved by using apache spark framework.

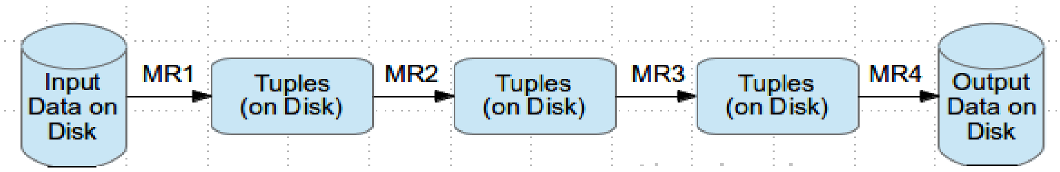


Figure 2: MapReduce iteration in Hadoop

Apache Spark is an open-source cluster computing framework developed in 2009 at AMPLab at University of California. Spark can read from any data source (relational, NoSQL, file systems, etc) and offers unified API for batch analytics, SQL queries, machine learning and graph processing, real-time analysis.

Spark is not only designed to run many more workloads, but it can do so much faster than older systems. Spark is 10 times faster than Hadoop MapReduce when data is read from disk and 100 times faster when data is read from memory. Spark is highly scalable. The largest Spark cluster has about 8,000 nodes.

Spark improves efficiency through in memory computation, general computation graph. It has rich API in java, python, scala and interactive shell where programmers write less line of codes.

II. LITERATURE REVIEW

Yanfeng Zhang et al. [4] paper discusses PrIter, which is the prioritized execution of iterative computations. PrIter stores intermediate data in memory for fast convergence or stores intermediate data in files for scaling to larger data sets. PrIter was evaluated on a local cluster of machines as well as on Amazon EC2 Cloud. The results show that PrIter achieves up to 50 × speedup over Hadoop for iterative algorithms type problems. In addition, PrIter is shown better performance for iterative computations than other distributed frameworks such as Spark and Piccolo

Yanfeng Zhang et al. [2] paper discusses iMapReduce is a distributive framework that significantly improves the performance of iterative computations by (1) reducing the creation new MapReduce jobs again and again, (2) shuffling of static data gets eliminated, and (3) asynchronous execution of map tasks is allowed, iMapReduce prototype shows that it can achieve up to 5 times speedup in implementing iterative algorithms.

Xu, X et al. [3] pointed that if TaskTracker could adjust to change of load as per its computing ability, results can be obtained faster.

Weizhong Zhao et al. [4] said that Parallel K-Means clustering based on MapReduce can process datasets efficiently using commodity hardware.

Matei Zaharia et al. [5] pointed out that the resilient distributed dataset (RDD), which represents a read-only collection of objects can be partitioned across multiple set of machines in cluster and can be rebuilt even if a partition is lost. If a partition of an RDD is lost, the RDD has enough information and uses other RDDs to rebuild just that partition by using lineage information. Lineage is the sequence of transformations used to build the current RDD. .
Matei Zaharia et al. [6] paper showed following results of spark: spark outperforms Hadoop by up to 20x in iterative machine learning and graph applications. The speedup comes from avoiding I/O and deserialization costs by storing data in memory as Java objects. Applications written in spark perform and scale well. In particular, spark speeds up an analytics report that was running on Hadoop by 40x. When nodes fail, Spark shows recovery strategy by rebuilding only the lost RDD partitions. Spark can to query a 1 TB dataset interactively with latencies of 5–7 seconds.

As per apache spark [7], spark runs much faster than hadoop which is evident from the figure below (see Figure 3) for logic regression.

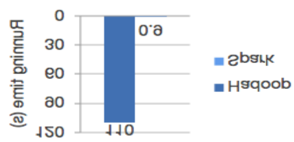


Figure 3: Logic Regression in Hadoop and Spark

Cite this article as: SUBHASH KUMAR. “BIGDATA ANALYTICS WITH SPARK.” *International Conference on Information Engineering, Management and Security (2015)*: 129-133. Print.

III. SPARK FRAMEWORK

A. RDD (Resilent Distributed Dataset)

It is an abstraction layer which is read only and represents collection of objects that can be stored in memory or disk in cluster. It can be rebuilt on failure. It has parallel functional transformation. RDDs support operations known as transformations, which create new dataset from an existing one, and actions, which after running a computation on the dataset return a value to the driver program. Some of the actions in spark are filter, count, union, join, sort, groupBy, groupByKey, pipe, cross, mapWith etc.

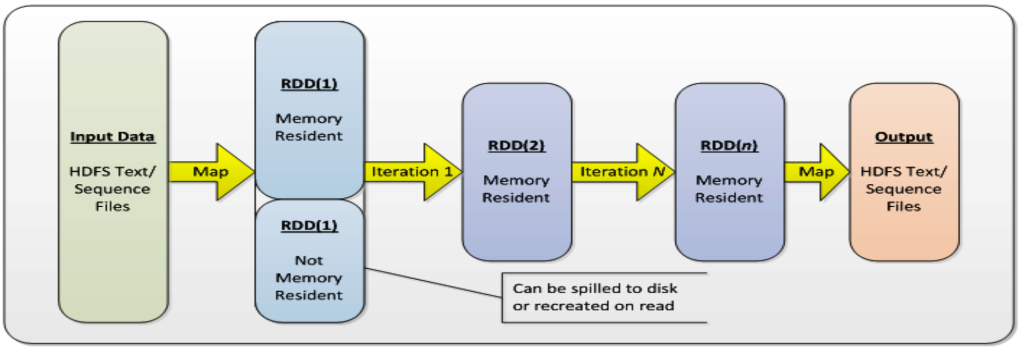


Figure 4: Iteration in RDD using SPARK

From Figure 4, the first map operation into RDD (1) is shown where, not all data could fit in the memory space so some data is passed to the hard disk. Data is first searched in the memory for the reading and also writing occurs in memory. This method makes system to be 100X faster than other methods that rely purely on disk storage.

Spark follows lazy loading that is it doesn't perform transformation on RDD immediately. Instead, it piles this transformation and forms batch which is then processed.

B. SPARK STACK

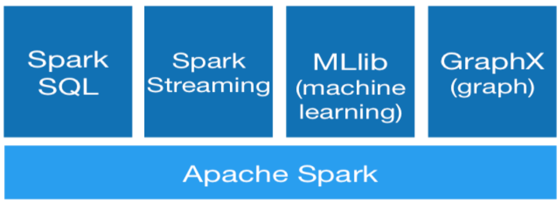


Figure 5: Spark Stack

Spark Stack consists of four major components Spark SQL, Spark Streaming, MLib, GraphX (see Figure 5)

C. SPARK SQL

The two useful components of Spark SQL are DataFrame and SQLContext. DataFrame provides an abstraction which can act as distributed SQL query engine. A DataFrame is a distributed collection of data which is organized into named columns. DataFrames can be converted to RDDs and vice versa. DataFrames can be created from different data sources such as: Hive tables, Existing RDDs, JSON datasets, structured data files, External databases. Spark SQL lets you query structured data as a distributed dataset (RDD) in Spark, with integrated APIs in Python, Scala and Java. This tight integration makes it easy to run SQL queries alongside complex analytic algorithms.

D. SPARK STREAMING

Spark Streaming allows one to process large data streams in real time. This helps to find fraud detection. Spark Streaming allows live streaming as well as post processing in batch. There is no other framework which can do both.
The live stream is divided into small batch of x second which is then passed to spark framework and it treats this batch as RDD and processes it in batch (see Figure 6).

Cite this article as: SUBHASH KUMAR. "BIGDATA ANALYTICS WITH SPARK." *International Conference on Information Engineering, Management and Security (2015)*: 129-133. Print.

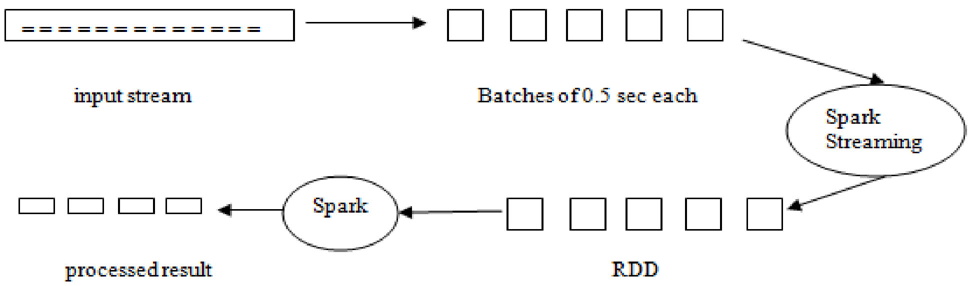


Figure 6: Illustration of Spark Streaming

E. MLLIB

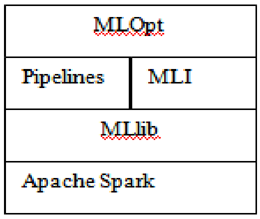


Figure 7: Building blocks of MLlib

MLOpt is declarative layer which automates hyperparameter tuning. Pipelines and MLI are API for simplifying development of machine learning such as distributed table, distributed matrices. MLlib is machine learning core library (see Figure 7). It consists of Gradient descent algorithm for optimization, K-Means algorithm for clustering, Logistic Regression for prediction, feature transformation etc.

F. GRAPHX

GraphX is the new Spark API for graph-parallel computation. GraphX extends the Spark RDD with graph concept where properties are attached to each vertex and edge. GraphX exposes fundamental operators such as subgraph, joinVertices etc. GraphX has collection of graph algorithms which simplifes graph analytics. With one can view the same data as both graphs and collections, join and transform graphs with RDD efficiently.

IV. SPARK RUNTIME

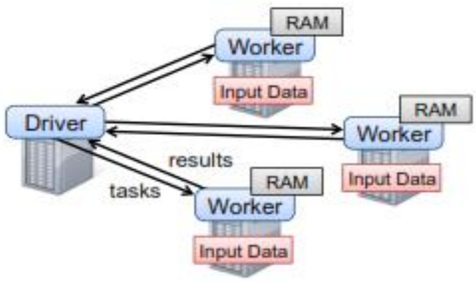


Figure 8: Spark Runtime

See Figure 8, where Driver program launches multiple worker threads that read data blocks from a distributed file system and persists computed RDD partitions in memory. Developers write a driver program which can connect to a cluster of workers, as shown in Figure 2. The driver defines one or more RDDs and invokes actions on them. The workers can store RDD partitions in RAM. A driver performs two types of operations on a dataset: action and transformation. action performs computation on dataset and returns value to the driver; transformation creates new dataset from an existing dataset.

V. PROGRAMMING ILLUSTRATION

Here illustration is shown in scala language which is functional programming language.
This program collects all errors having DB2 written in ERROR Log.

//RDD is created using hdfs
val txt = spark.textFile("hdfs://scrapper/user/alltweets.txt"")

Cite this article as: SUBHASH KUMAR. "BIGDATA ANALYTICS WITH SPARK." *International Conference on Information Engineering, Management and Security (2015)*: 129-133. Print.

```
//New RDD created by transformation which searches for ERROR
val errors = txt.filter(line => line.contains("ERROR"))
// Count all the errors. Action is performed
errors.count()
// Count errors mentioning DB2
errors.filter(line => line.contains("DB2")).count()
// Fetch the DB2errors as an array of strings
errors.filter(line => line.contains("DB2")).collect()
```

Another example for storing data in RAM using cache() method . It is illustrated below

step1
Go to the bin dir of spark installation
```
$ /home/spark-1.2.1//bin
```
step2
$ run the spark-shell which takes to scala prompt
```
./spark-shell
```

step3
Create RDD of TEST.txt
```
scala>val tf=sc.textFile("TEST.txt")
```

step4
Use transformation on RDD and creat new RDD
```
scala>val ramtxt=tf.filter(line=>line.contains("Data")
```

step5
Store this RDD in RAM for faster access
```
scala>ramtxt.cache()
```

CONCLUSION

Spark framework supports big data processing. This framework can give faster result than existing hadoop system. It has capability of doing in-memory computation using languages scala, java, python. One can work with cluster writing less lines of code in scala as it is functional programming language. Spark has rich set of libraries for data streaming, machine learning and spark sql. Spark framework improves performance by 10x when datasets are stored in hard disk and performance improves by 100x when data is on memory.

ACKNOWLEDGMENT

I would like to thank Knowledge Centre of St Xavier's College for providing the infrastructure for setting cluster. Also I would like to thank our Principal (Dr.) Agnelo Menezes and Dr Siby Abraham for supporting research activities.

REFERENCES

[1] Vernon Turner, John F. Gantz, David Reinsel and Stephen Minton (2014). White paper The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things. URL:http://idcdocserv.com/1678

[2] Yanfeng Zhang; Qinxin Gao; Lixin Gao; Cuirong Wang, "iMapReduce: A Distributed Computing Framework for Iterative Computation," Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW), 2011 IEEE International Symposium on, vol., no., pp.1112,1121,16-20May2011 doi: 10.1109/IPDPS.2011.260

[3] Xu, X.; Cao, L.; Wang, X. (2014), Adaptive Task Scheduling Strategy Based on Dynamic Workload Adjustment for Heterogeneous Hadoop Clusters Systems Journal, IEEE , vol.PP, no.99, pp.1,12. doi: 10.1109/JSYST.2014.2323112. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6832443&isnumber=4357939

[4] Yanfeng Zhang, Qixin Gao, Lixin Gao, Cuirong Wang (Sept.2013), PrIter: A Distributed Framework for Prioritizing Iterative Computations, Parallel and Distributed Systems, IEEE Transactions on, vol.24, no.9, pp.1884,1893.

[5] Matei Zaharia, Mosharaf Chowdhury, Michael J. Franklin, Scott Shenker, Ion Stoica University of California, Berkeley: Spark: Cluster Computing with Working Sets.

[6] Matei Zaharia, Mosharaf Chowdhury, Tathagata Das, Ankur Dave, Justin Ma, Murphy McCauley, Michael J. Franklin, Scott Shenker, Ion Stoica University of California, Berkeley: Resilient Distributed Datasets: A Fault-Tolerant Abstraction for In-Memory Cluster Computing.

Cite this article as: SUBHASH KUMAR. "BIGDATA ANALYTICS WITH SPARK." *International Conference on Information Engineering, Management and Security (2015)*: 129-133. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS022		**eAID**	ICIEMS.2015.022

A Survey on Pattern classification with missing data Using Dempster Shafer theory

M.Kowsalya[1], Dr.C.Yamini[2]
[1]Research Scholar, Department of Computer Science,
Sri Ramakrishna College of Arts and Science for Women,Coimbatore
[2]Associate Professor, Department of Computer Science,
Sri Ramakrishna College of Arts and Science for Women, Coimbatore

Abstract: *The Dempster-Shafer method is the theoretical basis for creating data classification systems. In this system testing is carried out using three popular (multiple attribute) benchmark datasets that have two, three and four classes. In each case, a subset of the available data is used for training to establish thresholds, limits or likelihoods of class membership for each attribute for each attribute of the test data. Classification of each data item is achieved by combination of these probabilities via Dempster's Rule of Combination. Results for the first two datasets show extremely high classification accuracy that is competitive with other popular methods. The third dataset is non-numerical and difficult to classify, but good results can be achieved provided the system and mass functions are designed carefully and the right attributes are chosen for combination. In all cases the Dempster-Shafer method provides comparable performance to other more popular algorithms, but the overhead of generating accurate mass functions increases the complexity with the addition of new attributes. Overall, the results suggest that the D-S approach provides a suitable framework for the design of classification systems and that automating the mass function design and calculation would increase the viability of the algorithm for complex classification problems.*

Keywords: Dempster-Shafer theory, data classification, Dempster's rule of combination.

1. INTRODUCTION

The ability to group complex data into a finite number of classes is important in data mining, and means that more useful decisions can be made based on the available information. For example, within the field of medical diagnosis, it is essential to utilise methods that can accurately differentiate between anomalous and normal data. In DST, evidence can be associated with multiple possible events, e.g., sets of events. The chief aims here are to describe the use of the Dempster-Shafer (D-S) theory as a framework for creating classifier systems, test the systems on three benchmark datasets, and compare the results with those for other techniques.As a result, evidence in DST can be meaningful at a higher level of abstraction without having to resort to assumptions about the events within the evidential set. Where the evidence is sufficient enough to permit the assignment of probabilities to single events, the Dempster-Shafer model collapses to the traditional probabilistic formulation. One of the most important features of Dempster-Shafer theory is that the

Cite this article as: M.Kowsalya, Dr.C.Yamini. "A Survey on Pattern classification with missing data Using Dempster Shafer theory." *International Conference on Information Engineering, Management and Security (2015)*: 134-138. Print.

model is designed to cope with varying levels of precision regarding the information and no further assumptions are needed to represent the information. It also allows for the direct representation of uncertainty of system responses where an imprecise input can be characterized by a set or an interval and the resulting output is a set or an interval.

2.Data classification

Data classification is the process of organizing data into categories for its most effective and efficient use.
A well-planned data classification system makes essential data easy to find and retrieve. This can be of particular importance for risk management, legal discovery, and compliance. Written procedures and guidelines for data classification should define what categories and criteria the organization will use to classify data and specify the roles and responsibilities of employees within the organization regarding data stewardship Once a data-classification scheme has been created, security standards that specify appropriate handling practices for each category and storage standards that define thedata's lifecyle requirements should be addressed.

3. Dempster Shafer theory

The drawbacks of pure probabilistic methods and of the certainty factor model have led us in recent years to consider alternate approaches. Particularly appealing is the mathematical theory of evidence developed by Arthur Dempster. We are convinced it merits careful study and interpretation in the context of expert systems. This theory was first set forth by Dempster in the 1960s and subsequently extended by Glenn Sharer. In 1976, the year after the first description of CF's appeared; Shafer published A Mathematical Theory of Evidence (Shafer, 1976). Its relevance to the issues addressed in the CF model was not immediately recognized, but recently researchers have begun to investigate applications of the theory to expert systems (Barnett, 1981; Friedman, 1981; Garvey et al., 1981). We believe that the advantage of the Dempster-Shafer theory over previous approaches is its ability to model the narrowing of the hypothesis set with the accumulation of evidence, a process that characterizes diagnostic reasoning in medicine and expert reasoning in general. An expert uses evidence that, instead of bearing on a single hypothesis in the original
Equal certainty. Because he attributes belief to subsets, as well as to individual elements of the hypothesis set, we believe that Shafer more accurately reflects the evidence-gathering process. Hypothesis set, often bears on a larger subset of this set. The functions and combining rule of the Dempster-Shafer theory are well suited to represent this type of evidence and its aggregation.

4. New Method for Classification of Incomplete Patterns

The new prototype-based credal classification (PCC) method provides multiple possible estimations of missing values according to class prototypes obtained by the training samples. For a c-class problem, it will produce c probable estimations. The object with each estimation is classified using any standard2 classifier. Then, it yields c pieces of classification results, but these results take different weighting factors depending on the distance between the object and the corresponding prototype. So the c classification results should be discounted with different weights, and the discounted results are globally fused for the credal classification of the object. If the c classification results are quite consistent on the decision of class of the object, the fusion result will naturally commit this object to the specific class that is supported by the classification results. However, it can happen that high conflict among the c classification results occurs which indicates that the class of this object is quite imprecise (ambiguous) only based on the known attribute values. In such conflicting case, it becomes very difficult to correctly classify the object in a particular (specific) class, and it becomes more prudent and reasonable to assign the object to a meta-class (partial imprecise class) in order to reduce the misclassification rate. By doing this, PCC is able to reveal the imprecision of the classification due to the missing values which is a nice and useful property. Indeed in some applications, especially those related to defense and security (like in target classification) the robust credal classification results are usually more preferable than the precise classification results subject potentially to a high risk of error. The classification of the uncertain object in meta-class can be eventually precisiated (refined) using some other (costly) techniques or with extra information sources if it is really necessary. So PCC approach prevents us to take erroneous fatal decision by robustifying the specificity of the classification result whenever it is necessary to do it.
A.Determination of c estimations of missing values in incomplete patterns
Let us consider a test data set $X = \{x1, . . . , xN\}$ to be classified using the training data set $Y = \{y1, . . . , yH\}$ in the frame of discernment $\Omega = \{\omega 1, . . . , \omega c\}$. Because we focus on 2 In our context, we call standard a classifier working with complete patterns. The classification of the incomplete data (test sample) in this work, one assumes that the test samples are all incomplete data (vector) with single or multiple missing values, and the training data set Y consists of a set of complete patterns. The prototype of each class i.e. $\{o1, . . . , oc\}$ is calculated using the training data at first, and og corresponds to class ωg. There exist many methods to produce the prototypes. For example, the K-means method can be applied for each class of the training data, and the clustering center is chosen for the prototype. The simple arithmetic average vector of the training data in each class can also be considered as the prototype, and this

Cite this article as: M.Kowsalya, Dr.C.Yamini. "A Survey on Pattern classification with missing data Using Dempster Shafer theory." *International Conference on Information Engineering, Management and Security (2015)*: 134-138. Print.

method is adopted here for its simplicity. Mathematically, the prototype is computed for g = 1, . . . , c by og = 1 Tg X yj∈ωg yj (4) where Tg is the number of the training samples in the class ωg.

5. Basic mathematical terminology and the TBM

The D-S theory begins by assuming a frame of discernment (Θ), which is a finite set of mutually exclusive propositions and hypotheses (alternatives) about some problem domain. It is the set of all states under consideration. For example, when diagnosing a patient, Θ would be the set consisting of all possible diseases. The power set 2 Θ is the set of all possible sub-sets of Θ including the empty set Φ. For example, if:

Θ = {a, b}
Then
2 = { {, },{ },Θ}. Θ φ a b

The individual elements of the power set represent propositions in the domain that may be of interest. For example, the proposition "the disease is infectious" gives rise to the set of elements of Θ that are infectious and contains all and only the states in which that proposition is true. The theory of evidence assigns a mass value m between 0 and 1 to each subset of the power set. This can be expressed mathematically as:

2: →[ ].1,0 Θ m

The function (3) is called the mass function (or sometimes the basic probability assignment) whenever it verifies two axioms: First, the mass of the empty set must be zero:
m(φ) = ,0
and second, the masses of the remaining members of the power set must sum to 1:

∑⊆ Θ = A m(A) 1 .

The quantity m(A) is the measure of the probability that is committed exactly to A [3]. In other words, m(A) expresses the proportion of available evidence that supports the claim that the actual state belongs to A but not to any subset of A. Given mass assignments for the power set, the upper and lower bounds of a probability interval can be determined since these are bounded by two measures that can be calculated from the mass, the degree of belief (bel) and the degree of plausibility (pl). The degree of belief function of a proposition A, bel(A), sums the mass values of all the non-empty subsets of A:

bel (A) m(B).

The degree of plausibility function of A, pl(A), sums the masses of all the sets that intersect A, i.e. it takes into account all the elements related to A (either supported by evidence or unknown):

6. Advantages and disadvantages of D-S

The systems described in this paper are all based on the theory presented in Sections 2.1 and 2.2, but D-S-based systems have a great deal of scope and flexibility as regards to system design, which means that classifiers can be created that are highly suited for solving any given problem. In particular, there are no fixed rules regarding how the mass functions should be constructed or how the data combination should be organized For example, consider the case where a car window has been broken and there are three suspects Jon, Mary, and Mike, and two witnesses, W1 and W2. W1 assigns a mass value of 0.9 to "Jon is guilty" and a mass value of 0.1 to "Mary is guilty". However, W2 assigns a mass value of 0.9 to "Mike is guilty" and a mass value of 0.1 to "Mary is guilty". Applying the DRC returns a value of 0.99 for K, which yields a value of 1 for "Mary is guilty". This is clearly counterintuitive since both witnesses assigned very small mass values to this hypothesis. The conflicting beliefs management problem is only a cause for concern when there are more than two classes, so the WBCD dataset used here presents no potential problem. Furthermore, the mass functions used with other two datasets are selected so that any conflicting beliefs are reduced (see Sections 5.2 and 6.2). This is possible since the problem is caused by conflicting mass values, not mass functions, so one can design mass functions and DRC combination strategies that minimize the problem. Some alternative combination rules that attempt to reduce the conflicting beliefs management problem have also been proposed, as in [7] and [8], but none have yet been accepted as a standard method.

7. Review of D-S applications

The D-S theory has previously been shown to be a powerful combination tool, but to date most of the research effort has been directed towards using it to unite the results from a number of separate classification techniques. For example, in [30] the results from a

Bayesian network classifier and a fuzzy logic-based classifier are combined and in [31] the D-S theory is used in conjunction with a neural network methodology and applied to a fault diagnosis problem in induction motors. The DRC acts as a data fusion tool, i.e. eight faulty conditions are first classified using the neutral network and the classification information is then converted to mass function assignments. These are then combined using DRC, which reduces the diagnostic uncertainty. Al-Ani and Deriche [32] also propose a classifier combination method based on the D-S approach. They propose that the success of the D-S methodology lies in its powerful ability to combine evidence measures from multiple classifiers. In other words, when the results of several classifiers are combined, the effects of their individual limitations as classifiers are significantly reduced. Valente and Hermansky [33] also suggest a DRC methodology that combines the outputs from various neural network classifiers, but in their work it is applied to a multi-stream speech recognition problem. As mentioned previously, the work here differs from the above approaches in that it is concerned with classification using the D-S theory alone; no other categorization techniques are employed at any stage in the classification process. This perspective is fairly novel as other works concerned with the D-S theory as a single classifier has mostly focused on adapting its methodology. For example, Parikh et al. [34] present a new method of implementing D-S for condition monitoring and fault diagnosis, using a predictive accuracy rate for the mass functions. The author's claim that this architecture performs better than traditional mass assignment techniques as it avoids the conflicting beliefs assignment problem. In other D-S related work, Chen and Venkataramanan [35] show that Bayesian inference requires much more information than the D-S theory, for example a priori and conditional probabilities. They postulate that the D-S method is tolerant of trusted but inaccurate evidence as long as most of the evidence is accurate.

8. The application of D-S to data classification

As discussed in Section 2.3, the D-S theory provides a general framework for creating classifier systems. This framework can be expressed as a series of steps that must be undertaken namely: 1. Define the frame of discernment (Θ). This is the set of all possible hypotheses related to the given dataset and identifies the classes to which the data must be assigned. 2. Determine which data attributes are important for establishing class membership and discard the others. In general, the frame of discernment and the selected attributes (their number and their data types) will provide loose guidelines for designing mass functions and the structure of the DRC combinations. 3. Examine the selected attributes and their data values within a subset of the data in order to design mass functions for each attribute. These functions will be used to assign mass values to the corresponding hypotheses based on the attribute values of the test data. 4. Design a DRC combination strategy based on the data structure. A single application of DRC combines the mass values of each attribute for each data item, but many applications can be used, and DRC can also be used to combine the results of previous applications. 5. Following combination, select a rule that converts the result to a decision. Several may be used on different steps, but the final one ultimately classifies the data.

9. Conclusions

This work has utilized the D-S theory (in particular mass functions and DRC) as a framework for creating classification algorithms, and has applied them to three standard benchmark datasets, the WBCD dataset, the Iris dataset, and part of the Duke Outage dataset. For the WBCD, the mass functions were created by considering threshold values in the training data and using a sigmoid model. In this case, classification was a simple one-step process. The accuracy proved to be much higher when all the data attributes were considered (97.6%), and this result was superior to other published results for other popular methods. Furthermore, the D-S method permitted the inclusion of data items that contained missing values in the dataset. Some of the other methods were unable to do this. This paper has hence demonstrated that the D-S approach works well with all three datasets provided the system is designed in the right way and the attributes are carefully selected. Attribute selection appears to influence overall performance considerably, for example, use of all the attributes worked well for the WBCD but not for the Duke Outage data. The D-S theory provides the framework for system design only, and in this sense allows the creation of systems that can be essentially tailored towards the specific problem domain of interest. This may be considered a disadvantage in that there are no strict guidelines for the detailed design of such systems, but it may also be thought of as an advantage, since the flexibility allows for the tweaking and refinement of the system until the desired output levels are reached, especially if this refinement process can be automated in some way. In particular, automating the attribute selection and mass function calculation processes may make the Dempster-Shafer approach an objective and accurate replacement for current state of the art classification systems.

REFERENCES

[1] R.J.LittleandD.B.Rubin, StatisticalAnalysisWithMissingData, 2nded.NewYork, NY, USA: Wiley, 2002.

[2] A.P. Dempster, "Upper and lower probabilities induced by a multivalued mapping", Ann. Math. Statist. 38 (1967) 325-339.

[3] G. Shafer, "A Mathematical Theory of Evidence", Princeton University Press, Princeton and London, 1976. [4]K. A. Lawrence, "Sensor and Data Fusion: A Tool for Information Assessment and Decision Making", SPIE, Washington, 2004.

[4] B. Tessem, "Approximations for efficient computation in the theory of evidence", Artificial Intelligence 61 (1993) 315-329.

[5] K.Jian,H.Chen,andS.Yuan,"Classificationforincompletedatausingclassifierensembles,"inProc.Int.Conf.NeuralNetw.Brain(ICNN&B'05),Beijing,China,Oct.,pp.55 9–563.

[6] K.Pelckmans,J.D.Brabanter,J.A.K.Suykens,andB.D.Moor,"Handlingmissingvaluesinsupportvectormachineclassifiers,"NeuralNetw.,vol.18,nos.5–6,pp.684–692,2005.

[7] P.ChanandO.J.Dunn,"Thetreatmentofmissingvaluesindiscriminantanalysis,"J.Amer.Statist.Assoc.,vol.6,no.338,pp.473–477,1972.

Cite this article as: M.Kowsalya, Dr.C.Yamini. "A Survey on Pattern classification with missing data Using Dempster Shafer theory." *International Conference on Information Engineering, Management and Security (2015)*: 134-138. Print.

[8] J.L.Schafer,AnalysisofIncompleteMultivariateData.London,U.K Chapman&Hall,1997.
[9] O.Troyanskayaetal.,"MissingvalueestimationmethodsforDNAmicroarrays,"Bioinformatics,vol.17,no.6,pp.520–525,2001.
[10] G.BatistaandM.C.Monard,"AstudyofKnearestneighbourasanimputationmethod,"inProc.2ndInt.Conf.HybridIntell.Syst.,2002,pp.251–260.
[11] J.Luengo,J.A.Saez,andF.Herrera,"Missingdataimputationforfuzzyrule-basedclassificationsystems,"SoftComput.,vol.16,no.5,pp.863–881,2012.
[12] [12].D.Li,J.Deogun,W.Spaulding,andB.Shuart,"Towardsmissingdataimputation:Astudyoffuzzyk-meansclusteringmethod,"inProc.4thInt.Conf.RoughSetsCurrentTrendsComput.(RSCTC04),Uppsala,Sweden,Jun.2004,pp.573–579.
[13] F.FessantandS.Midenet,"Self-organizingmapfordataestimationandcorrectioninsurveys,"NeuralComput.Appl.,vol.10,no.4,pp.300–310,2002.
[14] Y. Song, J. Huang, D. Zhou, H. Zha and C. Lee Giles, in: J. N. Kok et al. (Eds.), KNN: Informative K-Nearest Neighbor pattern classification, PKDD 2007, LNAI 4702, Springer-Verlag Berlin Heidelberg, 2007, pp. 248–264.

Cite this article as: M.Kowsalya, Dr.C.Yamini. "A Survey on Pattern classification with missing data Using Dempster Shafer theory." *International Conference on Information Engineering, Management and Security (2015)*: 134-138. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	**VOL**	01
Website	www.iciems.in	**eMail**	iciems@asdf.res.in
Received	10 - July - 2015	**Accepted**	31- July - 2015
Article ID	ICIEMS023	**eAID**	ICIEMS.2015.023

Study of an Orchestrator for Centralized and Distributed Networked Robotic Systems

Rameez Raja Chowdhary[1], Manju K. Chattopadhyay[1], Raj Kamal[2]

[1]Assistant Professor, School of Electronics, Devi Ahilya University, Indore-452001, India

[2]Professor, Department of Information Technology, Medicaps Institute of Science and Technology, Indore-453331, India

Abstract: *This paper presents an Orchestrator, developed for execution of given tasks on robotic nodes. The tasks execute in two modes, parallel and sequential. The orchestrator assigns task to multiple sets of robotic nodes. The nodes of a set perform the assigned tasks either synchronously or in parallel. Each set divides the given task in subtasks and each subtask further performs on different robotic nodes sequentially. Each robot has unique address. Robots possess capability to interact with each other using RF radio. Networked controlled robots (NCRs) imbibe two addition properties—fault tolerance and greater system efficiency. The experimental results of study of orchestration are also presented using the Orchestrator.*

Keywords: Robotics, Orchestration, Networks and Distributed Tasks.

I. INTRODUCTION

Networked controlled robots are subject to research in recent years. They find uses in civil, military and space applications. This research focuses on two types of NCRs, centralised and decentralised. A Robotic Systems' Network is a group of artificial autonomous systems which are mobile. The systems communicate among themselves either directly as in case of distributed network robotic systems [1] or through a central controller in case of centralised network robotic systems. Distributed autonomous robots are designed to perform collaborated mission [2], whose success depends on communication between individuals. Therefore, robots count sufficient knowledge of the network connectivity, and exploit this knowledge in order to best maintain the network connectivity while performing other tasks [3-5]. A broad challenge is to develop a model architecture that couples communication with control to enable such new capabilities, like Cloud Robotics, Global localization in Mobile Robot, Fleet management, Assert tracking and Covert surveillance [6, 7]. Orchestration deploys elements of control theory [8]. The usage of orchestration has been studied earlier in context of the service oriented architecture, virtualization, provisioning, converged infrastructure and dynamic data centre [9]. One service may be realized using orchestration through the cooperation of several services. Orchestration can also be defined as a type of cooperation in which the one service directly invokes other services. A new approach "Robotic Orchestration" is introduced, in this paper, which has advantages and minimum disadvantages of both approaches. An Orchestrator controlled robotic network performs Robotic Orchestration. It describes the automated arrangement, coordination, the management of complex robotic systems, and the services. The work focuses on developing a Robotic Orchestrator (Coordinator). This Orchestrator is able to invoke and coordinate other services by exploiting typical workflow patterns such as parallel composition, sequencing and choices or a manager which

Cite this article as: Rameez Raja Chowdhary, Manju K. Chattopadhyay, Raj Kamal. "Study of an Orchestrator for Centralized and Distributed Networked Robotic Systems." *International Conference on Information Engineering, Management and Security (2015)*: 139-143. Print.

controls and coordinates the functions and roles of the nodes (Slave robots). Orchestrator organizes and manages a set of activities in a network. Orchestrator assigns service to a robotic node. A robotic node performs the given service and after completing the service messages to the Orchestrator about the service completion. The paper is organised as follows. Section II presents the basic problem statement. Section III describes the hardware developed. Section IV gives the experimental results. Conclusions derived from the present study are given in Section V.

II. PROBLEM FORMUATION

This section describes the task flow pattern in network robotic system. Figure 1 shows the task flow pattern. The results of three coordinating ways are compared, after introduction of the task flow pattern and defining the experimental procedure.

Let H denote a set of k heterogeneous robots, and T denote a set of n tasks, that is,

$$H = \{h_1, h_2, \ldots, h_k\}, \qquad (1)$$

$$T = \{t_1, t_2, \ldots, t_n\}. \qquad (2)$$

The tasks are assigned to robotic node on fixed time interval like $(t_{s1}, t_{s2}, \ldots.)$ and Ts_i define the set of time interval,

$$Ts = \{t_{s1}, t_{s2}, \ldots, t_{sn}\}. \qquad (3)$$

Each time slots are equal $t_{s1} = t_{s2} = \ldots, = t_{sn}$.

Also, let A denote the allocation

$$A = \{a_1, a_2, \ldots, a_k\}, \qquad (4)$$

where a_s is a cluster of tasks,

$$\cup_{s=1}^{k}(a_s) = T\ , \qquad (5)$$

$$a_s \cap a_t = \emptyset,\ (s \neq t), \qquad (6)$$

and the cluster a_s is assigned to robot h_s.

The cost associated with A is given by [10]

$$C(A) = \sum_{s=1}^{k} c_s(a_s) \qquad (7)$$

where $c_s(a_s)$ is the minimum cost for robot h_s to complete the set of tasks a_s. In practice the cost function in (7) might be used to represent the total distance traveled or the total energy expended by the robots.

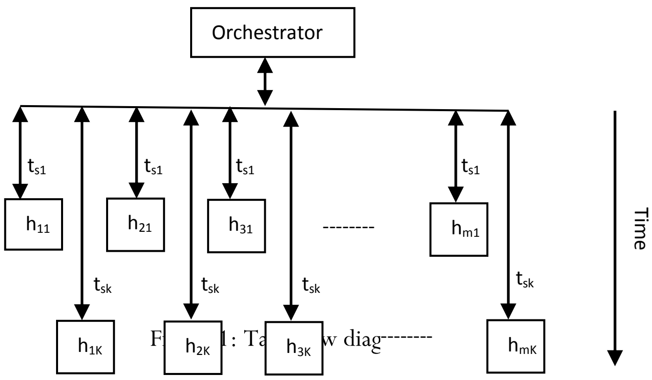


Figure 1: Task flow diagram

Robotic Orchestration in a networked robotic system is performed by defining a task of Robotic Relay Racing (RRR). RRR is similar to human relay racing in Olympics games. The two standard relays are the 4x100 meter relay and the 4x400 meter relay. A 4x400 relay race is a race in which four runners of each teams completes race between starting point to end point. Generally 4x400 relay starts in lanes for the first runner and first runner handoff the baton to second runner after first 100 meter. The second runner handoff the baton to third runner after second 100 meter and so on. The last runner completes the race to the finish line.

The Robotic relay race uses similar rules but some variation. The master (Robotic Orchestrator) controls the race. Each slave (Robotic node) takes command from orchestrator and after performing task give reply to orchestrator. The nodes are unable to communicate with each other.

The Robotic Relay Racing (RRR) is demonstrated by Figure 2. Here one Orchestrator (master) and four Robotic nodes (RN) are used to perform the experiments. The RN is divided in two groups '1' and '2' and each group has two members. The group '1' members are 1a and 1b, group '2' members are 2a and 2a. When Orchestrator starts the race 1a and 2a starts moving on and after reaching their finishing line both RN passing baton to their team members. Now RN 1b and 2b continue the race up to finishing line.

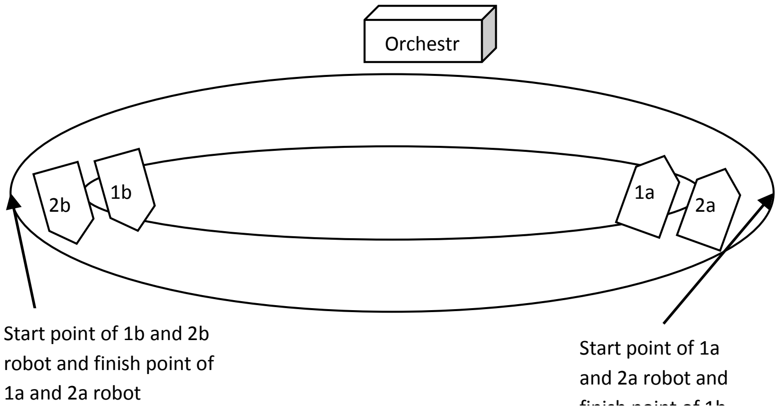


Figure 2: Robotics Relay Racing for Orchestrator control Robotic node.

The RRR experiment is performed by all three approaches discussed in this paper. The parameters used in experiments are shown by Table 1. The comparison is made by measuring the following three parameters.

1. Number of message passing in robotic network to perform given task.
2. Execution time of task.
3. Bandwidth for communication.

The approaches below have also been followed by other researchers [10-13] but not for Robotic Orchestration.

Experiment 1: Centralized RRR is performed under the following points.

1. The RN cannot communicate to each other directly. They can communicate to each other only through master.
2. All RN are using a single communication channel.
3. The master have task plan for each RN.
4. Each RN performing given task under monitoring of master.

TABLE I.
PARAMETERS USED FOR EXPERIMENTS.

S. No.	Parameters	Units
1	Number of robots	4 slave and 1 Master
2	Track Length	10m
3	Communication Range of Robots	12m
4	Max. Achievable speed of robots	$0.5ms^{-1}$
5	Normal speed of robots	$0.2\ ms^{-1}$
6	Min. Distance between robots to avoid collision	0.3m
7	Max. distance between robots to keep on track	1.1m

Experiment 2: Decentralized RRR is performed under the following points.

1. All RN communicate to each other through separate channel.
2. Each RN must know status of each of other RN in network to perform the collaborate mission.
3. The master only initiates the task. After that, the RN collaboratively performs the given task.

Experiment 3: Orchestrator RRR is performed under the following points.

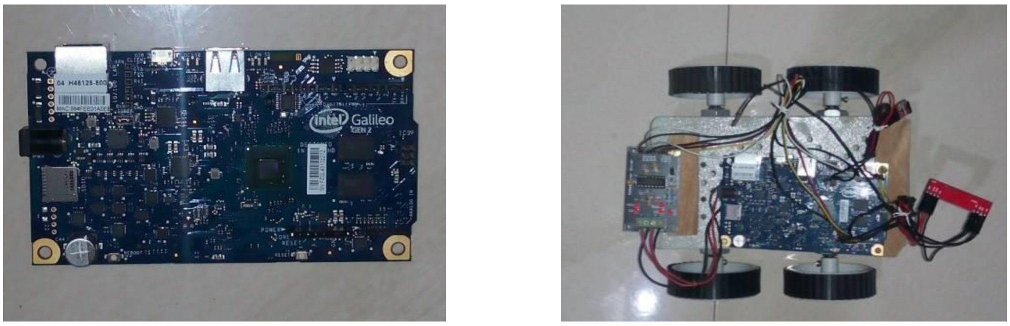
1. The orchestrator only assigns the tasks and monitors the status of task.
2. The task plan is preloaded in RN similar to a musical orchestra member.
3. The RN has limited communication capability to each other.
4. Each RN has two communications channel, one for Orchestrator and other for group member.

Cite this article as: Rameez Raja Chowdhary, Manju K. Chattopadhyay, Raj Kamal. "Study of an Orchestrator for Centralized and Distributed Networked Robotic Systems." *International Conference on Information Engineering, Management and Security (2015)*: 139-143. Print.

III. HARDWARE DISCRIPTION

Intel Galileo Gen 2 development board is shown in Figure 3(a) which is used for designing Orchestrator and RN. It is based on the Intel QuarkSoC X1000, a 32-bit Intel Pentium processor-class system on a chip (SoC), the genuine Intel processor and native I/O capabilities of the Intel Galileo board (Gen 2) provides a full-featured offering for a wide range of applications. The Intel Galileo Gen 2 board also provides a simpler and more cost-effective development environment compared to the Intel Atom processor- and Intel Core processor-based designs. The Intel Galileo board (Gen 2) is an open source hardware design [14].

The Orchestrator explained in Figure 2 is shown in Figure 3(b). RN ‘1a’ and ‘1b’ are shown in Figure 4(a) and RN ‘2a’ and ‘2b’ are shown in Figure 4(b). An Orchestrator is two wheel drive robot. Orchestrator uses two 60 rpm, 12V DC motors as a power train. The RF radio (TX/RX) is used for communication in robotic network. This radio frequency (RF) transmission system employs Amplitude Shift Keying (ASK) with transmitter/receiver (Tx/Rx) pair operating at 434 MHz. The transmitter module takes serial input and transmits these signals through RF. The transmitted signals are received by the receiver module placed away from the source of transmission.



(a) (b)

Figure 3: (a) Intel Galileo Gen 2 board. (b) Robotic Orchestrator for Robotic relay racing.

(a) (b)

Figure 4: Robotic nodes for Robotic relay racing. (a) Robotic node 1a and 1b. (b) Robotic node 2a and 2b.

IV. EXPERIMENTAL RESULTS

The experiments are performed for all three approaches according to the predefined points. Results are explained by three graphs. The graph in Figure 5(a) shows the task execution time for all three approaches. The centralized approach takes 90 second to complete the given task, distributed approach takes 55 second and orchestrator approach takes 56 second to complete the given task. The graph in Figure 5(b) shows number of message passing in network between RN and master for completing the task. The centralized approach passes 13 messages, distributed approach passes 8 messages and orchestrator approach passes 9 messages to complete the given task. Figure 5(c) shows number of communication channel required for completing the task according to predefine statement made.

Cite this article as: Rameez Raja Chowdhary, Manju K. Chattopadhyay, Raj Kamal. “Study of an Orchestrator for Centralized and Distributed Networked Robotic Systems.” *International Conference on Information Engineering, Management and Security (2015)*: 139-143. Print.

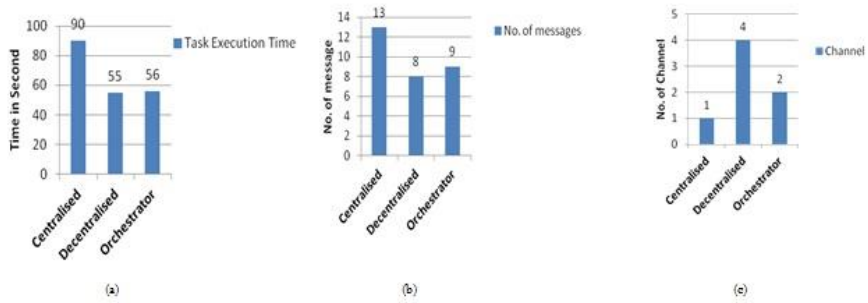


Figure 5: (a)Task execution time of all three approaches in second,(b)Number of message during task execution in all three approaches, (c) Number of communication channels used in all three approaches.

V. CONCLUSION

Experimental results show that the centralised approach has advantage of communication bandwidth but takes more time to execute the given task. This is because of number of messages passed in network is large. The decentralised approach executes the given task fast due to lesser accesses to communication channels. This reduces the communication overhead. The Orchestration approach lies between these two approaches. Orchestration reduces the number of messages passing on network robotic system and also reduces the task execution time.

The centralised approach has strong dependency on master. If master fails due to some reason the robotic system network is unable to complete the task. While in orchestration approach, each RN knows its task, so replacing an Orchestrator is not a tedious job. The orchestration approach in robotic is good for military control operation, disaster management and other areas. These applications need a monitoring authority for completing task. The task can be modified during the execution, whereas, modification is not possible in decentralised approach.

REFERENCES

[1] Z. Movahedi, M. Ayari, R. Langar, and G. Pujolle, "A survey of autonomic network architectures and evaluation criteria," IEEE Communications Surveys and Tutorials, vol. 14, no. 2, pp. 464-490, 2012.
[2] A. Khamis and A. ElGindy, " Minefield mapping using cooperative multirobot systems," Journal of Robotics, vol. 2012, pp. 1-17, Oct. 2012.
[3] V. Kumar, D. Rus, and S. Singh, "Robot and sensor networks for first responders," Pervasive Computing, IEEE , vol. 3, no.4, pp. 24-33, Oct.-Dec. 2004.
[4] V. T. Le, N. Bouraqadi, S. Stinckwich, V. Moraru, and A. Doniec, "Making networked robots connectivity-aware, Proceeding of the IEEE International Conference on Robotics and Automation (ICRA '09), pp. 3502-3507 , 2009.
[5] E. Cardozo, E. Guimaraes, L. Rocha, R. Souza, F. Paolieri, and F. Pinho, "A platform for networked robotics, Proceeding of the IEEE/RSJ International Conference Intelligent Robots and Systems (IROS), pp. 1000-1005, 2010.
[6] G. Hu, W. P. Tay and Y. Wen, "Cloud robotics: architecture, challenges and applications," Network, IEEE, vol. 26, no. 3, pp. 21-28, May-June 2012.
[7] T. He and S. Hirose, "Observation-driven Bayesian Filtering for Global Location Estimation in the Field Area, " Journal of Field Robotics, vol. 30, no. 4, pp. 489–518, 2013.
[8] R. Lundh, L. Karlsson, and A. Saffiotti, "Autonomous functional configuration of a robotic systems network". Journal of Robotics and Autonomous Systems, vol. 56, pp. 819–830, 2008.
[9] T. Erl. Service-Oriented Architecture: Concepts, Technology and Design. Pearson, 2005.
[10] K. Zhang and E. G. Collins, "Centralized and distributed task allocation in multi-robot teams via a stochastic clustering auction" ACM Transactions on Autonomous and Adaptive Systems, vol. 7, no. 2, pp. 21.1-21.22 July 2012.
[11] W. Ren and N. Sorensen," Distributed coordination architecture for multi-robot formation control," Journal of Robotic and Autonomous Systems, vol. 56, pp. 324-333, 2008.
[12] M. Koes, K. Sycara, and T. Nourbakhsh, "A constraint optimization framework for fractured robot teams.", Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS), pp. 491–493, 2006.
[13] M. Hoy, A. S. Matveev and A. V Savkin., "Collision free cooperative navigation of multiple wheeled robots in unknown cluttered environments, " Journal of Robotic and Autonomous Systems, vol. 60, pp.1253-1266, 2012.
[14] Intel Galileo Gen 2 development board, Datasheet, Intel Corporation, 2014.

Cite this article as: Rameez Raja Chowdhary, Manju K. Chattopadhyay, Raj Kamal. "Study of an Orchestrator for Centralized and Distributed Networked Robotic Systems." *International Conference on Information Engineering, Management and Security (2015)*: 139-143. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS024

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.024

Utilization of Rough Set Reduct Algorithm and Evolutionary Techniques for Medical Domain using Feature Selection

T.Keerthika[1], Dr.K.Premalatha[2]
[1] Assistant Professor, Department of Information Technology, Sri Krishna College of Engineering and Technology, Tamil Nadu, India
[2] Professor, Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Tamil Nadu, India

ABSTRACT: *With the real time data, results in increasing in size. Feature selection (FS) has been considered as the problem of selecting these input features that are most predictive of a given outcome. Also current methods are inadequate. By considering this scenario, this paper proposes the incremental techniques; in fact this has found unsuccessful application in tasks that involve datasets contain huge number of features, which could be impossible to process further. For achieving this, these evolutionary techniques such as Genetic Algorithm, Particle Swarm Optimization Algorithm and Ant Colony Algorithm are considered for comparative performance analysis in which the experimental results shows that feature selection is best for minimal reductions.*

Keywords: Feature selection, rough set theory, Genetic Algorithm, Particle Swarm Optimization, Ant Colony Algorithm

1. Introduction

The solution to the dimensionality reducing problem has been of prior importance and worked in a variety of fields like statistics, pattern recognition, discovery through knowledge and machine learning. two major techniques done for reducing the input dimensionality are:

- feature extraction
- feature selection.

The concept behind feature extraction is that the lower dimensionality is used when a primitive feature space is mapped onto a new space. Principle component analysis, partial least squares are the two important approach for the process of feature extraction. The various applications of feature extraction is applied in variety of fields that include literature, where image processing, visualization and signal processing plat an important role.

Cite this article as: T.Keerthika, Dr. K. Premalatha. "Utilization of Rough Set Reduct Algorithm and Evolutionary Techniques for Medical Domain using Feature Selection." *International Conference on Information Engineering, Management and Security (2015)*: 144-151. Print.

Unlike the feature extraction process, feature selection(FS)involves the selection methods of t-statistic, f–statistic, correlation, reparability correlation measure, or information gain which chooses the most appropriate and informative features from the original one using the methods above. The low accuracy and slow learning difficulties occurs because of redundancy and irrelevancy in the dataset. Finding the subset of features that are informative enough is NP complete. Heuristic algorithm is administered to invoke a search process through feature space. The complexity of learning an algorithm and defining its accuracy are some issues used for evaluating the selected subset.

The rough set(RS) [12,13,14] is a helping tool that reduces the problem of input dimensionality and finds a better solution for correcting the vague and uncertain datasets. The reduction of attributes is based on data dependencies. The RS theory partitions a dataset into some equivalent (indiscernibility) classes, and approximates uncertain and vague concepts based on the partitions. A function of approximation is used to calculate the measure of dependency. The measure of dependency is regarded as a heuristic in order to guide the process of FS. Proper approximations of concepts are very essential to obtain a significant measure which makes the initial partitions to be vital in this matter. For a given number of discrete dataset, finding the indiscernibility classes are feasible ,but in the case of real-valued attributes, one can't be sure if the two objects mentioned are the same, or by what relation they are same using the above mentioned indiscernibility relation. A team of research persons extended this RS theory by the usage of tolerant or similarity relation (termed tolerance-based rough set).

The similarity measure between two objects is delineated by a distance function of all attributes .when the similarity measure is exceeding a similarity threshold value, the objects are said to be similar. The important and challenging job is to find the best threshold boundary.

2. RELATED WORKS

A. ROUGH-SET BASED INCREMENTAL APPROACH:

In this approach [1], the approximations of a concept by a variable precision rough-set model (VPRS) usually vary under a dynamic information system environment. It is thus effective to carry out incremental updating approximations by utilizing previous data structures. This paper focuses on a new incremental method for updating approximations of VPRS while objects in the information system dynamically alter. It discusses properties of information granulation and approximations under the dynamic environment while objects in the universe evolve over time. The variation of an attributes domain is also considered to perform incremental updating for approximations under VPRS. Finally, an extensive experimental evaluation validates the efficiency of the proposed method for dynamic maintenance of VPRS approximations.

B. Novel Dynamic Incremental Rules Extraction Algorithm Based on Rough Set Theory:

In this paper, a novel incremental rules extraction algorithm which is called "RDBRST" (Rule Derivation Based on Rough Set And Search Tree) is proposed. It is one kind of width first heuristic search algorithms. The incremental rules are extracted and the existing rule set is updated based on this algorithm [2]. Incremental Induction of Decision Rules from Dominance-Based Rough Approximations. It is extended to handle preference-ordered domains of attributes (called criteria) within Variable Consistency Dominance-based Rough Set Approach. It deals, moreover, with the problem of missing values in the data set. The algorithm has been designed for medical applications which require: (i) a careful selection of the set of decision rules representing medical experience and (ii) an easy update of these decision rules because of data set evolving in time, and (iii) not only a high predictive capacity of the set of decision rules but also a thorough explanation of a proposed decision. To satisfy all these requirements, we propose an incremental algorithm for induction of a satisfactory set of decision rules and a post-processing technique on the generated set of rules.

C. A DISTANCE MEASURE APPROACH TO EXPLORING THE ROUGH SET BOUNDARY REGION FOR ATTRIBUTE REDUCTION

This paper examines a rough set FS technique which uses the information gathered from both the lower approximation dependency value and a distance metric which considers the number of objects in the boundary region and the distance of those objects from the lower approximation. The use of this measure in rough set feature selection can result in smaller subset sizes than those obtained using the dependency function alone. This demonstrates that there is much valuable information to be extracted from the boundary region [5].

D. INCREMENTAL LEARNING OF DECISION RULES BASED ON ROUGH SET THEORY

In this paper, based on the rough set theory, the concept of ∂-indiscernibility relation is put forward in order to transform an inconsistent decision table to one that is consistent, called ∂-decision table, as an initial preprocessing step[4]. Then, the ∂-decision

Cite this article as: T.Keerthika, Dr. K. Premalatha. "Utilization of Rough Set Reduct Algorithm and Evolutionary Techniques for Medical Domain using Feature Selection." *International Conference on Information Engineering, Management and Security (2015)*: 144-151. Print.

matrix is constructed. On the basis of this, by means of a decision function, an algorithm for incremental learning of rules is presented. The algorithm can also incrementally modify some numerical measures of a rule.

3. PROPOSED SYSTEM

The Proposed system idea is to develop a new feature selection mechanism based on Ant Colony Optimization to combat this difficulty. It also presents a new entropy based modification of the original rough set-based approach. These are applied to the problem of finding minimal rough set reducts, and evaluated experimentally.

- Feature selection methods, the Importance Score (IS) which is based on a greedy-like search and a genetic algorithm-based (GA) method, in order to better understand.
- This proposed work is applied in the medical domain to find the minimal reducts and experimentally with the Quick Reduct, Entropy Based Reduct, and other hybrid Rough Set methods such as Genetic Algorithm (GA), Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO).

Advantages:

- Reducing the dimensionality of the attributes reduces the complexity of the problem and allows researchers to focus more clearly on the relevant attributes.
- Simplifying data description may facilitate physicians to make a prompt diagnosis.
- Having fewer features means less data need to be collected, result in time-consuming and costly.

The proposed work can be explained with the help of the system flow diagram in Figure-1as below,

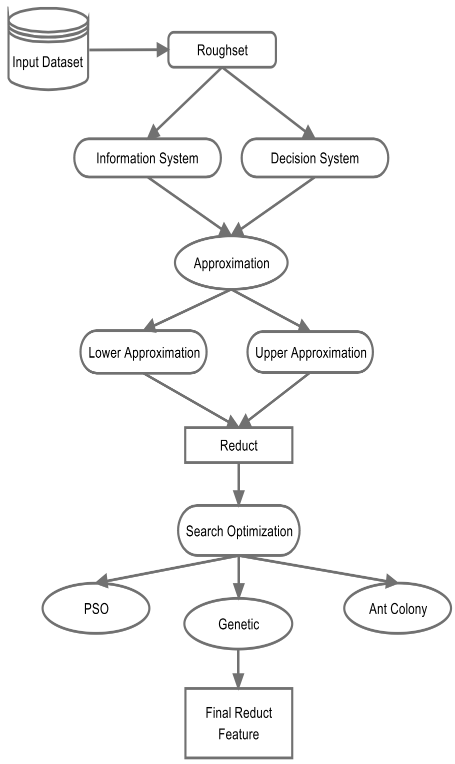


FIGURE 1. SYSTEM FLOW DIAGRAM

4. FEATURE SELECTION APPROACH

The main objective of FS (feature selection) is to pick out from the problem domain the minimal feature subset, such that it represents the original features with an outstanding accuracy given. FS plays a predominant role in real world issues and problems because of the irrelevant, noisy and misleading features of the data that are plenty in numbers. In case of reducing these irrelevant data's, the process of learning from the data technique can be beneficial for the users. The work of FS is to search a feature subset that is the most optimal(that varies depending on the problem to be solved) from the given n size of feature set by competing with candidate subset of size n. But this method is not feasible even though an exhaustive methodology is used.

The searching for the datasets are done randomly in order to cease this complexity. But in that case the extent of getting an optimal solution is drastically brought down.

Cite this article as: T.Keerthika, Dr. K. Premalatha. "Utilization of Rough Set Reduct Algorithm and Evolutionary Techniques for Medical Domain using Feature Selection." *International Conference on Information Engineering, Management and Security (2015)*: 144-151. Print.

The degree to which a feature subset or may be a feature may be useful is based on two important factors: 1. relevancy 2.redundancy. Relevancy depends on its ability to predict the decision feature(s), if not the datas are said to be irrelevant. Redundant feature must be correlating with other features. So an optimal search to find the best feature subset must be its ability to have a correlation between the decision features but must not be correlating apart from that.

When it comes to subset minimality and subset suitability, a tradeoff occurs with these non-exhaustive techniques and it becomes likely to choose between the two so that one will benefit over the other. Choosing this optimality is a challenging one. Involving situations when the inspection of many features is not possible, it is better to switch on to a subset feature that is much smaller and has a lesser accuracy amount.

For instance , classification rate that is a feature of modeling accuracy should be very high when the user is using selected features, by taking the expense of a non-minimal feature subset.

On the basis of evaluation procedure,there are two important classification in feature selection algorithm.the first one is the filter approach where the FS works independently and which is a separate pre-processor to any learning algorithm. This approach is applied in all the domains as they are very effective in filtering all irrelevant attributes before induction and no any specific induction algorithm is used.

The next is wrapper approach which involves tying up of evaluation procedure to a task of any learning algorithm as in the case of classification. This method employs an accuracy estimation that can search through the spaces of feature subset with the help of an induction algorithm that measures suitability of subsets. Wrapper are the ones that produce good results when compared with the other but faces the difficulties of a break down when large number of features are fed into it and also makes it expensive to run because of the learning algorithm used, that invokes the problem when large datasets are used.

5. Rough Set-Based Feature Selection APPROACH

Rough set theory (RST) can discover data dependencies. They can curtail the attributes found in the dataset by using only the data and not any additional information. this is a topic in trend that lures many researches to work on it and has been applied in various domains and fields over the past decade. Using RST it is possible to search for the right subset that is often termed as reduct when discretized attribute values are given in a dataset; the rest of the attributes can be taken out from the dataset with minimal loss of information. From the view of dimensionality, the one with the predictive nature of class attribute are often called the informative feature. Finding rough set reducts are put into two approaches: One for to estimate the degree of dependency and the other for discemibility matrix consideration. This section describes the fundamental ideas behind both of these approaches There are two main approaches to finding rough set reducts: those that consider the degree of dependency and those that are concerned with the discernibility matrix. This section describes the fundamental ideas behind both approaches. To illustrate the operation of these, an example dataset (Table 1) will be used.

Table 1. An example dataset

*x*U	a	b	c	d	e
0	1	0	2	2	0
1	0	1	1	1	2
2	2	0	0	1	1
3	1	1	0	2	2
4	1	0	2	0	1
5	2	2	0	1	1
6	2	1	1	1	2
7	0	1	1	0	1

A. Rough Set Attribute Reduction

Central to Rough Set Attribute Reduction (RSAR) is the concept of indiscernibility. Let $I = (U, A)$ be an information system, where U is a non-empty set of finite objects (the universe) and $A$is a non-empty finite set of attributes such that $a{:}U \rightarrow V_a$ for every $a \in A$. V_a is the set of values that attribute a may take. With any $P \in A$ there is an associated equivalence relation $IND(P)$:

B. Information and Decision Systems

An information system can be viewed as a table of data, consisting of objects (rows in the table) and attributes (columns). In medical datasets, for example, patients might be represented as objects and measurements such as blood pressure, form attributes. The attribute value for a particular patient is their specific reading for that measurement. Throughout this paper, the terms attribute, feature and variable are used interchangeably.

An information system may be extended by the inclusion of decision attributes. Such a system is termed a decision system. For example, the medical information system mentioned previously could be extended to include patient classification information, such as whether a patient is ill or healthy. A more abstract example of a decision system can be found in table 1. Here, the table consists of

Cite this article as: T.Keerthika, Dr. K. Premalatha. "Utilization of Rough Set Reduct Algorithm and Evolutionary Techniques for Medical Domain using Feature Selection." *International Conference on Information Engineering, Management and Security (2015)*: 144-151. Print.

four conditional features (a; b; c; d), a decision feature (e) and eight objects. A decision system is consistent if for every set of objects whose attribute values are the same, the corresponding decision attributes are identical.

6. A. ANT COLONY OPTIMIZATION FOR FEATURE SELECTION

Swarm Intelligence (SI) is the property of a system whereby the collective behaviors of simple agents interacting locally with their environment cause coherent functional global patterns to emanate. This provides a basis with which it is possible to explore collective (or distributed) problem solving without centralized control or the provision of a global model. Particle Swarm Optimization is one area of interest in SI, a population-based assumptive optimization technique. Here, the system is initialized with a population of random solutions, called particles. Optima are searched for by updating generations, with particles moving through the parameter space towards the current local and global optimum particles. The velocities of all particles are changed depending on the current optima, at each time step.

Ant Colony Optimization (ACO) is another area of interest within SI. In nature, it can be observed that real ants are capable of finding the shortest route between a food source and their nest without the use of visual information and hence possess no global world model, adapting to changes in the environment. The deposition of pheromone is the main factor in enabling real ants to find the shortest routes over a period of time.In this chemical, each ant probabilistically prefers to follow a direction. Over time the pheromone decays, which results in much less pheromone on less popular paths. Provided that over time the shortest route will have the higher rate of ant traversal, this path will be reinforced and the others will be diminished until all ants follow the same, shortest path (the "system" has converged to a single solution). There is a possibility that there are many equally short paths. In this situation, the rates of traversal of ants over various the short paths will be roughly the same, which results in these paths being maintained while the others are ignored. In addition, if there is a sudden change to the environment (e.g. a large obstacle appears on the shortest path), the ACO system responds to this and will eventually converge to a new solution. Based on this idea, artificial ants can be deployed to solve complex optimization problems via the use of artificial pheromone deposition.

ACO is particularly attractive for feature selection as there seems to be no heuristic that can guide search to the optimal minimal subset every time. Additionally, it can be the case that ants discover the best feature combinations as they proceed throughout the search space. This section discusses how ACO may be applied to the difficult problem of finding optimal feature subsets and, in particular, fuzzy-rough set-based reducts.

The ACO-suitable problem is formulated from feature selection task. ACO represents the problem as a graph where the nodes represent features and hte edges between them denote the choice of the next feature. The search for the optimal feature subset is an ant traversal through the graph where a minimum number of nodes are visited which satisfies the criteria for stopping the traversal. Illustrates this setup - the ant is currently at node *a* and has a choice of which feature to add next to its path (dotted lines).Next feature *b* is choosed based on transition rule, followed by c and *d*. With the arrival at *d*, the current subset *fa; b; c; dg* is determined to satisfy the traversal stopping criteria (e.g. a suitably high classification accuracy has been achieved with this subset, based on the assumption that the selected features are used to classify certain objects).The ant outputs this feature subset as a candidate for data reduction by terminating its traversal.

C. GENETIC ALGORITHM FOR FEATURE SELECTION

Genetic algorithm (GA) is a search heuristic, used to generate solutions to optimization problems following the techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. In the genetic algorithm,

- A population of strings (called chromosomes), which encode candidate solution to an optimization problem is taken.
- A proper fitness function is then constructed, and the fitness of the current population is evaluated.
- Two fittest chromosomes are chosen as the parents and (a) crossing over between them or (b) mutation of a parent is performed to produce new children and a new population.
- Again the fitness function for the new population is estimated.
- The process recurs as long as the fitness function keeps on improving or until the termination condition is attained.

The algorithm of a genetic programming begins with the population which is a set of randomly created individuals. Each individual represents a potential solution which is further represented as a binary tree. Each binary tree is constructed by all the possible compositions of the sets of functions and terminals. A fitness value of each tree is calculated by a suitable fitness function. According to the fitness value, a set of individuals with better fitness will be selected. These individuals are used to generate new population in next generation with genetic operators. Genetic operators generally also include reproduction, crossover, mutation and others that are used to evolve functional expressions.

After the evolution of multiple generations, we can obtain an individual having good fitness value. If the fitness value of such individual still does not satisfy the specified conditions of the solution, the process of evolution will be repeated until the specified conditions are satisfied.

D. PSO FOR FEATURE SELECTION

Particle swarm optimization (PSO) is an evolutionary computation technique the original intent was to graphically simulate the graceful but unpredictable movements of a flock of birds. The original version of PSO was formed from the modified initial simulation. To produce the standard ISO, later she introduced inertia weight into the particle swarm optimizer.A population of random solutions which is also called ‘particles’ was initialized by PSO. In S-dimensional space each particle is treated as a point. The *I*th particle is represented as $a_i=(a_{i1},a_{i2},a_{i3}\ldots a_{in})$. The best previous position (pbest, the position giving the best fitness $b_i=(b_{i1},b_{i2},\ldots.b_{in})$. The symbol ‘gbest’ is used to represent the index of the best particle among all the particles in the population. $E_i=(e_{i1},e_{i2},\ldots.e_{in})$ represents the rate of the position change(velocity) for the particle i. The following equation manipulates the particles value) of any particle is recorded and represented

$$e_{id}=w*e_{id}+c1*rand()*(b_{id}-a_{id})+c2*Rand()*(e_{gd}-a_{id})$$

$$a_{id}=a_{id}+e_{id}$$

Where $d = 1,2,\ldots, S$, w is the inertia weight, it denotes a positive linear function of time changing according to the generation iteration. The balance between global and local exploration is provided by suitable selection of inertia weight and it also results in less iteration on average to find a sufficiently optimal solution. The constants c1 and c2 in equation are known as acceleration constants which represent the weighting of the stochastic acceleration terms that pull each particle toward pbest and gbest positions. High values result in target regions, abrupt movement toward, or past while low values allow particles from target regions to roam far before being tugged back. The two random function in the range are rand () and Rand ().On each dimension, particle’s velocities are limited to maximum velocity, Vmax. Maximum velocity determines how large steps through the solution space is allowed to take for each particle. The particles may not explore sufficiently beyond locally good regions if Vmax is too small and it could become trapped in the local optima. Where as if Vmax is too high particles might fly past good solutions.

The first part of equation provides the “flying particles” with a degree of memory capability allowing the exploration of new search space areas. The second part represents the private thinking of the particle itself called as “cognition”. The third part provides the collaboration among the particles and it is called as “social”. PSO is used to calculate the particle’s new velocity according to its previous velocity and the distances of its current position from its own best experience (position) and the group’s best experience. Then according to the equation the particle flies toward a new position.

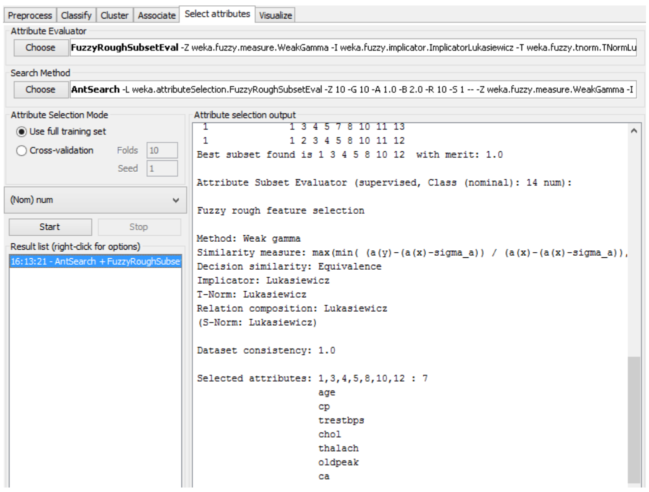
7. EXPERIMENTAL STUDY

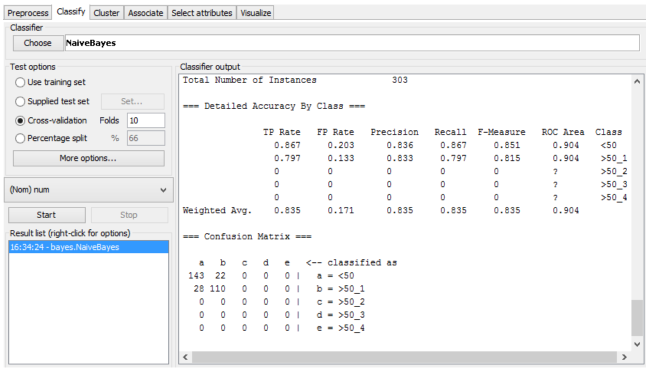
The performance of the reduct calculation approaches discussed in this paper has been tested with different medical datasets obtained from UCI machine learning data repository, [2] to evaluate the performance of proposed algorithm. Weka tool is being used for experimental purpose. Table 2 shows the details of datasets used in this paper.

Table 2. Detail of Data Sets Used for Experiment

Data Set Name	Total Number of Instances	Total Number of Features	Feature Reduction
Cleveland Heart	303	14	7
Lung Cancer	32	57	5

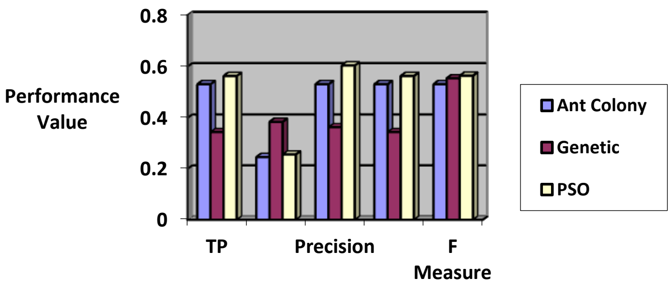
Sample screen shots:





PERFORMANCE ANALYSIS:

In order to obtain the optimal data reductions here, in this paper three types of optimization algorithms such as Genetic algorithm, PSO Algorithm, Ant colony optimization algorithm were used to analyze the performance as follows,



CONCLUSION AND FUTURE WORK

Feature selection is a most valuable preprocessing technique for applications involving huge amount of data. It mainly deals with the problem of selecting minimal attribute set that are most predictive to represent the original attributes in data set. This paper discussed the strengths and weaknesses of various existing feature selection methods. Rough Set Reduct algorithm used as a major preprocessor tool for feature selection. This paper starts with the fundamental concepts of rough set theory and explains basic techniques: Quick Reduct. These methods can produce close to the minimal reduct set. The swarm intelligence methods have been used to guide this method to find the minimal reducts. Here three different computational intelligence based reducts: Genetic algorithm, Ant colony optimization and PSO. Though these methods are performing well, there is no consistency since they are dealing with more random parameters. All these methods are analyzed using medical datasets. Experimental results on different data sets have shown the efficiency of the proposed approach. Comparative performance analysis in which the experimental result shows that feature selection is best for minimal reductions. When compare to other optimization algorithm, PSO algorithm produces higher performance value. As shown in the results, our proposed method exhibits consistent and better performance than the other methods. As an extension of this work the following may be done in future as comparing the results with some other evolutionary algorithm and performing disease prediction.

REFERENCES

[1] H.M. Chen, T.R. Li, D. Ruan, J.H. Lin, and C.X. Hu, "A Rough-Set Based Incremental Approach for Updating Approximations under Dynamic Maintenance Environments," IEEE Trans. Knowledge and Data Eng., vol. 25, no. 2, pp. 274-284, Feb. 2013.

[2] J.Y. Liang, F. Wang, C.Y. Dang, and Y.H. Qian, "An Efficient Rough Feature Selection Algorithm with a Multi-Granulation View," Int'l J. Approximate Reasoning, vol. 53, pp. 912-926, 2012.

[3] J.F. Pang and J.Y. Liang, "Evaluation of the Results of Multi- Attribute Group Decision-Making with Linguistic Information," Omega, vol. 40, pp. 294-301, 2012.

[4] Q.H. Hu, D.R. Yu, W. Pedrycz, and D.G. Chen, "Kernelized Fuzzy Rough Sets and Their Applications," IEEE Trans. Knowledge and Data Eng., vol. 23, no. 11, pp. 1649-1667, Nov. 2011.

[5] N. Parthalain, Q. Shen, and R. Jensen, "A Distance Measure Approach to Exploring the Rough Set Boundary Region for Attribute Reduction," IEEE Trans. Knowledge and Data Eng., vol. 22, no. 3, pp. 305-317, Mar. 2010.

Cite this article as: T.Keerthika, Dr. K. Premalatha. "Utilization of Rough Set Reduct Algorithm and Evolutionary Techniques for Medical Domain using Feature Selection." *International Conference on Information Engineering, Management and Security (2015)*: 144-151. Print.

[6] Y.H. Qian, J.Y. Liang, W. Pedrycz, and C.Y. Dang, "Positive Approximation: An Accelerator for Attribute Reduction in Rough Set Theory," Artificial Intelligence, vol. 174, pp. 597-618, 2010.

[7] W. Wei, J.Y. Liang, Y.H. Qian, F. Wang, and C.Y. Dang, "Comparative Study of Decision Performance of Decision Tables Induced by Attribute Reductions," Int'l J. General Systems, vol, 39, no. 8, pp. 813-838, 2010.

[8] S.Y. Zhao, E.C.C. Tsang, D.G Chen, and X.Z. Wang, "Building a Rule-Based Classifier-a Fuzzy-Rough Set Approach," IEEE Trans. Knowledge and Data Eng., vol. 22, no. 5, pp. 624-638, May 2010.

[9] M. Kryszkiewicz and P. Lasek, "FUN: Fast Discovery of Minimal Sets of Attributes Functionally Determining a Decision Attribute," Trans. Rough Sets, vol. 9, pp. 76-95, 2008.

[10] M.Z. Li, B. Yu, O. Rana, and Z.D. Wang, "Grid Service Discovery with Rough Sets," IEEE Trans. Knowledge and Data Eng., vol. 20, no. 6, pp. 851-862, June 2008.

[11] N. Suguna and Dr. K. Thanushkodi," A Novel Rough Set Reduct Algorithm for Medical Domain Based on Bee Colony Optimization," Journal of computing , Volume 2, Issue 6, June 2010, ISSN 2151-9617.

[12] Pawlak, Z. (1982) 'Rough Sets', International Journal of Com- puter and Information Sciences, Vol. 11, pp. 341–356.

[13] Pawlak, Z. (1991) Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers.

[14] Pawlak, Z. (1993) 'Rough Sets: Present State and The Future', Foundations of Computing and Decision Sciences, Vol. 18, pp. 157–166. Han, J., and Kamber, M.: Data mining concepts and techniques, Academic Press, 2001.

Cite this article as: T.Keerthika, Dr. K. Premalatha. "Utilization of Rough Set Reduct Algorithm and Evolutionary Techniques for Medical Domain using Feature Selection." *International Conference on Information Engineering, Management and Security (2015)*: 144-151. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS025		**eAID**	ICIEMS.2015.025

An Optimize Utilization of Carrier Channels for Secure Data Transmission, Retrieval and Storage in Distributed Cloud Network using Key Management with Genetic Algorithm: A Review

Dr.C.A.Dhote[1], Mr.Virendra P.Nikam[2]
[1]Professor,Badnera College of Engineering,Badnera
[2]Phd Scholar

Abstract: *Relay transmission can enhance coverage and throughput, whereas it can be vulnerable to eavesdropping attacks due to the additional transmission of the source message at the relay. Thus, whether or not one should use relay transmission for secure communication is an interesting and important problem. In this paper, we consider the transmission of a confidential message from a source to a destination in a decentralized wireless network in the presence of randomly distributed eavesdroppers. The source–destination pair can be potentially assisted by randomly distributed relays. For an arbitrary relay, we derive exact expressions of secure connection probability for both colluding and nonpolluting eavesdroppers. We further obtain lower bound expressions on the secure connection probability, which are accurate when the eavesdropper density is small. Using these lower bound expressions, we propose a relay selection strategy to improve the secure connection probability. By analytically comparing the secure connection probability for direct transmission and relay transmission, we address the important problem of whether or not to relay and discuss the condition for relay transmission in terms of the relay density and source–destination distance. These analytical results are accurate in the small eavesdropper density* regime.

Keywords: Attack effect, low-rate distributed denial of service (DDoS) attack, mathematical model, and shrew attack.

INTRODUCTION

The power grid has become a necessity in the modern society. Without a stable and reliable power grid, tens of millions of people's daily life will be degraded dramatically [1] . For instance, the India blackout in July 2012 affected more than 60 million people (about 9% of the world population) and plunged 20 of Indian 28 states into darkness [2]. Indeed, the traditional power grid, which is surprisingly still grounded on the design more than 100 years ago, can no longer be suitable for today's society [3] With the development of information system and communication technology, Many countries have been modernizing the aging power system into smart grid, which is featured with two-way transmission, high reliability, real-time demand response, self-healing, and security. Within smart grid, Advanced Metering Infrastructure (AMI) plays a vital role and is associated with people's daily life most closely [4] . AMI modernizes the electricity metering system by replacing old mechanical meters with smart meters, which provide two-way communications between utility companies and energy customers. With the AMI, people can not only read the meter data remotely, but also do some customized control and implement fine-coarse demand 106 Tsinghua Science and Technology, April 2014, 19(2):

Cite this article as: Dr.C.A.Dhote, Mr.Virendra P.Nikam. "An Optimize Utilization of Carrier Channels for Secure Data Transmission, Retrieval and Storage in Distributed Cloud Network using Key Management with Genetic Algorithm: A Review." *International Conference on Information Engineering, Management and Security (2015)*: 152-158. Print.

105-120 Response [5]. In addition, the real-time data collected from the smart meters can improve the reliability of the distribution grid by avoiding line congestion and generation overloads [6].

The utility companies can also provide faster diagnosis of outage and dynamical electricity price thanks to the AMI. Hence, AMI has attracted great attention from many stakeholders, including utility companies, energy markets, regulators,etc. AMI technologies are rapidly overtaking the traditional meter reading technologies and millions of smart meters are equipped in the household all over the world. For example, there are already more than 4.7million smart meters used for billing and other purposes in Ontario, Canada[7] According to the American Institute for Electric Efficiency (IEE), approximately 36 million smart meters have been installed in the United State by May 2012, and additional 30 million smart meters will be deployed in the next three years [8]However, rich information exchange and hierarchical semi-open network structure in AMI extend the attack surface for metering to entire public networks and introduce many vulnerabilities for cyber attacks [9, 10] . Among all the attacks to the AMI, energy theft in emerging economies has been a widespread practice, both in developing countries and developed countries. A World Bank report finds that up to 50% of electricity in developing countries is acquired via theft [11] . It is reported that each year over 6 billion dollars are lost due to the energy theft in the United States alone [12]. In 2009, the FBI reported a wide and organized energy-theft attempt that may have cost up to 400 million dollars annually to a utility following an AMI deployment [13] . In Canada, BC Hydro reports $100 million in losses every year [14] . Utility companies in India and Brazil incur losses around $4.5 billion and $5 billion due to electricity theft, respectively [15, 16]. There is even a video which shows how to crack the meter and cut the electricity bill in half in Youtube [17] . As a result, energy-theft issue becomes one of the most important concerns which prohibit the development of AMI. Due to the nature of non-technical loss during transmission of electrical energy, it is very difficult for the utility companies to detect and fight the people responsible for energy theft. The unique challenges for energy theft in AMI call for the development of effective detection techniques. However, so far, few studies have elaborated what have been achieved and what should be done for these challenges. As a result,
we are motivated to investigate energy-theft issue in AMI, which is of critical importance to the design of AMI information networks and has been considered as one of the highest priorities for the smart grid design. In this paper, we provide a state-of-the-art survey of existing energy-theft detection schemes in AMI

LITERATURE SURVEY

Tung-Hsiang Liu and *Long-Wen Chang* [20] has proposed a simple data hiding technique for binary images in *2004*.The proposed method embeds secure data at the edge portion of host binary image. Binary images consist of only two colors therefore changing any pixels in this image could be easily detected by human eyes. Therefore, data is stored in the edge portion of binary image; as the modification of edge pixels is more difficult to be recognized by human eyes. The Distance matrix mechanism is used to find the edge pixels of host binary image. Then the Weight mechanism is used to consider the connectivity of the neighborhood around changeable pixels for choosing the most suitable one. For the security and quality consideration, a random number generator is used to distribute the embedding data into the overall image. This method not only embeds large amounts of data into host binary image but also can maintain image quality In order to improve the capacity of the hidden secret data and to provide am imperceptible stego image quality *H.-C. Wu, N.-I. Wu, C.-S. Tsai* and *M.-S. Hwang* [21] has proposed a novel stenographic method based on Least Significant Bit (LSB) Replacement and Pixel Value Differencing (PVD) methods in *2005*.Pixel Value Differencing (PVD) method is used to discriminate between edge areas and smooth areas of cover image. In Wu and Tsai's steganographic method, a grey-valued cover image is partitioned into non- overlapping blocks of two consecutive pixels, states pi and p_i and p_{i+1}.From each block we can obtain a different value d_i by subtracting pi from $p_{i+1.}$ All possible different values of d_i range from -255 to 255, then $|d_i|$ ranges from 0 to 255. Therefore, the pixel p_i and p_{i+1} are located within the smooth area when the value $|d_i|$ is smaller and will hide less secret data. Otherwise, it is located on the edged area and embeds more data. From the aspect of human vision it has a larger tolerance that embeds more data into edge areas than smooth areas. The secret data is hidden into the smooth areas of cover image by LSB method while using the PVD method in the edge areas. As, this proposed method not only store data in the edge areas but also in the smooth areas; therefore it can hide much larger information and maintains a good visual quality of stego image.
In *2005 M. Carli M.C.Q. Fariasy, E. Drelie Gelascaz, R. Tedesco &A. Neri* [22] has proposed a no-reference video quality metric that blindly estimates the quality of a video. They had used Block based Spread Spectrum embedding method to insert a fragile mark into perceptually important areas of the video frames. They used a set of perceptual features to characterize the perceptual importance of a region that are Motion, Contrast and Color. The mark is extracted from the perceptually important areas of the decoded video on receiver side. Then a quality measure of the video is obtained by computing the degradation of the extracted mark. So, in this way quality of a compressed video is estimated by using simple embedding system on perceptually important areas of the video frame.
In *2007 Hsien-Wen Tseng, Feng-Rong Wu*, and *Chi-Pin Hsieh* [23] has proposed a novel method for hiding data in binary images. The binary cover image is partitioned into equal-sized, non-overlapping blocks and the watermark will be embedded into blocks by flipping pixels. For security consideration, the watermark data is firstly permuted into a meaningless bit sequence by using a secret key. The cover image is partitioned into blocks of predefined size n x n and then each block can be embedded one secret bit except the completely black or white blocks. The embedding rule is based on the odd-even information in a block. A Weight mechanism is used to select the most suitable pixel for flipping. Additionally boundary check is performed to improve the visual quality of stego image as well as to prevent boundary distortion. This method achieved a good visual quality for watermarked image and has high capacity of embedding.

In *2008 Beenish Mehboob* and *Rashid Aziz Faruqui* [24] discussed the art and science of Steganography in general and proposed a novel technique to hide data in a colorful image using least significant bit. Least Significant Bit or its variants are used to hide data in digital image. Digital Images are represented in bits. The idea of playing with 0's and 1's seem quite simple but a slight change in value may transform an image completely ,in other words it distorts image completely. Therefore this technique chops the data in 8 bits after the header and used LSB to hide data. So, they proved LSB method is the most recommended for hiding data than other techniques which require masking and filtering.

M.B. Ould Medeniand &El Mamoun Souidi [25] has proposed a novel stenographic method for gray level images on four pixel differencing and LSB substitution in *2010*. The proposed approach works by dividing the cover into blocks of equal sizes and split each pixel into two parts .Then it counts number of one's in most part and embeds the secret message in the least part according to the corresponding number of bits in most part. As shown in following fig. 2.1

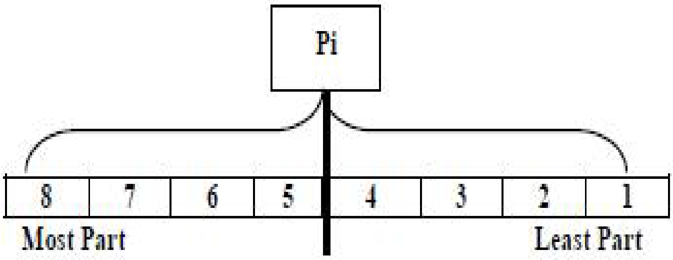


Figure 2.1: Split Process

$number-1-(MSB)$	$number-to-embeded$
$4or3$	$3bits$
2	$2bits$
$1or0$	1

TABLE: 2.1 Number of 1 and the Corresponding Number of Bits to Embed

Therefore, it embeds the message in the edge of the block depending on the number of ones in left four bits of the pixel. They used K-bit LSB substitution method for hiding the secret data into each pixel where K is decided by the number of one in the most part of pixel. This method gave best values for the PSNR measure which means that there were no big difference between the original and the stegno image.

In *2012 Tasnuva Mahjabin,Syed Monowar Hossain* and *Md. Shariful Haque* [26] has proposed a data hiding method based on PVD and LSB substitution to improve the capacity of the secret data as well as to make stegnalysis a complicated task they made an effort to implement a robust dynamic method of data hiding. An efficient and dynamic embedding algorithm was proposed here that not only hides secret data with an imperceptible visual quality and increased capacity but also make secret code breaking a good annoyance for the attacker. This method achieved an increased embedding capacity and lower image degradation with improved security as compared to LSB substitution method and some other existing methods of data hiding. This system used a dynamic method of image data hiding based on LSB Substitution method and Pixel Value Differencing method. The whole process of selecting eight pixels block for a sixteen pixels region and the embedding method for each eight pixels block is different for different cover images. That is, depending on the quality of the cover image the embedding procedure takes this decision in run time. This feature of this method provides security of the hidden secret data. In order to extract the secret data it is mandatory to know that the cover image is divided into regions of sixteen pixels and also the type of eight pixels block for these regions and type of method for each of these blocks. Moreover, if any one becomes aware of the techniques that have been used to insert data in one image, he cannot use the same technique to other images. For example, depending on the quality of the cover image the embedding technique can select horizontal block for inserting data in the first sixteen pixels region for one image whereas vertical eight pixels block for the other image. Thus the decision for steganalysis becomes difficult and this method becomes a secure one.

Ankit Chaudhary and *JaJdeep Vasavada* [27] has proposed an improved stenography approach for hiding text messages in RGB lossless images in *2012*.The security level is increased by randomly distributing the text message over the entire image *Ankit Chaudhary* and *JaJdeep Vasavada* [27] has proposed an improved stenography approach for hiding text messages in RGB lossless images in *2012*.The security level is increased by randomly distributing the text message over the entire image instead of clustering within specific image portions. The first step towards the random distribution of the message in image is using indicator values. They used MSB bits of Red, Green and Blue channel as pixel indicator values instead of utilizing an entire channel. The MSBs indicate in what sequence the message is hidden using the LSBs. In addition to this, this scheme is applied after applying compression to the original message; therefore it would be make it extremely difficult to break, even after suspicion of the message within an image. The scheme works as follows: The MSB remains unchanged when an LSB of a byte is utilized for storing a message. This scheme enables us to fully utilize all the LSBs of every channel of the cover image to store the hidden message and hence improve its capacity. Moreover the varying indicator values introduce a security aspect as it becomes increasingly difficult to decode the message even if its presence is suspected. They increased storage capacity by utilizing all the color channels for storing information and providing the source text message compression. The degradation of the images can be minimized by changing only one lease significant bitper color channel for hiding the message, incurring avery little change in the original image. So, this method increased the security level and improved the storage capacity while incurring minimal quality degradation.

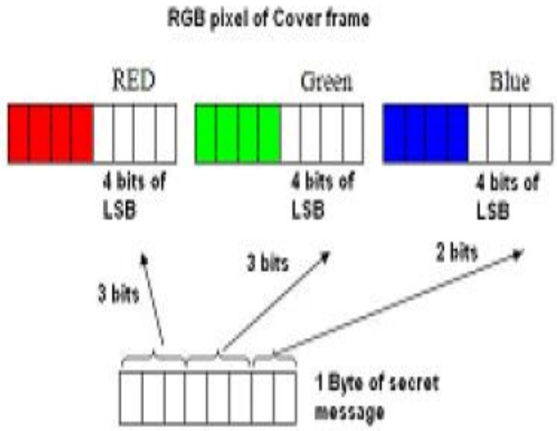


Figure: 2.2 shows secret data embedded in 4 bits of LSB in 3, 3, 2 order in corresponding RGB pixels of carrier frame

A hash function is used to select the position of insertion in LSB bits. The proposed technique takes eight bits of secret data at a time and conceal them in LSB of RGB (Red, Green and Blue) pixel value of the carrier frames in 3, 3, 2 order respectively. Such that out of eight (08) bits of message six (06) bits are inserted in R and G pixel and remaining two (02) bits are inserted in B pixel. After comparing the proposed technique with LSB technique it is found that the performance analysis of proposed technique is quite encouraging. The advantage of this method is that the size of the message does not matter in video stenography as the message can be embedded in multiple frames.

In *2012 Poonam V Bodhak* and *Baisa L Gunjal* [29] has proposed a method to hide data containing text in computer video file and to retrieve the hidden information. This can be designed by embedding the text file in a video file in such a way that the video does not lose its functionality using DCT & LSB Modification method. LSB is the lowest bit in a series of numbers in binary. The LSB based Steganography is one of the steganographic methods, used to embed the secret data in to the least significant bits of the pixel values in a cover image. DCT coefficients are used for JPEG compression. It separates the image into parts of differing importance. It transforms a signal or image from the spatial domain to the frequency domain. It can separate the image into high, middle and low frequency components.This method applies imperceptible modification. This proposed method strives for high security to an eavesdropper's inability to detect hidden information.
RigDas and *Themrichon Tuithung* [30] have proposed novel technique for image stenography based on Huffman Encoding in *2012*. Huffman Encoding is performed over the secret image/message before embedding and each bit of Huffman code of secret Image/message is embedded inside the cover image by altering the least significant bit (LSB) of each of the pixel's intensities of cover image. This paper presents a novel technique for image steganography based on Huffman Encoding. Two 8 bit gray level image of size M X N and P X Q are used as cover image and secret image respectively. As shown in fig: 2.3

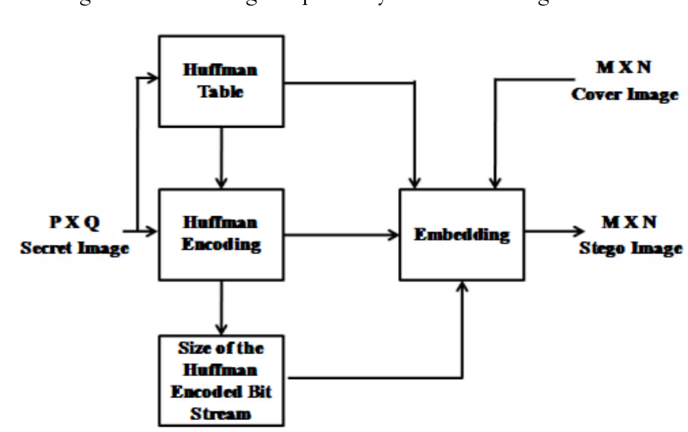


Figure: 2.3 Insertion of the Secret Image/Message into a Cover Image

Cite this article as: Dr.C.A.Dhote, Mr.Virendra P.Nikam. "An Optimize Utilization of Carrier Channels for Secure Data Transmission, Retrieval and Storage in Distributed Cloud Network using Key Management with Genetic Algorithm: A Review." *International Conference on Information Engineering, Management and Security (2015)*: 152-158. Print.

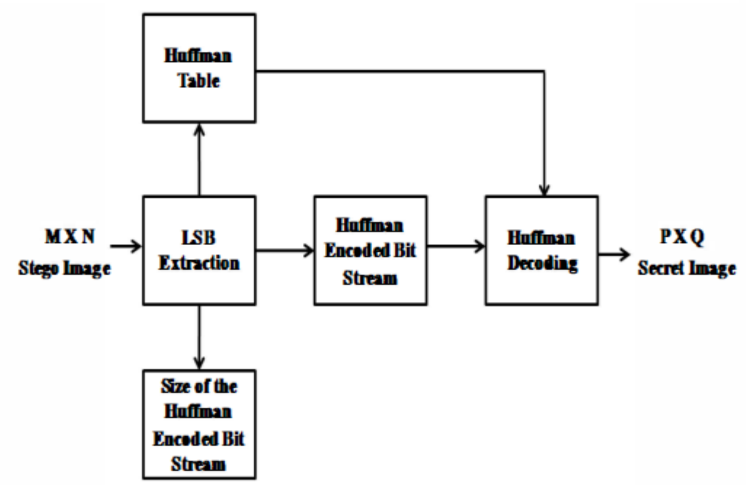


Figure: 2.4 Extraction of the Secret Image from the Stego Image

Huffman Encoding is performed over the secret image/message before embedding and each bit of Huffman code of secret image/message is embedded inside the cover image by altering the least significant bit (LSB) of each of the pixel's intensities of cover image. The size of the Huffman encoded bit stream and Huffman Table are also embedded inside the cover image, so that the Stego-Image becomes standalone information to the receiver.

In *2013 Ming Li,Michel K. Kulhandjian, Dimitris,A. Pados,,Stella N. Batalama,* and *Michael J. Medley* [31] has considered the problem of extracting blindly data embedded over a wide band in a spectrum (transform) domain of a digital medium (image, audio, video).We develop a novel multicarrier/signature iterative generalized least-squares (M-IGLS) core procedure to seek unknown data hidden in hosts via multicarrier spread-spectrum embedding. Neither the original host nor the embedding carriers are assumed available.

SUMMARY & DISCUSSION

Year	Author	Advantages
2004	Tung-Hsiang Liu Long-Wen Chang	Large amount of data can be stored in binary images as well as quality of an image is maintained.
2005	H.-C. Wu, N.-I. Wu, C.-S. Tsai M.-S. Hwang	Much larger information can be stored in images by using LSB method for storing data in smooth areas of image.
2005	M. Carli M.C.Q. Fariasy, E. Drelie Gelascaz, R. Tedesco, A. Neri	Quality of a compressed video is estimated by using simple embedding system.
2007	Hsien-Wen Tseng, Feng-Rong Wu,Chi-Pin Hsieh	This method achieved a good visual quality for watermarked image and has high capacity of embedding.

2010	M.B. Ould Medeni El Mamoun Souidi	K-bit LSB substitution method used here gave best values for the PSNR measure.
2012	Tasnuva Mahjabin, Syed Monowar Hossain Md.Shariful Haque	PVD & LSB methods used here which achieved an increased embedding capacity and lower image degradation with improved security.
2012	Ankit Chaudhary JaJdeep Vasavada	1-bit LSB substitution method used which increased the security level and improved the storage capacity
2012	Kousik Dasgupta, J.K.Mandal Paramartha Dutta	It allows embedding the large size of data in multiple frames. Therefore size of the message does not matter.
2012	Poonam V Bodhak Baisa L Gunjal	DCT & LSB methods used which provide high security to embedded data.
2012	RigDas ThemrichonTuithung	Huffman Encoding is used for secret message which again improves the security level of hiding data.
2013	Ming Li, Michel K. Kulhandjian, Dimitris, A. Pados, Stella N. Batalama, Michael J. Medley	M-IGLS procedure is used for extracting blindly data embedded over a wide band in a spectrum domain of a digital medium.

REFERENCES

[1] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding: A survey," *Proc. IEEE, Special Issue on Identification and Protection of Multimedia Information*, vol. 87, no. 7, pp. 1062–1078, Jul.1999.

[2] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*. San Francisco, CA, USA: Morgan-Kaufmann, 2002.

[3] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proc. IEEE, Special Issue on Identification and Protection of Multimedia Information*, vol. 87, pp. 1079–1107, Jul. 1999

[4] G. C.Langelaar, I. Setyawan, and R. L. Lagendijk, "Watermarking digital image and video data: A state-of-the-art overview," *IEEE Signal Process. Mag.*, vol. 17, no. 5, pp. 20–46, Sep. 2000.

[5] N. F. Johnson and S. Katzenbeisser, , S. Katzenbeisser and F. Petitcolas, Eds., "A survey of steganographic techniques," in *Information Hiding*. Norwood, MA, USA: Artech House, 2000, pp. 43–78.

[6] S. Wang and H. Wang, "Cyber warfare: Steganography vs. steganalysis," *Commun. ACM*, vol. 47, pp. 76–82, Oct. 2004.

[7] C. Cachin, "An information-theoretic model for steganography," in *Proc. 2nd Int. Workshop on Information Hiding*,Portland, OR, USA, Apr. 1998, pp. 306–318.

[8] G. J. Simmons, "The prisoner's problem and the subliminal channel," in *Advances in Cryptology: Proc. CRYPTO'83*, New York, NY, USA, 1984, pp. 51–67, Plenum.

[9] J. Fridrich, Steganography in Digital Media, Principles, Algorithms, and Applications. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[10] Y. Wang and P. Moulin, "Perfectly secure steganography: Capacity, error exponents, and code constructions," IEEE *Trans. Inf. Theory*, vol. 54, no. 6, pp. 2706–2722, Jun. 2008.

Cite this article as: Dr.C.A.Dhote, Mr.Virendra P.Nikam. "An Optimize Utilization of Carrier Channels for Secure Data Transmission, Retrieval and Storage in Distributed Cloud Network using Key Management with Genetic Algorithm: A Review." *International Conference on Information Engineering, Management and Security (2015)*: 152-158. Print.

[11] Federal Plan for Cyber Security and Information Assurance Research and Development Interagency Working Group on Cyber Security and Information Assurance, Apr. 2006.

[12] [H. S.Malvar and D. A. Florencio, "Improved spread spectrum: A new modulation technique for robust watermarking," *IEEE Trans. Signal Proc.*, vol. 51, no. 4, pp. 898–905, Apr. 2003.

[13] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shannon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.

[14] J. Hernandez, M. Amado, and F. Perez-Gonzalez, "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE Trans. Image Process.*, vol. 9, no. 1, pp.55–68, Jan. 2000

[15] C. Qiang and T. S. Huang, "An additive approach to transform-domain information hiding and optimum detection structure," *IEEE Trans. Multimedia*, vol. 3, no. 3, pp. 273–284, Sep. 2001.

[16] C. Fei, D. Kundur, and R. H. Kwong, "Analysis and design of watermarking algorithms for improved resistance to compression," *IEEE Trans. Image Process.*, vol. 13, no. 2, pp. 126–144, Feb. 2004.

[17] M. Gkizeli, D. A. Pados, and M. J. Medley, "SINR, bit error rate,and Shannon capacity optimized spread-spectrum steganography," in *Proc. IEEE Int. Conf. Image Proce. (ICIP)*, Singapore, Oct. 2004, pp. 1561–1564.

[18] M. Gkizeli, D. A. Pados, S. N. Batalama, andM. J.Medley, "Blind iterativerecovery of spread-spectrum steganographic messages," in *Proc. IEEE Int. Conf. Image Proc. (ICIP)*, Genova, Italy, Sep. 2005, vol. 2,pp. 11–14.

[19] M. Gkizeli, D. A. Pados, and M. J. Medley, "Optimal signature design for spread-spectrum steganography," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 391–405, Feb. 2007.

[20] Tung-Hsiang Liu and Long-Wen Chang, "An Adaptive Data Hiding Technique for Binary Images", *Proc.IEEE 17th Int.Conf. On Pattern Recognition (ICPR'04)* 2004.

[21] H.-C. Wu, N.-I. Wu, C.-S. Tsai and M.-S. Hwang," Image steganographic scheme based on pixel-value differencing and LSB replacement methods", *IEE Proc.-Vis. Image Signal Process.*, Vol. 152, No. 5, October 2005.

[22] M. Carli , M.C.Q. Fariasy, E. Drelie Gelascaz, R. Tedesco, A. Neri, "QUALITY ASSESSMENT USING DATA HIDING ON PERCEPTUALLY IMPORTANT" *IEEE* AREAS0-7803-9134-9/05/$20.00 ©2005.

[23] Hsien-Wen Tseng, Feng-Rong Wu,and Chi-Pin Hsieh," Data Hiding for Binary Images Using Weight Mechanism",*IEEE* 2007.

[24] Beenish Mehboob and Rashid Aziz Faruqui," A StegnographyImplementation", *IEEE* 2008

[25] M.B. Ould MEDENI, El Mamoun SOUIDI," A Novel Steganographic Method for Gray-Level Images With four-pixel Differencing and LSB Substitution "*IEEE* 2010

[26] Tasnuva Mahjabin, Syed Monowar Hossain, Md. Shariful Haque," A Block Based Data Hiding Method in Images Using Pixel Value Differencing and LSB Substitution Method", *IEEE* 2012.

[27] Ankit Chaudhary, JaJdeep Vasavada,"A Hash Based Approach for Secure Keyless Image Steganography in Lossless RGB Images", *IEEE* 2012.

Cite this article as: Dr.C.A.Dhote, Mr.Virendra P.Nikam. "An Optimize Utilization of Carrier Channels for Secure Data Transmission, Retrieval and Storage in Distributed Cloud Network using Key Management with Genetic Algorithm: A Review." *International Conference on Information Engineering, Management and Security (2015)*: 152-158. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	VOL	01
Website	www.iciems.in	eMail	iciems@asdf.res.in
Received	10 - July - 2015	Accepted	31- July - 2015
Article ID	ICIEMS026	eAID	ICIEMS.2015.026

Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment

Dhanalakshmi G[1], Rajeswari R[1], Vignesh G[2]
[1]G.C.T., Coimbatore
[2]BHEL, Tiruchirapalli

ABSTRACT: *This paper proposes an optimal phasor measurement unit (PMU) placement model considering power system controlled islanding so that the power network remains observable under controlled islanding condition as well as normal operation condition. The optimization objectives of proposed model are to minimize the number of installed PMUs and to maximize the measurement redundancy. These two objectives are combined together with a weighting variable so that the optimal solution with minimum PMU number and maximum measurement redundancy would be obtained from the model.At last, IEEE-14 bus standard systems and the Tamil Nadu state power grid (83 bus system) are employed to test the presented model. Results are presented to demonstrate the effectiveness of the proposed method.*

Keywords: Controlled islanding, integer linear programming, measurement redundancy, optimal phasor measurement unit (PMU) placement.

I INTRODUCTION

Synchronized phasor measurement unit (PMU) is essentially a digital recorder with synchronized capability. It can be a stand-alone physical unit or a functional unit within another protective device. By measuring the magnitude and phase angles of currents and voltages a single PMU can provide real-time information about power system events in its area, and multiple PMU can enable coordinated system-wide measurements. PMU also can time-stamp, record, and store the phasor measurements of power system events. This capability has made PMU become the foundation of various kinds of wide area protection and control schemes.
A lot of PMU potential applications in power system monitoring, protection, and control have been studied since it was introduced in mid-1980s. Specially, in recent years, PMUs have been and extensively used or proposed to be used in many applications in the area of power system protection and control with the cost reduction of PMUs and improvement of communication technologies in power system [11]. Synchrophasors are precise measurements of the power systems and are obtained from PMUs. PMUs measure voltage, current, and frequency in terms of magnitude and phasor angle at a very high speed (usually 30 measurements per second). Each phasor measurement recorded by PMU devices is time-stamped based on universal standard time, such that phasors measured by different PMUs installed in different locations can be synchronized by aligning time stamps. However, PMU and its associated communication facilities are costly. Furthermore, the voltage phasor of the bus incident to the bus with PMU installed can be computed with branch parameter and branch current phasor measurement [5]. So it is neither economical nor necessary to install PMUs at all system buses. Thus, one of the important issues is to find the optimal number and placement of PMUs. Optimal PMU

Cite this article as: Dhanalakshmi G, Rajeswari R, Vignesh G. "Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment." *International Conference on Information Engineering, Management and Security (2015)*: 159-167. Print.

placement (OPP) is firstly attempted in [6], formulating as a combinatorial optimization problem of minimizing the PMU number for system observability. In [7], an integer programming formulation of OPP problem is proposed with the presence of conventional measurements. A generalized integer linear programming (ILP) formulation for OPP is presented in [8]. Generally, the existing OPP models concerns about the determination of minimum number and optimal location set of PMUs, ensuring thattheentirepowersystemremainsasingleobservableisland [1]. In another word, these models can only handle the cases in which the power system is operated as a single and integrated network. However, some severe faults may lead parts of the network to angle, frequency or voltage instability. In that case, trying to maintain system integrity and operate the system entirely interconnected is very difficult and may cause propagation of local weaknesses to other parts of the system [11]. As a solution, controlled islanding (CI) is employed by system operators, in which the interconnected power system is separated into several planned islands prior to catastrophic events [12], [13]. After system splitting, wide area blackout can be avoided because the local instability is isolated and prevented from further spreading [14]. In order to operate each island with power balancing and stability after controlled islanding, it is essential to provide an OPP scheme which can keep the network observable for the post-islanding condition as well as normal condition.

In this paper, an ILP model of OPP considering controlled islanding (OPP-CI) is proposed. This model is able to determine the minimal number and optimal location set of PMUs in order to provide the full network observability in normal operation as well as in controlled islanding scenario. To distinguish multiple optimal solutions, measurement redundancy is incorporated into the optimization objective. The performance of the proposed new model is assessed using several IEEE-14 bubs standard systems and a Tamil Nadu state power grid system.

II INTEGER LINEAR PROGRAMMING

Integer Linear Programming(ILP) is a mathematical optimization method for getting an optimal outcome from a given mathematical objective function, subject to some linear inequality constraints. In this thesis ILP is used for finding the minimum set of PMUs for a given power grid to achieve its complete observability. The objective of the PMU placement problem is that a bus will be reached by at least one

PMU. The detailed description of ILP was reported in Refs. Two assumptions are made before applying ILP for PMU placement. First, there is no constraint on the number measuring channels for the PMU, i.e., a PMU can measure the current phasors from any number of branches that are connected to it. Second, there are no

problems with the availability of the communication system. i.e., all buses are well equipped with communication facilities for the transfer of data from PMUs .

The Program for objective function and constraints of IEEE-14 Bus test systems are mentioned below

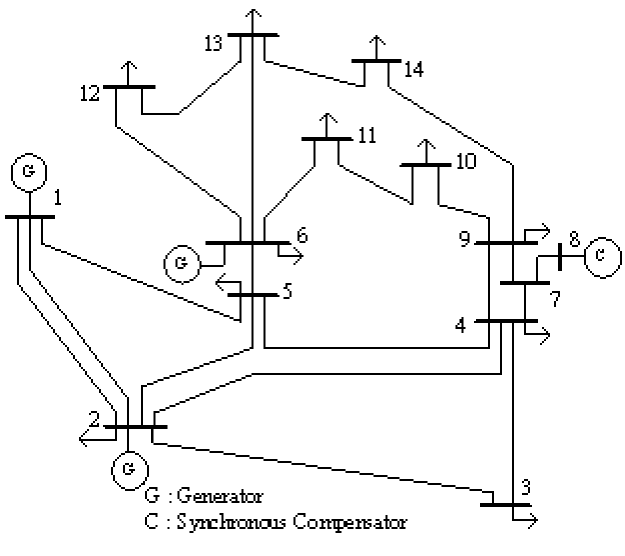


Figure 1 IEEE14 bus system

. The first stage of ILP for complete observability is to create a matrix U_{PMU}, in such a way that, its entries are defined as

$$U_{i,j} = 1, \text{ (if i and j are connected)} \tag{1}$$

In a power system network, the PMU placement at a bus can be seen as a binary decision variable defined as

$$u_i = \begin{cases} 1 & \text{if PMU is placed at bus } i \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

For a system with buses, therefore, the optimal PMU placement problem can be formulated as an integer linear programming problem as follows:

Cite this article as: Dhanalakshmi G, Rajeswari R, Vignesh G. "Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment." *International Conference on Information Engineering, Management and Security (2015)*: 159-167. Print.

$$\min F_1 = \sum_{i=1}^{n} c_i u_i \tag{3}$$

subject to constraints

$$f_i = \sum_{j=1}^{n} a_{i,j} u_j \geq 1 \qquad i = 1, 2, \ldots, n \tag{4}$$

where

- C_i is the cost of installing a PMU at bus . Without loss of generality, cost of PMU installation at each bus is assumed to be equal to 1 per unit in the present study.
- f_i refers to the number of times that the th bus is observed through PMU measurements.
- $a_{i,j}$ is the *i-j* th entry of network connectivity matrix defined as

$$a_{i,j} = \begin{cases} 1 & \text{if } i = j \text{ or if } i \text{ and } j \text{ are connected} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

For example, with (3), minimizing the number of PMUs for the IEEE 14-bus system (Fig. 1) can be formulated as follows:
Constraints function:

```
function [c ceq ]=fourteencons (x)
c(1)=-(u(1)+u(2)+u(5))+1;
c(2)=-(u(1)+u(2)+u(3)+u(4)+u(5))+1;
c(3)=-(u(2)+u(3)+u(4))+1;
c(4)=-(u(2)+u(3)+u(4)+u(5)+u(7)+u(9))+1;
c(5)=-(u(1)+u(2)+u(4)+u(5)+u(6))+1;
c(6)=-(u(5)+u(6)+u(11)+u(12)+u(13))+1;
c(7)=-(u(4)+u(7)+u(8)+u(9))+1;
c(8)=-(u(7)+u(8))+1;
c(9)=-(u(4)+u(7)+u(9)+u(10)+u(14))+1;
c(10)=-(u(9)+u(10)+u(11))+1;
c(11)=-(u(6)+u(10)+u(11))+1;
c(12)=-(u(6)+u(12)+u(13))+1;
c(13)=-(u(6)+u(12)+u(13)+u(14))+1;
c(14)=-(u(9)+u(13)+u(14))+1;
        ceq=[ ];
```

In this thesis, an ILP model of OPP considering controlled islanding (OPP-CI) is proposed. This model is able to determine the minimal number and optimal location set of PMUs inorder to provide the full network observability in normal operation as well as in controlled islanding scenario.
Compared to (4), the observability constraints of OPP-CI model are modified as follows:

$$f_i = \sum_{j=1}^{n} a_{i,j}^{CI} u_j \geq 1 i = 1, 2, \ldots, n \tag{6}$$

where $a_{i,j}^{CI}$ is the binary entry in the connectivity matrix for post-islanding network, which is defined as

$$a_{i,j}^{CI} = \begin{cases} 0 & \text{if line } i - j \text{ is opened in } CI \text{ process} \\ a_{i,j} & \text{otherwise.} \end{cases} \tag{7}$$

III CONCEPTS OF ISLANDING AND REDUNDANCY MEASUREMENT

Cascading failures are the most significant threats for power system security. Cascading failures together with additional line tripping can lead the system to uncontrolled splitting [11]. Formation of uncontrolled islands with significant power imbalance is the main reason for system blackouts. In order to avoid catastrophic
wide area blackouts due to cascading failures, controlled islanding has been considered as an effective defense strategy. The main advantages of controlled islanding of power systems can be listed as follows [11]:

- It can separate weak and vulnerable areas from other stable parts of the system.
- Compared to the whole system, small subsystems are easier to be handled and controlled under dynamic and emergency conditions.

Cite this article as: Dhanalakshmi G, Rajeswari R, Vignesh G. "Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment." *International Conference on Information Engineering, Management and Security (2015)*: 159-167. Print.

After establishment of planned islands, there exist some factors which may threat the stability and integrity of each island, such as power imbalance, line overloading, voltage, angle and frequency instabilities, etc. [11]. Therefore, to maintain static and dynamic stability, necessary load shedding and other control actions may be needed in each island, which always require real-time information throughout the island. In addition, real-time measurements in different islands should be collected and analyzed together to determine whether and how the power system can be restored to normal operation. To ensure the effectiveness of all the above actions, it is essential to keep each island totally observable through properly placed PMUs. In other words, the optimal placement of PMUs should be carried out in such a manner that the network remains observable under controlled islanding condition as well as normal operation condition.

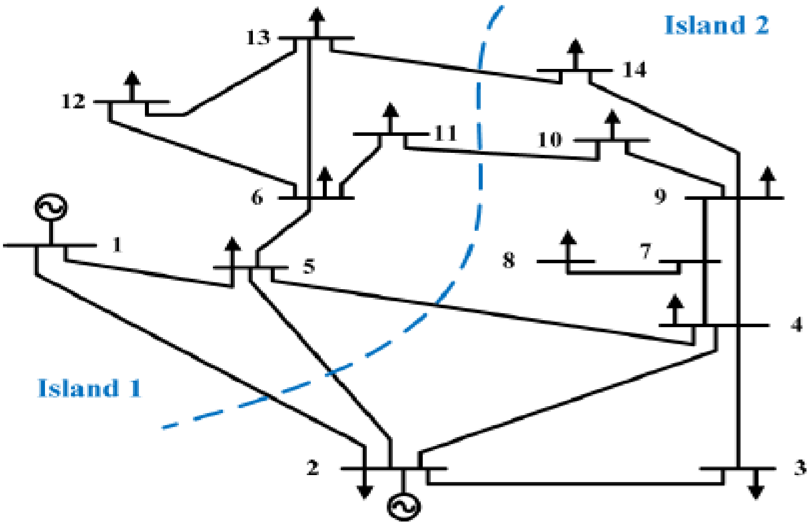


Figure 2 IEEE-14 Bus system under CI condition

For example, with (6), minimizing the number of PMUs for the IEEE 14-bus system (Fig. 2) can be formulated as follows:

$$
\begin{aligned}
f_1 &= u_1 + u_5 \geq 1 \\
f_2 &= u_2 + u_3 + u_4 \geq 1 \\
f_3 &= u_2 + u_3 + u_4 \geq 1 \\
f_4 &= u_2 + u_3 + u_4 + u_7 + u_9 \geq 1 \\
f_5 &= u_1 + u_5 + u_6 \geq 1 \\
f_6 &= u_5 + u_6 + u_{11} + u_{12} + u_{13} \geq 1 \\
f_7 &= u_4 + u_7 + u_8 + u_9 \geq 1 \\
f_8 &= u_7 + u_8 \geq 1 \\
f_9 &= u_4 + u_7 + u_9 + u_{10} + u_{14} \geq 1 \\
f_{10} &= u_9 + u_{10} \geq 1 \\
f_{11} &= u_6 + u_{11} \geq 1 \\
f_{12} &= u_6 + u_{12} + u_{13} \geq 1 \\
f_{13} &= u_6 + u_{12} + u_{13} \geq 1 \\
f_{14} &= u_9 + u_{14} \geq 1.
\end{aligned} \tag{8}
$$

In this thesis, thus, maximizing the measurement redundancy is considered as an additional objective to pick out the most suitable OPP scheme for power systems. Conventionally, measurement redundancy is defined as the ratio of the number of measurements (including direct measurements and indirect measurements) to the number of states [7]. Considering that the most important state variables in state estimation are bus voltage phasors, the measurement redundancy can be redefined as the ratio of the number of voltage measurements to the number of system buses. Moreover, the measurement redundancy under islanding operation scenario as well as normal operation should be considered.

To keep consistency with (3) which is a minimization problem, the objective function of maximizing measurement redundancy is formulated as a minimization problem as well:

$$
\min F_2 = \frac{1}{n} \sum_{i=1}^{n} \left[\omega_1 (m_i^N - t_i^N) + (1 - \omega_1)(m_i^I - t_i^I) \right] \tag{9}
$$

Cite this article as: Dhanalakshmi G, Rajeswari R, Vignesh G. "Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment." *International Conference on Information Engineering, Management and Security (2015)*: 159-167. Print.

where *n* is the total number of system buses; constant m_i^N is the maximum number of times that the i^{th} bus can be observed in normal operation, which equals to the number of its incident lines plus one; variable t_i^N represents the number of times that the i^{th} bus is observed by the solved OPP scheme in normal operation; m_i^N and t_i^N refer to the corresponding constant and variable in islanding operation condition, respectively ω_1 and (1- ω_1); and are weighting factors assigned to the two components of the objective function. Since there is greater probability for a power system to be operated in normal condition than in islanding condition, in this study ω_1 and (1- ω_1) are set at 0.7 and 0.3, respectively.

IV RESULTS AND DISCUSSION

TAMIL NADU STATE POWER GRID

The Tamil Nadu state Indian power grid consists of 83 buses of UHV, EHV and HV which are interconnected by 126 branches. The single line diagram of the power grid is depicted in figure 3 and their bus details are given in Table I. The ILP described in above has been applied to the grid for finding the optimal locations of the PMUs for the complete observability

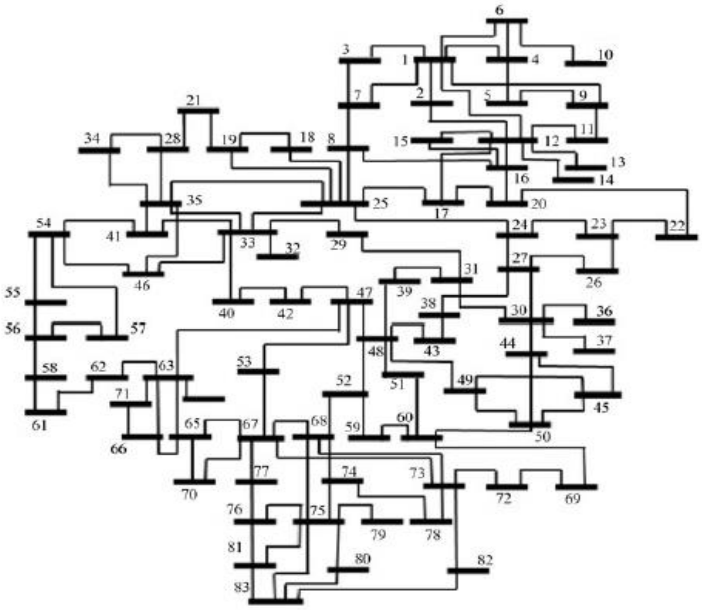


Figure 3 Single line diagram of TN State Indian Power Grid

Table.I Bus details for TN State Indian Power Grid

Bus No.	Bus Name	Bus No.	Bus Name	Bus No.	Bus Name	Bus No.	Bus Name	Bus No.	Bus Name
1	Chennai	18	V. Mangalam	35	Mettur TPS	52	Alagarkoil	69	Paramkudi
2	Gummidipoondi	19	Hosur	36	Bahoor	53	RGPuram	70	Theni
3	Almathy	20	Kalpakkam	37	Villianur	54	Arasur	71	Kadamparai
4	Ennore	21	Karimangalam	38	Eachengadu	55	Pykara	72	Sipcot
5	Monali	22	Acharapakkam	39	Peranbalur	56	Kundah3	73	Tuticorin
6	Tondiarpet	23	Villupuram	40	Unjanai	57	Kundah2	74	Sathur
7	Mosur	24	TV malai	41	Gobi	58	Kundah1	75	Kayathur
8	Thiruvalam	25	Singarapettai	42	P. Chandai	59	Valthur	76	Viranam
9	Korattur	26	Cuddalore	43	Samaypuram	60	Karaikudi	77	Kodikurchi

Cite this article as: Dhanalakshmi G, Rajeswari R, Vignesh G. "Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment." *International Conference on Information Engineering, Management and Security (2015)*: 159-167. Print.

10	Mylapore	27	Neyveli TS1	44	Kdalangudu	61	Thudiyalur	78	Sterlite
11	Koyambedu	28	Mettur	45	Nallur	62	O.K.Mandabam	79	Auto
12	Budur	29	D. Kurchi	46	Ingur	63	Udumalpet	80	Udayathur
13	Kadperi	30	Neyveli TS2	47	Pugalur	64	Ponnapuram	81	Thirunelveli
14	Hyundai	31	STCMS	48	Trichy	65	Sembatti	82	S.R.Pudur
15	Tharamani	32	SAIL	49	Thanjavur	66	Myvadi	83	Sankaneri
16	SP Koil	33	Salem	50	Thiruvarur	67	Madurai		
17	Arni	34	M. Tunnel	51	Pudukottai	68	Pasumalai		

To establish a planned islands, there exist some factors which may threat the stability and integrity of each island, such as power imbalance, line overloading, voltage, angle and frequency instabilities, etc. Therefore, to maintain static and dynamic stability, necessary load shedding and other control actions may be needed in each island, which always require real-time information throughout the island. In addition, real-time measurements in different islands should be collected and analyzed together to determine whether and how the power system can be restored to normal operation. Hence the generation capacity of all the 83 buses in Tamil Nadu state grid is found as listed in table II

The one line diagram of TN State Power Grid having 83 buses is shown in figure.3. Here the OPP schemes are solved for normal and CI conditions.

TN STATE POWER GRID - NORMAL OPERATION

Figure 3 shows the single line diagram of TN State Power Grid. To determine the OPP for normal operation, entire bus is considered as a single island and their observability constraints are determined.

This observability constraints are solved by using ILP in MATLAB and OPP is determined. The results of solved ILP for TN State Power Grid shows that, the system is made completely observable by placing 20 PMUs in buses **5, 7, 10, 12, 19, 22, 27, 30, 32, 34, 40, 45, 48, 54, 58, 60, 63, 67, 73, 75.**

Table II Bus details for TN State Grid with Generation Capasity

Bus No.	Bus Name	MW	Bus No.	Bus Name	MW	Bus No.	Bus Name	MW	Bus No.	Bus Name	MW	Bus No.	Bus Name	MW
1	Chennai	2026	18	V. Mangalam	18.6	35	Mettur TPS	1440	52	Alagarkoil	10	69	Paramkudi	9
2	Gummidipoondi	130	19	Hosur	8	36	Bahoor	36	53	RGPuram	7.5	70	Theni	19
3	Almathy	22	20	Kalpakkam	470	37	Villianur	24.5	54	Arasur	5	71	Kadamparai	400
4	Ennore	450	21	Karimangalam	10	38	Eachengadu	10	55	Pykara	253.2	72	Sipcot	10
5	Monali	43	22	Acharapakkam	10	39	Peranbalur	22	56	Kundah3	180	73	Tuticorin	1050.5
6	Tondiarpet	1500	23	Villupuram	77.5	40	Unjanai	10	57	Kundah2	175	74	Sathur	10

Cite this article as: Dhanalakshmi G, Rajeswari R, Vignesh G. "Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment." *International Conference on Information Engineering, Management and Security (2015)*: 159-167. Print.

7	Mosur	10	24	TV malai	84.3	41	Gobi	4	58	Kundah1	60	75	Kayathur	130
8	Thiruvalam	5	25	Singarapettai	10	42	P. Chandai	15	59	Valthur	187.2	76	Viranam	1320
9	Korattur	12	26	Cuddalore	71.5	43	Samaypuram	48	60	Karaikudi	19	77	Kodikurchi	10
10	Mylapore	9	27	Neyveli TS1	1020	44	Kdalangudu	101	61	Thudiyalur	7.5	78	Sterlite	160
11	Koyambedu	0.25	28	Mettur	370	45	Nallur	340.5	62	O.K.Mandabam	10	79	Auto	3.18
12	Budur	70	29	D. Kurchi	7.5	46	Ingur	6.4	63	Udumalpet	25.7	80	Udayathur	7.5
13	Kadperi	400	30	Neyveli TS2	1750	47	Pugalur	47.12	64	Ponnapuram	182.5	81	Thirunelveli	96.7
14	Hyundai	200	31	STCMS	270	48	Trichy	19.5	65	Sembatti	7.5	82	S.R.Pudur	10
15	Tharamani	12	32	SAIL	60	49	Thanjavur	152.66	66	Myvadi	5.625	83	Sankaneri	1.25
16	SP Koil	50	33	Salem	14	50	Thiruvarur	109.38	67	Madurai	117.5			
17	Arni	31	34	M. Tunnel	12.5	51	Pudukottai	35.5	68	Pasumalai	10			

6.2.2 TN STATE GRID - CONTROLLED ISLANDING (CI) CONDITIONS

In islanding conditions, the whole system is separated into two subsystems based on measurement of generation and distribution capacity of all the buses. The islanding are done by proper load shedding so as to make the generation capacity and distribution around the island remains equal. Thus 4 different cases of islanding are chosen for TN State Power Grid.

These different cases leads to multiple solution for OPP scheme. Therefore, maximizing the measurement redundancy is considered to pick out the most suitable OPP scheme for power systems. ILP is solved for all the cases to determine the OPP and measurement redundancy is found to chose the most feasible solution. The different cases of islanding are shown in fig 4, 5, and 6 The results obtained for OPP scheme are shown in table III and the comparison on measurement redundancy of different OPP solutions for TN State Power Grid is listed in table IV.

CASE 1

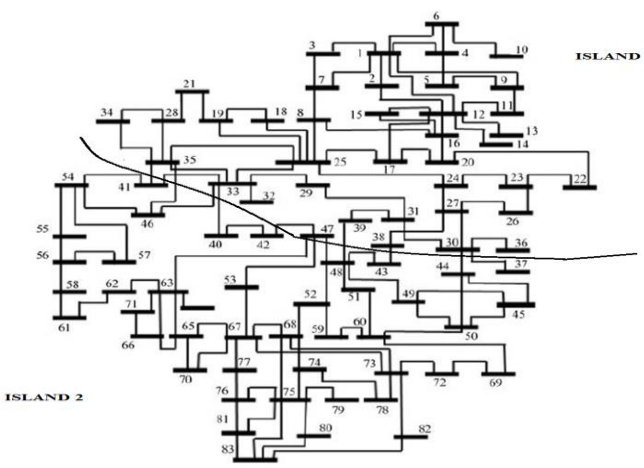


Figure 4 TN State Power Grid under CI - Case 1

Cite this article as: Dhanalakshmi G, Rajeswari R, Vignesh G. "Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment." *International Conference on Information Engineering, Management and Security (2015)*: 159-167. Print.

CASE 2

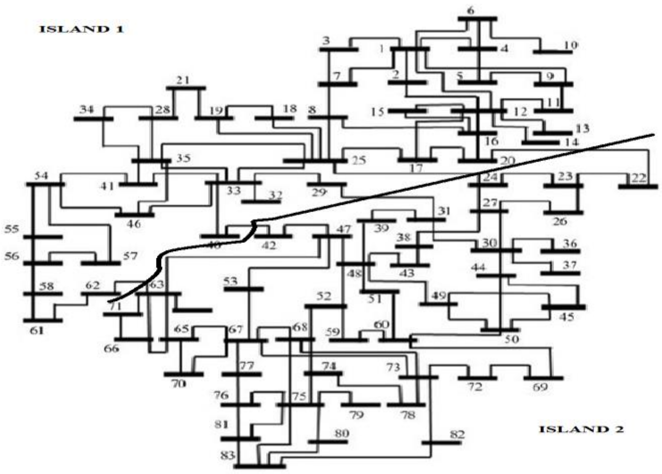


Figure 5 TN State Power Grid under CI - Case 2

CASE 3

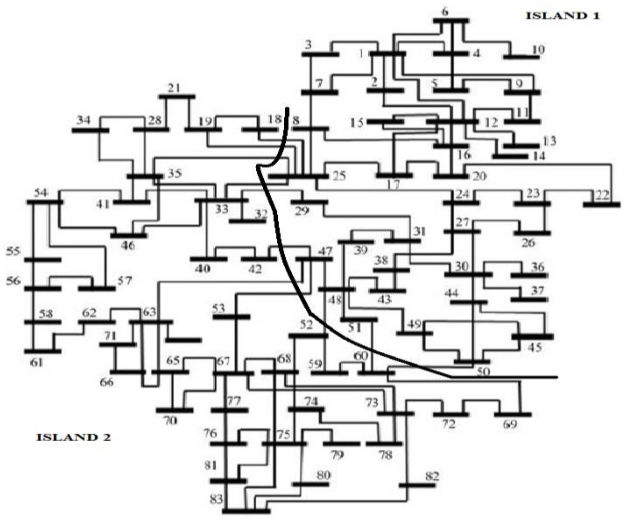


Figure 6 TN State Power Grid under CI - Case 3

Table III Results for Solved OPP in Controlled Islanding conditions

S.No	Cases	Results For OPP
1	Case 1	5, 7, 10, 12, 19, 22, 27, 30, 32, 34, 40, 45, 48, 54, 58, 60, 63, 67, 73, 75.
2	Case 2	1, 6, 12, 16, 19, 23, 30, 33, 34, 42, 43, 45, 48, 55, 56, 60, 62, 63, 67, 73, 75.
3	Case 3	1, 6, 8, 12, 19, 22, 27, 29, 30, 33, 35, 42, 48, 50, 54, 58, 60, 63, 67, 68, 73, 75.

Cite this article as: Dhanalakshmi G, Rajeswari R, Vignesh G. "Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment." *International Conference on Information Engineering, Management and Security (2015)*: 159-167. Print.

Table IV Comparison on measurement redundancy of different OPP solutions for IEEE-14 bus system

OPP Solutions	Measurement Redundancy Difference		
	Normal Operation	*Islanding Operation*	*Value of F_2*
Case 1	2.8313	2.5904	2.7590
Case 2	2.6506	2.5663	2.6253
Case 3	2.5060	2.4578	2.4916

Therefore, for the TN State Power Grid system, Case 3 is the most suitable solution because it has smaller value of redundancy factor than other ones, as shown in table IV

V CONCLUSION

Smart Grid(SG) can deliver reliable electric power to consumers with efficient utilization of power network than that provided by the traditional power system. SG is essential for a developing and highly populated country like India. One of the key requirements for the implementation of SG is the complete observability of the power grid, which can be achieved by using PMUs.

An effective OPP scheme should ensure complete observability of a power network under various operation conditions. To avoid wide-area blackout following cascading failures, power system might be operated in controlled islanding mode. In this thesis, an OPP model considering controlled islanding of power system is proposed. The proposed model guarantees complete observability of power network for normal condition as well as controlled islanding condition. By introducing the measurement redundancy into the optimization objective, our OPP-CI model can find the globally optimal solution with the minimum number of PMUs and maximum measurement redundancy. At last, case studies on IEEE-14 Bus standard test systems and Tamil Nadu State Power Grid (83 Bus System) practical system provide verification of the effectiveness of the presented OPP models.

This investigation can be applied to remaining all the National Power Grid of India so that the OPP schemes under Normal and Controlled Islanding conditions can be determined. Thus wide-area blackout following cascading failures can be avoided.

REFERENCES

[1] Lei Huang, Yuanzhang Sun and JianXu., "Optimal PMU Placement Considering Controlled Islanding of Power System", IEEE Transactions on Power Systems, 2013 IEEE, 0885-8950 .

[2] A. Enshaee, R. A. Hooshmand, and F. H. Fesharaki, "A new method for optimal placement of phasor measurement units to maintain full network observability under various contingencies," Elect. Power Syst. Res., vol. 89, no.1,pp. 1–10, Aug.2012.

[3] Gou,B., "Generalized Integer Linear Programming Formulation for Optimal PMU Placement". Power Systems , IEEE Transactions on, 2008. 23(3): 1099-1101.

[4] B. Gou, "Optimal placement of PMUs by integer linear programming," IEEE Trans. Power Syst., vol. 23, no. 3, pp. 1525–1526, Aug. 2008.

[5] Pathirikkat GOPAKUMAR,G .Surya CHANDRA,M .jayabharata REDDY, "Optimal placement of PMUs for the smart grid implementation in Indian power Grid - A case study". Front. Energy 2013, 7(3): 358-372

[6] F. Aminifar, A. Khodaei, M. Fotuhi-Firuzabad, and M. Shahidehpour, "Contingency-constrained PMU placement in power networks," IEEE Trans.Power Syst., vol. 25, no. 1, pp. 516–523, Feb. 2010.

[7] M. R. Aghamohammadi and A. Shahmohammadi, "Intentional islanding using a new algorithm based on ant search mechanism," Int. J. Elect. Power Energy Syst., vol. 35, no. 1, pp. 138–147, Feb. 2012.

[8] G. Xu, V. Vittal, A. Meklin, and J. E. Thalman, "Controlled islanding demonstrations on the WECC system," IEEE Trans. Power Syst., vol. 26, no. 1, pp. 334–343, Feb. 2011.

[9] S. S. Ahmed, N. C. Sarker, A. B. Khairuddin, M. R. B. A. Ghani, and H. Ahmad, "A scheme for controlled islanding to prevent subsequent blackout,"IEEE Trans. Power Syst., vol. 18, no. 1, pp. 136–143, Feb. 2003.

[10] R. F. Nuqui and A. G. Phadke, "Phasor measurement unit placement techniques for complete and incomplete observability," IEEE Trans. Power Del., vol. 20, no. 4, pp. 2381–2388, Oct. 2005.

[11] Bindeshwar Singh, N.K. Sharma, A.N. Tiwari, K.S. Verma, and S.N. Singh "Applications of phasor measurement units(PMUs) in electric power system networks incorporated with FACTS controllers ", International Journal of Engineering,Science and Technology Vol. 3, No. 3, 2011, pp. 64-82.

[12] A.G. Phadke and J. S. Thorp " History And Applications Of Phasor Measurements",IEEE,142440178X/06/2006.

[13] A.G. Phadke , "Synchronized Phasor Measurements - A Historical Overview", 0-7803-7525-4/02/ 2002 IEEE

Cite this article as: Dhanalakshmi G, Rajeswari R, Vignesh G. "Optimal Pmu Placement For Tamilnadu Grid Under Controlled Islanding Environment." *International Conference on Information Engineering, Management and Security (2015)*: 159-167. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS027

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.027

Disease Diagnosis using Meta-Learning Framework

Utkarsh Pathak[1], PrakhyaAgarwal[1], PoornalathaG[1]
[1]Information and Communication Technology Department
Manipal Institute of Technology, Manipal University
Manipal, Karnataka 576-104, India

Abstract: *Data mining techniques have been widely used in clinical decision support systems for prediction and diagnosis of various diseases with good accuracy. These techniques have been very effective in designing clinical support systems because of their ability to discover hidden patterns and relationships in medical data. The main objective of this paper is to develop and implement a framework which provides considerable classification results for users who have no prior data mining knowledge. We also propose a suitable prediction model to enhance the reliability of medical examinations and treatments for diseases. We analyzed different medical records for certain disease and based on the hypothesis made on the training dataset, applied it on the test dataset and achieved disease with a good accuracy. We focus on minimizing the system dependence on user input while providing the ability of a guided search for a suitable learning algorithm through performance metrics.*

Keywords: Meta-learning framework, Dataset features, classifier.

I. INTRODUCTION

As one introduces new dataset to the system, one important step is selecting which classifier will serve with one of the best accuracies for that data. An initial assessment is time consuming since one has to decide which classifier is most suited in the given context. Thus, selecting a suitable classifier for the dataset is a complex task. Even an experienced analyst might find it very difficult to find it out. Moreover, some hidden knowledge could be present in data which adds to the problem. Here, we take up an approach which involves comparing the new problem with a set of problems for which the classifier performances are already known. First, using the meta-features that are extracted from the dataset, the dataset is plot in the space. Next, identification of the dataset which resembles the most to the new dataset is carried out using distance computation Consequently the same classifier and settings that are obtained from the near neighbour are expected to achieve similar performances on the new dataset. Thus making a structure which unites the tools important to investigate new datasets and make predictions using the learning algorithm's performance would greatly aid the novice user. This outcomes in a critical pace up and an expanded dependability on the choice of the learning algorithm. The tool we discuss is proposed in [1] and the datasets used are all of .arff format and provided by the Weka Framework. We have added a functionality of prediction; where the user uploads the test and train datasets and the prediction is done on the class attribute. The rest of the paper is organized as follows: The literature survey is discussed in section II, details regarding the proposed model is given in section III, the results obtained are given in section IV, conclusions and future scope is provided in section V, followed by the references at the end.

Cite this article as: Utkarsh Pathak, PrakhyaAgarwal, Poornalatha G. "Disease Diagnosis using Meta-Learning Framework." *International Conference on Information Engineering, Management and Security (2015)*: 168-172. Print.

II. LITERATURE SURVEY

Aha [2] proposes a system that constructs rules which describe how the performance of classification algorithms is determined by the characteristics of the dataset. Rendell et al.[3] describe a system VBMS, which predicts the algorithms that perform better for a given classification problem using the problem characteristics (number of examples and number of attributes). The main limitation of VBMS is that the training process runs every time a new classification task is presented to it, which makes it slow. The approach applied in the Consultant expert system relies heavily on a close interaction with the user. Consultant poses questions to the user and tries to determine the nature of the problem from the answers. It does not use any knowledge about the actual data. Schaffer [4] proposes a brute force method for selecting the appropriate learner: execute all available learners for the problem at hand and estimate their accuracy using cross validation. The system selects the learner that achieves the highest score. This method has a high demand of computational resources. Statlog [5] extracts several characteristics from datasets and uses them together with the performance of inducers (estimated as the predictive accuracy) on the datasets to create a meta-learning problem. It then employs machine learning techniques to derive rules that map dataset characteristics to inducer performance. The limitations of the system include the fact that it considers a limited number of data sets. Moreover, it incorporates a small set of data characteristics and uses accuracy as the sole performance measure. The use of our framework is inspired by the work done in [1] which discusses the benefits of meta-data and feature selection for mining purposes. We have used the framework as the basis of our proposed model and also added the feature to predict the diagnosis.

III. PROPOSED MODEL

In this section, we present the formal working of our framework shown in fig 1. The essential characteristic of the proposed model is

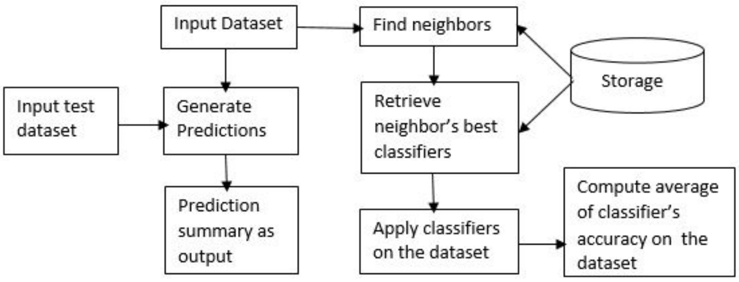


Fig. 1. Framework Model

to recommend a precise learning algorithm for a dataset submitted to the framework. The framework should achieve results with just the knowledge of the neighbour's best classifiers. The first step is to store the meta-data of the dataset. These include the total number of attributes of a dataset, the number of nominal attributes, the number of Boolean attributes and the number of continuous (numeric) attributes, the maximum number of distinct values for nominal attributes, the minimum number of distinct values for nominal attributes, the mean of distinct values for nominal attributes, the standard deviation of distinct values for nominal attributes and the mean entropy of discrete variables. Similarly for continuous attributes, it includes the mean skewness of continuous variables, which measures the asymmetry of the probability distribution, and the mean kurtosis of continuous variables representing the peak of the probability distribution. Finally, the dimensionality of the dataset is stored; It contains the overall size, represented by the number of instances, and imbalance rate information. The next step includes computing distance between the analyzed dataset and the datasets stored in the framework. The distance is computed by using the dataset metafeatures (all numeric values) as coordinates of the dataset. By representing a dataset as a point in a vector space, the distance can be evaluated using any metric defined on a vector space. The first distance computation strategy considered is the normalized Euclidean distance(E). The Euclidean distance is the ordinary distance between two points in space, as given by the Pythagorean formula. The next step is the neighbour selection step, after the distance computation phase, a list of distances is obtained, we select the Top 3 (i.e the three least distances) and we analyse the classifiers which yielded the best result on them and store the classifiers name. In the next step, we use the classifiers obtained from the last phase and use it on the analyzed dataset; we compute the average accuracy and output it to the user. Finally, we have added prediction functionality where the user uploads the test and train dataset and we predict the disease with considerable accuracy. The classifier used for prediction is J48 (studies show that J48 is more reliable than other classifiers).

Cite this article as: Utkarsh Pathak, PrakhyaAgarwal, Poornalatha G. "Disease Diagnosis using Meta-Learning Framework." *International Conference on Information Engineering, Management and Security (2015)*: 168-172. Print.

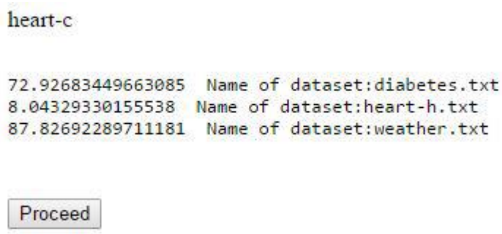


Fig. 2. Neighbours of heart-c.arff

TABLE I. ACCURACIES IN PERCENTAGE

Dataset	J48	NaiveBayes	BayesNet	SMO	Neighbour$_a pproach$
heart-c.arff	77.55	83.49	83.49	84.15	81.73
heart-h.arff	80.95	83.67	85.03	82.65	82.08
diabetes.arff	73.82	76.3	74.34	77.34	75.17
contact-lenses.arff	83.33	70.83	70.83	70.83	75.0

IV. RESULTS

The system offers a web application which first authenticates a user; if not a valid member, a registration process is provided. Once the user is authenticated, he/she can upload a test dataset. Now, firstly the meta-features are extracted and listing of the data set features and the minimum and maximum values for each of these features is done. Next, the neighbours for the uploaded dataset is computed (i.e the top three neighbours). In the next step, the classifiers which yielded the best result on the respective neighbours is applied on the analyzed dataset and an average accuracy is computed. As mentioned earlier, the datasets used in our framework are all of (.arff) format and can be found at [6] and all the datasets mentioned at [6] are used as potential neighbours in our framework.

The Fig 1. shows neighbours results for the dataset heartc. arff ; the top three neighbour result lists out heart-h.arff ;. Now, in Table I. we have listed out the name of some of the datasets and correspondingly the accuracy (in %) obtained by the classifiers namely J48, NaiveBayes, BayesNet and SMO and the neighbour classification as the last column; In the Fig 2., we have plotted the classifier accuracy (in %) for all the datasets listed in the table. Note, our neighbour approach outperforms some of the classifiers in every dataset.

In Table II, we have computed the average classifier accuracy of every classifier over the course of all the four datasets and we find out that our neighbour approach outperforms the popular Bayesian Network Model (i.e BayesNet classifier) and performs almost as efficiently as all the othe classifiers. Now, for the prediction functionality, the user has to upload train and test datasets (see Fig 3.) and the dataset uploaded is for Prostrate Tumor and a classifier (J48) is applied on the training dataset. This step builds the decision boundary or the hypothesis model which is then applied on the test dataset (on the class atrribute) for prediction. The accuracy of prediction depends mainly on the accuracy of classification on the training dataset.

The next step comprises the display of result of the classification along with the detailed summary and the confusion matrix is presented to the user as output (shown in Fig 4.) which lists out that out of 34 samples, 9 are normal and 25 are malignant which is correct.

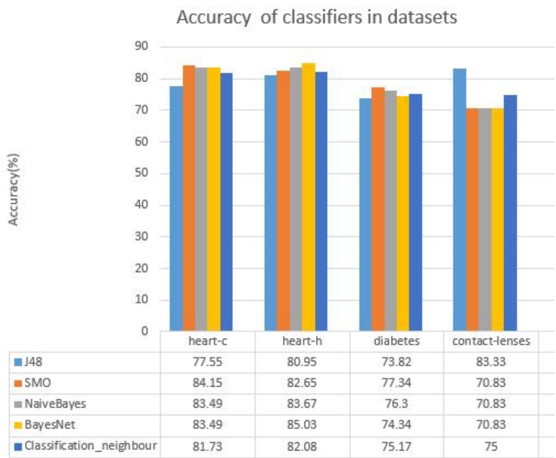


Fig. 3. Bar-graph showing accuracies of classifiers on different datasets

Cite this article as: Utkarsh Pathak, PrakhyaAgarwal, Poornalatha G. "Disease Diagnosis using Meta-Learning Framework." *International Conference on Information Engineering, Management and Security (2015)*: 168-172. Print.

TABLE II. AVERAGE ACCURACY IN PERCENTAGE

J48	NaiveBayes	BayesNet	SMO	Neighbour$_a pproach$
78.91	78.57	78.42	78.74	78.50

sification along with the detailed summary and the confusion matrix is presented to the user as output (shown in Fig 4.) which lists out that out of 34 samples, 9 are normal and 25 are malignant which is correct.

V. CONCLUSIONS AND FUTURE SCOPE

The successful application of data mining in highly visible fields like e-business, stock marketing and retail has led to its application in other industries and sectors. Among these sectors just discovering is healthcare and disease prediction. In our work, we have used a framework for classification which is done by using the classifiers which yielded the best results on the neighbours of our test dataset. Moreover the prediction of disease functionality has also been added which makes this model highly beneficial as the complex task of predicting a disease based on patterns on similar data has been done with sufficient accuracy and diligence.

The user is presented with the option of doing classification

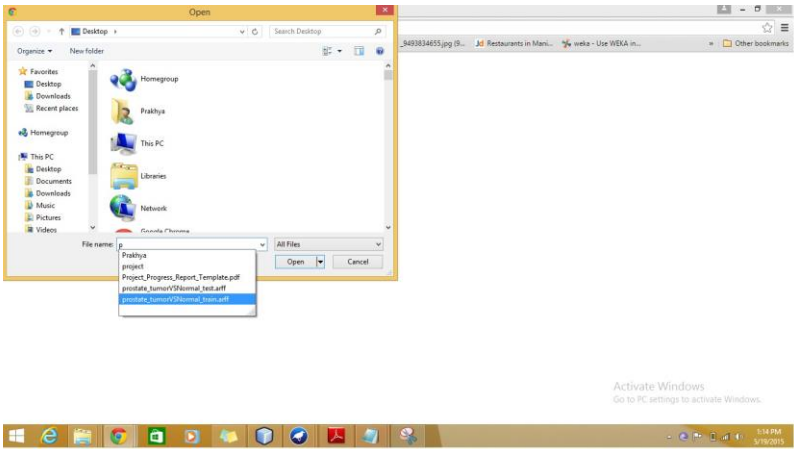


Fig. 4. Test and Train file for Prediction

```
Results
======

Correctly Classified Instances           9               26.4706 %
Incorrectly Classified Instances        25               73.5294 %
Kappa statistic                          0
K&B Relative Info Score              -1645.2619 %
K&B Information Score                  -16.4482 bits     -0.4838 bits/instance
Class complexity | order 0              33.5651 bits      0.9872 bits/instance
Class complexity | scheme            26850      bits    789.7059 bits/instance
Complexity improvement     (Sf)     -26816.4349 bits   -788.7187 bits/instance
Mean absolute error                      0.7353
Root mean squared error                  0.8575
Relative absolute error                148.4018 %
Root relative squared error            173.0394 %
Coverage of cases (0.95 level)          26.4706 %
Mean rel. region size (0.95 level)      50      %
Total Number of Instances               34

=== Confusion Matrix ===

  a  b   <-- classified as
  0 25 |  a = Tumor
  0  9 |  b = Normal
```

Fig. 5. Summary and Confusion Matrix

using the classifiers or using the neighbour approach. In our work, we found that most of the test dataset yielded a better result with the neighbour approach than the accuracy achieved by the worst classifier; thus the user can achieve healthy classification result even if he is devoid of any prior data mining knowledge.

Cite this article as: Utkarsh Pathak, PrakhyaAgarwal, Poornalatha G. "Disease Diagnosis using Meta-Learning Framework." *International Conference on Information Engineering, Management and Security (2015)*: 168-172. Print.

However there is scope for further improvement; selecting the neighbours is a complex and tricky task and the number of neighbours to be found out for every test dataset is an open problem (we have taken 3 closest neighbours), the number of classifiers used could be incremented to achieve even greater accuracy.

In the prediction technique, there is a lack of extensive train and test datasets of most of the diseases as the task of accumulating the data and narrowing the number of attribute(i.e feature selection) to a limited number of attributes which affect the class attribute is a very complex task. However, the availability of the real dataset would greatly help us to learn more about disease diagnosis and prediction. Medical diagnosis is regarded as an important yet complicated task that needs to be executed accurately and efficiently. The automation of this system would be extremely advantageous. There is a shortage of resource persons and manpower at almost every hospital, therefore an automatic medical diagnosis system would probably be exceedingly beneficial by getting positive results even from novice or inexperienced users.

ACKNOWLEDGMENT

This project consumed huge amount of work, research and dedication. Still, implementation would not have been possible if we did not have a support of many individuals and organization. Therefore we would like to extend our sincere gratitude to all of them.

REFERENCES

[1] Potolea, Rodica and Cacoveanu, Silviu and Lemnaru, Camelia, Metalearning framework for prediction strategy evaluation, Enterprise Information Systems,Springer. page 280-295, 2011.

[2] Aha, David W, Generalizing from case studies: A case study, Proc. of the 9th International Conference on Machine Learning. Page 1-10, 1992.

[3] Rendell, Larry A and Sheshu, Raj and Tcheng, David K, Layered Concept-Learning and Dynamically Variable Bias Management, IJCAI. page 308-314, 1987.

[4] Schaffer, Cullen, Selecting a classification method by cross-validation, Machine Learning, Springer. page 135-143, 1993.

[5] Michie, Donald and Spiegelhalter, David J and Taylor, Charles C, Machine learning, neural and statistical classification, Citeseer. 1994.

[6] Sample Weka Datasets, http://storm.cis.fordham.edu/ gweiss/datamining/ datasets.html, (last accessed May 2015).

[7] Dataset Repository, http://datam.i2r.a-star.edu.sg/datasets/index.html, (last accessed May 2015).

Cite this article as: Utkarsh Pathak, PrakhyaAgarwal, Poornalatha G. "Disease Diagnosis using Meta-Learning Framework." *International Conference on Information Engineering, Management and Security (2015)*: 168-172. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS028		**eAID**	ICIEMS.2015.028

Soft computing Applications to power systems

Comparison of Numerical Techniques Applied to Shunt Connected Reactive Power Control Device

Dr. Minal Salunke[1], Anupama Aili and Manasa Aili[2]
[1]Senior Professor, Electrical and Electronics Dept.,
B.V. Bhoomaraddi College of Engineering and Technology, Hubli, India
[2]Student, E&E Dept., B.V. Bhoomaraddi College of Engineering and Technology, Hubli, India

Abstract: *This paper presents the application of different numerical techniques to model one of the FACTS controlled device TCR (Thyristor controlled reactor) for power enhancement in transmission lines with the help of power electronics concepts. The results are verified with Saber RD student edition for circuit simulation.*

Keywords**:** Numerical Techniques, TCR, FACTS devices

I. INTRODUCTION

By definition, capacitors generate and reactors (inductors) absorb reactive power when connected to an ac power source. They have been used with mechanical switches for (coarsely) controlled var generation and absorption since the early days of ac power transmission. Continuously variable var generation or absorption for dynamic system compensation was originally provided by over-or under excited rotating synchronous machines and later, by saturating reactors in conjunction with fixed capacitors [2]. Since the early 1970s high power, line-commutated thyristors in conjunction with capacitors and reactors have been employed in various circuit configurations to produce variable reactive output. These in effect provide a variable shunt impedance by synchronously switching shunt capacitors and /or reactors "in" and "out" of the network. Using appropriate switch control, the var output can be controlled continuously from maximum capacitive to maximum inductive output at a given bus voltage. More recently gate turn-off thyristors and other power semiconductors with internal turn-off capability have been used in switching converter circuits to generate and absorb reactive power without the use of the ac capacitors or reactors. These perform as ideal synchronous compensators (condensers), in which the magnitude of the internally generated ac voltage is varied to control the var output. All of the different semiconductor power circuits, with their internal control enabling them to produce var output proportional to an input reference, are collectively termed by the joint IEEE and CIGRE definition, static var generators (SVC). Thus, a static var compensator (SVC) is, by the IEEE CIGRE co-definition, a static var generator whose output is varied so as to maintain or control specific parameters (e.g., voltage, frequency) of the electric power system. A TCR is one of the most important building blocks of thyristor-based SVCs. Although it can be used alone, it is most often employed in conjunction with fixed or thyristor-switched capacitors to provide rapid, continuous control of reactive power over the entire selected lagging-to-leading range [2].

Cite this article as: Dr. Minal Salunke, Anupama Aili and Manasa Aili. "Soft computing Applications to power systems." *International Conference on Information Engineering, Management and Security (2015)*: 173-178. Print.

II. Objectives

Objectives of this paper are:

- To motivate the study of numerical methods through discussion of engineering applications.
- To determine the performance of impedance type var generator-the thyristor controlled reactor by applying different numerical methods.
- To verify the program results using saber simulation tool
- To compare the simulations results with GNU plots obtained from code blocks using C as coding language
- To verify the decrease in sinusoidal property of TCR current and increase in transmittable power in transmission line.

III. NUMERICAL METHODS

Calculus is a branch of mathematics involving or leading to calculations dealing with continuously varying functions
Calculus is a subject that falls into two parts:

- Differential calculus (or differentiation)
- Integral calculus (or integration)

The equations which are composed of an unknown function and its derivative are called differential equations. When the function involves one independent variable, the equation is called as ordinary differential equation. Differential equations are classified as their order. If the highest derivative is a first derivative, then it is a first –order equation. A second order equation would include a second derivative.

A. Different methods to solve differential equation

In this section, a brief review of various numerical techniques commonly employed in the stability study is presented. To solve a differential equation, the numerical techniques employed are [1]

1. Forward Euler's Method
2. Backward Euler's Method
3. R-K 4th Order Method

Given a differential equation,

$$\frac{dy}{dx} = f(x,t) \qquad (1)$$

1. Forward Euler method:
The algorithm is given by:

$$X_{n+1} = X_n + f(x_n, t_n)h \qquad (2)$$

Where 'h' is the step size.
For an RL series circuit with v(t) as the source voltage.

$$\frac{dy}{dx} = \frac{di_L}{dt} = \left(\frac{v - i_L R}{L}\right) \qquad (3)$$

For the nth interval

i_l (4) $n) + Kh$

Where $K = \frac{V_n}{L} - \frac{R i_{L(n)}}{L}$

2. Backward Euler Method:
The algorithm is given by:

$$X_{n+1} = X_n + \left(X_{n+1,} t_{n+1}\right)h \qquad (5)$$

For an RL series circuit with v(t) as the source voltage,

$$\frac{dy}{dx} = \frac{di_L}{dt} = \left(\frac{v - i_L R}{L}\right) \qquad (6)$$

For the nth interval, $\boldsymbol{i_{L(n+1)}} = \frac{\boldsymbol{i_{L(n)}L + vh}}{\boldsymbol{L + Rh}}$ (7) (3.2.3)

3. Runge –Kutta(R-K) 4th order approximation Method
The value of *x* at the end of the interval is given as:

$$x_{n+1} = x_n + \frac{(k_1 + 2k_2 + 2k_3 + k_4)}{6} \qquad (8)$$

Cite this article as: Dr. Minal Salunke, Anupama Aili and Manasa Aili. "Soft computing Applications to power systems." *International Conference on Information Engineering, Management and Security (2015)*: 173-178. Print.

Where,

$$k_1 = f(x_1, t_1)\Delta t$$
$$k_2 = f[(x_1 + 0.5k_1), (t_1 + 0.5\Delta t)]\Delta t$$
$$k_3 = f[(x_1 + 0.5k_2), (t_1 + 0.5\Delta t)]\Delta t$$
$$k_4 = f[(x_n + hk_3), (t_n + h)]$$

It is to be noted that R-K method employs slope of the curve at predetermined points within the interval to calculate the value of *x* at the end of the interval.

I. MODELING OF TCR

With increased power transfer, transient and dynamic stability is of increasing importance for secure operation of power systems. A power electronic based system and other static equipment that provides control of one or more transmission parameters are called FACTS controllers.

A basic single-phase TCR comprises an anti-parallel–connected pair of thyristors valves, *T*1 and *T*2, in series with a linear air-core reactor, as illustrated in Fig.1. The anti-parallel–connected thyristor pair acts like a bidirectional switch, with thyristor valve *T*1 conducting in positive half-cycles and thyristor valve *T*2 conducting in negative half-cycles of the supply voltage. The firing angle of the thyristors is measured from the zero crossing of the voltage appearing across its terminals.

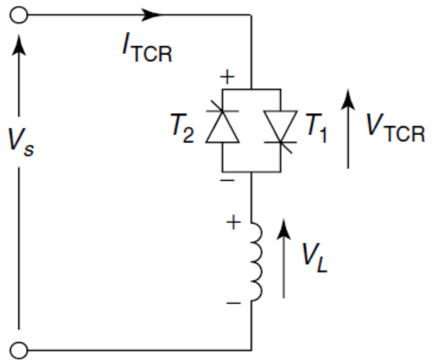


Fig.1 A TCR

For modeling and analysis of TCR, the most practical available method is time domain simulation in which the nonlinear differential equations are solved using numerical method considering the time step. The main aim of these devices is to decrease the line reactance so that the power transmitted to the load is increased.

$$\frac{di_{(TCR)}}{dt} = \frac{V_S - iR}{L} \quad (9)$$

With $v(t) = v_m \sin(\omega t)$ as sine function, R=3Ω, L=0.01H and with firing angles, 90°,110°,150° Euler's forward method is applied to the above differential equation and the results obtained are analyzed.

The basic equation of power transmission is given by:

$$P = \frac{V_1 V_2}{X} \sin \delta \quad (10)$$

Where, V_1 and V_2 are voltages at both ends, δ is the angle between V_1 and V_2, X is the total line reactance.

The controllable range of the TCR firing angle, α, extends from 90° to 180°. A firing angle of 90° results in full thyristor conduction with a continuous sinusoidal current flow in the TCR. As firing angle is increased above 90°, non-sinusoidal current flows and magnitude of fundamental frequency of current reduces. This is equivalent to increase in the inductance value which in turn decreases the line reactance thereby reducing its capacity to draw reactive power and hence enhance the transmittable power.

Cite this article as: Dr. Minal Salunke, Anupama Aili and Manasa Aili. "Soft computing Applications to power systems." *International Conference on Information Engineering, Management and Security (2015)*: 173-178. Print.

I. RESULTS

As the firing angle is varied from 90° to close to 180°, the current flows in the form of discontinuous pulses symmetrically located in the positive and negative half-cycles. The results are displayed in fig. 2, 4 and 6 below for different firing angles. The results are verified with Saber RD student edition for circuit simulation shown in fig. 3, 5 and 7 below for different firing angles

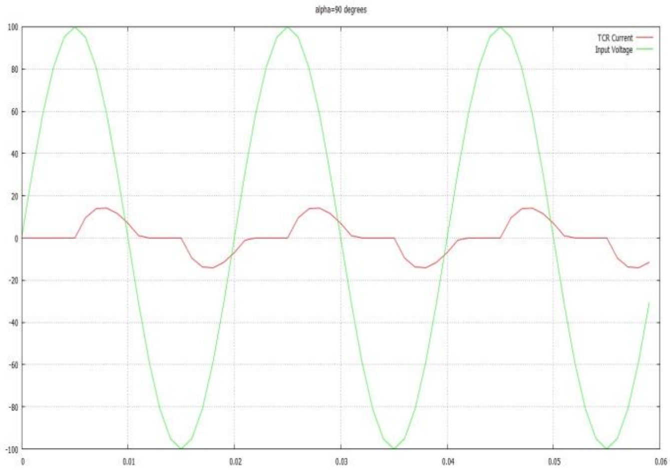


Fig.2 GNU plots of voltage and current for α=90° in a TCR

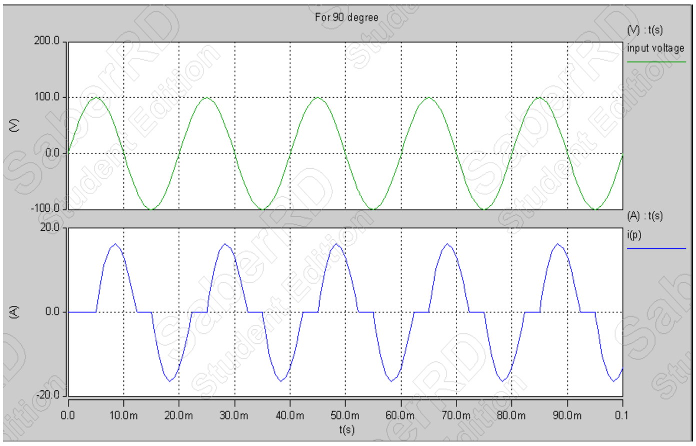


Fig.3 Voltage and current for α=90° in a TCR

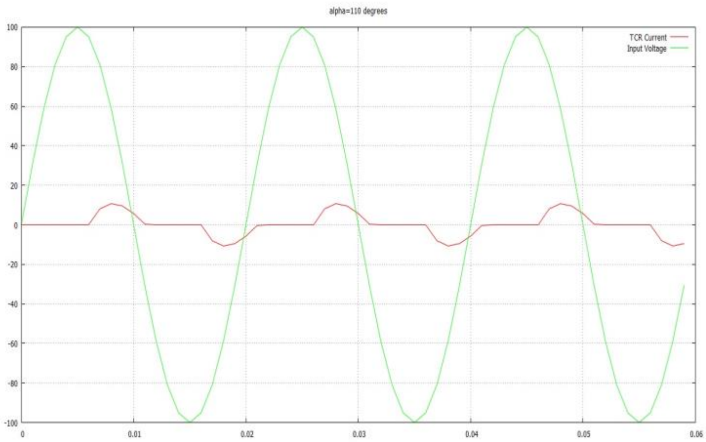


Fig.4 GNU plots of voltage and current for α=110° in a TCR

Cite this article as: Dr. Minal Salunke, Anupama Aili and Manasa Aili. "Soft computing Applications to power systems." *International Conference on Information Engineering, Management and Security (2015)*: 173-178. Print.

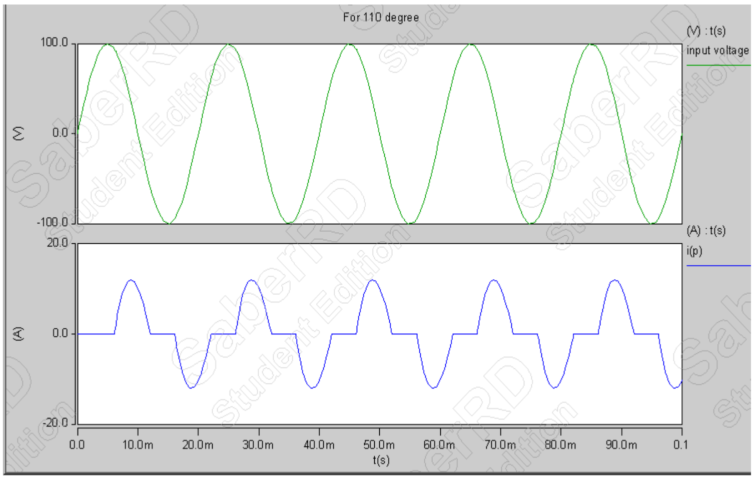


Fig.5 Voltage and current for α=110° in a TCR

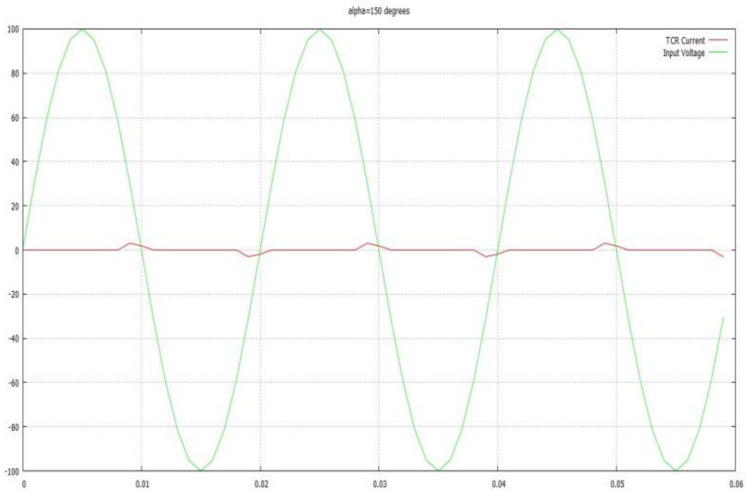


Fig.6 GNU plots of voltage and current for α=150° in a TCR

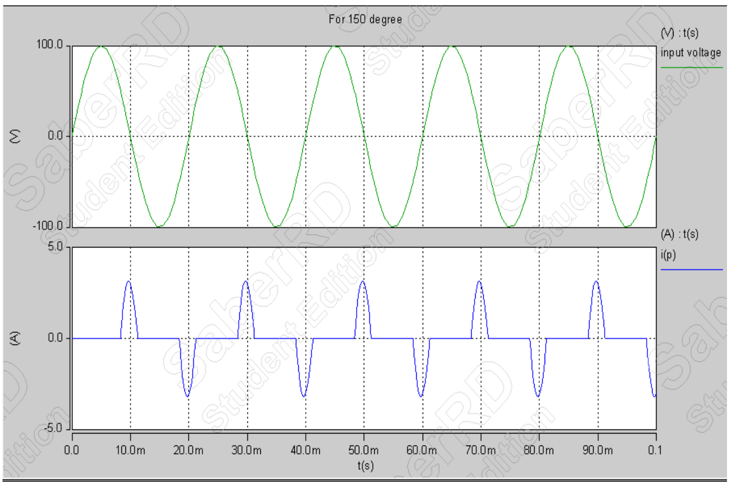


Fig.7 Voltage and current for α=150° in a TCR

Cite this article as: Dr. Minal Salunke, Anupama Aili and Manasa Aili. "Soft computing Applications to power systems." *International Conference on Information Engineering, Management and Security (2015)*: 173-178. Print.

IV. CONCLUSION

- It is evident that the current in the reactor can be varied continuously by the method of delay angle control from maximum $\alpha = \frac{\pi}{2}$ (to minimum $\alpha = \pi$)
- Increasing the value of firing above $\frac{\pi}{2}$ causes the TCR current waveform to become non-sinusoidal, with its fundamental frequency component reducing in magnitude. This, in turn, is equivalent to an increase in the reactor, reducing its ability to draw reactive power from the network at the point of connection.

REFERENCES

[1] Dr. K.N. Shubhanga, *"State-space Modelling and Numerical Integration Techniques: An Introduction"*, July 20, 2014.
[2] R.MohanMathur and Rajiv.K.Verma, *Thyristor based FACTS controllers for electrical transmission systems*, Wiley India Pvt. Ltd.
[3] Narain G. Hingorani and Laszlo Gyugyi, *"Understanding FACTS-concepts and technology of FACTS"*, 2013.
[4] Enrique Acha , "*FACTS-Modelling and Simulation in power Networks",2012*
[5] A.M. Padma Reddy, *"Systematic approach to Data Structure Using C"*, Revised Edition, 2007

Cite this article as: Dr. Minal Salunke, Anupama Aili and Manasa Aili. "Soft computing Applications to power systems." *International Conference on Information Engineering, Management and Security (2015)*: 173-178. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS029

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.029

Titanium Alloy Subjected to Tensile testing under Ambient and Cryogenic Conditions using Acoustic Emission Techniques

Dr.S.Sundaram, G. Vetri Chelvan
Department of Mechanical Engineering,
Vidyaa Vikas College of Engineering and Technology, Tiruchengode-637 214, Tamilnadu, India

Abstract: Titanium (Ti) alloys are strategic aerospace materials used in relatively severe working environment. Owing to the excellent properties such as high rigidity to weight ratio, elevated temperature strength, corrosion resistance and toughness in ambient as well as cryogenic environment, Titanium alloys find high technology applications in aerospace industries. As the Ti alloy finds application in aircraft engines, compressor blades and gas turbines, it is necessary to characterize the performance of this material under stress conditions. Acoustic Emission (AE) is a high sensitivity technique for detecting active microscopic events in a material under stress. The processes that are capable of changing the internal structure of a material such as dislocation motion, directional diffusion, creep, grain boundary sliding and twinning which are usually associated in plastic deformation and fracture are the sources of Acoustic Emission. Thus, using AE signals, it is possible to evaluate the performance of material under stress. The data acquired can be used to predict the performance of products made of Ti alloy. With this view, the acoustic emission response of Ti alloy subjected to tensile testing under ambient and cryogenic conditions have been studied.

Keywords: Titanium (Ti) alloys, corrosion, Acoustic Emission, corrosion resistance, strength, etc.

I. INTRODUCTION

Performance of a product is largely dependent on design, manufacturing and maintenance. Material characteristics influence significantly on the three aspects. Any material under stress will respond according to the nature of stress and environment. Material response is to be carefully monitored, especially in the case of critical parts such as part in aerospace and related applications. The response of the material can be assessed in terms of observable / visible feature such elongation, contraction or other deformation related visible feature. These are macroscopic effects. However, if the response of the material is to be evaluated even in microscopic level, then suitable indicators have to be monitored. Acoustic Emission is a relative indicator for identifying the status of material under stress. Acoustic Emission (AE) as the transient elastic wave generated by the rapid release of energy from a localized source or sources within material when subjected to a state of stress. This energy release is associated with the abrupt redistribution of internal stresses and as a result of this, a stress wave is propagated through the material. The definition of AE given above Indicates that processes that are capable of changing the internal structure of a material such as dislocation motion, directional diffusion, creep, grain boundary sliding and twinning, which result in plastic deformation, phase transformations, vacancy coalescence and decohesion of inclusions and fracture are sources of acoustic emission; of processes said above, only plastic deformation and fracture are of

Cite this article as: Dr.S.Sundaram, G. Vetri Chelvan. "Titanium Alloy Subjected to Tensile testing under Ambient and Cryogenic Conditions using Acoustic Emission Techniques." *International Conference on Information Engineering, Management and Security (2015)*: 179-185. Print.

significance in metal cutting. Out of the four plastic deformation process mentioned, generally, dislocation motion is the dominant mechanism in crystalline materials which are widely used in practice [3]. AE study is used as a condition monitoring process by different researchers [13] and [14].Studies the acoustic behavior of screws under tensile load. The significance of the results for the in-process monitoring of screws is explained in this work. AE technique for the integrity evaluation of aerospace pressure chambers made of M250 Maraging steels is carried out by [4]. Due to the excellent properties to titanium alloy, such as good ductility, high temperature strength, corrosion resistance and lower density, this alloy finds under high technology application aerospace, chemical and petrochemical industries, it is necessary to acquire a thorough knowledge on the behavior of the material. The acoustic response of titanium alloy subjected to a tensile testing reveals the behavior of the material during fracture. Only limited literature is available in this area. The AE response of the material can indicate microstructure – property relationships. [7] Studied the AE produced during tensile straining and fracture to have a better understanding of the titanium aluminide alloy behavior. [12] Investigated the effects of matrix microstructure and interfacial properties on the fatigue and fracture behavior of a metastable titanium matrix composite. The damage behavior of the titanium matrix composite, during monotonic and cycle loading were studied through AE. [1] Conducted AE studies to locate and to observe the damage of the titanium matrix composite. The results were supported by SEM analysis carried out on the fractured surfaces.The relationship between microstructure and AE of Ti-641 – 4v has been studied by [11]. Different microstructures of Ti-641-4v alloy have been obtained through different grain sizes and different heart treatment procedures. The AE response of these different microstructures subjected to mechanical deformation rest has been studied. A detailed study on the micro fracture mechanism in fracture toughness test of Ti-8A1-1Mo-1v alloy was examined by AE wave analysis by [5]. The widespread use of cryogenic fluids for several industrial applications such as frozen food, metal industry, space application, superconductors and biomedical applications has to be studied. Suitable materials have to be selected, in such a way that selected material should have toughness, ductility and weld ability at this low temperature. Titanium by its inherent properties meets the requirements of cryogenic technology. As the titanium alloy finds application in aerospace and cryogenic in industries, the behavior of this material under both the working conditions has to be investigated.

2.0 ACOUSTIC EMISSION TECHNIQUE

2.1 Principle

The Acoustic Emission Technique (AET) is relatively recent entry in the field of non-destructive evaluation which has particularly shown a very high potential for material characterization and damage assessment in conventional as well as non-conventional material.

2.2 Definition

Acoustic emission (AE) is the class of phenomenon where transient elastic waves are generated by the rapid release of energy from localized sources within a material, or the transient elastic waves so generated. In other words, AE refers to the stress waves generated by dynamic processes in materials. Emission occurs as a release of a series of short impulsive energy packets. The energy thus released travels as a spherical wave front and can be picked from the surface of a material using highly sensitive transducers, (usually electro mechanical type). The picked energy is converted into electrical signal which on suitable processing and analysis can reveal valuable information about the source causing the energy release. The flow chart of typical AE system is shown in Figure 1.

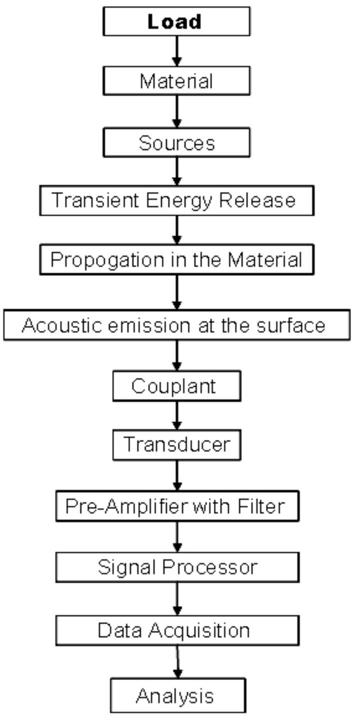


Figure 1: Flow Chart of Acoustic Emission System

Cite this article as: Dr.S.Sundaram, G. Vetri Chelvan. "Titanium Alloy Subjected to Tensile testing under Ambient and Cryogenic Conditions using Acoustic Emission Techniques." *International Conference on Information Engineering, Management and Security (2015)*: 179-185. Print.

The load applied on the material results in the transient energy release from the source. It obviously travels as a spherical wave front. As these pressure wave propagate through the material it undergoes distortion and attenuation. The volume and characteristics of AE generated are dependent on the nature. Type and characteristics of the source: the main characteristics being its initial severity, present state, local metallurgical structure and the environment. These are converted in to electrical signals by mounting Piezo electric transducer in suitable locations on the material by pasting them with complaints. As the stress waves pass through the compliant and transducers they further undergo distortions depending upon their transfer function characteristics. In order to increase strength of the signals, a preamplifier with filter leads the AE signals to the signal processor where ambient noise and unwanted frequency comportment of the signal are eliminated. This also helps to increase the signal to noise ratio. It further leads to data acquisition unit for analysis.

2.3 Sources Of Acoustic Emission

Sources of AE include many different mechanisms of deformation and fracture. Sources that have been identified in metals include moving dislocation, slip, twinning, grain boundary sliding, crack initiation, crack growth etc. Other mechanisms like leaks, cavitations, friction, growth of magnetic domain wells, phase transformations also fall within definition and are detectable by Acoustic emission equipment. These sources are termed as secondary or pseudo sources.

Acoustic emission signals are transient in nature (burst emission). The transducer output can be modeled crudely as a decaying sinusoid. This model is applicable only for signal which can be identified as individual bursts with discernible time gap between two successive events. If the burst rate is very high, events may occur very close to one another and sometimes even overlapping, in which case it termed as 'continuous emission". Thus we can broadly divide AE signals into burst type and continuous type. The characteristics of these are compared in Table 1.

Table 1: Comparison of AE Signals

Characteristics	Burst Type	Continuous Type
Ringtown Counts	Low	High
Rise Time	Reduced rise Time	More rise time
Event duration	Shorter	Longer
Frequency	High	Low
Event Rate	Low	High

3.0 TITANIUM ALLOY

Titanium is a low density metallic element that is abundant in the earth's crust. Metallic titanium became available in the early 20th Century. Titanium is relatively expensive compared with other common metals (iron, copper, aluminum, Magnesium) but frequently, by virtue of its attractive properties, may be more cost effective than these other metals. Ti alloys can be cast, rolled, forged and otherwise produced in a variety of mill product forms or special shapes. The initial use of Ti alloy is aircraft engines, compressor blades (Pratt & Whitney Aircraft J – 57, Rolls – Royce Avon) and then as disks, (Pratt & Whitney Aircraft JT – 3D). In fact, the existence of Ti alloys made possible the fan – type gas turbine engines now in use. Ti and Ti alloys are noted primarily for outstanding strength –to – weight ratios, elevated temperature properties and corrosion resistance. They also possess high rigidity – to – weight ratios, good fatigue strength and toughness and, in some cases, excellent cryogenic properties. Because of these characteristics and improved fabrication technology, Titanium and its alloys are now important materials for aircraft and processing equipment.

Ti and Ti alloys are classified into three major categories according to the predominant phases present in the microstructure. These are

- Alpha Alloys
- Alpha + Beta Alloys
- Beta Alloys

4.0 EXPERIMENTAL SET UP

Acoustic Emission (AE) is the stress waves produced by sudden movements in stressed materials, the classic source of AE are defect related deformation. Elastic waves are generated due to the local changes in source region. These waves propagate as mechanical disturbances through the structure causing a time varying surface displacements/in this experiment using Titanium Alloy under tensile load in Universal Testing Machine (UTM) has been carried out for the AE Study.

The tensile specimens used have been machined out from Titanium Alloy (Ti-641-4V) Plates, which were in received condition, are subjected to annealing before the machining process. The axis of the specimen has been kept coinciding with the rolling direction of the plate. Care has been exercised to ensure that the radius of curvature at the gauge and has been as smooth as possible. The major interest in this class of specimens is toe study AE signature at the pre – yielding onset of yielding. As such no strict quality control

measures have been affected to control the width or thickness of the specimen within close tolerance at the gauge length portion. The tensile specimen is
subjected to the required loading / load cycle on Servo controlled SCHENCK TREBEL RM 250 Universal Testing Machine, 250 KN Capacity with the provision for fatigue cycling as well as displacement controlled loading machine cells for necessary screening if the genuine emission events from the specimens under test are to be acquired.

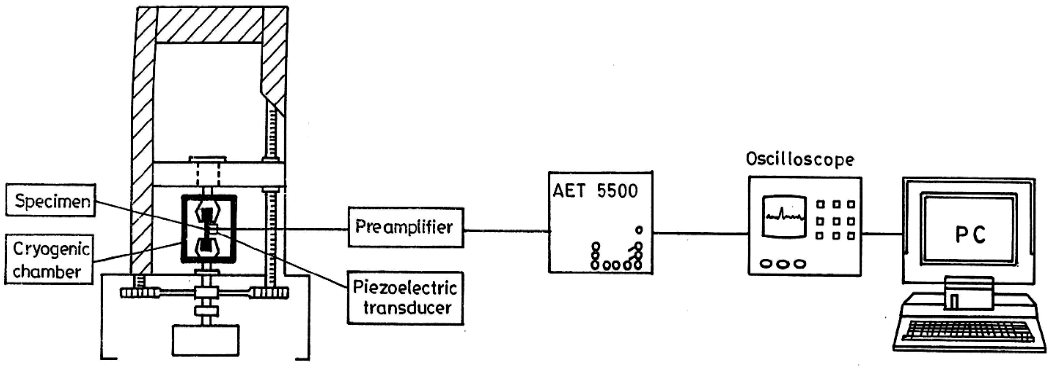
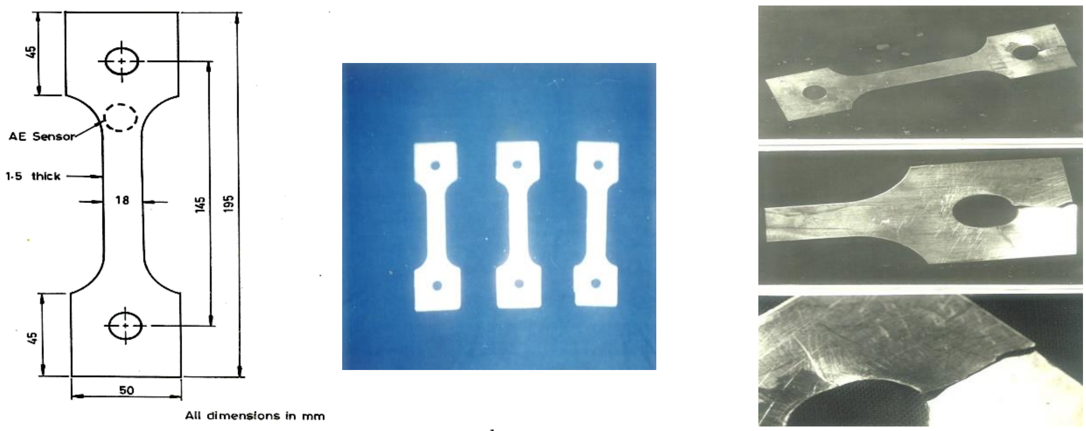


Figure 2: Schematic diagram of the experimental setup

Tensile test specimen of 1.5 mm thickness and 18 mm width, made of 6% Aluminum and 4% Vanadium, 9% of Titanium alloy is tested with tensile loading in an universal testing machine of 250 KN capacity. It is seen that as the load increases there is a steady improvement in activity; around certain critical/threshold load, the test material becomes active, associated with deformation and dislocation movements, the material exhibiting permanent set. This is associated with emission of continuous AE signal. As the load applied increases, the material tries to attain the threshold/permissible stress beyond which degradation/failure sets-in. Hence, as the material is subjected to increasing loading, acoustic stability is attained; with the result the material may even become inactive over the permissible stress region. Beyond that stress, failure of material sets-in associated with localized burst signals. An Oscilloscope has been used to acquire the amplified and filtered data. The 2 K pts resolution of the Oscilloscope and the storage duration of second/signal limited the signal rate and precision. The trigger threshold was adjusted exactly to just above the noise level. The signal exceeding this threshold (together with the elongation, load and speed of the crosshead) was recorded and the signal from the sensor was magnified by a pre-amplifier for increasing load.



(a) (b) (c)

Figure 3: (a) Tensile test Specimen (b) Typical Tensile specimen (c) Typical fractured tensile specimen

The raw signals were monitored using a data-logger. The acquired signal was analyzed separately using a suitable PC based data acquisition system at a sampling frequency of l MHz for spectrum analysis. The time duration signal consisting of 25 and 50 observations were considered of the purpose of analysis. The recorded data were used to calculate the AE rate and the frequency spectra. The AE signals were obtained under the ambient and cryogenic conditions.

Cite this article as: Dr.S.Sundaram, G. Vetri Chelvan. "Titanium Alloy Subjected to Tensile testing under Ambient and Cryogenic Conditions using Acoustic Emission Techniques." *International Conference on Information Engineering, Management and Security (2015)*: 179-185. Print.

5.0 RESULTS AND DISCUSSION

The titanium (Alpha + Beta phase) material was tested to uni-axial tensile loading. Typical observed load – extension characteristics of the test material are illustrated in 5.1. It is seen that up to around 13% (2mm) elastic behaviors can be observed; beyond that elongation the material undergoes plastic deformation associated with viscous yielding. This is indicated by the occurrence of staircase type load-extension characteristics. The response of the titanium alloy to tensile loading was monitored on-line by sending the acoustic emission signal emanation from the test specimen by suitably positioning a broad band AE sensor the recorded signal was characterized in terms of rms value and dominating frequency.

Typical observed r.m.s value of the acoustic emission signal monitored is illustrated in fig 5.2. It can be seen that with tensile load on, there is a gradual reduction in r.m.s value indicating that slow degradation of the material with applied load/stress. The r.m.s characteristics is of zigzag nature; form this it can be inferred that after certain cumulative yielding the test material experiences localized bursting associated with a reduction of r.m.s value. Titanium alloy is relatively low strain hardening material; hence it experiences higher order strain before failure. During stressing the test material experiences straining of localized lump of material after certain order of strain, this may experience discrete burst/fissures resulting the reduced acoustic emission i.e. Titanium alloy under tensile loading experiences localized yielding and burst depending upon the load sequence, till failure, of course with continuous reduction in r.m.s value.

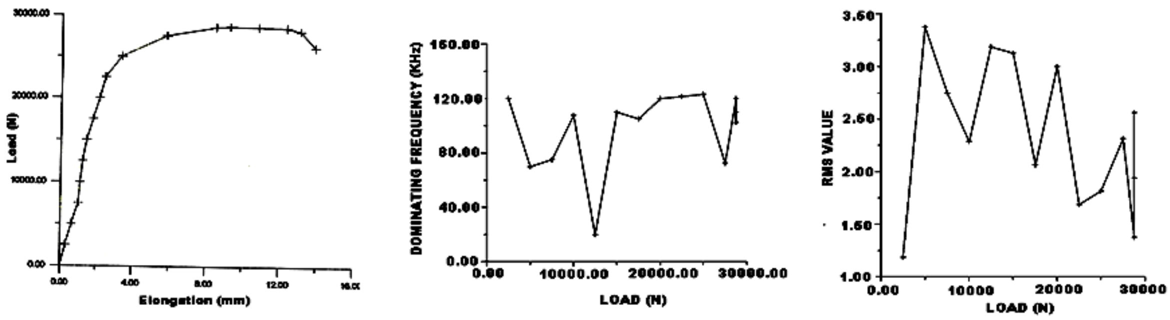
From the recorded raw acoustic emission signal, the dominant frequency was noted. Typical variation of the domination frequency with load is illustrated in fig 4. Referring to the illustration on r.m.s value and dominant frequency, it can be seen that especially with higher testing load, there is a reduction in the r.m.s value of the acquired acoustic emission signal; the signal acquires for higher loads, indicate the dominant frequency of 120kHz i.e. around 20000N of loading, the material exhibits more of burst emission, indicating there by the occurrence of localized cracking associated distressing the material.

Observation on characteristics of raw signal acquired. Typical illustrations on the acoustic emission signal acquired with different applied load are presented in fig 5. With applied load of 250 the acoustic emission signal comprises many different frequency components, with the dominant around 120KHz. As the load is increased to 5000N, only few frequency components were observed with a shift in the dominant frequency towards a lower magnitude. This indicates the occurrence of a relatively more continuous emission (also indicated by occurrence of higher r.m.s value). A summary of observations is indicated below.

7500	Higher power, high amplitude over 75 KHz frequencies, continuous mode, mixed mode of emission-associated with localized burst emission.
10000N	Similar trend continued.
12500N	Occurrence of dominant low frequency, continuous mode of emission has indicated by raise in rms value.
15000N	Reduced power-mixed mode of emission
17500N	Increased power (75KHz, 110-120kHz), mixed mode of emission.
20000N	Mixed mode, increased mode of signal emission.
22500N	Reduction in power, light shift in peak frequency, mixed mode of emission.
25000N	Reduction in power, light shift in peak frequency, mixed mode of emission.
27000N	Increased power, shift in peak frequency, increasing order of rms
26000N	Reduced power, failure.

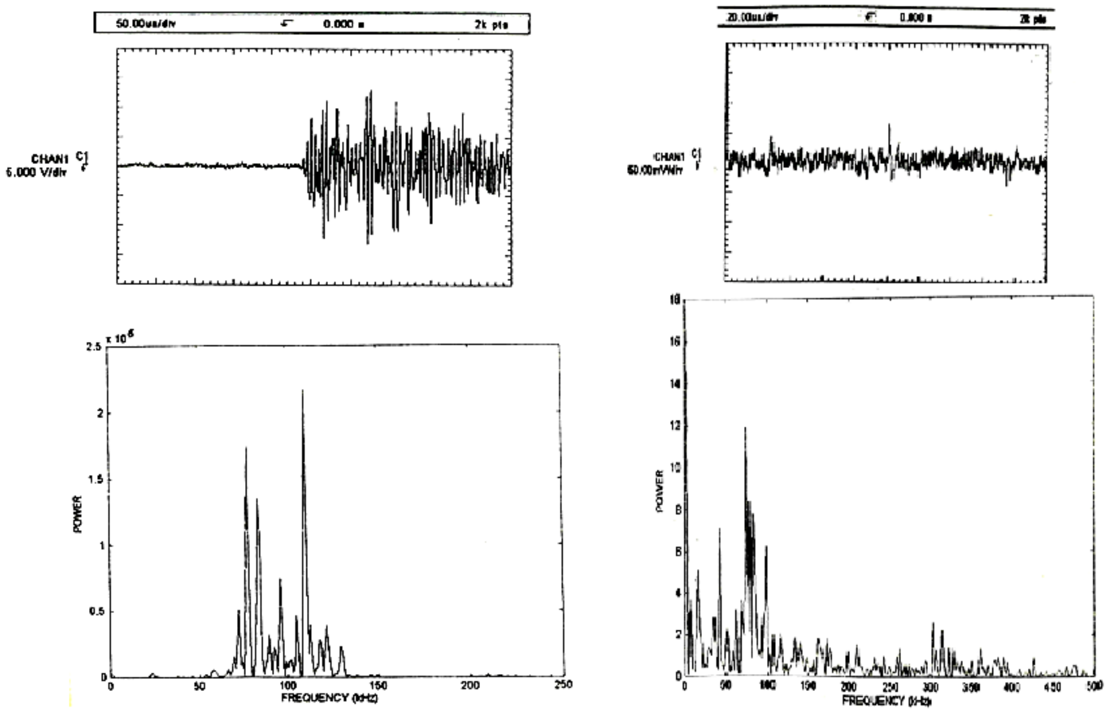
Summing up; the continuous monitoring of AE has illustrated the deformation of the material, occurrence of local fissures during early phase of loading and continuous deformation as illustrated by higher r.m.s value; further, barring few load stage, the monitored AE signal contains a dominant frequency around 100-120kH. This can be the typical frequency of acoustic emission for titanium alloy tested.

Fractured surfaces of test samples were observed through JEOL make Scanning Electron Microscope. The typical Scanning Electron Microscope of fractured surface is shown in Fig.5.16 (a-h). The higher ductility of titanium alloy is clearly illustrated through a textured macrograph, with elongated grains and occurrence of dimpled zones. Further, the flow of material around the dimples indicated that the material has undergone viscous yielding during tensile loading. Closer observation of localized regimes indicate possibility of failure initiation around spherical, second phase particles. Observations also indicate the occurrence of the cracking of the material prior to failure.



(a) (b) (c)

Figure 4: (a) Load Extension Characteristic (b) Variation of the R.M.S value with the load (c) Variation of the Dominating frequency with the load



(a) (b)

Figure 5: (a) FFT Spectra for power vs. frequency at 15000 N Applied load for ambient Condition and (b) (a) FFT Spectra for power vs. frequency at 15000 N Applied load for Cryogenic Condition

6.0 CONCLUSION

Study on AE response of titanium alloy, has been carried out with a view to develop an integrity evaluation methodology applicable to aerospace material. Necessary criteria has been evolved and applied to aerospace related application for real time integrity monitoring. The following are the significant conclusions emerging from the studies.

Observation on the tensile specimens bearing possible surface defects indicated that the AE response of titanium alloy subjected to tensile under ambient condition. The acoustic emission acquisition data indicated mixed mode of signal emission ambient condition. This might be due to occurrence of local fissure during early phase of loading and continuous deformation illustrated by high r.m.s value. Further barring few load stages, the monitored AE signal contains a dominating frequency around 100-120 kHz. This can be a typical frequency of acoustic emission of the titanium alloy tested.

REFERENCES

[1] Bakuckas, J.G.Jr. Prosser. W.H., Johnson W.S., "Monitoring damage growth in titanium matrix composites using AE". Journal of composite materials V28 n4, 1994 pp 305 – 328.

[2] Hlaws D. Timmerhaus and Thomas M. Flynn "Cryogenic process engineering". The international cryogenic monographic series 1989 pp 1-37, 38 &39.

Cite this article as: Dr.S.Sundaram, G. Vetri Chelvan. "Titanium Alloy Subjected to Tensile testing under Ambient and Cryogenic Conditions using Acoustic Emission Techniques." *International Conference on Information Engineering, Management and Security (2015)*: 179-185. Print.

[3] Krishnamurthy .R :Acoustic emission for condition monitoring in manufacturing "14th Tribology and condition monitoring proceedings of the national seminar" (June 19,1993) pp- 125- 137.
[4] Krishnamurthy.R , Chelladurai. T and Acharya .A.R "An approach for the integrity assessment of M250 maraging steel pressurized systems". Journal of Acoustic emission vol 8 No 1-2 (January- June 1989) pp S88-S92.
[5] Mashino .S ; Mashimo .Y , Horiya .T, Shiwa .M, T , "Analysis of microfracture mechanism of Titanium alloy by Acoustic emission technique". Material science and Engg.A; Structural materials; Properties ; V A213 n-1-2,1996, pp 66-70.
[6] Matthew J Donachie, JR, Compiled by "Titanium and Titanium Alloy" American Society for Metals Source Book 1982. PP – 3-19 and pp 100-139.
[7] Roman .I , Ward C.H , "AE characterization of the deformation and fracture of a Ti / 341 alloy" . Scripta Metallurgica et Materialia vol 27, n4 Aug 15, 1992. pp 413-418.
[8] Randall F. Barron, "Cryogenic systems". Second edn. 1985 pp 3-8, 43-55 and 591-623.
[9] Schneider, D., Ollendorf, H., Schwarz, T, "Non-destructive evaluation of the mechanical behaviour of TiN- coated steels by laser-induced ultrasonic surface waves:. Material science and processing v 61 n3 Sep 1995, Pp 277-284.
[10] Sieniawski, Jan, Gieron Miroslaw, Siaja, Waldemar., "The application of resistance and AE measurement in fatigue tests of two phase alpha plus beta Titanium alloy". Journal of material processing Technology v 53, n 1-2, Aug 1995, pp 363-372.
[11] Shaw, L., Miracle D., "On the relationship between microstructure and acoustic emission in Ti 6A1 -4v". Journal of Material science v 30, n 17, September 1995, pp 4286-4298.
[12] Soboyejo, W.O.Ramasundaram, P, Rabeen B, Parks, J.W., "Investigation of the effects of matrix microstructure and interfacial properties of the fatigue and behaviour of a meta stable titanium matrix composite". American society of Mechanical Engineers, Applied Mechanics Division, ADM, V117 (1993) pp 33-44.
[13] Volker Hanel and Wolfgang Thelen "Monitoring Screws under Tensile load using AE Analysis". IEEE Transactions of Instrumentation and Measurement, Vol. 45, No.2, April 1996. pp 547/555.
[14] Kulandaivelu.P., Sundaram S and Senthil Kumar P., " Wear monitoring of single point cutting tool using Acoustic Emission Techniques", Sadhana – Academy Proceeding in Engineering Science,V38, Part 2, (2013)pp 211 – 234.

Cite this article as: Dr.S.Sundaram, G. Vetri Chelvan. "Titanium Alloy Subjected to Tensile testing under Ambient and Cryogenic Conditions using Acoustic Emission Techniques." *International Conference on Information Engineering, Management and Security (2015)*: 179-185. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS030		**eAID**	ICIEMS.2015.030

Akkhara-Muni: An instance for classifying PALI characters

Neha Gautam, R. S. Sharma, Garima Hazrati
Department of Computer Science Engineering, UCE, RTU

Abstract: *Handwritten Recognition and Archaeology are significant facts of antediluvian epoch and scripts which are not tranquil to learn. For example, Pali written in Sinhala, Khmer, Burmese, Devanagari, Lao and many more which are having prodigious influence in the Buddhism culture since they hoard the lessons. Akkhara-Muni which constitutes Pali alphabet recognition been presented in this paper done in a few steps trailed by OCR for classification with the results that portray need of ancient scripts to be recognized and classification accuracy leads to 85.4%.*

Keywords: Handwritten Recognition, Archaeology, Pali, OCR, Akkhara-Muni

1. Introduction

Man and machine are the two wheels of chariot. Humans need to get command on the machine in order to move with the wind. Keyboard is one of the most used devic-es now-a-days by every sphere of people from child to grandparents, student to high officials and also from hardware to virtual type. Handwritten is that field of image processing which is growing on a large scale to accomplish the requirements of global relations. Handwritten recognition is challenging though it is still a time and cost-saving work if done in daily activities and one of the most important uses of this technique is in archaeological department to learn ancient scripts written thousands of years ago. Scripts are defined as "markers" of a civilization keeping the record of lives in picto-grams to phonograms as they are hoping to preserve heritage, culture, and history of the region understanding their importance [1]. A lot of work is done on the western scripts whereas eastern scripts are not having a notified work till yet and to learn about civilization history one needs to know about early writings. Optical character recognition (OCR) - a well-known technique with some modification that is used to recognize akkhara-muni (Pali) providing help to the archaeologists so that they can understand Buddhism and their teachings written in ancient period offering new thoughts to youth as past always pace an innovative way towards development.

1.1 Pali:

Language of prehistoric period- PALI has been inscribed in many scripts and lan-guages besides being popular in southern countries for teachings of Buddhism. Pali alphabet basically consists of 41 letters: 6 vowels, 2 diphthongs, 32 consonants and 1 accessory nasal sound called as nigahitta [2]. Consonants are divided into 25 mutes, 6 semi-vowels, 1 sibilant and 1 aspirant followed by vowels comprises of

Cite this article as: Neha Gautam, R. S. Sharma, Garima Hazrati. "Akkhara-Muni: An instance for classifying PALI characters." *International Conference on Information Engineering, Management and Security (2015)*: 186-188. Print.

long and short. Ian James, developer of a modified Latin font, and forms get their existence from ancient Brahmi and Pallava (ancestors of the Indic scripts) and later on named as Akkhara Muni, (Letters of the Sage). In Sri Lanka, Pali was used not only for the writing of Buddhist scriptures, but also to record the history of the country [3,4].

2 Related Work

Several other works were done by OCR on many other languages and scripts giving top rated accuracies setting standard for others. In 2004 U. Pal and B.B. Chaudhari [5] describes work done on 12 major Indian scripts through OCR. In 2007 V.N. Manjunath Aradhya , G. Hemantha Kumar, S. Noushath [6] show a multi-lingual OCR leading to a good accuracy. Afterwards, Apurva A. Desai [7] use OCR tech-nique for recognizing gujrati handwritten digits in 2010 resulting 82% success rate and likewise Amit Choudhary, Rahul Mishra, Savita Ahlawat [8] routine this with binarization method to judge capability of OCR for english cahartces in 2013 giving accuracy of 85.62%.
Pali characters recognition using devnagri was shown by kiran s mantri, s p ramteke, s r suralkar[7] in 2012, by comprises following features like image pre-processing, feature extraction and classification algorithms that have been traversed to design software (OCR) with high performance. The recognition rate is 100% that has been done using simple feed forward multilayer perceptions also proposes a back propagation learning algorithm that is used to guide each network with the characters in that particular group.
Another work proposes a recognition system that has taken Pali cards of Bud-dhadasa Indapanno was presented by Tanasanee Phienthrakuland and Wanwisa Chevakulmongkol [8] in 2013. Its handwritten images have been refined by contrast adjusting, grayscale converting and noise removing. Basically the features of every single character are removed by the zoning method where average of all accuracies considered in groups comes out to be approximately 81.73%.

3 Our Work

Optical Character Recognition (OCR) attempted first in 1870, developing from era of 40's while transforming from first generation to third now in present with a wide variety of applications in numerous fields from banking to education, archaeology to space science. Entire work is done on local database of Akkhara-muni script collected online by numerous resources and is shown as:

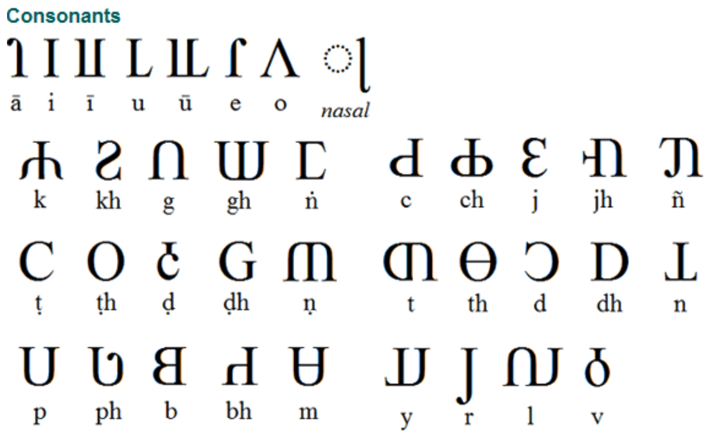


Fig. 1. Akkhara-Muni- A PALI Alphabet (Src. [2])

OCR is combination of several components illustrated in figure below:

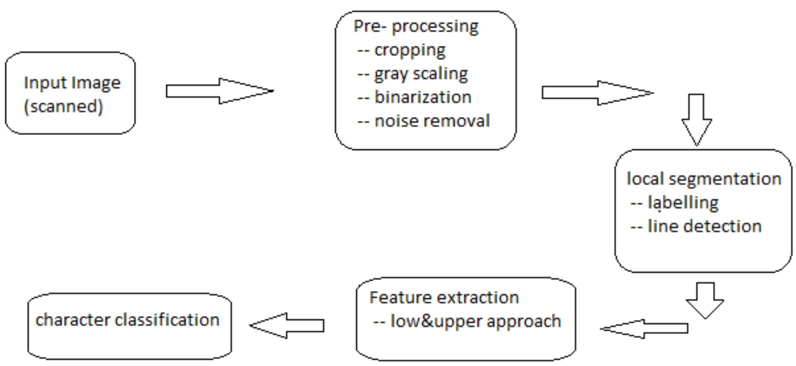


Fig. 2. Projected Accost for classification

Cite this article as: Neha Gautam, R. S. Sharma, Garima Hazrati. "Akkhara-Muni: An instance for classifying PALI characters." *International Conference on Information Engineering, Management and Security (2015)*: 186-188. Print.

Scanned input image is pre-processed by cropping it to let it fitted in proper dimen-sions trailed by gray scaling then binarization with noise removal by using various MATLAB functions. Subsequently local segmentation of pre-processed image is done by labeling and line detection. Later on, feature extraction is done using lower and upper approach resulting in character recognition with 85.4% success rate.

4 Results

The overhead approach was tested on 60 sets of digits and samples used over here for experiments are collected online and trained in MATLAB. In all, we can say that this approach gave 85.4% success rate. The recognition rate for 2 vowels is high and 2 consonants are low as they are misidentified. The outcome we have obtained here is not very good if only OCR is taken but as script taken here akkhra-muni and no trained data sets available though we can't find any other work for comparison there-fore we think that our result is good in this zone.

5 Conclusion:

In this approach Optical character recognition (OCR) is proposed with some other techniques for classification of Akkhara-Muni. Techniques like optical scanning, binarization, segmentation and feature extraction followed by classification of charac-ters. The overall performance is 85.4%, but it can't be considered as a benchmark. Any model of classification is based on their feature extractions which are also need-ed to be used here for improving is performance. This attempt is unique as a whole. It offers a great value for scripts that are needed to be recognized and there is not much research done in this field of ancient scripts. We are currently working for other an-cient Indic script.

REFERENCES

[1] www.ancient.eu
[2] "An Elementary Pali Course" book by Ven Narada, Thera pg. no. 9-12
[3] www.skyknowledge.com
[4] www.ancient.eu/www.ancientscripts.com
[5] U. Pal, B.B. Chaudhari, "Indian scripts character recognition: A Survey" Pattern Recogni-tion Society. Published by Elsevier Ltd 2004
[6] V.N. Manjunath Aradhya , G. Hemantha Kumar, S. Noushath, "Multilingual OCR system for South Indian scripts and English documents: An approach based on Fourier transform and principal component analysis" Engineering Applications Artificial Intelligence Else-vier Ltd 2007
[7] Apurva A. Desai, "Gujrati handwritten OCR through neural network" Pattern Recognition Elsevier Ltd. 2010
[8] Amit Choudhary, Rahul Mishra, Savita Ahlawat, "Off- line handwritten character recogni-tion using feature extracted from binarization technique" The Authors. Published by Else-vier B.V 2013
[9] Kiran S. Mantri, R. S. Ramteke, S. R. Suralkar, "Pali Character Recognition System" at IJAIR 2012 ISSN: 2278-7844
[10] Tanasanee Phienthrakuland Wanwisa Chevakulmongkol, "Handwritten Recognition on Pali Cards of Buddhadasa Indapanno" International Computer Science and Engineering Conference (ICSEC): ICSEC 2013

Cite this article as: Neha Gautam, R. S. Sharma, Garima Hazrati. "Akkhara-Muni: An instance for classifying PALI characters." *International Conference on Information Engineering, Management and Security (2015)*: 186-188. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	VOL	01
Website	www.iciems.in	eMail	iciems@asdf.res.in
Received	10 - July - 2015	Accepted	31- July - 2015
Article ID	ICIEMS031	eAID	ICIEMS.2015.031

Software Application Generator: An ER Model-based Software Product Building Tool

Mr. Souradeep Sarkar[1], Mr. Debasish Hati[2], Mr. Prasun Kumar Mitra[3]
[1,2,3] Lecturer, CST Department, Technique Polytechnic Institute,
Hooghly,WB-712102,India

Abstract: *Software Application Generator (SAG) is a very powerful software product building tool. It helps developer to build an entity relationship model based software product using modern web technology. ER model is one of the most popular methodologies for designing relational database. Asp.net, JSP and PHP are most popular technology for developing web application. Several commercial products have been developed to support ER-model and those web technologies. Inspired by these technologies, we have developed an educational prototype SAG that supports the following task:*

- *Drawing ER diagram visually and translate it to relational database schema automatically.*
- *Develop different type of template easily from ER model.*
- *Develop skeleton code of selective technology automatically.*
- *Develop skeleton test plan automatically.*
- *Deploy generated code to corresponding environment automatically.*

In this paper, we describe the architecture of SAG and its implementation details to illustrate how such tool can be developed.

Keywords: SAG, ER modeler, Application Composer, Application Generator, Test plan Generator, Application Deployment.

I. INTRODUCTION

SAG is very powerful tool for developing a software product based on ER models and web technologies. It contains five non overlapping modules: ER modeler, Application Composer, Application Generator, Test plan Generator and Application Deployment. It is implemented in java, therefore it can be executed in all environment when a java virtual machine available.

A. ER Modeler: This module is used to design ER diagram and translate it to a relational database schema. It contains two essential components: entity and relationship. Entity represents object that are involved in enterprise such as student, professor and course in a university. Relationship represents the associations among these objects such as student takes course. ER model supports the following set of attractive features:

- Representing ER-Diagram in Binary Repository: We have defined binary repository to convert an ER diagram in binary repository format and vice versa.

Cite this article as: Souradeep Sarkar, Debasish Hati, Prasun Kumar Mitra. "Software Application Generator: An ER Model-based Software Product Building Tool." *International Conference on Information Engineering, Management and Security (2015)*: 189-194. Print.

- Draw ER-Diagram semantically: It recognizes ER diagram components such entity sets and relationship sets semantically. This greatly facilitates the diagram to change the layout of an ER diagrams.
- Validation Verification: It supports the verification of validation of an ER diagram and ensures that only well formed ER diagram can be exported in binary file and translated to relational database schema.
- Automatic translation to relational models: After an ER diagram is created; one can simply translate the ER diagram to a relational database schema.

B. Application Composer:

Application Composer module is used to develop template which contain elements of web forms. We consider five types of templates: Login, Role selection, Maintain, Master detail and Associate. Login template is a general purpose template which mainly use for user authentication. Login template contains one textbox (for user name), one hidden textbox (for password), one button (for submit) and one label (for showing error message) elements. Role-Selection template is used for select the role of user. Role-Selection template contains two textboxes (for user ID and user name), one Drop Down List (for role of user), two buttons (for submit and exit) and one label (for showing error message). Maintain template is mainly used for perform basic operations like insert, update, delete and view. Maintain template contain two type of elements: entity element and form element. Entity element contains a textbox and a label for each attribute of entity of ER model which will be selected by developer. Form element contains one datagrid (for display table value), four buttons with form option true (for insert, update, delete and view operation), one button (for exit) element. Master detail template is used for update details of master. The criteria for build master detail template is that there will be two entity and a relationship between them and a common attribute among them. Master detail template contains two type of elements: entity element and form element. Entity element contains a textbox and a label for each attribute of target entity of ER model which will be selected by developer. Form element contains one datagrid (for display table value), five buttons (for submit, add to grid, delete from grid, ok and cancel) elements. Associate template is used for update the associate table from original table. The criteria for build associate template is that there will be three entities and two relationships between them and two commoun attributes for each of two end entities with target entities. Associate template contain two type of elements: entity element and form element. Entity element contain a textbox and a label for each attribute of target entity of ER model which will be selected by developer. Form element contain two listboxes (for contain the value of original and associate table), four buttons (for add to list,remove from list,ok and cancel) elements. Application Composer supports the following set of attractive features:

- Representing Template in XML file: We have defined a XML schema to convert template into xml format and vice versa.
- Design Template Semantically: Application Composer recognize template component such as entity elements and form elements. Entity element will be derived from the entity of ER model which are build by ER modeler. Form elements are general purpose web form elements, which will be selected by developer. This greatly facilitates the design to change the layout of a template.
- Validity Verification: Application Composer supports the verification of validity of a template and ensures that only well form template can be exported in xml file.

C. Application Generator:

Application Generator is used to generate code. Befor generate code, first select xml file which contains all templates. This xml file is constructed by Application Composer. Then select the web technology in which envornment the code will be generated. In ASP technology, we generate one .aspx and one .aspx.vb file for each login, role selection, master detail and associate template. For maintain template five .aspx and five .aspx.vb for main, insert, update, delete, view page. For menu page one .aspx and one .aspx.vb file will be generated. One .vb file will be generated for database connection class in App_Code folder. One configaration file will be generated by default.

D. Testplan Generator:

Testplan Generator is used to generate testplan. Before generate testplan, select xml file which contain all templates. This xml file is constructed by Application Composer. It generate one each document that contains one table for template or menu page. One row of table contain a test case for each element of template. It maintain the standard format of test plan.

E. Application Deployer:

Applicatin Deployer is used to deploy the code which is generated by Application Generator. It helps the user to deploy the code automatically. First user select the code which will be deployed. Next they select the server in which the deployment will be done. Then the deployment process will be done and the browser will be opened with default url.

Cite this article as: Souradeep Sarkar, Debasish Hati, Prasun Kumar Mitra. "Software Application Generator: An ER Model-based Software Product Building Tool." *International Conference on Information Engineering, Management and Security (2015)*: 189-194. Print.

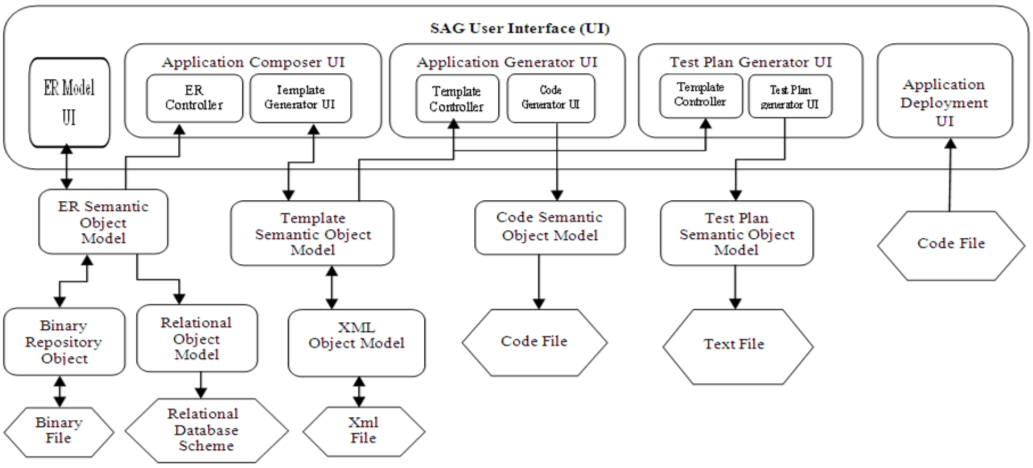


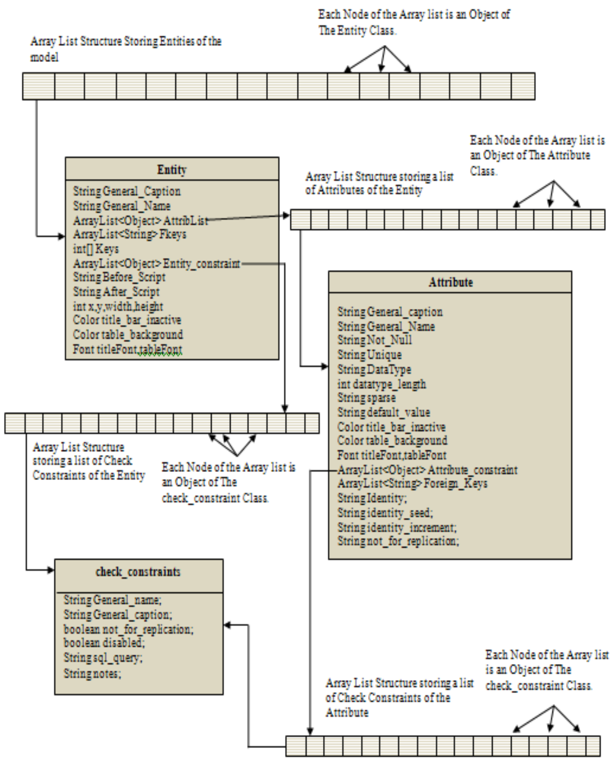
Figure 1. System Architecture of SAG Tool

II. SYSTEM ARCHITECTURE

The overal system architecture of SAG Tool is illustrated in figure 1. Basically, it consists of the following modules.

- **ER Model User Interface:** It provides a user friendly graphical user interface to support the interaction between ER diagram designers and ER Modeler.
- **ER Semantic Object Model:** It is the key internal data structure that represents the complete semantic information of an ER diagram. An ER sementic object model is created either for ER diagram or from a binary file.

Application Composer User Interface: It provides a user friendly user interface to support the interaction between developer and application composer. It is divided into two sub modules : ER Controller and Template Generator User Interface. ER Controller is used to manipulate the existing ER diagram.Template Generator User Interface is used to manipulate the templates of application.



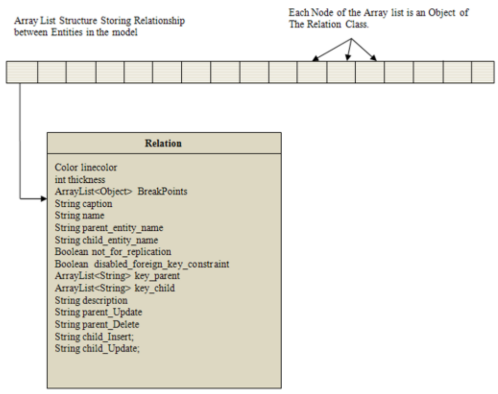


Figure 3. Local Data Structure for storing relationship collection.

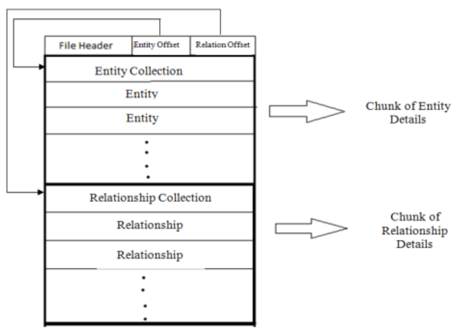


Figure 2. Local Data Structure for storing entity collection.

Figure 4. Data structure of storing the ER Model

Cite this article as: Souradeep Sarkar, Debasish Hati, Prasun Kumar Mitra. "Software Application Generator: An ER Model-based Software Product Building Tool." *International Conference on Information Engineering, Management and Security (2015)*: 189-194. Print.

- **Template Semantic Object Model:** It is the key internal data structure that represents the complete sementic information of templates of application. Template semantic object model is created either from application composer or from an XML file.
- **XML Object Model:** It is the intermediate data structure that map sql and Template semantic object model to XML file. It is used to generate a XML file or is constructed from a XML file.
- **Application Generator User Interface:** It provides a user friendly user interface to support the interaction between developer and application generator. It is divided into two sub modules : Template Controller and Code Generator User Interface. Template Controller is used to manipulate the existing templates of application. Code Generator User Interface is used to manipulate the codes of application.
- **Code Semantic Object Model:** It is the key internal data structure that represents the complete semantic information of codes of application. Code Semantic Object Model is created from application generator. It is used to generate a code file.
- **Testplan Generator User Interface:** It provides a user friendly user interface to support the interaction between developer and testplan generator. It is divided into two sub modules : Template Controller and Testplan User Interface. Template Controller is user to manipulate the existing templates of application. Testplan Generator User Interface is used to manipulate the testplans of application.
- **Testplan Semantic Object Model:** It is the key internal data structure that represents the complete semantic information of testplans of application. Testplan Semantic Object Model is created from testplan generator. It is used to generate a text file.
- **Application Deployment User Interface:** It provies a user friendly user interface to support the interaction between deployer and Application Deployment.

II. IMPLEMENTATION

A. ER Modeler:

To support interoperability between different ER diagram we have defined data structure to convert an ER diagram into memory object. The data structure of entity and relationship is shown in Figure 2 and Figure 3 respectively. The data structure of storing the ER Model is shown in Figure 4.
To translate an ER diagram to a relational database schema, ER Modeler follow the following steps:

a) For each entity set A implement a table using CREATE table SQL command.
b) For each relationship implement a ALTER table SQL command to set foreign key concept.

B. Application Composer:

To support interoperability between different Template generating tools, we have defined a XML schema defination to templates of application into XML file and vice versa. The following features is maintained by the application composer.

- Login and role selection template is general pupose template. It is fixed for all applications.
- Each Maintain template will be generated for each Entity.
- Each Master detail template will be generated for two entities and a relationship between them.
- Each Associate template will be generated for three entities and two relationships between them.

C. Application Generator:

To support interoperability between different Code generation tools, we have defined a data structure to convert templates of application into corresponding code. The data structure is shown in Figure 5.To generate code the following points will be considerd

- Code will be generated according to selected technology.
- Code will be generated for each template.
- Code for menu template will be generated by default.
- Code will maintained standard skeleton. Programmer can run this code in any environment by minimum change.
- Some additional code file will be generated by default depending on web technology.

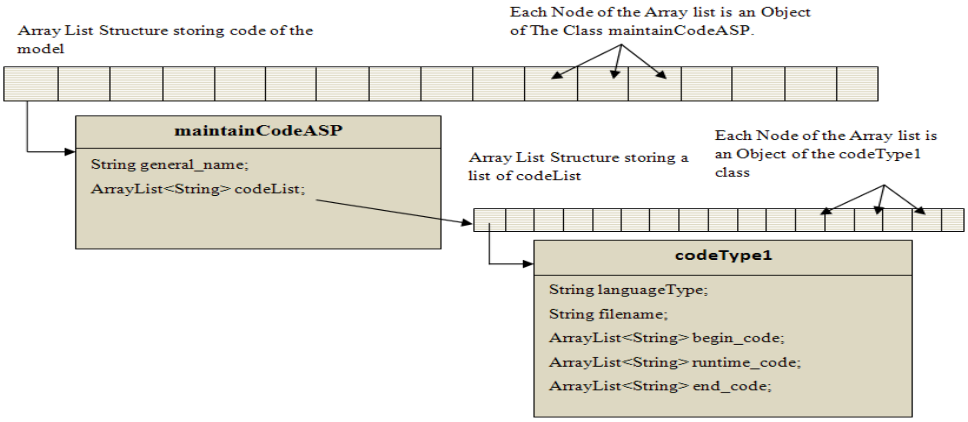


Figure 5. Local Data Structure for storing code of application .

D. Testplan Generator:

To support interoperability between different Testplan Generator tools, we have defined a data structure to convert templates of application into corresponding testplan. The data structure is shown in Figure 6.To generate testplan the following points will be considerd

- Testplan will be generated for each template.
- Testplan for menu template will be generated by default.
- Testplan will maintained standard skeleton. Developer can generate final testplan by minimum change.

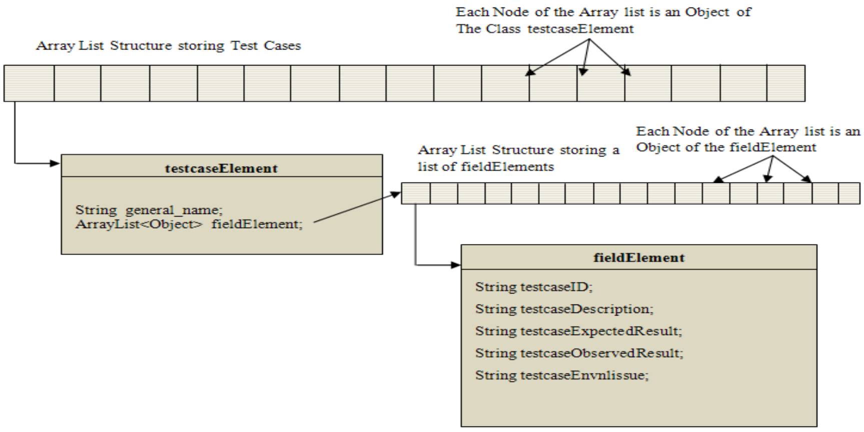


Figure 6. Local Data Structure for storing Testcase of Application .

E.Application Deployer:

Application Deployer is a user friendly interface. It guides user to deploy the application in selected web server environment. All the server setup is previously completed. Here our application code file will be posted in the default location of selected environment and then the application willl be run with home page. The following sets will be consider in application deployer:

- Select the Server Environment.
- Select the location where the application file will be exists. All the files of the application will be listed in the application file list box.
- Select the files which will be deployed and bring into deploy file list box.
- Press the deployment button.

III. CONCLUSION

We have developed an ER-model and web technology based software application product generating tool for educational purposes. This tool incorporates object oriented technolgoy, XML. The verification process guarantee the semantic correctness for ER diagrams. The automatic translation from ER- diagrams to relational data schemas is practically useful. The automatic generation of template give relief from the time consuming form design task. The automatic generation of code give relief from the complex and time consuming task. Using this tools developer generate a software product in five minutes using different technology. The automatic testplan generator helps the tester to test this product according to all testcase. It helps the developer to design the testplan and guarantee that all test cases will be considerd. The automatic deployment helps the tester to test the product is run correctly in server environment in short time.

Cite this article as: Souradeep Sarkar, Debasish Hati, Prasun Kumar Mitra. “Software Application Generator: An ER Model-based Software Product Building Tool.” *International Conference on Information Engineering, Management and Security (2015)*: 189-194. Print.

V. OUR FUTURE WORK

We are extending ER modeler to support Complex type of ER diagram. We are currently extending Application Composer to support other complex type of templates such as tree template. We are exteding Application Generator to support more web technology language and different database server environment. We are extending Testplan Generator to support black box testing and white box testing test plan. We are extending Application Deployer to support the application in Cloud computing environment.

REFERENCES

[1] Florescu, D. & Kossmann, D. (1999), Storing and Querying XML Data Using an RDBMS, IEEE Data Engineering Bulletin (22), 1999, 27-34

[2] Fernandez, F., Tan, C. & Suciu, D. (2000), SilkRoute, Trading between Relations and XML, Proc. 9th International World Wide Web Conference, Netherlands, 2000, 723-745

[3] Kleiner, C. & Lipeck, U. (2001), Automatic Generation of XML DTDs from Conceptual Database Schemas, Proc.Workshop of the Annual Conference of the German and Austrian Computer Societies, Austria, 2001, 396-405

[4] Shanmugasundaram, J.,Tufte, K., He, G., Zhang, C., DeWitt, D. & Naughton, J. (1999), Relational Databases for Querying XML Documents: Limitations and Opportunities, Proc. International Conference on Very Large Data Bases (VLDB), Scotland, 1999, 302-314

[5] Lee, D., Mani, M. & Chu, W. (2003), Schema Conversion Methods Between XML and Relational Models, Knowledge Transformation for the Semantic Web, IOS Press, Netherlands, 2003, 1-17

[6] A Practical Approach for Automated Test Case Generation. Published in: Computer Software Applications Conference, 2006. COMPSAC '06. 30th Annual International, Page(s): 183 – 188. IEEE Computer Society.

[7] Tallman, Owen H. Project Gabriel: Automated Software Deployment in a Large Commercial Network. Digital Technical Journal Vol. 7 No. 2, 1995, pp56-70.

[8] Paul E. Ammann, Paul E. Black, and WilliamMajurski. Using Model Checking to Generate Tests from Specifications. In Proceedings of the Second IEEE International Conference on Formal Engineering Methods (ICFEM'98), pages 46–54.IEEE Computer Society, 1998

[9] Kim B. Bruce. Foundations of Object-Oriented Languages: Types and Semantics. MIT Press, 2002.

[10] Luciano Baresi and Mauro Pezz`e. An introduction to software testing. Electr. Notes Theor. Comput. Sci.,148(1): 89–111, 2006.

Cite this article as: Souradeep Sarkar, Debasish Hati, Prasun Kumar Mitra. "Software Application Generator: An ER Model-based Software Product Building Tool." *International Conference on Information Engineering, Management and Security (2015)*: 189-194. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS032

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.032

Application of Color Segregation in Visual Cryptography using Halftone Technique and RGB Color Model

Mr. Prasun Kumar Mitra[1], Mr. Souradeep Sarkar[2], Mr. Debasish Hati[3]
[1,2,3]Lecturer, CST Department, Technique Polytechnic Institute
Hooghly,WB-712102,India

Abstract: *Visual Cryptography is a special encryption technique to hide information in images in such a way that it can be decrypted by the human vision if the correct key image is used. This experiment describes a secret visual cryptography scheme for color images based on halftone technique. Firstly, a chromatic image is decomposed into three monochromatic images in tones of Red, Green and Blue. Secondly, these three images are transformed into binary images by halftone technique. Finally, the traditional binary secret sharing scheme is used to get the sharing images. This scheme provides a more efficient way to hide natural images in different shares. Furthermore, the size of the shares does not vary when the number of colors appearing in the secret image differs.*

Keywords: Visual cryptography, secret sharing, color image, halftone technique.

I. INTRODUCTION

Visual cryptography is a cryptographic technique which allows visual information (pictures, text etc.) to be encrypted in such a way that the decryption can be performed by the human visual system without the aid of computers. As network technology has been greatly advanced, much information is transmitted via the Internet conveniently and rapidly. At the same time, the security issue is a crucial problem in the transmission process. For example, the information may be intercepted from transmission process. This method aims to build a cryptosystem that would be able to encrypt any image in any standard format, so that the encrypted image when perceived by the naked eye or intercepted by any person with malicious intentions during the time of transmission of the image is unable to decipher the image.

Visual Cryptography Scheme (VCS), introduced by Naor and Shamir in 1994, is a type of secret sharing techniques for images. The idea of VCS is to split an image into a collection of random shares (printed on transparencies) which separately reveal no information about the original secret image other than the size of it. The image is composed of black and white pixels, and can be recovered by superimposing a threshold number of shares without any computation involved. Here is an example using a dithered black-and-white Lena image as the original secret image (Fig. 1).

Cite this article as: Prasun Kumar Mitra, Souradeep Sarkar, Debasish Hati. "Application of Color Segregation in Visual Cryptography using Halftone Technique and RGB Color Model." *International Conference on Information Engineering, Management and Security (2015)*: 195-198. Print.


Fig. 1. Original Secret Image


Fig. 2. Dithered

By applying the Naor-Shamir 2-out-of-2 visual cryptography algorithm, two shares (printed on transparencies) are created, which separately reveal no information about the original image. It can only be recovered when both of the shares are obtained and superimposed. Fig. 3 shows the two shares and the superimposition of them. Note that the size of the images is expanded by a factor of 4.

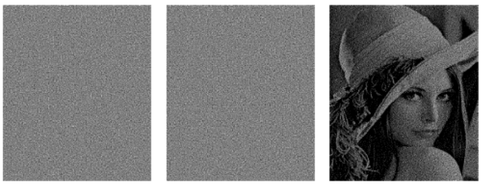


Fig. 3. Two Shares and the Superimposition of the Shares

The technology makes use of the human vision system to perform the OR logical operation on the superimposed pixels of the shares. When the pixels are small enough and packed in high density, the human vision system will average out the colors of surrounding pixels and produce a smoothed mental image in a human's mind. For example, a block of 2 × 2 pixels shown below will be viewed as a gray-like dot as the two black pixels and the two nearby white pixels are averaged out. If we print the 2×2 pixel blocks shown in Fig. 4 separately onto two transparencies and superimpose them. This effect is equivalent to performing a pixel-wise OR logical operation on each of the four pairs of pixels between these two transparencies. The result is shown in Fig. 5. One of the unique and desirable properties of VCS is that the secret recovery process can easily be carried out by superimposing a number of shares (i.e. transparencies) without requiring any computation.

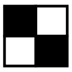


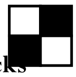
Fig. 4. Two 2 × 2 pixel blocks



Fig. 5. Superimposed Image

Besides black-and-white images, a natural extension of this research problem is to perform secret sharing on color images. Hou proposed three VCS for color images. Among them, the first one uses four shares to split a secret image. The four shares are called black mask, C (Cyan) share, M (Magenta) share and Y (Yellow) share. This scheme reproduces the best quality among the three in terms of image contrast during secret image recovery process. It is also the only one supporting a practically useful feature called two-level security control. This feature allows an authority to keep a particular share, the black mask, secret and release the other three shares to the public, without worrying about exposing the concealed image. In particular, the author claimed that this scheme is secure as long as the black mask is kept secret. There would have no information leaked even if all the other three shares, namely C, M, Y shares, are exposed regardless of the color composition of the original secret image.

Advantage of Visual Cryptography:

- Simple to implement.
- Encryption doesn't require any NP-Hard problem dependency.
- Decryption algorithm not required (Use a human Visual System).
- So a person unknown to cryptography can decrypt the message.
- We can send cipher text through FAX or E-MAIL.
- Infinite Computation Power can't predict the mess.

Cite this article as: Prasun Kumar Mitra, Souradeep Sarkar, Debasish Hati. "Application of Color Segregation in Visual Cryptography using Halftone Technique and RGB Color Model." *International Conference on Information Engineering, Management and Security (2015)*: 195-198. Print.

II. RELATED WORKS

The most traditional visual cryptography schemes are used for black and white images. Recently, some visual cryptography schemes for gray or color images have been proposed.

Verheul and Tilborg present a secret sharing scheme for images with c colors. The principle of this scheme is to transform one pixel of image to b sub-pixels, and each sub-pixel is divided into c color regions. In each sub-pixel, there is exactly one color region colored, and all the other color regions are black. The color of one pixel depends on the interrelations between the stacked sub-pixels. A major disadvantage of this scheme is that the number of colors and the number of sub-pixels determine the resolution of the revealed secret image. If the number of colors is large, coloring the sub-pixels will become a very difficult task.

Naor and Shamir propose a secret sharing scheme, which reconstructs a message with two colors by arranging the colored or transparent sub-pixels. Both approaches assign a color to a sub-pixel at a certain position, which means that displaying m colors uses m-1 sub-pixels. The resulting pixels contain one colored sub-pixel and the rest of the sub-pixels are black. Therefore the more colors are used; the worse the contrast of the images becomes significantly. Their approaches cannot be applied to the extended visual cryptography either. Rijmen and Preneel presented a scheme which enable multicolor with relatively less sub-pixels (24 colors with m = 4). However each sheet must contain color random images, which means applying this approach to the extended visual cryptography is impossible.

For this reason, Chang, Tsai and Chen recently proposed a new secret color image-sharing scheme based on the modified visual cryptography. In that scheme, through a predefined Color Index Table (CIT) and a few computations they can decode the secret image precisely. Using the concept of modified visual cryptography, the recovered secret image has the same resolution as the original secret image in their scheme. However, the number of sub-pixels in their scheme is also in proportion to the number of colors appearing in the secret image, i.e., the more colors the secret image has, the larger the shares will become. Another disadvantage is that additional space is needed to store the Color Index Table (CIT).

III. EXPERIMENTAL RESULTS

Our experiment is based on the RGB color model and the halftone technique. Firstly, a chromatic image is decomposed into three monochromatic images in tones of red, green and blue. Secondly, these three images are transformed into binary images by halftone technique. Finally, the traditional binary secret sharing scheme is used to get the sharing images. Halftone technique is a method to display a gray image with black-and-white spots. Figure 6 shows the basic principle of the halftone technique. The more black spots the image includes, the more the image will be alike the true gray image. Construct to the other two binary images shown in Fig. 6(c) and (d), Fig. 6(b) is closest to the true gray image.

(a) The true gray image (b) Binary image1 (c) Binary image2 (d) Binary image3

Fig. 6: Basic principle of halftone technique

In this paper, we use the Floyd-Steinberg Algorithm to get the halftone images. The algorithm is given below:

For an 8-bit gray scale image, the gray value of the image is from 0(black) to 255(white).

Letting b=0,
w=255,
t= int [(b+w)/2] =128.

Assuming g is the gray value of the image, which location is P (x, y), e is the difference between the computed value and the correct value. Then the Floyd-Steinberg Algorithm can be described as following:

```
If  g > t then
     print white;
     e=g-w;
else
    print black;
    e=g-b;
    (3/8 × e) is added to P (x+1, y);
    (3/8 × e) is added to P (x, y+1);
    (1/4 × e) is added to P (x+1, y+1);
End if
```

For example, a point with the gray value of 130 in an image should be gray point. Since the intensity of general image changes continuously, so the values of adjacent pixels are likely close to 130, and the surrounding region is also gray. According to the Algorithm, the number 130 is bigger than 128, then a white point is printed on the new image. But 130 are away from the real white 255. While -46 (-125 multiplied by 3/8) added to adjacent pixel, the value of adjacent pixel is close to 0; the adjacent pixel comes to black. Next time, e also become positive, the adjacent pixel comes to white, so a white one after a black one, gray is demonstrated. If not transmitting the error, the pixel in the new image is white. Take another example; if the gray value of a point is 250, it should be white in gray image, and e equals to -5, it has little impact on the adjacent pixel. This certificates the correctness of the algorithm. In

Cite this article as: Prasun Kumar Mitra, Souradeep Sarkar, Debasish Hati. "Application of Color Segregation in Visual Cryptography using Halftone Technique and RGB Color Model." *International Conference on Information Engineering, Management and Security (2015)*: 195-198. Print.

the experiment, First a color image is decomposed into three basic components R, G and B. Then the above Floyd-Steinberg Algorithm is used to get the halftone images of the corresponding components. After that we get the halftoned red, halftoned green and halftoned blue images.

Fig. 7 (a)

Fig. 7 (b)

Fig. 7 (c)

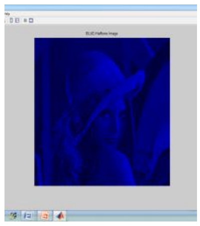


Fig. 7 (d)

Fig. 7 (a) Original Inputted Chromatic Image **Fig. 7 (b) Halftone Image of Red**
Fig. 7 (c) Halftone Image of Green **Fig. 7 (d) Halftone Image of Blue**

If we compose these three monochromatic images into a chromatic image, we can get the following image.

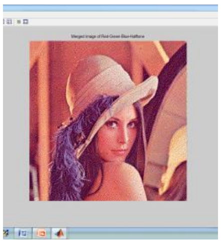


Fig. 7 (e) Merged image of R-G-B Halftones

We can consider every monochromatic image as a secret image and can use the traditional binary image-sharing scheme to divide it into three secret shares with same color, and then, we can choose any three different colors of which to compose them into three colored shares. The original secret information will be visible by stacking any 2 or 3 transparencies, but none secret information will be revealed by only one transparency.

IV. FUTURE SCOPE OF FURTHER IMPROVEMENT

Our future work is to generate shares in such a way so that it can be hidden within different cover images. It will look like some picture, not just a share. So, the original secret shares will be transmitted as hidden within different pictures. Finally, by super imposing these shares, the original secret image will be generated.

ACKNOWLEDGMENT

We would like to thank our respected Executive Director Mr.S.N.Basu ,Principal Dr. Abhijit Chakrabortty and Administration of Technique Polytechnic Institute for motivating us in this research work. We would also like to thank all the members of Technique Polytechnic Institute for their support and co-operation. We thank all mighty God and our parents for their blessings in our life.

REFERENCES

[1] R. Floyd and L. Steinberg, "An adaptive algorithm for spatial greyscale". Journal of the Society for Information Display. pp. 36-37, 1976.

[2] M. Mese and P. P. Vaidyanathan, "Optimized Halftoning Using Dot Diffusion and Methods for Inverse Halftoning". IEEE Trans. On image processing. Vol. 9, No. 4, pp. 691-709, 2000.

[3] Visual Cryptography and Its Applications by Jonathan Weir, WeiQi Yan & Ventus Publishing ApS : 2012.ISBN : 978-87-403-0126-7.

[4] Visual Cryptography and Secret Image Sharing by Stelvio Cimato and Ching-Nung Yang.CRC Press - 2012 by Taylor & Francis Group, LLC

Cite this article as: Prasun Kumar Mitra, Souradeep Sarkar, Debasish Hati. "Application of Color Segregation in Visual Cryptography using Halftone Technique and RGB Color Model." *International Conference on Information Engineering, Management and Security (2015)*: 195-198. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS033		**eAID**	ICIEMS.2015.033

Optimization of the critical loop in Renormalization CABAC decoder

Karthikeyan.C[1], Dr.Rangachar[2]
[1]Assistant Professor, ECE Department, MNM Jain Engg.College, Chennai
[1]Research scholar, Hindustan University, Chennai-603103
[2]Senior Professor, Dean for school of electrical science, Hindustan University, Chennai, INDIA

Abstract: *Context-based adaptive binary arithmetic coding (CABAC) is needed in the present days for high speed H.264/AVC decoder. The high speed is achieved by decoding one symbol per clock cycle using parallelism and pipelining techniques. In this paper we present an innovative hardware implementation of the renormalization which is a part of CABAC binary arithmetic decoder. The renormalization of range and value is specified as a sequential loop process that shifts only one bit per cycle until the range and value are renormalized. To speed up this process, a special hardware technique is used. The hardware will take one clock cycle to shift n bit data. The proposed hardware is coded using HDL language and synthesized using Xilinx CAD tool.*

Keywords: CABAC, renormalization, H.264, AVC, MPEG2 etc

I. INTRODUCTION

For multimedia coding applications, ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group (MPEG) jointly developed the latest video standard H.264/AVC (ITU-T Recommendation H.264:2003). Compared with existing video coding standards this provides more than twice the compression ratio while maintaining video coding quality. The higher throughput is due to the adoption of many new techniques, such as multiple reference frames, weighted prediction, deblocking filtering and context-based adaptive entropy coding. There are two approaches available for context- based adaptive entropy coding namely context-based adaptive variable length coding (CAVLC) and context-based adaptive binary arithmetic coding (CABAC). The CABAC coding achieves better compression efficiency better than CAVLC, but it brings higher computation complexity during decoding.

The compression efficiency is up to 50% over a wide range of bit rates and video resolutions compared to previous standards (e.g. MPEG2 or H.263). The downside is that the decoder complexity also increased; it is about four times higher [2]. Using a DSP processor to decode a single bin, it takes 30 to 40 cycles. In order to improve the video decoding, the throughput of a video coder using CABAC reaches almost 150 Mbin/s, which makes it difficult to implement in a programmable processor. Therefore, an efficient hardware decoder [3] is important for low-power and real-time H.264 codec applications. The decoding process of CABAC is bit-serial and has strong data dependency because the next bin process is depended on the previous bit decoding result. This data dependency makes the designer to exploit parallelism during decoding is difficult. The context models [5] of the current syntax element (SE) are closely related to the results of its neighboring macro blocks (MBs) or blocks, which leads to frequent memory access. The researchers are addressing these issues for exploring the parallelism and optimize memory access.

Cite this article as: Karthikeyan C, Dr. Rangachar. "Optimization of the critical loop in Renormalization CABAC decoder." *International Conference on Information Engineering, Management and Security (2015)*: 199-203. Print.

Figure 1 shows H.264/AVC's basic coding structure for encoding one macro block, a sub block of a frame of the video stream. The decoder is used inside the encoder to obtain best perceptual quality at the decoder side. To reduce block artifacts an adaptive deblocking filter is used in the motion compensation loop. This combined with multiple reference frames and sub-pixel inter and intra mode motion compensation gives very strong compression results.

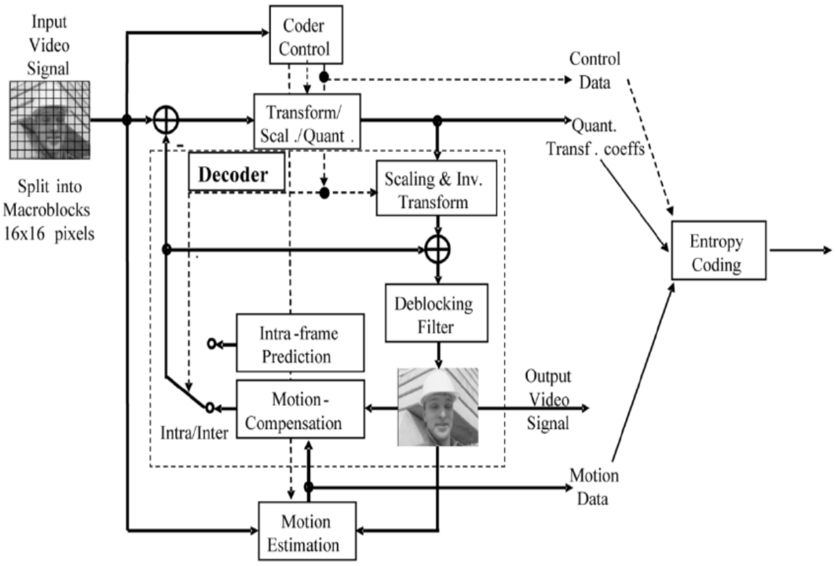


Figure. 1. H.264/AVC macro block encoder with functional blocks and data flows.

The decoder is a central part of the encoder. In section II, we introduce the primary steps of CABAC encoding and decoding process. In Section III, we describe the basic scheme of our CABAC decoder architecture. We present an overview of the framework of our renormalization hardware architecture. In this section IV, we focus on the simulation and synthesize of the proposed architecture. In Section V, we summarize the conclusions and future work.

II. CABAC ENCODER AND DECODER

In this section the basic principles of CABAC encoding and decoding process are discussed. The CABAC encoding and decoding process consists of three elementary steps.

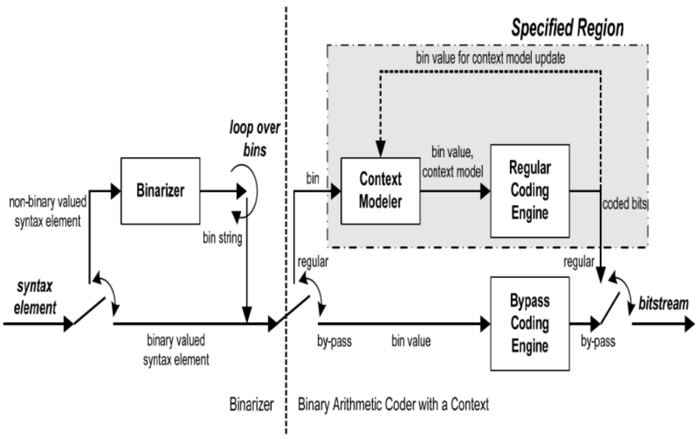


Figure 2 CABAC Encoder diagram

Figure 2 shows the encoding procedure of CABAC [9]. In the first step a given binary valued syntax element is uniquely mapped to a binary sequence, called bin string by the binarizer unit. When the input itself is in binary format this initial step is bypassed. For each element of the bin string or for each binary valued syntax element, one or two subsequent steps may follow depending on the coding mode. In the regular coding mode, prior to the actual arithmetic coding process the given binary decision which, in the sequel, referred to as a bin, enters the context modeling stage, where a probability model is selected such that the corresponding choice may depend on previously encoded syntax elements or bins. After the assignment of a context model the bin value along with its associated model is passed to the regular coding engine, where the final stage of arithmetic encoding together with a subsequent model updating

Cite this article as: Karthikeyan C, Dr. Rangachar. "Optimization of the critical loop in Renormalization CABAC decoder." *International Conference on Information Engineering, Management and Security (2015)*: 199-203. Print.

takes place. Bypass coding mode is chosen for selected bins in order to allow a speedup of the whole encoding process by means of simplified coding engine without the usage of an explicitly assigned model.
The CABAC encoder consists of three elementary steps: binarization, context modeling and binary arithmetic coding [4]. These incoming data are the coefficients from the transformations in Figure 1 together with some context information. In the second step a fitting probability model, based on the context, is selected for each binary symbol. This model drives the arithmetic coder (step three) by providing an estimate of the probability density function (PDF) of the symbol that will be encoded. The better this estimate, the better the compression. CABAC uses in total 399 models to model the PDFs of each syntax element such as macro block type, motion vector data, texture data, etc. The models are kept 'up to date' during encoding through the use of an *adaptive* coder [6] which estimates the PDF based on previously coded syntax elements.
There are three major data dependencies are extracted as follows: Renormalization is dependent on range update.

- Probability transition is dependent on bin decision
- Context switching is dependent on decoded bin

These three data dependency relations lead to three recursive computation loops, which can hardly be sped up by pipelining [7],[10], and thus largely limit the system performance. The following table I illustrates the frequency and the necessary operation to the internal variables. If the decoded symbol is the least probable symbol (LPS), it takes more cycles to evaluate the next coding range and coding offset required for the next symbol decoding. The coding range should always be modified and the offset should also be decremented. To find the shift amount n, we also need to count the leading zeros of the codeword. On the contrary, the consequent operations are much simpler when the decoded symbol is the most probable symbol
Table I Update variable after one symbol decoding

Sl.No	Case	MPS decoding	LPS decoding	Inference
1	Frequency	Frequent	None	No renormalization
	Range	R_{MPS}	-	
	offset	No change	-	
2	Frequency	Rare	Always	Renormalization
	Shift amount	1	Arbitrary	
	Coding range	R_{MPS}<<1	R_{LPS} << n	
	Coding offset	Offset<<1	(Offset-R_{MPS}) << n	

The following is the renormalization process in the arithmetic decoding engine

1. It accepts bit inputs from slice data and the variables codIRange and codIOffset.
2. After the renormalization process it outputs the updated variables codIRange and codIOffset.
3. A flowchart of the renormalization is shown in Figure 4. The current value of codIRange is first compared to 0x0100:
   - If codIRange is larger than or equal to 0x0100, no renormalization is needed and the RenormD process is finished.
   - Otherwise (codIRange is less than 0x0100), the renormalization loop is entered. Within this loop, the value of codIRange is doubled, i.e., left-shifted by 1 and a single bit is shifted into codIOffset by using read_bits(1).

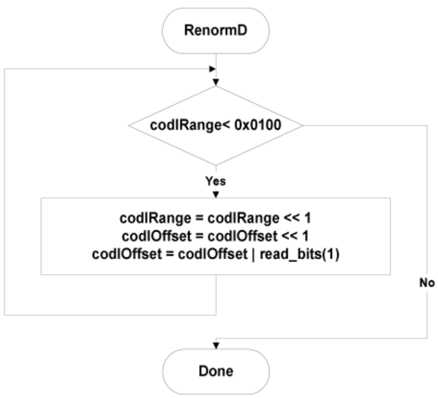


Figure 4 Flowchart of renormalization

Cite this article as: Karthikeyan C, Dr. Rangachar. "Optimization of the critical loop in Renormalization CABAC decoder." *International Conference on Information Engineering, Management and Security (2015)*: 199-203. Print.

III HARDWARE IMPLEMENTATION OF RENORMALIZATION

Re-normalization engine based on a head-one detector The last step of the decode decision engine flow is renormalization. To keep the precision of the whole decoding process, the refined codIOffset and codIRange have to be renormalized to ensure that the codIRange is not less than 256. For example, if the refined codIRange is 9'b000001010, the codIRange should be shifted five bits while the codIOffset reads five bits from the bit stream during the renormalization process. Based on the principle of renormalization, we find that if we locate the first appearing '1' inside the codIRange, we can successfully decide the number of bits of the codIRange to shift and of the codIOffset to read. Moreover, the renormalization process is part of the critical timing path in CABAC hardware decoder implementation.

To improve clock frequency, this path must be kept as short as possible. Thus, a parallel 'head-one detector' re-normalization architecture is proposed in the figure 5. Nine bits of the codIRange are split into three parts (3-bit vector), each of which determines whether there is a '1' among three input bits.

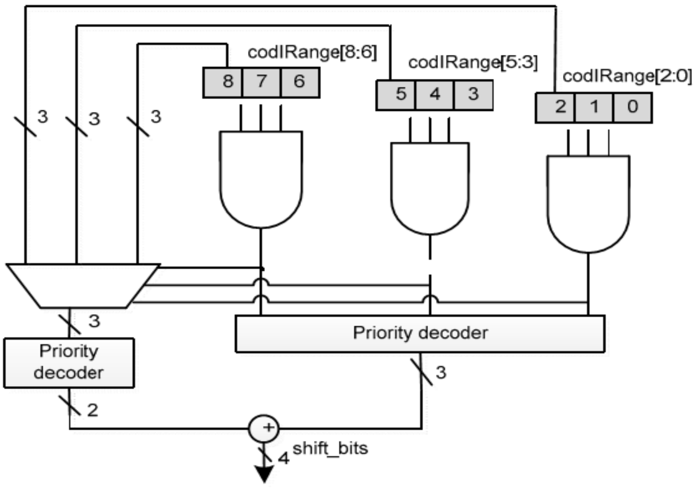


Figure. 5 Re-normalization engine based on a head-one detector

IV RESULT AND DISCUSSION

The proposed architecture is coded using HDL language. We have used structural level implementation and the simulation result of renormalization of given data is shown in the figure 6

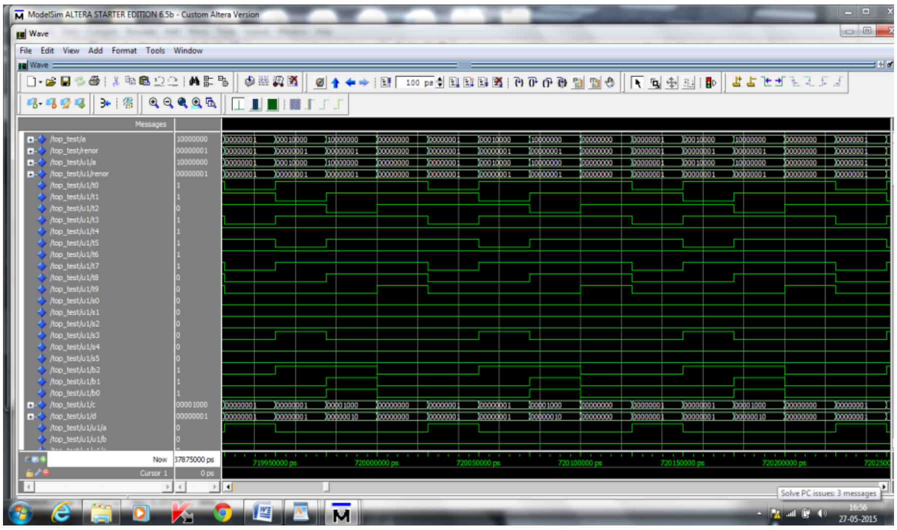


Figure 6 simulation result of renormalization process

The above code is further synthesized using Xilinx EDA tools. The device used for synthesize is vertex 4 200k FPGA. The RTL diagram is shown in the figure 7. The device utilization summery is shown in the table II.

Table II Device utilization summery

Cite this article as: Karthikeyan C, Dr. Rangachar. "Optimization of the critical loop in Renormalization CABAC decoder." *International Conference on Information Engineering, Management and Security (2015)*: 199-203. Print.

Sl.No	Description	Utilized	Available	% of utilization
1	Slices	15	89088	0%
2	4 input LUTs	26	178176	0%
3	Bounded IOBs	16	960	1%
4	Maximum combinational path delay 8.48 ns			

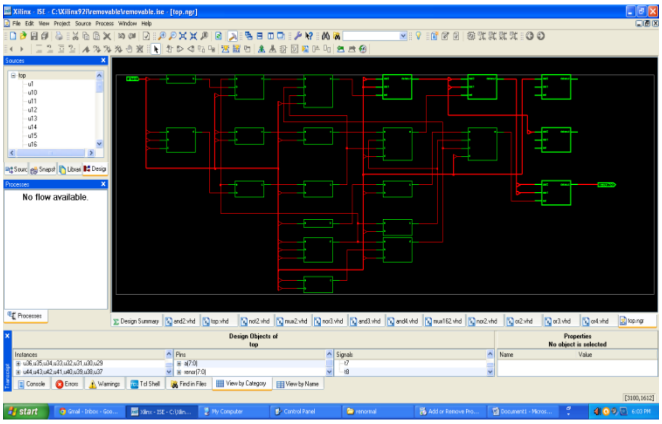


Figure 7 RTL View of renormalization

CONCLUSION

In this work we have presented a novel FPGA-design for renormalization engine which is present in CABAC decoder. CABAC decoder uses leading one detector for the renormalization. We have proposed a hardware which will have one clock cycle to find the leading one in the given bit stream. The proposed hardware is simulated and synthesized using CAD tools. The maximum frequency of operation is 117 MHz.

REFERENCES

[1] T. Wiegand, G. J. Sullivan, and G. B. A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560–576, July 2003.

[2] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: tools, performance, and complexity," IEEE Circuits and Systems Magazine, vol. 4, no. 1, pp. 7–28, 2004.

[3] Wei Yu, Yun He, "A high performance cabac decoding architecture", IEEE Trans. Consum. Electron., vol. 51, no. 4, pp. 1352-1359, Nov. 2005.

[4] D. Marpe, H. Schwarz, G. Bl¨attermann, G. Heising, and T. Wiegand, "Context-based adaptive binary arithmetic coding in JVT/H.26L," Proc. IEEE International Conference on Image Processing (ICIP'02), vol. 2, pp. 513–516, September 2002.

[5] D. Marpe, H. Schwarz, and T. Wiegand, "Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 620–636, July 2003.

[6] R. R. Osorio, J. D. Bruguera, "A New Architecture for fast Arithmetic Coding in H.264 Advanced Video Coder", IEEE Proc. Euromicro Conference on Digital System Design, pp. 298-305, Sept. 2005.

[7] Yuan, T.C., "A Novel Pipeline Architecture for H.264/AVC CABAC Decoder", IEEE Asia Pacific Conf. on Circuit and Systems, p.208-311, 2008

[8] Kuo, M.Y., Li, Y., Lee, C.Y., "An Area-Efficient High-Accuracy Prediction-Based CABAC Decoder Architecture for H.264/AVC", IEEE Int. Symp. on Circuit and Systems, p.160-163, 2011.

[9] Liao, Y.H., Li, G.L., Chang, T.S., "A highly efficient VLSI architecture for H.264/AVC level 5.1 CABAC decoder", IEEE Trans. Circ. Syst. Video Technol., 22(2): 272-281, 2012.

[10] Shi, B., Zheng, W., Lee, H.S., Li, D.X., Zhang, M., Pipelined Architecture Design of H.264/AVC CABAC Real-Time Decoding. 4th IEEE Int. Conf. on Circuits and Systems for Communications, p.492-496, 2008.

Cite this article as: Karthikeyan C, Dr. Rangachar. "Optimization of the critical loop in Renormalization CABAC decoder." *International Conference on Information Engineering, Management and Security (2015)*: 199-203. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS034		**eAID**	ICIEMS.2015.034

Polarization Modulation for Communication

Mr. Jameer Manur[1], Ms. Mohini Nagardeolekar[2], Prof. Joydeep Bagchi[2], Prof. Milind Patil[1]
[1]Deptment. of Electronics and Telecommunication,
Vishwakarma Institute of Information Technology, Pune, India.
[2]Radio Physics Lab, Inter-University Centre for Astronomy and Astrophysics, Pune, India.

Abstract- *Polarization is one of the most fundamental property of Electromagnetic (EM) wave, which describes it completely. The study of polarization of light through Faraday Rotation Effect, rotation of plane of polarized wave when traveling through Terbium Gallium Garnet (TGG) crystals placed inside solenoid, subjected to a strong axial magnetic field can be a novel approach in communication. Experiment shows conversion of polarization modulated light into intensity modulated light, and phase shifted demodulated waveform w. r. t. input modulating signal. Insertion of properly matched and tuned circuit before coil and amplifier after demodulation leads to better reception of signal.*

Keywords- Faraday rotation, Polarization Modulation, Analog communication.

I. INTRODUCTION

In radio communication, modulating signal can be modified based on the properties of wave, such as, amplitude (AM), frequency (FM) and phase (PM). Similarly, modulation of light can be done by varying its properties; like intensity (ASK), frequency (FSK), phase (PSK), wavelength or polarization. All these techniques use State of Polarization (SOP) of fully polarized light-wave as an information carrying parameter. The greatest difference between radio-wave AM techniques and optical modulation techniques is that the output of radio receiver is usually proportional to amplitude of incoming signal, while output of optical receiver is proportional to intensity of incoming signal, because of the use of square-law photo detectors. Amplitude of a signal contains both magnitude and phase information, while intensity of a signal is the square of its magnitude. Amplitude modulation of light can be done by internal or external way. Internal modulation can be achieved by varying supply source (voltage or current) of artificial source of radiation, for eg., laser, while external modulation can be through controlling the light absorption in semiconductors by varying voltage or current [13]. The different effects used to modulate the light are: Kerr and Pockels Effect (Electro-optic), Faraday Effect (Magneto-optic) and Acousto-optic Effect [1] [13].

In Faraday Effect, the phase modulation of two mutually perpendicular components of linearly polarized light results into polarization modulation, which is then transformed into Amplitude Modulation by analyzer. The Faraday Effect originates from the effect of static magnetic field on the motion of electrons in the presence of light and Lorentz force [1]. The magneto-optic modulators are based on the rotation of optical polarization as light propagates along the magnetic field in a material, by the Faraday

Cite this article as: Jameer Manur, Mohini Nagardeolekar, Joydeep Bagchi, Milind Patil. "Polarization Modulation for Communication." *International Conference on Information Engineering, Management and Security (2015)*: 204-207. Print.

Effect. These modulators can be made of Faraday material (for eg., Terbium Gallium Garnet-TGG crystals)placed inside an electric coil (solenoid) with modulating electric current. Electro-optic modulators can be of transverse and longitudinal type, which require several thousands of drive voltage to achieve a desired phase shift. In Acousto-optic modulator, the frequency response is limited by the acoustic transit time across the finite width of light beam, acoustic propagation loss in that medium and frequency response of acoustic transducer. There is a need of sharply focused laser beam to reduce the acoustic transit time. Also, Acousto-optic effect has limitation of width of frequency band < $(1-2)*10^6$ [13]. To overcome these limitations we tried to modulate the light beam by varying its SOP using Faraday Effect and hence using Magneto-optic modulators. Polarization of a laser beam is modulated or switched between two orthogonal states. Like intensity modulation, polarization modulation does not require sophisticated stabilized laser. Since, the optical sensitivity of a polarization modulated light is about twice as high as intensity modulated signal, it provides better receiver Signal to noise ratio (SNR) [1]. Polarization Shift-keying (POLSK) modulation is an abstract scheme of modulation, due its applicability to fiber communication; microwave communication and light wave inter satellite communication. POLSK modulation needs 3dB less power than ASK and FSK in absence of laser phase noise [2].

II. THEORY

Christian Huygens was the first to suggest that light was not a scalar quantity, based on his work on the propagation of light through crystals. (This vectorial nature of light is called as Polarization.) If we follow mechanics and equate an optical medium to an isotropic elastic medium, it should be capable of supporting three independent oscillations (optical disturbances): $u_x(r, t)$, $u_y(r, t)$ and $u_z(r, t)$.

In a Cartesian system $u_x(r, t)$ and $u_y(r, t)$ are the transverse components, while $u_z(r, t)$ is longitudinal component of light wave.

$$u_x(r, t) = u_{0x} \cos(\omega t - kr + \delta_x) \quad (1a)$$

$$u_y(r, t) = u_{0y} \cos(\omega t - kr + \delta_y) \quad (1b)$$

$$u_z(r, t) = u_{0z} \cos(\omega t - kr + \delta_z) \quad (1c)$$

From Fresnel and Argo's investigation (1818) on Young's interference experiment using polarized light, it is concluded that Eq.(1c),does not exist, i.e., light consists of transverse components only [4] [15]. Plane wave propagation in the z-direction, the components of optical field in x-y plane are represented as:

$$E_x(z, t) = E_{0x} \cos(\omega t - kz + \delta_x) \quad (2a)$$

$$E_y(z, t) = E_{0y} \cos(\omega t - kz + \delta_y) \quad (2b)$$

Where, $\tau = \omega t - kz$ is a propagator, E_{0x} and E_{0y} are maximum amplitudes of x and y components. δ_x and δ_y are phases of x and y components. In the x y-plane, we can construct a vector as:

$$E(z, t) = E_x(z, t)\cdot i + E_y(z, t) \cdot j \quad (3)$$

Where, i and j are unit vectors in x and y directions respectively.

III. EXPERIMENTAL SETUP

The system shown in Fig.1 includes the Laser (Red, 632.8nm), as a light source. The light is then passed through a Polarizer, which then is rotated due to magnetic field induced in the magnetic coil or Solenoid (1200 turns) with the Faraday material TGG crystals are placed inside. The rotated light is then analyzed through another polarizer, called as 'Analyzer'. The exact detection had been done at Detector. In the whole system, the laser light acts as a carrier, whereas the input signal, applied to solenoid acts as the modulating signal. The modulating signal produces the magnetic field in the solenoid in the direction of propagation of light beam. This magnetic field rotates the plane of polarization of the incoming light, due to Faraday Effect, i.e. the original state of polarization (SOP) of light have been changed, which ultimately results into the polarization modulation. The rotation of the plane of polarization through an angle θ can be calculated by:

$$\theta = V \cdot B \cdot L$$

where, 'V' is the property of material known as Verdet constant, 'B' is the induced magnetic field in the solenoid (in gauss) and 'L' is the length of the optical beam through a material, placed inside the solenoid.

Helmholtz coil is special coil to generate a uniform magnetic field at the center of the coil. It consist of pairs of circular coil, when charges are in motion the magnetic field is generated around it. Both circular coils spaced on radius apart and wound so that the current flows through both coils in the same direction. The Helmholtz coil is designed in the lab as shown. Each coil has 500 turns witch generate around 120 Gauss of magnetic field at 1A of current. The crystal is placed at the center of the coil with help of spider arrangement, the laser is pass through the crystal and signal get modulate in the presence of magnetic field. Helmholtz coil used in many applications from canceling the Earth's magnetic fields to generate uniform magnetic field for many experiments, it can generates static, time varying DC or AC. It also used for measurement of a permanent magnetic moment, removing or ruling background magnetic field, it also used in TV production for picture alignment procedures.

Cite this article as: Jameer Manur, Mohini Nagardeolekar, Joydeep Bagchi, Milind Patil. "Polarization Modulation for Communication." *International Conference on Information Engineering, Management and Security (2015)*: 204-207. Print.

Figure 1. Block diagram of a system

IV. OBSERVATIONS

The sinusoidal signal of 1 kHz, 25Vp−p was applied as an input to the solenoid; the output is sinusoidal signal of 1 kHz, 100mVp−p, which is then amplified. Fig. 2 shows the phase difference between input and output signal. Here, the polarized laser light (Red) carries sinusoidal signal combines with sinusoidal (modulating) signal at solenoid al, which then gets rotated due to presence of TGG crystals, i.e. SOP of original light has got changed, i.e. 'Polarization Modulation'. This modulated light when passes through analyzer, the polarization modulated light gets converted into intensity modulated light. This light is then demodulated using the photo-detector. The demodulated signal shows the phase difference w. r. t. the input signal. Except the loss in energy of signal, the nature of the original signal has been successfully retrieved at the detector side. The solenoid gives a stable response for the frequency range 10Hz-100 kHz. Insertion of impedance matching and tuning circuit before solenoid leads to a better reception of signal. To test it further, we used audio signal as a modulating wave. An audio signal in form of '.wav' file format was sent through a solenoid for around 30sec. The modulated audio is received at photo detector. Demodulated signal is amplified through an audio amplifier. The Fig.3 shows an audio input applied, with its time sequence (above) and spectrogram (below), while Fig.4 is the detected output.

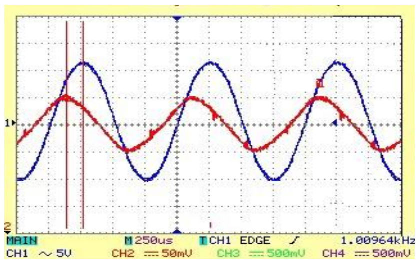


Figure 2. Graph showing input signal (blue) and phase shifted output signal (red), $\Delta t = 125\mu$sec

From the time sequence plots, the variation of magnitude at a particular time instant is nearly same for both the plots. From spectrogram of input audio, the highest frequency components is ~3kHz, and lowest of ~1kHz. The same levels have been observed at output too. Along with this, sudden drop in input (at t ~20sec) has been perfectly reflected in the demodulated signal. The level of magnitude in spectrogram, for input is 72.78 dB and for output it is 70.76 dB. The grayish part at the background of spectrogram shows the energy at particular time instant, darker the region, more is the energy. The output spectrogram shows small amount of loss in energy.

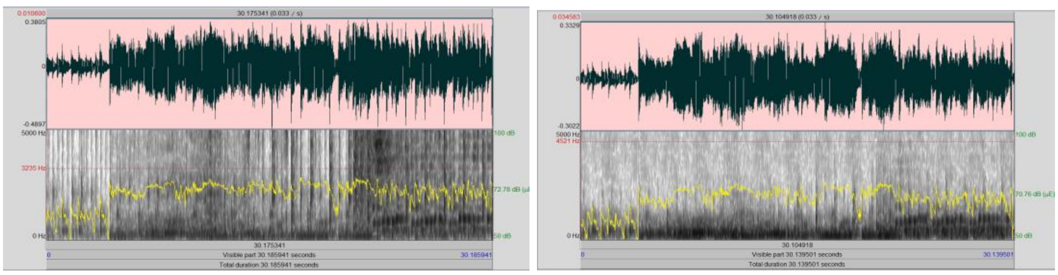


Figure 3. The time sequence (above) and spectrogram (below) of the transmitted (Left) and received (right) audio signal

V. ACKNOWLEDGEMENT

This work has been supported by Inter-University Centre for Astronomy and Astrophysics (IUCAA), Pune, India. Dr. Joydeep Bagchi is a Professor in IUCAA, Pune. All authors would like to acknowledge IUCAA for availing the Radio Physics Lab for the experimentation.

REFERENCES

[1] Patrick LiKamWa, Shi-Kay Yao, Eddie H. Young, "Modulators and Demodulators, Optics," *OSA, Applied Optics*, Vol. 24 No. 12,15 June 1985

[2] "Modulation of Light," Encyclopedia.

Cite this article as: Jameer Manur, Mohini Nagardeolekar, Joydeep Bagchi, Milind Patil. "Polarization Modulation for Communication." *International Conference on Information Engineering, Management and Security (2015)*: 204-207. Print.

[3] Serigo Benedetto, Pierluigi Poggolini, "Theory of Polarization Shift Keying Modulation," *IEEE transactions on communication*, Vol. 40 No. 4, Apr 1992

[4] Edward Collette, Beth Schaefer, "Visualization and Calculation of Polarized Light," *OSA, Applied Optics*, Vol. 47 No. 22/ Aug, 2008.

[5] A book by Edward Collet on "Polarized Light: Fundamentals and Application"

[6] Shigeru Kuwano, Katsutoshi Tsukamoto, Norohiko Morinaga, "A Polarization Modulation (PLM), Optical coherent Detection System and its Phase Noise Canceling Effect," *Electronics and Communication in Japan*/part-1,Vol. 75 No. 11,2 Feb, 1992.

[7] Zain ul Abidin , Pei Xiao, Muhammad Amin, Vincent Fusco, "Circular Polarization Modulation for Digital Communication stems".

[8] Andreas Edlioth, John Thiclecke , "A mathematical model for a polarization based orientation measurement principle in time of arrival radio localization system," *IEEE, International Conference on Indoor Position and Indoor Navigation*,Sep 2010.

[9] Jeffrey A. Davis, M.Azad Islam, Roger A.Lilly, "Time-dependent effects in Faraday rotation in pulsed magnetic fields: an explanation," *OSA, Applied Optics*, Vol. 24 No. 12, 15 Jun, 1985.

[10] Aloke Jain, Jayant Kumar, Fumin Zhou, Lian Li, "A simple experiment for determining Verdet constants using alternating current magnetic fields," *American Association of Physics Teachers, Am. J. Phys.* 67(8), Aug 1999.

[11] Bin Cao, Qin-Yu Zhang and Lin Jin, "Polarization division multiple access with polarization modulation for LOS wireless communications," *EURASIP Journal on Wireless Communications and Networking*, 2011:77.

[12] Z. Kavehvash, F. Massoumian, "Coding and decoding in point-to-point communication using the polarization of light beam," *OSA, Applied Optics*, Vol 47 No. 14, 10 May 2008.

[13] Sridhara K., "Free space optical communication," International *Journal of Latest Research in Science and Technology*, Vol. 1, Issue3, Pg. No. 202-205, Sep-Oct 2012.

Cite this article as: Jameer Manur, Mohini Nagardeolekar, Joydeep Bagchi, Milind Patil. "Polarization Modulation for Communication." *International Conference on Information Engineering, Management and Security (2015)*: 204-207. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	VOL	01
Website	www.iciems.in	eMail	iciems@asdf.res.in
Received	10 - July - 2015	Accepted	31- July - 2015
Article ID	ICIEMS035	eAID	ICIEMS.2015.035

Report on Design of distributed Energy Efficient and reliable Routing Protocol for Wireless Sensor Networks

Jayashri Gaurkar[1], Kanchan Dhote[2]
[1,2] P.G Student , Dept.of Electronics Communication Engg.,
Head of Electronics and Communication Engg.,TGPCET, NAGPUR, INDIA

Abstract: *A wireless sensor network (WSN) is a collection of nodes organized into a cooperative network, which are small energy constrained devices The efficient use of energy source in a sensor node is most desirable criteria for prolong the life time of wireless sensor network. So designing efficient routing for reducing energy consumption is the important factor. The energy consumed in cluster head (CH) selection phase of a random cluster based wireless sensor network (WSN) has been assumed as an insignificant factor in the previous research works. In this paper, the consumption of energy during the transmission of data from sensor nodes to the sink has been calculated. Routing protocols in WSNs along with the most energy efficient protocol named LEACH (low energy adaptive clustering hierarchy) and STR protocol (shortest tree routing protocol) along with its advantages and disadvantages are discussed here. In this paper we improved the energy consumption of the node to get the parameters result such as energy, delay, throughput, jitter and pdr.*

Keywords: Wireless sensor network, hierarchical routing, cluster based routing protocol, LEACH, Shortest path routing

I. INTRODUCTION

A wireless sensor network is a collection of nodes organized into a cooperative network. Each node consists of processing capability (one or more microcontrollers, CPUs or DSP chips), may contain multiple types of memory (program, data and flash memories), have a RF transceiver (usually with a single omnidirectional antenna), have a power source (e.g., batteries and solar cells), and accommodate various sensors and actuators. The nodes communicate wirelessly and often self-organize after being deployed in an ad hoc fashion. Systems of 1000s or even 10,000 nodes are anticipated. Such systems can revolutionize the way we live and work. Wireless sensor networks are comprised of large numbers of low-cost, low-power and multifunctional sensor nodes. Thus, it is predicted that wireless sensor networks will become conventional in our daily life. A wireless sensor network (WSN) is typically composed of a large number of low-cost sensor nodes, which work collectively to carry out some real-time sensing and monitoring tasks within a specific area. The main constituent of a WSN are multiple number of sensor nodes and at least one sink node. The number of sensor nodes depends upon the application's requirement. Energy efficiency is one of the most important factors in WSNs. Hierarchical (clustering) techniques can aid in reducing useful energy consumption. LEACH is a classical clustering hierarchical protocol, which incorporates randomized rotation of the high-energy cluster head position among the sensors to avoid draining the energy of any one sensor in the network. In this way, the energy load of being a cluster head is evenly distributed among the nodes.

Cite this article as: Jayashri Gaurkar, Kanchan Dhote. "Report on Design of distributed Energy Efficient and reliable Routing Protocol for Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 208-214. Print.

The subsequent LEACH-optimized protocols in literatures mostly improve the capability of cluster head distribution, so that the energy consumption of the whole network is reduced and the system lifetime is prolonged. However, all these hierarchical protocols only consider the algorithm of cluster formation and give little consideration on aggregated data transmission. Since the aggregated data is important, the reliability of transmission should be guaranteed. The proposed model is on LEACH and STR protocol and also compared with only STR protocol and no protocol to get the parameters result such as energy, delay, throughput, jitter and pdr. LEACH is used for the cluster head formation and STR protocol is used for finding the shortest tree routing nodes with minimum score of the node and transmits the data by dividing into two way and then data aggregation is done in the next cluster head near to the neighbour to destination node. In this way we can minimise the energy consumption of the node.

The remainder of this paper is organized as follows: Section 2 presents the related works .Section 3 describes the proposed protocol in details. In section 4 we make analysis and simulation, comparing with the previous multipath routing protocols. Finally, section 5 draws conclusions and shows the future works.

II. RELATED WORK

Routing Protocol for Wireless Sensor Network

Recent advances in wireless sensor networks have lead to many new protocols specifically designed for sensor networks where energy awareness is an essential consideration. But approaches like Direct Communication and Minimum Transmission Energy do not guarantee balanced energy distribution among the sensor nodes. In Direct Communication Protocol each sensor node transmits information directly to the base station, regardless of distance. On the other hand, in case of Minimum Transmission Energy routing protocol data is transmitted through intermediate nodes. we classified most important energy efficient routing techniques based on various clustering attributes like cluster formation and data gathering process.

A.LEACH

W. R. Heinzelman, A. P. Chandrakasan and H. Balakrishnan proposed Low Energy Adaptive Clustering Hierarchy (LEACH) protocol in 2000. It is one of the most popular hierarchical routing algorithms for sensor networks. This protocol incorporates the formation of clusters and cluster heads (CHs) for the respective clusters in which all the other sensor nodes send the data to the cluster head (CH). The received data is then aggregated and is sent to the base-station (BS) periodically by the cluster head which reduces the amount of data that is to be transmitted to the basestation. The role of the cluster head (CH) is rotated among the other sensor nodes in the cluster so as to evenly distribute the power load between the sensor nodes in a particular cluster. A TDMA/CDMA MAC is used for avoiding the collisions among the clusters and within the clusters.

Working Principle: The LEACH protocol functions in two different phases. The setup phase and the steady state phase . The formation of clusters and selection of the cluster heads is done during the setup phase and the aggregated data is transmitted to the base-station during the steady state phase which is of greater duration than the setup phase. During the setup phase, a random number r, between 0 and 1, is selected by the sensor nodes. If this random number is less than a threshold value T(n), that sensor node is selected as the cluster head. The threshold value T(n) is calculated as follows :

$$T(n) = p / [1-p(r \bmod(1/p))] \text{ if } n \in G$$

Where, p is the predetermined number of sensor nodes, r is the random number and G is the set of nodes that are involved in the CH selection that have not been selected as cluster heads in the last (1/p) round. After the selection, the cluster heads sends an advertisement to all the other sensor nodes in the network. The formation of clusters is based upon the signal strength of this advertisement. After the cluster formation, a TDMA schedule is created assigning time slots to the sensor nodes for data transmission. After the cluster formation and the selection of the cluster heads, the network goes into steady state phase where the aggregated data from the sensor nodes is sent to the base-station by the cluster heads. The network again goes back into the setup phase after a predetermined time period to select a new set of cluster heads as to rotate the role of the cluster heads among the nodes of a cluster.

The network lifetime is increased as the load of power dissipation is evenly distributed among the nodes in the cluster. Also the amount of data to be transmitted is less which in turn reduces the latency of the network. The LEACH protocol is not suitable for networks deployed in large areas. Also the predetermined cluster heads may not be uniformly distributed. The path taken by the aggregated data to reach the base station is not optimal.

B.PEGASIS

The Power-Efficient Gathering in Sensor Information Systems (PEGASIS) proposed in is an improvement over the LEACH protocol. It is a near optimal chain-based protocol. The idea of cluster formation and cluster head is discarded in PEGASIS. Instead of multiple nodes, a single node in the chain communicates with the base-station. The sensor nodes in this protocol only communicate with a single node closest to them and communication with the base-station is done in rounds so that the power dissipation in communicating with the base-station is distributed evenly among all the nodes.

Working Principle: PEGASIS assumes that all the sensor nodes maintain a database of the location of all the other nodes in the network. Each node determines the distance of its neighboring nodes using the signal strength and adjusts the signal strength only to communicate with the closest node. In PEGASIS, the sensor nodes closest to each other are in the chain and they form a path to transmit the aggregated data to the base-station. The chain is constructed using Greedy algorithms. Each sensor node sends the sensed

Cite this article as: Jayashri Gaurkar, Kanchan Dhote. "Report on Design of distributed Energy Efficient and reliable Routing Protocol for Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 208-214. Print.

data to its closest node in the chain. The data is aggregated at each node in the chain and finally only the aggregated data is sent to the base-station. The lifetime of each node is increased as they only have to communicate with their closest node which, as a result increases the network lifetime. Delay is caused in data transmission from the distant node in the chain. There is significant overhead as the nodes need the know-how about the other node location and the path for transmitting data.To overcome the problem of delay occurrence in transmitting the aggregated data to the base-station (BS) an extension to PEGASIS, called Hierarchical-
PEGASIS was introduced in which the transmission of the data was allowed only by the spatially separated sensor nodes. This ensured parallel data transmission and reduced the delay.

C. TEEN

Threshold-sensitive Energy Efficient sensor Network (TEEN) is a hierarchy based routing protocols proposed in, for time-critical applications. The region is sensed continuously by the sensor nodes but the sensed data is transmitted less frequently. The cluster heads (CHs) broadcasts a hard threshold, which is the threshold value of the sensed data and a soft threshold, which is a small change in the hard threshold value of the sensed data to all the other sensor nodes in a cluster. The soft threshold instigates the sensor node to switch on its transmitter and transmit the data.
Working Principle: In TEEN, a hard threshold value and a soft threshold value is sent to all the other sensor nodes by their respective cluster heads (CHs). The sensor nodes begin to transmit data by switching on their transmitters when they sense a change in the soft threshold value. Transmission of data occurs only when the sensed data is in the range of interest of the user.

Adv/Disadv: TEEN protocol reduces the number of transmissions by only transmitting the data only when the sensed data is of interest to the user. The major disadvantage of the TEEN protocol is that, if the threshold values are not received, the sensor nodes will not communicate and the user will not receive any data either.

D. APTEEN

Def: Adaptive Periodic Threshold-sensitive Energy Efficient sensor Network (APTEEN), is a hybrid protocol which was proposed in, is also for time critical applications. In APTEEN, according to the user needs and the application type, the threshold values used in TEEN are changed at some specific time intervals.
Working Principle: In APTEEN, few parameters such as the Attributes (A), Hard Threshold (HT), Soft threshold (ST), Schedule and Count Time (CT) are sent to the other sensor nodes in the cluster by the respective cluster heads (CHs). When the sensed data value is greater than the HT, the data is transmitted only when there is a change in that value. Each sensor node in the cluster is given a time slot using a modified TDMA schedule for transmission. When a sensor node does not transmit data for a time period equal to the CT, it is forced to sense again and retransmit the data.
APTEEN is flexible, as the power consumption is controlled by the user by changing the count time (CT) and the threshold values. The implementation of the threshold values and the count time (CT) is complex. Also the overhead increases.

III.Proposed Protocol

A.Introduction

The proposed model is on LEACH and STR protocol and also compared with only STR protocol and no protocol to get the parameters result such as energy, delay, throughput, jitter and pdr. LEACH is used for the cluster head formation and STR protocol is used for finding the shortest tree routing nodes with minimum score of the node and transmits the data by dividing into two way and then data aggregation is done in the next cluster head near to the neighbour to destination node. In this way we can minimise the energy consumption of the node.

B.Cluster Formation

LEACH is one of the first hierarchical routing protocols for WSNs. The idea proposed in LEACH has inspired many other hierarchical routing protocols. Clustering is the method by which sensor nodes in a network organize themselves into hierarchical structures. By doing this, sensor nodes can use the scarce network resources such as radio resource, battery power more efficiently. Within a particular cluster, data aggregation and fusion are performed at cluster- head to reduce the amount of data transmitting to the base station. Cluster formation is usually based on remaining energy of sensor nodes and sensor's proximity to cluster-head. Non cluster-head nodes choose their cluster-head right after deployment and transmit data to the cluster-head. The role of cluster-head is to forward these data and its own data to the base station after performing data aggregation and fusion.

C.LEACH (Low-Energy Adaptive Clustering Hierarchy)

Low Energy Adaptive Clustering Hierarchy (LEACH) proposed by Wendi B. Heinzelman, et al. is the first hierarchical, self-organizing, adaptive cluster-based routing protocol for wireless sensor networks which partitions the nodes into clusters, in each cluster a dedicated node with extra privileges called Cluster Head (CH) is responsible for creating and manipulating a TDMA (Time division multiple access) schedule and sending aggregated data from nodes to the BS where these data is needed using CDMA (Code division multiple access). Remaining nodes are cluster members.In Adaptive clustering, cluster heads change as nodes move in order to keep the network fully connected.

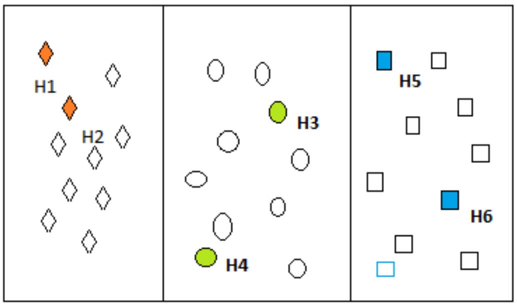


Fig. Multiple Cluster-head in small region.

Key Features of LEACH (Low-Energy Adaptive Clustering Hierarchy)

- Localized coordination and control for cluster set up and operation.
- Local compression to reduce global communication
- Randomized rotation of the cluster heads and the corresponding clusters.
- Random Death of nodes : there is no one section of the environment that is not being "sensed" as nodes die, as occurs in the other protocols.

D.STR protocol (Shortest tree routing protocol)

By formation of cluster head, with highest energy node using LEACH protocol then we uses STR protocol. The STR protocol is used for finding the shortest tree routing nodes with minimum score of the node and transmits the data by dividing into two way and then data aggregation is done in the next cluster head near to the neighbour to destination node. In this way we can minimise the energy consumption of the node.

E. Principle of Proposed Protocol

We proposed the model of LEACH and STR protocol and also compared with only STR protocol and no protocol to get the parameters result such as energy, delay, throughput, jitter and pdr. LEACH is used for the cluster head formation with highest energy node and STR protocol is used for finding the shortest tree routing node with minimum score of the node and transmits the data by dividing into two way and then data aggregation is done in the next cluster head near to the neighbour to destination node. In this way we can minimise the energy consumption of the node.

Consider the above sensor network in which shows the multiple cluster head in small region (fig.). Here we define the no of nodes (nn), which is 30 nodes and cluster head (CH_n) in the network with three group of cluster heads consisting of each cluster heads has 10 sensor nodes including two cluster heads. These cluster heads are formed by LEACH protocol with selecting highest energy of the sensor node into hierarchical structures and lebel it as (CH_n) and (CH_{n+1}) such that $E(CH_n) > E(CH_{n+1})$.This work done by the LEACH protocol.

Then the STR protocol is used after creation of cluster head. We have to find the score of the nodes. This way we will find two shortest tree routing nodes which has minimum score of the nodes and communication will start between the nodes i.e from source to destination. The data will be transmitted by the two shortest path by dividing the data into two ways and then that data aggregates to the next cluster head near to the neighbour of destination node. In this way we can minimise the energy consumption of the node. Also we get result of the parameter energy, delay, throughput, jitter and pdr which shown the next chapter.

Now the number of nodes is 30 in the sensor network by making group of 10 node there is formation of cluster. Each cluster has two cluster heads which has maximum energy node. Then we will find two shortest tree routing nodes which has minimum score of the nodes and communication will start between the nodes i.e from source to destination. The data will be transmitted by the two shortest path by dividing the data into two ways and then that data aggregates to the next cluster head near to the neighbour of destination node. In this way we can minimise the energy consumption of the node. While running this model we can run it in three way as:

1) STR protocol+ Cluster head (LEACH)
2) Only STR protocol
3) Normal network (no protocol used)

In STR protocol+ Cluster head (LEACH), communication is done between source and destination with the help of two cluster heads. And also use STR protocol for shortest path so that data is divided into two ways and then that data aggregates to the next cluster head near to the neighbour of destination node. In this way we can minimise the energy consumption of the node.

In Only STR protocol, communication is done between source and destination without the help of two cluster heads. And then STR

Cite this article as: Jayashri Gaurkar, Kanchan Dhote. "Report on Design of distributed Energy Efficient and reliable Routing Protocol for Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 208-214. Print.

protocol is used for finding shortest tree routing node considering minimum score of the node we will transmit the data from source to destination.

In normal network no protocol is used that means the communication is done directly between source to destination. From the fig.6.1 working of the proposed protocol is as below

- In the cluster1, H1 and H2 are the two cluster heads having E(H1)>E(H2). In the cluster2, H3 and H4 are the two cluster heads having E(H3)>E(H4). In the cluster3, H5 and H6 are the two cluster heads having E(H5)>E(H6).
- If Data transmission done between first two cluster. Cluster head will find shortest path by using STR protocol.Then, data from the source will go to the H1 then, H1 will find such two nodes which having minimum score. And minimum score will find by ratio of minimum distance from H1 and high energy of the node.
- Suppose that node are S1 and S2.Then H1 will send the 50% data to S1 to H3 and other 50% data will send to S2 to H3. In this way data will aggregated to the cluster head H3 and then transmitted to the destination.

IV. Simulation

The performance of LEACH and STR protocol is being evaluated by analysis comparing with only STR protocol and normal network. For analysation we used NS2 and tested our protocols. For performance evaluations following parameters are taken into account: Energy, Network Throughput, Delay, Jitter, Pdr.

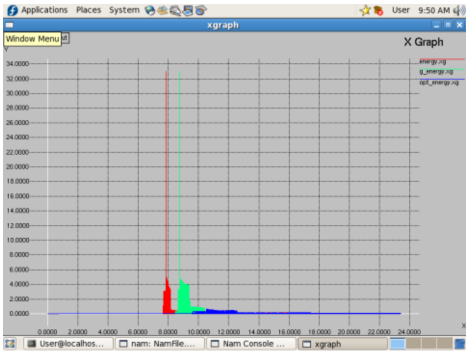


Fig.1 Comparison between energy for Normal protocol, Only STR Protocol and STR Protocol +Cluster Head (LEACH)

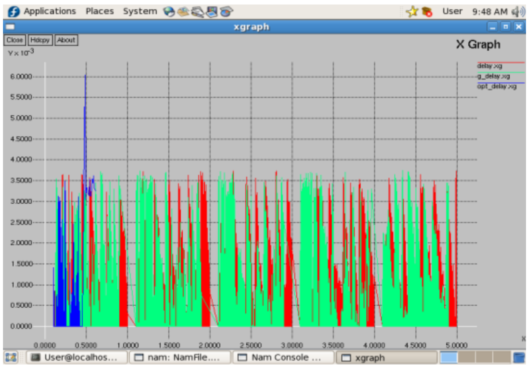
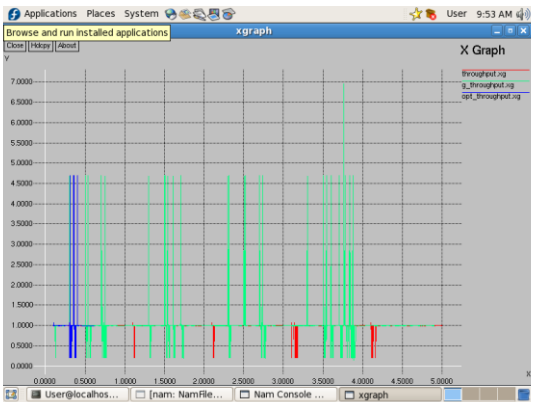


Fig.2 Comparison between dealy for Normal protocol, Only STR Protocol and STR Protocol +Cluster Head (LEACH)



Cite this article as: Jayashri Gaurkar, Kanchan Dhote. "Report on Design of distributed Energy Efficient and reliable Routing Protocol for Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 208-214. Print.

Fig.3 Comparison between throughput for Normal protocol, Only STR Protocol and STR Protocol +Cluster Head (LEACH)

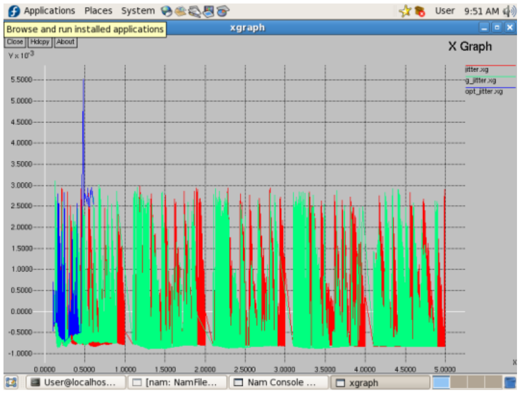


Fig.4 Comparison between jitter for Normal protocol, Only STR Protocol and STR Protocol +Cluster Head (LEACH)

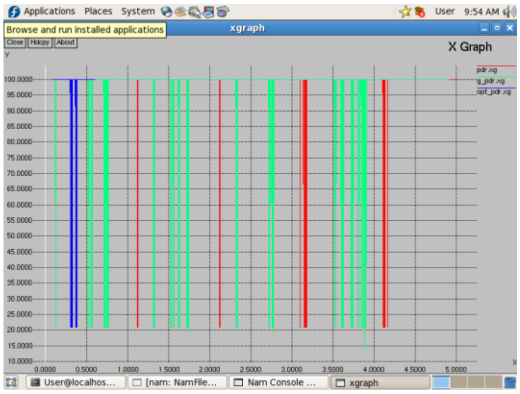


Fig.5 Comparison between pdr for Normal protocol, Only STR Protocol and STR Protocol +Cluster Head (LEACH)

In the given graphs, red color indicates normal network, green color indicates only STR protocol and blue color indicates STR protocol and cluster head. The comparison of parameter are as shown below:

Table .1 Comparison of all parameter for Normal protocol, Only STR Protocol and STR Protocol +Cluster Head (LEACH)

Sr. No.	Parameter	Energy consumption		
		Normal protocol (E_N)	Only STR Protocol (E_{STR})	STR Protocol +Cluster Head (LEACH) (E_{opt})
1	Energy	High	Medium	Low
2	Delay	High	Medium	Low
3	Throughput	Low	Medium	high
4	Jitter	High	Medium	Low
5	Pdr	Low	Medium	high

Cite this article as: Jayashri Gaurkar, Kanchan Dhote. "Report on Design of distributed Energy Efficient and reliable Routing Protocol for Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 208-214. Print.

V. CONCLUSION

The main goal of this thesis was to minimize the energy consumption of the node. The proposed system consists of two main parts: clustering and shortest path routing. In this thesis we have presented a performance analysis of routing protocols for wireless network communication. These protocols have been implemented in NS-2 and are analyzed on the basis of four crucial parameters: energy , delay, throughput, jitter and pdr. After analyzing the graphs , we conclude that STR protocol and cluster head is better in comparison to only STR protocol and normal network. After applying the clustering and shortest path algorithm energy consumption of all parameters are improved.

As a future scope we can say that energy efficiency is one of the major design issues in wireless sensor networks. As most of the energy is consumed in communication than any other task so the need is to develop energy efficient routing protocols. Many of these protocols have been developed in the past few years but still improvements are required in routing protocols in terms of QoS parameters.

REFERENCES

[1] Ning Sun, Young-bok Cho, Sang-ho Lee,"A Distributed Energy Efficient and Reliable Routing Protocol for Wireless Sensor Networks" IEEE International Conference on Computational Science and Engineering IEEE International Conference on Computational Science and Engineering CSE/I-SPAN 2011

[2] Rahul Goyal," A Review On Energy Efficient Clustering Routing Protocol In Wireless Sensor Network" IJRET: International Journal of Research in Engineering and Technology Volume: 03 Issue: 06 | Jun-2014

[3] Mohammad Masdari1 and Maryam Tanabi2, "Multipath Routing protocols in Wireless Sensor Networks: A Survey and Analysis" International Journal of Future Generation Communication and Networking Vol.6, No.6 (2013), pp.181-192

[4] Shivani Duggal, Ms.Preeti Arora, "Design of Shortest Path Algorithm Based on Adjacency Matrix Reduction with LEACH as Benchmark" International Journal of Scientific Research Engineering & Technology (IJSRET)Volume 2 Issue 5 pp 267-273 August 2013

[5] Satvir Singh, Meenaxi," A Survey on Energy Efficient Routing in Wireless Sensor Networks" IEEE Volume 3, Issue 7, July 2013

[6] Shalli Rani, Jyoteesh Malhotra, Rajneesh Talwar," EEICCP—Energy Efficient Protocol for Wireless Sensor Networks" *Wireless Sensor Network*, Published Online July 2013

[7] K. Vinoth Kumar1, S.Karthikeyan2," Multihop Energy Efficient Reliable and Fault Tolerant Routing Protocol for Wireless Sensor Networks" ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 2, February 2013

[8] Rabia Noor Enam1, Syed Misbahuddin2(IEEE Member), Mumtazul Imam(IEEE Member), "Energy Efficient Round Rotation Method for a Random Cluster Based WSN" IEEE Department of Computer Engineering 2012

[9] Neha Rathi1, Jyoti Saraswat2and Partha Pratim Bhattacharya3 ,"A Review On Routing Protocols For Application In Wireless Sensor Networks" International Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.5, September 2012

[10] Monica R Mundada, Savan Kiran, Shivanand Khobanna, Raja Nahusha Varsha and Seira Ann George," A Study On An Energy Efficient Routing Protocol In Wireless Sensor Network" International Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.3, May 2012

[11] Meenakshi Diwakar and Sushil Kumar," An Energy Efficient Level Based Clustering Routing Protocol For Wireless Sensor Networks" International Journal Of Advanced Smart Sensor Network Systems (IJASSN), Vol 2, No.2, April 2012

[12] Sourabh Jain, Praveen Kaushik, Jyoti Singhai," Energy Efficient Maximum Lifetime Routing For Wireless Sensor Network" International Journal Of Advanced Smart Sensor Network Systems (IJASSN), Vol 2, No.1, January 2012

[13] G.Kalpana, Dr.T.Bhuvaneswari," A Survey on Energy Efficient Routing Protocols for Wireless Sensor Networks" 2nd National Conference on Information and Communication Technology (NCICT) 2011

[14] Vinay Kumar[1], Sanjeev Jain[2] and Sudarshan Tiwari, " Energy Efficient Clustering Algorithms in Wireless Sensor Networks: A Survey" IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 2, September 2011

[15] Md. Golam Rashed, M. Hasnat Kabir, Shaikh Enayet Ullah," WEP: An Energy Efficient Protocol For Cluster Based Heterogeneous Wireless Sensor Network", International Journal of Distributed and Parallel Systems (IJDPS) Vol.2, No.2, March 2011

Cite this article as: Jayashri Gaurkar, Kanchan Dhote. "Report on Design of distributed Energy Efficient and reliable Routing Protocol for Wireless Sensor Networks." *International Conference on Information Engineering, Management and Security (2015)*: 208-214. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS036		**eAID**	ICIEMS.2015.036

Effective Fusion Mechanism for Multimodal Biometric System-Palmprint and Fingerprint

Ms. Bhagyashree Madhukar Kale[1], Prof. P.G. Gawande[1]
[1]Dept. of Electronics and Telecommunication, Vishwakarma Institute of Information Technology, Pune, India.

Abstract: *Security is main issue in every where today. This paper proposes the multimodal biometrics system (MBS) for identity verification using three traits i.e., finger print, iris and palm print. The proposed MBS system is designed for applications like authentication where the training database contains a finger print images and palm print for each individual. Palm print is chosen as a biometric trait as no two palm print match unless they are of the same person also palm has a good vascular pattern making it a good identifying factor for an individual as compared to other biometric traits. The images captured by the designed hardware are preprocessed using Image enhancement techniques and Features are extracted by Gaussian kernel, Gabor Filter and Principal Component analysis. These feature vectors are fused at feature level and later matched by using Euclidean distance or Manhattan distance. Quality measures are also found for these above modalities.*

I. INTRODUCTION

Biometrics refers to metrics related to human characteristics. This system use biological information about person to create a detailed record of their personal characteristics. Nowadays biometric system provides highly secure and authenticated is low-cost, non-invasive and provides easy access to Biometric Authentication called multimodal biometric system (MBS) system. Biometric system has categorized into two main types- 1) Unimodal Biometric System (UBS), 2) Multimodal Biometric System (MBS). UBS has developed from many years but it suffers from various problems like enrollment problems, insufficient accuracy .The biometric system uses two or more biometric modalities for authentication. In this paper, we propose to an effective fusion mechanism for MBS. We use modalities like Fingerprint, Palm print. After capture, preprocessing is done and features are extracted and later fused at feature level. These features compared with stored feature template to give the accurate decision like accept or reject.

II. METHODOLOGY

First database of iris, palm print, fingerprint is collected. These are pre-processed individually and features are extracted which are fused to form a single feature vector. This vector is stored at the server side. Matching is then done i.e. test image is compared with stored feature vector, based on the minimum Euclidean distance.

Cite this article as: Ms. Bhagyashree Madhukar Kale, Prof. P.G. Gawande. "Effective Fusion Mechanism for Multimodal Biometric System-Palmprint and Fingerprint." *International Conference on Information Engineering, Management and Security (2015)*: 215-218. Print.

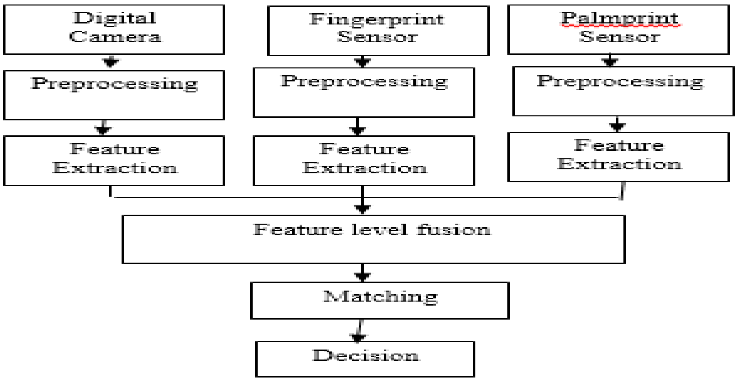


Figure 1. Block Diagram

A. Data acquisition and preprocessing

1) Palm print

A system is designed for capture of palm vein images based on principle of near infra-red imaging. The system consists of a CCD camera and a lightening system. The CCD camera is modified to see only IR light which involves removal of infrared filter inside the camera as shown in Fig 2. Finally to make this IR-only camera the visible light needs to be filtered out for the same filter is made for the front of the lens out of fully exposed 35mm color negative film. The red tinged glass in Fig 2 (c) The lightning system consists of an Matrix arrangement of Near IR LED`s (SFH4550) 850nm. As matrix arrangement of LED`s offer best distribution of intensity with uniformity matrix arrangement is chosen. A variable power supply having rating 0-12V and 0.4A is designed; with the help of an adjustable potentiometer the intensity is varied.

The camera i-Ball Face to Face c12.0 (640x480) is mounted on a wooden stand. The LED assembly is placed below. A simply paper material box is placed over the LED assembly over which the palm is placed. With the help of a switch the LED assembly can be turned ON and OFF. The camera has a USB interface hence the image captured are stored on the computer. Fig 2(a) shows the system for palm print capture. The image captured by the system has clear visibility of palm print as shown in Fig 2(b).

The above system could also be used to capture palm vein images. The captured hardware design is very cost effective and the palm vein images captured when combined with authentication algorithms provide good accuracy.

Palm vein images are low in contrast hence the data values are remapped by contrast adjustment. Gray values between low to high are mapped onto bottom to top. Depending on the Gamma correction factor [7] the mapping of values between input and output image may be linear. Further after preprocessing palm vein images boundary of the hand is found out by border tracing algorithm which locates the position of fingers to find Region of interest (ROI) by Euler's distance diagram.

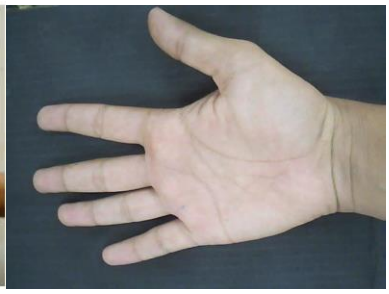


Fig.1 i-ball camera　　Fig.2 a) Palm print capture system　　Fig.2 b) Palm print

2) Fingerprint

A contactless capture of Fingerprint is performed. One hundered subjects both male and female took part in the experiment. The capture of fingerprint involves using Xpro night vision web camera shown in Fig 1. It`s a 20.0Mega Pixel Camera with a high quality CMOS sensor and five glass optical coating lens. The frame captures size of 640x480. It is driver free for windows XP and above. It`s compatible with MAC with 2.0 USB port and windows. The camera also has USB port and frame rate 30F/S @ VGA. Fig 1(b) shows a fingerprint image captured by the camera. A completely contactless fingerprint capture makes it a highly hygienic process.

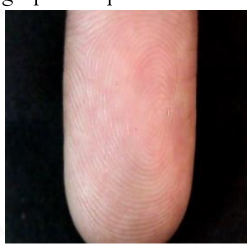


Fig. 3a) Xpro camera　　Fig. 3b) captured fingerprint

Cite this article as: Ms. Bhagyashree Madhukar Kale, Prof. P.G. Gawande. "Effective Fusion Mechanism for Multimodal Biometric System-Palmprint and Fingerprint." *International Conference on Information Engineering, Management and Security (2015)*: 215-218. Print.

B. Feature Extraction.

Feature Extraction is transforming the input data into the set of features. Gaussian kernel with Principal Component Analysis, Gabor filter is used for palmprint and fingerprint feature extraction respectively.

1) Palm print Features:

Principal lines are extracted using Gaussian Kernel and canny edge detector. Features are selected like length of two principal lines and orientation between them for each individual. Then PCA applied for feature extraction.

PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. Its objective is to reduce the large dimensionality of the data space to smaller intrinsic dimensionality of feature space. PCA involves [8] calculation of Eigen value decomposition of data covariance matrix after mean centering the data.

2) Fingerprint Features:

Gabor filter impulse response is defined by a sinusoidal wave (a plane wave for 2D Gabor filters) multiplied by a Gaussian function.

Gabor filters [5] are characterized by spatial frequencies and orientation capabilities hence they are used for finger print feature extraction. Spatial frequencies and orientation are important characteristics of textures in images. Gabor filters are used for finger print feature extraction. A Gabor filter is obtained by combining a sinusoid with a Gaussian. Gabor filter centered at origin with frequency θ and orientation ϕ is,

C. Fusion

In feature level fusion, different feature vectors extracted from proposed biometric trait are combined together into a single feature vector .This process undergoes two stages which are feature normalization and feature selection.

III. EXPERIMENTS AND RESULTS

The methods are implemented by MatLab 8.2® on a computer with Intel® Core(TM) 2-Duo T6570 CPU @ 2.10GHz, 4GB RAM, Windows 7, 64-bit operating system. The following figures show the experimental results. Experiments are performed on 100 persons including 60 male and 40 female. The age distribution is from 20 years to 60 years. 10 images of each person were captured of all three modalities i.e. Fingerprint and palm print. Hence in total the database has 2000 images of 100 persons. For training, 600 images were used for each modality and 400 images were used for testing. In Fig 4 (a) is the original input image of size 640x480 captured from Xpro web camera. Fig 4(b) represents the histogram of the original image. Histogram equalization is applied to the original image which expands the range of gray level near Histogram maxima; hence this transformation improves the detectability of many image features. In Fig 5 (a) is a fingerprint image 640x480 captured by touchless fingerprint capture camera Xpro 20.0MP.

Fig.4a) Fingerprint Fig. 4b) Palm print

Palm print images require ROI extraction, then principal lines are extracted and length is calculated.

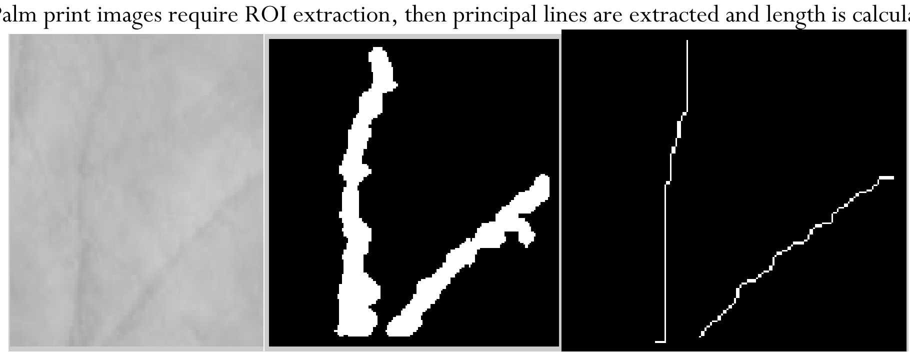


Fig.5a) ROI extracted Fig.5b) Principal line detected Fig.5c) Thinned principal lines

Cite this article as: Ms. Bhagyashree Madhukar Kale, Prof. P.G. Gawande. “Effective Fusion Mechanism for Multimodal Biometric System-Palmprint and Fingerprint.” *International Conference on Information Engineering, Management and Security (2015)*: 215-218. Print.

The feature vector matrix of fingerprint and palm print is of the order 13900x1, 10816x1 respectively. Feature level fusion is applied to the above feature vectors and a fused matrix is obtained of order 217738x900. Testing is carried out using 100 images which gives accuracy of 99%.

REFERENCES

[1] Wei Li1, Lei Zhang, David Zhang and Jingqi Yan1,"Principal Line Based ICP Alignment for Palmprint Verification",*IEEE Transactions on International Conference on Image processing 2009*

[2] D. Zhang, A.W.K. Kong, J. You, and M. Wong, "On-line palmprint identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1041-1050, 2003.

[3] Basanti B. Sawant1, M. Talib2, Sagar S. Jondhale3 , Pradeep M. Patil4 "Palmprint Recognition Using Radon Transform" *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 2, Issue 12, December 2013

[4] http://in.mathworks.com/help/images/contrast- adjustment.html

[5] D. J. Gabor, "Theory of communication," *IEEE, vol. 93, no. 26, pp. 429–457, 1946.*

[6] T. A. Albert **,**S. Ganesan "Applications of Principal Component Analysis in Multimodal Biometric Fusion System" *European Journal of Scientific Research* ISSN 1450-216X Vol.67 No.2 (2012), pp. 248-259

[7] P. A. Johnson, F. Hua, and S. Schuckers "Comparison of Quality- Based Fusion of Face and Iris Biometrics" Foundation under Grant Nos. 0325333 & 0325640, *the Department of Homeland Security, and the Center for Identification Technology Research*,.2011

[8] Mr. J.Aravinth1, Dr.S.Valarmathy "A Novel Feature Extraction Techniques for Multimodal Score Fusion Using Density Based Gaussian Mixture Model Approach" *International Journal of Emerging Technology and Advanced Engineering,* ISSN 2250-2459, Volume 2, Issue1, January 2012

Cite this article as: Ms. Bhagyashree Madhukar Kale, Prof. P.G. Gawande. "Effective Fusion Mechanism for Multimodal Biometric System-Palmprint and Fingerprint." *International Conference on Information Engineering, Management and Security (2015)*: 215-218. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS037		**eAID**	ICIEMS.2015.037

A Survey on Quality Assessment And Enhancement of Document Images

Pooja Sharma[1], Shanu Sharma[2]
[1,2] Department of Computer Science & Engineering, Amity University, Noida, India

Abstract: *With the advancement of technology, there has been a tremendous rise in the volume of captured and distributed content. Image acquisition can be done with the help of scanners, cameras, smart phones, tablets etc. Document retrieval and recognition systems require high quality document images but most of the time, the images acquired suffer from various degradations like blur, uneven illumination, low resolution etc. To reduce the processing time and get good results, we require methods to evaluate and improve the quality of such images. This paper reviews the quality assessment methods and enhancement techniques for document images. It presents a survey of the work that has been performed in the field of document image quality assessment and enhancement.*

Keywords: Document image processing, Quality assessment , OCR accuracy, Camera captured documents

1. Introduction

Capturing of documents using a handheld camera is preferred to acquiring images by scanning these days for no contact images, flexibility and low cost involved. But quality of camera captured documents is lower than scanned documents as camera acquired documents contain various degradations such as blur, perspective distortions, character smear, uneven lightning etc. But the application areas cannot be ignored such as digitization of books, digitization of historical documents, finding text in scene images, mobile OCR etc. Document image quality assessment is a measure of text distortion. It is necessary to evaluate the quality of a document image to save the processing time by any document recognition system. Low quality documents give poor results. Document images contain various degradations such as blur, uneven illumination, perspective distortion, low resolution, smear etc. Hence there is requirement of enhancement algorithms for improving the quality of degraded images for enhancing the readability of the documents as well as improve the performance of document processing systems. Therefore this paper is an attempt to review the various quality assessment and enhancement methods for document images. It is divided into three sections. Section II covers quality assessment methods and metrics while section III focuses on quality assessment techniques whereas section IV concludes the work with direction of future work.

2. Quality Assessment of Document Images

Quality assessment refers to the process of evaluating the image quality with some metrics based on different features. There is a mandatory need of maximizing the readability of the document images. Since some images are degraded in the process of acquisition,

Cite this article as: Pooja Sharma, Shanu Sharma. "A Survey on Quality Assessment And Enhancement of Document Images." *International Conference on Information Engineering, Management and Security (2015)*: 219-223. Print.

we need assessment methods so that enhancement can be performed. While some of the metrics correlate OCR accuracy with the degradations, other view it through human perception. Hence we have various methods for quality assessment in document images. There are broadly three image quality assessment methods: full reference, reduced reference and no reference methods. The full reference methods need a high quality image to assess the quality in comparison, reduced reference methods require some properties of the high quality image based on which the quality of an image can be assessed. Whereas no reference methods pose no requirement for a high quality image as they evaluate quality solely based on the properties of the input image. We have a greater need for no reference methods as we do not have a high quality image always available as required strictly in case of full reference methods while some. The major assessment methods reviewed from the previous work are as follows:

Xujun *et al.* [1] presents a method to estimate image quality by measuring the degree of degradation. It deals with blurring as well as several other degradations found in camera captured document images. The method predicts the impact of degradations on OCR error rate.

Jayant *et al* [2] created a dataset of camera captured documents containing various blur levels. A case study was also performed with three quality estimation methods for the prediction of OCR quality and their advantages and disadvantages were evaluated.

Tayo *et al.* [3] presents an objective user centred function that evaluates the distortion level in a document image. It focuses on human perception of quality which is required in case of historical images.

Rusinol *et al.* [4] focuses on the estimation of focus quality as a metric for quality estimation to predict OCR accuracy. Gradient based features are found to correlate best with OCR engines response.

Amina [5] presents a blur image quality metric based on MMD which is a nonlinear transform and the blur is treated as convolution noise.

Taegeun *et al*. [6] proposes a no reference method for predicting perceptual quality of camera captured images based on spectral statistics..

Ke *et al.* [7] presents a no reference image quality metric which combines two existing quality assessment models, one being the free energy based and the other based on structural degradation.

Qing *et al.* [8] focuses on a no reference blur assessment method related to perception utilizing the feature termed as GPS(Gradient Profile Sharpness)

Ming-Jun *et al.* [9] concentrates on no reference image quality assessment based on multi resolution decomposition methods. It takes into account natural scene statistics.

Karen *et al.* [10] brings out a new no reference image contrast feature to help in selecting the best operating parameters for image enhancement algorithm.

Table 1 lists out the various quality assessment methods and their unique focus areas as some of the methods take blur into consideration, while some deal with a mix of degradations.

Table 1. Summary of Various Quality Assessment methods

S.No	Author	Year	Source	Major Findings
1	Xujun *et al*	2011	IEEE	It presents a new no-reference image QA method that evaluates the impact of degradations in an image on OCR error rate.
2	Jayant *et al*	2013	CBDAR	It provides with a publically available database of camera captured documents containing various levels of blur. 3 image QA approaches are evaluated to bring out the best.
3	Tayo *et al*	2012	IEEE	It proposes a user centred objective function to evaluate level of degradation in a document image. This is a human perception of quality metric.
4	Rusinol *et al*	2014	IEEE	It focuses on a QA metric that estimates focus quality for processing in OCR and considers out-of-focus blur.
5	Amina *et al*	2011	IEEE	It presents a no-reference blur image QA metric based on MMD.
6	Taegeun *et al*	2014	IEEE	It proposes a no-reference QA metric based on spectral statistics.
7	Ke *et al*	2013	IEEE	It presents a no-reference QA metric that combines 2 previous QA models.
8	Qing *et al*	2013	IEEE	It concentrates on a no reference image blur QA metric which is perceptual and based on GPS.
9	Ming-Jun *et al*	2009	IEEE	It presents a no-reference QA metric that takes into account natural scene statistics.
10	Karen *et al*	2013	IEEE	It focuses on a no-reference QA metric that provides new features for selecting optimal parameters.

Cite this article as: Pooja Sharma, Shanu Sharma. "A Survey on Quality Assessment And Enhancement of Document Images." *International Conference on Information Engineering, Management and Security (2015)*: 219-223. Print.

3. Quality Enhancement of Document Images

Images acquired from handheld cameras suffer from a lot of degradations such as blur, uneven illumination, perspective distortion, low resolution, character smear etc. Degradations lead to low readability, deteriorated recognition and hampers the performance of a document processing system. Hence it becomes necessary to enhance these images in order to remove these degradations or reduce their effect so as to achieve higher OCR accuracy in most of the cases. While some of the researchers have worked on improving the binarization technique which is an initial step before OCR, others have worked on specific issues like blur, illumination, resolution etc. The previous work done in this area is summarized as follows:

Bukhari *et al.* [11] proposes a new adaptive binarization technique for distorted camera captured documents. Global binarization is found to give better results for degraded camera captured documents.
Utpal *et al*[12] presents an image enhancement method for historical documents based on ICA to improve their OCR accuracy.
Fabian *et al.* [13] focuses on enhancement of ancient handwritings whose images are taken through multispectral scan and then their dimensionality is reduced utilizing LDA; the method is then evaluated using OCR.
Soo-Chang *et al.* [14] presents an image enhancement method that focuses on removal of uneven illumination using EMD(Empirical Mode Decomposition).
Aishwarya *et al.* [15] presents a method for enhancing the region of an image using specular reflection. This method focuses on the scene images and deals with the degradations that are caused by light.
Henry *et al*. [16] presents a technique for image enhancement of camera captured document images. It addresses poor illumination, fading and noise and helps in achieving improved OCR accuracy.
Zhan *et al.* [17] proposes a resolution enhancement algorithm based on interpolation to improve recognition rate in a document with low resolution.
Chethan *et al.* [18] presents an image enhancement technique for lightning correction and improve data quality for camera captured documents using homomorphic and morphological filtering.
Khairunnisa *et al.* [19] proposes an enhancement technique for a low contrast image and focuses on insufficient light using fuzzy technique by modifying the membership function. The method is found to perform in minimum processing time when compared to others.
Nibal *et al.* [20] proposes an image enhancement method for removing blur by making use of sparse based methods using non local means for document images.
Jian *et al.* [21] proposes an image enhancement algorithm based on watershed segmentation for solutions such as lightning correction and color correction as well as text sharpening.
Gaofeng *et al.* [22] brings out a new method for correcting geometric distortions in camera captured document images with the help of two structured beams.
Anand *et al.* [23] focuses on solution of OCR issues related to camera captured images and also provides a solution for the detection of misaligned text.

Table 2 lists out the various quality enhancement techniques, what degradations they deal with and what is the use for each technique.

S.No	Author	Year	Source	Major Findings
1	Bukhari *et al*	2011	Journal	It proposes a new adaptive binarization technique to improve OCR processing for distorted document images.
2	Utpal *et al*	2013	IEEE	It presents an image enhancement method for historical documents based on ICA to improve OCR accuracy.
3	Fabian *et al*	2012	IEEE	It discusses an image enhancement technique using LDA for multispectral scans to improve OCR accuracy.
4	Soo-Chang *et al*	2014	IEEE	It focuses on an image enhancement technique to correct uneven illumination in text images using EMD.
5	Aishwarya *et al*	2011	IEEE	It presents an image enhancement techniques for scene images utilizing specular reflection.

Cite this article as: Pooja Sharma, Shanu Sharma. "A Survey on Quality Assessment And Enhancement of Document Images." *International Conference on Information Engineering, Management and Security (2015)*: 219-223. Print.

6	Henry *et al*	2014	IEEE	It presents an image enhancement technique called retinex to address poor lightning,noise and fading to improve OCR accuracy.
7	Zhan *et al*	2013	IEEE	It proposes a resolution enhancement technique using interpolation to improve recognition.
8	Chethan *et al*	2013	IEEE	It addresses uneven illumination, color abnormalities and poor data quality using homomorphic filters.
9	Khairunnisa *et al*	2009	IEEE	It deals with inadequate lightning by modifying membership function in fuzzy logic and has minimum processing time.
10	Nibal *et al*	2013	IEEE	It provides a technique to remove blur using non local means to improve OCR accuracy.
11	Jian *et al*	2007	CBDAR	It corrects poor illumination, uneven color and performs text sharpening.
12	Gaofeng *et al*	2014	CVPR	It provides a method to correct geometric distortions using 2 structured beams.
13	Anand *et al*	2009	Springer	It provides a solution for detecting misaligned text.

4. Conclusion

There exist various quality metrics for document images and OCR accuracy is found to be an important quality indicator for document images. Since camera captured document images suffer from degradations such as blur, uneven lightning, low resolution etc. Though there have been some enhancement techniques developed to enhance the OCR accuracy of these images, most of them address blur. Therefore there is a requirement of more enhancement techniques that deal with various degradations and help in improving the OCR accuracy rate of camera captured document images. Further work can be done in the direction of creating enhancement algorithms that deal with various types of degradations and doing work related to out of focus camera captured images as major work has been done for motion blur and out of focus images still need more attention to be enhanced and get better results.

REFERENCES

[1] Xujun Peng, Huaigu Cao, K.Subramainian, R.Prasad, P.Natarajan, (2011) “Automated image quality assessment for camera-captured OCR”, 18th IEEE International Conference on Image Processing(ICIP), pp.2621-2624

[2] Jayant Kumar, Peng Ye, D.Doermann, (2013)“A Dataset for Quality Assessment of Camera Captured Document Images ”,International Workshop on Camera Based Document Analysis and Recognition (CBDAR), pp.39-44.

[3] T.Obafemi-Ajayi, G.Agam, (2012) “Character-Based Automated Human Perception Quality Assessment in Document images”,IEEE Transactions on System,Man and Cybermet ics,pp.584-595

[4] M.Rusinol, J.Chazalon, J.-M. Ogier, (2014)“Combining Focus Measure Operators to Predict OCR Accuracy in Mobile-Captured Document Images”, 11th IAPR International Workshop on Document Analysis Systems(DAS),pp.181-185

[5] A.Serir, (2011)“No-reference blurred image quality assessment”, 3rd European Workshop on Visual Information Processing (EUVIP), pp.168-173

[6] Taegeun Oh, Park Jincheol, K.Seshadrinathan, Sanghoon Lee, A.C. Bovik,(2014) “No-Reference Sharpness Assessment of Camera-Shaken Images by Analysis of Spectral Structure,” IEEE Transactions on Image Processing, vol.23, no.12,pp.5428-5439

[7] Ke Gu, Guangtao Zhai, Xiaokang Yang, Wenjun Zhang, Longfei Liang, (2013)“No-reference image quality assessment metric by combining free energy theory and structural degradation model,” IEEE International Conference on Multimedia and Expo(ICME), pp.1-6

[8] Qing Yan, Yi Xu, Xiaokang Yang,(2013) “No-reference image blur assessment based on gradient profile sharpness,”IEEE International Symposium on Broadband Multimedia Systems and Broadcasting(BMSB) , pp.1-4

[9] Ming-Jun Chen, A.C. Bovik, (2009)“No-reference image blur assessment using mult iscale gradient,” International Workshop on Quality of Mult imedia Experience(QoMEx) , pp.70-74

[10] K.Panetta, Chen Gao, S.Agaian (2013)“No reference color image contrast and quality measures," IEEE Transactions on Consumer Electronics, pp.643-651

[11] S.S. Bukhari, F. Shafait, T.M. Breuel, (2009) “Adaptive binarization of unconstrained hand-held camera-captured document images”, Journal of Universal Computer Science,vol.15,no.18,pp.3343- 3363

[12] U.Garain, A.Jain, A.Maity, B.Chanda, (2008) “Machine reading of camera-held low quality text images: An ICA-based image enhancement approach for improving OCR accuracy”, 19th International Conference on Pattern recognition(ICPR),pp.1-4

[13] F.Hollaus, M.Diem, R.Sablatnig (2014) “Improving OCR Accuracy by Applying Enhancement Techniques on Multispectral Images”,22nd International Conference on Pattern Recognition(ICPR),pp.3080-3085

[14] Soo-Chang Pei, M.Tzeng, Yu-Zhe Hsiao,(2011) “Enhancement of uneven lightning text image using line-based Empirical Mode Decomposition”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.1249-1252

Cite this article as: Pooja Sharma, Shanu Sharma. “A Survey on Quality Assessment And Enhancement of Document Images.” *International Conference on Information Engineering, Management and Security (2015)*: 219-223. Print.

[15] A.Visvanathan, T.Chattopadhyay, U.Bhattacharya,(2013) "Enhancement of camera captured text images with specular reflection", Fourth National Conference on Image Processing and Graphics(NCVPRIPG) ,pp.1-4

[16] H.Kiragu, E.Mwangi, (2012) " An improved enhancement of degraded binary text document images using morphological and single scale retinex operations", IET Conference on Image Processing(IPR), pp.1-6

[17] Zhan Li, Jiangao Luo,(2011) " Resolution Enhancement from Document Images for Text Extraction",5th FTRA International Conference on Multimedia and Ubiquitous Engineering(MUE), pp.251-256

[18] H.K. Chethan, G.H. Kumar(2011) "Illumination Reduction for Camera Captured Documents Using Block Processing and Homomorphic Filtering Technique", International Conference on Communicat ion Systems and Network Technologies(CSNT), pp.350-354

[19] K.Hasikin, N.A.M. Isa, (2012) " Enhancement of the Low Contrast Image Using Fuzzy Set Theory",14th International Conference on Computer Modelling and Simulation(UKSim), pp.371-376

[20] N.Nayef, P.Gomez-Kramer,J.-M. Ogier,(2014) " Deblurring of Document Images Based on Sparse Representations Enhanced by Non-local Means",22nd International Conference on Pattern Recognition(ICPR), pp.4441-4446

[21] Jian Fan,(2007) "Enhancement of Camera -captured Document Images with Watershed Segmentation",CBDAR,Curitiba,Brazil, pp.87-93

[22] Gaofeng Meng, Ying Wang, Shenquan Qu, Shiming Xiang, Chunhong Pan(2014) , "Active Flattening of Curved Document Images via Two Structured Beams ",IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp.3890-3897

[23] Anand Joshi, Mi Jhang, Ritesh Kadmawala, Karthik Dantu, Sameera Poduri, Gaurav S. Sukhatme,(2009) "OCRdroid: A Framework to Digitize Text Using Mobile Phones ", ICST International Conference on Mobile Computing, Applications and Services (MobiCASE), pp.273-292

Cite this article as: Pooja Sharma, Shanu Sharma. "A Survey on Quality Assessment And Enhancement of Document Images." *International Conference on Information Engineering, Management and Security (2015)*: 219-223. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS038		**eAID**	ICIEMS.2015.038

Video Depiction of KeyFrames- A Review

Deepika Bajaj[1], Shanu Sharma[2]
[1,2] CSE Department, Amity University, Noida, Uttar Pradesh, India

Abstract: *Nowadays, there are numerous, unstructured and voluminous videos which leads to high collection of data on web. Searching and navigating through these videos for meaningful information is a time consuming task, whereas a good 'summarized video' can provides a user determined information about particular video sequence in definite time limits. So, there is great need of extraction of semantic and useful information from videos for various multimedia applications. The video summarization is the novel and promising method of detecting relevant and informative data from videos and also aims to provide effective and efficient storage of relevant information. This technique of summarization leads to abstraction of most representative and relevant scenes from videos and concatenates to display as one successive and uninterrupted video and thus has been powered up the rapidly progressing research domain. In this paper, all latest and enhanced approaches of video shot detection have been discussed and summarized.*

Keywords- key frame extraction, face detection, shot boundary detection, feature extraction.

I.INTRODUCTION

In recent times, the quantity of videos has been increased day by day. Videos are ample, unorganized and redundant data streams which contain images, graphics and textual information such as label, keywords, etc. Almost, every field such as entertainment, news channels or advertisements, etc. [1] involves wide use of videos. However, people use to spend their large amount of time to download huge videos in order to evaluate that these videos are relevant or irrelevant. Browsing through large amount of videos is a time consuming, tough and very tiring job for human beings. Therefore, it is hard and painful work to extract the meaningful content or desired scenes from videos. Video summarization is the best and efficient solution to transform huge and amorphous videos into organized, structured and systemized manner. Summarization of videos is process of creating concise, clear, succinct and meaningful information. Generally, it contains following three complexities-

☐ The summarized video should involve the important and desired parts or scenes from the video which means it should be as concise and brief as possible. For example, video summarization approach has the ability to provide semantic portion to user as per his requirements from the huge videos.

☐ The summarized video should maintain the semantic meaning which means it should represents the good continuous association between scenes.

☐ The summary of video should not contain any redundancy/duplicity among scenes.

Cite this article as: Deepika Bajaj, Shanu Sharma. "Video Depiction of KeyFrames- A Review." *International Conference on Information Engineering, Management and Security (2015)*: 224-230. Print.

In order to generate optimal video summaries, the two different ways and means for creating efficient and effective retrieval of video are:- summarization and highlights [2]. The prior aims to deliver a summarized and concise storyline demonstration of a video while the later aims to extract the affective data from the video. Video summarization is a technique that is important and suitable in context of searching and retrieving desired part of video. Summarizing the large videos into small videos is very helpful to people who can get the main concept or idea about the movie without watching the full video. This aims to provide the new view mode to audience. For example, by watching movie trailer, the people get to know that movie is romantic, comedy or action movie. Therefore, this technique is very significant and useful in classifying the videos and movies as well and explains the main concept and idea of video.
The summarized videos have wide variety of advantages in various fields for various purposes. Summarizing videos has wide variety of applications such as video indexing, estimating the rating of movie, etc. [3]. The parsing of videos is shown below-

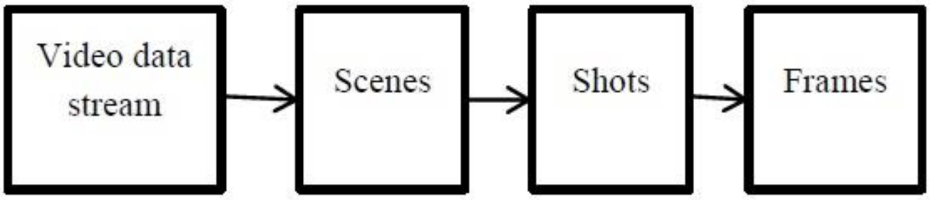


Fig1: Flowchart of parsing of video

A. Shot detection and scene detection in videos

Initially, input is taken as "video" and it consists of distinct scenes, then from those scenes further different shots are analyzed and extraction of frames is performed. The frame extraction leads to generate the most representative frames (keyframe extraction) from already extracted frames. A scene can be defined as "a section/ division of movie or video in which the set is permanent, time is uninterrupted and the action is established in one place". The scene can also be described as a collection of video shots that satisfy certain similarity along with semantic and meaningful information [4]. A scene must abide three rule i.e. unitary space, time and action.

According to Bordwell et al. [5], three types of scenes exists in videos i.e. action scene, conversational scene and suspense scene. In order to detect scene, there are many approaches and procedures such as graph based segmentation method, color histogram based method etc. A scene can contain various transitions in videos. Transitions are the variations from one scene to another. During the changing of shot or scene, discontinuities are found. There are four types of transition [6]:

1) Hard cut- This is the most simple and basic kind of transition. When one shot replaces the other instantly, then hard cut occurs. In normal feature film, there are thousands number of hard cuts. Cuts are useful for the continuous and enhanced movement of the movie.

2) Fade- Fade is further divided into two parts-

☐ *Fade in* – It is basically used in the starting of the movie. Fade-in occurs when scene changes gradually from black /solid color to picture/image.

☐ *Fade out-* It is basically used in the ending of the movie. Fade-out occurs when scene changes gradually from current image to solid paint.

Cite this article as: Deepika Bajaj, Shanu Sharma. "Video Depiction of KeyFrames- A Review." *International Conference on Information Engineering, Management and Security (2015)*: 224-230. Print.

Fig2: Hard cut effect and fade effect

3) Dissolve- It is also known as overlapping. Dissolve occurs when there is replacement of one shot with the next gradually. One shot disappears as the next shot appears. For some time both shots overlap each other and used to represent the passage of time.

Fig3: Dissolve effect

4) Wipe- Wipes are dynamic kind of transition. When one shot pushes the other off frame, then wipe takes place.

Fig-4 Wipe effect

Scenes are the combination of extracted shots. For each detected shots, variable count of keyframes can be determined. The keyframe which is extracted first is always lies near a shot boundary. These frames can be further combined to form semantic and most representative scenes.

B. Key Frame Extraction

Key frames extraction is the most significant step in abstraction of videos. Extraction of key frames leads to set of meaningful images from the sequences of videos. Many researches are still working for the better and improved automatic system of key frame extraction. This step is quite effective and efficient in providing the summarization of video.
Below some definitions to extract keyframes are listed [7]:

☐ Reference Frame- refers to first frame in every shot.
☐ General Frames- refers to rest of the frames other than reference frames;
☐ "Dynamic shot factor" max (i)- The max x2 histogram within shot i;
☐ Static shot and dynamic shot- if max(i) is larger than mean calculated then it is considered to be dynamic shot otherwise it is static shot.
☐ *Fm* (*k*) : The *k*th frame within the current shot, k=1,2,3… *Fm*(*k*) (*Fm*(*k*) is the total number of the current shot).

The following steps are used generally in extraction of key frames:

☐ Find the difference between reference frame and general frame.
☐ Look which shot has the maximum difference.
☐ Determine the shot type i.e. static or dynamic shot.
☐ Then, see the position of key frame.

II. Video Shot Boundary Detection (SBD) Algorithms

Todate, many scholars and researchers are doing work to develop more reliable and accurate algorithms that can results into more precise shot boundaries. Earlier technology was more focused on cut detection whereas latest approaches are more concentrated on gradual transitions detection. Broadly categorizing or classifying the shot boundary detection methods are listed below:

(a) Basic approach
(b) Feature based approach
(c) Segmentation based approach
(d) Texture based approach

☐ *Pixel based difference-* This is the basic approach towards shot boundary detection. In this method, intensity of pixels is evaluated by taking two consecutive frames and comparing pixel by pixel. When the intensity of pixels is more than threshold, then it is referred to scene change [8]. A number of algorithms have been implemented for calculation of pixel difference. 3 X 3 averaging filter is used in [9] [10]. The limitations in this method is setting threshold manually. Hence, this method is slow and setting threshold manually is not acceptable concept.

☐ *Statistical based difference-* It is modified and overcome the limitations of pixel based approach. It is sensitive to noise. In this method, each frame is divided into blocks and further each block"s mean and standard deviation is evaluated. Each block some characteristics or features of each pixel in that particular block between consecutive frames are equated [6]. The limitations of this approach are it is sensitive to noise and quite complex as it includes statistical parameters and calculations.

☐ *Transform based difference-* It includes various transformation methods like Discrete Cosine Transformation (DCT) coefficients. Using various transformation approaches, it computes compression difference [6].

☐ *Histogram based difference-* Histograms are the most significant and important technique to find the shot boundaries in videos. In this method, find the difference between frame n and frame n+1 that results to change of color content within shots. Histograms are based on certain concept such as bin to bin, distance formula and intersection. In bin to bin, calculate the difference between color component (R, G, B) of two consecutive frames. If the difference is greater than threshold, then a shot is declared. Distance based approach include Chi- square distance, Manhattan distance, Swain and Ballard distance. Color based histogram is also one of the advanced version of histogram based detection algorithms.

☐ *Edge change ratio-* It is a feature based approach. First of all, edge operator is selected in order to find out the edges. After applying canny edge detector [11], calculate no of pixels in each edge. Edges of successive frames are identified. Then, in order to find new edges are appeared in image or old edges are vanished, edge pixels are combined with neighboring pixels in other image.

☐ *Graph theory based-* It is a latest and feature based approach to evaluate the shot boundary. In this procedure, "color" feature is defined and HSV color model is used to extract color content. Then, difference between frames is computed and apply graph based algorithm on frames of videos which are further divided into numerous different sets. Cut and gradual shots are detected through this algorithm. Cut and gradual changes [12] refers to when two consecutive frames appropriate to different set have different characters on those frames, then it is said to be cut sand gradual shot.

Cite this article as: Deepika Bajaj, Shanu Sharma. "Video Depiction of KeyFrames- A Review." *International Conference on Information Engineering, Management and Security (2015)*: 224-230. Print.

☐ *Information theory based-* It is a texture based approach and wavelet transformation is used to extract the texture and then define the difference based on MI (mutual information) and co-occurrence MI of texture feature. To analyze the image in different scales i.e. information of high frequency defines basically the texture feature and information of low frequency defines the color feature. So, to measure the dissimilarity, the MI and Information entropy is used. The wavelet coefficient is evaluated by discrete wavelet transform along with some advantages of less complexity and orthogonality.
Performance of video shot boundary detection methods can be evaluated by following measures [13]:
(a) Recall- This measure is also known as function of true positive or sensitivity. It is the ratio of detections of correct experimental to the detection of correct and missed.

Recall= correct
correct+missed
(b) Precision- It is the ratio of detections of correct experimental to the detection of correct and false.

Precision=correct
correct+false
(c) F-measure- It is defined as:

F-measure=2*Recall+Precision
Recall+Precision

III. Extraction Of Features

Feature extraction is the most important and essential step in summarization of videos. Depending on various parameters, features are extracted and are classified as low level features and high level features. Low level features are extracted on the basis of texture, color and shape. Every domain have distinct technology for evaluating the graphical and pictorial information and moreover, it has own pros and cons. Sometimes, the information present in video and the understanding of an individual differs due to understanding of graphic information which is basically depending on low level features. This difference of understanding is known as semantic gap. The problem of semantic gap is the biggest complexity that the technical society faces. In order to overcome, feedback of users is collected and updates the required information but then again this information may not be attainable every time and it is somehow a supervised approach. Therefore, now high level features comes into picture in order to deal with the shortcomings of semantic gap.
High level features acts as the bridge and filling the gap between user understanding and information exists. Human plays a vital role in understanding the graphic content of Human determined research. To recognize the face and upper portion data of human is very useful and supportive in understanding gender and age of person. The high level features are listed below:-

(a) Processing of face
(b) Motion magnitude
(c) Duration of shot
(d) Identification of gender

(a) *Processing of face*- Many studies have been done in face processing domain from last many years. Processing of face is very important and basic element in field of security, networking and surveillance videos. The facial expressions and motions of human face help us to recognize the emotions, fitness quality and social interaction information. Processing of face contains mainly three types of steps-

(i) Detection of face
(ii) Clustering of face
(iii) Recognition of face

(i) Detection of face- It is very basic and first step in face processing [14]. A number of algorithms have been introduced for the purpose of detecting faces in videos. The importance and significance of person in video depend on the number of occurrences of particular face in particular video. In [15] Li et al. provides promising results for faces along with different measures in video. Using this method, extraction of face features provides the size and position of all face detected and the total of hits [16]. The technique used for face detection is [17] local successive mean quantization transform (SMQT) and Sparse network of Winnows classifier (SNoW) is experimentally verified to be optimal choice for face detection process. Face detection is a challenging task due to following factors such as postures of faces, intensity, different backgrounds and invariant level of contrast.

The biggest challenge in detecting faces from videos is that characters do not face the camera in full front. So, this problem gives rise to the concept of head turning and head rotation [18]. This concept can easily deal with above mentioned limitation of face direction.

Everingham et al. [19] proposed an algorithm in which characters in videos and films are label automatically. Foucher et al. [18] used spectral clustering to detect faces of actors.
(ii) Clustering of face- It is performed by considering the degree of interaction between characters and can be computed by Mutual information (MI). MI is a significant and beneficial process to find out the similarities between characters.
(b) Motion magnitude- The information of magnitude and orientation is contained in vector. Motion magnitude is quite useful for indicating highlights. The fast motion frames as well as slow motion frames both are involved in highlights. In [20], there is a scene in movie in which a actress‟s was crying ,so therefore a scene was slow moving scene and can be considered as emotional scene. In order to identify fast and slow motion frames, we need to perform two steps- first, normalize the
(iii) Recognition of face- The process has two pros which have to be deal in face recognition. The first one is face detection process also involve frontal face along with profile between characters. And, second is the huge number of faces detected for recognition process is a heavy task to be considered. The faces are tagged with face number, shot number and frame number when the faces are obtained after face detection process. The recognition of face is implemented with eigenfaces technique [13] is highly appropriate because our foremost concern is effectiveness and speediness along with easiness. In [16], multiple classes face recognition procedure is implemented which leads to reduce the computing time. To improve the speed of process, interaction score computing is calculated and along with it interaction graph and phenograph between characters in videos are represented.
(b) Motion magnitude- The information of magnitude and orientation is contained in vector. Motion magnitude is quite useful for indicating highlights. The fast motion frames as well as slow motion frames both are involved in highlights. In [20], there is a scene in movie in which a actress‟s was crying ,so therefore a scene was slow moving scene and can be considered as emotional scene. In order to identify fast and slow motion frames, we need to perform two steps- first, normalize the average motion „a‟ of every block within the i-th frame. After normalizing the average motion, in second step, imply both fast frames as well as slow motion frames and suppress average motion frames to calculate the highlight score.
(c)Duration of shot- It is most important parameter in summarization of video and also specifies and defines the time limit of shot appeared in video. To attract viewer‟s attention, directors use long duration shots [20] and highlights extracted from the videos also involve long duration shots. The duration of shot and motion magnitude and highlights have relationship and association between them. Low motion magnitude and long shot duration refers to highlight love scenes and emotional scenes and high motion magnitude and short shot duration are used to represent an action scene.
(d)Identification of gender- It is very important to identify the gender i.e. male or female in videos and parameters such as eyes, nose, and chin all this helps in identification of gender. The two basic methods that are defined for identification of gender are- (a) related to facial features and (b) observe the relationship between facial features [3]. A lot of researchers and scholars have been developed numerous algorithms and techniques for detecting gender. The principal component analysis (PCA) and neural network have provided very fine results. In order to detect frontal features from face, a selection of genetic features subset was used.

IV. DISCUSSION AND CONCLUSION

Video summarization has many applications in various domains such as scene tagging, retrieving scenes, indexing of videos, highlighting useful and semantic scenes, etc. This paper summarizes all the recent methods and techniques used for video summarization. Many researchers and scholars have done lots of work in current research area and still work is going on high peak. These technologies can be advanced and refined in order to detect shot boundary more accurately. The extraction of keyframes needs to be more accurate and represents the semantic and meaningful portion from videos. It is an essential step so that no important information should be missed. Furthermore, research and work can be done in this area to develop more efficient and fast video shot detection algorithms. This application provides the user with most representative and needful information from video as per to his desired requirement and delivers the new mode of viewing the video. Moreover, summarization of videos saves time, energy and provides ease to an individual.

REFERENCES

[1] [1] *Yihong Gong and Xin Liu,* "Generating Optimal Video Summaries", Multimedia and Expo, 2000. ICME 2000, IEEE International Conference 2000 (Volume:3), pp. 1559 – 1562.

[2] [2] Chen, Oscal T.-C. ,Jhen Jhan Gu, Chih-Chang Chen and Ping-Tsung Lu, "Automatic Highlights Extraction for Drama Video Using Music Emotion and Human Face Features", 4th International Conference on Awareness Science and Technology (iCAST) 2012, pp.104 – 108.

[3] [3] Muhammad Nabeel Asghar et.al, "A Framework for Feature Based Dynamic Intravideo Indexing", Proceedings of the 19th International Conference on Automation & Computing, Brunel University, London, UK, 13-14 September 2013.

[4] [4] Ruxandra Tapu, Bogdan Mocanu and Ermina Tapu ; "Automatic Scene / DVD Chapter Extraction in Hollywoodian Movies" in IEEE, 2013.

[5] [5] D. Bordwell and K. Thompson "Film Art: An Introduction, 5th edition". *New York: McGraw-Hill*, 1997.

[6] [6] Nikita Sao and Ravi Mishra**, "**A survey based on Video Shot Boundary Detection techniques", International Journal of Advanced Research in Computer and Communication Engineering, (Volume: 3, Issue: 4), April 2014.

[7] [7] Ganesh. I. Rathod and Dipali. A. Nikam ," An Algorithm for Shot Boundary Detection and Key Frame Extraction Using Histogram Difference", International Journal of Emerging Technology and Advanced Engineering,(Volume: 3, Issue: 8),August. 2013, pp.155-163.

[8] [8] Mohini Deokar and Ruhi Kabra, " Video Shot Detection Techniques Brief Overview", International Journal of Engineering Research and General Science (Volume:2, Issue:6), October-November, 2014, pp-817-820.

[9] [9] Ravi Mishra ,S.K.Singhai,M. Sharma , "Comparative study of block matching algorithm and dual tree complex wavelet transform for shot detection in videos", Electronic system, signal processing and computing technologies(ICESC), 2014 International Conference, Jan 2014.

Cite this article as: Deepika Bajaj, Shanu Sharma. "Video Depiction of KeyFrames- A Review." *International Conference on Information Engineering, Management and Security (2015)*: 224-230. Print.

[10] [10] Zhe Ming Lu and Yong Shi ―Fast Video Shot Boundary Detection Based on SVD and Pattern Matching‖Image processing IEEE Transactions (Volume:22 , Issue: 12), Dec. 2013.
[11] [11] Rainer Lienhart, "Comparison of Automatic Shot Boundary Detection Algorithms", [Online]. Available: http://www.vis.uky.edu/~cheung/courses/ee639_fall04/readings/spie99.pdf
[12] [12] Ravi Mishra ,S.K.Singhai,M. Sharma, "A Comparative based study of Different Video-Shot Boundary Detection algorithms", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) (Volume :2, Issue: 1), Jan. 2013.
[13] [13] Zuzana Cˇernekova´et.al., "Information Theory-Based Shot Cut/Fade Detection and Video Summarization", Circuits and Systems for Video Technology, IEEE Transactions on (Volume:16, Issue: 1),jan 2006, pp-82-91.
[14] [14] *Songhao Zhu and Yuncai Liu,* "A Novel Scheme for Video Scenes Segmentation and Semantic Representation", IEEE, 2008.
[15] [15] L.Chaisorn, T. S. Chua and C. H. Lee, "The segmentation of news video into story units", IEEE proceeding on International Conference on Multimedia and Expo, pp. 73–76, 2002.
[16] [16] *Yi-Chong Zeng,* "Automatic Extraction of Useful Scenario Information for Dramatic Videos", ICICS, 2013.
[17] [17] Hari R, Roopesh C. P. and Wilscy M., "Human Face Based Approach For Video Summarization", in IEEE Recent Advances in Intelligent Computational Systems (RAICS), 2013, pp. 245-250.
[18] [18] S. Foucher, and L. Gagnon, "Automatic detection and clustering of actor faces based spectral clustering techniques," *Canadian Conf. on Computer and Robot Vision*, pp.113-122, May 2007.
[19] [19] M. Everingham, J. Sivic, and A. Zisserman, "Hello! My name is …Buffy – automatic naming of characters in TB video," *in Proceedings of BMVC*, pp. 889–908, 2006.
[20] [20] L. D. Giannetti, "*Understanding Movies*," 10th edition, Prentice Hall, 2004.

Cite this article as: Deepika Bajaj, Shanu Sharma. "Video Depiction of KeyFrames- A Review." *International Conference on Information Engineering, Management and Security (2015)*: 224-230. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS039

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.039

Combining parameters for detection of ventricular fibrillation

Reji Thankachan[1], Aswathy R Krishnan[2]
[1]Faculty of Department of Electronics and Communication
[2]Student of M.Tech Signal Processing
[1, 2] College of Engineering Karunagapally, kollam, Keral, India

Abstract: *The irregularity in human heart beat indicate the disorders associated with it.Identifying the cause of arrhythmias is important in its treatment.The method of correcting shockable arrhythmias using Automatic External Defibrillators(AED) are of great demand in these days .The detection of arrhythmias and defibrillating the irregularity within minimum time fromthe ventricular fibrillation can considerably minimizes the death rate.Instances were reported in which the external defibrillator failed to act on time or interrupted the non-shockable rhythm as shockable. Research in this area is necessary to develop an effective algorithm that reduces this misrepresentation.The error rate can be minimised by combining various ECG parameters. It is preferable to use right combination of minimum number of parameters to minimise computation time and error rate. This paper uses ten ECG parameters in a combined manner for differentiating ventricular fibrillation and non ventricular fibrillation using Gaussian Support vector machine classifier.This project is simulated using MATLAB. On simulation the Balanced Error Rate is found to be around 7.5 percentage.*

Keywords- Defibrillation,Support vector machine, ECG parameters.

I. INTRODUCTION

Cardiac arrest is a major cause of higher death rate in present day.Initiation of timely treatment is important in life saving. Reason for cardiac arrest determine the type of treatment technique.Defibrillation is a technique for treatment of shockable arrhythmias. Ventricular fibrillation is a class of shockable rhythm which can be effectively treated using defibrillation . It is done by applying electric shock either invasively or noninvasively to heart. Non invasive defibrillators are considered as faster first aid treatment procedure to restore the normal heart rate. The vital component in defibrillator is detection algorithm. Diverse detection algorithms are there that takes into consideration of either time domain, frequency domain ,complexity , shape features of ECG signals. These parameters are used in classification of different scenarios of VF ,nonVF,ventricular tachycardia ,shockable and nonshockableECG signals.When individual parameter based algorithm is implemented in real time, its performance decrease from the investigated values. Combination of different ECG parameters in detection technique can improve the efficiency[1]. The next vital component in defibrillator is classifier. Classifiers like neural network,support vector machine, KNN classifier etc can be used in classification. The premier work in this field was first introduced by Nitish V Thakor, which uses width,height and area of beat[2] in analysing the arrhythmia.This techniques main drawback is the overload of data collection and time consuming analysis. Later Thakor introduced new temporal based algorithm called TCI[3] and sequential hypothesis was used in classification. This algorithm operates on binary

Cite this article as: Reji Thankachan, Aswathy R Krishnan. “Combining parameters for detection of ventricular fibrillation.” *International Conference on Information Engineering, Management and Security (2015)*: 231-235. Print.

signal and hence easier than previous work.A spectral domain algorithm was proposed by Kuo called VFleak[4] algorithm,which involve a narrow band rejection filtering in mean frequency of ECG.Another spectral algorithm(SPEC)[5] measures the amount of energy contained in different set of frequency ranges. Y. S. Zhu and Thakor introduced complexity based algorithm(CPLX)[6]that is based on repetitive patterns .Amann proposed two complexity based algorithms, first one uses Hilbert transform called Hilbert transform algorithm(HILB) [7]which has best sensitivity for a given specified specificity,second one is time-delay reconstruction (PSR) [8],at a given specificity it has greater IROC .Both this algorithms identify random behaviour.The main drawbacks of both is that its doesn't take in to consideration of shape of ECG signal which bear information.H.Li put forward Sample entropy algorithm[9],which uses entropy values.

This is useful in analysing short and noisy ECG.Modified and efficient temporal parameter than previously defined TCI named threshold crossing sample count were proposed by M.A Arafat (TCSC)[10]which involve comparison with threshold and counting its crossings. In real time,algorithms exhibits advantages and disadvantages. An efficient algorithm must work effectively in real time and can be easily implementable . The shock should not be given to a person with nonshockable arrhythmias.Faster, accurate and efficient algorithm is required in Automatic external defibrillators. Combining parameters with improved efficiency and less computation time is an effective approach for decreasing the error rate.This paper involve joining such features and improving performance.

II. DATABASE

The database of ECG signal are from the the CUDB[11] which are available at the PhysioNet repository. The files contain 8-minute long recordings of VF,normal sinus, flutter etc.The signals are sampled at 250 HZ.Using the annotation provided in recordings,10-second long each type of ECG signal are taken in .mat format. The description about the records including the sampling frequency,nature of irregularity ,age of patient are listed in database .The ECG signals are recorded in time/data format .

III. PREPROCESSING

The ECG is recorded using electrode placed on surface of human body. The signals are of millivolts range. The recording technique is highly sensitive to noise and other artifacts. The irregularity can be due to misplacement of electrodes,muscle movement, non cardiac noises, powerline interference of about 60 Hz from other electromagnetic devices. The noises and other artifacts are removed by four preprocessing methods[1].

1) Mean subtraction Subtracting mean value from the signal
2) Moving average filter A moving averaging Filter to eliminate 60 HZ powerline interference
3) High pass filter Baseline drift is minimised using high pass filter of cut off frequency 1 HZ
4) Low pass filtering To reduce muscle noise above 30 HZ second order butterworth low pass filter is used.

IV. ECG PARAMETERS

The preprocessed signal is analysed for duration of 8 -sec

1) Wavelet based algorithm:[12] To find the maximum of absolute values of hyperbola located in wavelet space,wavelet transform followed by fourier transform is defibrillation process.

2) Hurst index :[13] The detailed coefficients of wavelet decomposition of ECG signal up to 4 levels are calculated. From this the Hurst index is calculated by equation:

$$\log 2\ (\text{variance}(\ j\text{thdetailedcoe f f icients})) = (2H + 1)\ j \qquad (1)$$

3) Threshold crossing interval:[3] Average intervals between threshold crossings are investigated.

$$N = \frac{Number\ of\ samples\ that\ crossV_0}{Total\ number\ of\ samples} \times 100 \qquad (2)$$

$$N = \frac{1}{L_e - 2} \times \sum_{i=1}^{L_e-2} N_i \qquad (3)$$

4) Standard Exponential:the ratio[12] between decreasing exponential placed in maximum amplitude with ECG.Decreasing function is given by

$$E_s(t) = M \times exp(\frac{(-|t - t_m|)}{\tau} \qquad (4)$$

5) Threshold crossing sample count:[10] Number of samples that cross the threshold values.On an Le duration episode ,TCSC value is evaluated for average Le-2 values.

$$TCI = \frac{1000}{(N-1) + \frac{t_2}{t_1+t_2} + \frac{t_3}{t_3+t_4}} \qquad (5)$$

Cite this article as: Reji Thankachan, Aswathy R Krishnan. "Combining parameters for detection of ventricular fibrillation." *International Conference on Information Engineering, Management and Security (2015)*: 231-235. Print.

6) Modified Exponential:An improved form of STE, called MEA [12],Identifying crossing points by placing ECG at relative maximum. Exponential function is given by equation

$$E_i(t) = \begin{cases} A_i \times exp\frac{-|t-t_{m,i}|}{\tau} & t_{m,i} \leq t \leq t_{i,j} \\ \text{given ECG signal} & t_{c,i} \leq t \leq t_{m,i+1} \end{cases} \quad (6)$$

7) Mean absolute value:[?]averaging 2-sec Le □1 consecutive values.

$$MAV = \frac{1}{N} \times \sum_{n=0}^{N-1} |x(n)| \quad (7)$$

8) VF filter:[4]finds frequency with maximum amplitude,if it in frequency range of ventricular fibrillation then VF parameter is calculated. Period is given by the equation

$$T = \frac{1}{f_p} \times f_{sample} \quad (8)$$

where fp is the frequency correspond to maximum amplitude 3-second ECG and fsamp is the sample frequency (250Hz).Parameter is calculated using through the correlation technique and is given by equation

$$VF = \frac{\sum_{j=T/2}^{N-1} |x_j + x_{j-T/2}|}{\sum_{j=T/2}^{N-1} |x_j| + |x_{j-T/2}|} \quad (9)$$

where x j is the jth signal sample; T is the period calculated in equation [8] and N is number of samples in 2-sec ECG.

9) Median frequency(FM):[14]Duration of cardiac is calculated using this parameter. This will determine if the defibrillation process will be successful.

$$FM = \frac{\sum_{i=1}^{n}(f_i.p_i)}{\sum_{i=1}^{n} p_i} \quad (10)$$

10) Spectral Algorithm:[12] using Fourier analysis the energy contained in different frequency bands are estimated.

V. SVM CLASSIFIER

SVM is an efficient classification technique which creates a optimal separating hyperplane with maximum margin.The ECG parameters calculated are nonlinear in nature. So Gaussian SVM is used for classification. All the features estimated are used for training the Support Vector Machine classifier. In nonlinear data SVM map the two classes using kernel method[15],here gaussian kernal is used and form discriminant function that minimizes the cost function

$$\min_{V,a_i,\xi_i} \frac{1}{2} \times \| X \|^2 + R\sum_{i=1}^{M} \xi_i \quad (11)$$

$$\text{subject to} \quad y_i(\langle \phi(v_i), x \rangle) + a) - 1 + \xi_i \geq 0 \quad (12)$$

$$\xi_i \geq 0, i = 1...M \quad (13)$$

where f(vi) is maps the nonlinear data to a higher dimensional space , xi represent the slack variable and R determine the shape of discrimination function and is the regularization parameter.As the parameters are nonlinear Gaussian type SVM is used for classification.

VI. PERFORMANCE EVALUATION

The performance of the detection algorithm is evaluated using Sensitivity(SE):the proportion of correctly identified VF,Specificity(SP):[1] the proportion of correctly identified nonVF.

$$SE = \frac{TruePositive}{TruePositive + FalseNegative} \quad (14)$$

$$SP = \frac{TrueNegative}{TrueNegative + FalsePositive} \quad (15)$$

Cite this article as: Reji Thankachan, Aswathy R Krishnan. "Combining parameters for detection of ventricular fibrillation." *International Conference on Information Engineering, Management and Security (2015)*: 231-235. Print.

True positive(TP) indicate the number of signals correctly classified as VF,True negative(TN) is number of signals correctly detected as NVF,False positive(FP) is number of NVF classified as VF and False negative (FN) is number of VF classified as NVF. The efficiency of classifier is estimated by Balanced Error Rate(BER) ,positive predictivity (PP)and accuracy(AC).

$$BER = \frac{1}{2} \times (\frac{FN}{PC} + \frac{FP}{NC}) \qquad (16)$$

$$PP = \frac{TP}{TP+FP} \qquad (17)$$

$$ACC = \frac{TN+TP}{PC+NC} \qquad (18)$$

where where PC = TP + FN and NC = TN + FP.As indicated in previous paper [1] as the number of parameters increases the BER decreases.The 10 parameters are calculated for VF and non VF ECG signals The SVM is trained using these parameters.

TABLE I
TESTED SAMPLE OUTPUTS

Tested sample	Number
NonVF	33
VF	33
TP	30
FP	2
TN	31
FN	3

On testing samples and evaluating the performance is given by

TABLE II
PERFORMANCE EVALUATION

Parameter	Value (in percentage)
SE	90.9
SP	94
ACC	92.42
PP	93.75
BER	7.5

On comparing with previous studies stated in [1] which uses 13 set of ECG parameters,the BER is always above 8. It also stated that the BER can be decreased by increasing the number of ECG parameters.This paper aims at combining 10 parameters and there by reducing BER. On evaluation, the BER is found to be around 7.5 which is less compared to the BER in[1] and by using lesser number of parameters. The accuracy has improved even without the use of feature selection in classifier. Meanwhile sensitivity and Specificity decreased than in [1].

VII. CONCLUSION

The detection algorithm using combination of 10 parameters leads to better BER and accuracy than previous studies which uses 13 parameters.The Specificity and sensitivity parameter needs to be further improved for real time application . This can be done by using complexity parameter along with temporal, spectral and wavelet based parameters.Avoiding those parameters that have less detection performance and combining those of high performance can improve performance and computation time.

ACKNOWLEDGEMENT

We would like to thank all the staffs of Department of Electronics and Communication , College of Engineering Karunagapally.

REFERENCES

[1] F.A.Atienza,E.Morgado,L.F.Martinez,A.G.Alberola and J.L. Alvarez,"Detection of Life-Threatening Arrhythmias Using Feature Selection and Support Vector Machines"IEEE Tran. Biomed.Eng vol. 61, no. 3, March 2014

[2] Nithish .V.Takor,"From Holter Monitors to Automatic Defibrillators: Developments in Ambulatory Arrhythmia Monitoring" ,IEEE Trans. Biomed.Engg vol. bme-31, no. 12, December- 1984

Cite this article as: Reji Thankachan, Aswathy R Krishnan. "Combining parameters for detection of ventricular fibrillation." *International Conference on Information Engineering, Management and Security (2015)*: 231-235. Print.

[3] N. Thakor,K.P.Pan and Y. S. Zhu,, "Ventricular tachycardia and fibrillation detection by a sequential hypothesis testing algorithm," IEEE Trans. Biomed. Engg.,vol. 37, no. 9, pp. 837-43, Sep. 1990.
[4] JCTB Moraes, M. Blechner, E.V. Costa and , F.N Vilani,,"Ventricular Fibrillation Detection using a Leakage/Complexity Measure Method "IEEE Computers in Cardiology2002;29:213-216.
[5] S. Barro, J. Mira, D. Cabello, and R. Ruiz, Algorithmic sequential decision making in the frequency domain for life threatening centricular arrhythmias and aimitative artifacts: A diagnostic system, J.Bio.Engg vol. 11, no. 4, pp. 320328, Jul. 1989.
[6] X. S. Zhang,Wang, N. Thakor, and Y. S. Zhu, , "Detecting ventricular tachycardia and fibrillation by complexity measure," IEEE Trans. Biomed. Eng.,vol. 46, no. 5, pp. 548-55, May 1999
[7] A Amann, Tratnig and K Unterkofler,"A New Ventricular Fibrillation Detection Algorithm for Automated External Defibrillators"Computers in Cardiology 2005;32:559562.
[8] A. Amann, R. Tratnig, and K. Unterkofler,"Detecting ventricular fibrillation by time-delay methods,"IEEE Trans. Biomed. Eng.,,vol. 54, no. 1, pp. 174 -177, Jan. 2007.
[9] H. Li, C. Hu, and W. Han,H. Meng, Detecting ventricular fibrillation by fast algorithm of dynamic sample entropy, IEEE Int. Conf. , Dec. 2009, pp. 11051110.
[10] M. Arafat, A. Chowdhury, and M. Hasan, "A simple time domain algorithm for the detection of ventricular fibrillation in electrocardiogram," Sig., Img. Vid. Proc.,vol. 5, pp. 1-10, 2011.
[11] http://www.physionet.org
[12] Unterkofler, Amann, and Tratnig,,"Reliability of old and new ventricular fibrillation detection algorithms for automated external defibrillators", Biomed. Eng. Online,, vol. 4, no. 60, Oct. 2005.
[13] Yan Sun, Kap Luk Chan and Shankar Muthu Krishnan."Life-threatening ventricular arrhythmia recognition by nonlinear descriptor"BioMed.Engi.OnLine 2005.
[14] C. Brown,H.Werman and Dzwonczyk , "The median frequency of the ECG during ventricular fibrillation: Its use in an algorithm for estimating the duration of cardiac arrest,"IEEE Tran. Biomed. Eng.,vol. 37, no. 6, pp. 640-646, June. 1990.
[15] F. Alonso-Atienza, E. Morgado, L. Fernandez-Martinez, A. Garcia- Alberola, and J. Rojo-Alvarez, "Combination of ECG parameters with support vector machines for the detection of life-threatening arrhythmias," in Proc. Comput. Cardiol.,pp. 385-388.

Cite this article as: Reji Thankachan, Aswathy R Krishnan. "Combining parameters for detection of ventricular fibrillation." *International Conference on Information Engineering, Management and Security (2015)*: 231-235. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS040		**eAID**	ICIEMS.2015.040

Embedded Based On Medical Technology

V.Bhuvaneswar[1], K.Aarthi[2], K.Eswar[3]
[1, 2, 3] VIVEKANANDHA COLLEGE OF TECHNOLOGY
FOR WOMEN

Abstract: *Every surgical item used during surgery (e.g., sponges) must be accounted for after surgery to ensure that none of these items is left inside the patient. Despite the numerous precautions in place, in approximately 1 in 1500 cases, something gets left behind inside the patient's body This paper presents ASSIST, an automated system for surgical instrument and sponge tracking that increases the safety of surgical procedures. ASSIST utilizes RFID technology to aid in accounting for all items used during surgery. The design takes into account safety, simplicity, ease of deployment, and ease of use. An initial evaluation utilizing RFID-tagged sponges demonstrates that ASSIST can reliably track surgical sponges with minimal impact to current operating room procedures. Sources of error that can impact the reliability of the system are also discussed.*

Key words: *ASSIST*-Automated System for Surgical Instrument and Sponge Tracking, *RFID*-Radio Frequency Identification

INTRODUCTION

One dangerous medical error that can occur during surgery is unintentionally leaving a surgical instrument or sponge inside a patient. Commonly known as retained foreign object this error can lead to in ammation, obstruction, perforation,sepsis, and sometimes death. The problem is thought to be avoidable when stringent manual counting guidelines are followed by Operating Room (OR) personnel. While these guidelines are very effective in reducing the risk, the problem persists .Some estimates report that the incidence can be as high as 1 in 1500 surgeries.Human error is not the only drawback of manual counting During sponge counting, nurses are unable to provide support for the surgeon as they are focused on accurately counting sponges. Each sponge count takes a couple of minutes, with at least three counts per surgical procedure. Under these counting procedures, the nurse is inevitably distracted from her primary role for a significant part of the time. Also when a miscount is found, there is a significant increase in the OR time since an x-ray of the patient is many times required. Some hospitals x-ray every patient after any open cavity operation, which requires a radiologist to be available after every surgery and unnecessarily exposes the majority of patients to radiation . This paper presents ASSIST, an automated system for surgical instrument and sponge tracking that increases the safety of surgical procedures. With ASSIST, RFID (Radio Frequency Identification) technology is used to detect and uniquely identify each surgical item at various stages during surgery. The use of low frequency RFID enables reliable detection of tags even when soaked in body uids, in the vicinity of metallic objects such as surgical tools, or inside a patient's body. A software-based inventory component keeps track of every item, and enables users to quickly identify the state of the procedure through a color-coded interface While ASSIST can track any item tagged with an RFID device, our initial design focuses on retained sponges (a.k.a. gossypiboma) as they constitute the majority of retained foreign body cases. A check-in station verifies the content of a package and registers each tagged item in a database to keep an inventory.

A check-out station, consisting of a smart bucket where sponges are discarded, accounts for used sponges. The number of items in-use is

Cite this article as: V.Bhuvaneswari, K.Aarthi, K.Eswari. "Embedded Based On Medical Technology." *International Conference on Information Engineering, Management and Security (2015)*:236-243. Print.

displayed to the user at all times. If zero, all items are accounted for . Otherwise a patient scanner is available to detect whether a missing sponge is still inside the patient ASSIST has several benefits over current practice.

First, it eliminates false-positives, i.e., counts that appear to be complete when a sponge is still missing, which are estimated to account for more than 80% of cases of retained foreign objects. Second, ASSIST eliminates unnecessary exposure to x-rays as only those patients that appear to have unaccounted items need to be x-rayed. Third, it reduces the operational cost to the hospital by reducing nurse and OR time required for each intervention. Finally, real time information on the state of retained foreign objects allows the surgeon to close an incision without delay Our initial investigation shows that our system can reliably account for 100% of tagged sponges during surgery. This high level of reliability is attained by RFID verification during check-in, and continuous counting with multiple orthogonal antennas at the check-out bucket. The read rate is maximized by utilizing random antenna selection between consecutive reads. The measured check-in time for a 10- sponge packet is just 2 seconds, while regular check-out time is between 1 to 5 seconds when several sponges are thrown into the bucket at the same time. Experiments also show that we can detect missing sponges inside an in vivo porcine model, with an RFID patient scanner, in less than 5 seconds .

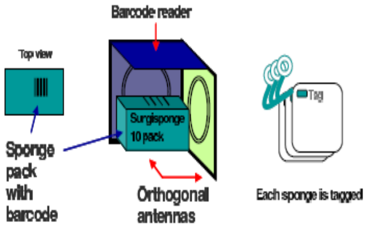


Fig. 1: Check-In Station

BACK GROUND

In healthcare, RFID has two major areas of application administrative and direct-patient. Administrative applications include supply chain, smart shelving, and equipment and/or pharmaceuticals tracking. RFID technology is becoming well established in these areas due to efficiencies established in commercial and industrial sectors. Direct-patient applications, on the other hand, are still in their infancy largely because they can have a direct impact on the patient's health. . One example is VeriChip, approved by the FDA in2004, and offers implantable RFID chips containing personal medical information to help limit medical treatment errors ASSIST will also require FDA approval as it directly affects patients' health and safety Two general variants of RFID technology exist; active tags, which require an internal power source, and passive tags, which rely on the incoming radio frequency signal to power up and respond to commands. Both tag systems operate in different frequencies; low, high, ultra high, and microwave. Low frequency (LF) tags work in the 125-148 KHz range and utilize magnetic fields for communication. They range and utilize magnetic fields for communication. They metals, and most electromagnetic (EM) noise sources. High frequency (HF) tags work at 13.56 MHz, can read some distance through uids, and are susceptible to noise and antenna detuning in the presence of metals. Higher frequency
RFID devices utilize electric fields for communication which are attenuated by uids and are therefore not suited for tracking uid soaked sponges. Finally, there are two families of anticollision protocols: binary tree and Aloha. The binary tree provides a deterministic approach to read every available tag, while Aloha derived protocols are purely stochastic and rely on probability to ensure that every tag is read. RFID is well suited for automating sponge counting in the OR. It allows us to uniquely identify each tag and reliably account for each one during surgery. We chose passive tags as they do not require an internal energy source, last much longer than active tags, can be made in much smaller sizes and can be manufactured at a significantly lower cost. Low frequency tags enable us to search for items within a patient with high reliability as organs, bones, and body uids are transparent to magnetic fields. In addition, we choose tags that utilize a binary tree anticollision protocol as we need to account for every item before and after surgery with the highest certainty.

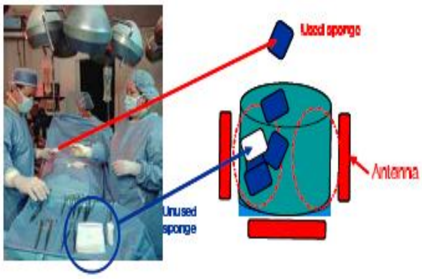


Fig. 2: Check-Out Station

Cite this article as: V.Bhuvaneswari, K.Aarthi, K.Eswari. "Embedded Based On Medical Technology." *International Conference on Information Engineering, Management and Security (2015)*:236-243. Print.

APPROCH

Our approach is to develop an electronic inventory that can keep track of every surgical item used during surgery RFID is used for non-line-of-sight identification (a unique serial number for each sponge can be received by wireless means). A check-in station, a check-out station, and a patient scanner are used by OR personnel to track and/or find sponges throughout the surgery. All of these components are controlled via a software system that utilizes a color-coded interface

A.CHECK-IN STATION

The check-in station consists of two orthogonal RFID antennas surmounted on top by a perpendicular UPC barcode reader. A design is depicted in Figure 1. The check-in station is small, simple, and occupies very little space. The UPC barcode identifies a package and the number of items within it. Before accepting the package, ASSIST verifies number of items in it, which is gathered from barcode information. If the package contains a bad RFID tag, the reader will recognize the discrepancy and subsequently direct the removal of the package. In addition the barcode provides us with a description of the type of sponge associated with each RFID tag. Note that this barcode can be easily replaced by an RFID-based Electronic Product Code (EPC) .

It is also possible to store the barcode information in the RFID tags already available inside the sponges. While this is possible, it would increase the price of the tags as they would require read/write memory to hold this information instead of just a unique read-only ID. In addition, sponge manufacturers would have to ensure that the RFID tags that they attach to the sponges are programmed with the correct information. Further, the same sponge may be packaged in different quantities, which requires different codes to be inserted in tags that are destined for different packages. As such, it is simpler to attach that information to the packaging means of a barcode or EPC.

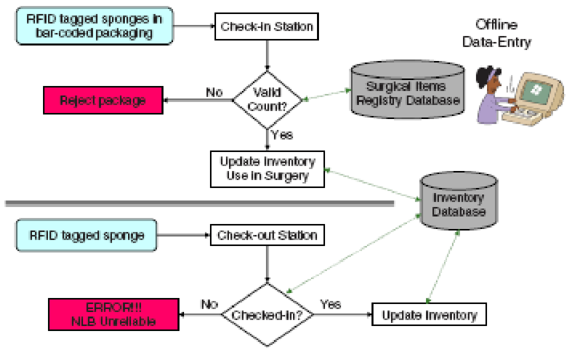


Fig. 3: Control Flow Overview

When a package is accepted, every item is registered into the system inventory where it can be tracked throughout the surgery. As previously stated, each sponge is distinguishable by its RFID tag which has a unique serial number.

B. CHECK-OUT STATION

A stainless-steel bucket, commonly known as a kick bucket, is the normal depository for used sponges in the OR. It is small, simple, and convenient. Our approach is to provide a new kickbucket with a similar form factor, enabling the OR to replace existing ones with minimum disruption.The ASSIST kickbucket is equipped with an RFID reader, five orthogonal antennas (four sides plus the bottom), and the means to talk to the software component that keeps track of the inventory (i.e. wireless serial port). Since every RFID tag is equipped with a unique serial number, no sponge is counted more than once. A design is depicted in Figure 2.

The ASSIST kickbucket continuously scans for sponges and updates the inventory accordingly. While 5 antennas may seem excessive, our initial results suggest that we require all of them to ensure that all the sponges are read reliably without any significant involvement from OR personnel. Further design refinement may result in a more economical design, but such considerations are left for a subsequent design phase.Although the ASSIST kickbucket is intended for use throughout the surgery, the OR staff can also use another, non-RFID equipped kickbucket to dispose sponges, and transfer them all (possibly in a bag) into the ASSIST kickbucket at the end of the surgical procedure.

C. PATIENT SCANNER

When a sponge is considered "missing", current practice OR procedure dictates that an x-ray of the patient be taken. This requires an x-ray machine to be brought into the OR, or that the patient be moved to an x-ray equipped location.

Cite this article as: V.Bhuvaneswari, K.Aarthi, K.Eswari. "Embedded Based On Medical Technology." *International Conference on Information Engineering, Management and Security (2015)*:236-243. Print.

Fig. 4: GUI

A radiologist needs to be called to read the x-ray. The x-ray may not reveal the missing sponge, especially in cases where the sponge is likely to be surrounded by bones (e.g., heart surgery). A patient scanner can be used to detect, and somewhat localize, missing sponges in the patient. We use a blanket embedded with a large antenna that allows us to reach deeper into the patient, increasing the chance of finding any missing sponge. Smaller antennas may also be embedded in the blanket to localize the missing sponge. Any sponge found by the patient scanner is identified as ”In-Patient” in the software inventory. While very convenient, some medical devices may react inappropriately to the presence of an RFID patient scanner. For example, a recent study by the FDA shows that RFID can interfere with pacemakers and Implantable Cardiac Defibrillators. These patients can still benefit from ASSIST accurate inventory system, but must resolve to other means of finding missing sponges when necessary (e.g., x-ray).

SOFTWARE SYSTEM

It is routine for most OR facilities to utilize computer monitors during surgical procedures. Our approach is to incorporate our inventory software-based approach with the current OR computer infrastructure. In general, the software system should be multi-layered, extensible, and fault tolerant enforcing when possible the policies in place to ensure compliance with the new procedure ASSIST uses two databases; the surgical items registry,and the inventory. The registry contains barcode information for each package (i.e. description, quantity). The inventory keeps track of all items that have been checked-in as they pass through the system and makes sure that every item is accounted for after the surgery is complete. Any missing item will be re ected in the inventory, at which time the patient must be scanned by the patient\ scanner and/or x-ray.An overview of the different databases and the ow of the procedure is shown in Figure 3. The Graphical User Interface ties together all the different components of the system. A snapshot of the ASSIST GUI

Fig. 5: Equipment

Antennas	One Cycle	Five Cycles
1	63.9%	71.7%
2	63.9%	71.7%
3	84.4%	94.4%
4	91.4%	98.6%
5	97.2%	100.0%

TABLE I:

Antennas	One Cycle	Five Cycles
1	63.9%	71.7%
2	74.7%	91.7%
3	84.4%	94.4%
4	91.4%	98.6%
5	97.2%	100.0%

Number of antennas required to reliably find 40 sponges in check-out bucket.

can be seen in Figure 4. It contains three-panels, each with color-coded identifiers which allow OR staff to quickly and easily identify any problem and assess the state of the procedure.

EXPERIMENTAL RESULTS

Our prototype system comprises of a Frosch Electronics Low Frequency RFID reader with a built-in multiplexer, seven 18x18 cm PCB antennas,a 30x40 cm at cable ribbon antenna for the patient scanner, a Symbol Omni- Directional barcode scanner, and approximately 100 passive low-frequency hitag 1/S RFID tags. Figure 5 shows some of the parts used in our experiments. All RFID tags where glued inside Kendall 4x4 raytec sponges (one in each) Two PCB antennas were arranged orthogonal to one another in the check-in station, with the barcode scanner on top. The other five PCB antennas were arranged around the sides of the ASSIST kick-bucket, while leaving the top open. Tag orientations inside the kick-bucket are random. Mutual inductance between tags that are in close proximity can change their frequency response and make them unre sponsive to the tag reader. The packaged sponges can be arranged in such a way that they do not interfere with one another but there is no such control over the position of the used sponges in the bucket. Our initial experiments examine the reliability and feasibility of the system using dry sponges with RFID tags. First, we identified the number of antennas required to reliably detect sponges in the check-out bucket. We deposited 40 tagged sponges in the bucket, and noted the number of sponges found when using different numbers of antennas and

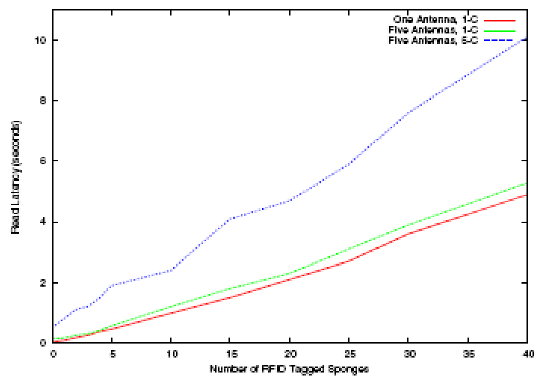


Fig. 6: Time to read sponges with one antenna, 5 antennas, and 5 cycles with 5 antennas

performing one or multiple read cycles. A read cycle consists of one read from each of the antennas, so five cycles with one antenna is identical to reading five consecutive times from the same antenna. The experiment was repeated 20 times, revolving the sponges each time so that they ended up in different positions. Table I shows that we required all five antennas, and multiple read cycles, to reliably detect all sponges in the bucket.

Reading from five antennas multiple times will introduce some unavoidable latency. Figure 6 shows how the read latency increases as more sponges are deposited in the bucket. Reading from multiple antennas resulted in less than 10% (1/2 second) additional latency. The overhead is low because the anti-collision procedure turns off tags as it reads them, and so consecutive reads from other antennas will find a smaller set of tags to read from. Performing 5. cycles introduces another few seconds of latency. As we are using a multiplexer, only one antenna at a time can send the refresh signal to make sure that the tags stay in silent mode. Tags that are not aligned with that antenna do not hear this signal and so they turn on again, and are read again by another antenna in consecutive cycles.

The system was then tested on a more realistic setting at the Johns Hopkins Minimally Invasive Surgery Training Center (MISTIC) center. ASSIST was used in five in-vivo pig surgeries, on three different sessions. Each operation consisted of an open abdominal and/or thoracic cavity surgery where organs were explored, and sometimes extracted, as they would be in a regular surgery. RFID tagged sponges where used in each intervention. Sponges were deliberately left inside the ”patient” after surgery, and the system correctly determined the number of sponges missing each time.

Using one antenna could not reliably read the complete pool of sponges as the RFID tags lied in random orientations with only a subset appropriately aligned to the field lines of a single antenna at any one time. We had to redistribute the sponges several times in order to read all tags in one cycle

Test Detail Result

Time to check-in 10 sponge pack 2 seconds
Time to find 7 sponges thrown Empty Bucket 2 seconds
into bucket as bucket gets 25 sponges 3 seconds
filled (online reading) 45 sponges 6 seconds
60 sponges 10 seconds
Time to find all sponges 25 sponges 6 seconds
simultaneously in the bucket 45 sponges 12 seconds
(of ine reading) 60 sponges 31 seconds

TABLE II:

Test	Detail	Result
Time to check-in	10 sponge pack	2 seconds
Time to find 7 sponges thrown into bucket as bucket gets filled (online reading)	Empty Bucket	2 seconds
	25 sponges	3 seconds
	45 sponges	6 seconds
	60 sponges	10 seconds
Time to find all sponges simultaneously in the bucket (offline reading)	25 sponges	6 seconds
	45 sponges	12 seconds
	60 sponges	31 seconds

Time that ASSIST takes to detect and account for sponges in different scenarios ,

We then performed a series of ”hide-and-seek” tests where one person hid sponges in different locations within the animal, and another looked for the missing sponges with the patient scanner. Sponges where hidden around the stomach, in the pelvic hole, and behind the sternum and ribs, including areas around the heart. On each occasion, all of the missing sponges were successfully found in less than 5 seconds of scanning. Reverse experiments were also performed, where multiple (>5) sponges were left in the patient around the abdomen and the chest, and the patient scanner was used to find all sponges before a count was performed in the bucket, i.e., the person scanning did not know how many sponges were missing.

Each time, the final count of the bag revealed all sponges were successfully found without having to re scan the patient .Once the animals were euthanized, we collected enough blood to test the reliability of the system after the sponges were immersed in blood. This way, we could perform a realistic evaluation of the reliability of the system under realistic surgical conditions (bearing in mind that wireless signals can be affected by uids and metals). A different number of blood-soaked sponges were used on three different occasions: 25, 45, and 60. The sponges were redistributed randomly on each trial We found that, with one cycle through all five antennas, the detection rate was about 95%. With multiple (5) repeated cycles the detection rate was between 98% and 100%. While dry sponges did achieve 100% detection rate, the blood soaked ones were not always all found. One solution was to agitate the sponges (i.e. shake the bag). We found that a small movement of the bag quickly resulted in 100% detection rate However, we also found that randomly selecting the antenna between consecutive reads produced 100% read rate, without having to move the sponges. This finding basically reacts the benefit of anticollision from adjacent reads.

Table II shows a summary of what a user would experience when using our system. The average time to check-in a sponge packet is just under 2 seconds. The average time to check-out sponges varies depending on the number of sponges in the bucket, with a steep increase in latency when 60 sponges were inside the bucket.Finally, we performed experiments where metal instru ments were ”inadvertently” dropped in the bucket (a few surgical scissors). We had to move the sponges on 2 occa sions, out of 15 tries, to be able to detect all sponges in the bucket. However, while it did take some additional effort to find them, all of the sponges were reliably, and accurately,accounted for by the ASSIST system

SOURCES OF ERROR

ASSIST provides an alternate mechanism to keep track of surgical sponges during surgery. Several key advantages exist over current practice, including no false-positives, the ability to eliminate false-negatives, time savings, added confidence, and elimination of unnecessary x-ray exposure. However, as with other systems[10], ASSIST is susceptible to a variety of issues. While the likelihood of some of these conditions occurring is very small, it is important to be aware and understand what can go wrong

- Protocol violation: An item that was never checked in may be used during surgery, at which point the final count cannot be trusted. We can distinguish any unchecked item at check-out and report the problem The nurse must identify the number of sponges that are likely to have been missed during check-in (i.e. typical quantity in package) and make sure the Never Checked-In count in the ASSIST software re ects this number. The patient should also be scanned with a patient scanner and/or x-ray.

- Damaged RFID Tag: Any damaged RFID tag will not be able to be checked-in by the system. The barcode scanner prevents this type of error from becoming a problem as it will make sure that all of the tags are readable before they are used in surgery. However, there is a finite probability that a tag that is working when checked-in stops working during the procedure. This will result in a false negative. In this case, a manual count can be performed to verify the number of checked-out sponges

- Unreliable Manufacturing Process: While ASSIST can detect and/or race a ag for single-errors like a damaged RFID tag, double-errors can occur in an unreliable manufacturing process. For example, if a 10-pack of sponges contains 11 sponges, and one of these sponges is either untagged or with a damaged tag, ASSIST will accept the package, but only 10 sponges will be Entered into the inventory.While extremely unlikely (as it requires that there is a packaging error and that one of the RFIDs in the package is damaged or missing), we should be aware of the possibility as this kind of error can ultimately result in an unreliable system. One possible solution is to have the OR nurse manually count sponges, during check-in, to verify the expected quantity in the box.

REATED WORK

A recent case-control study showed that the risk of retained foreign objects increases during emergencies, un planned changes in procedure, and in patients with high body-mass index. However, more astonishingly, they show that 88% of analyzed cases involved a final count that was erroneously thought to be correct (false positive). Other case control studies show a similar trend.A variety of approaches have been proposed to mitigate the problem of retained foreign objects following surgery. sponges[12], to reduce human-error during manual counting.

This method is simple and does not significantly change the current procedure. Another approach, patented by Sur gicount Medical, uses a printed barcode on each sponge with a unique identification number. This requires nurses tomanually check-in and check-out every sponge with a line of-sight barcode reader, a process that must be made more difficult by the presence of blood .

Others have looked at the problem of finding sponges inside a patient. In one approach, an electronic tag is inserted in every sponge.As in our approach, they choose a LF (in their case 50 to 70 kHz) tags. Six sponges were hidden in the abdomen and the thorax of 50 cadavers weighting between 45 and 190 kg, and 100% of the tags where found. Another low frequency solution, at 145 KHz, is commercially marketed by RF Surgical Systems. Note that these systems lack unique identification means as in RFID therefore, they cannot provide an accurate inventory of the surgical instruments before and after the procedure Another study evaluated the use of a wand to find high frequency (HF, 13.56MHz) RFID tagged sponges in 8 patients. A sponge was hidden in the abdomen of each pa tient. The RFID wand reliably detected each of the sponges, within five seconds. However, more studies are needed with HF RFID to validate its use for detecting sponges inside patients. High frequency signals are attenuated by body uids and can be obstructed by bone, so the depth of field is likely to be insufficient for patients with large body mass.An alternate system has been proposed by Clearcount

Medical Solutions .In their approach, HF RFID tagged sponges are tracked using a stand-alone device that has a scanner on one side and a waste disposal bucket on the other The sponges are detected as they pass into the bucket rather than being detected inside the bucket, which has potential for errors when many sponges are thrown in together. The high frequency tags can be read at a faster rate than low frequency ones which helps to ensure that all of the sponges are detected as they enter the bucket but the trade-off is poorer penetration of body uids leading to difficulty detecting the sponges inside the patient and also when wadded together As this is a stand-alone system it does not integrate into the OR software but simply offers the user information on how many sponges have been checked in and checked out. There is no published data available on this system, so it is not possible to compare it's reliability with ASSIST

CONCLUSION

ASSIST provides the OR staff with a reliable and accurate record of the status of each surgical item throughout the surgical procedure. Human error is minimized as the system eliminates false-positive (mistakenly complete count) and reduces false-negatives (the mistaken belief that a sponge is missing). Also, we cut down the time and effort that is otherwise needed for manual counting during for each intervention. Sources of error were identified, and the system was evaluated under realistic surgical conditions.

REFERENCES

[1] A. Macario, D. Morris, and S. Morris, "Initial clinical evaluation of a handheld device for detecting retained surgical gauze sponges using radiofrequency identification technology. *Arch Surg*, 2006

[2] C. Fabian, "Electronic tagging of surgical sponges to prevent their accidental retention." *Surgery* , vol. 137, no. 3, pp. 298–301, 2005.

[3] K. Finkenzeller, RFID handbook: fundamentals and applications in contactless smart cards and identification, 2nd ed., 2003.

[4] A. Rogers, E. Jones, and D. Oleynikov, "Radio frequency identification (rfid) applied to surgical sponges."*Surg Endosc*, 2007.

[5] S.Sarma, D. Brock, and D. Engels, "Radio frequency identification and the electronic product code," *IEEE Micro* , vol. 21, no. 6, pp. 50–54, 2001.

[6] "ClearCount Medical Solutions." Online available[http://www.clearcount.com]

Cite this article as: V.Bhuvaneswari, K.Aarthi, K.Eswari. "Embedded Based On Medical Technology." *International Conference on Information Engineering, Management and Security (2015)*:236-243. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		VOL	01
Website	www.iciems.in		eMail	iciems@asdf.res.in
Received	10 - July - 2015		Accepted	31- July - 2015
Article ID	ICIEMS041		eAID	ICIEMS.2015.041

Review on Demosaicking via Directional Linear Minimum Mean Square Error Estimation

Nidhi Chauhan[1], Shanu Sharma[2]
[1, 2] Computer Science Department, Amity University, Noida

Abstract: *The digital cameras test scenes utilize a color channel cluster of mosaic example (e.g., the Bayer example). The demosaicking of the color samples in [4] is basic to the picture essence. It displays another shading demosaicking procedure of ideal directional sifting of the green- red and green-blue contrast signals. Expecting it that the essential distinction signals (PDS) between the green and the red/blue channels are low pass, the missing green samples are robustly decided in both even and additionally vertical bearing by the direct least mean square-mistake estimation (LMMSE) procedure. These directional evaluations will be then ideally intertwined to further enhance the green gauges. In conclusion, guided by the demosaicked full-determination green channel, the other two shading channels are recreated from the LMMSE separated and intertwined PDS.*

Keywords: Color Demosaicking, Bayer color filter array (CFA), directional filtering, linear minimum mean square-error estimation (LMMSE).

I. INTRODUCTION

Demosaicking is a technique used to find the value of each pixel of interpolated image. Today we found that the digital cameras resolutions are more and more but the logic is same. Color filter array (CFA) is used in the digital cameras sample scenes (eg: Bayer Pattern). Image quality is more concern with the demoaicking of color samples. New technique of color demosaicking i.e. Directional Linear Minimum Mean Square Error Estimation is used to determine the color samples with the difference of green - red and green - blue samples. Directional Linear Minimum Mean Square Error Estimation (DLMMSE) technique is used to evaluate primary difference signals (PDS) between green and red/blue channels in horizontal as well as vertical directions. These estimations are further used to enhance the value of the green channel by fusing together. At the end both the red and the blue are calculated by demosaicked full resolution green channel estimated above. By performing this method we found the PSNR and the visual perception to be more accurate [10]. If we consider there is a square grid in which red, blue and green color are arranged in such a manner that each square of four pixels has on red , one blue and two green filter as shown in Figure 1. It is because our eye is more perceptive to green color as compared to red or blue. There are many different techniques to measure Peak Signal to Noise Ratio (PSNR). Few of them are Pixel Doubling Interpolation, Bilinear Interpolation, Gradient Based Interpolation and High Quality Linear Interpolation. J. E. Adams et al [3] describes basic methods which can be used in for demosaicking explained with advantages and disadvantages of individuals. The four methods explained in these papers are pixel doubling interpolation, bilinear interpolation, gradient based interpolation and high

Cite this article as: Nidhi Chauhan, Shanu Sharma. "Review on Demosaicking via Directional Linear Minimum Mean Square Error Estimation." *International Conference on Information Engineering, Management and Security (2015)*: 244-250. Print.

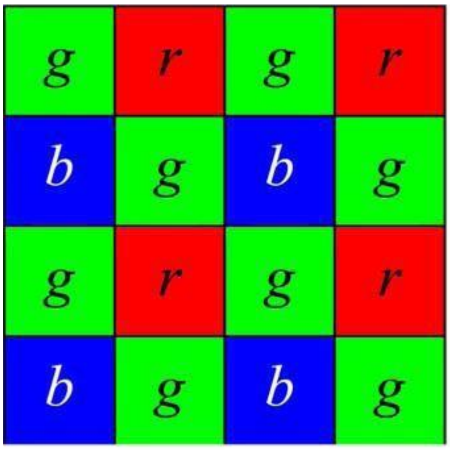


Figure 1: Bayer Pattern

quality linear interpolation. All these techniques use a simple Bayer pattern color filter array.B. K. Gunturk, Y. Altunbasak and R. M. Mersereau et al [7] explains the advanced cameras utilization color channel clusters to test red, green, and blue hues as per a particular sample. At the place of every pixel one and only coloring specimen is taken, and the estimations of the other hues must be interjected utilizing neighboring tests. This shading plane interjection is known as demosaicking; it is one of the critical undertakings in a computerized camera pipeline Many digital cameras uses single sensor array which is used to capture an image. On every pixel, stand out of the three essential hues (red, green and blue) is examined. Here we are utilizing Bayer color filter array (CFA). To reproduce a full color picture the missing shading examples need to be interjected by a procedure called color demosaicking. The nature of these reproduced color pictures relies on the picture substance and the demosaicking strategy we are utilizing.

Previously we have used the nearest neighbor techniques, bilinear interpose and couple of other different methods to demosaicked the image. Those methods were easily used and therefore suffer from some problems like blurring, blocking and zipper effect at the edges. The smooth hue transition (SHT) methods interpose the luminance channel i.e. green and chrominance channels i.e. red and blue. After we get the green channel by bilinear interpolation, we get the red and blue channels by bi-linearly interposing the red hue (the ratio of red to green) and blue hue (the ratio of blue to green). Although the SHT methods have the interrelation between the red, the blue and the green channels, they will result in large interpolation errors in red and blue channels when green values unexpectedly modified.

Many demosaicking methods avoid interpolating at edges, as human visual systems are very perceptive to the edge organization in an image. At each pixel the gradient is calculated, and the color interpolation is done by using directionally based on the determined gradient. Directional filtering is very prominent way for color demosaicking that gives competing results. The best known directional interpolation scheme is the second order Laplacian filter [2]. When the green examples are filled in the picture, then after the red and the blue specimens are interjected comparably with the rectification of the second request slopes of the green channel. One more slope based demosaicking system displayed by Chang et al [7] is utilized as a situated of slopes in diverse bearings in the 5x5 area focused at the pixel to be interjected. A subset of the arrangement of these slopes is chosen by versatile limit. Another type of color demosaicking techniques is iterative method, which we can also be used with the gradient-based methods. Iterative demosaicking process consist of two-step i.e. of an enhancement step and a reconstruction step [13]. Based on its eight neighbors, we calculate eight directional derivatives at each pixel . Another iterative demosaicking strategy was given by Gunturk [7]. Investigating the way that the three color channels of a characteristic picture are profoundly connected. Firstly the picture is inserted by utilizing Bilinear or other demosaicking routines, and afterward green channel is upgraded by the high recurrence data of red and blue channels. Finally a wavelet based iterative procedure was utilized to upgrade the high recurrence points of interest of the red and blue channels as indicated by the green channel.

In all sort of color demosaicking strategies inclination examination plays an essential or we can say a focal part in reproducing sharp edges. On the other hand, the inclination estimation may not be hearty when the info sign surpasses the Nyquist recurrence. This demonstrates that the fundamental driver of shading relics in demosaicked pictures [4]. A typical proceeding in color demosaicking is to endeavor the relationship between the color channels. As we realize that the three shading channels for a characteristic picture are very interrelated, the distinction between the green channel and the red or blue channel constitutes a smooth (low-pass) process. Besides, we watch that this shading distinction sign is not related to the interjection slips of inclination guided shading demosaicking routines, which is essentially band-pass process.

II. MINIMUM MEAN SQUARE ERROR

These considerations provide a principle for calculating the color difference signals by minimum mean square-error determination method [11], which provides a good approximation to the optimal determination in mean square-error. The minimum mean square-error determination are obtained in both vertical as well as horizontal direction, and then fused optimally to remove the demosaicking

Cite this article as: Nidhi Chauhan, Shanu Sharma. "Review on Demosaicking via Directional Linear Minimum Mean Square Error Estimation." *International Conference on Information Engineering, Management and Security (2015)*: 244-250. Print.

noise. At the end, the full resolution for three color channels are rebuilt from the minimum mean square-error determination filtered difference signals. The practical results proves that this new color demosaicking technique gives better improvement both in PSNR measure and visual perception. This technique is basically used for Bayer pattern which is mostly found everywhere, but it can be elongated for other CFA patterns also [5]. To take an favor of spectral interrelations, we calculated the red-green and the blue-green difference images, called primary difference signals (PDSPDS), instead of precisely getting the missing color samples.

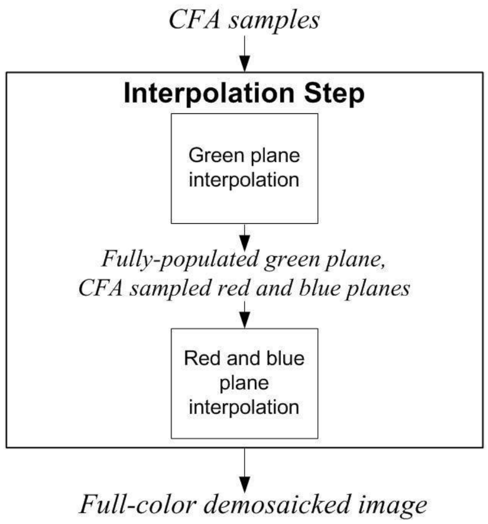


Figure 2: Algorithm for demosaicking

A linear model as shown in Figure 2 is used to show and calculate the red-green and the blue-green PDS signals. Observed PDS values are showed as the sum of color interpolation error (IE) and true PDS. Based on the second order measurements of these components, a minimum mean square error detection method is used to calculate the PDS from the strident measurements. From the calculated PDS, we determined a full resolution green image. The red and the blue channels are then after recovered.

III. PRIMARY DIFFERENCE SIGNAL AND DIRECTIONAL DEMOSAICKING NOISE

To estimate high-frequency countenance ahead the Nyquist frequency of Color Filter Array, a color demosaicking scheme depends on some additional properties for the input color signals. A very common property is the interrelation between the sampled primary color channels: i.e. the red, the green, and the blue. To exploit this property in demosaicking method, let us find the relationships between the green and the red channels, and between the green and the blue channels. There are number of reasons why the green channel has a major role in our determination of omitted color samples. Firstly, the green channel has double of the total samples as the other two channels in the Bayer mosaic pattern. Secondly, the perceptiveness of the human eye at the green wavelength. Thirdly, the green channel is close to the red and to the blue than the difference between the red and the blue in wavelengths [6]. We assume that the difference images between the green and the red channels and between the green and the blue channels are low-pass signals, that are referred one after the other as primary difference signals (PDS) and are denoted by

$$\Delta g\ ,\ r(n) = Gn - Rn$$
$$\Delta ,b(n) = Gn - Bn$$

Where n is known as the position index of the pixel.

We calculate PDS Δg,r and Δg ,b instead of particular color channels directly because random processes Δg ,r and Δg ,b have some statistical properties that can be extracted for demosaicking. We are interested in how the demosaicking noise relates to Δg ,r and Δg ,b. The most effective color demosaicking filters is the second-order directional Laplacian filter which is depends on a robust basis that Δg ,r and Δg ,b are constant in either horizontal direction or vertical direction. The major question is that which Interpolation heading is to be chosen

Using the interpolated missing green and red values, we obtain two conclusion of the arbitrary mechanism *Delta* g, r in horizontal and vertical directions as shown in Figure 3, respectively

IΔhg, r (i) = IG hi - R hi , G is interpolated

as opposed to settling on a brutal choice, we can make two different determinations of a missing essential shading specimen in level and vertical bearings, and afterward consolidating the two determinations.

We have seen the design of the Bayer sample. A section and a column of exchanging green and red specimens cross at a red inspecting location, where the missing green quality is to be resolved. Correspondingly we go for blue channel too. Similarly we go for blue channel also. We denote the red sample at the center of the window as *R*0. Its interlaced red and green neighbors in horizontal

Cite this article as: Nidhi Chauhan, Shanu Sharma. "Review on Demosaicking via Directional Linear Minimum Mean Square Error Estimation." *International Conference on Information Engineering, Management and Security (2015)*: 244-250. Print.

direction are labeled as $Rh\ i$, i²…,-4,-2,2,4 and $Gh\ i$, i ² …,-3,-1,1,3 respectively; similarly, the red and green neighbors of $R0$ in vertical direction are $Rh0$, and Rvi.

To get some coarse dimensions of PDS Δg,r and Δg,b, we first interpose the omitted green samples at red and blue pixels and then interpose the missing red and blue samples at green samples. Any of the prevailed interpolation methods for color demosaicking may be used.

For any red original sample Rhi or Rhj, the corresponding omitted green sample is interpolated as

$$IG_{hi} = 1/2\ (G_{hi-1} + G_{hi+1}) + 1/4\ (2.\ R_{hi} - R_{hi-2} - R_{hi+2})$$
$$IG_{vj} = 1/2\ (G_{vj-1} + G_{vj+1}) + 1/4\ (2.\ R_{vj} - R_{vj-2} - R_{vj+2})$$

Similarly, for any original green sample Ghi or Gvj, the corresponding missing red sample is interpolated as

$$IR_{hi} = 1/2\ (R_{hi-1} + R_{hi+1}) + 1/4\ (2.\ G_{hi} - G_{hi-2} - G_{hi+2})$$
$$IR_{vj} = 1/2\ (R_{vj-1} + R_{vj+1}) + 1/4\ (2.\ G_{vj} - G_{vj-2} - G_{vj+2})$$

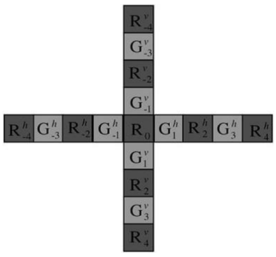


Figure 3:Row and column of mosaic data that intersect at red sample position.

$$= G\ hi - IR\ hi\ ,\ R \text{ is interpolated}$$
and
$$I\Delta vg,\ r\ (i) = IG\ vi - R\ vi\ ,\ G \text{ is interpolated}$$
$$= G\ vi - IR\ vi\ ,\ R \text{ is interpolated}$$

The measurement errors associated with $I\Delta_{hg,\ r}$ and $I\Delta_{vg,\ r}$ are

$$E_{hg,\ r} = \Delta_{g,\ r} - I\Delta_{hg,\ r}$$
$$E_{vg,\ r} = \Delta_{g,\ r} - I\Delta_{vg,\ r}\ .$$

We regard $I\Delta_{hg,\ r}$ and $I\Delta_{vg,\ r}$ to be two observations of $\Delta_{g,\ r}$ and, accordingly, $E_{hg,\ r}$ and $E_{vg,\ r}$ to be the correlative directional demosaicking noises, and rewrite and as

$$I\Delta_{hg,\ r} = \Delta_{g,\ r} - E_{hg,\ r}$$
$$I\Delta_{vg,\ r} = \Delta_{g,\ r} - E_{vg,\ r}$$

Now, the we have to find an optimal estimate of Δg,r from the two observation sequences $I\Delta_{hg,\ r}$ and $I\Delta_{vg,\ r}$, and then acquire the omitted green values The evaluated algorithm will be developed in the next step.

To abridge the notations, we express by x the true PDS signal Δg ,r and by y the linked observation $I\ \Delta h$,r or $I\ \Delta v$,r and by v, the linked demosaicking noise $E_{hg,r}$ or $E_{vg,r}$ namely

$$y(n) = x(n) + v(n)$$

The optimal minimum mean square-error estimation (MMSE) of x is

nd $E_{vg,\ r}$.

IV. INTERPOLATION OF OMITTED RED (BLUE) SAMPLES AT THE BLUE (RED) SAMPLE LOCATIONS

We firstly interpolate the omitted red sample at a blue pixel B_n. R_{nnw} , R_{nsw}, R_{nne}, R_{nse} are the four nearest red neighbors of the blue sample position B_n, where the superscripts are directional notations for northwestern, southwestern northeastern and southeastern.Note that R_{nnw} , R_{nsw} , R_{nne} , R_{nse} and B_n are all original samples in the Bayer Pattern. The calculated green samples at these locations are expressed by

$G1nwn$, $G1swn$,$G1nen$, $G1sen$, $G1n$ respectively as shown in Figure5.

At that point, the omitted red specimen is assessed as

$$R1n = G1n - \Delta n,\ gr$$

Thus the missing blue samples at red sample positions R_n can be interjected. The four green-blue distinction values in the northwestern, southwestern, northeastern and southeastern of R_n are accessible.

$$x1 = M[x/y] = \int xp(x/y)dx$$

Then again, the MMSE estimation is extremely extreme, if conceivable by any means, on the grounds that p(x/y) is sometimes known practically speaking. Rather, we utilize the LMMSE strategy to gauge x from y, which is a decent estimate to MMSE however more

Cite this article as: Nidhi Chauhan, Shanu Sharma. “Review on Demosaicking via Directional Linear Minimum Mean Square Error Estimation.” *International Conference on Information Engineering, Management and Security (2015)*: 244-250. Print.

manageable to proficient usage. Especially, if x(n) and v(n) are provincially Gaussian forms (a sensible supposition for normal picture signals), then The LMMSE of x is computed as

$$x = M[x] + cov(x,y)/var(y)*(y-M[y])$$

we examine that the demosaicking noises $E_{hg,r}$ and $E_{vg,r}$ are zero-mean random processes, and they are almost uncrated with $\Delta_{g,\ r}$. Consequently, we can simplify the above equation to

$$x1 = \mu_x + \{S_x\text{^}2/(S_x\text{^}2+ S_y\text{^}2)*(y- \mu_x)$$

$$\text{Where } \mu_x = M[x],\ S_x\text{^}2 = Var(x),\ S_y\text{^}2 = Var(y).$$

Uniformly, we can characterize the difference signal Δg,b between the green and blue and its two assessments $I\Delta_{hg,b}$ and $I\Delta_{vg,b}$, in horizontal and vertical directions. The *n* relating estimation lapses $E_{hg,\ b}$ and $E_{vg,\ b}$ have the same means as those of $E_{hg,\ r}$ a

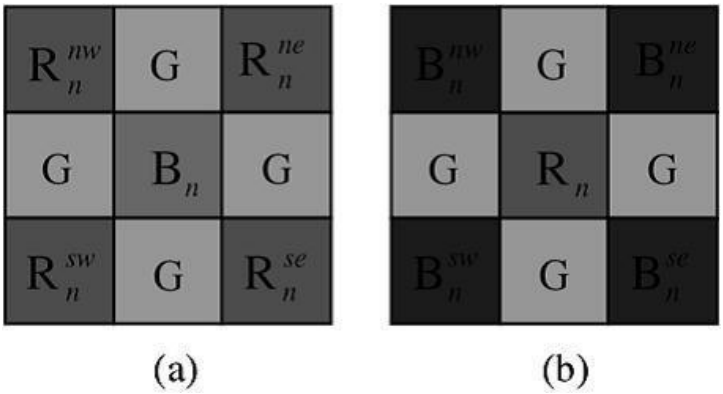


Figure 5: (a) Blue sample and its four nearest red neighbors. (b) Red sample and its four nearest blue neighbors

V. INTERPOLATION OF OMITTED RED/BLUE SAMPLES AT THE GREEN SAMPLE LOCATIONS

When the omitted red/blue samples at the blue/red positions are filled, we come across at the four cases. As earlier, the samples are calculated ones if marked with 1 and original ones otherwise. Because of the uniformity between the red and the blue samples in these four cases, we only need to discuss case (1). Given the green assessments $G1_{nn}$, $G1_{sn}$, $G1_{en}$ and $G1_{wn}$ at the positions $R1_{nn}$, $R1_{sn}$, $R1_{en}$ and $R1_{wn}$ we have the related four green–red difference values, denoted by $\Delta_{nn,\ gr}$,$\Delta_{sn,\ gr}$,$\Delta_{en,\ gr}$,and $\Delta_{wn,\ gr}$ shown in

Figure 6. As in the preceding step, we calculate the bilinear

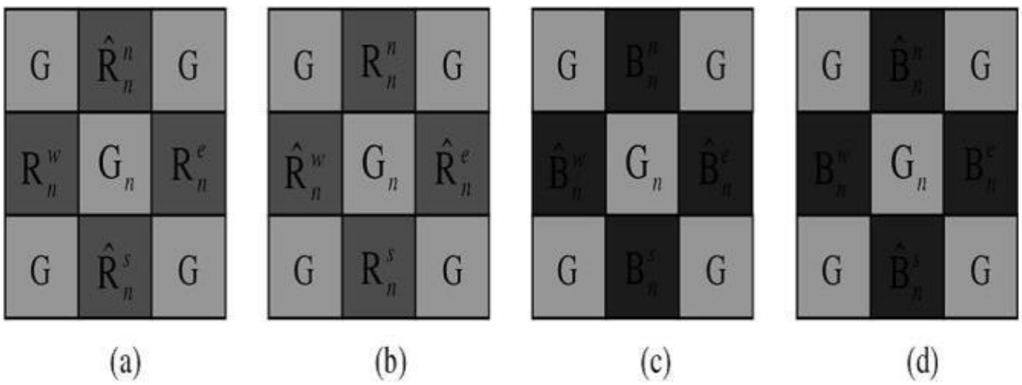


Figure 6: (a)(b) Green sample and its two original and two estimated red neighbors. (c)(d) Green sample and its two original and two estimated blue neighbors

$$\Delta_{n,\ gr} = (\Delta_{nn,\ gr} + \Delta_{sn,\ gr} + \Delta_{en,\ gr} + \Delta_{wn,\ gr})\ /\ 4$$

Then, the omitted red sample at green sample position G_n is estimated to be

$$R1n = G1n - \Delta n,gr$$

Likely, the omitted blue sample at a green position is estimated as

$$B1n = G1n - \Delta n,\ gb$$

Cite this article as: Nidhi Chauhan, Shanu Sharma. "Review on Demosaicking via Directional Linear Minimum Mean Square Error Estimation." *International Conference on Information Engineering, Management and Security (2015)*: 244-250. Print.

All the omitted red and blue samples have been filled till now. The full color picture is remade. The exhibited demosaicking plan first adventures the relationship between the green and red/blue channels to acquire great appraisals of the omitted green specimens and then gauges the missing red and blue average of the green-red differences [8]. tests by a basic and quick bilinear normal operation on the green-red and green-blue PDS signals [11].

VI. EXPERIMENTAL RESULTS

We implemented the proposed LMMSE color demosaicking algorithm, and tested it on a large number of natural color images. In this section, we present our experimental results for the three images in Fig. 7 shown below, and compare them with the methods of Hamilton *et al.* [11], Chang *et al.* [7], and Gunturk *et al.* [9], which are among the best schemes. The results reported in the recent paper of [9] were better than the previously published algorithms, especially for the red and blue channels. Table I below shows the PSNR values of the tested images. The proposed LMMSE-based demosaicking algorithm appears to produce visually more pleasant color images with color artifacts greatly suppressed.

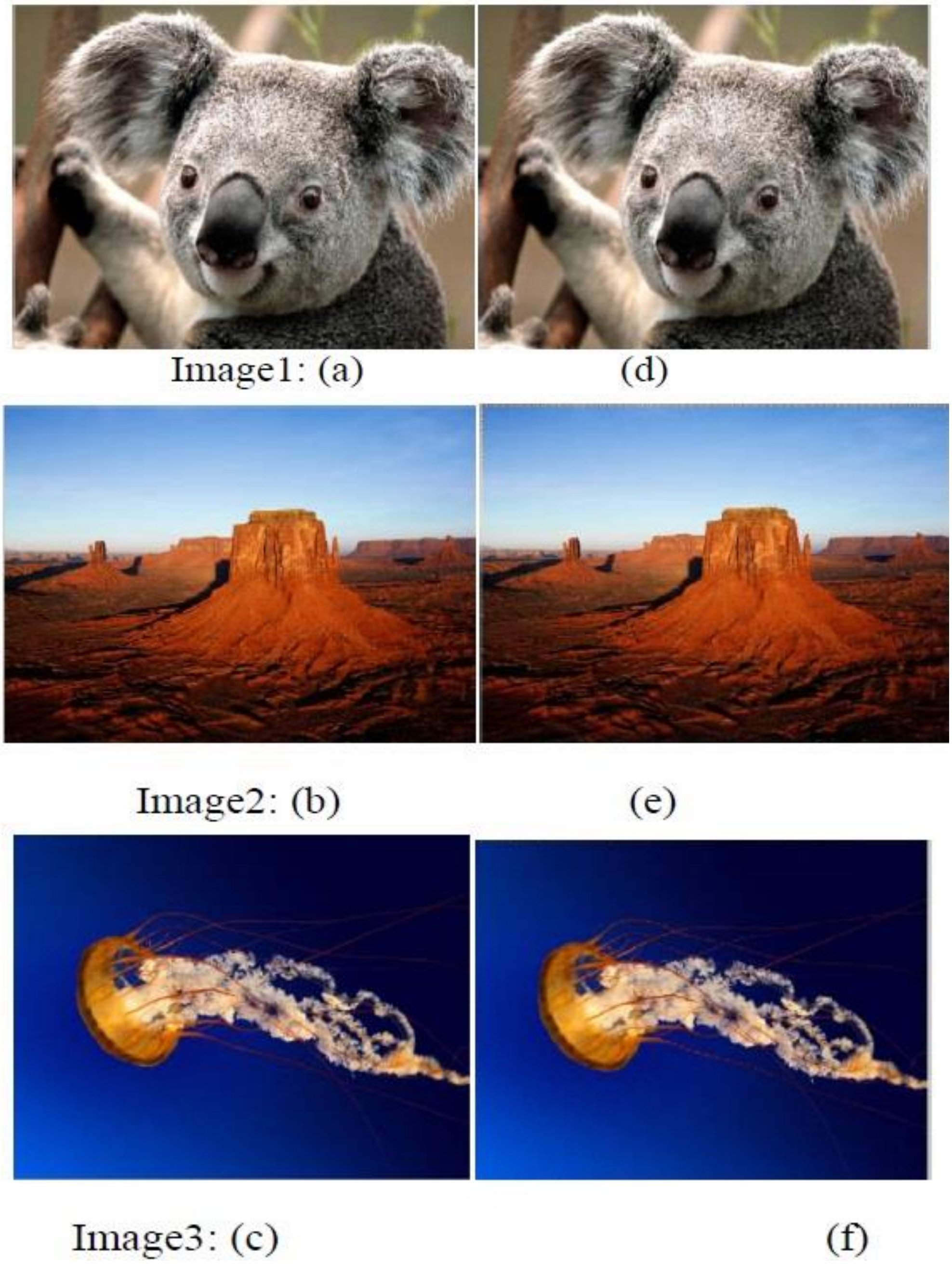


Figure 7: Image (a)-(c) input and (d)-(f) are output

Cite this article as: Nidhi Chauhan, Shanu Sharma. "Review on Demosaicking via Directional Linear Minimum Mean Square Error Estimation." *International Conference on Information Engineering, Management and Security (2015)*: 244-250. Print.

Table I: PSNR of Tested images

IMAGE NO.	PSNR FOR RED CHANNEL	PSNR FOR GREEN CHANNEL	PSNR FOR BLUE CHANNEL
1	40.47	41.31	37.39
2	42.90	43.33	40.06
3	43.01	44.40	41.90

REFERENCES

[1] J. E. Adams, "Intersections between color plane interpolation and other image processing functions in electronic photography," Proceedings of SPIE, vol. 2416, pp. 144-151, 1995.
[2] J. E. Adams and J. F. Hamilton Jr., "Adaptive color plane interpolation in single color electronic camera," U. S. Patent, 5 506 619, 1996.
[3] J. E. Adams, "Design of practical color filter array interpolation algorithms for digital cameras," Proceedings of SPIE, vol. 3028, pp. 117-125, 1997.
[4] Weiran Tang, Oscar C. Au, Xing Wen, Yi Yang and Lu Fang, "LMMSE Frequency Merging for Demosaicking", IEEE, vol. 90, pp. 123-132, 2002.
[5] B. E. Bayer and Eastman Kodak Company, "Color Imaging Array," US patent 3 971 065, 1975.
[6] D. R. Cok and Eastman Kodak Company, "Signal Processing method and apparatus for producing interpolated chrominance values in a sampled color image signal," US patent 4642678, 1987.
[7] B. K. Gunturk, Y. Altunbasak and R. M. Mersereau, "Color plane interpolation using alternating projections," IEEE Trans. ImageProcessing, vol. 11, pp. 997-1013, 2002.
[8] W. T. Freeman, "Method and apparatus for reconstructing missing color samples," U. S. Patents, 4 663 655.
[9] E. Chang, S. Cheung and D. Y. Pan, "Color filter array recovery using a threshold-based variable number of gradients," Proceedings of SPIE, vol. 3650, pp. 36-43, 1999.
[10] Lei Zhang and Xiaolin Wu, "Color Demosaicking via Directional Linear Minimum Mean square- Error Detection," IEEE Trans on Image Processing, Vol. 14, No. 12, Dec 2005.
[11] Fan Zhang, Xiaolin Wu, Xiaokang Yang, Wenjun Zhang, and Lei Zhang, "Robust Color Demosaicking With Adaptation to Varying Spectral Correlations," IEEE Trans on Image Processing, Vol. 18, No. 12, Dec 2009.

Cite this article as: Nidhi Chauhan, Shanu Sharma. "Review on Demosaicking via Directional Linear Minimum Mean Square Error Estimation." *International Conference on Information Engineering, Management and Security (2015)*: 244-250. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS042		**eAID**	ICIEMS.2015.042

Microphysical Parameters Analysis of Cloud using X & Ka band Dual Polarized Doppler Weather Radar

Mr. Anurag Tirthgirikar[1], Prof. Milind Patil[1], Dr. Kaustav Chakravarty[2]
[1]Deptment. of Electronics and Telecommunication, Vishwakarma Institute of Information Technology, Pune, India.
[2]Physical Meteorology & Aerosol Division, Indian Institute of Tropical Meteorology, Pashan, Pune, India.

Abstract: *The study of microphysical parameter of clouds is studied and analyzed using X &Ka band Dual Polarized Doppler Weather Radar. Field work is done on operational characteristics of both the radars stationed at Mandhardevi and is used to analyze cloud pattern behavior over Mahabaleshwar region during monsoon period. X band radar is used for analysis of precipitation characteristics of cloud for convective and stratiform rainfall pattern. Ka band radar is used for study and analysis of cloud dynamics which include cloud height behavior during a rain event over Mahabaleshwar region. Disdrometer is used to measure the drop size distribution for convective and stratiform rain events over Mahabaleshwar.*

Keywords: Radar, Reflectivity, Disdrometer, Cloud dynamics.

I. INTRODUCTION

India being a tropical country has four seasons throughout the year, out of which monsoon is one of the major season as the India's agrarian economy mostly depends upon monsoon for rains which is important factor for agricultural activities. Monsoon season occurs in Indian during phase of June to September which is further briefly classified into two sub types i.e. South-West monsoon & North-East monsoon. In this project we are concentrating on South-West monsoon season winds which interact first with Western Ghats region.

In present study we are going to study the rainfall pattern over Western Ghats region. Here we are concentrating on cloud behavioral pattern over Mahabaleshwar region. RADAR observations are firstly analyzed from both X & Ka band Dual Polarized Doppler Weather RADAR owned by Indian Institute of Tropical Meteorology (IITM), Pune a premier research institute under Ministry of Earth Sciences, Govt. of India. Study mainly focusses on two types of rainfall pattern i.e. convective rainfall[3] and stratiform rainfall[4] over High Altitude Cloud Physics Lab (HACPL) which is stationed at Mahabaleshwar, X & Ka band RADAR are stationed at a place Mandhardevi which is at same altitude as that of HACPL. Disdrometer analysis is also been taken into consideration for this study of microphysical parameters of cloud.[8]

II. THEORY

A Doppler radar is a specialized radar that makes use of the Doppler Effect to produce velocity data about objects at a distance. It tdoes this by beaming a microwave signal towards a desired target and listening for its reflection, then analyzing how the frequency of the returned signal has been altered by the object's motion. This variation gives direct and highly accurate measurements of the radial

Cite this article as: Anurag Tirthgirikar, Milind Patil, Kaustav Chakravarty. "Microphysical Parameters Analysis of Cloud using X & Ka band Dual Polarized Doppler Weather Radar." *International Conference on Information Engineering, Management and Security (2015)*:251-254. Print.

component of a target's velocity relative to the radar. Doppler radars are used in aviation, sounding satellites, meteorology, etc. Most modern weather radars use the pulse-Doppler technique to examine the motion of precipitation, but it is only a part of the processing of their data.[11] Below is the (1) determining how Doppler shift can be calculated, this Doppler shift can be used to find velocity or direction of microphysical parameters of cloud.

$$fr = ft\left(\frac{1+v/c}{1-v/c}\right) \qquad (1)$$

Here, c is the speed of light, v is the target velocity gives the shifted frequency (f_r) as a function of the original frequency (f_t). Doppler Weather Radar are usually classified into two main category i.e. Single Polarized & Dual Polarized Doppler Weather Radar. In this study we are using X & Ka band Dual Polarized Doppler Weather Radar.

III. EXPERIMENTAL SETUP

The experimental setup consist of X & Ka band Dual Polarized Doppler Weather Radar which operates with two scanning strategy i.e. Plan Position Indicator (PPI) scan where radar is set up at a fix elevation angle w.r.t horizon and the azimuth angle is varied and a scan of 360° is taken. Another scanning strategy is Range Height Indicator (RHI) scan in which azimuth is kept constant at certain specific angle as in our case pointing towards Mahabaleshwar which lies at 237° from the azimuth angle and elevation angle is varied from 0° to 90°. In this study we have concentrated on RHI scan strategy which helps us to give vertical structure of cloud over Mahabaleshwar region. RHI scan plots can be seen for backscattered signal using EDGE software, by analyzing these plots we can differentiate between a precipitating rain event and a non-precipitating rain event. Based on this analysis we can classify rain events during monsoon and can study how a rain transition occurs. EDGE software gives us a basic data file which consist of all the products or as we call it as radar parameters which include reflectivity (DBZ), velocity (VEL), differential reflectivity (ZDR), range, elevation, time and many more products. This base file is of .vols format for X band and .rpp for Ka band but the problem with these file format is that these cannot be analyzed by any software. So these file formats are converted into common data file format such as comma separated value (.csv) / Net CDF (.NC) file formats so this basic file format data can be used for analysis in simulation software such as MATLAB. This common formats can then be accessed by extracting data from this file about the products using MATLAB coding. MATLAB is then used for plotting various radar product behavior and by analyzing these plots we can derive our conclusions linking the plot results with theoretical atmospheric science phenomena.

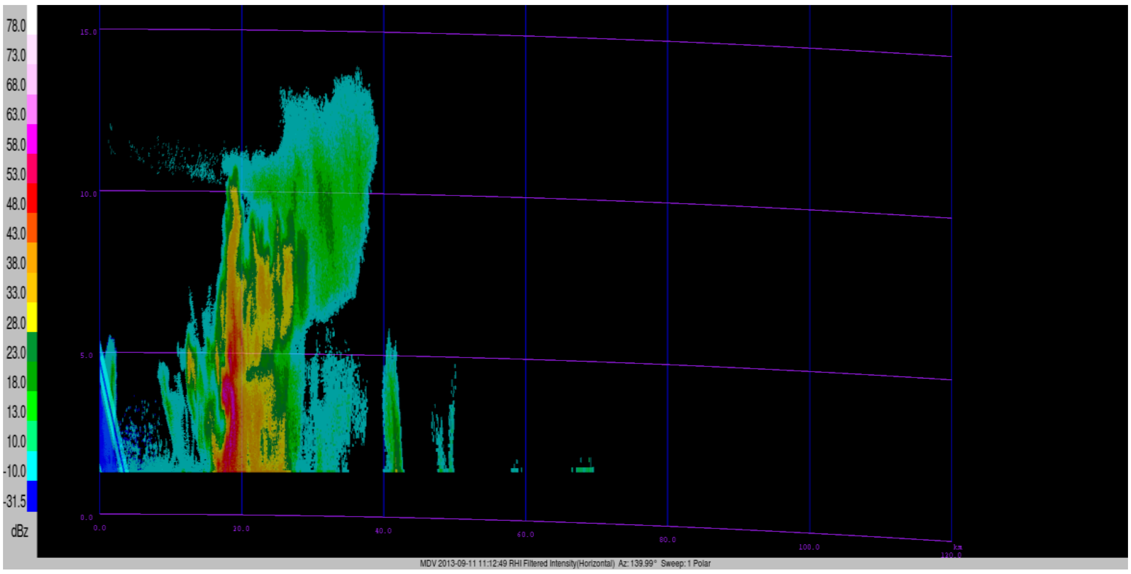


Figure 1: RHI scan plot generated by EDGE software

IV. OBSERVATIONS

The observations taken with help X band radar data analysis using MATLAB give us detail about the precipitation characteristics of cloud and help us to classify whether the rain event is a convective rain event or a stratiform rain event. The below plots are plotted with gate width on X axis and vertical height on Y axis.[1]

Cite this article as: Anurag Tirthgirikar, Milind Patil, Kaustav Chakravarty. "Microphysical Parameters Analysis of Cloud using X & Ka band Dual Polarized Doppler Weather Radar." *International Conference on Information Engineering, Management and Security (2015)*:251-254. Print.

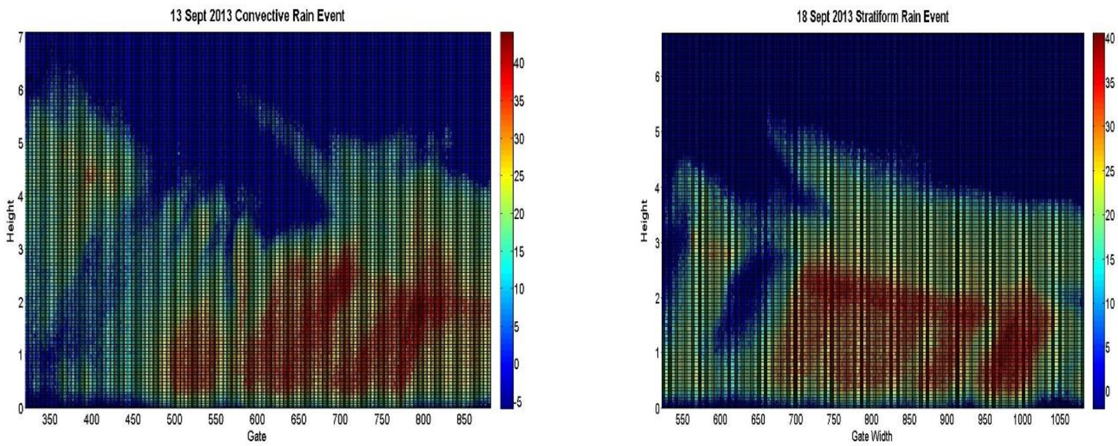


Figure 2: Convective & Stratiform rain event contour plots

Reflectivity is varying factor in both of above graphs which is varying according to color bar scale from -5 to 40 dBz. Red color signifies maximum reflectivity at that particular height and a series of band of such red gates make up a bright band, by observing bright band we can classify whether a rain event is convective[2] one or stratiform one. A convective rain event has vertical bright band structure in any rain event and a stratiform rain event has a horizontal bright band structure.[4]

The observation taken with help of Disdrometer help us to study the variation of raindrop size distribution in convective and stratiform rain event which is one of the important microphysical parameter in clouds.[6]

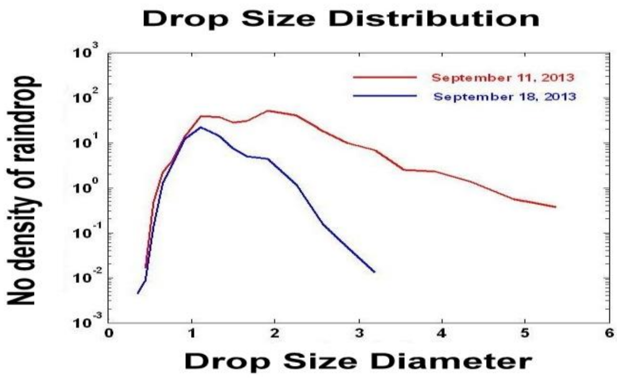


Figure 3: Rain drop size distribution using Disdrometer data

Here we can see graph is plotted between drop size diameter in mm v/s no density of raindrop. The red line in graph depicts about convective rain event we can see that drop size density and diameter of drop is quite higher than that of blue line depicting a stratiform rain event. This shows that raindrops present in a convective rain event are usually larger and denser than its counterpart stratiform rain event.[5][7]

The observation done with Ka band radar help us to find detail about cloud dynamics. In this study we have done analysis on clouds height transition over Mahabaleshwar region during a rain event.

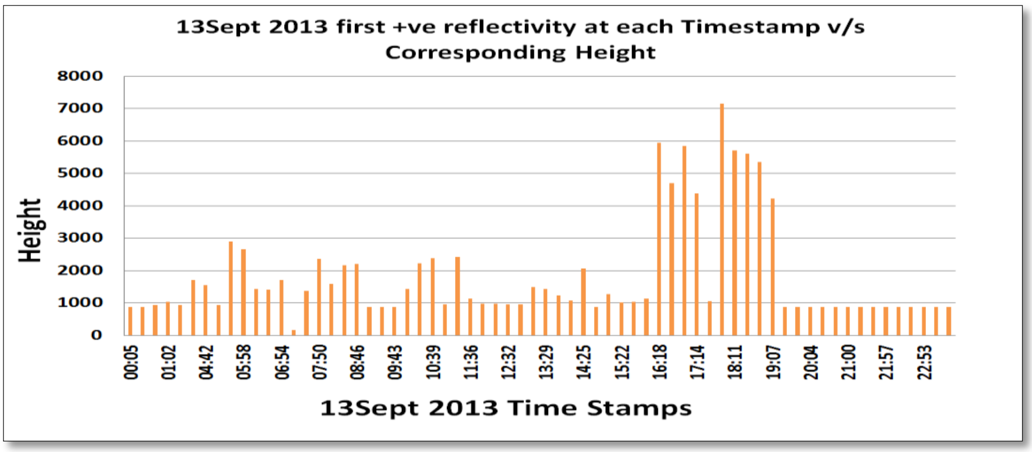


Figure 4: Clouds height transition over Mahabaleshwar

Cite this article as: Anurag Tirthgirikar, Milind Patil, Kaustav Chakravarty. "Microphysical Parameters Analysis of Cloud using X & Ka band Dual Polarized Doppler Weather Radar." *International Conference on Information Engineering, Management and Security (2015)*:251-254. Print.

The cloud height is taken at each time stamp (RHI scan time) measuring first positive reflectivity value using Ka band radar data. The graph is plotted between time stamps on X axis and vertical height in kms. The time stamps considered here are for a whole day in a rain event. As we can see that cloud height is varying from 1-3 km before the actual transition of rain event.[10] When the transition begins we can see that cloud height suddenly increase to 5-8 km this phenomena occurs when clouds after accumulation of raindrop usually move upwards which clearly tells us that this transition is of updraft of clouds before rain. As the cloud transit at greater height due to temperature variation the accumulated rain drop start to percolate in form of rain and suddenly cloud height is decreases and rainfall occurs this describes that cloud lower their height during a rainfall transition in rain event.[9]

V. ACKNOWLEDGEMENT

This work is supported by Indian Institute of Tropical Meteorology (IITM), Pune, India. The authors would like to acknowledge IITM to avail the facility to conduct study and analysis on their RADAR and Disdrometer.

REFERENCES

[1] André Tremblay, "The Stratiform and Convective Components of Surface Precipitation," *J. Atmos. Sci.*, vol. 62, pp. 1513–1528, May 2005.

[2] Sergey Y. Matrosov, Robert Cifelli and David Gochis, "Measurements of Heavy Convective Rainfall in the Presence of Hail in Flood-Prone Areas Using an X-Band Polarimetric Radar." *J. Appl. Meteor. Climatol.,* vol. **52**, pp. 395–407, February 2013.

[3] Yali Luo, Hui Wang, Renhe Zhang, Weimiao Qian and Zhengzhao Luo, "Comparison of Rainfall Characteristics and Convective Properties of Monsoon Precipitation Systems over South China and the Yangtze and Huai River Basin." *J. Climate*, vol. 26, pp. 110–132, January 2013.

[4] https://en.wikipedia.org/wiki/Precipitation_types

[5] Ali Tokay, Walter A. Petersen, Patrick Gatlin and Matthew Wingo,"Comparison of Raindrop Size Distribution Measurements by Collocated Disdrometers." *J. Atmos. Oceanic Technol.*, vol. 30, pp.1672–1690, August 2013.

[6] Ali Tokay, Anton Kruger and Witold F. Krajewski, "Comparison of Drop Size Distribution Measurements by Impact and Optical Disdrometers." *J. Appl. Meteor.*, vol. 40, pp. 2083–2097, November 2001.

[7] Terry J. Schuur, Alexander V. Ryzhkov, Dusan S. Zrnić and Michael Schönhuber, "Drop Size Distributions Measured by a 2D Video Disdrometer: Comparison with Dual-Polarization Radar Data." *J. Appl. Meteor.*, vol. 40, pp. 1019–1034, June 2001.

[8] Katja Friedrich, Evan A. Kalina, Forrest J. Masters and Carlos R. Lopez, "Drop-Size Distributions in Thunderstorms Measured by Optical Disdrometers during VORTEX2." *Mon. Wea. Rev.*, vol. 141, pp. 1182–1203, April 2013.

[9] Paquita Zuidema, David Painemal, Simon de Szoeke and Chris Fairall, "Stratocumulus Cloud-Top Height Estimates and Their Climatic Implications." *J. Climate.*, vol. 22, pp. 4652–4666, September 2009.

[10] Evgueni Kassianov, Charles N. Long and Jason Christy, "Cloud-Base-Height Estimation from Paired Ground-Based Hemispherical Observations." *J. Appl. Meteor.*, vol. 44, pp. 1221–1233, August 2005.

[11] https://en.wikipedia.org/wiki/Doppler_effect

Cite this article as: Anurag Tirthgirikar, Milind Patil, Kaustav Chakravarty. "Microphysical Parameters Analysis of Cloud using X & Ka band Dual Polarized Doppler Weather Radar." *International Conference on Information Engineering, Management and Security (2015)*:251-254. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS043		**eAID**	ICIEMS.2015.043

RESOLUTE MOBILE CULPRIT IDENTIFIER AND ACQUIRER

S.Srikiran Rao
Department of Electronics Communication Engineering , Easwari Engineering College

Abstract: *Smartphone plays an important role in ones day to day life. We have many smartphone tracker apps to find it when it get stolen but all those app fail to operate in switch off condition and even if we find the stolen smartphone we fail to identify the person who was responsible for the theft. Here in this paper we focused on finding our stolen smartphone and even the person who was responsible for the theft ,by using his fingerprint even when the smartphone is switched off by using touchscreen fingerprint sensor, GSM ,RTC,GPS.*

Keywords: GPS, GSM, RTC, finger print sensor, code convertor

I INTRODUCTION

Smartphones is one of the essential gadget in today's day to day life .The smartphone theft had been increased during recent years, so to reduce the increase in smartphone theft we need to eliminate the root cause, the people who is responsible behind this action. During recent years the smart phone developers have come up with many antitheft or phone tracking apps but all those facilities fail to locate or find the culprit who stole the mobile. So to help the smartphone developers and users to locate the phone and culprit we came up with this solution.

II CONSTRUCTION

REAL TIME CLOCK
It works like a counter which is capable of running in background even when the smartphone is switched Off, we use it because of its accuracy, and its power and it can be run effetely by using low power battery or by the backup battery
GSM
Global System for Mobile communication (GSM), this technology helps in sending SMS to the backup mobile number, it's a second generation (2G) technology used in this device. This network technology is divided into BSS-base station subsystem NSS-Network switching subsystem

Cite this article as: S.Srikiran Rao. "RESOLUTE MOBILE CULPRIT IDENTIFIER AND ACQUIRER." *International Conference on Information Engineering, Management and Security (2015)*: 255-259. Print.

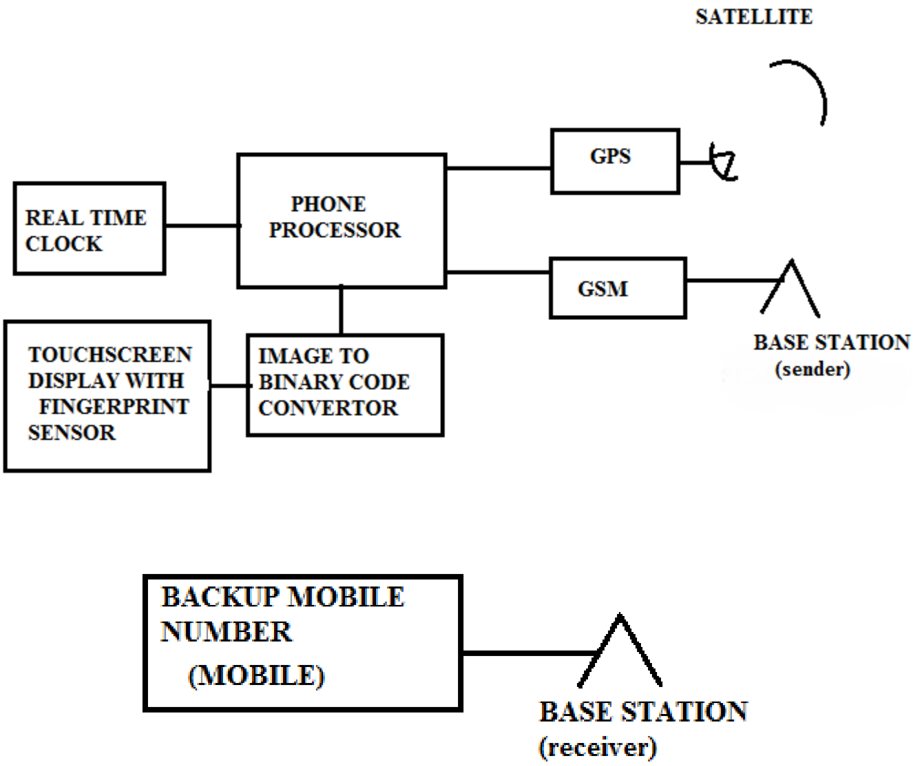


Fig 1: block diagram of mobile culprit identifier

NSS-Network switching subsystem
OSS-Operation support system
GPS
This technology is used for navigation, tracking purpose with help of the satellite. In this the smartphone uses this technology to get the GPS coordinates from the nearer satellite. The location coordinates will be in latitude and longitude.
TOUCHSCREEN DISPLAY FINGERPRINT SENSOR
Its main operation is to take necessary fingerprints of the person who stole the mobile. In this we use the whole smartphone screen as fingerprint sensor. So that it will be useful to take fingerprints. This rebuilds the image of the fingerprint.

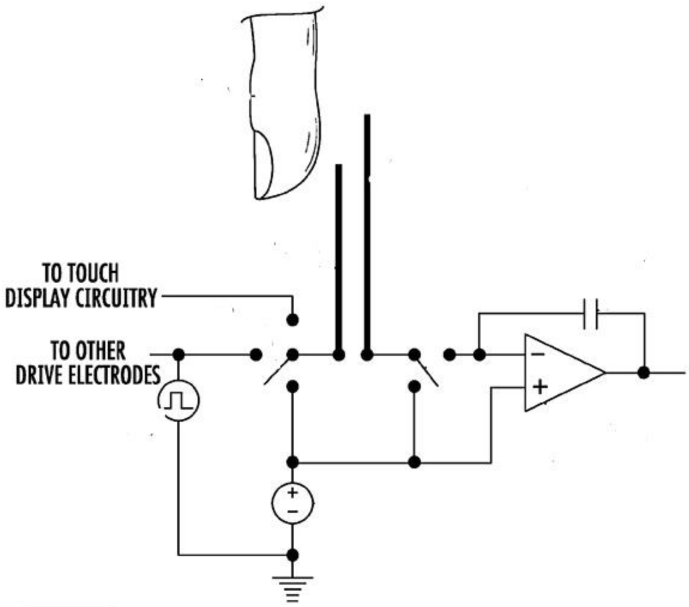


Fig 2: fingerprint sensor on the display screen.

Cite this article as: S.Srikiran Rao. "RESOLUTE MOBILE CULPRIT IDENTIFIER AND ACQUIRER." *International Conference on Information Engineering, Management and Security (2015)*: 255-259. Print.

III WORKING

In this when the smart phone device gets switched off .the RTC in the device starts to operate with help of backup battery In the device. RTC starts its counting operation and activates the GSM for every one hour. This GSM when activated checks for the base station. It verifies the present base station and the registered base station. If it is in registered base station location for the particular time period no action is taken place. If the device is not in registered location for the

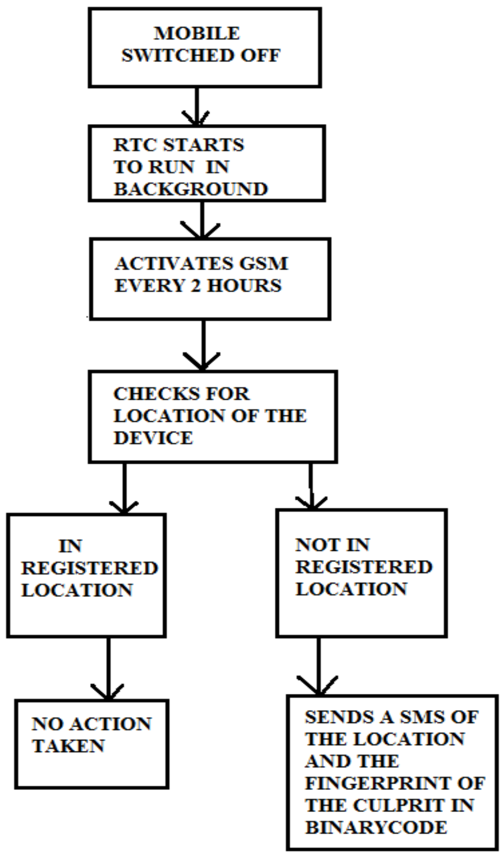


Fig 3: flow diagram of the mobile acquirer and culprit identifier.

Particular time period. Then the GPS gets activated and gets the necessary location coordinates of the phone and along with the fingerprint of the person who operated the phone after stolen. This fingerprint is converted into binary code like 0 and 1 by using code converter. After converting into binary number those code along with the GPS location is sent to the backup mobile number as SMS. So that by using those coordinates we can find the location and with help of the binary code we can convert into image and with help of this fingerprint image we can very it with id (voter ID, other ID issued by government).

IV RESULT

Before using the device the user of the smartphone must registered his details such as the location and timing of the smartphone in his device.

TIME	LOCATION
8am-10am	Gsm location a
10am -8pm	Gsm location b
8pm-10pm	Gsm location c

Table1: represents the location and time period of the device registered

Cite this article as: S.Srikiran Rao. "RESOLUTE MOBILE CULPRIT IDENTIFIER AND ACQUIRER." *International Conference on Information Engineering, Management and Security (2015)*: 255-259. Print.

After entering the location and timing details, then the user have to enter the backup mobile number to which the details like location coordinates and the binary coded fingerprint image will be sent as SMS.

Fig 4: entering of back up mobile numbers

So, now when our mobile get stolen, and switched off by the person or switched off due to no battery. This activates the RTC and it starts it counting process , now here it get GPS coordinates or the positions of the stolen device along with the finger print which is converted into binary form and is sent as SMS to the backup mobile number as shown in figure

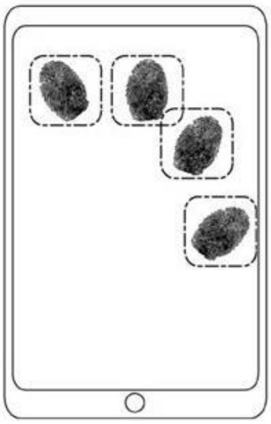


Fig 5: shows fingerprint of the culprit on the display.

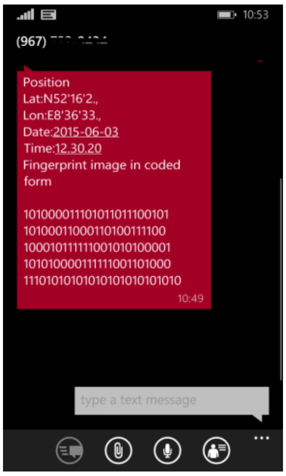


Fig 6: message received to the backup number from the stolen smart phone.

Fig 7: shows the match found to the fingerprints on the smartphone screen.

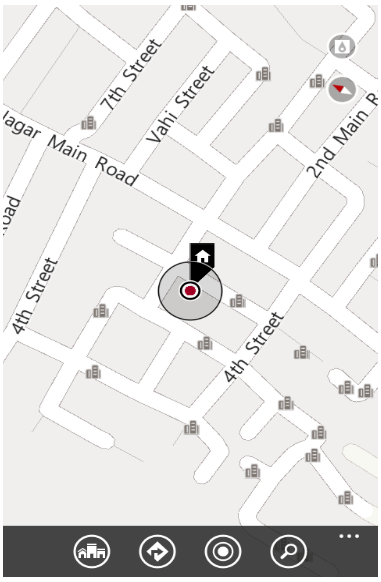


Fig 8: shows the location of the stolen mobile phone, which can be tracked using the GPS coordinates.

V REFERENCE

[1] http://en.wikipedia.org/wiki/Real-time clock.
[2] https://learn.adafruit.com/ds1307-real-time-clock-breakout-board-kit/what-is-an-rtc.
[3] http://en.wikipedia.org/wiki/GSM.
[4] http://www.gps.gov/
[5] https://www.google.co.in/webhp?sourceid=chrome-instant&ion=1&espv=2&ie=UTF 8#q=intel%20phone%20processor.

Cite this article as: S.Srikiran Rao. "RESOLUTE MOBILE CULPRIT IDENTIFIER AND ACQUIRER." *International Conference on Information Engineering, Management and Security (2015)*: 255-259. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	VOL	01
Website	www.iciems.in	eMail	iciems@asdf.res.in
Received	10 - July - 2015	Accepted	31- July - 2015
Article ID	ICIEMS044	eAID	ICIEMS.2015.044

A Novel Proactive Secret Sharing

Sayantan Mandal[1], Prof.V. Ch. Venkaiah[2]
[1,2] School of Computer and Information
Sciences, University of Hyderabad, Hyderabad-500046, India

Abstract: *A (t, n) secret sharing scheme that divides a secret into n shares in such a way that any t or more than t shares can reconstruct the secret; but fewer than t shares cannot reconstruct the secret. A share renewal protocol is to protect a secret in long-lived system by distributing it to a group of n participants and refreshing their shares periodically in this fixed group. In this paper we propose a share renewal protocol without the presence of a trusted party. Shareholders renew their shares among themselves using old shares without disclosing their shares at any stage of the protocol and able to reconstruct the secret without changing it. The protocol is robust and maintain privacy of shares.*

I. INTRODUCTION

In Cryptography, a secret sharing scheme refers to any method for distributing a secret among a group of participants, each of which is allocated a share of the secret. The secret can only be reconstructed when an authorized subset of participants combine their shares.In a secret sharing scheme, a secret s is divided into n shares by a dealer and shared among n shareholders in such a way that any t or more than t shares can reconstruct this secret; but fewer than t shares cannot reconstruct the secret s. Such a scheme is called a (t, n) secret sharing scheme.

Some of the Secret Sharing schemes assume long-lived shares; however the protection provided by these schemes may be insufficient after some time. Several faults might occur, shares can gradually be corrupted/compromised, hardware failure or damage may take place . One way to overcome this problem is to refresh the shares periodically. Thus the concept of proactive secret sharing was introduced.The goal of the proactive security scheme is to prevent the adversary from learning the secret or from destroying it. In particular any group of t non-faulty shareholders should be able to reconstruct the secret whenever it is necessary. The term pro-active refers to the fact that it's not necessary for a breach of security to occur before secrets are refreshed, the refreshment is done periodically and hence, proactively.

The core property of pro-active secret sharing is to renew the existing shares without changing the secret, so that previous exposures of shares will not damage the secret. This should be performed without, of course, any information-leak or any secret change. Proactive security for secret sharing was first suggested by Herzberg et al. [6] where they presented, among other things, a proactive polynomial secret sharing scheme. Proactive security refers to security and availability of secret in the presence of a mobile adversary. Herzberg et al. [6] further specialized this notion to robust secret sharing schemes and gave a detailed efficient proactive secret sharing scheme. Robust means that in any time period, the shareholders can reconstruct the secret value correctly.

A. Related Work

Secret sharing schemes were introduced by both Blakley [2] and Shamir [4] independently in 1979. Shamir's (t,n) secret sharing scheme is a threshold scheme based on polynomial interpolation. It allows a dealer D to distribute a secret s to n players, such that at least t less than n players are required to reconstruct the secret. The protocol is information theoretically secure, i.e., any fewer than t

Cite this article as: Sayantan Mandal, V. Ch. Venkaiah. "A Novel Proactive Secret Sharing." *International Conference on Information Engineering, Management and Security (2015)*:260-265. Print.

players cannot gain any information about the secret by themselves. In 1985, Chor et al. [5] extended the notion of the original secret sharing and presented the concept of verifiable secret sharing (VSS). The property of verifiability allows the shareholders to verify their shares for consistency. The proactive secret sharing(PSS) has been studied extensively in the literature [6], [7], [3],[8]. All these PSS schemes initiate the secret using a trusted party.Proactive secret sharing scheme based on Verifiable Secret Sharing (VSS) provides strong security against an active attacker. It combines the secret sharing scheme with a periodical share update process to ensure the overall security of a system.Through update mechanism, old shares become useless. Even to steal a secret; however, an attacker needs to intrude on at least t participants during the same time period if security is maintained in a (t; n) threshold secret sharing scheme.

1) Motivation: Trusted Party is required to create and renew the shares. Also these schemes make use of polynomial generation in both the share distribution and share renewal phases; thereby increasing the computational complexity. Our motivation is to design of an proactive secret sharing without a trusted party. The scheme proposed in this report is independent of trusted party and uses the polynomial generation only for share distribution and not for share renewal phase. Thus our scheme has the advantage that shares will be renewed without Dealers involvement. And polynomial is generated once in share distribution phase, share renewal phase make use of old share to generate new share without changing the secret.

2) Overview: Section II, introduces the basic security model and definitions. Section III presents our share renewal scheme without a trusted party. Section IV looks at the security issues and discusses the complexity of the proposed scheme. Conclusion remarks are in Section V.

II. MODEL AND DEFINITION

A. Model

1) Existing System: In all the existing proactive secret sharing scheme found in the literature are based on the presence of the trusted party.

[Initialization]

Dealer generates a polynomial of degree t-1.

f(x) = s + d1x + : : : + dt□1xt□1

where di, 1 _ i _ t □ 1 are the random elements of the field Fq: Then, the dealer sends the share si = f(i) mod q to the server i. Since this is Shamir's secret-sharing scheme [4], any t servers can reconstruct the secret by using Lagrange interpolation, while t □ 1 cannot get any information.

[Share Updation]

Each server i generates a polynomial of degree t-1 by using random numbers di1 : : : dit□1. where f(x) = S + di1x + : : : + dit□1xt□1 This satisfies fi(0) = 0. Server i then sends the value sij= fi(j) mod q to server j, which updates its new share s(new) J as follows :

$$s_j^{(new)} = s_j^{(old)} + s_{1j} + \ldots + s_{nj} \text{ mod q}$$

Since the new shares lie on the polynomial

f(new)(x) =f(old)(x) + f1(x) + ::: + fn(x)

The new polynomial still maintains the secret value S at x = 0 , because f(new)(0) = S + 0 + ::: + 0 = S.

Observations:

•In both the [Initialization] and [Share Updation] phases, the dealer and the participants generate polynomials satisfying constraints f(0) = S and fi(0) = 0, 1 _ i _ n.

• Dealer control and coordinates all the activities like generation, distribution and updation of shares, the scheme is heavily dependent upon dealer.

B. Definition

A share renewal protocol is secure if following properties hold:

•Privacy : No information about the secret S is revealed.

•Robust : Correct re- constructibility is possible at any time when k shares are present.

III. PROPOSED SHARE RENEWAL PROTOCOL WITHOUT TRUSTED PARTY AGAINST PASSIVE ADVERSARY

A. Overview of the Scheme

In this section we construct our own share renewal scheme without the presence of a Dealer. Every participant Pi chooses a random value si and distributes it among all n participants the secret S is a function as

$$S = \sum_{i=1}^{n} s_i \delta_L^i, \text{ where n is the set}$$

of participants and _Il is the Lagrange coefficient according to L. Then each shareholder uses its old share values to obtain new shares without changing the secret S. The difference from previous schemes is that no participant uses the Shamir's scheme to distribute

Cite this article as: Sayantan Mandal, V. Ch. Venkaiah. "A Novel Proactive Secret Sharing." *International Conference on Information Engineering, Management and Security (2015)*:260-265. Print.

share. Each participant uses its old share to create new shares without disclosing his original share. The scheme consists of three phases. We have basically divided our proposed scheme into three
phases:

- •Initialization phase
- •Distribution phase
- •Updation phase

Each phase of our scheme is discussed below in detail. It may be noted that the Dealer is not involved in any phase of the proposed scheme. All the computation such as secret generation, share distribution and share renewal are done by the shareholders without revealing any information about the secret or their share. In addition to that its implementation is less complicated compared to other existing schemes.

1) [Initialization]: In this phase each participant creates share value to distribute shares to other participants.

Algorithm 1 Initialization

1: Let P_1,P_2 $\ldots P_n$ be the n shareholders.

2: Each P_i chooses a random polynomial $f_i(x) = \sum_{k=0}^{t-1} a_{ik}z^k$ of degree t-1 where a_{ik},$1 \leq i \leq n$, are random elements from the field F_q.

3: Each participant P_i sends $f_i(j)$ to each participant P_j, $1 \leq i, j \leq n$.

4: Each P_i creates a set A_i=$\{A_{ik}\}$ where A_{ik} = $g^{a_{ik}}$ mod q, where g is an element of the finite field F_q and sends it to each participant P_j, $1 \leq j \leq n$. Once P_i,$1 \leq i \leq n$ receives messages from other participants it verifies each share.

5: Upon verifying the shares, the share received from each participants are added to obtain a final share.

6: That is, the share of the j^{th} participant is $\sum_{i=1}^{n} f_i(j) mod\ q$=$P_j(s_j)$.

2) [Distribution]: In this section distribution of share takes place among the participants.

Algorithm 2 Distribution

1: Each i^{th} shareholder $i \in [1 \ldots$n] randomly splits its current share $P_i(S_i)$ obtained after initialization phase, into sum of two shares such that $P_i(S_i) = s_1 + s_2$.

2: Each participant P_i now secretly send its second part of the share ($S = s_1 + s_2$), that is s_2 to P_{i+1} Now at this point each P_i knows its own share i.e $(P_i(s_1), P_i(s_2))$ and share received from P_{i-1}, i.e $P_{i-1}(s_2)$.

3: Each P_i now takes its first part of the share, $P_i(s_1)$ adds it to the share received from P_{i-1} and passes the sum secretly to P_{i+1}.

4: Taking the share received in step 2 i.e $P_{i-1}(s_2)$ and adding it to the computed value received from P_{i+1} in step 3.
Note: It may be observed that neither P_i nor P_{i+1}, $1 \leq i \leq n-1$ can figure out the share of the other.

5: Each P_i now adds its own share value $P_i(S_i)$to the computed value received at step 4 and passes this value to P_{i+1}.

6: The above steps will leave each P_i with a good mix of sum of shares from other participants, these are the coefficients for polynomial of degree $t-1$. Shareholder among themselves randomly pick $t-1$ coefficient values from n participants.

3) [Update]: Now we have the new coefficient values required to create a new polynomial to update the old polynomial without changing the secret.

Algorithm 3 Update

1: We now form a polynomial $h(i)$ of degree $t-1$ whose free coefficient is zero $(P_i(0)=0)$ and whose coefficient values are the one calculated from above steps. Substituting 0 in old share - $f(i)$ to the sum of the new n shares. Mathematically speaking: $h(i) = f(i) + \sum_{c=1}^{n} P_c(i)$

B. Example

[Initialization Phase]

We hereby give an example of working of the Protocols

• Let n = 4, p = 11

•Each participant chooses a random polynomial of degree t □ 1 with fi(0) = si, where si is a random value chosen by each participant.

•Four participant A,B,C and D chooses the following polynomial and create shares.

fa(x) = 5 + 7x + 2x2 mod 11

fb(x) = 4 + 2x + 5x2 mod 11

fc(x) = 10 + 7x + 2x2 mod 11

fd(x) = 2 + 3x + 2x2 mod 11

• Each participant now create n shares and distribute to each participant Pj , 1 _ j _ n

•Using the Lagrange0s interpolation the shares created are:

•participant A creates f(1) = 3 , f(2) = 5 ,f(3) = 0, f(4) = 10

•participant B creates f(1) = 0 , f(2) = 6 ,f(3) = 0 , f(4) = 4

• participant C creates f(1) = 8 , f(2) = 10 ,f(3) = 5, f(4) = 4

• participant D creates f(1) = 7 , f(2) = 5 ,f(3) = 7, f(4) = 2

• Each participant Pi sends _ij = fi(j) to each participant Pj , 1 _ i; j _ n, where _ij is the corresponding share w.r.t to the participant .

[Verifying shares using VSS]

Before proceeding to the share renewal protocol. each participant can verify whether the shares given to him by the other participant is correct or not.

From the above example, polynomial f(x) generates four shares which are distributed among four participants A,B,C and D respectively. Now say A wants to verify his share received from C then it initiates the following steps: Verification of shares : Let the encrypted values of the coefficients of the

polynomial be :

E(a0) = g10mod 11 , E(a1) = g7mod 11 , E(a2) = g2mod 11, where g is an element of the finite field Fp: Using Feldmen verifiable secret sharing scheme, we have for the first shareholder (i = 1): E(f(1)) = g8mod 11 is equal to:

E(f(i)) = E(a0)(E(a1)E(i1))_: : :_(E(at□1)E(it□1))

= E(a0 _ (a1 _ (11)) + (a2 _ (12)))

= g10+7+2 mod 11

= g8 (share verified for i = 1)

Similarly other shareholders can also verify their share values. Once share are verified we proceed to share renewal scheme. Computing Final Shares After verification of shares the share received by each Pi from other Pj are added to form final shares.

•A computes (3 + 0 + 8 + 7) mod 11 = 7

• B computes (5 + 6 + 10 + 5) mod 11 = 4

• C computes (0 + 0 + 5 + 7) mod 11 = 1

• D computes (10 + 4 + 4 + 2) mod 11 = 9

Note: As said above secret

$$S = \sum_{i=1}^{n} s_i \delta_L^i,$$

of participants and S^i_L is the Lagrange coefficient according to L. Thus each Pi now sends its si also t - 1 coefficients are randomly picked from above computed shares and a new polynomial is formed.

• Adding si gives 5 + 4 + 10 + 2 = 21 mod 11 = 10

• New polynomial becomes f(x) = 10 + 4x + 9x2

Cite this article as: Sayantan Mandal, V. Ch. Venkaiah. “A Novel Proactive Secret Sharing.” *International Conference on Information Engineering, Management and Security (2015)*:260-265. Print.

[Distribution]
Let the random partition of the shares A,B,C and D be
•A = 4 + 3 (A1 + A2)
•B = 2 + 2 (B1 + B2)
•C = 0 + 1 (C1 + C2)
• D = 6 + 3 (D1 + D2)

Step 2 Secretly, After exchanging second parts of their shares with the other participants we have the following scenario
•A knows A1;A2;D2
B knows B1;B2;A2
C knows C1;C2;B2
D knows D1;D2;C2
Step 3 After adding the first part of the share with the second part of the received share from the other participant, we have
• A calculates A1 + D2 and gives to B
B calculates B1 + A2 and gives to C
C calculates C1 + B2 and gives to D
D calculates D1 + C2 and gives to A
That is
A computes 4 + 3 = 7 and gives to B
B computes 2 + 3 = 5 and gives to C
C computes 0 + 2 = 2 and gives to D
D computes 6 + 3 = 9 and gives to A
Step 4 Now each participant takes the share value received in step 2 and add it to value received in step 3. That is,
• A adds D2 to (D1+C2) getting (D2+D1+C2) =
(D+C2), since A does not know C2, he cannot derive
D.
• Similarly B adds A2 to (A1+D2) getting (A+D2)
• Similarly C adds B2 to (B1+A2) getting (B +A2)
• Finally D adds C2 to (C1 + B2) getting (C + B2)
That is
A computes 9 + 1 = 10
B computes 7 + 3 = 10
C computes 4 + 3 = 7
D computes 1 + 2 = 3
Step 5 The participant then adds their own share to the sum arrived in the above step and passes the sum to the next participant.
That is
•A adds A to (D + C2) getting (A + D + C2) and passes to B.
• B adds B to (A + D2) getting (B + A + D2) and passes to C.
• C adds C to (B + A2) getting (C + B + A2) and passes to D.
• D adds D to (C + B2) getting (D + C + B2) and passes to A.
That is,
A computes 7 + 10 = 17 passes to B
B computes 4 + 10 = 14 passes to C
C computes 1 + 7 = 8 passes to D
D computes 9 + 3 = 12 passes to A

Privacy check : At this stage of Algorithm
• A knows A1, A2, D2, (D1 + C2), (D + C + B2), A knows D2 however he cannot derive D, since D1 and C2 are not known individually to A.
• B knows B1, B2, A2, (A1 + D2),(A + D + C2), B knows A2, however he cannot derive A, since A1 and D2 are not known individually to B.
• C knows C1, C2, B2, (B1 + A2), (A + B + D2) C knows C2, however he cannot derive B, since B1 and A2 are not known individually to C.
• D knows D1, D2, C2, (C1 + B2), (C + B + A2) D knows C2, however he cannot derive C, since C1 and B2 are not known individually to D.
C. Update
•With the arrival of new shares old shares are erased and new shares are kept. Thus new shares are
A:12 , B:17 , C:14 , D:8
• Above steps generate coefficients for new polynomial, Now we want to form polynomial of degree two, any two coefficients. Old polynomial was

f(x) = 10 + 4x + 9x2

f(0) = 10 + 4(0) + 9(02) = 10

•Now we form new polynomial from above coefficient

g(x) = 0 + 12x + 14x2

• we add above two polynomials

h(x) = f(0) + g(x)

h(x) = 10 + 12x + 14x2

• Thus h(x) is our new polynomial, we can repeat the above Distribution Algorithm with new shares to renew the shares.

IV. SECURITY CHECK

As per the Definition we can call a share renewal protocol secure if it satisfies properties :Privacy, Robustness.

A. Privacy Check

As seen in the Distribution phase of proposed share renewal protocol at no stage any shareholder gets any information about share value of any other shareholder. As each shareholder randomly splits its share value into sum of two share and throughout the Distribution phase only part of share value is used to communicate among shareholder.Thus our proposed protocol complies with privacy constraint.

B. Robustness Check

The new shares computed at the end of Update phase corresponds to secret S.The algorithm allows to re-construct the original secret whenever t - 1 shares are present.In other words any subset of t - 1 share will give us secret S.

1) Correctness: Let K be set of $k-1$ shares after k-th update phase. Let K= $\{ y_1^t, y_2^t \dots y_{k+1}^t \}$ also assume $a_1, a_2 \dots a_{k+1}$ as the coefficients of polynomial such that $\sum_{i=1}^{k+1} a_i y_i^t$ would give us the secret using Shamir 's scheme.

$$\begin{aligned}
\textstyle\sum_i^{k+1} a_i y_i^t &= \textstyle\sum_i^{k+1} a_i \left\{ y_i^{t-1} + \sum_{j=1}^{n} \delta_j(i) \right\} \\
&= \textstyle\sum_{i=1}^{k+1} a_i y_i^{t-1} + \sum_{j=1}^{n} \sum_{i=1}^{k+1} a_i \delta_j(i) \\
&= x + \textstyle\sum_{j=1}^{n} \delta_j(0) \text{ (by interpolation)} \\
&= x \qquad\qquad (\forall j, \delta_j(0) = 0)
\end{aligned}$$

V. CONCLUSION

In this paper, we have designed a share renewal protocol where shares can be renewed without a trusted party against passive attacker.The scheme is both robust and maintain privacy. shareholder can verify their share before beginning of the share renewal scheme.We assume secure channel exist among the shareholder to exchange share value.The correctness of the properties of the scheme is also discussed The proposed scheme is secure and practical.

REFERENCES

[1] Dahlia Malkhi,:An advance course in computer and network security.The Hebrew University of Jerusalem, Lecture Notes

[2] Blakley, G. R.: Safeguarding cryptographic keys. In: Proc. of AFIPS National Computer Conference. vol. 48, pp. 313317 (1979)

[3] Cachin, C., Kursawe, K., Lysyanskaya, A., Strobl, R: Asynchronous verifiable secret sharing and proactive cryptosystem. In: Proc. 9th ACM Conference on Computer and Communications Security, pp. 8897 (2002)

[4] A. Shamir, How to share a secret, Communications of the ACM 22 (11) (1979) 612613.

[5] B. Chor, S. Goldwasser, S. Micali, B. Awerbuch, Verifiable secret sharing and achieving simultaneity in the presence of faults, Proceedings of the 26th IEEE Symposium on Foundations of Computer Science, 2123 October, Oregon, Portland, IEEE Computer Society, 1985, pp. 383395.

[6] Herzberg, A., Jarecki, S., Krawczyk, H., Yung, M.: Proactive secret sharing or: How to cope with perpetual leakage. In: Proceedings of CRYPTO 95. LNCS, vol. 963, pp. 339352. Springer, Heidelberg (1995)

[7] Zhou, L., Schneider, F. B., Ranesse, R. V.: APSS: Proactive secret sharing in asynchronous systems. ACM Transaction on Information and System Security 8 (3), 259286 (2005)

[8] Choudhury, A., Patra, A.: Brief announcement: efficient optimally resilient statistical AVSS and its applications. In: Proceedings of PODC 12, pp. 103104 (2012)

[9] Pedersen, T.: Non-interactive and information-theoretic secure verifiable secret sharing. In: Proceedings of CRYPTO 91. LNCS, vol. 576, pp. 129140. Springer, Heidelberg (1992)

Cite this article as: Sayantan Mandal, V. Ch. Venkaiah. “A Novel Proactive Secret Sharing.” *International Conference on Information Engineering, Management and Security (2015)*:260-265. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July -2015		**Accepted**	31- July- 2015
Article ID	ICIEMS045		**eAID**	ICIEMS.2015.045

ADVANCED ATTACK AGAINST WIRELESS NETWORKS WEP, WPA/WPA2-PERSONAL AND WPA/WPA2-ENTERPRISE

MuthuPavithran. S[1], Pavithran. S[2],
[1,2] Department of computer science and engineering
Velammal Engineering College, Chennai, India

ABSTRACT: *In the emerging world of electronics the wireless devices are used by millions of people in their day to day life. Every person is constantly in contact with the cyberspace. Thus, ensuring the proper encryption facility is a major undertaking to offer dependable communication. The aim of this paper is to transmit a wireless penetration test and compares the encrypted key of a wireless network with a file that contains the captured packets as alphanumeric letters with the help of Kali Linux. This paper shows penetration tests in WEP and WPA/WPA2 protocols, and also the methods to develop these protocols using various attacks and to supply tools that separate the vulnerable access point protocol for the web administrators to protect their networks.*

I. INTRODUCTION

As the technologies and, along with it, the threats facing the wireless communications have risen in numbers with the rapidly increasing number of deployments, there is a need for protection. Nonetheless, the risk if often surpassed by the benefits and convenience of wireless technologies, which have been a big component in the scatter of these devices within homes, agencies and enterprises spanning the world. The popularity of wireless technologies has created an acute involvement in other popular wireless protocols such as Wi-Fi interest. Wi-Fi has been manifesting itself to attack, research and vulnerabilities within the protocols and the execution of those protocols in devices. With this growth in wireless technologies, these nets have become increasingly attractive to Hackers who will seek for data to steal or compromise functionality. While the traditional security measures are less efficient the wireless attack surface presents a singular and difficult challenge. Most of the wireless nets are much unprotected so it is vulnerable to assault. When we consider Wi-Fi most of the people have consciousness about two major encryption techniques (WEP) Wired Equivalency Protocol and (WPA) Wi-Fi Protected Access which were frequently employed. WPA is modern and securing when compared to WEP.

LITERATURE SURVEY

[1].Test and confirm the plausibility of WEP attack in a university wireless LAN ,also suggests some mitigation techniques.[2].Analyzes wireless protocol enhancements to existing handshake mechanism in WPA by using Eliptic Curve Cryptography.[3]. Analyzes functional intrusion detection system that combines them in order to offer resilient detection of the most common attacks in 802.11

Cite this article as: MuthuPavithran S, Pavithran S. "ADVANCED ATTACK AGAINST WIRELESS NETWORKS WEP, WPA/WPA2-PERSONAL AND WPA/WPA2-ENTERPRISE." *International Conference on Information Engineering, Management and Security (2015)*: 266-271. Print.

networks.[4].Explains WEP and RC4 used in WEP.FMS attacks and PTW attacks are described.[5].How to secure our wireless world and gives steps to take care by the user for not affected by the attacker.[6].Describes 2 attacks on IEEE 802.11 WEP,WPA.Deals with TKIP to encrypt traffic. How to deal with ARP request and response and to send with custom network.[7].Effective security protocols right from evolution to existing scenario and discusses various pros and cond of security protocols in WLAN with respect to its countermeasure techniques on various attacks.[8].Challenges and solutions for emergent security technologies ,WIFI.[9].Solution for WPA2 shortcomings and thus provide protection to wireless networks from several attacks.[10].Suggetions on wifi protected access2(WPA2)protocol vulnarabilities might be mitigated and addressed through the enhancement new protocols.

II. BASIC ENCRYPTION

A. Enable WEP encryption settings

Wired Equivalent Privacy (WEP) is an IEEE 802.11 wireless protocol, which provides the security algorithm for the data during the wireless transmission. WEP uses a 24 bit initialization vector (IV) to form the stream chipper and the CRC -32 checksum for integrity of the wireless transmission. 64 bit WEP uses a 40 bit key, 128 bit WEP uses the 104 bit key, 256 bit WEP uses 232 bit key size.

B. WPA/WPA2 Encryption

WiFi Protected Access (WPA) is a data encoding method for WLAN based on the 802.11 standard. WPA was developed after WEP to provide a stronger encryption by configuring two different ways –pre-shared key mode and enterprise mode. The TKIP (Temporal Key Integrity Protocol) is applied to migrate vulnerability by increasing the size of the IV and by using the mixed use. In WPA and WPA2 encryption keys (TK) are derived during the four way handshake. It serves to go through the sequence counter for security against the reply attack. Temporal Keys are changed for every 10,000 packets and this makes TKIP protected network more resistant to cryptanalytic attacks involving the key reuse.

C. WPA/WPA2 Enterprise

Whenever a user connects it dynamically creates the PMK every time in the WPA based network enterprise mode. The PMK is generated by the authentication server and then transmitted down to the client. The AP and the server address over a protocol called RADIUS. The determination to admit or reject the user can be served by the host. Since the AP acts as a relay to forward the packets from clients that are for authentication purposes.

III. SCANNING

As we categorize the tools into passive and active tools.

A. Active scanning

Probe request packets are periodically sent by the tools that carry out the active scanning. These packages are used by the clients whenever they are awaiting for a network. I.e. The client may post the broadcast probe requests. Beacon packets are charged by the Access Points every tenth of a second. Beacon packets are sometimes accessed by active scanners.

B. Passive scans

Passive scanning is too known as Monitor Mode. They listen to all bundles on a gifted channel and then study them.

IV. SNIFFING AND CRACKING TOOL

Aircrack-ng is developed by Christophe Devine, which causes a packet sniffer, packet injector, WEP and WPA/WPA2-PSK cracker and analyzer for 802.11 wirelesses LAN and it will go with whatever wireless network interface controller which supports raw monitoring and sniff 802.11a, 802.11b, 802.11g.

Aircrack-ng

It uses WEP and WPA/WPA2-PSK cracking tools.

Aireplay-ng

It is employed for traffic generation, fake authentication, packet replay, and ARP request injection.

Airodump-ng

It is applied to capture packet of a raw 802.11 builds.

V. ATTACKING WEP-PROTECTED 802.11 NETWORK

Before we attack, we demand to recognize the Mac filtering most of AP`s allow you to set up list of trusted MAC addresses. Any packets transmitted from other IPs will get cut. MAC addresses are very static things, fired into the chip and are immutable. We can simply steal Mac from a person who is already on the web. To answer this we need to hunt down a passive scanner on the network it will that will give the list of addresses whose are connected to that network (CLIENT). We need to wait for the user to disconnect from the net because we can tie to that network with his destination. There is some other direction to make this it's called "DOSING". E.g.: ifconfig wlan0 HW ether 00:11:22:33:44:55.

A. Dictionary attack against WEP

A dictionary attack on wep involves feeding a cracking utility a dictionary and pcap file. The instrument serves to check words in dictionary with words in the backup file, it won't come out until it's found or words are blended.It's a stranded way to translate a password into WEP keys. We need to be given at least three different algorithms (NEESUS, DATACOM, MD5, APPLE).

B. Cryptographic attack against WEP

Cite this article as: MuthuPavithran S, Pavithran S. "ADVANCED ATTACK AGAINST WIRELESS NETWORKS WEP, WPA/WPA2-PERSONAL AND WPA/WPA2-ENTERPRISE." *International Conference on Information Engineering, Management and Security (2015)*: 266-271. Print.

This attempt is present even if the WEP key completely random. Rc4 is the stream cipher used in WEP and then it makes it used in the WEP it make a perfect prey for this vulnerability the main trouble is how WEP uses the initiation vectors (IV) in each WEP packet. When the WEP encrypt the packet, it prepends the IV to secret key before feeding the key into rc4 this shows that the first three bytes is carrying the secret key used in the every packet.

C. Break WEP when the client is bound

Put our card into monitor mode

#airmon-ng start wlan1

Then we need to start the airodump on the specific channel and in specific BSSID to capture the packets and stored in a file.

#airodump-ng –channel –bssid –write filename wlan1

In our target access point there is a client is attached we need to use its MAC address to inject the ARP packet to generate more traffic and besides we can get more packet to crack and besides in a faster manner.

#aireplay-ng –arpreplay –h –client address –b –access point address wlan1

Move to the airodump window we can realize that the information package will get increased like a skyrocket. We need more than 40,000 packets we can begin cracking the key of 104 bit WEP key. The 40,000 packet have the 50% of chance to breach the key the more packets it will increase the probability of finding the key. Then fire the aircrack-ng to crack.

#aircrack-ng. /captured file.cap -0

D. Break WEP without client attached

.First step: we need to capture the entire packet from the access point so we are using the airodump-ng tool helps to capture packet by selecting particular Mac address and its channel and with the network interface and it`s saved in the file called the pcap file.

#airodump-ng –channel –bssid –write file wlan1

Second step: now we are starting to do fake authentication attack which leaves us to associate to its target access point and utilize either two types of authentication open and shared key which will help to produce a fake an authentication to the AP for in order to communicate with the AP

#aireplay-ng –fakeauth 0 –o 1 –e ESSID –a –accesspoint –h –attackerid wlan1

It's gone bad due to Mac is filtering use Mac spoofing method

Third step: Now we are proceeding to perform the fragmentation attack which is the most advanced cracking method. It is employed for the retrieval of key stream from the data packet. It can turn the few bytes of key stream into more or equal 1500bytes of key stream in a few moments. It helps in attack by multiplying an attacker's key stream by the factor up to 16 on each round. The common initial key stream sources are SNAP header. The first three bytes of a SNAP header are 0xAA, 0xAA; 0x03.It is used to make the three bytes of key stream which is enough to take up the fragmentation attack. Then XOR the first three bytes of a SNAP header with first three bytes of captured packet, it will result in three bytes of key stream, then craft an ARP broadcast packet break this packet into 12 three-byte fragments then encrypt and the beam. Each fragment can reuse the same three bytes of key stream. After transmitting then look for the 36 Byte packet that is sent by an AP. This is the ARP packets relayed from the AP. When you have crafted the package in the first place you know the 36-bytes of spare text. Then XOR the encrypted packet with the plaintext, now you recovered the 36 bytes of key stream, then try to craft long ARP packet you can also padded NULLs while crafting. Until you get to full bytes of key stream

Aireplay-ng –fragment –b –access point –h –attacker wlan1.

Fourth step: now we are proceeding to do the chop chop attack, it's a modifying an encrypted packet one byte at a time and played back to an AP. If it has a modified packet chop chop can slowly decrypt the packet, it is protected by WEP regardless of key size as said before, even an AP will generate some packets when no node is attached then remove the final act from what we captured from AP. Then adjust the checksum by assuming the byte is 1. Retransmit it towards a multicast address. If the AP relay the packet then we assumed checksum was correct, then you guessed plain text value was correct now we have recovered the one byte of plain text and key stream. If the AP dose not relay the packet, then our hypothesis was incorrect, so try to keep on guessing for 256 attacks. At the conclusion of the attack, we have the patent text and key stream.

Aireplay-ng –chopchop –b access point –h attacker wlan1

Fifth step: we are going to craft the ARP packet we need for the output of the any one of the attack chop chop or fragmentation. By injecting particular, ARP packet that will cause the AP to generate more traffic at present we are starting to get the ARP packet.

#packet forge-ng –arp –a access point –h attacker –k 255.255.255.255.255 –l 255.255.255.255 –y fileof.xor –w file

Most of the network will accept the ARP packet crafting. If it is fails check the output of the chop chop attack of the plain text and tailor value to the Subnet then The resultant will be encrypted using the key stream and IV in the.Cursor file

Sixth step: now we are starting to inject the crafted ARP packet that is replaying the encrypted ARP packet what we are set up with the assistance of the aireplay-ng after injecting we can view the information packet is increasing in the airodump-ng

Aireplay-ng –interactive –f –r./forged_arp wlan1

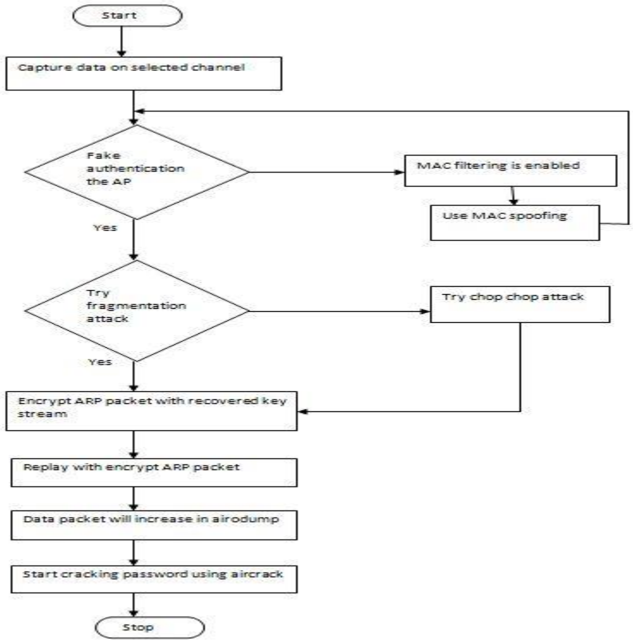
Seventh step: Now we are going to crack the key from the pcap file by passing the argument to the aircrack-ng

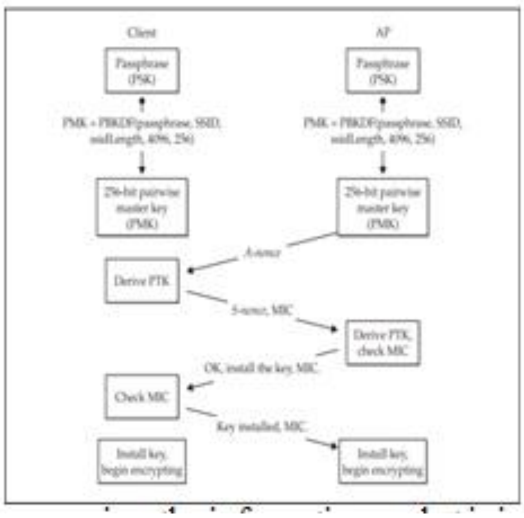
Aircrack-ng./File. cap -0

A. How to defend WEP attack

The best room to defend from this attack you needs to switch to WPA2 with CCMP.

Cite this article as: MuthuPavithran S, Pavithran S. "ADVANCED ATTACK AGAINST WIRELESS NETWORKS WEP, WPA/WPA2-PERSONAL AND WPA/WPA2-ENTERPRISE." *International Conference on Information Engineering, Management and Security (2015)*: 266-271. Print.





VI. ATTACKING WPA-PROTEDTED 802.11 NETWORKS

As we seen before WPA/WPA2 improve our wireless network protection, notwithstanding the extra protection also comes with a cost. Even though WPA was developed with high complexity, it holds its own flaws that are starting to take advantage, attacking the authentication that gives direct access to the wireless web. When attacking WPA-PSK authentication, the attacker can also hold the power to decrypt traffic since the PMK is recovered. Encryption attacks are just emerging against WPA networks. These approaches provide the power to decrypt traffic but do not permit the attacker to fully join the mesh as a lawful user.

A. Breaking authentication WPA-PSK (pre-shared key)

WPA-pre-shared key is also called as WPA- personal. This method is a shared secret among all devices on the network for authentication. The four way handshake that allows the client and the access point to negotiate the key used to encrypt the traffic sent over the wireless. We are starting to crack the key, and this handshake will be executed while a client getting try to connect to the any specific access level. From the diagram, we see that the access level (AP) first sends the A-nonce and, the S-nonce sent by the customer.And so the client MAC address and the AP`s MAC address are mailed, and MIC to verify with the exclusion of the SSID.These values can start in the four way handshake. Can see through wireshark. Sometime they are replicated across the chassis.

B. Passive sniffing

Cite this article as: MuthuPavithran S, Pavithran S. "ADVANCED ATTACK AGAINST WIRELESS NETWORKS WEP, WPA/WPA2-PERSONAL AND WPA/WPA2-ENTERPRISE." *International Conference on Information Engineering, Management and Security (2015)*: 266-271. Print.

Equally we have already witnessed in the scanning we can seize the four way hands shake in the passive sniffing (scanning). This four way handshake will occur if we are in the right groove and at the proper time. Equally we have already known about the airodump-ng of the aircrack - ng suite, it is the lightweight sniffer. Before that we must match our card is in monitor mode & locked onto a particular channel, and we are saving sniffer data into a file, we`ll stay board by just targeting a single groove.
#airmong-ng start wlan0
#airodump-ng –channel –write –bssid wlan0
These commands will put our card into monitor mode that will lock our card onto the channel. The AP is transmitted, and the transmitted data will lock into the file, remember that.

C. Active sniffing

We have more serious things to do than wait around for new user to join. Alternatively we can sound off the user off and then follow him to reconnect. To kick off the user we use the de-authentication attack for that we need aireplay-ng. We catch the four way handshake.

D. Cracking pre shared key

#aircrack-ng –w wordlist.txt –some.cap
Cracking WPA-PSK can be done by offline brute-force attack. It is challenging the character position for the pre-shared key and can be between 8 and 63 printable ASCII characters and the chosen passphrase is hashed 4096 times before using it within the PMK.

Even though aircrack-ng was a powerful it causes its own limitations, so to improvise the methods we use capacity that needs the limited number of frames than aircrack-ng to crack the key in offline. The great problem because WPA-PSK the PMK of this is not just hashed of the pre shared key, but also the SSID. This implies that even the different network can receive the same pre shared key, but the PMK will differ. Thither is a way to create they our own hash table, by using the genpmk
#genpmk –f wordlist –d wordlist.genpmk –s <BSSID>
Now we are going to increase the speed of the cracking the password. By utilizing the customized field-programmable gate array, it's applied to perform the simple logical operation at incredible velocity. More incredible speed was achieved by using the improvised graphical process units (GPUS) it is merely called as video card which handles the graphic version. Only the help of NVDIAs CUDA (computed unified device architecture) and the c developer can offload the tasks to the video card to leverage its GPU for password checking.

E. Decrypting WPA-PSK captures

After we crack the key, we can able to read other users packets. But there is problem that every user has a unique pair wise transient key (ptk) that was generated when they associated with the network. Even though we have the PMK, we can decode the packets sent and receive from the user by using the wire shark tool have built in methods to decrypt the packets we need that not only the PMK and SSID of the target user. Because of the most of the encryption is mixed up with the SSID.

F. How to Defend WPA/WPA2 personal attack

Most of the home based Wi-Fi networks attacks are increased because the users are using their mobile numbers, birthdates or some favorite names and soon which can be easily guessed by the hackers. Thus, we necessitate to use maximum number of alpha-numeric-special character and we demand to switch it every week. The word should not connect to your personal. Everyone must bear their own updated firewall.

VII. BREAKING AUTHENTICATION: WPA ENTERPRISE

This case of authentication will be practiced in most of the establishments. Because it offers better security and also economical. WPA enterprise supports a kind of authentication schemes with the usage of EAP (extensible authentication protocol). Some of the authentication schemes are considered more safe.

A. LEAP (Lightweight Extensive Authentication Protocol):

It is one of the Cisco`s proprietary. The EAP types are established on the MS-CHAPv2 it's a challenge-response protocol. The customer gets connected to the network, sending its username and the authentication server return the 8-byte challenge. The client works out the NT hash of the password and uses them as seed material to encrypt the challenges using DES. When the server delivers the same computation and verifies the solution. LEAP is seemed to be a decent protocol. Its major downfall is the challenge and responses are communicated in the open. If we can sniff a user authenticating, we can set up an offline approach to obtain the user`s password.

A. PEAP and EAP-TTLS:

Protected EAP and EAP-TTLS tunneled transport layer security. They provide the best authentication by establishing a TLS tunnel between client and the authentication server, then passing their information within the tunnels using less secure inner authentication protocol. This case of authentication protocol worked in the networks so sniffing on this mesh is less viable. They are protected from the eavesdropping attack. This tunnel additionally provides the customer to secure the authentication server identity by TLS certificate via trusted certificate authority.

B. Attacking PEAP and EAP-TTLS

If we start attacking against the tunnel we won't win because the tunnel is extremely dependable, but if we found vulnerability in that tunnel we can proceed some sort of attack, but mostly we can't able to get results if we discover the vulnerability in the implementation or misconfiguration in the certificate validation on the customer side by skipping the validation. Most of the admin will not notice this shape and then now its vulnerability to access point impersonation attacks and human being in the middle attack. To launch, we need an access point with the same SSID of the target network with better signal we ensure that the net must be same as the objective web.It must attract the guest to connect this network before that we need the radius server to reply to the customer request.There is an open source server free RADIUS that will have any inner authentication protocol sent by a client and respond to it. To start RADIUS we use #radiusd. At once the server has started it works in the background when the client connects inner authentication protocol, which will save every request and responds, even the username and password in the log files. If the client used some other authentication like MSCHAP we need to use sleep to crack it and we can plug in to their net.

C. EAP-TLS

This method was very safe, because it uses client and host credentials to authenticate the user along the web. Thither is a large problem with the managing the certificates for all other users in the system because of computer storage management. Most of the company doesn't suffer the level of PKI required. The working of EAP-TLS uses the server that sends the client certificate which is verified and the public key is applied to encrypt the further message. At once the customer transmits the authentication server certificate, which is verified by the host. After the verification the client and server will generate the random key. This is utilized to initialize the symmetric cipher to encrypt the data from the TLS session. On the EAP success message the PMK is transmitted from the RADIUS server to the AP.

D. Attacking EAP-TLS

It is pretty much impossible to attack this EAP-TLS protocol; we can't say that EAP-TLS have flaws. Till today at that place is no weakness are found because this protocol is very robust as we require to defeat in a practical way by stealing the private Key. Most of the pin or key stored in the smart cards or it uses the RSA secure ID token. The largest trouble is if you gained the access to the network you can't decrypt anyone else`s traffic because it use the unique PMK.

E. How to defend WPA/WPA2 enterprise attack

Every day new vulnerability has been establish in every technology because it was contrived by humankind. And at that place is the way to protect them from those injuries. And then every system should update their technology day to day by their administrator. It is possible only if the administrator updates his knowledge every day. The most significant thing is during the design and conformation of the organization's networks some admin misses some configuration that makes the major fault in their net. The users of the network also receive knowledge of security they should not miss use it.

VIII. CONCLUSION

In this paper, we proposed several advanced attacks against wireless network These varieties of approaches will pass off in most of the wireless nets that will have many losses like money, important data or whole network may be compromised.Then we must aware of these approaches with the supporter of network administrator of every system. In future these kind of wireless attacks plays a major function in top cyber complains in nation because every day wireless technology is emerging tremendously. Therefore each and every individual must aware of these varieties of attacks to have a secured wireless communication.

IX. ACKNOWLEDGMENT

For the help and encouragement, we would wish to thank many of our college professors especially Dr. A. Balaji Ganesh, professor, department of electrical and electronics at Velammal engineering college, Dr. V. Vijaya Chamundeeswari, head of the computer science at Velammal college of engineering, Mr. R. Prakash, research associate department of electronics instrumentation at Velammal engineering college, Mr. Siva V Girish, research associate department of electronics instrumentation at Velammal engineering college, Mrs.R. Ramyadevi, assistant professor, department of computer science at velammal engineering college and Mrs. S Deepa. Class adviser and assistant professor, department of computer science at velammal engineering college.

REFERENCE

[1] WarrenHarrop, GrenvilleArmitage, Transparent IP Layer Interception on Enterprise 802.11 b/g networks.
[2] Kshitiz Saxena, The Analyses of Wireless Encryption to handshake mechanism in WPA-NITTTR, 2010.
[3] Guzman Santaf'e-An anomaly based Intrusion detection system for IEEE 802.11 networks, S2ISec information security labs, 2006.
[4] LIU Yong-lie, JIN Zhigang, Research on attack to WEP Protocol, Nov-2010.
[5] 5.Popuri Manoj Kumar,Attacks and Mitigations over wireless network networks-Insecure Wireless world,The International Journal of Engineering & Sciences.-2013.
[6] 6.Martin Beck,Practical attacks against WEP & WPA ,Erik Tews TU-Darmstadt-Nov,2008.
[7] 7.Latha.P.H,Vasantha.R-Review of existing security protocols techniques and their performance analysis in WAN-IJETCAS-2014.
[8] 8.Paul S.Henry and Hui Luo,AT&T Labs-Research wifi:what is next? December 2002.
[9] 9.R.N.Rajotiya,PridhiArora,-Enhancing security of wifi network,The International Journal of Computer Application,June 2012.
[10] 10.Paul Arna,Benefits and vulnarabilities of wifi protocol access2 (WPA2)-INFS612-Fall,2006.

Cite this article as: MuthuPavithran S, Pavithran S. "ADVANCED ATTACK AGAINST WIRELESS NETWORKS WEP, WPA/WPA2-PERSONAL AND WPA/WPA2-ENTERPRISE." *International Conference on Information Engineering, Management and Security (2015)*: 266-271. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS046

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.046

DESIGN AND FABRICATION OF PNEUMATIC JACK FOR AUTOMOBILE

S.Sathiyaraj[1], V.Selvakumar[2]
[1,2] FINAL YEAR MECHANICAL ENGINEE RING,
JAYALAKSHMI INSTITUTE OF TECHNOLOGY, THOPPUR, DHARMAPURI

ABSTRACT: *The main target of project is to improve version of a mini pneumatic jack. This will be more efficient for the user. This machine is pneumatic powered which has low co-efficient of friction. A pneumatic cylinder erected provides power to lift up the Jacky. This is a pneumatic powered machine and requires no other means of power to operate. The required components are Compressor, Pneumatic cylinder, Solenoid, Control circuit and Jack.*

INTRODUCTION

NEED FOR AUTOMATION:
Automation can be achieved through computers, hydraulics, pneumatics, robotics, etc., of these sources, pneumatics form an attractive medium for low cost automation. The main advantages of all pneumatic systems are economy and simplicity. Automation plays an important role in mass production. Nowadays almost all the manufacturing processes are being made automatic in order to deliver the products at a faster rate. The following reasons affirms the benefits of automation,

- To achieve mass production
- To reduce man power
- To increase the efficiency of the plant
- To reduce the work load
- To reduce the production cost
- To reduce the production time
- To reduce the material handling
- To reduce the fatigue of workers
- To achieve good product quality
- Less maintenance

Cite this article as: S.SATHIYARAJ, V.SELVAKUMAR. "DESIGN AND FABRICATION OF PNEUMATIC JACK FOR AUTOMOBILE." *International Conference on Information Engineering, Management and Security (2015)*: 272-281. Print.

PNEUMATICS

The word 'pneuma' comes from Greek and means wind. The word pneumatics is the study of air movement and its phenomena is derived from the word pneuma. Today pneumatics is mainly understood to means the application of air as a working medium in industry especially the driving and controlling of machines and equipment.
Pneumatics has for some considerable time between used for carrying out the simplest mechanical tasks in more recent times has Played a more important role in the development of pneumatic technology for automation.
Pneumatic systems operate on a supply of compressed air which must be made available in sufficient quantity and at a pressure to suit the capacity of the system. When the pneumatic system is being adopted for the first time, however it wills indeed the necessary to deal with the question of compressed air supply.The key part of any facility for supply of compressed air is by means using reciprocating compressor. A compressor is a machine that takes in air, gas at a certain pressure and delivered the air at a high pressure.Compressor capacity is the actual quantity of air compressed and delivered and the volume expressed is that of that of the air at intake conditions namely at atmosphere pressure and normal ambient temperature.
The compressibility of the air was first investigated by Robot Boyle in 1962 and that found that the product of pressure and volumes of particular quantity of gas.

The usual written as

$$PV = C \quad \text{(or)}\ P_iV_i = P2V2$$

In this equation the pressure is the absolute pressured which for free is about 14.7Psi and is of courage capable of maintaining a column of mercury, nearly 30 inches high in an ordinary barometer. Any gas can be used in pneumatic system but air is the mostly used system now a days.

SELECTION OF PNEUMATICS

Mechanization is broadly defined as the replacement of manual effort by mechanical power. Pneumatic is an attractive medium for low Cost mechanization particularly for sequential (or) repetitive operations. Many factories and plants already have a compressed air system, which is capable of providing the power (or) energy requirements and control system (although equally pneumatic control systems may be economic and can be advantageously applied to other forms of power).
The main advantages of an all pneumatic system are usually Economic and simplicity the latter reducing maintenance to a low level. It can have out standing advantages in terms of safety.

PNEUMATIC POWER

Pneumatic systems use pressurized gases to transmit and control power. Pneumatic systems typically use air as the fluid medium because air is safe, low cost and readily available.

THE ADVANTAGES OF PNEUMATICS

1. Air used in pneumatic systems can be directly exhausted back
In to the surrounding environment and hence the need of special reservoirs and no-leak system designs are eliminated.
2. Pneumatic systems are simple and economical
3. Control of pneumatic systems is easier

THE DISADVANTAGES OF PNEUMATICS

1. Pneumatic systems exhibit spongy characteristics due to compressibility of air.
2. Pneumatic pressures are quite low due to compressor design limitations(less that 250 psi).
PRODUCTION OF COMPRESSED AIR
Pneumatic systems operate on a supply of compressed air, which must be made available. In sufficient quantity and at a pressure to suit the capacity of the system. When pneumatic system is being adopted for the first time, however it wills indeed the necessary to deal with the question of compressed air supply. The key part of any facility for supply of compressed air is by means using reciprocating compressor. A compressor is a machine that takes in air, gas at a certain pressure and delivered the air at a high pressure. Compressor capacity is the actual quantity of air compressed and delivered and the volume expressed is that of the air At intake conditions namely at atmosphere pressure and normal ambient temperature. Clean condition of the suction air is one of the factors, which decides the life of a compressor. Warm and moist suction air will result increased precipitation of condense from the compressed air.

COMPRESSOR MAY BE CLASSIFIED IN TWO GENERAL TYPES

1. Positive displacement compressor

Cite this article as: S.SATHIYARAJ, V.SELVAKUMAR. "DESIGN AND FABRICATION OF PNEUMATIC JACK FOR AUTOMOBILE." *International Conference on Information Engineering, Management and Security (2015)*: 272-281. Print.

2. Turbo compressor
Positive displacement compressors are most frequently employed for
Compressed air plant and have proved highly successful and supply air for pneumatic control application.
The types of positive compressor
1. Reciprocating type compressor
2. Rotary type compressor

Turbo compressors are employed where large of air required at low discharge pressures. They cannot attain pressure necessary for pneumatic control application unless built in multistage designs and are seldom encountered in pneumatic service.

RECIPROCATING COMPRESSORS

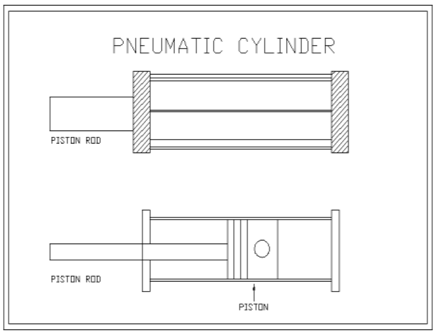
Built for either stationary (or) portable service the reciprocating compressor is by far the most common type. Reciprocating compressors lap be had is sizes from the smallest capacities to deliver more than $500m^3$/min.In single stage compressor, the air pressure may be of 6 bar machines discharge of pressure is up to 15bars.Discharge pressure in the range of 250bars can be obtained with high pressure reciprocating compressors that of three & four stages. Single stage and 1200 stage models are particularly suitable For applications, with preference going to the two stage design as soon as the discharge pressure exceeds 6 bars, because it in capable of matching the performance of single stage machine at lower costs per driving powers in the range.

ULTIMATE AIM

The pneumatic jack can be widely used in low cost automation in manufacturing industries. The weight lifting is quick and effortless, which reduces the physical fatigue (tiredness) felt by the worker.

PNEUMATIC CYLINDER

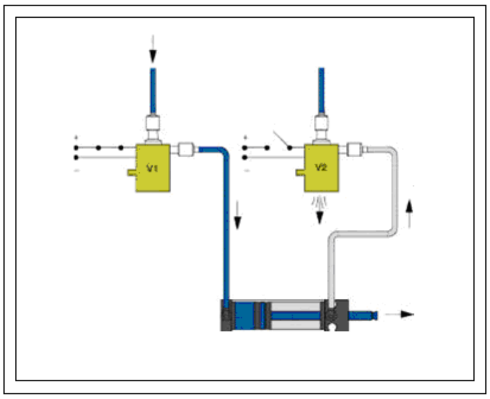
Pneumatic cylinders impart a force by converting the potential energy of compressed gas into kinetic energy. This is achieved by the compressed gas being able to expand, without external energy input, which itself occurs due to the pressure gradient established by the compressed gas being at a greater pressure than the atmospheric pressure. This air expansion forces a piston to move in the desired direction.



Pneumatic cylinders can be moved both inwards and outwards by compressed air. Cylinders of this type are called double-action
Cylinders.

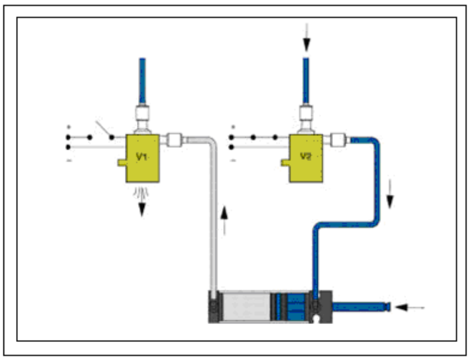
Cylinders also exist which can only be moved pneumatically in one direction. The return movement is caused by a spring. Cylinders of this type are called "single-action cylinders". The compressor cylinder is a single-action cylinder.

In order to move a cylinder in both directions, two of the valves contained in the kit are required.



Cite this article as: S.SATHIYARAJ, V.SELVAKUMAR. "DESIGN AND FABRICATION OF PNEUMATIC JACK FOR AUTOMOBILE." *International Conference on Information Engineering, Management and Security (2015)*: 272-281. Print.

To move the cylinder outwards, valve V1 must be open (the coil is supplied with electric current) and valve V2 closed (no current flowing).



To move the cylinder inwards, valve V2 is open and valve V1 closed. The diagram also makes it clear why vent "R" on the valve is required. Without this vent, the cylinder would be unable to move as the same pressure would be exerted on both sides of the piston and the air would not be able to escape. The pneumatic system uses manually or electrically operated valves to control direction of movement. Directional control valves can be operated by hand lever or electric solenoid to maintain an adjustable travel rate. The internal porting or spool of the directional control valve regulates airflow.

To extend the cylinder piston, air flows into the directional valve pressure port, through the flow control valve, and into the cylinder. As pressure builds in one end of the cylinder and the rod starts to extend, air exhausts out the opposite end of the cylinder. The flow control valve on the end of the cylinder restricts exiting airflow, which builds pressure to slow rod movement.

The exhausting air passes through the flow control valve and the directional control valve located at the end of the cylinder and exhausts to the atmosphere. When the cylinder retracts, the flow control valve at the end of the cylinder controls the flow, and the first valve allows air freely through.

Some cylinders have a cushion on one or both ends of travel. This cushion is a flow control valve that does not operate until the cylinder piston reaches a certain point in the cylinder. Then, the cushion restricts airflow to slow the cylinder movement. This allows it to move to the end of its travel at a slower speed. This adjustment is normally on the end of the cylinder head. See the air piping schematic to see what specific controls are provided with this equipment.

Because pneumatic systems always contain moisture from the air, the system should not be allowed to freeze. Freezing can damage the seals and control surfaces, allowing air leakage past valves, or locking a valve from operating.

Check valves may be inserted in the line to be sure the cylinder will stay in the desired position and not drift. This is useful in case some part is leaking, or there is a loss of air pressure in the plant system.

NEEDS FOR PNEUMATIC POWER

Pneumatic system use pressurized gases to transmit and control power as the name implies pneumatic systems typically use air as fluid medium because air is a safe, low cost and readily available fluid. It is particularly safe environments where an electrical spark could ignite leaks from the system components.
There are several reasons for considering the use of pneumatic system instead of hydraulic system liquid exhibit greater inertia than gases. Therefore in hydraulic system the weight of the oil is a potential problem. To design and development a material handling system for automation or semi automation of industries by using pneumatic control system which is used for low cost automation.

3.2 VALVES

SOLENOID VALVE

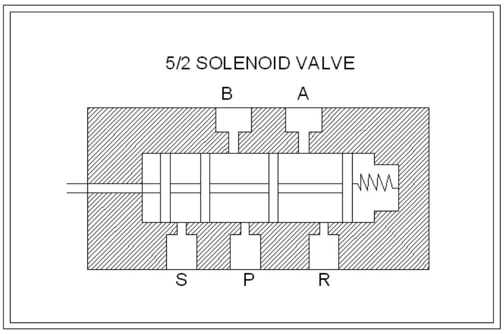
The directional valve is one of the important parts of a pneumatic system. Commonly known as DCV; this valve is used to control the direction of air flow in the pneumatic system. The directional valve does this by changing the position of its internal movable parts.

This valve was selected for speedy operation and to reduce the manual effort and also for the modification of the machine into automatic machine by means of using a solenoid valve.

Cite this article as: S.SATHIYARAJ, V.SELVAKUMAR. “DESIGN AND FABRICATION OF PNEUMATIC JACK FOR AUTOMOBILE.” *International Conference on Information Engineering, Management and Security (2015)*: 272-281. Print.

A solenoid is an electrical device that converts electrical energy into straight line motion and force. These are also used to operate a mechanical operation which in turn operates the valve mechanism. Solenoid is one is which the plunger is pulled when the solenoid is energized.

The name of the parts of the solenoid should be learned so that they can be recognized when called upon to make repairs,to do service work or to install them.



PARTS OF A SOLENOID VALVE

1. Coil

The solenoid coil is made of copper wire. The layers of wire are separated by insulating layer. The entire solenoid coil is covered with a varnish that is not affected by solvents, moisture, cutting oil or often fluids. Coils are rated in various voltages such as 115 volts AC,230volts AC,460volts Ac,575 Volts AC.6Volts DC,12Volts DC,
24 Volts DC,115 Volts DC &230Volts DC.they are designed for such
Frequencies as 50Hz to 60Hz.

2. Frame

The solenoid frame serves several purposes. Since it is made of laminated sheets, it is magnetized when the current passes through the coil. The magnetized coils attract the metal plunger to move. The frame has provisions for attaching the mounting. They are usually bolted or welded to the frame. The frame has provisions for receivers, the plunger. The wear strips are mounted to the solenoid frame, and are made of materials such as metal or impregnated less Fiber cloth

3. Solenoid plunger

The solenoid plunger is the mover mechanism of the solenoid. The plunger is made of steel laminations which are riveted together under high pressure, so that there will be no movement of the lamination with respect to one another. At the top of the plunger a pin hole is placed for making a connection to some device. The solenoid plunger is moved by a magnetic force in one direction and is usually returned by spring action.Solenoid operated valves are usually provided with cover either the solenoid or the entire valve. This protects the solenoid from dirt and other foreign matter, and protects the actuator. In many applications it is necessary to use explosion proof solenoids.

WORKING OF SOLENOID VALVE:

The solenoid valve has 5 openings. These ensure easy exhausting of 5/2Valve.the spool of the 5/2 valve slide inside the main bore according to spool position: the ports get connected and disconnected.
The working principle is as follows.

Position-1

When the spool is actuated towards outer direction port 'P' gets
Connected to 'B' and 'S' remains closed while 'A'gets connected to 'R'.

Position-2

When the spool is pushed in the inner direction port 'P' and 'A'
Gets connected to each other and 'B' to 'S' while port 'R'remains closed.

SOLINOID VALVE (OR) CUT OFF VALVE:

The control valve is used to control the flow direction is called cut off valve or solenoid valve. This solenoid cutoff valve is controlled by the electronic control unit. In our project separate solenoid valve is used for flow direction of vice cylinder. It is used to flow the air from compressor to the single acting cylinder.

3.2.2 Flow control valve:

In any fluid power circuit, flow control valve is used to control the speed of actuator. The flow control can be achieved by varying the area of flow through which the air in passing.
When area is increased, more quantity of air will be sent to actuator as a result its speed will increase. If the quantity of air entering into the actuator is reduced, the speed of the actuator is reduced.

3.2.3 Pressure control valve:

The main function of the pressure control valve is to limit (or) Control the pressure required in a pneumatic circuit.
Depending upon the method of controlling they are classified as

1. Pressure relief valve
2. Pressure reducing valve

3.3.5 Hoses:
Hoses used in this pneumatic system are made up of polyurethane. These hose can with stand at a maximum pressure level of 10 $x10^5 N/m^2$.

3.3.6. Connectors:
In our system there are two type of connectors used. One is the
Hose connector and the other is the reducer. Hose connectors normally comprise an adopt hose nipple and cap nut. These types of connectors are made up of brass (or) aluminum (or) hardened pneumatic steel.

3.4 CONTROL UNIT:
The pneumatic jack machine. Air-operated device used for many small operations. It is a portable one. Compressed air is the source of energy for this device. The compressed air is allowed. Here the compressed air from the compressor firstly enters the Control unit. In the control unit the pressure of the air is controlled.

3.5 PRESSURE GAUGE:
Pressure gauges are usually fitted with the regulators. So the air Pressure adjusted in the regulator is indicated in the pressure Gauge, is the line pressure of the air taken to the cylinder. 3.6.JACK A jack is a mechanical device used as a lifting device to lift heavy loads or apply great forces. Jacks employ a screw thread or hydraulic cylinder to apply very high linear forces.
A mechanical jack is a device which lifts heavy equipment. The most common form is a car jack, floor jack or garage jack which lifts vehicles so that maintenance can be performed.
A pneumatic jack is a hydraulic jack that is actuated by compressed air - for example, air from a compressor - instead of human work. This eliminates the need for the user to actuate the hydraulic mechanism, saving effort and potentially increasing speed. Sometimes, such jacks are also able to be operated by the normal hydraulic actuation method, thereby retaining functionality, even if a source of compressed air is not available.

DESIGN OF EQUIPMENT AND DRAWING

4.1 **PNEUMATIC COMPONENTS AND ITS SPECIFICATION**

The pneumatic jack machine consists of the following components to full fill the requirements of complete operation of the machine.
1. Double acting pneumatic cylinder
2. Solenoid valve
3. Flow control valve
4. Connectors
5. Hoses

1. Double acting pneumatic cylinder:

Technical Data

Stroke length: cylinder stroke length 100mm =0.1m
Piston rod : 10mm =10 $X10^{-3}$m
Quantity : 1
Seals : Nitride (Buna-N) Eastover
End cones : Cast iron
Piston : EN-8
Media : Air
Temperature: 0-80°C
Pressure Range: $8N/m^2$

2. Solenoid Valve

Technical data

Size : $0.635x10^{-2}$m
Part size : G$0.635x10^{-2}$m
Maximum pressure: 0-10 $x10^5 N/m^2$
Range

Quantity: 1

3. Flow control valve:

Technical data

Port size: 0.635 x 10^{-2} m
Pressure: 0-8 x10^{5} N/m^{2}
Media : Air
Quantity: 1

4. Connectors

Technical data

Max working pressure : 10 x10^{5} N/m^{2}
Temperature : 0-100°C
Fluid media :Air
Material :Brass

5. Hoses

Technical data

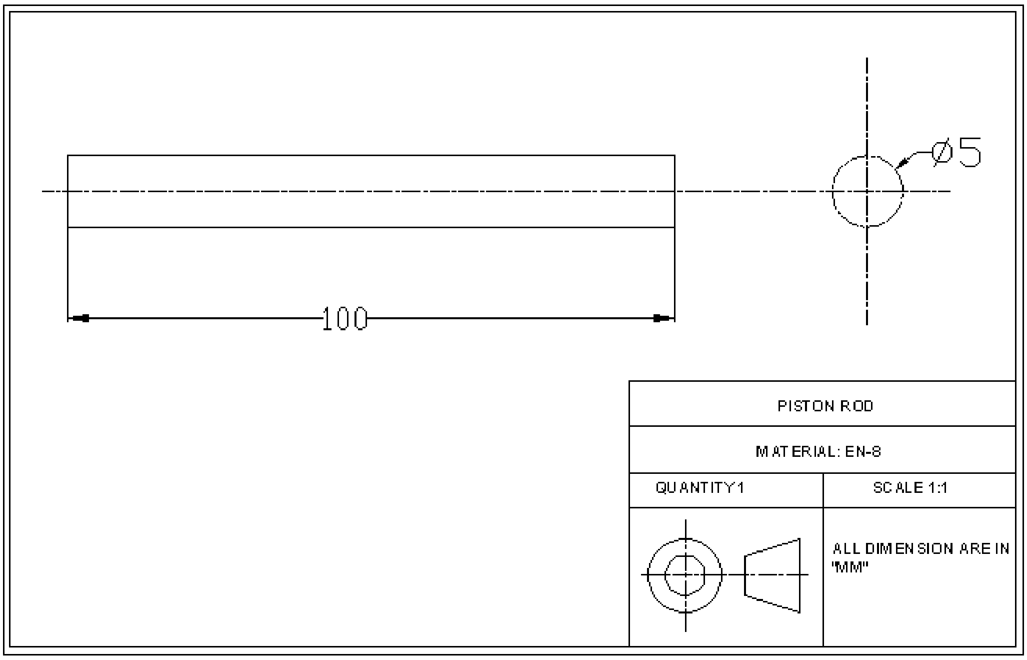
Max pressure : 10 x10^{5} $_{N/m}{}^{2}$
Outer diameter : 6mm =6 x 10^{-3} m
Inner diameter : 3.5mm =3.5 x10^{-3}m

Pneumatic unit

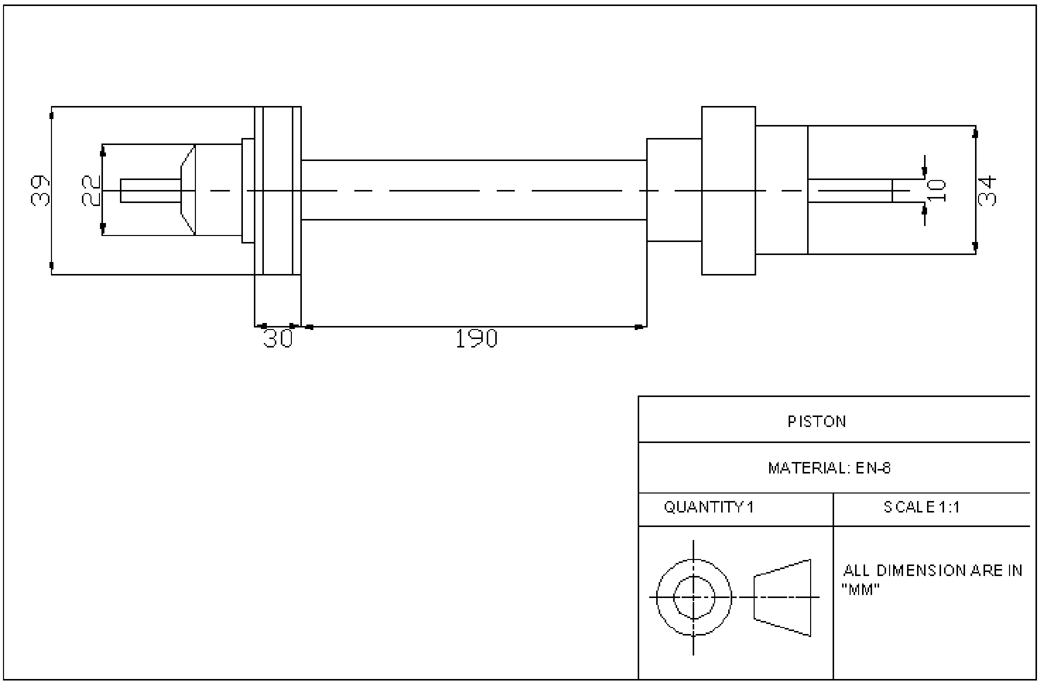
Type of cylinder : Double acting cylinder
Type of valve : flow control valve &solenoid valve
Max air pressure : 8 x10^{5} N/m^{2}

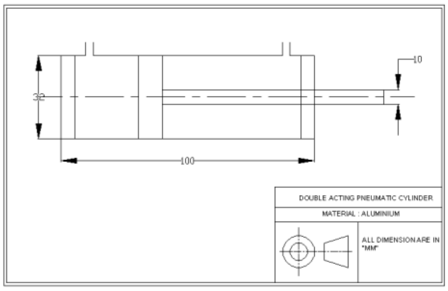
4.3 DESIGN CALCULATION

Max pressure applied in the cylinder (p) : 8 x10^{5}N/m^{2}
Area of cylinder (A) : (3.14/4*(D^{2})
: 80.38mm^{2}
: 80.38 X $10^{-4}$$m^{2}$
Force exerted in the piston (F) : Pressures applied X area Of cylinder.

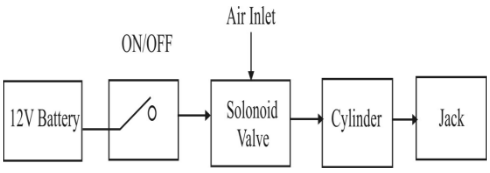


Cite this article as: S.SATHIYARAJ, V.SELVAKUMAR. "DESIGN AND FABRICATION OF PNEUMATIC JACK FOR AUTOMOBILE." *International Conference on Information Engineering, Management and Security (2015)*: 272-281. Print.

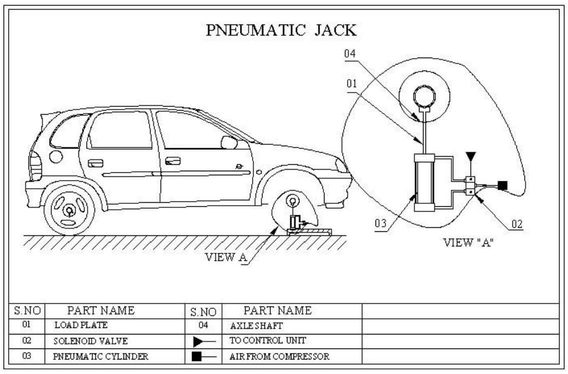




BLOCK DIAGRAM



FABRICATION OF PNEUMATIC JACK



METHOD OF FABRICATION

Here the pneumatic jack is worked with the help of pneumatic power. The name of jack is "pneumatic jack" To carry the vehicle load for working in the automobile workshop and in the service station.

WORKING PRINCIPLE

The working medium adopted is compressed air. The compressed air is transmitted through tubes to pneumatic cylinder where power is converted into reciprocating motion. The reciprocating motion is obtained by using an electrically controlled solenoid valve. The input to the solenoid valve is given through the control unit. The reciprocating motion transmitted to the jack through the piston which moves on the cylinder. The jack is placed under the vehicle chassis, where the vehicle to be lifted. The vehicle can be lifted when the solenoid valve is switched. The vehicle over the jack gets the reciprocating motion through the piston which is connected to the jack. Thus using a pneumatic jack the vehicle can be lifted with ease in operation.

- Power can be easily transmission
- Less loss in transmission
- A single compressor can supply power to many pneumatic Jacky.
- Low cost
- Easy to work and reduces the manual stress

DEMERITS
Need separate compressor.

APPLICATIONS

Used in automobile service stations and can also used in vehicles instead of screw jack.

FACTORS DETERMINING THE CHOICE OF MATERIALS

The various factors which determine the choice of material are discussed below.

1. Properties:

The material selected must posses the necessary properties for the proposed application. The various requirements to be satisfied
Can be weight, surface finish, rigidity, ability to withstand environmental attack from chemicals, service life, reliability etc.

The following four types of principle properties of materials decisively affect their selection

a. Physical
b. Mechanical
c. From manufacturing point of view
d. Chemical

The various physical properties concerned are melting point, thermal
Conductivity, specific heat, coefficient of thermal expansion, specific gravity, electrical conductivity, magnetic purposes etc.

The various Mechanical properties Concerned are strength in tensile,
Compressive shear, bending, torsional and buckling load, fatigue resistance, impact resistance, eleastic limit, endurance limit, and modulus of elasticity, hardness, wear resistance and sliding properties.

The various properties concerned from the manufacturing point of view are,

- Cast ability
- Weld ability
- Surface properties
- Shrinkage
- Deep drawing etc.

2. Manufacturing case:

Sometimes the demand for lowest possible manufacturing cost or surface qualities obtainable by the application of suitable coating substances may demand the use of special materials.

3. Quality Required:

This generally affects the manufacturing process and ultimately the material. For example, it would never be desirable to go casting of a less number of components which can be fabricated much more economically by welding or hand forging the steel.

Cite this article as: S.SATHIYARAJ, V.SELVAKUMAR. "DESIGN AND FABRICATION OF PNEUMATIC JACK FOR AUTOMOBILE." *International Conference on Information Engineering, Management and Security (2015)*: 272-281. Print.

4. Availability of Material:

Some materials may be scarce or in short supply.it then becomes obligatory for the designer to use some other material which though may not be a perfect substitute for the material designed.the delivery of materials and the delivery date of product should also be kept in mind.

5. Space consideration:

Sometimes high strength materials have to be selected because the forces involved are high and space limitations are there.

6. Cost:

As in any other problem, in selection of material the cost of material plays an important part and should not be ignored.
Some times factors like scrap utilization, appearance,and non-maintenance of the designed part are involved in the selection of proper materials.

CONCLUSION

The project carried out by us made an impressing task in the field of automobile and automobile workshops. It is very usefully for the workers to work in the automobile workshop are in the service station.
This project has also reduced the cost involved in the concern. Project has been designed to perform the entire requirement task which has also been provided.

Cite this article as: S.SATHIYARAJ, V.SELVAKUMAR. "DESIGN AND FABRICATION OF PNEUMATIC JACK FOR AUTOMOBILE." *International Conference on Information Engineering, Management and Security (2015)*: 272-281. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9	VOL	01
Website	www.iciems.in	eMail	iciems@asdf.res.in
Received	10 - July - 2015	Accepted	31- July - 2015
Article ID	ICIEMS047	eAID	ICIEMS.2015.047

Some Tuning Methods of PID Controller For Different Processes

R. G. Rakshasmare[1], G. A. Kamble[2], R. H. Chile[3]

[1]M.tech, S.G.G.S Institute of Engineering and Technology, Vishnupuri, Nanded - 431606, India.

[2]M.tech, College Of Engineering, Autonomous institute Pune, India

[3]Professor, is with S.G.G.S Institute of Engineering and Technology, Vishnupuri, Nanded-431606, India

Abstract: *Proportional, Integral and Derivative (PID) con-trollers are the most widely used controller in the chemical process industries because of their simplicity, robustness and successful practical application. Many tuning methods have been proposed for PID controllers such as Ziegler-Nichols, Tyreus-Luyben, Cohen-Coon, IMC, IMC based PID, FuzzyPID. Our purpose in this study is comparison of these tuning methods for single input single output (SISO) systems using computer simulation. Comparative analysis of performance evaluation of different controller are performed. Such as percentage of overshoot, settling time, rise time has been used as the criterion for comparison. These tuning methods have been implemented for first, second and third order systems with dead time and for two cases of set point tracking and load rejection response has considered.*

I. INTRODUCTION

A proportional-integral-derivative controller (PID con-troller) is a control loop feedback mechanism (controller) widely used in industrial control systems. A PID controller calculates an error value as the difference between a mea-sured process variable and a desired set point. The controller attempts to minimize the error by adjusting the process through use of a manipulated variable. The field of Fuzzy control has been making rapid progress in recent years. Fuzzy logic control has been widely exploited for nonlinear, high order and time delay system [2]. This paper has two main contributions. Firstly, a PID controller has been designed for higher order system using Ziegler-Nichols frequency response method and its performance has been observed. The Ziegler Nichols tuned controller parameters are fine tuned to get satisfactory closed loop performance. Secondly, for the same system a fuzzy logic controller has been proposed with simple approach and smaller number of rules (four rules) as it gives the same performance as by the larger rule set [1], [3], [7], [8], [9], [10]. Simulation results for a higher order system have been demonstrated. A performance comparison between Ziegler Nichols tuned PID controller, IMC-based PID controller,Tyreus-Luyben,Cohen-Coon PID Controller and the proposed fuzzy logic controller is presented. In this study we have compared the performances of these tuning methods. For simulation study first, second and third order systems with dead time have been employed and it was assumed that the dynamics of system is known. Simulation study has been performed for two cases of set point tracking and load rejection.The paper has been organized as follows, Section-II explains generalized model of PID controller.

Cite this article as: R. G. Rakshasmare, G. A. Kamble, R. H. Chile. "Some Tuning Methods of PID Controller ForDifferent Processes." *International Conference on Information Engineering, Management and Security (2015)*: 282-288. Print.

Section-III describes the design consideration for a higher order system. Section IV presents design of PID controller using Z-N technique. Section V presents design of fuzzy logic controller using simple approach and smaller rule base. Section VI finally conclusion close the paper.

II. GENERALISED MODEL OF PID CONTROLLER

The PID control logic is widely used in the process control industry. PID Controllers have traditionally chosen by the control system engineers due to their flexibility and reliability. A PID controller has proportinal, integral and derivative terms that can be represented in transfer function form as

$$K(s) = K_p + \frac{K_i}{s} + K_d s$$

where,

K_p represents the Proportional gain. K_i represents the Integral gain.
K_d represents the Derivative gain.

III. DESIGN CONSIDERATION

A PID controller is being designed for a first,second and higher order system with transfer function,

1) First order plus dead time model(FOPDT).
   $T(s) = e^{(0:3s)}=(s + 1)$
   where,dead time( )=0.3 sampling time(T_s)=0.05($_1$)=1
2) second order plus dead time model(SOPDT).
   $T(s) = e^{(0:3s)}=(0:4s + 1)(0:5s + 1)$
   where,dead time( )=0.3 sampling time(T_s)=0.05($_1$)=0.4 ($_2$)=0.5
3) Higher order plus dead time model.
   $T(s) = 0:0404e^{(0:1s)}=s^3 + 3:27s^2 + 3:61s + 0:07107$

Fig.shows the simulink model of the PID controller and the plant with unity feedback. i) PID controller using Z-N technique (ii) fuzzy controller so that the closed loop system exhibit small overshoot Mp and settling time ts with zero steady state error ess.

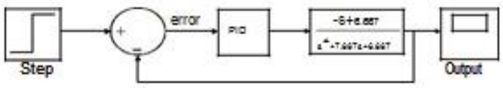


Fig. 1. A simple PID Controller system block diagram

TABLE I
ZEIGLER-NICHOLS METHOD

Controller	K_p	K_i	K_d
P	$0:5\ K_u$	–	–
PI	$0:455\ K_u$	$0:833\ P_u$	–
PD	$0:71\ K_u$	–	$0:15\ P_u$
PID	$0:6\ K_u$	$0:5\ P_u$	$0:125\ P_u$

IV. DESIGN OF PID CONTROLLER FOR DIFFERENT

TUNING METHOD

A. Ziegler-Nichols Method

Frequency response method suggested by Zeigler-Nichols is applied for design of PID controller.

Cite this article as: R. G. Rakshasmare, G. A. Kamble, R. H. Chile. "Some Tuning Methods of PID Controller ForDifferent Processes." *International Conference on Information Engineering, Management and Security (2015)*: 282-288. Print.

By setting T_i=1 and T_d=0 and using the proportional control action(K_p)only, the value of gain is increased from 0 to a critical value K_u at which the output first exhibits oscillations.P_u is the corresponding period of oscillation. The unit step response for different values of gain K_p were observed. The step response for the K_p=7.65 is shown in figure below: The above response clearly shows that sustained

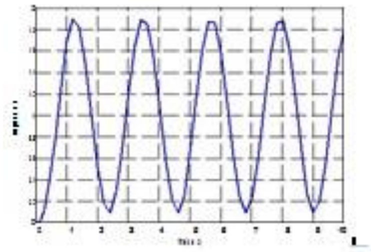


Fig. 2. Step response for K_p=7.65

oscillation occurs for K_p = K_u=7.65. The ultimate period Pu obtained from the time response is 3.14.K_u and P_u are Zeigler-Nichols parameters which can be calculated for plant by inserting the plant in setup with a step input and gain K and tuning the gain K upto which the plant output is sustained oscillations. At that time,gain K will be equal to K_u and P_u will be the time difference between two consecutive peaks.

B. Tyreus Luyben Method

The Tyreus-Luyben procedure is quite similar to the Ziegler-Nichols method but the final controller settings are different. Also this method only proposes settings for PI and PID controllers. These settings that are based on ultimate gain and period are given in below table

TABLE II
TYREUS-LUYBEN METHOD

Controller	K_p	i	d
PI	0:31 K_u	2:2 p_u	
PID	0:31 K_u	2:2 P_u	0:152 P_u

C. Cohen-Coon Method

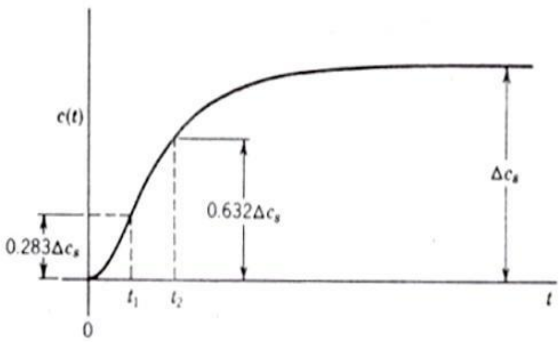
In this method the process reaction curve is obtained first, by an open loop test as shown in Figure , and then the process dynamics is approximated by a first order plus dead time model, with following parameters:

$_m$= 3=2(t_2 t_1)

d_m= $_2$ $_m$

t_1 = time at which = 0.283 Cs t_2 = time at which = 0.632 Cs C = the plant output.

This method is proposed by Dr C. L. Smith provides a good approximation to process reaction curve by first order plus dead time model After determining of three parameters of k_m , $_m$ and d, the controller parameters can be obtained, using Cohen-Coon relations given in Table 2.3 . These relations were developed empirically to provide closed loop response with a decay ratio.



Cite this article as: R. G. Rakshasmare, G. A. Kamble, R. H. Chile. "Some Tuning Methods of PID Controller ForDifferent Processes." *International Conference on Information Engineering, Management and Security (2015)*: 282-288. Print.

TABLE III
COHEN-COON CONTROLLER SETTING

Controller	Kp	i	d
P	1 m / km d (1 + d / 3 m)	–	–
PI	1 m / km d (9 / 10 + d / 12 m)	d (30+3m d m / 9+20 dm)	–
PID	1 m / km d (3 / 4 + d / 4 m)	d (32+6 d m / 13+8 d m)	d (4 / 11+2 d m)

D. Internal Model Controller

The main advantage to IMC is that it provides a transparent framework for control-system design and tuning. The IMC control structure can be formulated in the standard feedback control structure. For many processes, this standard feedback control structure will result in a PID controller (sometimes cascaded with a first-order lag). This is pleasing because we can use standard equipment and algorithms (i.e., PID controllers) to implement an "advanced" control concept.
The IMC design procedure is exactly that of the open-loop "control" design procedure. Remember that a factorization of the process model was performed so that the resulting controller would be stable. If the controller is stable and the process is stable, then the overall control system is stable.

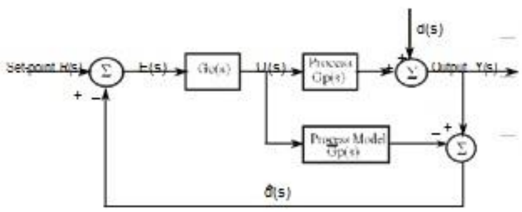


Fig. 4. Schematic of the IMC scheme

1) IMC Design Procedure: The assumption we are making is that the model is perfect, so the relationship between the output, y, and the setpoint, r, is given by equation

$y(s) = G_p(s)q(s)r(s)$. Model uncertainty is handled by adjusting the "filter factor" for robustness (tolerance of model uncertainty) and speed of response. The IMC design procedure consists of the following steps.

Develop a process model, $\hat{G}_p(s)$
Factor the process model into invertible ("good stuff") and noninvertible ("bad stuff"-time delays and RHP zeros) portions, usually using an all-pass factorization.

$$\hat{G}_p(s) = \hat{G}_p^{+}(s)\hat{G}_p^{-}(s)$$

This factorization is performed so that the resulting controller will be stable.

Invert the invertible portion of the process model (the good stuff) and cascade with a filter that makes the controller q(s) proper.

$$q(s) = \hat{G}_p^{-1}(s)f(s)$$

Cite this article as: R. G. Rakshasmare, G. A. Kamble, R. H. Chile. "Some Tuning Methods of PID Controller ForDifferent Processes." *International Conference on Information Engineering, Management and Security (2015)*: 282-288. Print.

For a focus on step setpoint changes, the following form is often used:

$$f(s) = \frac{1}{(\ s+1)^n}$$

and n is chosen to make the controller proper (or semiproper).

For good rejection of step input load disturbances, the form used is,

$$f(s) = \frac{s+1}{(\ s+1)^n}$$

where is selected to cancel the slow process time constant.

E. DESIGN OF FUZZY LOGIC CONTROLLER (FLC)

Simulink model of the fuzzy controller and the plant with unity feedback is shown in Fig For a two input fuzzy con-troller, 3,5,7,9 or 11 membership functions for each input are mostly used [7]. In this paper, only two fuzzy membership functions are used for the two inputs error e and change in error e˙ membership functions for the output parameter are shown in Fig., here N means Negative, Z means Zero and P means Positive.

TABLE IV
FUZZY RULE

e/ e	N	P
N	N	Z
P	N	P

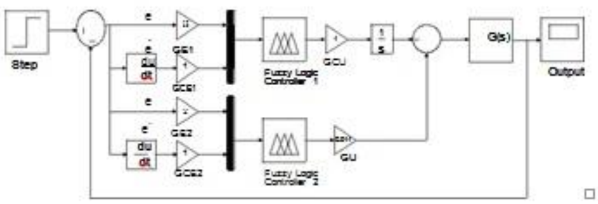


Fig. 5. System with fuzzy logic controller

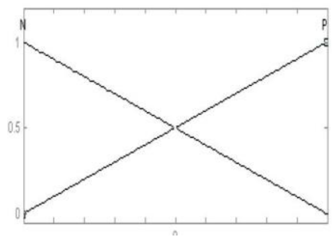


Fig. 6. Membership function for inputs e and e˙

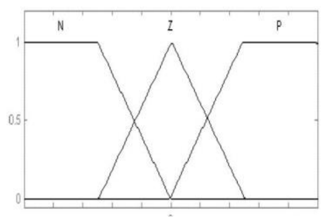


Fig. 7. Membership function for outputs

V. RESULT AND CONCLUSION

This paper covered an overview of PID controller, design of PID controller using Z-N, T-L, C-C, IMC technique and design of fuzzy logic controller for first, second, higher order processes. Simulation results using Matlab simulink are discussed for Ziegler Nichols, Tyres Luyben, Cohen-Coon, IMC based PID controller and the Fuzzy logic controller. Ziegler Nichols technique gives high overshoot and settling time with zero steady state error. Initial controller parameters obtained using Z-N formula need to be adjusted repeatedly through computer simulation to get satisfactory performance. IMC based PID controller gives zero steady state error and smaller overshoot and settling time than Ziegler Nichols tuned PID controller but it is not applicable for higher order. The Fuzzy Logic controller gives no overshoot, zero steady state error and smaller settling time than obtained using Ziegler Nichols tuned PID controller and IMC based PID controller. The simulation results shown in table 5.1,5.2,5.3 confirms that the implemented Fuzzy logic controller with simple design approach and smaller rule base can provide better performance comparing with the Ziegler Nichols tuned PID controller, IMC based PID controller, Tyres-Luyben tuned PID controller, Cohen-Coon tuned PID controller.

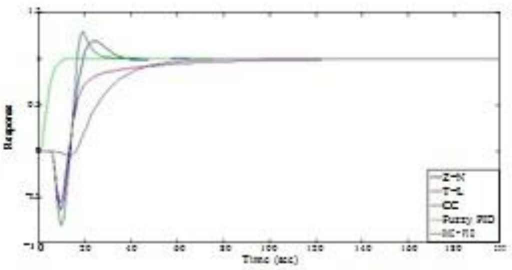


Fig. 8. Simulation plot for first order plus dead time model

TABLE V
FIRST ORDER PLUS DEAD TIME MODEL

Method	%overshoot	Ts	Tr
Z-N	18	33.32	18
C-C	28	25:48	2.32
T-L	0:0693	193.78	81:32
IMC	2:24	8:25	3:027
FUZZY PID	0:0	10:31	5:68

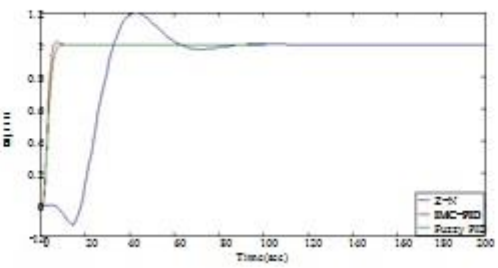


Fig. 9. Simulation plot for second order plus dead time model

TABLE VI
SECOND ORDER PLUS DEAD TIME MODEL

Method	%overshoot	Ts	Tr
Z-N	19:95	77.81	11.03
IMC	2	5:716	2:73
FUZZY PID	0:04	6:7	3:729

Cite this article as: R. G. Rakshasmare, G. A. Kamble, R. H. Chile. "Some Tuning Methods of PID Controller ForDifferent Processes." *International Conference on Information Engineering, Management and Security (2015)*: 282-288. Print.

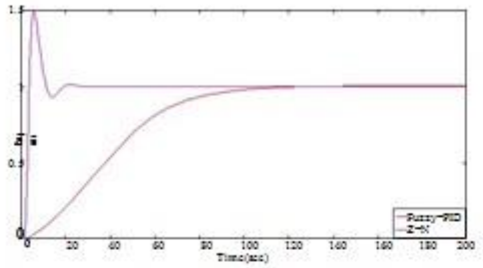


Fig. 10. Simulation plot For Higher order system

TABLE VII
HIGHER ORDER PLUS DEAD TIME MODEL

Method	%overshoot	Ts	Tr
Z-N	47:09	18.46	1.39
FUZZY PID	0	102	62

REFERENCES

[1] Gopal, Control Systems: Principles and Design. McGraw-Hill Education (India) Pvt Limited, 2002.
[2] B. Bequette, Process control: modeling, design, and simulation. 2002.
[3] M. Morari and E. Zafiriou, Robust process control. Prentice Hall, 1989.
[4] C. C. Lee, "Fuzzy logic in control systems: fuzzy logic controller. I,"
[5] Systems, Man and Cybernetics, IEEE Transactions on, vol. 20, pp. 404– 418, Mar 1990.
[6] K. J. Astrom and T. Hagglund, PID Controllers: Theory, design, and tuning. International Society for Measurement and Control, Research Triangle Park, NC., 1995.
[7] R. Bandyopadhyay and D. Patranabis, "A new autotuning algorithm for PID controllers using dead-beat format," ISA Transactions, vol. 40, no. 3, pp. 255 – 266, 2001.
[8] H.-X. Li and H. B. Gatland, "Conventional fuzzy control and its enhancement," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, vol. 26, pp. 791–797, Oct 1996.
[9] Z. J. Khan, S. R. Vaishnav, "Design and performance of pid and fuzzy logic controller with smaller rule set for higher order system," oct 2007.
[10] M. Shahrokhi and A. Zomorrodi, "Comparison of PID controller tuning methods,"
[11] K. H. Ang, G. Chong, and Y. Li, "PID control system analysis, design, and technology," Control Systems Technology, IEEE Transactions on, vol. 13, pp. 559–576, July 2005.
a. Visioli, Practical PID Control. Advances in Industrial Control, Springer, 2006.
[12] K. Ogata, Modern Control Engineering. Upper Saddle River, NJ, USA: Prentice Hall PTR, 4th ed., 2001.
[13] D. E. Rivera, M. Morari, and S. Skogestad, "Internal model control: PID controller design," Industrial & engineering chemistry process design and development, vol. 25, no. 1, pp. 252–265, 1986.
[14] K. Astrom and T. Hagglund, "The future of PID control," Control Engineering Practice, vol. 9, no. 11, pp. 1163 – 1175, 2001.
[15] G. K. I. Mann, B. G. Hu, and R. Gosine, "Time-domain based design and analysis of new pid tuning rules," Control Theory and Applications, IEE Proceedings, vol. 148, pp. 251–261, May 2001.
[16] W. L. Luyben, Process Modeling, Simulation, and Control for Chemical Engineers. McGraw-Hill, 1990.
[17] F. G. Shinskey, Process Control Systems: application, design, and tuning. McGraw-Hill, 1996.

Cite this article as: R. G. Rakshasmare, G. A. Kamble, R. H. Chile. "Some Tuning Methods of PID Controller ForDifferent Processes." *International Conference on Information Engineering, Management and Security (2015)*: 282-288. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS048

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.048

ONLINE SIGNATURE RECOGNITION USING NEURAL NETWORK

Long CAI[1], Kokula Krishna Hari Kunasekaran[2], Vignesh R[3]
[1]University of Hong Kong, HKSAR, Hong Kong
[2]Techno Forum Research and Development Center, India
[3]Life Member, Association of Scientists, Developers and Faculties, India

ABSTRACT: *Here it discusses about some of the features of signature data and their extraction from the raw data set collected from ATVS signature database. The features include time duration, sign changes of dx/dt and dy/dt, average jerk, number of pen-up pen-down etc. Signature features are pre-processed and brought to a value having same decimal point and trained using back propagation neural network. For signature data of 10 users and accuracy rate of 86% is obtained.*

KEYWORDS: PATTERN RECOGNITION, FALSE ACCEPTANCE RATE, FALSE REJECTION RATE, NEURAL NETWORK, BACK PROPAGATION, CONFUSION MATRIX.

1. INTRODUCTION

Authentication of user is becoming very important to do business transactions, accessing data and for security purpose. Many different techniques are applying for authentication purpose. User IDs and passwords, PIN codes, ATM card, PAN card are many different ways which are common today, but the problem of such systems are that, they need remembering different PINs or passwords, or they need to be kept secret from others access. Signature is a behavioural biometric. Automatic signature authentication is now getting popular in research areas because of its acceptance in legal and social areas and its widespread use in authentication purpose.
Since signature for different individuals vary with the variation of individuals, so it is a very robust biometric to authenticate a user. Signature verification is a very difficult pattern recognition problem. Since intra class variations occur, even experts get difficulty to recognize the forgery signature. And also it is very easy to forge a signature. Signature is believed to be a reflex action which produces its dynamic properties unconsciously.

Biometrics can be broadly divided into physiological and behavioural biometrics. Physiological biometrics is fingerprint; iris, face recognition etc and behavioural biometrics are signature, voice, hand writing etc. Authentication of signature is done by detecting forgeries. Forgeries can be divided into random forgery, simple simulated forgery and simulated skilled forgery. Fig1 shows an example of forgery signatures. Random forgeries are produced randomly without any information about the name and person for whom signature is produced. Random forgery are generated when forger do not have any available access to the signature. They may

Cite this article as: Long CAI, Kokula Krishna Hari Kunasekaran, Vignesh R. "ONLINE SIGNATURE RECOGNITION USING NEURAL NETWORK." *International Conference on Information Engineering, Management and Security (2015)*: 289-293. Print.

have different shape and size from the authentic signature. Simple forgery may have same semantic meaning like authentic one but overall shape and size may differ. Skilled forgery signatures are the signatures which are given by a person by vigorous practice and a kin observation.

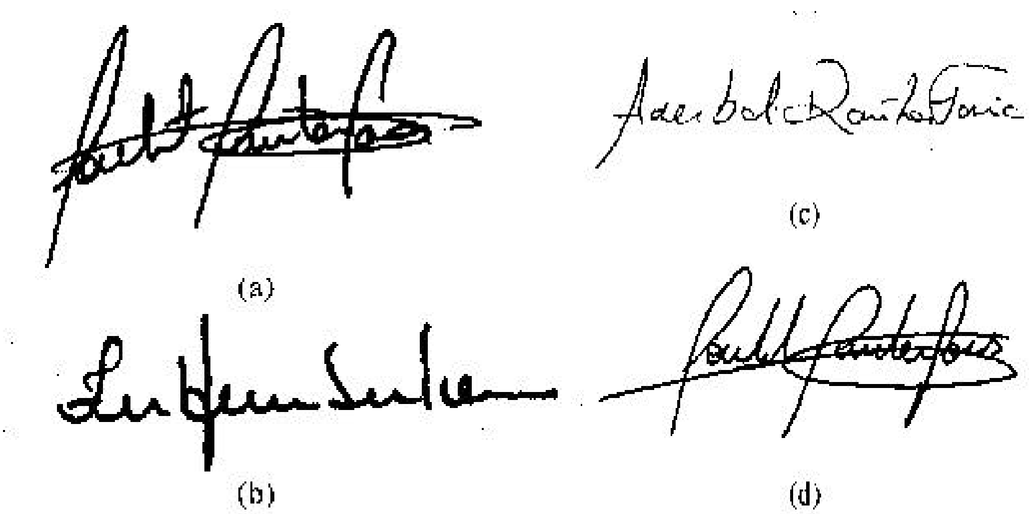


Fig 1: (a) genuine signature; (b) random forgery; (c) simulated simple forgery; (d) simulated skilled forgery

For online signature verification, signature data are generally taken using capacitive tablet or PDA which gives the x-y coordinate, pressure readings etc. From these raw data different features can be computed. Signature authentication problem can be solved by two ways i.e. dynamic and static. According to that features can be classified as dynamic features and static features. Dynamic data can be obtained by electronic tablet or PDA, and static images are obtained either by camera or scanning the photo of the signatures. Dynamic features [1] are functions of time and static features are time independent. Even if a skilled forger produces the same looking signature like the authentic one, but they can't produce the same pressure produced by the authentic one, so it is difficult to forge. The use of pen dynamics over shape of signature would be very useful in forgery detection because dynamic features of a signature are not readily available for forger as in the shape of offline signature.

Handwritten signatures can be represented by multiple modals i.e. global and local, shape based and time based. Local shape based signature and their advantages are discussed in [2]. Three signature databases are collected and analysed for a different period of time. Different reasons for inaccuracies are also discussed. Global feature based technique is applied in [4]. Three global features i.e. projection moment, upper envelope based characteristics and lower envelope based characteristic are used and then multiple neural network is applied for the classification purpose. To remove noise they have applied median filter. A fusion of local and regional features is discussed in [5]. The local function based features are classified with DTW and regional features are classified with HMM. In [7] signature verification based on logarithmic spectrum is done. Principle components of the logarithmic spectrum are compared with the reference signature and similarity value is calculated between the enrolled and reference signature. A stroke based method for shape and dynamics of signature is discussed in [8]. Two levels strategy i.e. soft and hard rule are implemented in it.

The signature authentication can be done by two methods that are verification and recognition. In verification, the features of the test signature are compared with the limited number of stored features of the signature, and need to verify if the signature belongs to that particular person or not. But in recognition we are not known to the output signature, we have to recognize the given signature belongs to the given database or not. In recognition feature matching is done with the entire database.

Online signature verification can be broadly classified into two groups based on their feature extraction: parametric approach and function based approach. In parametric approach a set of parameters (e.g. Speed, displacement, position, pen up pen down, wavelet transform etc.) extracted can be used as a feature to form a signature pattern, and those feature patterns can be used as reference and test signature to examine the authentication of the signature. In function based approach the features are the function of time (e.g. velocity, acceleration, pressure, direction of pen movement etc.). Online signatures are characterized as a time function.

In any verification task there are two types of error involved i.e. false rejection rate (FRR) and false acceptance rate (FAR). False rejection is when the authentic signature is rejected and false acceptance means when a forged signature is accepted to be authentic. When percentage of false rejection rate is equal to the percentage of false acceptance rate we call it equal error rate. Equal error rate is the measure of the performance of a biometric system. Average error rate is the average of FAR and FRR. The authenticity of test signature is evaluated by matching it with that of reference signature. There are many techniques available for matching e.g. dynamic time warping (DTW) [9], hidden Markov model (HMM), support vector machine (SVM) and neural network (NN). When functions are considered, the matching technique must take into account the variation of duration of signature. A method of similarity measure

Cite this article as: Long CAI, Kokula Krishna Hari Kunasekaran, Vignesh R. "ONLINE SIGNATURE RECOGNITION USING NEURAL NETWORK." *International Conference on Information Engineering, Management and Security (2015)*: 289-293. Print.

for signature verification and recognition using symbolic representation is done in [3]. In [6] a method of string matching or dynamic time warping is done. Local features and stroke based global features are extracted and compared the results by varying the values of absolute and relative speed of the signature. Signature feature vectors are formed in symbolic interval value and test data are inserted to check if it lies within that interval or not. Interval value is set by calculating mean, variance and standard deviation. Writer dependent and feature dependent threshold are set in the database.

- Dynamic time warping technique is generally used for function based parameter. But the time complexity of DTW is more of the order of (O^2).
- HMM performs stochastic matching using probability distribution of the features. They can compute both similarity and variability of the pattern. But they require a large dataset to train and are complex.
- Support vector machine classifies one class of data from the other by finding the hyper plane that maximizes the separation between classes. SVM have algorithmic complexity, and requires large storage for large scale task
- Neural network have ability of generalization. They can be used to detect nonlinear equations for dependent and independent variables. They can train large amount of database. Easily implemented in parallel architecture.

2. FEATURE EXTRACTION

The database is collected from ATVS signature sub corpus. The database contains raw data values of the signature i.e. x-coordinate, y-coordinate, time stamp, pen up pen down and pressure signal. From these data the features are extracted.

Total duration of signature: it is the time taken to complete a signature. It can be calculated as the difference between the last time stamp and the first time stamp.

Number of pen ups: the number of times pen is removed from the pad/paper.

Sign changes of dx/dt and dy/dt: dx/dt and dy/dt may be positive or negative value. So when it changes the value from positive to negative or negative to positive it is counted.

Average jerk: jerk is change in acceleration with respect to time. Average jerk is the mean of the jerk.

Standard deviation of acceleration in y-direction: standard deviation of a_y ; where

Velocity in y-direction $v_y = dy/dt$

Acceleration in y-direction $a_y = dv_y/dt$

Standard deviation of velocity in y-direction: standard deviation of v_y ; where

Velocity in y-direction $v_y = dy/dt$

Number of local maxima in x direction: local maxima can be calculated from change in x with respect to time.

Standard deviation of acceleration in x-direction: standard deviation of a_x ; where

Velocity in y-direction $v_x = dx/dt$

Acceleration in y-direction $a_x = dv_x/dt$

Standard deviation of velocity in x-direction: standard deviation of v_x ; where

Velocity in y-direction $v_x = dx/dt$

3. METHODOLOGY

In neural network approach, the main procedure to implement is: first of all the features need to be extracted and then the network is to be trained to learn the relationship between the pattern and its class. During training, validation of the network is to be checked by few features to see if the network is giving a satisfactory result or not. After validation, the network is to be tested using features which are completely unknown for the network, to see the performance of the network. Data base is collected from ATVS signature sub corpus [11]. Database consists of 25 signature data for each individual. From the raw data set consisting of information of x coordinate, y coordinate, time stamps, pressure and pen up and pen down, features were extracted. Then the database is divided into three parts for training, validation and testing. For training first 15 signature were extracted, for validation next 5 signature were extracted and for testing also remaining 5 signature were extracted and randomized them. The matching technique used is neural network approach. At first the extracted data are brought to 10th decimal point. Then the data is normalized so that the pattern values are between 0 and 1. During training weights are updated to minimize the difference between the desired output and the actual output i.e. error. The fixed weights after training can be used for the task in pattern recognition and classification. The neural network structure used have three layers: one input layer, one hidden layer and one output layer. Activation provides the measure of confidence of corresponding decision of the classifier. Commonly used activation functions are sigmoid functions. The activation function used here is log sigmoid function (logsig). It gives the activation label between 0 and 1. If activation is 1 it means confidence is high and if it is 0 means confidence is zero.

3.1 Vector matrix form of back propagation algorithm

1. An input pattern is presented and calculated the outputs of the network at all the internal layers
2. For each of the layers, the sensitivity vector is calculated according to

$D^{(s)} = G(v^{(s)})(d_q - x_{out}^{(s)})$ for output layer

$D^{(s-1)} = G(v^{(s-1)})W^{(s)T}D^{(s)}$ for all hidden layers

3. The synaptic weights are updated for the network according to

Cite this article as: Long CAI, Kokula Krishna Hari Kunasekaran, Vignesh R. "ONLINE SIGNATURE RECOGNITION USING NEURAL NETWORK." *International Conference on Information Engineering, Management and Security (2015)*: 289-293. Print.

$W^{(s))}(k+1) = W^{(s)} + \alpha^{(s)}D^{(s)}x_{out}^{(s-1)T}$

4. Continue steps 1 through 3 until the network reaches the desired mapping accuracy

d_q = desired output
x_{out} = actual output of the neural network
k = iteration number

4. RESULT

Neural network is a generalization tool. The reason why the neural network approach is chosen among the number of other classification method is that, neural network is easy to use and can solve complex problems with ease. From this work it is realized that, when variation of data is more, neural network finds it difficult to generalize. That is why normalization of database is important. When introduced pre-processing of data by converting them to 10^{th} decimal point or same decimal point value the generalization becomes more and classification error decreases. Confusion matrix is a table which helps the visualization of the performance of supervised machine learning. The diagonal boxes in the table from left (up) to right (down) gives the true positive classification. And other boxes show the true negative classification. From the confusion matrix in fig2, when genuine signatures were taken the True positive rate found is 86% i.e. false rejection rate (FRR) is 14%. From confusion matrix in fig3, when forgery signatures were taken, the false acceptance rate is 12%.

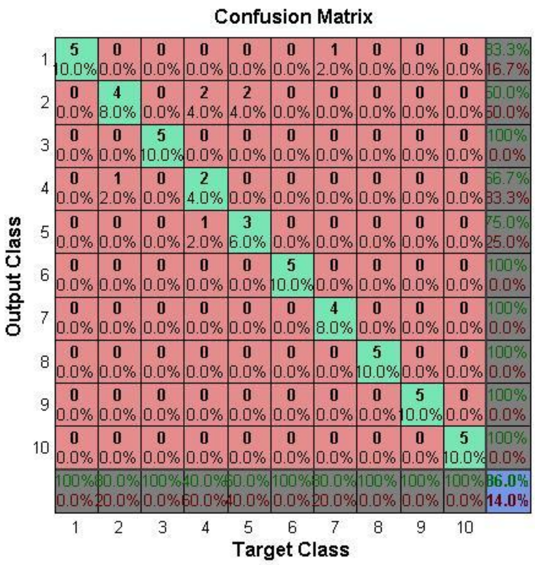


Fig2: confusion matrix for genuine signature data of 10 users

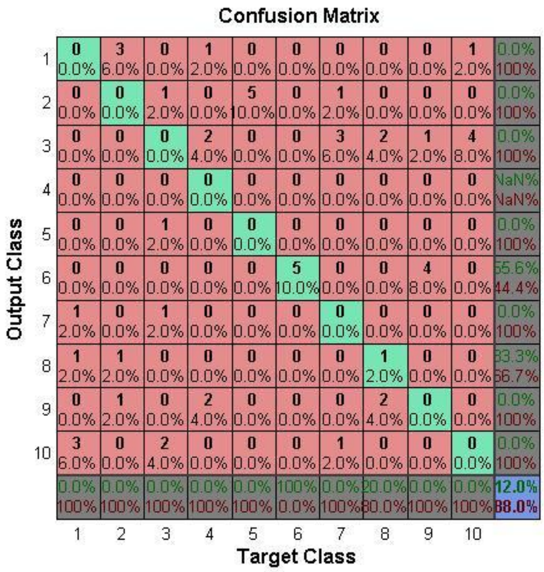


Fig3: confusion matrix for forgery signature data of 10 users

Cite this article as: Long CAI, Kokula Krishna Hari Kunasekaran, Vignesh R. "ONLINE SIGNATURE RECOGNITION USING NEURAL NETWORK." *International Conference on Information Engineering, Management and Security (2015)*: 289-293. Print.

From the confusion plot the accuracy of the system can be calculated by
Accuracy = (100 – (FAR + FRR)/2) %
The accuracy rate of the system is 86 %

5. CONCLUSION

The main objective of this work is to construct a signature recognition system so that to get a maximum accuracy label. To do this, some features were extracted and formed pattern from them. The features acts as an input pattern to the neural network and corresponding targets are constructed. In neural network, the patterns are trained according to the target, where weights are updated to get a minimum error. When a stopping condition is reached the iteration stops. In neural network training many times of trial and error is done to get satisfactory neural network architecture. Parameters like hidden nodes in a neural network, initial learning rate parameter, learning rate schedule are needed to be adjusted again and again. Neural network architecture is dependent on these parameters. In this work the nodes in the hidden layers are 80. Initial learning rate parameter is 1 and learning rate scheduled at 300. These parameters give a false acceptance rate of 12% and a false rejection rate of 14%. And accuracy rate when calculated gives an accuracy rate of 87%.

Since neural network has a generalization capability, once trained its weight need not be changed again. In testing it gives the result from the trained architecture itself. The main drawback of neural network training is that, for larger dataset it is very difficult to adjust the parameters by trial and error method. And the time consumption is more.

REFERENCES

[1] W. Nelson and E. Kishon, "Use of Dynamic Features for Signature Verification," Proc. IEEE Int'l Conf. Systems, Man and Cybernetics, vol. 1, pp. 201-205, 1991.
[2] V.S. Nalwa, "Automatic On-Line Signature Verification," Proc. Third Asian Conf. Computer Vision, vol. 1, pp. 10-15, 1997.
[3] D.S. Guru and H.N. Prakash "Online Signature Verification and Recognition: An Approach Based on Symbolic Representation". IEEE transactions on pattern analysis and machine intelligence, vol. 31, no. 6, june 2009.
[4] R. Bajaj and S. Chaudhary, "Signature Verification Using Multiple Neural Classifiers," Pattern Recognition, vol. 30, pp. 1-87, 1997.
[5] J.F. Aguilar, S. Krawczyk, J.O. Garcia, and A.K. Jain, "Fusion of Local and Regional Approaches for On-Line Signature Verification," Proc. Int'l Workshop Biometric Recognition System, pp. 188-196, 2005.
[6] A.K. Jain, F. Griess, and S. Colonnel, "On-Line Signature Verification," Pattern Recognition, vol. 35, pp. 2963-2972, 2002.
[7] Q.-Z. Wu, S.-Y. Lee, and I.-C. Jou, "On-Line Signature Verification Based on Logarithmic Spectrum," Pattern Recognition, vol. 31, no. 12, pp. 1865-1871, 1998.
[8] L. Bovino, S. Impdevo, G. Pirlo, and L. Sarcinella, "Multiexpert Verification of Hand-Written Signature," Proc. Int'l Conf. Document Analysis and Recognition, pp. 932-936, 2003.
[9] F.-Z. Marcos, "On-Line Signature Recognition Based on VQDTW," Pattern Recognition, vol. 40, no. 3, pp. 981-992, 2007.
[10] J.F. Aguilar, S. Krawczyk, J.O. Garcia, and A.K. Jain, "Fusion of Local and Regional Approaches for On-Line Signature Verification," Proc. Int'l Workshop Biometric Recognition System, pp. 188-196, 2005.
[11] DESCRIPTION OF ATVS-SSig DB, National Laboratory of Pattern Recognition (NLPR),Institute of Automation, Chinese Academy of Sciences(CASIA)

Cite this article as: Long CAI, Kokula Krishna Hari Kunasekaran, Vignesh R. "ONLINE SIGNATURE RECOGNITION USING NEURAL NETWORK." *International Conference on Information Engineering, Management and Security (2015)*: 289-293. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS049

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.049

RAGA ANALYSIS AND CLASSIFICATION OF INSTRUMENTAL MUSIC

Prabha Kumari[1], Prof.(Dr.) Y.H.Dandawate[2], Angha Bidkar[3]
[1,2,3] Vishwakarma Institute of Information Technology, Pune, S.NO.34 , Kondhwa, Budruk

Abstract: *Raga played by Indian instrument is the actually soul of the Indian classical music. Indian classical music is famous around all over the world for its particular structure and well soundness. Our work is related to analyze and classify the instrumental music according to their features. This will help to non-professional and music learner for understanding and acquire knowledge about the music using the system intelligence. There are various features for analysis of music but our approach is towards the spectral and temporal features. For extraction of feature we prefer MIR toolbox and MATLAB function, it is really a very helpful tool for these purpose. At very first we just collect clips of ragas and find out the spectral and temporal features. These features show the better result. We are using four ragas namely:- Bhairav, Bhairavi, Todi and Yaman. For classification we use different types of classifier just like KNN classifier and SVM classifier they gives approximate 87% and 92% accuracy respectively..*

I. INTRODUCTION

Indian classical music is known for its perfect technical soundness and well defined structure known as ragas . Each raga is based on some specific combination of swara(notes). Any raga should have at least five notes out of seven and it is also possible that two raga has same notes but the aarohan, avrohan and pakad is different so one can identify it properly. Expert person can understand the raga very easily, but for learner it is very difficult to classify and identify the raga. So this method is helpful for both professional and non-professional one. Proposed work is related to classification and analysis of Instrumental Indian classical music. We are using MATLAB tool for processing music segment and find out information related to raga analysis and classification. **M**usic **I**nformation **R**etrieval toolbox (MIR)[2] is also helpful to find out the features for comparison. This software widely use in western music and now days implemented in Indian Classical music. Work focused on four ragas namely Bhairav, Bhairavi, Todi and Yaman and the selected instrumental music is mixed polyphonic to find out the spectral and temporal features like brightness, RMS energy, spectral flux spectrum chromogram, histogram etc. For classification we preferred KNN and SVM classifier [2], [3]. Swaras are the frequency generated by instrumentally or vocally. Actually these seven swaras represents the absolute frequencies ratio with respect to each other and these are very similar to SOLFEGE in western music The seven swaras are namely: Shadja (Sa), Rishabh (Re), Gandhara (Ga), Madhyama (Ma), Panchama (Pa), Dhaivata(Dh), and Nishad (Ni).Out of seven, two swaras i.e Sa and Pa has only pure form while other five has both pure and impure form in structural elements of raga. The purpose behind this work is to design a computer based education of Indian classical music for everyone.

Cite this article as: Prabha Kumari, Y.H.Dandawate, Angha Bidkar. "RAGA ANALYSIS AND CLASSIFICATION OF INSTRUMENTAL MUSIC." *International Conference on Information Engineering, Management and Security (2015)*: 294-301. Print.

Following are the basic terms related to Indian classical music.
a) Raga:

Indian music is famous for its ragas specialty. Raga represents the color of emotion and it has some specific melodic phrases. Raga has definite pattern of notes (swaras). Two ragas may be same notes but the pattern of distribution is different in case of each raga that is why two ragas have same notes but they tuned differently.
b) *Aarohan ,Avrohan and Pakad*
Aarohan and Avrohan is the ascending and descending progression of swaras respectively. Pakad is a small sequence of swaras in a raga that acts as a mark for the raga and an artist often calls and recalls the pakad over a performance.
c) Vadi , Samvadi and Jati

The Vadi is the most noticeable notes used in a raga and next to vadi, samvadi is the second most noticeable notes in a raga. Often Vadi as the Monarch of swaras for that raga. There are three jati in Indian classical music namely: Audhav-has five notes, Shadav-has six notes and Sampoorna has seven notes. Ragas have combination of Jati, e.g. Audhav- Audhav, Audhav – Shadav
The rest of the paper isorganized as follows: The literature survey is described in section II. Methodology explained in section III. The proposed approach details in section IV. Section V describes the details regarding hardware and its related issues. The experimentation and results are projected in section VI and section VII concludes the paper

II .LITERATURE SURVEY

For literature we have been following the related papers for, how we can extract these features the method of analysis and classification of ragas, to distinguish the different characteristics of raga and their structure and methodology. Pranay Dighe *et.al* [1] has proposed work to vigorous programmed analysis of Indian classical music through machine learning and signal processing toolbox like MIR toolbox. They developed idea to perform scale-independent raga identification by using a random forest classifier on swara histograms and succeeded state-of-the-art results for the same. The average accuracy is approximate 97% and some algorithm is also used by them especially for computation of swara based features. By referring this paper we acknowledged about the features extraction of ragas and using these features for the classification of two ragas. Although this paper has done good work in classification of ragas but no means for temporal properties of ragas have been provided by the authorss. The same has been included in proposed work. Gopala Krishna Kaluti and Rajeshwari Shreedhar *et.al*[2][3] has proposed and music recognition and classification their work related to implement computational model for raga recognition. In this work they examined the raga and identifying these ragas naturally. They mainly focus on the pitch of the notes and through which it is recognizable of notes .The maximum accuracy is approximate 94%. Through this paper we acquire knowledge about we can implement this technique for both western music and Indian classical music. Although this paper has done excellent work in raga identification and classification but no any description about the analysis of features of ragas have been done by the authors.
The same work has been included in proposed work. Sourabh Deshmukh *et. al*. [3] has proposed ethno musico logical identification of singer and ragas. He has used features for analysis and all those features are analysed in time or frequency domain and for classification he has used various types of classifier and maximum accuracy is 93%. V. Sivaranjani [4] has proposed work for mainly pitch analysis and retrieves information of music which is related to pitch. Soubhik Chakraborty *et.al.* [5] has proposed work related to scientifically verify a raga on the basis of vadi and samvadi. They have done statistical analysis of ragas and not using properties like temporal or spectral which are used by proposed work. Prasad Reddy *et.al* [6] suggested Automatic Raaga Identification System for Carnatic Music mainly for melkartha raga identification. They have used Hidden Markov Model and also introduce special algorithm for pakad matching for this purpose Michael K.[7] has recommended exploration for North Indian classical music, his process is related to analysis of the raga in north Indian classification the north Indian classical music. He used ethno musicological method for raga analysis and positing it. Kris West [9] worked on the analysis of audio music using multiple processes for classification and implemented so called search –by – example method. He used novel machine learning algorithm (MVCART) that is totally based on the classic Decision Tree algorithm. Preeti Rao *et.al* [10] proposed identification of melodic motifs in raga. They used machine learning method for phrase classification on data that is manually segmented. They used HMM and Dynamic time warping method for classification. Rajshri Pendekar*et.al* [11] worked on raga and swara identification using harmonium, they used onset detection for determination of spectral flux for frequency estimation. The programing used here is dynamic in nature this technique is mainly used for template matching.
Although all the papers which we referred here for literature survey is related to music but there are very few work has been done in Indian instrumental music.

I. METODOLOGY

A. DATA BASE
The steps followed in proposed methodology are shown in fig.1.

Cite this article as: Prabha Kumari, Y.H.Dandawate, Angha Bidkar. "RAGA ANALYSIS AND CLASSIFICATION OF INSTRUMENTAL MUSIC." *International Conference on Information Engineering, Management and Security (2015)*: 294-301. Print.

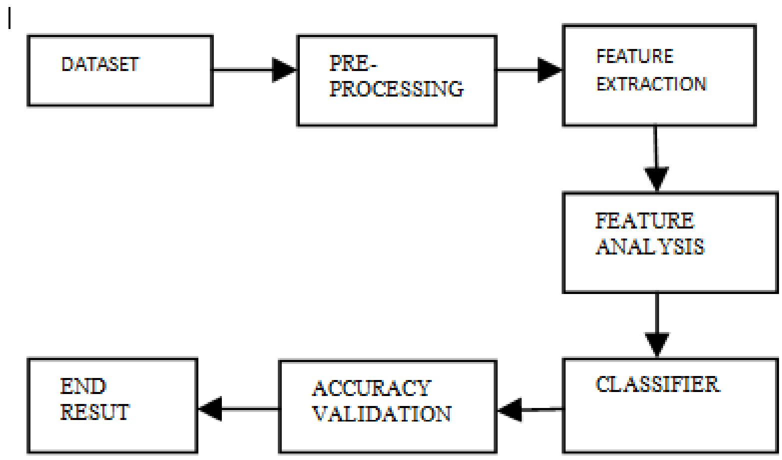


Figure 1: Block diagram of methodology

We collected the data base from live concerts recording, original CDs, IIT Kanpur, and download from internet. We have selected four ragas for our work; details about each raga is depicted below:
a) RAGA BHAIRAV : Bhairav is a raga in Hindustani classical music. Bhairav is its name from Bhiarava, a name of Shiva. It is traditionally played before sunrise. Bhairav consider the most important Hindustani raga

- ☐ Aaroh-Sa Re(k) Ga Ma Pa Dha(k) Ni Sa
- ☐ Avroh-Sa Ni Dha(k) Pa Ma Ga Re(k) Sa
- ☐ Pakad – S G MP Dha(K) P
- ☐ Vadi – Dha(k)
- ☐ Samvadi – Re(k)

b) RAGA BHAIRAVI**:** Raga Bhairavi is sampoorna raga because it contains all seven notes. Bhairavi is the most ancient raga said to have been ubiquitous about 1500 years ago. Bhairavi is its name from Goddess Bhairavi(parvati). It is a morning raga.

- ☐ Aaroh- Sa Re Ga Ma Pa Dha Ni Sa
- ☐ Avroh – Sa Ni Dha Pa Ma Ga Re Sa
- ☐ Vadi – Ma
- ☐ Samvadi – Sa

c) RAGA TODI: "Todi' is a Hindustani classical raga. Its name comes from the THAAT, one of the ten modes of Hindustani classical music. Todi is a late morning raga.

- ☐ Aaroh – Sa Re(k) Ga(k) Ma Dha(k) Ni Sa
- ☐ Avroh – Sa Ni Dha(k) Ma Ga(k) Re(k) Sa
- ☐ Vadi – Dha(k)
- ☐ Samvadi – Ga(k)

d) RAGA YAMAN: Raga Yaman is also known as Emaan. It is a sampoorna Hindustani classical music. Yaman is night raga, it originates from the Persian mode other says it is vedic origins as raga Yamuna which over the time altered as Yaman.

- ☐ Aaroh – Sa Re Ga Ma(t) Pa Dha Ni Sa
- ☐ Avroh – Sa Ni Dha Pa Ma(t) Ga Re Sa
- ☐ Vadi - Ga
- ☐ Samvadi – Ni

B. PRE- PROCESSING

Preprocessing is the second step of methodology, in this step we take all the ragas and first converts these files into .WAV files. After conversion we cut each file into a specific duration i.e approximates 60sec so that we can get information about each and every note properly. For these purpose we use Virtual Dj Pro software. After pre –processing we fed the data into feature extraction.

Cite this article as: Prabha Kumari, Y.H.Dandawate, Angha Bidkar. "RAGA ANALYSIS AND CLASSIFICATION OF INSTRUMENTAL MUSIC." *International Conference on Information Engineering, Management and Security (2015)*: 294-301. Print.

C. FEATURE EXTRACTION AND ANALYSIS

Here we mainly focused on temporal as well spectral properties of raga wave file.

a) ONSET

For onset detection we use spectral flux method for ragas, in this method we calculated change in spectral energy from one frame to the next frame by using spectral flux. Actually spectral flux is the squared of normalized difference between consecutive spectral distribution.

$$\sum_{K=0}^{K=1}(|X_n(k)| - |X_{n-1}(k)| \qquad (1)$$

where n-1 and n are the frame indices and (k) is the FFT of nth frame. So through Onset detection we can easily find out the energy distribution of the specific notes.

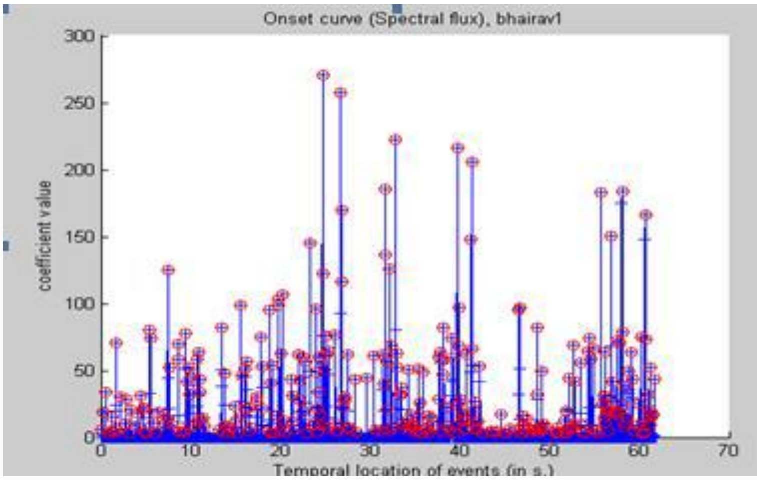


Fig.2. Onset curve of raga bhairav

Above figure shows the detail description of each formant and it provides easy access to the value of formant.

b) CHROMAGRAM

Chromagram denotes the distributions of energy along pitches. There are 12 semitones in western music which is equivalent to notes of Hindustani classical music. These semitones octaves are C, C#, D, D#, E, F, F#, G, G#, A, A# and B. these semitones are fixed with absolute frequency value and most interesting thing is that musical octave has a special property is that the current semitone is equivalent to one octave below or above to it so semitone has the property of repeatability it repeats in each octave above and below. By the table which is given below we can relate the semitones of western music with notes of Indian classical music. Although chromagram is find out through MIR toolbox but sometimes this is not provides exact notation of swara.

Table 1: Symbol of notes of Hindutani and western music.

Symbols of Indian swara notation	Symbols of Western music notation
Sa	C
Re(komal)	C#,Db
Re	D
Ga(komal)	D#Eb
Ga	E
Ma	F
Ma(Tvra)	F#Gb
Pa	G
Dha(komal)	G#Ab
Dha	A
Ni (komal)	A#Bb
Ni	B

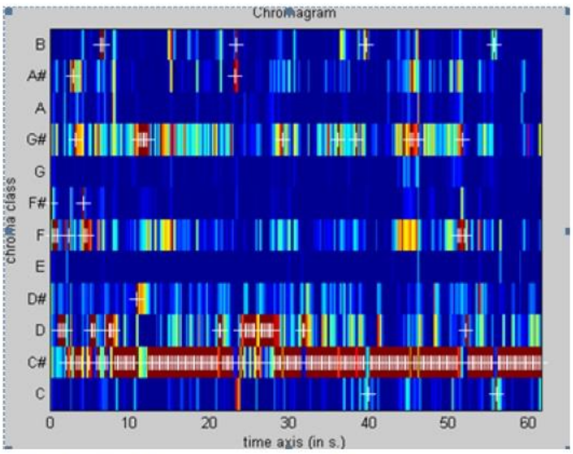


Fig.3.Distribution of chroma on the notes of the raga

Cite this article as: Prabha Kumari, Y.H.Dandawate, Angha Bidkar. "RAGA ANALYSIS AND CLASSIFICATION OF INSTRUMENTAL MUSIC." *International Conference on Information Engineering, Management and Security (2015)*: 294-301. Print.

In the above figure represents distribution of raga Chromagram is useful for extraction of notes of ragas. By the help of above diagram we can identify the notes which is used in raga bhairav .

D. PITCH DETECTION

Pitch detection is processed for the finding the frequency of pitch. It plays important role because it is linked with periodicity of the audio wave signal and these periodic signal are made up of fundamental repeated frequencies, these fundamental repeated frequencies are multiples of a joint fundamental frequency. There are various approach for finding the pitch of the audio wave namely: spatial domain, time domain frequency domain. Here we apply MIR toolbox to find out the pitch of the signal.

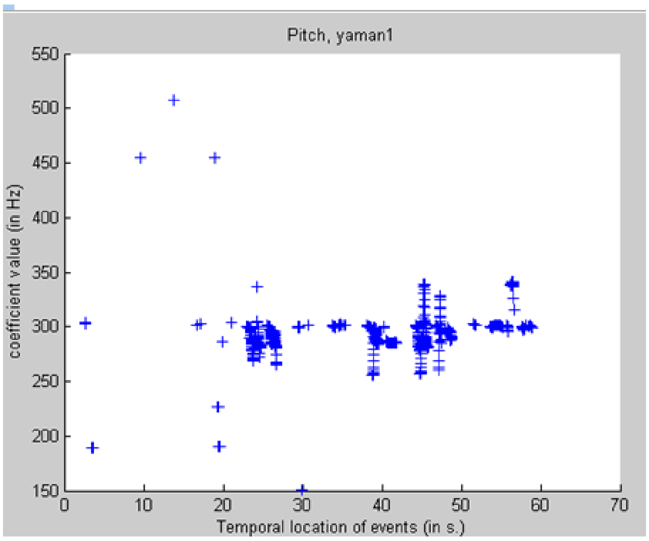


Fig.4. Pitch detection of raga Yaman

E.CLASSIFICATION

For classification of ragas we are using KNN and SVM respectively. We depicted each classification in detail given below.

a) K-NEAREST NEIGHBOR (KNN)

This method is used for pattern recognition and classification. Here some steps by which we apply this classifier for raga classification

- ☐ Take feature vector of both test and train ragas.
- ☐ Now compare the features of both test and trained ragas.
- ☐ Find the Euclidean distance between test vector and trained vector of ragas
- ☐ The test segment assigned only those who have most common category among k nearest.
- ☐ KNN-classification can be classified by following equation

$$C* = \sum_i \delta wi\,(c, (f(x \qquad (2)$$

where c is the class level i.e raga identity and fi(x) is the class label for the ith neighbor of x and (c, fi(x)).

- ☐ This method gives 87% accuracy.

b) SUPPORT VECTOR MACHINE(SVM)

- ☐ SVM is supervised learning method and it is used for classification.
- ☐ This method is used for data analysis and classification.
- ☐ In this method we categories data into two classes.
- ☐ SVM classifies data by finding the best hyper plane that separates all data into two classes .
- ☐ SVM gives approximate 93% accuraccy

IV.EXPERIMENTAL RESULT

We are using four ragas for our experiment and we find the final result of our work with accuracy 87.5% and 93% respectively. Here we used KNN and SVM classifier for classification approach. Through all the properties we have used here are give better result in classification also .Although we are using KNN and SVM classifier, graphically it seems easy to classify the Indian classical music.We are taking some spectral properties of raga for analysis and we can find out the true and wrong raga. Let us see with comparison between observation and actual ragas.We get fundamental frequency of the different raga by using this methology which is also play an important role in classification of raga.

Comparision between actual and observe notes of raga bhairav and bhairavi using chromagram

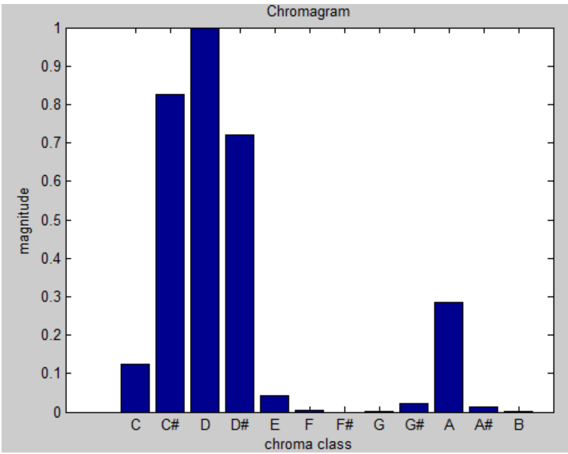


Fig.4 Chromagram of raga Bhairav

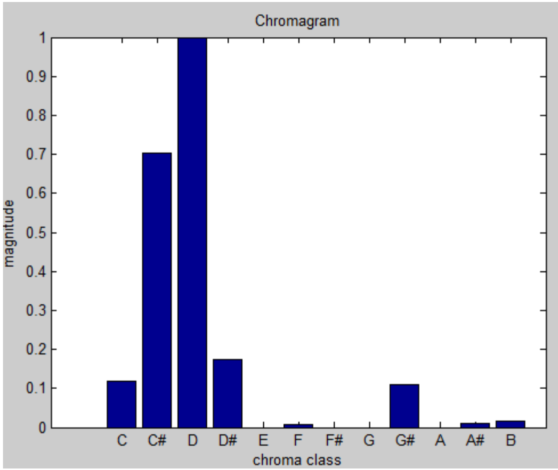


Fig.5Chromagram of Raga Bhairavi

Table no. 1 comparision between observed and actual notes

NOTES	OBSERVED	ACTUAL
VADI	Dha(k)	Dha(komal)
SAMVADI	Re(k)	Re(komal)

Table no. 2Comparision between actual and observed notes

NOTES	OBSERVED	ACTUAL
VADI	Dha(k)	Ma
SAMVADI	Sa	Sa

Although chromagram is graphical representation of chroma ,we can also find out the energy distribution among the notes of the raga from above diagram we can discriminate the notes and the energy of the notes. This is also helpful to provide the difference between vadi and samvadi, which is shown in the above table for both raga Bhairav and Bhairavi.

Cite this article as: Prabha Kumari, Y.H.Dandawate, Angha Bidkar. "RAGA ANALYSIS AND CLASSIFICATION OF INSTRUMENTAL MUSIC." *International Conference on Information Engineering, Management and Security (2015)*: 294-301. Print.

HISTOGRAM OF RAGA

Through histogram graph here we can find out the Vadi and Samvadi notes and by the help of graph and using the value of vadi andsamvadi it is while easy to detect the raga, we can also find out which type of raga has been played through this process

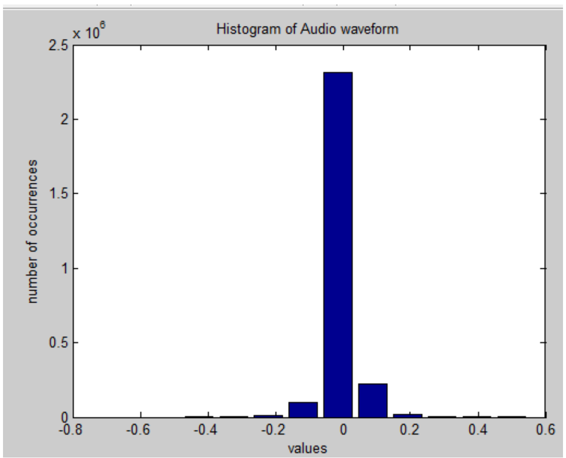


Fig.8 Histogram of raga Yaman the heighest bar shows the value of Vadi and nex to Vadi is Samvadi.

We know that the value of vadi is heighest among all notes. Now from above fig it is cleared that the value of vadi is approximate 2.25 which is nearly to raga yaman.

Table no.5 Raga database and accuracy after classification

Raga	Test sample	Accurately classify	accuracy
Bhairav	15	12	80%
Bhairavi	10	8	80%
Todi	15	10	67%
Yaman	16	12	75%
Total	56	42	75%

Classifier	Accuracy
SVM	92%
K-NN	87%

Table no.6 Accuracy of different classifier

IV. CONCLUSION AND FUTURE WORK

In the proposed work, an acoustic and signal processing based approach is used for analysis and classification of Indian instrumental music. We have tested our algorithm over the database of 60 wave files.Our work is concerned with the discrimination between four different raga using KNN and SVM classifier. Through this work we are emphasized the spectral and temporal properties of instrumental raga .The KNN classifier can significantly support in classification of different ragas with an average accuracy as high. Although this work has done for analysis and classification but we can also use this methodology for detection of raga also. This classification provide approximate 87.5% accuracy.

The main purpose behind this proposed work is designed to a model for music learner and music lover by the help of this method a person can easily classify the difference between two different ragas. This technique is also very helpful in music recommendation system. However it is also helpful in music synthesis and invention of new raga. Automatic tagging is also possible through this work. Future work also lies to improve the dataset uses as maximum as possible and short out all the imperfections which could not improve through this proposed work to increase the accuracy.

Cite this article as: Prabha Kumari, Y.H.Dandawate, Angha Bidkar. "RAGA ANALYSIS AND CLASSIFICATION OF INSTRUMENTAL MUSIC." *International Conference on Information Engineering, Management and Security (2015)*: 294-301. Print.

REFERENCES

[1] Dighe P.,KarnikH.,Raj B.,"Swara Histogram Based Structural analysis and Identification of Indian Classical Ragas ",*Mellon University, Pittsburgh PA, USA*

[2] Kaluti G., Gulati S.,Rao P.,"A Survey of Raaga Recognition Techniques and Improvements to the State-of-the-Art".*Proc of sound and music computing 2011Podova Italy.*

[3] Deshmukh S., Bhirud S., "Analysis and application of audio features extraction and classification method to be used for North Indian Classical Music's singer identification problem". *International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 2, February 2014*.

[4] Sivaranjani V., "Pitch Analysis for Musical Instruments using Cepstrum*", International Journal of Advanced Computer Technology (IJACT)* .

[5] Chakraborty S.,Ranganayakulu R.,Chauhan s.,Saulanki S.,Mahto k., "Given a raga recording, can we scientifically verify which school of Vadi-Samvadi selection is supported by the artist*?", International Journal of Computers & Technology Volume 4 No. 2, March-April, 2013.*

[6] Reddy P., Rao T.,Sudha K.,K.K., "Automatic Raaga Identification System for Carnatic Music Using Hidden Markov Model", *Global Journal of Computer Science and Technology Volume 11 Issue 22 Version 1.0 December 2011*

[7] Michaela K., "Spector, A. Z. an Exploration into North Indian Classical Music : Raga, AlifLaila",

[8] West k., "Novel technique for Audio Music Classification and Search"

[9] Rajeswari Sridhar and T.V. Geetha, "Swara identification of Carnatic music*", IEEE Computer Society press, proceeding ICIT 2006.*

[10] Rao P.,Ros J.,GanguliK.,PanditV.,Ishwar V.,Bellur, A.Murthy H., "Classification Of Melodic In Motifs In Raag Music With Time Series Matching",*Appeared In Journal Of New Music Research, Volume 43issue 1,31 March 2014.*

[11] Pendhekar R.,Mahajan S.,Majumdar R.,Ganoo P., "Harmonium Raga Recognition ",International Journal of Machine Learning and Computing, Vol. 3, No. 4, August 2013.

Cite this article as: Prabha Kumari, Y.H.Dandawate, Angha Bidkar. "RAGA ANALYSIS AND CLASSIFICATION OF INSTRUMENTAL MUSIC." *International Conference on Information Engineering, Management and Security (2015)*: 294-301. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS050

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.050

Simulation and design of PFC boost converter with constant output voltage and EMI filter

Rohit B. Chachda[1], Syed A. Naveed[2]

[1,2]Department of Electronics and Telecommunication,
MGM's JNEC Auranagabad, Maharashtra, India.

Abstract: *This paper perform simulation and design of PFC boost converter with EMI Filter which provides approximate unity power factor with constant output voltage .EMI filter is design for the purpose of products must fulfill with international electromagnetic compatibility(EMC) standards which have been developed to control conducted and radiated emissions from electrical and electronics systems. Power factor stage is require to make the input current waveform in phase to the voltage and in view of power supply like a simple resistor. The MATLAB Simulink model for proposed system is implemented and the experimental results are achieved. These experimental results are match up with the simulation results.*

Keywords: Power Factor Correction, Electromagnetic Interference, Diode Rectifier, Boost Converter

1.INTRODUCTION

Switch mode power supplies (SMPS) design by including full wave rectifier having large energy storage capacitor. When mains instantaneous voltage exceeds voltage across capacitor, SMPSs draws current. The capacitor delivers energy to the power supply for remaining portion of the AC cycle. As a outcome, High harmonic content in input current waveform of basic SMPS and hence reduces power factor. Filter is used to remove harmonics but it expensive.. Extra circuits are required to neutralize the effect of the brief current pulses. Putting a current regulated boost chopper stage after the off-line rectifier can correct the power factor.

In 2001, the European Union fixed the standard IEC/EN61000-3-2 to fixed limits on the harmonics of the AC input current up to the 40th harmonic for apparatus above 75 W. To get these requirements, modern SMPSs normally include an additional power factor stage (PFC). Due to switching action in chopper stage switch mode power supplies generate high frequency noise i.e. electromagnetic interference(EMI). EMI produced due to the current being switched on and off sharply. Hence EMI filters and RF shielding are needed to reduce the interference.

Literature deals with EMI concerns in Power Electronic Converters are given by "Ref. [4]". Design of Boost PFC Converter using genetic algorithms is given in "Ref. [5]".Analysis of EMI Conduction in boost PFC Converter is existing given in "Ref. [3]". A technique for EMI analysis in PFC rectifier is given in "Ref. [4]". Soft switching methods in PWM converters are presented in "Ref. [2]"and model of inductor design is presented in "Ref. [1]" .In the literature stated above, the hardware of boost converter using Atmel microcontroller is not existing. This paper design hardware and MATLAB Simulink model for microcontroller based boost converter as well as use of single phase model in three phase circuit.

Cite this article as: Rohit B. Chachda, Syed A. Naveed. "Simulation and design of PFC boost converter with constant output voltage and EMI filter." *International Conference on Information Engineering, Management and Security (2015)*: 302-306. Print.

2.EXPERIMENTAL SECTION

This paper work is focused in the area of active PFC approach and boost topology is employed for research on AC-DC PFC pre-regulator system for the improvement of quality of power.

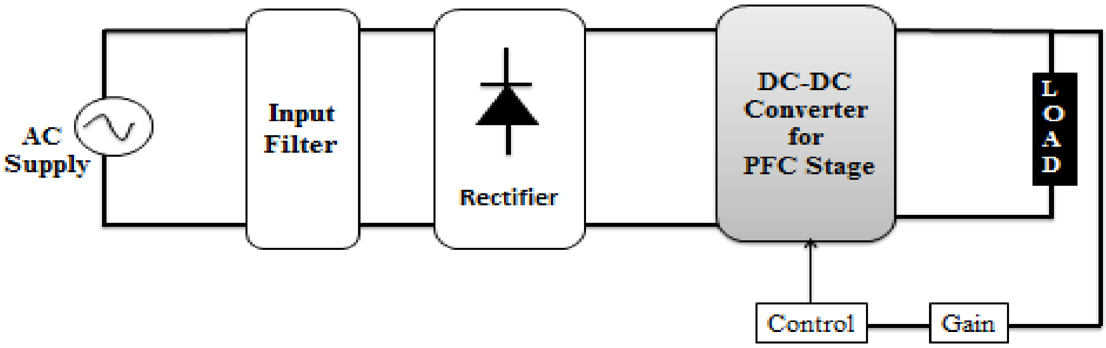


Fig: 1 proposed system

Fig. 1 shows that the PFC technique improves the input current taken from the mains supply and minimizes the DC bus voltage ripple. The target of PFC is to make the input current waveform in phase to the voltage and in view of power supply like a simple resistor.

A] Design of DC to DC converter for PFC stage

Fig. 2 shows the Boost Power Factor Correction converter. It involves boost inductor, switching device, diode rectifier, boost output voltage and boost diode capacitor.

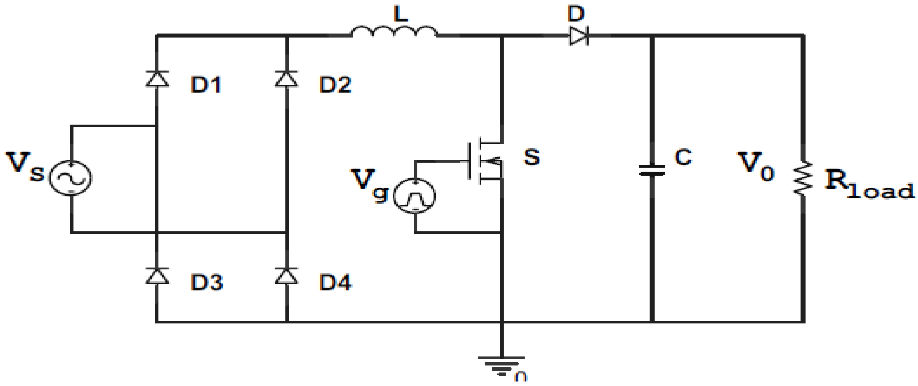


Fig. 2 Boost PFC Converter

The detail explanations of the proposed Boost PFC converter are as follows

MOSFET is used for switching purpose .Inductor and capacitor are select by using following equation “Ref. [6]”.

$$L = \frac{Vs*D}{\Delta I_l * f}\ ,\ C = \frac{Io*D}{f*\Delta Vc}$$

Output voltage equation for boost converter is $Vo = \frac{Vd}{1-\propto}$ Where $\propto$ = delay angle of the boost converter. The output voltage ideally rises from 0 to infinity , as Threshold angle rise from 0 to 1.Hence it is called as boost converter.

B] Design of EMI filter

The Electromagnetic Interference is spread in two types first one radiation and second one is the conduction. Electromagnetic noise is created in the source because of rapid current and voltage changes, and spread through the coupling mechanisms. Since breaking a coupling path is important at either the start or the end of the circuit. Hence to break coupling path EMI filter is used between ac source and bridge rectifier.

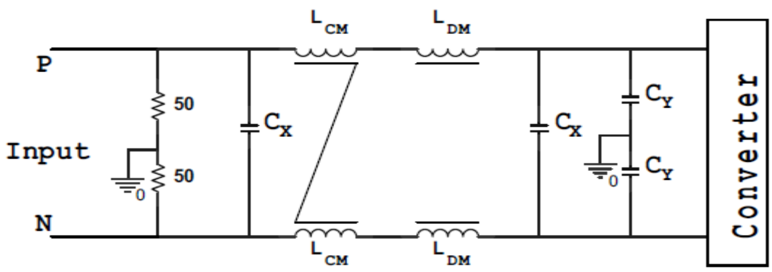


Fig. 3 EMI Filter

EMI filters can help in bypassing EMI or improving RF immunity. The filter consists of inductors and capacitors as shown in figure3. In EMI filter inductor is used to reduce the di/dt rate during its turn-off and capacitor is used to for purpose of decoupling. Resistors is used to control rise time of high speed signal.

3.SIMULATION RESULT

MATLAB Simulink is used to simulate this proposed system. The simulation circuit of PFC boost converter with constant output voltageand EMI filter is shown in Figure 4.Interference is create by using an additional source connected in series with main ac source.

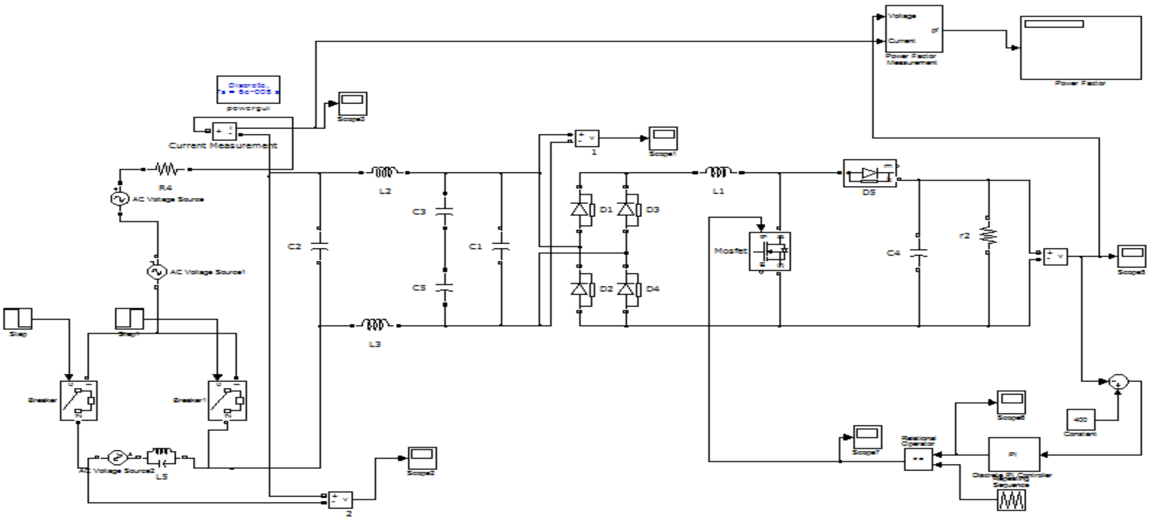


Fig.4 : PFC boost converter with constant output voltage and EMI filter

Distorted input Voltage before EMI is shown in Figure 4(a). The voltage waveform follows EMI filter is shown in Figure 4(b).Control pulses for the MOSFET are shown in Figure 4(c). Error signal for closed loop system is shown in Figure 4(D). The output voltage of closed loop system is shown in Figure 4(e). FFT analysis of input current waveform is shown in figure in 4(f).

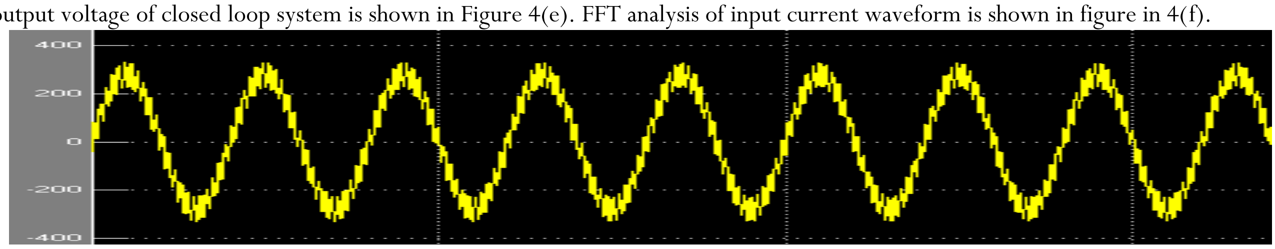


FIG. 4(A): Input voltage before EMI filter

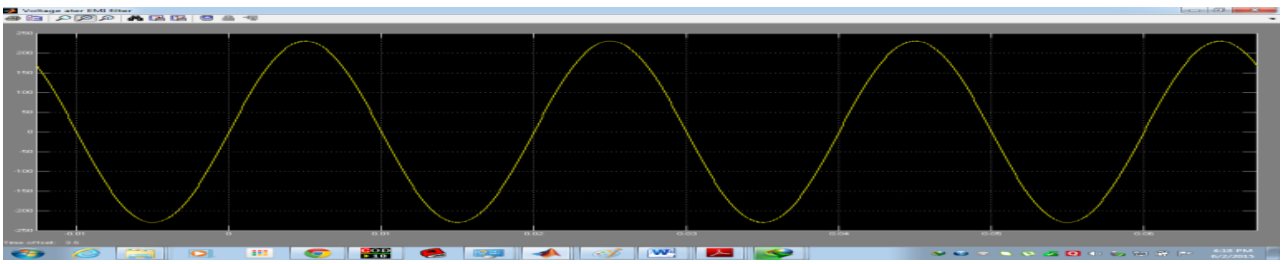


FIG. 4(B): voltage after EMI filter

Cite this article as: Rohit B. Chachda, Syed A. Naveed. "Simulation and design of PFC boost converter with constant output voltage and EMI filter." *International Conference on Information Engineering, Management and Security (2015)*: 302-306. Print.

Fig. 4(c): Driving pulses for mosfet

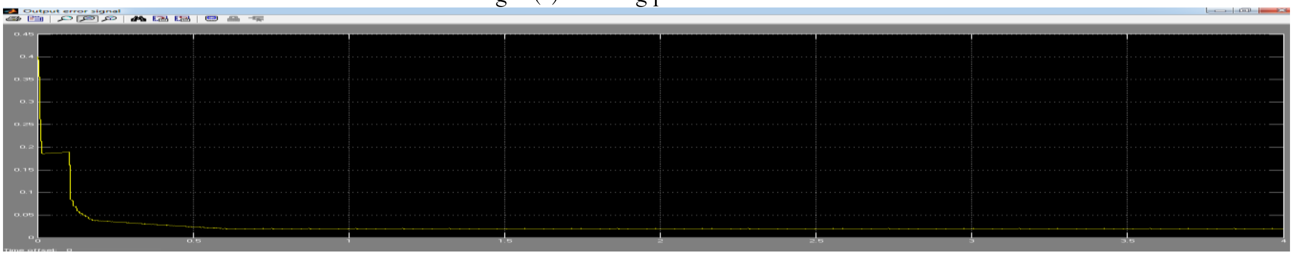


Fig. 4(d): Error signal for closed loop system

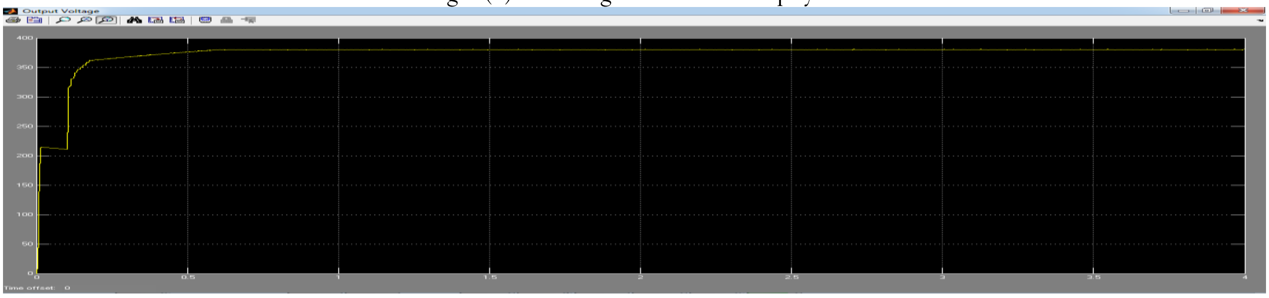


FIG.4 (e): Output voltage with closed loop system

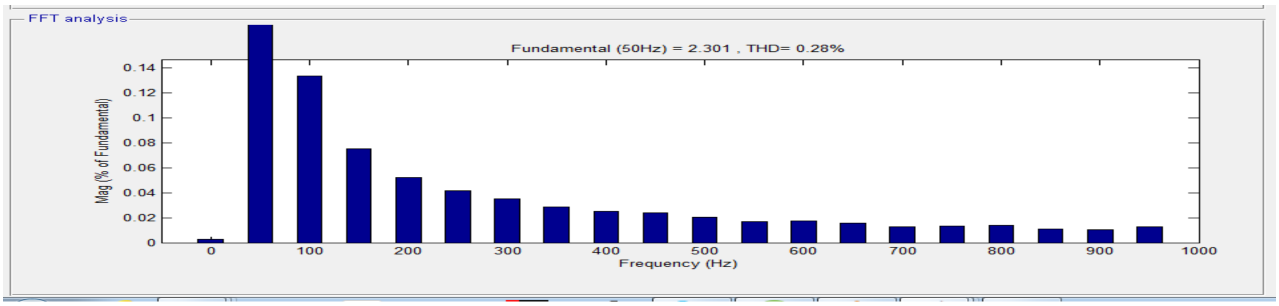


FIG.4(f): FFT analysis of input current waveform

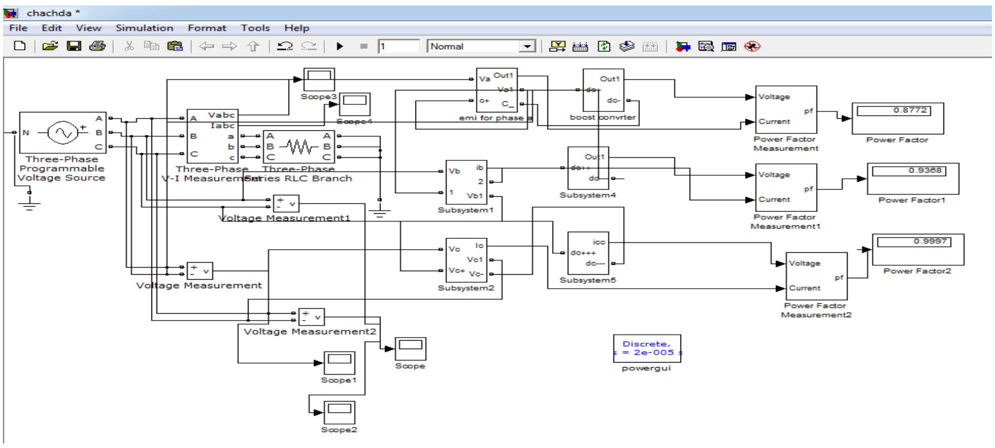


Fig.(5): USE OF SINGLE PHASE MODEL IN THREE PHASE CIRCUIT

Cite this article as: Rohit B. Chachda, Syed A. Naveed. "Simulation and design of PFC boost converter with constant output voltage and EMI filter." *International Conference on Information Engineering, Management and Security (2015)*: 302-306. Print.

Single phase circuit used in each phase of common three phase source and observe the result. It can be seen that it provides approximately unity power factor and constant boosted output voltage for each phase.

4.EXPEREMENT RESULT

Hardware of proposed system is shown in Fig. 6. Hardware consist of four stages i.e. EMI filter, signal conditioning, boost converter and microcontroller with driver circuit. Microcontroller PIC16F877 is used to generate control pulses that amplified by driver circuit upto 20V which is used to applied the gate of MOSFET. The experimental results are gained and presented here. The constant boosted output voltage is shown on LCD

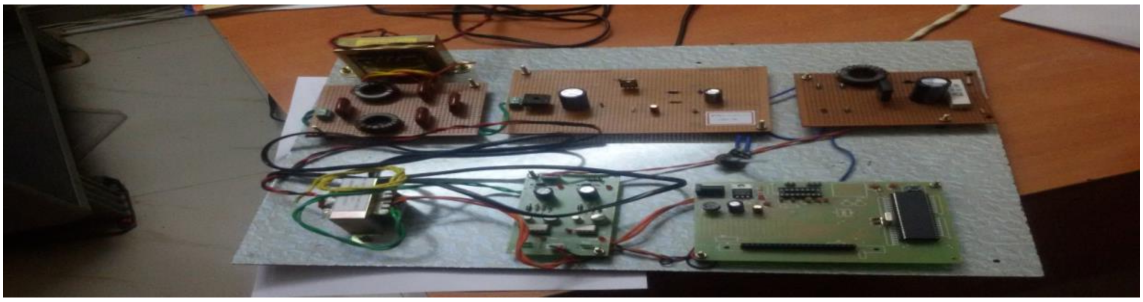


FIG.6:Hardware model of circuit

5.CONCLUSION

This paper gives the result of constant dc boosted output voltage with less harmonics and approximately unity power factor. Simulation and design of PFC boost converter with constant output voltage and EMI filter is studied, simulated and fabricated. From the simulation results, it is cleared that the best power factor can be achieved. In Simulink, this model provide approximately unity power factor while using it in each phase of common three phase source. The simulation studies prove that this model is a alternative solution for power factor improvement. The circuit is tested with resistive load. The experimental results given in this paper. The experimental results match up with the simulation results.

6.ACKNOWLEDGMENTS

The author would like to offer special thanks to mentor Dr S A Naveed from our institute for directing the special concerns related with this topic and organizing this issue .I also owes great deal to my colleagues for contributing their efforts in enthusiastic manner.

REFERENCES

[1] Evans, P. and Heffernan, W., "Inductor design concepts for high frequency applications", EPE FlRENZE, pp. 1.019-1- 1.024, 2010.

[2] Hua, Guichao and Lee, F.C., "Soft switching techniques in PWM converters", iEEE iECON, pp. 637--643, 2011.

[3] Zhang, A.W., Zhang, M.T. and Lee, F.C. et al., "Conducted EMI analysis of a boost PFC Circuit", Proc. iEEE Applied Power Electronics coni, APEC-2009, Vol. 1.

[4] Crebier, J.e., Brunello, M. and Ferrieux, J.P., "A new method for EMI study in PFC rectifiers", Proc. iEEE Power Electronics Spc. Coni, PESC-20J J, pp. 855-860.

[5] Busquets, S., Monge, G. Soremekun et al., "Design of a boost power factor correctionconverter using genetic algorithms", proc. iEEE Applied Power Electronics Con!(APEC-2002), Vol. 2, pp. 1177-1182, 2011.

[6] Abdul Fathah, Prof. B.D. Subudhi "Design of a Boost Converter",2013.

[7] Pratap Ranjan Mohanty, Prof. Anup Kumar Panda "Power Factor Correction (Pfc) Of Ac-Dc System Using Boost-Converter" ,2014

[8] P. Ram Mohan, M. Vijaya Kumar and O. V. Raghava Reddy" Simulation of a Boost PFC Converter with Electro Magnetic Interference Filter", World Academy of Science, Engineering and Technology Vol:2 2008-10-22

[9] www.wikipedia.org

Cite this article as: Rohit B. Chachda, Syed A. Naveed. "Simulation and design of PFC boost converter with constant output voltage and EMI filter." *International Conference on Information Engineering, Management and Security (2015)*: 302-306. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS051

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.051

Simulation and Design of Solar Feed EZ-Source Inverter

Vinay Y. Somwanshi[1], Syed A. Naveed[2]
[1,2] Department of Electronics and Telecommunication,
MGM's JNEC Aurangabad, Maharashtra, India.

Abstract: *Z-Source inverters are used mainly for buck-boost energy conversion with the help of passive elements. In Z-source inverter topology, impedance network is used to couples the source and the inverter to achieve voltage boost and inversion. Embedded EZ-Source inverter and Z-Source inverter can produce the same gain. Input to the Embedded EZ-Source inverter can be obtained from solar cell. The ripple content in the output voltage of solar cell is filtered and pure DC is given to the three-phase inverter which converts DC in to three-phase balanced AC. The output of the Embedded EZ Source inverter is used to control the harmonics present in the load. The limitations of the conventional Z-source inverter can be overcome by using Embedded Z-source inverters. It can produce smoother current or voltage across the dc input source and it has low harmonic distortions. Simulations are carried out using MATLAB-SIMULINK. Hardware implementation and Microcontroller programming can be done in the lab.*

Keywords: EZ-source inverters, voltage boost, Z-Source inverter and Pulse Width Modulation (PWM), Total Harmonic Distortion (THD).

I. INTRODUCTION

The circuit which is used to convert the direct current (DC) input signal into alternating current (AC) output signal is called as inverter. Traditionally there exist two types of inverters namely Voltage Source Inverter, Current Source Inverter. For dc to ac and ac to dc power conversion voltage source inverter performs buck and boost operation respectively. In the case of current source inverter it performs buck operation for ac to dc power conversion and boost operation for dc to ac power conversion that is both the V-source converter and the I-source converter have the common problem that they can operate either a boost or a buck converters and cannot be a buck–boost converter. The limitations of the traditional voltage source and current source converters can be eliminated by introducing Z-source inverter. Z-Source inverter utilize LC impedance network which performs both buck-boost energy conversions. It can be used in implementing dc-to-ac, ac-to-dc, ac-to-ac, and dc-to-dc power conversion. The Z-source network boosts the input voltage level for inverter and also provides filtered output. Thus this technique improves efficiency of inverter. The structure Z-source consist two inductors and two capacitors arranged in X shape for buck boost operation in single stage conversion. To avoid chopping in source current an additional LC filter is placed before diode therefore the cost of system would rise and due to addition of LC filter the system becomes more complex. In the case of EZ source inverter, source is places in series with inductors which smoothens the source current and chopping currents get filtered without any additional LC filter. For system design of EZ-Source inverter two PV panels are required. These PV panels work as a dc source which embedded within LC impedance network. The two DC source generate varying DC voltages whose values are depend on atmospheric conditions. Voltage and current filtering can be done by using LC components.

Cite this article as: Vinay Y. Somwanshi, Syed A. Naveed. "Simulation and Design of Solar Feed EZ-Source Inverter." *International Conference on Information Engineering, Management and Security (2015)*: 307-311. Print.

II. Z-SOURCE INVERTER

Fig.1. shows the circuit of Z-source inverter where the input dc source and three-phase inverter bridge is connected by X-shape network which is LC impedance network. Any two switches from the same phase-leg of inverter can be turned on safely to introduce a shoot-through state which is nothing but the short-circuit state. This condition can be brought by addition of LC impedance network. The inductive element ($L1$, $L2$, or both) used to limit the current paths from dc front-end. Due to insertion of shoot-through state, the Z-source inverter can provides voltage-boosting capability. Consider the inverter-state equations during shoot-through and non shoot-through states, expressed by (1) and (4) with a balanced network assumed ($L1 = L2 = L$ and $C1 = C2 = C$). These equations are used to derive the gain of inverter.

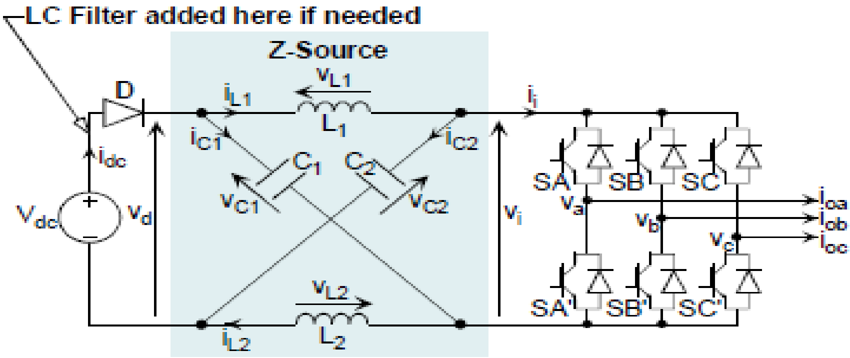


Fig.1. Z-source inverter

Conditions of Shoot-Through state:

(Sx = Sx1 = ON, x = A, B, or C; D = OFF)

$vL = VC$; $vi = 0$; $vd = 2VC$; $vD = Vdc - 2VC$---- (1)

$iL = - Ic$; $ii = iL - iC$; $idc = 0$---- (2)

Conditions of Non Shoot-Through state:

(Sx $\neq$Sx1, x = A, B, or C; D = ON)

$vL = Vdc - VC$;$vi = 2VC - Vdc$; $vd = Vdc$; $vD = 0$ ---(3)

$idc = iL + iC$; $ii = iL - iC$; $idc \neq 0$--- (4)

The following expressions are for VC capacitive voltage, peak dc-link voltage ^vi, and peak ac output voltage vx.

$VC = (1 - T0/T)/(1 - 2T0/T) * Vdc$-----(5)

$\hat{}vi = Vdc / (1 - 2T0/T) = BVdc$------(6)

$vx = M\, vi/2 = B(M\, Vdc/2)$------(7)

where T0/T denotes the shoot-through ratio ($T0/T < 0.5$) per switching period, M is the modulation index used for traditional inverter control, and B is boost factor given by expression $B = 1/(1 - 2T0/T)$.

III. EMBEDDED EZ-SOURCE INVERTER

Fig.2. shows the circuit of Embedded EZ-Source inverter. It consist L1 and L2 inductive elements and C1 and C2 capacitive elements which forms LC impedance network.

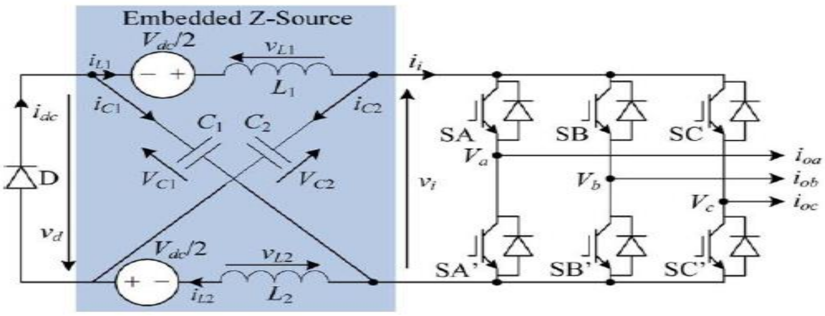


Fig.2. EZ-source inverter

Two DC sources embedded within the X-shaped LC impedance network. In the case of voltage type EZ-source inverter inductive elements L1 and L2 used for filtering the currents and in case of current type EZ-source inverters capacitive elements C1 and C2 used

Cite this article as: Vinay Y. Somwanshi, Syed A. Naveed. "Simulation and Design of Solar Feed EZ-Source Inverter." *International Conference on Information Engineering, Management and Security (2015)*: 307-311. Print.

for voltage filtering in current type EZ-source inverters keeping the voltage or current gain of the inverter constant. The input DC can be taken from the solar cell given to the Z-source. The filtered, ripple free, pure DC given to the three phase inverter which converts pure DC into three phase balanced AC.

IV.EXPERIMENTAL RESULTS

The simulation of proposed system is done in MATLAB SIMULINK. The output voltage from two DC sources fed to the Z filter of Embedded EZ-Source Inverter. The filtered DC output from Z filter is given to the Embedded EZ-Source Inverter which converts pure DC into three phase balanced AC.

The simulation circuit diagram of Embedded EZ-Source inverter is shown in fig.3. The triggering pulses for switches of Embedded EZ-Source inverter is shown in fig.4. The line current waveform is shown in fig.5, fig.6 and fig.7 represents line voltage waveforms. The output from the Embedded EZ-Source Inverter is given to the induction motor which is asynchronous machine. Three phase induction motor is taken as a load. Fig.8 shows the plot of rotor speed of induction motor with respect to time. It represents the speed of induction motor increases and settles down above 1600 RPM (Revolution Per Minute). For calculation of total harmonic distortion FFT analysis for current of the system is taken. It shows that THD value is 3.93% shows in fig.9. Thus total harmonic distortion is low in Embedded EZ-Source inverter system

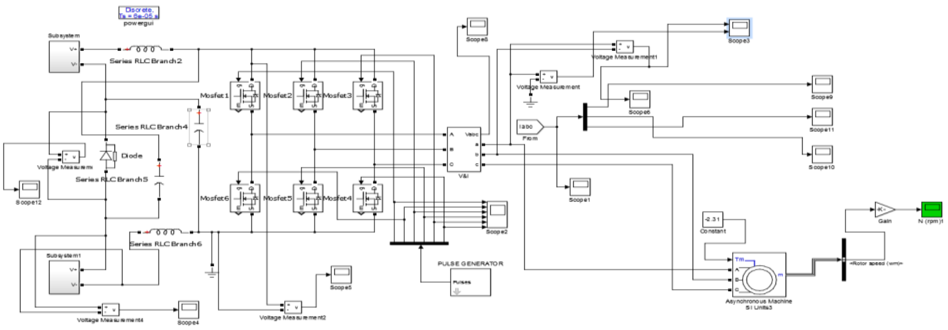


Fig.3. Simulation circuit diagram

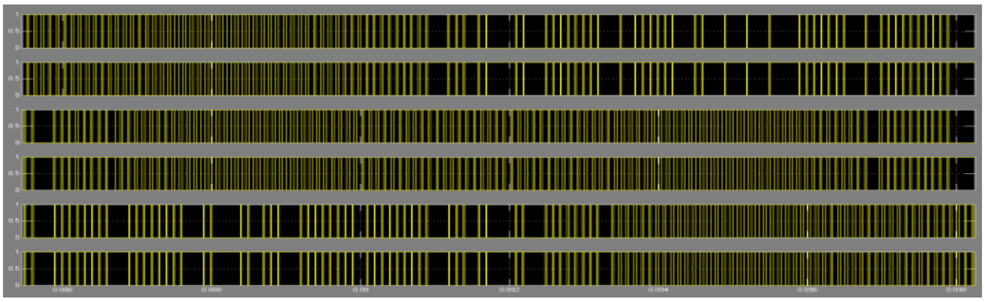


Fig.4. Triggering pulses

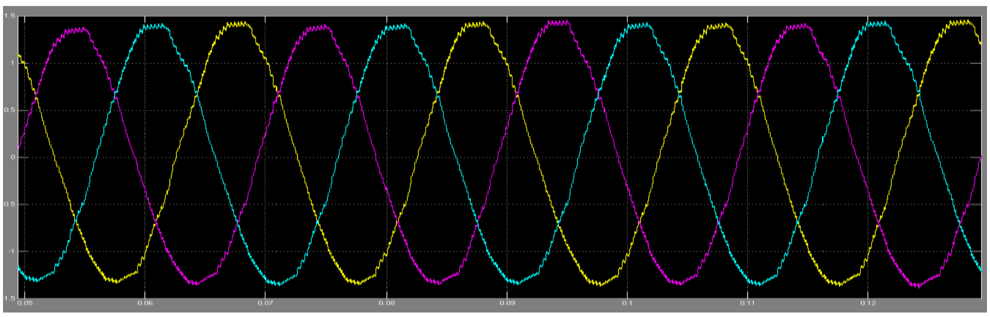


Fig.5. Line current waveform

Cite this article as: Vinay Y. Somwanshi, Syed A. Naveed. "Simulation and Design of Solar Feed EZ-Source Inverter." *International Conference on Information Engineering, Management and Security (2015)*: 307-311. Print.

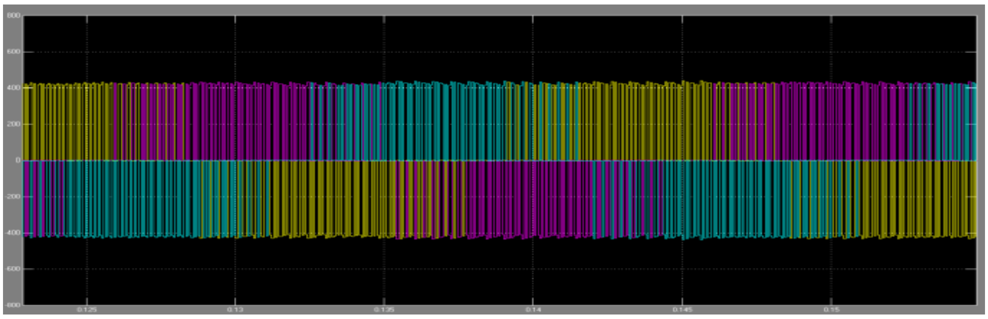


Fig.6. Line voltage waveform

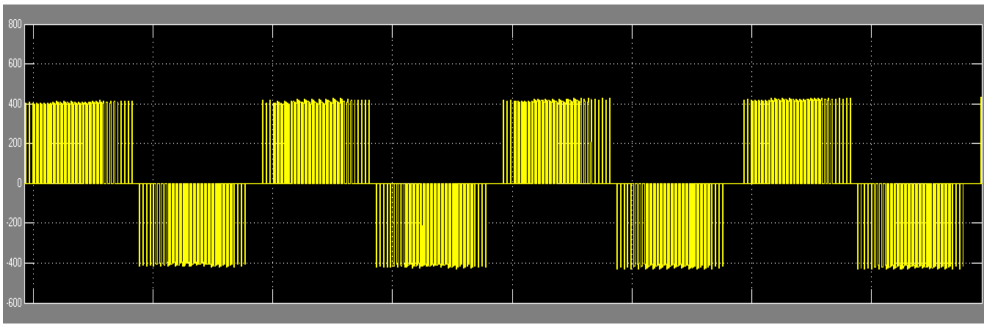


Fig.7. Phase to Phase voltage waveform

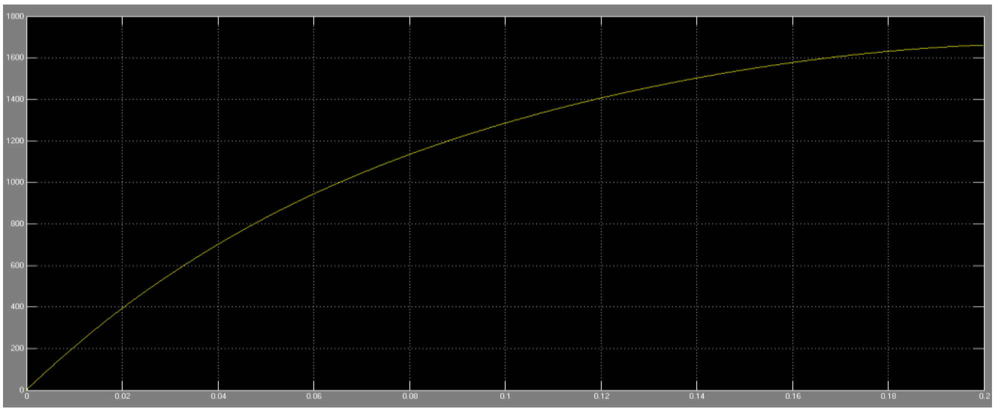


Fig.8. Motor Rotor speed

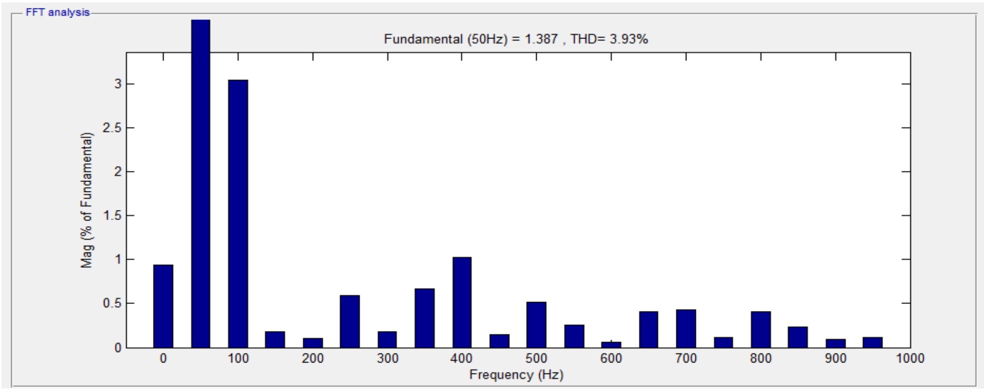


Fig.9. FFT analysis and total harmonic distortion (THD)

Hardware Setup of EZ-Source Inverter

Cite this article as: Vinay Y. Somwanshi, Syed A. Naveed. "Simulation and Design of Solar Feed EZ-Source Inverter." *International Conference on Information Engineering, Management and Security (2015)*: 307-311. Print.

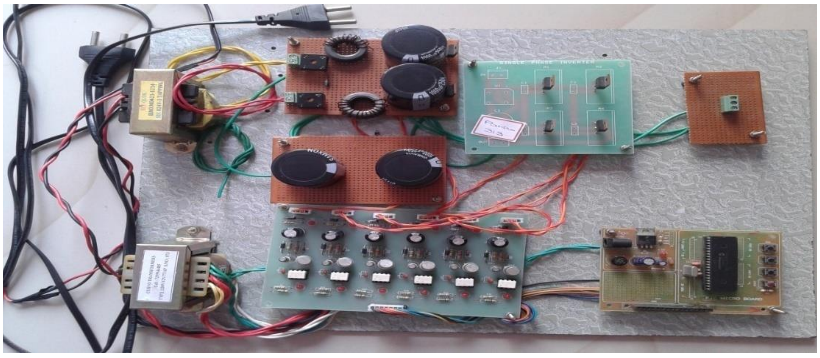


Fig.10 Proposed system hardware

Fig.10 shows the complete hardware setup of EZ-source inverter. It consist power supply unit, Z-source impedance network, induction motor as a load, driver circuit for inverter and 6 switches three phase inverter. Projections for solar input are provided in the system. DC Input for the Z-source is provided by the solar plates, filtered DC from Z-source fed to three phase inverter which generate balanced AC to run induction motor. The Microcontroller PIC16F877A used to generate control pulses for three phase inverter. The driver circuit is used to amplify these triggering pulses and given to the inverter. The balanced three phase ac provided by three phase inverter which then given to the motor loads.

V. CONCLUSION

The simulation of EZ-Source inverter system is carried out in MATLAB SIMULINK software. From the experimental results are presented in this paper, we observed that the given system provides voltage boosting and current filtering using impedance network also it has advantages of reduced harmonic distortions of about 3.93%. Its hardware implementation has been done and results are obtained by using PV panels at input and motor load at the output. Thus EZ-source inverter has great advantages over current source and voltage source inverters.

VI. ACKNOWLEDGMENT

I wish to thank my project guide Dr. S. A. Naveed for providing me a great technical support with his knowledge and experience also I would like to express my gratitude to my colleagues for their kind co-operation and encouragement which help me in completion of this project.

REFERENCES

[1] S. Kamalakkannan, S. Deve gowda, “Digital simulation of renewable energy source controlled ez-source Inverter system”, 2012 IEEE.

[2] P.Kannan, “Harmonic Analysis and Design of Embedded Z-Source Inverter for Induction Motor Drives”, Volume-4 Number-1 Issue-14 March-2014.

[3] K.Spandana, G.N.Sreenivas, “Performance analysis of partially parallel

[4] Embedded z (ez) - source inverter with reduced switches fed im drives”, volume- 1, issue-3, jan.-2014.

[5] https://en.wikipedia.org/wiki/Z-source_inverter.

[6] Blaabjerg, Frede; Loh, Poh Chiang; Gao, F., “Embedded EZ-Source Inverters”, (pp. 1-8). IEEE. 10.1109/08IAS.2008.300.

[7] N.Gurusakthi, R.Sivaprasad, “Performance Enhancement of EZ-Source Inverter Using Induction Motor”, Volume 4, Issue 4, April-2013.

[8] Fang Zheng Peng, “Z-Source Inverter”, vol. 39, no. 2, march/april 2003.

[9] https://en.wikipedia.org/wiki/LC_circuit

Cite this article as: Vinay Y. Somwanshi, Syed A. Naveed. “Simulation and Design of Solar Feed EZ-Source Inverter.” *International Conference on Information Engineering, Management and Security (2015)*: 307-311. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS052

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.052

Measurement and Analysis of Air Pollution

Manohar R.Bodkhe[1], Dr. R.D. Kokate[2]
[1,2] Department of Electronics and Telecommunication
MGM's JNEC Aurangabad, Maharashtra ,India.

ABSTRACT: *In this paper I have developed prototype module using semiconductor sensors at the emission release of vehicles which detects the level of pollutants and also indicates this level with a meter. When the pollution emission level shoot the breeze which is already set to threshold level, there will be a buzz in the vehicle to indicate that the limit has been violated and the vehicle will stop after.*

Keywords: Air Pollution Sensors, Threshold Level, GPS, GSM, Solenoid valve, Atmega328.

I. INTROCUTION

In new era of the 21st century there is a time where the importance for Environmental awareness is incite. One of the main reasons regarding the environment is air pollution which is insidious to all living beings on earth as well as environment . Air pollution imparts to the green houses gases, which causes the green house effect, whose side effects are now well known to all of us after the findings about the hole in the ozone layer. Air pollutants that have serious impact on human health affecting the respiratory system and lungs, they are also carried to the blood and pumped all around the body. These pollutants are also stuck on soil, plants, and in the water, further contributing to human vulnerability and also affecting the sea life. Apart from industries vehicles are the major contributors to air pollution. Carbon dioxides and nitrogen are the main pollutants from vehicles which can be easily detected these days with the help of semi conductor gas sensors. Therefore this paper gives an idea to reduce the air pollution from vehicles which would be very helpful to us.literature survey gives the background information and a brief note about the various research activities, on gas sensors and monitoring systems that can be run at various places and in vehicles. Section II discusses about the various blocks of the proposed system. Section IV concludes the paper with an idea to implement the same as a real time project.

LITERATURE SURVEY

The first emission norms were introduced in India in 1991 for petrol and 1992 for diesel vehicles. These were followed by making the Catalytic converter mandatory for petrol vehicles and the introduction of unleaded petrol in the market. All new vehicles manufactured after the implementation of the norms have to be compliant with the regulations with upgrade of embedded modules . In India since October 2010, Bharat stage III norms have been enforced for all kinds of automobile sector . In 13 major cities, Bharat stage IV emission norms are in place since April 2010. The phasing out of 2 stroke engine for two wheelers, the stoppage of production of various old model cars & introduction of advanced electronic systems with automatic control strategies have been due to the regulations related to vehicular emissions. The standards, based on European regulations were first introduced in 2000.

Cite this article as: Manohar R.Bodkhe, Dr. R.D. Kokate. "Measurement and Analysis of Air Pollution." *International Conference on Information Engineering, Management and Security (2015)*: 312-317. Print.

Progressively stringent norms have been rolled out since then. The standardized values for the emission levels are referred as given in [1]. The sensing of the emitted gases is done using various sensors and devices. The past decade, has seen several research activities that have been taking place to develop semiconductor gas sensors [2].

II. PROPOSED SYSTEM

Focusing on the main parameters like CO, CO2 and NO2 A system with advanced electronic controller At mega328 along with GSM,GPS System, solenoid valve and semiconductor sensors is built. In the paper [3], the quality of air in the car cabin was analyzed using semiconductor (MOS) gas sensors. In this paper, the semiconductor sensors have been used to detect the pollutant level of the vehicles. This paper concentrates mainly on three blocks; smoke detector, microcontroller and fuel injector. The smoke detector detects the pollutants (CO, NOx, etc.) continuously. The microcontroller compares the level of pollutants with the stipulated level allowed by the government. When the pollutant level exceeds the standardized limit, it sends a signal to the fuel injector. On receiving a signal from the controller, the fuel injector stops the fuel supply to the engine after a particular period of time.

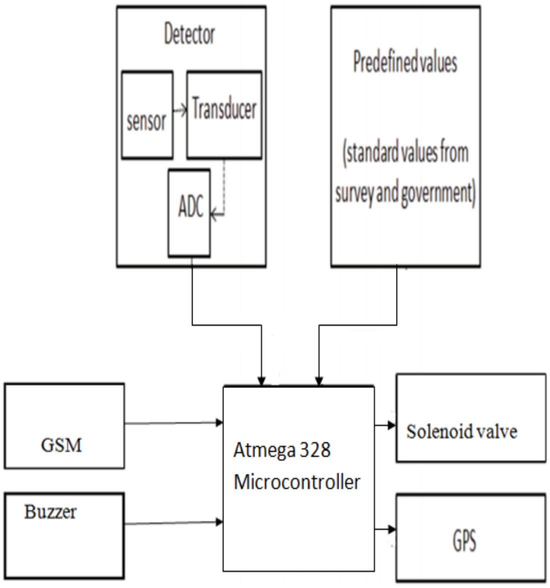


Figure.1.Block diagram proposed system

1.Microcontroller ATMEGA 328

In this paper, At mega 328 is used, which is an 8bit micro controller. It consists of three inbuilt timer/counter which will be used for the timer configuration. The microcontroller is programmed to do three functions namely comparison, timer and triggering circuit. The microcontroller takes in two inputs; one from the smoke sensor's output and another being the pre-defined threshold value specified by the government. When the smoke sensor output is more than the threshold value, the microcontroller triggers the timer circuit and an alarm is set off to inform the driver of the vehicle, about the same and also indicate that the vehicle will come to a halt as soon as the timer runs out.

Apart from the timer being triggered, a trigger is also given to the GPS, which helps in locating the nearest service station. Once the timer runs out, a trigger pulse is generated by the microcontroller which is fed to the fuel injector, which in turn stops the flow of fuel to the engine, as a result of which, the vehicle comes to a halt.

2. Alpha-numeric LCD display

A liquid crystal display (LCD) is a flat panel display, electronic visual display, based on Liquid Crystal Technology. A liquid crystal display consists of an array of tiny segments (called pixels) that can be manipulated to present information. Liquid crystals do not emit light directly instead they use light modulating techniques.

3. Global Position System for Mobiles (GSM)

Global system for mobile communication (GSM) is a globally accepted standard for digital cellular communication. GSM is the name of a standardization group established in 1982 to create a common European mobile telephone standard that would formulate specifications for a pan-European mobile cellular radio system operating at 900 MHz. It is estimated that many countries outside of Europe will join the GSM partnership.

Cite this article as: Manohar R.Bodkhe, Dr. R.D. Kokate. "Measurement and Analysis of Air Pollution." *International Conference on Information Engineering, Management and Security (2015)*: 312-317. Print.

4. Global positioning system (GPS)

GPS-634R is a highly integrated smart GPS module with a ceramic GPS patch antenna. The module is with 51 channel acquisition engine and 14 channel track engine, which is capable of receiving signals from up to 65 GPS satellites and transferring them into the precise position and timing information that can be read over either UART port or RS232 serial port. Small size and high end GPS functionality are at lower power consumption, both of the LVTTL-level and RS232 signal interface are provided on the interface connection.

5. Smoke sensor

The detector consists of blocks namely smoke sensor, transducer . The smoke sensor is the main component of the detector which is embedded on to the exhaust of the vehicle. In this paper, carbon monoxide sensor (MQ-7) which can measure CO concentrations ranging from 10 to 10,000 ppm is considered. This sensor, basically finds usage in sensing carbon monoxide concentrations in (ppm), in the exhaust of cars as shown in figure. 5 and gives an analog output. The MQ -7 gas sensor is mainly made up of SnO_2, whose conductivity varies with the cleanliness of air i.e. it has a lower conductivity in clean air and vice versa.

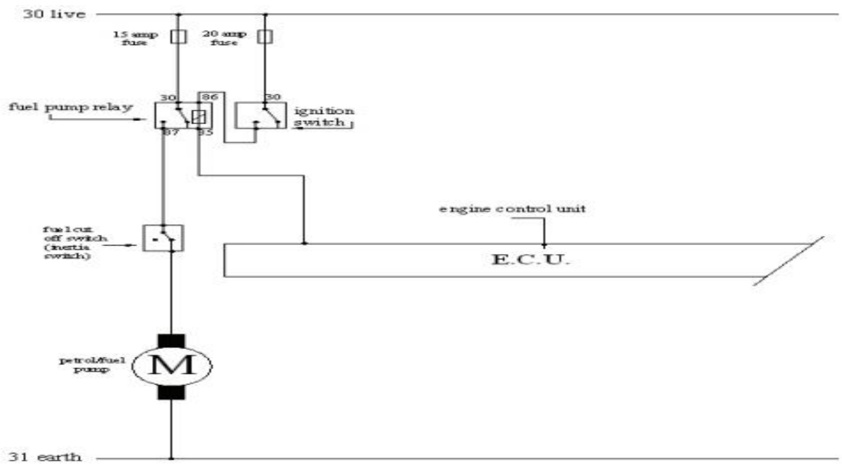


Figure 2: Fuel pump wiring circuit

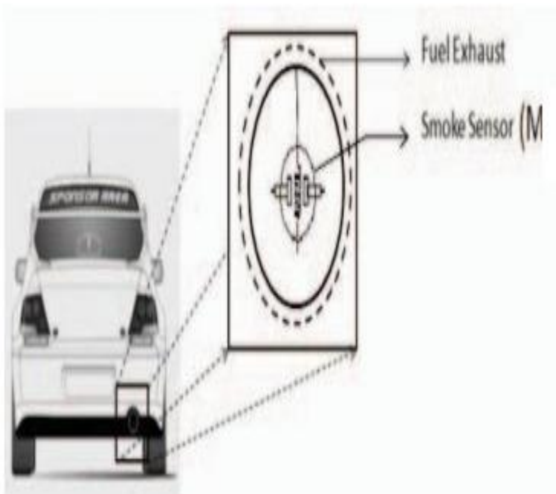


Figure 3: Smoke sensor

Cite this article as: Manohar R.Bodkhe, Dr. R.D. Kokate. "Measurement and Analysis of Air Pollution." *International Conference on Information Engineering, Management and Security (2015)*: 312-317. Print.

Measurement and Control: Flowchart

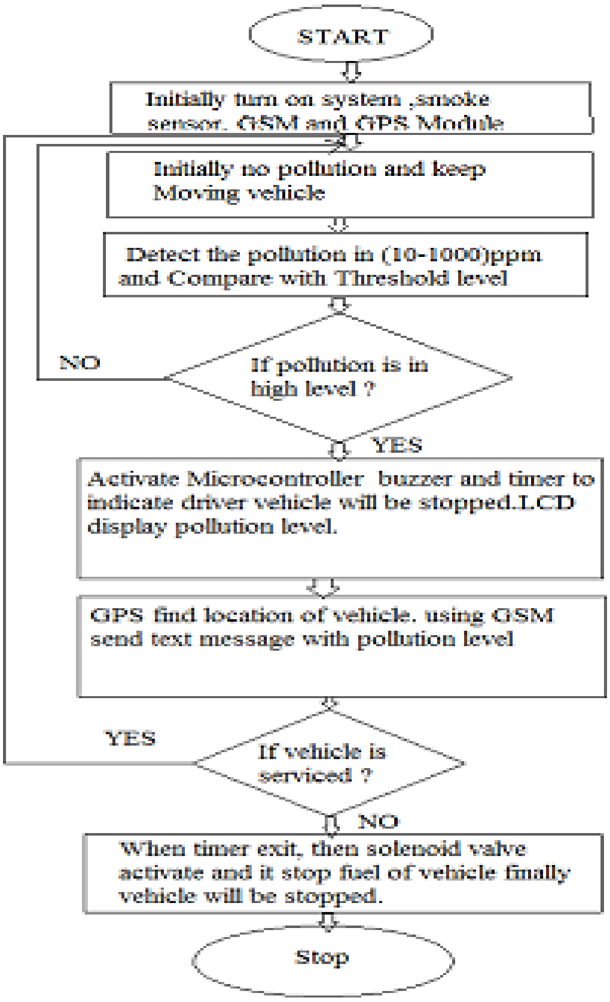


Figure 4: Flowchart of working sequence

Hardware of corresponding proposed system:

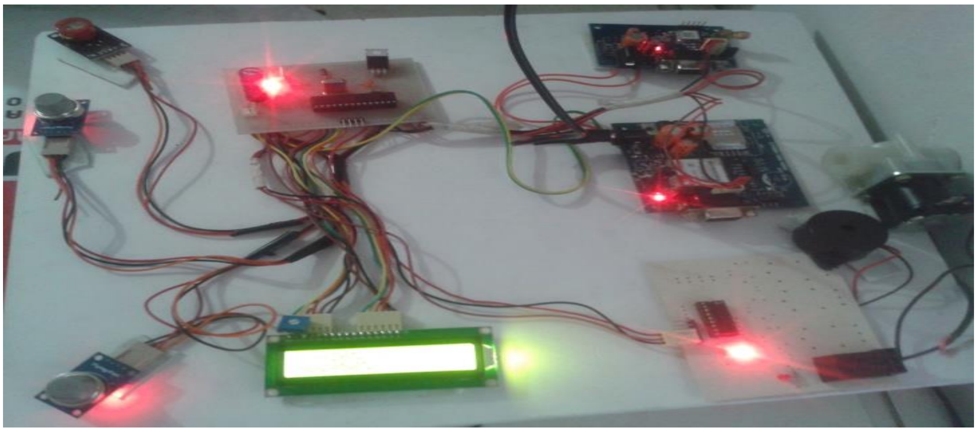


Figure 5. Hardware of proposed system

Cite this article as: Manohar R.Bodkhe, Dr. R.D. Kokate. "Measurement and Analysis of Air Pollution." *International Conference on Information Engineering, Management and Security (2015)*: 312-317. Print.

IV. RESULT AND ANALYSIS:

In this paper I have detected air pollution at Aurangabad city in Maharashtra from India and Corresponding readings for different locations in Aurangabad city has been collected with this project and Readings are shown as below in table (a) with suitable multicolored plot with figure (b) and Result of GSM Messages which show latitude and longitude location as well as Air pollution Co, Co2 and No2 in (c) and LCD images in figure (D).

Sr.no.	CO (PPM)	CO2 (PPM)	NO2 (PPM)	CORDINATES LOCATION (PPM)	PLACE Aurangabad
1	384	47	67	19.869169 75.331657	Railway station
2	25	46	157	19.861160 75.310203	Kanchan wadi Corner
3	355	37	56	19.854396 75.302116	Waluj MIDC
4	496	42	44	19.837240 75.253410	Baba Petrol pump
5	20	279	45	19.874151 75.315330	Sant Ekanath Rangmandir
6	18	28	178	19.869129 75.331680	Kranti chowk
7	15	83	155	19.885116 75.358207	Central naka

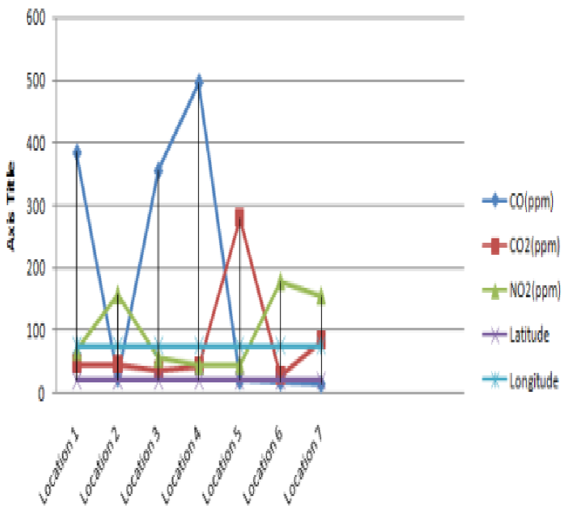


Figure 6.(A): Readings are shown in table

Figure 6.(B): Result of multicolored plot

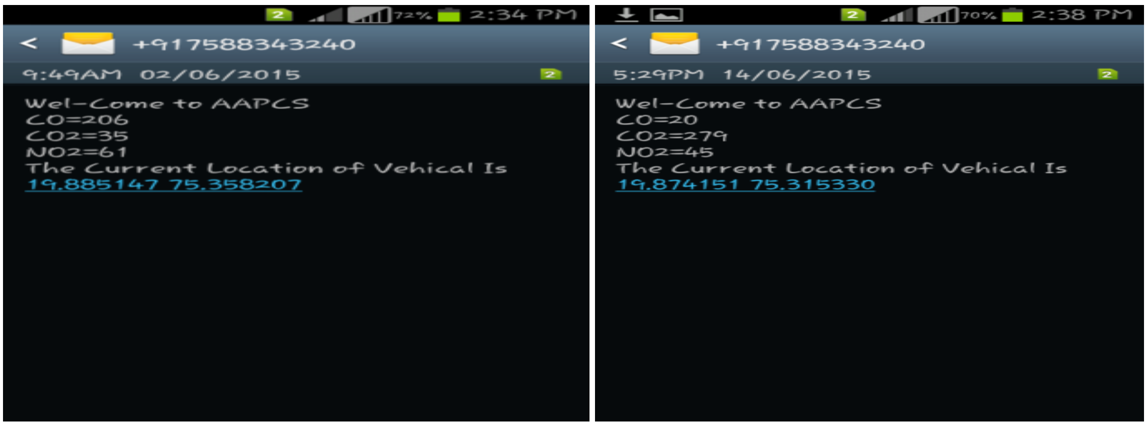


Figure 6. (C): Result of GSM Message

Cite this article as: Manohar R.Bodkhe, Dr. R.D. Kokate. "Measurement and Analysis of Air Pollution." *International Conference on Information Engineering, Management and Security (2015)*: 312-317. Print.

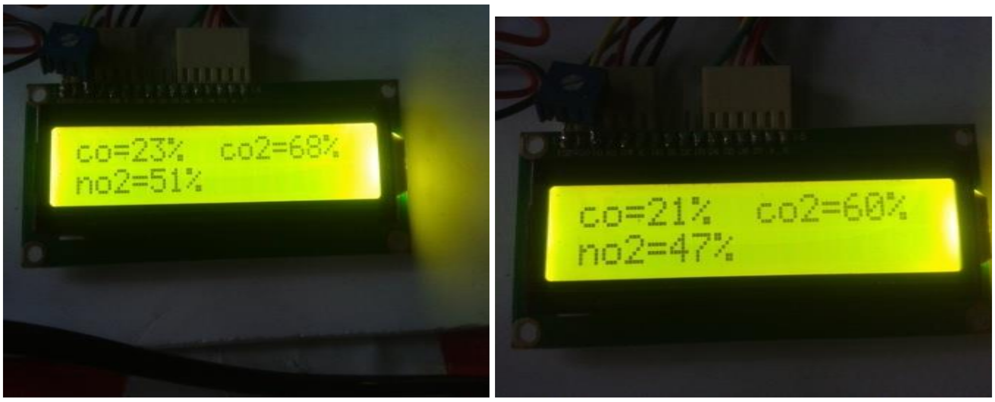


Figure 6. (D): Result of LCD images

V. CONCLUSION

In this paper here main focus is given on two basic conceptual things. The detection of pollution level and its indication to driver being the first. The remote user notification via GPS and GSM.Technology with latest modules this system can be integrated on mass scale to fulfill the dream of the concept 'smart city'. Along with the time much more advance controller and its subsidiary interfacing devices can be upgraded this technology can also be embedded with internet as a part of 'Internet of kings ' concept and using this system Minimize the air pollution .

VI.ACKNOWLEDGMENTS

The author would like to offer special thanks to mentor Dr.R.D.Kokate from our institute for directing the special concerns related with this topic and organizing this issue .I also owes great deal to my colleagues for contributing their efforts in enthusiastic manner

REFERENCES

[1] Siva Shankar Chandrasekharan, Sudharshan Muthukumar & Sabeshkumar Rajendran " Automated Control System for Air Pollution Detection in Vehicles" 2013 4thinternational Conference on Intelligent Systems, Modeling and Simulation, 2013 IEEE 2166-0662/13.
[2] George F. Fine, Leon M. Cavanagh, Ayo Afonja and Russell Binions " Metal Oxide Semi-ConductorGas Sensors in Environmental Monitoring", Sensors 2010, 10, 5469-5502; doi:10.3390/s100605469
[3] K. Galatsis, W. Wlodarsla, K. Kalantar -Zadeh and A. Trinchi, "Investigation of gas sensors for vehicle cabin air quality monitoring," vol. 42, pp. 167-175, 2002.
[4] "Trade of Motor Mechanic"; Module 5; Unit 2 Electronic Fuel injection; Phase 2 by FÁS Learning Innovation Unit with Martin mcmahon & CDX Global; Curriculum Revision 2.2 16 -01-07.
[5] LIU Zhen-ya, WANG Zhen-dong, CHEN Rong, "Intelligent Residential Security Alarm and Remote Control System Based On Single
[6] chipcomputer,"vol.42,pp. 143-166,2008. Digiinternational inc,"xbee/xbee-proper modules", available .
[7] http://ftp1.Digi.com/support/documentation/90000982_b.pdf
[8] Atmegacorporation,"atmega328datasheet", available
[9] Http://www.atmel.com/dyn/resources/prod_documents/doc2466pdf
[10] http://wikipedia.org/wiki/Bharat_Stage_emission_standards

Cite this article as: Manohar R.Bodkhe, Dr. R.D. Kokate. "Measurement and Analysis of Air Pollution." *International Conference on Information Engineering, Management and Security (2015)*: 312-317. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS053		**eAID**	ICIEMS.2015.053

Smart-phone Based Home/Office Automation With Environment Monitoring

SUYOG SHRIPAD PANDE[1], AARTI R. SALUNKE[2]
[1]M.E. Student, [2]Assistant Professor,
[1,2] Jawaharlal Nehru Engineering College, Aurangabad, (M.S.), India

Abstract: *This paper presents the design of Smart-phone Based Home/Office Automation System with environment monitoring with low cost and wireless remote control. General idea of home/office automation shows the quality of human being at office/house. Prime focus of this technology is to control the office/household equipment's like light, fan, door etc. automatically. In hazardous condition, it is useful for old aged and handicapped persons. Also, the smart home concept in the system improves the standard living at Office/home. The main control system implements wireless Bluetooth technology to provide remote access from PC/laptop or smart phone. The design remains the existing electrical switches and provides more safety control on the switches with low voltage activating method. The system intended to control electrical appliances and devices in office/house with relatively low cost design also for physical environment monitoring, user-friendly interface and ease of installation.*

Keywords: LPC2148 development board, Bluetooth device, Smart phone, Sensors, Controlled devices.

I. INTRODUCTION

Focusing on the use of home area networks to improve disabled people's autonomy at home, In this paper presents a smart phone based accessible office/home appliance control [1].In recent years the popularity of home/office automation has been increasing due to higher affordability and simplicity by connecting through Smart-phone. Home automation include controlling of lights, fans, appliances, security locks for gates and doors, etc., which are used to improve comfort, energy efficiency and security for office/home. Office/Home automation is useful for elderly and disabled, who can control the things by staying at one place without the help of others and can increase the life quality of them. Office/home automation system provides the integration among all the electrical and electronic devices in office/house. The techniques used in office/home automation systems include controlling of electronic and electrical devices, such as entertainment systems, security systems, air conditioners, lawn watering systems, domestic robots, etc., As information technology has been integrated with the office/home appliances and systems, they are able to communicate in an integrated manner which results in energy saving and safety benefits. As the wireless technology is emerging day by day, several different connections are introduced such as Bluetooth, WIFI, ZIGBEE and GSM. Each of these connections has their unique specifications. Among the above mentioned wireless connections, Bluetooth is chosen with its suitable capabilities for designing this project. Bluetooth with globally available frequencies of 2400Hz is able to provide connectivity up to 100 meters and speed up to 3Mbps depending on different Bluetooth device classes [2].

Cite this article as: SUYOG SHRIPAD PANDE, AARTI R. SALUNKE. "Smart-phone Based Home/Office Automation With Environment Monitoring." *International Conference on Information Engineering, Management and Security (2015)*: 318-322. Print.

II. LITERATURE REVIEW

Based on the study of HAS project done by researchers and developers, [4] implemented Micro-controller in wireless HAS. For wireless connection, the system implemented a FM transmitter and receiver to establish a RF connection. The simplex connection between control board and controller limited that only one type of input (voice) to the system. [5] Implemented GSM, Internet and voice as wireless HAS. The system implemented microprocessor and GSM SMS control method by a GSM modem. The system [5] mentioned as low cost but the cost of GSM modem and micro-controller is not considered. Also, long term cost by the GSM is not fully accepted by every user. Project [3], [6-10] are Bluetooth based HAS design architecture. Where reference [6-8] proposed a Bluetooth based HAS that controls home appliances by a PC's GUI, but it does not provide portable remote function. For system [6-8], all the controls are performed only at the GUI on PC. [3], [9-10] are designed with cellular phone remote control to the system. Reference [3] implemented Arduino Bluetooth board in their HAS project with cell phone remote control. The project stated as low cost HAS system but the cost of Arduino BT board is not the best cost efficient solution. Moreover, the cell phone control is implemented by Symbian OS application. It does limit the users of the system as the Symbian based cell phones in market nowadays are very less. While reference [9] did not mentioned the specific type of phone's OS implemented for their phone application. Meanwhile reference [10] mentioned the phone control is designed in JAVA application but it also did not mention the specific phone's OS for the application. From the overall papers reviews, HAS according to [3-10] never mentioned about the existing physical electrical switches in their system.

III. SYSTEM OVERVIEW

Figure 1 shows the block diagram of the Smart-phone based Office/Home Automation with Environment Monitoring System i.e., control function of the system. The system is directly connected to the electrical and electronic devices present in the home such as fan, light, door etc., The Bluetooth connection is established between the system and the application which was designed and installed in the Android device.

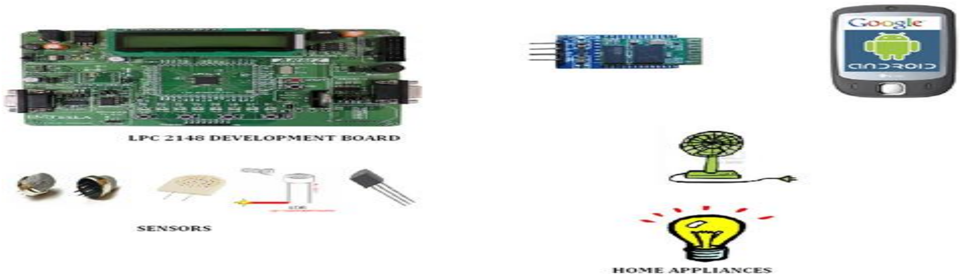


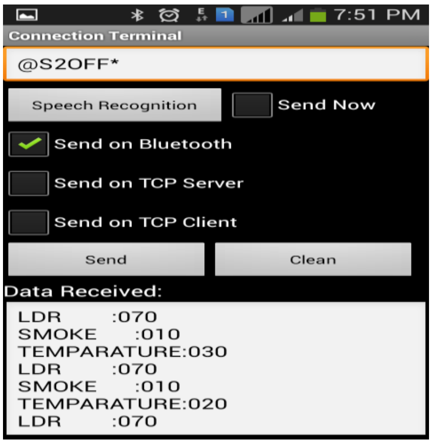
Fig 1: Functional Block Diagram of System

In order to improve the standard of living, the controlling of the home appliances is done by the Android application installed in Android device. The users can easily access the Android application by giving the commands on the touch-screen of Android device. This method is very much useful for the persons who are physically disabled and can't move on their own to the switches to turn on the appliances. The temperature, LDR and Smoke values can be measured using the sensors that are connected to the main control board. The indication from the sensors reminds or helps the user to turn on/ off the fan in the house. The on/off status of office/home appliances, temperature, LDR and Smoke readings are synchronized with the Android application present in the Android device. The monitoring of sensor reading is done in real-time; any changes in the sensor readings will be transmitted to the Android application present in Android device.

IV. HARDWARE DESIGN

In this section we discuss about the hardware construction of the main control board. Figure 2 shows the hardware blocks present in main control board. ARM7 Micro-controller, ARM7 is considered for designing of this hardware due to its capability of performing serial communication using Bluetooth connection with the Android device. As we know, the temperature, Smoke and LDR sensors are considered for getting the temperature, Smoke and LDR levels in the room. The Bluetooth module, BLUETOOTH BT24LT is chosen for establishing the connection between the Android device and the main control board due to its low cost. The electrical current is directly connected to the main control board. The voltage regulator is constructed by Zero Crossing detection and optocopuler circuit which consists of transformer, rectifier and regulator. 3.3V to 5V DC output is needed for the specific components in the main control board.

Cite this article as: SUYOG SHRIPAD PANDE, AARTI R. SALUNKE. "Smart-phone Based Home/Office Automation With Environment Monitoring." *International Conference on Information Engineering, Management and Security (2015)*: 318-322. Print.



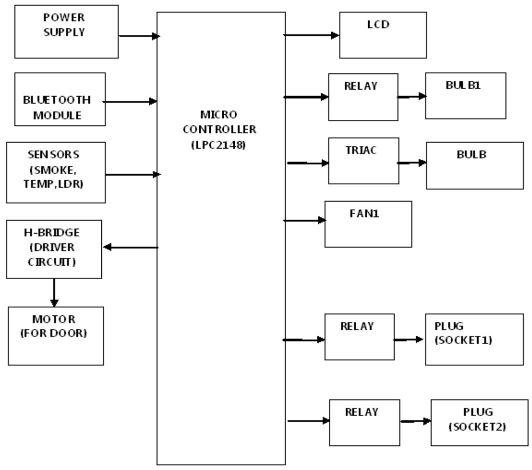


Fig 2: Main Control Board Hardware Block Diagram

The system designed is directly installed beside the electrical switches on the wall. The installation of this systems does not need any wiring re-installation and wiring on the wall, but the existing switches in directly connected to the OptoCopular circuit inside the main control board. Depending on the requirement, multiple control boards can be installed in home. With these low cost components, the main control board is constructed in small size but still performs the strong functions of the system.

V. SOFTWARE DESIGN

Software design section is divided into two sections (1) Main function of the system designed in ARM7 LPC2148 micro-controller and (2) Designing of Android application. Figure 3 illustrates the control flow in ARM7 LPC2148 micro-controller. The input to the main control board is detected by ARM7 LPC2148 micro-controller. Any input to ARM7 micro-controller will cause an interrupt to the main function loop of ARM7 LPC 2148. This will cause a change in the output peripherals connected to main control board.
The Android application is designed using Eclipse, ADK and JDK. Figure 4 illustrates the Android application i.e., installed and tested using the Android device which has Android 4.1.2.The application is simple to use, user can turn on and off the appliances that are connected to main control board by simply giving commands.

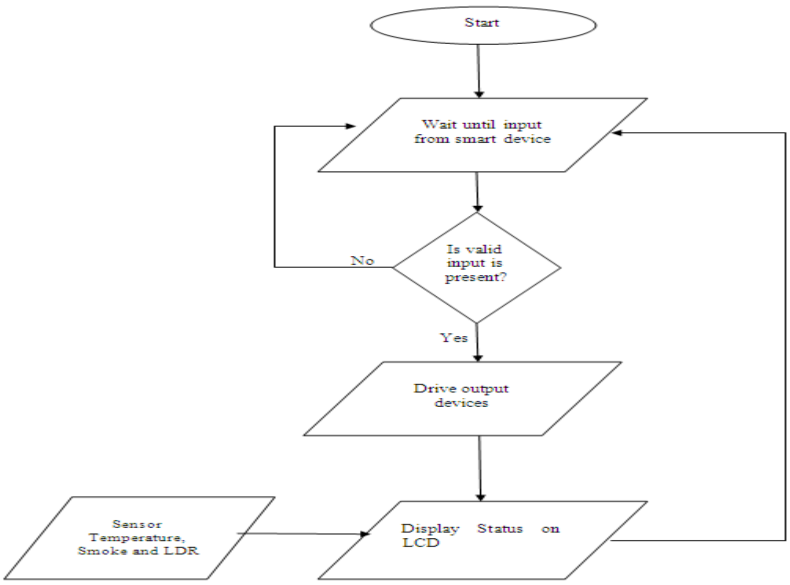


Fig 3: Control Flow in ARM7 LPC2148 Microcontroller

Cite this article as: SUYOG SHRIPAD PANDE, AARTI R. SALUNKE. “Smart-phone Based Home/Office Automation With Environment Monitoring.” *International Conference on Information Engineering, Management and Security (2015)*: 318-322. Print.

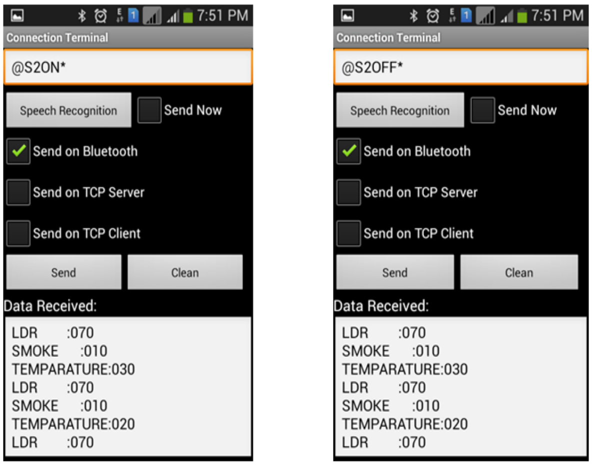


Fig 4: Android application testing

VI. RESULTS

This system is tested and verified in the real time environment. The below pictures will you understand how perfectly the system is working. Picture 1 is taken when the system is turned off. When the system is turned on then the bulb glows with the low intensity as show in picture 2. When we need to increase the intensity of light then we have to give command on Android Application then the intensity of the light will change. Pictures 3 show the bulb intensity at higher level. In the similar way, we can also control the Fan speed. The temperature, Smoke and LDR values are displayed on the LCD present in the system and also in the Android Application.

Fig5: Output of System

VII. CONCLUSION

In conclusion, this system is designed at low cost and is used to improve the standard of living in office/home. The wireless connectivity through the Android device provides help to the people especially to elderly and disabled. The implementation of the Bluetooth connection in control board allows the system to install in simple way. The control board can be directly installed besides the electrical switches. For future work, the Android application will be implemented with speech recognition to control appliances with voice commands. All the voice commands given to the Android device will be transmitted to the main control board after signal processing. All the future work can be implemented on the same system by changing the application in the Android device.

Cite this article as: SUYOG SHRIPAD PANDE, AARTI R. SALUNKE. "Smart-phone Based Home/Office Automation With Environment Monitoring." *International Conference on Information Engineering, Management and Security (2015)*: 318-322. Print.

REFERENCES

[1] Laisa C.P. Costa, Nicholas S.Almedia, Ana G.D. Correa "Accessible Display Design to Control Home Area Network", IEEE Transaction on consumer Electronics,Vol. 59,No 2,May 2013.

[2] Bluetooth official website http://www.bluetooth .com

[3] R. Piyare and M. Tazil, "Bluetooth Based Home Automation System using Cell Phone," in Consumer Electronics, pp. 192-195, 2011.

[4] Kailash Pati Dutta, Pankaj Rai and Vineet Shekher, "Microcontroller Based Voice Activated Wireless Automation System: VSRD-IJEECE", Vol. 2(8), 642-649, 2012.

[5] Baris Yuksekkaya, A. Alper Kayalar, M. Bilgehan Tosun, M. Kaan Ozcan, and Ali Ziya Alkar, "A GSM, Internet and Speech Controlled Wireless Internet Home Automation System", IEEE Transactions on Consumer Electronics, Vol. 52, No. 3, AUGUST 2006.

[6] N.Sriskanthan and Tan Karande, "Bluetooth Based Home Automation Systems," Journal of Microprocessors and Microsystems, Vol. 26, pp. 281-289, 2002.

[7] Kwang Yeol Lee & Jae Weon Choi, "Remote-Controlled Home Automation System via Bluetooth Home Network" in SICE Annual Conference in Fukui, Vol. 3, pp. 2824-2829, 2003.

[8] Wijetunge S.P., Wijetunge U.S., Peiris G.R.V, Aluthgedara C.S. & Samarasinghe A.T.L.K., "Design and Implementation of a Bluetooth based General Purpose Controlling Module", in IEEE, pp. 206-211, 2008.

[9] Sandeep Kumar & Mohammed A Qadeer, "Universal Digital Device Automation and Control", in IEEE, pp. 490-494, 2009.

[10] Hiroshi Kanma, Noboru Wakabayashi, Ritsuko Kanazawa & Hirimichi Ito, "Home Appliance Control System over Bluetooth with a Cellular Phone", in IEEE, pp. 1049-1053, 2003.

Cite this article as: SUYOG SHRIPAD PANDE, AARTI R. SALUNKE. "Smart-phone Based Home/Office Automation With Environment Monitoring." *International Conference on Information Engineering, Management and Security (2015)*: 318-322. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9
Website	www.iciems.in
Received	10 - July - 2015
Article ID	ICIEMS054

VOL	01
eMail	iciems@asdf.res.in
Accepted	31- July - 2015
eAID	ICIEMS.2015.054

ENHANCING MAP-REDUCE JOB EXECUTION ON GEODISTRIBUTED DATA ACROSS DATACENTERS

JEBILLA P[1], Dr P JAYASHREE[2]
[1]M.E Computer Science & Engineering
[2]Associate Professor, Dept of Computer Science and Technology
[1,2] Anna University-MIT Campus.

Abstract: *In Recent era the size of the data set we need to handle grows rapidly. Efficiently analyzing Big data has always been an issue in our current era. Cloud computing along with the implementations of MapReduce framework provides a parallel processing model and associated implementation to process huge amount of data. In cloud, in many scenarios the input data set are geographically distributed across datacenters. This paper deals with enhancing the MapReduce job execution on the geo distributed data. Possible execution paths are analyzed. A Data Transformation Graph is used to determine the schedules for the job sequences which are optimized using the Shortest Path Algorithm. The proposed model deals with the extending of the Existing Dijkstra's Algorithm to consider the node weight in addition to the edge weight. Ozzie workflow and mapper side joins are used to reduce the execution time and cost. As shown by the comparisons the execution time of the MapReduce job execution has been enhanced.*

Keywords: cloud computing, data center, map reduce, geodistribution

I. INTRODUCTION

Analyzing Big data is so complex that the traditional data processing applications are inadequate. We face many challenges when we need to analyze, capture, search, storage, sharing, transfer the large data sets. The data sets inherently arise due to the applications generating and retaining more information to improve operation and monitoring. Many applications such as social networking sites generates increasing amount of data. MapReduce framework[2] adopted by Apache Hadoop [3]has become a part of the standard toolkit for processing the large set of data sets using the cloud services provided by various cloud vendors.

A. Geodistribution

In today's hyper connected world more and more data is being created every day. Users have started to look beyond the illusion that the resources are omnipresent. More and more applications relying on cloud platforms are geodistributed. Data is stored near the respective sources which can be distributed for frequent access. Data can also be replicated across data centers for availability to limit the cost of costly updates. But how to analyze these data sets more efficiently is a major problem. Map Reduce is used to handle these datasets. These subdata sets can handled by gathering them into a single datacentre or by executing instance of MapReduce job separately on each subdata set in respective datacenters and then aggregate the results or to perform the MapReduce jobs as a single geodistributed where mappers and reducers may be deployed in different data centers. MapReduce operations perform poorly as they were not designed to work on multiple data centers [4][5]. MapReduce executing on geodistributed dataset may follow any

Cite this article as: JEBILLA P, Dr P JAYASHREE. "ENHANCING MAP-REDUCE JOB EXECUTION ON GEODISTRIBUTED DATA ACROSS DATACENTERS." *International Conference on Information Engineering, Management and Security (2015)*: 323-327. Print.

combinations of the above options. There can be many execution paths for performing a MapReduce job on a geodistributed dataset and performance can differ.

II. BACKGROUND AND RELATED WORKS

A. MAPREDUCE FRAMEWORK

MapReduce framework such as Apache Hadoop have become a part of the standard toolkit for processing large dataset using cloud resources, and are provided by most cloud vendors. MapReduce works by dividing input data into chunks and process these in a series of parallelizable steps. Sequence of Mapreduce jobs are executed on a given input by applying first job on the given data and then the second job on the output of the fir st job. In many scenarios distinct MapReduce jobs are executed in sequence rather than to perform the job iteratively. When performing a sequence of MapReduce jobs the number of possible execution path increases.

A simple MapReduce program consists of two functions as Map and Reduce such as,

Map <key1; val1> -> list(<key2; val2>)

Reduce <key2; list(val2)> -> ist(val3)

The map function takes the input data and it outputs a set of <key, value> pair. The reducer accepts the set as input and emits set of values .Execution one of this function is a Phase. A MapReduce job includes map and reduce function, input data. The execution of a single job can include any number of mappers and reducers. The input files divided into chunks are inputted to the mapper. These are known as the splits. A partitioning function is used to assign the mapper phase output to the reducer. The Mapreduce framework parallelizes the execution of all functions and ensures fault tolerance.

B. APACHE HADOOP

Apache Hadoop is an open source software framework for distributed storage and distributed processing of very large datasets. The core part of Hadoop consists of a storage called Hadoop Distributed File System (HDFS) and a processing part using MapReduce. Hadoop splits the files into large blocks and distributes these block among the nodes in the cluster and the MapReduce transfers the code to the node to process the data. Hadoop is bundled with the HDFS which is used to store the input of the map phase and also the output of the reduce phase. However, it does not store the intermediate results which are the output of map phase. They are stored on the individual local file systems of the nodes. The Hadoop follows a master-slave model where the master is responsible for accepting jobs dividing those into tasks that includes mappers and reducers and to assign those tasks to slave worker nodes. The hadoop cluster has a single namenode plus a cluster of datanodes. The namenode is the center piece of the HDFS file system which keeps the directory tree of all files in the filesystem and also tracks where across the cluster the data is present. Datanodes are used to store the data and there are number of datanodes. The data is replicated among the datanodes. Above this filesystem comes the MapReduce engine which consists of Job Tracker, to which the client submit the mapreduce jobs. The job tracker service within the hadoop that farms out the mapreduce tasks to specific nodes in the cluster, ideally to the nodes that have the same data. Each worker node has a Task Tracker that processes the assigned tasks. A heartbeat from Task tracker is sent to Job Tracker periodically to update its status.

C. HDFS

The Hadoop Distributed File System is a distributed file system which has been designed to run on commodity hardware. It is designed to be highly fault tolerance and can be deployed on low cost hardware and is also designed to store huge amount of data and it provides high throughput access to the data. HDFS provides fast and scalable access to the information loaded on the clusters and it stores data reliably as it has replication of data across clusters.

D. PREVIOUSLY PROPOSED EXTENSIONS

Numerous efforts have been proposed to improve the efficiency of the execution of the MapReduce jobs. Yang et al.[8] introduced an extra MapReduce phase known as the merge which works after the map and the reduce phase and also extends the MapReduce model for heterogeneous data. Chang et al. [9] Introduced approximation algorithm that was used to order the MapReduce jobs to minimize the overall job completion time. Zaharia et al. [6] worked on to improve the performance of the Hadoop by making it aware of the heterogeneity of the network. MapReduce Online [7] made some modification in MapReduce which allows the MapReduce components to start executing before the data is fully materialized. It allows the reducers to start its process before the complete output of the mapper is available. Amrith Dhananjayan et al. [16] presented an application of the Lyapuno stability theory for load balancing Data Center Networks modeled as discrete event systems. Gaochao Xu et al. [17] introduces a load balance model for the public cloud based on the cloud partitioning concept with a switch mechanism to choose different strategies for different situations. Jehn-Ruey Jiang et al.[11] proposed to extend the well known Dijkstra's shortest path algorithm to consider not only the edge weight but also the node weight in the SDN networks. David B. Wilson et.al [12] described a new forward-backward variant of Dijkstra's and Single source shortest path which allows some of the edges to be scanned backward .

II. MODEL

Sequences of operations are performed on the geodistributed data center. The MapReduce jobs J1; J2; . . . ; Jm give rise to 2 x m operations as each MapReduce job consists of two phase map and reduce. The state of the data before a phase is identified as a stage which starts with 0. The input data is in stage 0 and the final output after the m MapReduce jobs are done it is stage 2m. To move a data from stage s to s + 1, a MapReduce phase is applied to the data partitions and the same amount of data partitions are produced as output.

Cite this article as: JEBILLA P, Dr P JAYASHREE. "ENHANCING MAP-REDUCE JOB EXECUTION ON GEODISTRIBUTED DATA ACROSS DATACENTERS." *International Conference on Information Engineering, Management and Security (2015)*: 323-327. Print.

A. G-MR Overview

G-MR is a hadoop based framework system which is designed to efficiently process the geodistributed data sets and can efficiently perform a sequence of MapReduce jobs on the geodistributed data across data centers. The G-MR employs a Data Transformation Graph (DTG) that is used to determine the optimized execution path for performing sequence of MapReduce jobs. The DTG is used to optimize the execution time of the MapReduce jobs. The optimized execution path obtained may be different from the optimum execution path.

There are generally three main execution paths for performing the MapReduce job on the geodistributed data which are identified as COPY, GEO and MULTI. The COPY execution path is identified as copying all the input data to a single data center prior to the MapReduce job execution. When individual MapReduce jobs are executed on each data center on the corresponding input and the output is aggregated to produce the final result is the MULTI execution path. The other option is to perform the MapReduce job as a single geodistributed operation with mappers and reducers which are distributed across the datacenters. Determining an optimized execution path for a given scenario is not a straightforward process and it becomes even more difficult when sequence of MapReduce job is involved.

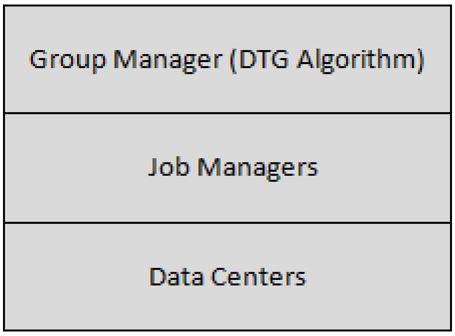


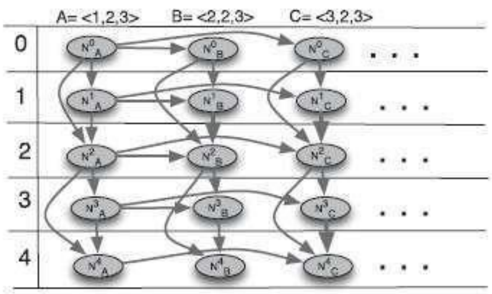
Fig. 1. Architecture of G-MR.

B. ARCHITECTURE AND EXECUTION

G-MR consists of a single component named Group Manger and many components named as Job Manager which is deployed on each participating datacenters. The Group manager determines the optimized execution path while the job Manager manages the MapReduce job that has to be performed within the corresponding datacenter. Fig 1 overviews the architecture of G-MR. The Group Manager executes the DTG algorithm to determine the optimized execution path. The Group Manager informs the corresponding Job Managers about the MapReduce jobs it has to perform in the corresponding datacenters and also regarding the subset of the data on which the job has to be executed.

C. DTG Algorithm

The DTG Algorithm involves constructing a Data Transformation Graph which represents the possible execution paths for performing the MapReduce jobs on the given input dataset. In the existing G-MR system the DTG graph constructed has the edge weight that denotes either the execution time or the cost. The execution time is taken into account. A given node in the graph denotes number of MapReduce phases that has been applied on the input data and its location. The DTG algorithm was used to determine the optimized solution in terms of either the execution time or the cost which involved both the maintaining the node and for transferring the data. Edge between nodes denotes the data copy operations. A node in the graph is described as NsΔ, where s is the number of MapReduce operations in the sequence it is applied and Δ describes the current distribution of data. dk denotes that Psk is located in data center DCdk .Fig 2 shows a simple DTG for a sequence of two mapreduce jobs. Each MapReduce job is represented by three stages in a DTG numbered from s=0 to 2 where s=0 denotes data prior to map phase, s=1 denotes data after map phase and prior to reduce phase and s=2 denotes data after reduce phase. The Execution path is defined as the path from the starting node to the ending node in the DTG. The Data Transformation graph is constructed first based on which the optimized way for executing the MapReduce jobs are determined. This was done using the well know Dijkstra's algorithm [10]. We have used the extended Dijkstra's Algorithm to find the optimized execution path. The extended Dijkstra's algorithm uses the node weight in addition to the edge weight. The node weight denotes the activeness of the node. It could either denote the time or the cost as the edge weight. Based on the activeness of the node, each node is assigned a node weight. The node weight is included only for the outgoing edges in the intermediate nodes and not on the starting and the ending node. This helps in finding the most optimized execution path to perform sequences of mapreduce jobs.



Cite this article as: JEBILLA P, Dr P JAYASHREE. "ENHANCING MAP-REDUCE JOB EXECUTION ON GEODISTRIBUTED DATA ACROSS DATACENTERS." *International Conference on Information Engineering, Management and Security (2015)*: 323-327. Print.

Fig. 2. DTG for a sequence of two MapReduce jobs.

D. Optimizing MapReduce Operation

We have also used the oozie workflow scheduler to manage the MapReduce jobs. It is a server based workflow engine specialized in running the workflow jobs. This Oozie workflow jobs are collection of actions grouped in the form of Directed Acyclic Graph (DAG). The workflow actions start the job in remote systems, which upon completion will call back oozie to notify the action completion and to proceed with the next action. Oozie workflow contains control flow nodes (start, end, decision, fork, join, kill) and action nodes (map-reduce, pig, etc). The control flow nodes define the start and the end of the workflow and Action nodes provides a mechanism to trigger the execution of the processing task. Oozie can identify the completion of the processing task using polling and callback functions. Each task is provided with a callback URL which has to be invoked by the task at the time of the completion else Oozie polls the processing task for completion if failure happens. MapReduce paradigm needs to analyse massive amount of data. Applications of such type need to process large number of data sets which in turn needs to perform several join operations. Query evaluation is a part of the analysis of the large datasets. Joins is most important form of query. Algorithms have been broken into two categories- Two Ways Join and Multi- way joins [13]. A two-way join is performed on the given two datasets P and Q which is defined as a combination of tuples p and q such that p.a = q.b where a and b are values in column P and Q. The multi-way join is defined as a combination of tuplesof n number of datasets. The Reduce side joins looks like a natural way to join the datasets which uses the built-in framework to sort the intermediate keys before they reach the reducer. But this is very time consuming. Hadoop offers another way to join datasets before they reach the mapper. We have used the map-side join [14] to join the datasets to perform the MapReduce jobs. Here the sorting order and the number of partitions must be identical in all the datasets to be joined.

IV. RESULT AND COMPARITIVE ANALYSIS

Big data analysis which is a major issue is overcome by using the Hadoop framework and the MapReduce framework. The Data Transformation graph was introduced to find the all possible execution paths from the source to the destination. The existing Dijkstra's algorithm was used to find the optimized execution path previously. Here we have considered a single datacenter for our implementation. We have extended the Dijkstra's algorithm to include the node weight in addition to the edge weight. The edge weight corresponds to the execution time and the node weight corresponds to the time based on the activeness of the node through which the data is obtained. We also have implemented the map join algorithm to combine the datasets obtained from various datacenters before the mapper phase which helps in reducing the execution time. The execution time of the mapreduce jobs has been enhanced.

V. CONCLUSION

We have modeled a system which enhances the execution of the map reduce jobs which are executed on the geo distributes data across the datacenters. A G-MR system, a MapReduce framework which can efficiently execute a sequence of mapreduce jobs on the geo distributed data relies on the DTG to perform the sequence of job minimizing the execution time. The extended Dijkstra's algorithm is used to enhance the performance of the DTG in the execution of the sequence of the mapreduce jobs.

REFERENCES

[1] Chamikara Jayalath, Julian Stephen, and Patrick Eugster "From the cloud to the Atmosphere: Running MapReduce across data centers" IEEE TRANSACTIONS ON COMPUTERS, VOL. 63, NO. 1, JANUARY 2014.

[2] J.Dean and S. Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters," Proc. Sixth Conf. Symp. Operating Systems Design and Implementation (OSDI), 2004.

[3] "Hadoop," Apache Software Foundation, 13.

[4] "Hadoop Across Data-Centers,

[5] " http://mailarchives.apache.org/mod_mbox/hadoopgeneral/2010 01.mbox//%3C4B4B6AC7.6020801@lifeless.net%3 E, 2013.

[6] Jairam Chandar 'Join Algorithms using Map/Reduce'

[7] M. Zaharia, A. Konwinski, A.D. Joseph, R.H. Katz, and I. Stoica, "Improving MapReduce Performance in Heterogeneous Environments," Proc. Eighth USENIX Conf. Operating Systems Design and Implementation (OSDI), 2008.

[8] T. Condie, N. Conway, P. Alvaro, J. Hellerstein, K. Elmeleegy, and R. Sears, "MapReduce Online," Proc. Seventh USENIX Conf. Networked Systems Design and Implementation (NSDI), 2010.

[9] H.-C. Yang, A. Dasdan, R.-L. Hsiao, and D.S. Parker, "Map-Reduce-Merge: Simplified Relational Data Processing on Large Clusters," Proc. ACM SIGMOD Int'l Conf. Management of Data,2007.

[10] H. Chang, M. Kodialam, R. Kompella, T. Lakshman, M. Lee, and S. Mukherjee, "Scheduling in MapReduce-like Systems for Fast Completion Time," Proc. IEEE INFOCOM, 2011.

Cite this article as: JEBILLA P, Dr P JAYASHREE. "ENHANCING MAP-REDUCE JOB EXECUTION ON GEODISTRIBUTED DATA ACROSS DATACENTERS." *International Conference on Information Engineering, Management and Security (2015)*: 323-327. Print.

[11] E.W. Dijkstra, “A Note on Two Problems in Connection with Graphs,” Numerische Math., vol. 11, pp. 1-269, 1959.

[12] Jehn-Ruey Jiang, Hsin-Wen Huang, Ji-Hau Liao, and Szu-Yuan Chen “Extending Dijkstra’s Shortest Path Algorithm for Software Defined Networking

[13] David B. Wilson, Uri Zwick “A forwardbackward single-source shortest paths algorithm” 2013 IEEE 54th Annual Symposium on Foundations of Computer Science

[14] Jairam Chandar “Join Algorithms using Map/Reduce”[A. Pigul “Comparative Study Parallel Join Algorithms for MapReduce environment” Saint Petersburg State University.

[15] Chung-Cheng Li, Kuochen Wang, 2014 ‘An SLA aware Load Balancing Scheme for Cloud Data centers’ Information Networking (ICOIN), International Conference.pp 58-63

[16] Amrith Dhananjayan1, Kiam Tian Seow1 and Chuan Heng Foh2 , 2013 ‘Lyapunov Stability Analysis of Load Balancing in Datacenter Networks’ Globecom Workshop - The 5th IEEE International Workshop on Management of Emerging Networksand Services.pp 912-916.

[17] Gaochao Xu, Junjie Pang, and Xiaodong Fu, February 2013 ‘A Load Balancing Model Based on Cloud Partitioning for the Public Cloud’ TSINGHUA SCIENCE AND TECHNOLOGY ISSN 1 007 - 02 14 04/12 Volume 18, Number 1. pp 34-39.

Cite this article as: JEBILLA P, Dr P JAYASHREE. “ENHANCING MAP-REDUCE JOB EXECUTION ON GEODISTRIBUTED DATA ACROSS DATACENTERS.” *International Conference on Information Engineering, Management and Security (2015)*: 323-327. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS055		**eAID**	ICIEMS.2015.055

A Review Paper On Latest Trends In Distributed Smart Grid Technology

Upendra Vishnu Kulkarni[1], Aruna P Phatale[2], S A Naveed[3] (Member, IEEE)
[1, 2, 3] Department of Electronics & Telecommunication, MGM'S Jawaharlal Nehru Engineering College
Aurangabad, Maharashtra, India

Abstract: *In current era of Internet of Things, overall internet is made intelligent by making each individual 'smart' at its own place. Use of distributed energy sources including both renewable and non renewable in smart grid technology has gain lot of importance in today's era facing energy crisis. These distributed resources provide greatest flexibility whenever continuous and uniform deployment of power flow is required, across all networks. Interest in the potential of smart grid to transform the way societies generate, distribute and use electricity has increased dramatically over the past decade. This Paper identifies and illustrates the radical, competing and influencing priorities in Distributed Smart Grid Technology. Following latest researches primarily, the goal of paper to cover maximum aspects of to make system design-integration-implementation approach more powerful and perfect is reviewed.*

Keywords: Adaptive, GUI, self healing, management, DSG, Big Data

I. INTRODUCTION

The electricity or power is considered to be backbone of each and every enterprise and has witnessed many recent developments in research and infrastructure, significant socio-economic and other non-tangible benefits to the community at large.Overall interconnection of advanced system components well known as Smart grid. "Ref.[5]" The arrival of power system deregulation and corresponding parallel growing vertically integrated utility business model is a second important development which shaped the direction of electric power technology in drastic but fruitful manner. It is a dynamic interactive, real-time infrastructure that responds to the challenges of designing and building the power system of the future.

"Ref.[14]"To illustrate the diversity of terminology as in figure 1, the paper summarizes subfields with examples and definitions of Smart Grid. Miscellaneous Operating Level model of Smart Grid Technology can be seen in terms of diagrams and definitions mentioned .Smart distribution and utilization systems are suitably referred with their detailed abbreviations in subsections from corresponding figures respectively. "Ref.[15]" This paper includes the concept of Distributed smart grid and its various manifestations to allow for an appreciation of the complexity, potential benefits, and challenges involved in this exciting field to particular extent.

Cite this article as: Upendra Vishnu Kulkarni, Aruna P Phatale, S A Naveed. "A Review Paper On Latest Trends In Distributed Smart Grid Technology." *International Conference on Information Engineering, Management and Security (2015)*: 328-333. Print.

Fig.1 Customer-Vendor insight of DSG

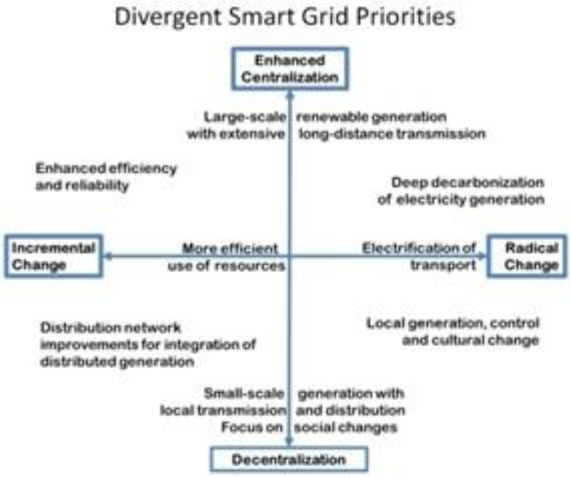


Fig. 2 Divergent Smart Grid Priorities

Topics such as real-time energy control approach for any smart building energy management systems, optimal opinion of energy-efficient buildings with distributed energy resources, various management strategies at device, big data, system and environment levels with suitable diagrams in short. These concepts needed to be followed thoroughly along with Divergent Smart Grid Priorities as shown in fig 2.

Different kind of facts is stated related with this Distributed Smart Grid (DSG) with addressing challenges and related innovative technologies, products and services. Further short note on DSG technology activities in China, India, and around the globe on this roadmap is done.

II. CONTRIBUTARY INFORMATION

"Ref.[13]"Currently available literature on the term "smart grid" shows a vast array of publications and latest valuable trends in distributed system to become backbone of "Internet of Things". Today the power industry is taking advantage of advanced computer, communication, and control technologies throughout the 21th century to enhance its 360 degrees of capabilities.

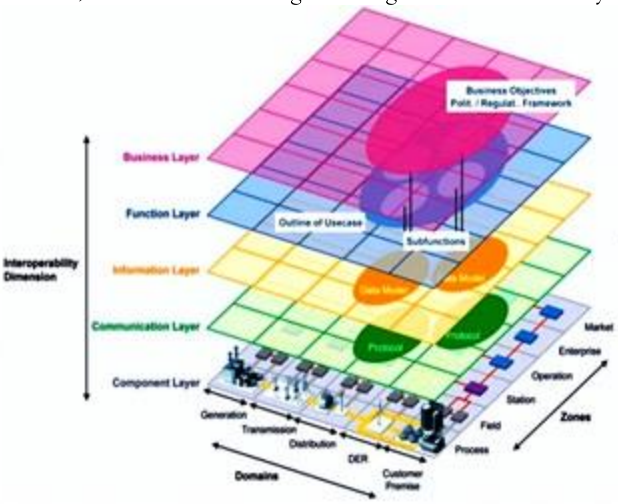


Fig.3 Miscellaneous Operating Level model Fig4. Distributed Parallel Network of Energy & Information Technology

III. THE MOTIVATION FOR DISTRIBUTED SMART GRID

"Ref.[15]"Presently the stakeholders (utilities, vendors, manufacturers, regulators, consumers and their advocates, and governments) recognize the need to address challenging issues that motivate developing and implementing the Distributed smart grid and its

Cite this article as: Upendra Vishnu Kulkarni, Aruna P Phatale, S A Naveed. "A Review Paper On Latest Trends In Distributed Smart Grid Technology." *International Conference on Information Engineering, Management and Security (2015)*: 328-333. Print.

peripheral elements due to limitations of simple smart grid system. In every situation the priority of local drivers and challenges might differ from one point of view to another when a partial list of issues given below is concerned.

a. Aging and underinvested infrastructure

b. Fulfilling the demand of Electricity throughout the world

c. Adoption of alternative energy sources to reduce pollution.

d. Optimization in flexibility of the distance between generation sites and load centers in either or parallel with large numbers of small, decentralized generation.

e. Intermittent and fluctuating energy availability of renewable energy sources

f. Need for securing electric supply and dataflow associated with it.

IV. DISTRIBUTED SMART GRID FEATURES

"Ref.[6]"Many Distributed smart grid possess following attributes as representative of its functions:

1) *Efficient*—satisfy consumer's electricity demand without adding infrastructure.

2) *Accommodating*—works with any kind of energy source, platform, virtual online mode

3) *Motivating*—enabling real-time communication using graphical user interface (GUI)

4) *Opportunistic*— ability to capitalize on plug-and-play innovation wherever and whenever required

5) *Quality focused*—capable of delivering the power quality necessary, free of sags, spikes, disturbances, and interruptions to power our increasingly digital economy and the data centers, computers, and electronics necessary to make it run.

6) *Resilient*—increasing decentralized process helps to improve flexibility, maintenances and reinforce the smart grid security protocols.

V. SOME IMPERFECTIONS IN DISTRIBUTED SMART GRID

"Ref.[3]"As with many new innovative technological developments, care must be taken to address particular concerns and issues that present itself to forward progress, adoption, and acceptance of this enterprises .With current Big Data Concept huge information related with this technology is also growing exponentially and securing it is a big issue .

a. *Stakeholder Engagement:* Well-knowing the fact that the benefits of each component of the smart grid to the customers that are the potential key to service success they worked seriously on it.

b. *Security:* As with integrated approach information in turning into big data, creating cyber-security vulnerabilities, tackling security risks gained utmost importance.

c. *Privacy:* The concern like consumer acceptance, privacy violations needs to be addressed appropriately.

Costliness: New cutting edge inventions will be costly if we put them into practice

VI. SMART DISTRIBUTION AND UTILIZATION SYSTEMS

"Ref. [10] [11]"Given that the origins of many power system issues are typically based in the electrical distribution system, the point of departure for grid enhancement and modernization is to be found at the bottom of the supply chain.

"Ref.[9]"While the distribution system shown in figure 5 are major part of the electric power system, it comes as a surprise that there is not a corresponding appreciable level of embedded intelligence with the only information available—that from the feeder at the substation. This makes it difficult to optimize the operation of the distribution system and to recreate and recover from abnormal

Cite this article as: Upendra Vishnu Kulkarni, Aruna P Phatale, S A Naveed. "A Review Paper On Latest Trends In Distributed Smart Grid Technology." *International Conference on Information Engineering, Management and Security (2015)*: 328-333. Print.

events. Distributed Energy Resources(DER) are able to create self contained cells (micro-grids), which can in turn help to assure energy supply in distribution grids even when the transmission grid has a blackout . Moreover, it also facilitates more effective utilization and life extension of existing distribution system infrastructure integrating DERs and some of their functional requirements are as below.

i) *Distributed intelligence:* possess high-speed data processing capability; make decisions locally through distributed intelligence offered by low-cost embedded computing facility.

ii) *Visualization:* As per high-priority requirement we can recognize available and controllable resources to maximize economic and reliability benefits. It can also visualize in terms of devices and entities as shown in figure 4 and figure 6 respectively.

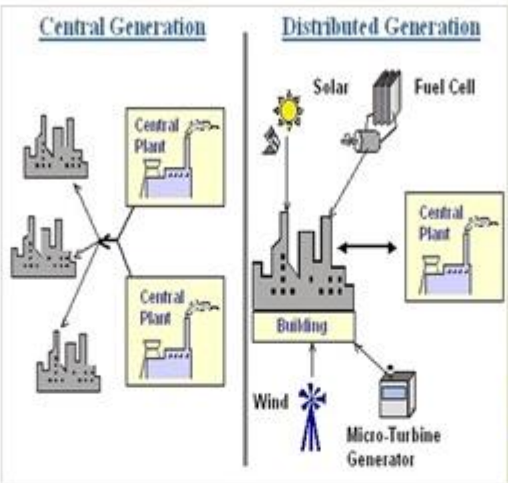


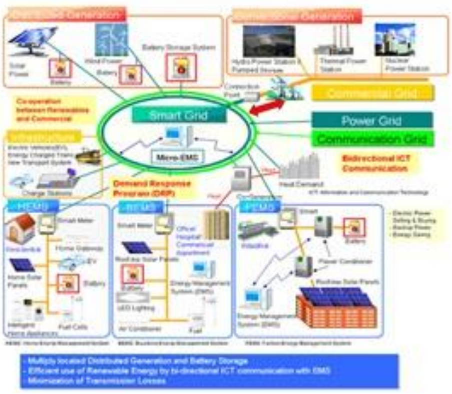
Fig.5 Centralized Vs Distributed

Fig.6 Hitachi for Energy and IT Power Generation

iii) *Forecasting and prediction:* Due to immaturity and low penetration of DERs which increases the uncertainty associated with their performance it is challenging.

iv) *Interoperability:* "Ref.[3]" Among networks, systems, devices, applications or components to externally exchange and readily use information securely and effectively standards need to be followed with drastic variation.

"Ref.[5][8]"As in figure 7, a Distributed management system for energy, outage, number ON/OFF devices in certain time, fault location are important component of the DSG to increase efficiency of overall system . It allows remote meter configuration, communicating dynamic tariffs, power quality monitoring, and potential load control. The fig. 7 considers distribution network active management and future development trends in technologies and methods, where centralized and decentralized management frameworks and applying agent-based coordination are shown. Smart grid initiatives, developments, plans, and example technologies are taken shortly.



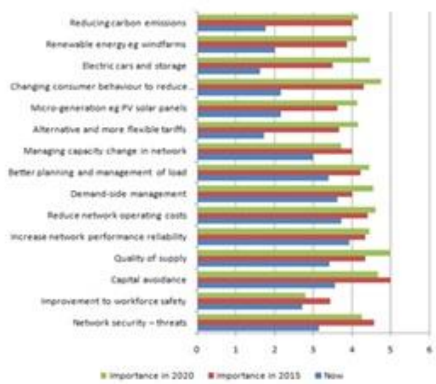


Fig.7 Overview of Distributed Smart Grid

Fig.8 Evaluative Graph of various aspects and parameters for Distributed Smart Grid as future perspective

Cite this article as: Upendra Vishnu Kulkarni, Aruna P Phatale, S A Naveed. "A Review Paper On Latest Trends In Distributed Smart Grid Technology." *International Conference on Information Engineering, Management and Security (2015)*: 328-333. Print.

VII. DISTRIBUTED SMART GRID INITIATIVES, DEVELOPMENTS, PLANS, AND EXAMPLE TECHNOLOGIES

1. Recent Advancements on Smart Grids in China :

"Ref.[3]"Xu, Xue, and Wong from Hong Kong Polytechnic University, China; State Grid Electric Power System Research Institute, China; and University of Western Australia, Australia, respectively, discuss China's 12th Five-Year Plan (2011 2015)by renewable energy accounting for15% of national primary energy consumption by 2020.The State Grid Corporation of China (SGCC) established the plan to implement smart power grids in China by 2020 with further phases, policy and strategy.

2. Smart Grid Status in India :

"Ref.[12]"Government of India has launched the scheme namely Faster Adoption and Manufacturing of (Hybrid &) Electric Vehicles (FAME India) under National Electric Mobility Mission Plan (NEMMP) 2020 in the Union Budget for 2015-16 with an initial outlayofRs.75Cr. The scheme will provide a major push for early adoption and market creation of both hybrid and electric technologies vehicles in the country.

3. Activities Worldwide:

"Ref. [12]" International Renewable Energy Agency (IRENA), an intergovernmental organization that supports the spread of renewable energy worldwide is expected to finalize details this week of a road map to install 160 Giga watts (GW) of battery storage worldwide in 2030. The volume of battery storage is expected to soar on the back of increasing electric vehicle penetration specially; lithium-ion battery-based storage is due to rise from 100MW in 2012 to around 25GW in 2020, and 150GW in 2030 applications, allowing the operational layer to be secured.

VIII. FEW FUTURE RESEARCH TRENDS

"Ref.[12]"To implement DSG successfully, the smart grid requires careful attention to the multitude of new needs for applied research as given as.

a) Optimal sizing and placement of distribution system resources
b) Optimal predictor-corrector resource dispatching
c) Integrated forecasting suites
d) Optimal adaptive reconfiguration of DNs
e) DN state estimation and observation

"Ref.[4][6]" The figure 8 gives the detailed importance of various aspects and parameters for Distributed Smart Grid as future perspective in comparison with the past, present and future .The era under observation is shown with respect to earlier study. The graph comprises of environmental health issues in joint with infrastructure needed for energy sources, building the network and its numerous components, human resources for research and associated management techniques, Quality security threats and much more. All these issues are interlinked, interdependent and as they reside in vast variety of different fields they should be of main concern.

"Ref. [10]" Distributed and embedded intelligence are important for emphasizing self-healing, optimizing operation. Functional and integration requirements of Distributed Energy Resources, Infrastructure .Management Systems, Metering ,Smart buildings and Smart devices should be of main concern with respect to recent initiatives, developments, technologies, and research.

IX. CONCLUSIONS

This paper has given short over view on the latest aspects of Distributed smart grid and closely related articles. Integrated communication using intermediate interfaces allows real time control of equipments, exchange of data measurements and optimization through atomization with certain parameters at all levels. Interest in the potential of smart grid to transform the way societies generate, distribute, and use electricity has increased dramatically over the past decade and transformed it into Distributed Smart Grid. With the above precisely selected issues this review paper has identified and illustrated the radical, competing and influencing priorities in Distributed Smart Grid Technology and until now whatever technology is developed ;still bigger challenges will be there that need to be resolved very efficiently based upon the new cutting edge technology , always as we move along the journey.

REFERENCES

[1] Peng Li ,Song Guo And Zixue Cheng "Joint Optimization of Electricity and Communication Cost for Meter Data Collection in Smart Grid" Received 3 May 2013; revised 7 July 2013; accepted 8 July 2013. Date of publication 18 July 2013; date of current version 21 January 2014.

[2] Yu Wang, Shiwen Mao and R. M. Nelms" Online Algorithm for Optimal Real-Time Energy Distribution in the Smart Grid" Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849-5201, USA

Cite this article as: Upendra Vishnu Kulkarni, Aruna P Phatale, S A Naveed. "A Review Paper On Latest Trends In Distributed Smart Grid Technology." *International Conference on Information Engineering, Management and Security (2015)*: 328-333. Print.

[3] Mohamed E. El-hawary (2014) The Smart Grid—State-of-the-art and Future Trends, Electric Power Components and Systems, 42:3-4, 239-250, DOI: 10.1080/15325008.2013.868558 http://dx.doi.org/10.1080/15325008.2013.868558

[4] Jennie C. Stephens [1], Elizabeth J. Wilson [2], Tarla R. Peterson [3]and James Meadow croft [4]"Getting Smart? Climate Change and the Electric Grid" Received: 7 July 2013 / Revised: 23 August 2013 / Accepted: 23 August 2013 Published: 5 September 2013Challenges 2013, 4(2), 201-216; doi: 10.3390/challe4020201.

[5] IEEE power & energy magazine march/april 2015

[6] http://smartgrid.ieee.org/conferences/event/196-ieee-international-symposium-on-power-line- communications and-its-applications-isplc

[7] http://www.smartgridupdate.blogspot.in/

[8] http://energy.gov/oe/services/technology-development/smart-grid

[9] http://indiasmartgrid.org/en/technology/Pages/Distributed-Generation.aspx

[10] https://www.navigantresearch.com/research/smart-grid-10-trends-to-watch-in-2015-and-beyond

[11] https://blogs.siemens.com/smartgridwatch/stories/23619

[12] SMART GRID Bulletin Volume 2, Issue 3 March 2015 (www.indiasmartgrid.org)

[13] Jayakrishnan R. Pillai 'Smart Grid– Projects & Future Challenges presentation', Department of Energy Technology Aalborg University

[14] Vikram Gandotra 'Smart Grids IIT Delhi Introduction to Smart Grids presentation 27/04/2013

[15] Nur Asyik Hidayatullah, Blagojce Stojcevski, Akhtar Kalam'Smart Grid and Renewable Energy', 2011, 2, 216-229 doi:10.4236/sgre.2011.23025

Cite this article as: Upendra Vishnu Kulkarni, Aruna P Phatale, S A Naveed. "A Review Paper On Latest Trends In Distributed Smart Grid Technology." *International Conference on Information Engineering, Management and Security (2015)*: 328-333. Print.

International Conference on Information Engineering, Management and Security
2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS056		**eAID**	ICIEMS.2015.056

An Efficient Energy Management System for Customers Using Renewable Energy Sources

Snehal D. Solunke[1], Vishwashri A. Kulkarni[2]
[1,2] Department of Electronics and Telecommunication,MGM's JNEC
Aurangabad,Maharashtra, India.

Abstract: *Life as we know it would be impossible without electricity. But as our supply of fossil fuels, energy source for electricity, continues to dwindle, we face a majorchallenge. Renewable energy that we get from the environment is a commendable option for providing green energy to homes. A high efficiency power management system for solar energy harvesting applications is proposed. Admiring efforts have been made recently in transforming the power grid into smart grid by means of unifying extensive information and communication infrastructures. In this paper a versatile stochastic optimization approach has been proposed for smart grid, to meet the need of the residential customers in an efficient energy management comprising high penetration of renewable and distributed energy sources, market based online electricity pricing, large scale storage of energy and high quality demand side management, mitigating monetary expense.The concept Boost converter used in this paper increases the throughput of solar panel voltage. Smart energy meter, an added advantage of the proposed system provides a two way communication between the utility board and residential unit.*

Keywords: Power grid, smart grid, smart energy meter, solar energy, boost converter, solar panel.

I. INTRODUCTION

An increasing Global warming, currently occurring on this 4.6 billion years old earth, is a very critical issue to be addressed by the modern society that has been enjoying economical growth by consumption of fossil energies Since the Industrial Revolution in Great Britain, much carbon dioxide (CO2) has been emitted as a result of the combustion of petroleum and coal. In the past 200 years, the carbon dioxide concentration in the atmosphere has increased by as much as 25%. Now the entire earth is, so to speak, situ three of the most prominent issues facing the world today are escalating climate change, energy security and meeting the increasing global demand for electrical energy generated from renewable sources.Renewable sources are also called Echo friendly technologies are very important due to their pollution free energy generation and having sustainable growth. The electrical power grid is by nature is a complex adaptive system and it regards with significant amount of uncertainties [1]. The existing grid faces some sensitive problems which are major factors of concern. They can be specifically mentioned as follows:
1) Limited delivery system.
2) High cost of power outage and power quality interruption.
3) In efficiency at managing peak load. 4) Increase in global warming and hazardous emissions.

Cite this article as: Snehal D. Solunke, Vishwashri A. Kulkarni. "An Efficient Energy Management System for Customers Using Renewable Energy Sources." *International Conference on Information Engineering, Management and Security (2015)*: 334-337. Print.

This adverse impact of global warming and greenhouse effects is indeed a curse to the entire existence of life on earth. So for the sake of saving our earth it is very important to switch towards renewable energy sources. Since smart grid encapsulates this peculiar features of utilizing this renewable energy sources, smart grid integrate large amount of renewable generation in specific solar to meet our overwhelming electricity energy demands [2]. Smart grid introduces another worth mentioning feature say smart meter. Smart meter is a device that is connected to power distribution system which embeds a scheduling unit that helps in implementing shifting of workload [7], [9]. It also helps in receiving periodically the updated pricing information from various utility companies and the scheduling unit incorporated within it arranges the various household appliances for operation during different time intervals

II. PV TECHNOLOGY

Photovoltaic (PV) systems involve the direct conversion of sunlight into electricity with no intervening heat engine [7]. PV devices are solid state; therefore, they are rugged and simple in design and require very little maintenance. PV systems produce no emissions, are reliable, and require minimal maintenance to operate. They can produce electricity from microwatts to few megawatts. (Figure 1)

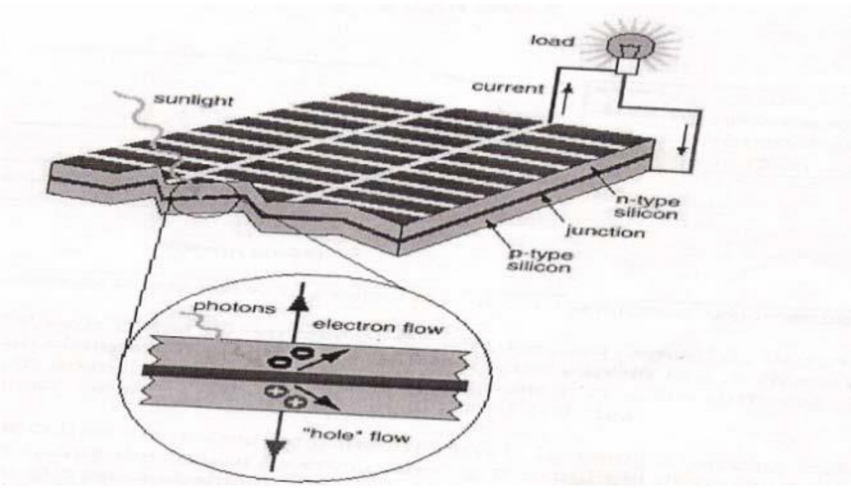


Figure 1.Electricity generation by solar (PV) Module

III. SYSTEM ARCHITECTURE

Owing to uncertainties under the real time pricing environment this paper mainly focuses in minimizing the electricity expenses of customers through optimally scheduling the operation and energy consumption for each and every appliance. Figure 2 depicts the architecture of energy management system used in residence. The following subsections are included in this model.

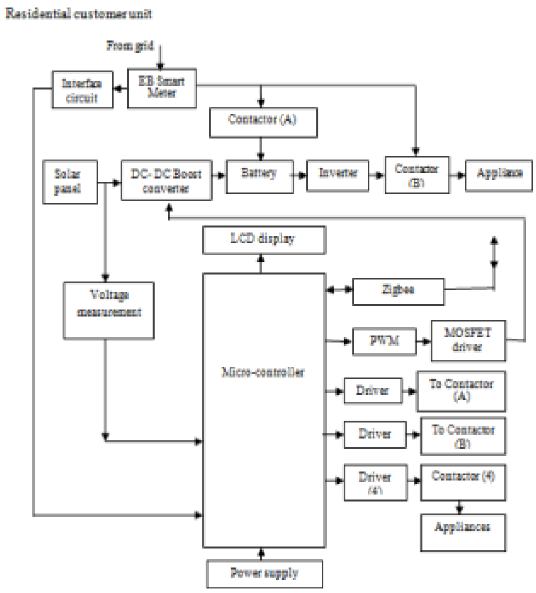


Figure 2. System Architecture.

Cite this article as: Snehal D. Solunke, Vishwashri A. Kulkarni. "An Efficient Energy Management System for Customers Using Renewable Energy Sources." *International Conference on Information Engineering, Management and Security (2015)*: 334-337. Print.

A. Utility control unit

The utility control unit performs the usage patterns and the other analytics associated with it. In this system, there persist a mutual usage data communication between the smart meter and the control unit. A cluster of meters is together assigned with a single control unit which helps in optimal utilization of the same.

B. Residential customer unit

The residential unit includes the below mentioned household appliances: air conditioner, washers, coolers, etc. The complete functionality of this residential unit is achieved by the inclusion of the following modules as shown in Fig.3.

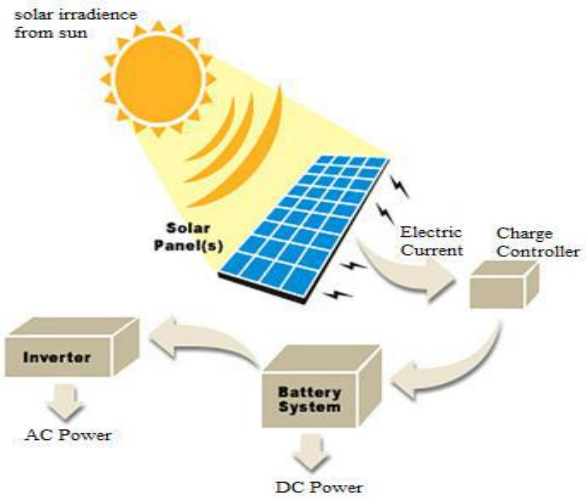


Figure 3.Electricity generation for residential unit using solar panel

1) Renewable energy generation: The resources are fast depleting with proportionate increase in energy demand, which ultimately results in providing a significant importance in the generation of renewable energy. Renewable energy source that is adopted in this paper is photovoltaic cells which promise a clean and cost effective factor that simultaneously turns out to be an opaque medium against harmful reactions to our ecosystem. PV cell uses the technique photoelectric effect which helps in converting the sunlight of certain wavelengths into direct current.

Generally a single PV cell cannot suffice the residential customer needs, so it is mandatory to connect them in series and parallel for the purpose of achieving the targeted voltage and current levels. As the PV cell generated voltage is intermittent by nature, the boost converters are included which aids in boosting the voltage, thus helps the PV cells in maintaining an ever sustainable voltage.

2) Boost converters: A boost converter or step up converter is basically a DC-to-DC power converter which functions with an aim of producing an output voltage greater than its input voltage. Power is a factor that must be conserved and hence as a result of these criteria, always output current is lower than the source current.

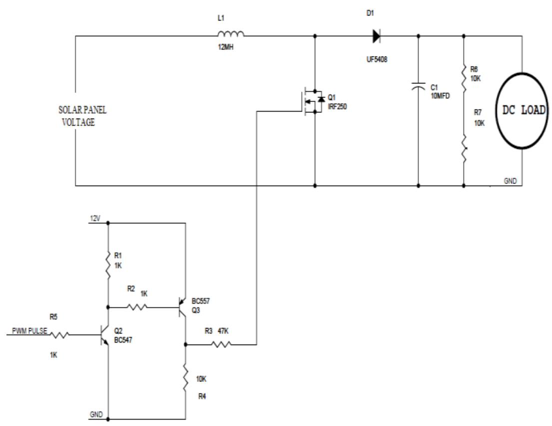


Figure 4. Boost Converter.

The role of maximum power point converter is felt by the presence of boost converters which is ultimately merged with the output load of solar panel. The voltage from the solar panel is given to the controller which in turn generates the relevant PWM waves that is further fed back to the boost converter's switch control input. The core functionality of boost converter is storing current in its boost inductor when the switch is closed or just delivering the currents from its inductor to the load when the switch opens and it is shown in Figure 4. The output of boost converter is fed to energy storage block.

Cite this article as: Snehal D. Solunke, Vishwashri A. Kulkarni. "An Efficient Energy Management System for Customers Using Renewable Energy Sources." *International Conference on Information Engineering, Management and Security (2015)*: 334-337. Print.

3) Energy storage: There most frequently exists an energy mismatch between the PV profile system and the residential energy demand. To synchronize this mismatch the concept of energy storage is introduced, which is used to store the excessively generated electricity at daytime, which forms as a supplement form of power usage that can be released at night time to meet the residential customer needs. This phenomenon paves way for appreciable amount of reduction in tapping of electricity from grid.
Overcharging and discharging will affect the operational life of a battery and hence it must be protected from these phenomena by incorporating a controller that helps in regulating the charging and discharging cycles [8].
4) Inverters: Generally battery produces the DC power, but the household appliances requires AC power for its operation, for that reason the inverter is used to convert this 12V DC into 230V AC household voltage. The voltages from these inverters are used by appliances through contactors.

IV. SMART METER

A smart meter is an energy meter that records utilization of electric energy in short intervals of time and sends the relevant details to the utility unit for scrutinizing and billing purpose, in predetermined time intervals. A two-way communication between meter and the central system is aided by smart meter. The Advanced Metering Infrastructure (AMI) is a technique incorporated by smart meter and this notable feature is absent in traditional AMR. This helps to reduce the monetary expense charged to customers in the real time pricing arena. The following worthy features are encapsulated in smart meter: eliminates the payment of bill in person at the EB office, provides consent for fetching details regarding updated pricing power utilization, accuracy of bill is verified, inclusion of multiple buildings to the wireless methodology does not alter the uniqueness of the network.

V. CONCLUSION

The concept introduced in this paper provides an efficient energy management for household appliances based on real time pricing released by utility companies in predetermined time intervals. The methodology proposed in this paper is to store the excessively generated renewable energy for future use and thereby to charge the battery at times of low electricity price and simultaneously discharging them during peak pricing time to minimize the monetary expense. It can be concluded that power scheduling approach using RTP combined with the IBR pricing model is a better way compared with the RTP alone pricing scheme. As a future vision of this paper super capacitors are used as energy storage that can charge faster than batteries, last longer and overcome physical toll that wears down the batteries in charging and discharging. This concept will help to the consumers for lower billing rate and also to save the energy that is produced from the natural resources.

REFERENCES

[1] [1] YuanxiongGuo, Miao Pian, and Yuguang Fang, “Optimal Power Management Of Residential Customers in the Smart Grid,” IEEE Trans. Parallel and Distributed Systems, vol. 23, no. 9, Sep 2012.
[2] [2] Xiaodao Chen, Tongquan Wei, Shiyan Hu, “Uncertainty Aware Household Appliance Scheduling Considering Dynamic Electricity Pricing in Smart Home,” IEEE Trans. Smart Grid, vol. 4, pp. 932-941, June 2013.
[3] [3] B. Roberts and C. Sandberg, “The Role of Energy Storage in Development of Smart Grids,” Proc. IEEE, vol. 99, no. 6, pp. 1139-1144, June 2011.
[4] [4] P. Ribeiro, B. Johnson, M. Crow, A. Arsoy, and Y. Liu, “Energy Storage Systems for Advanced Power Applications,” Proc. IEEE, Vol. 89 no. 12, pp. 1744-1756, Dec. 2001.
[5] [5] T. Markel and A. Simpson, “Plug-in Hybrid Electric Vehicle Energy Storage System Design,” Proc. Advanced Automotive Battery Conf., 2006.
[6] [6] A. Khodaei, M. Shahidehpour, and S. Bahramiradh, “SCUC with Hourly Demand Response Considering Intertemporal Load Characteristics,” IEEE Trans. Smart Grid, vol. 2, no. 3, pp. 564-571, 2011.
[7] [7] A. Mohsenian-Rad, V. Wong, J. Jatskevich, and R.Schober, “Optimal and Autonomous Incentive-based Energy consumption Scheduling Algorithm for Smart Grid,” Proc. Innov.Smart Grid Technol. (ISGT), 2010.
[8] [8] “How a PV System Works,” Florida Solar Energy Center, http:// www.fsec.ucf.edu/en/consumer/solar_electricity/basics/how_pv_sysem_works.htm, 2012.
[9] [9] A. Vojdani, “Smart Integration,” IEEE Power Energy Mag., vol. 6, pp. 71-79, 2008.
[10] [10] K. Moslehi and R. Kumar, “A Reliability Perspective of the Smart Grid,” IEEE Trans. Smart Grid, vol. 1, no. 1, pp. 57-64, 2010.
[11] [11] B. Russel and C. Benner, “Intelligent Systems for Improved Reliability and Failure Diagnosis in Distribution Systems,” IEEE Trans. Smart Grid, vol. 1, no. 1, pp. 57-64, 2010.
[12] [12] A.H. Mohsenian-Rad and A. Leon-Garci, “Energy- Information Transmission Traseoff in Green Cloud Computing,” Proc. IEEE Globecom’10, Mar. 2010.
[13] [13] P. Centolella, “The Integration of Price Responsive Demand into Regional Transmission Organization (RTO) Wholesale Power Markets and System Operations,” Energy, vol 35, no. 4, pp. 1568-1574, 2010.
[14] [14] R. Anderson, A. Boulanger, W. Powell, and W. Scott, “Adaptive Stochastic Control for the Smart Grid,” Proc. IEEE, vol. 99 no. 6, pp. 1098-1115, June 2011.

Cite this article as: Snehal D. Solunke, Vishwashri A. Kulkarni. “An Efficient Energy Management System for Customers Using Renewable Energy Sources.” *International Conference on Information Engineering, Management and Security (2015)*: 334-337. Print.

International Conference on Information Engineering, Management and Security 2015 [ICIEMS 2015]

ISBN	978-81-929742-7-9		**VOL**	01
Website	www.iciems.in		**eMail**	iciems@asdf.res.in
Received	10 - July - 2015		**Accepted**	31- July - 2015
Article ID	ICIEMS057		**eAID**	ICIEMS.2015.057

AN INTERACTIVE IMPLEMENTATION ON A SMART PHONE FOR DISABLED PERSONS TO ACCESS HOME APPLICATIONS

M.Shabnam Banu[1], M.S.Fathima Bevi[2], K.Thilagavathi[3]
[1,2,3] UG Student, Department of ECE, AVS College of Technology, Salem, India

ABSTRACT: *The main objective of this work is to present a display design for accessible interaction in home area networks. Recently the social inclusion and technical aid to assure autonomy to people with disabilities are getting attention all over the world. The interface was implemented over a Tablet that controls domestic devices through a home network controller prototype In order to evaluate the design, a research was conducted, interviewing people with disabilities in Brazil. This research consolidated a feasible interface to control home area networks. The system is cheap, reasonably easy to configure and run. It consumes very low power. Wireless home automation system is used for speech command. Voice recognition is used to support for blind people. Objective of the project is to help the disabled people. Simulation of the project is done by IAR embedded work bench which is implemented in MSP430 microcontroller. Home appliances are controlled by voice using MSP 430 microcontroller. HM2007 is a voice recognition system that has two modules. The first module contains the mike through which voice input is given that is then stored in the HM2007 voice recognition system. The voice recognition system transfers the voice command to the next module with the help of Zigbee which acts as the transmitter during transfer..*

1.INTRODUCTION

Home automation refers to the use of computer and information technology to control home appliances and features (such as windows or lighting). Systems can range from simple remote control of lighting through to complex computer/micro-controller based networks with varying degrees of intelligence and automation. Home automation is adopted for reasons of easy, security and energy efficiency. Much software is needed while we are operating in tablet. This provides inconvenience to the persons who are physically challenged. So through voice recognition home automation is achieved. Using this voice recognition technique physically challenged persons can operate the home appliances more easily.

The objective of the work is to provide home automation for the physically challenged persons. Home automation provides controlling of the home appliances. Home automation is adopted for security and energy efficiency. This provides a voice recognition method which is easily operated by physically challenged persons. The interface was implemented over the tablet that control domestic through a home network controller prototype. Home automation can also provide a remote interface to home appliances or the automation system itself, via telephone line, wireless transmission or the internet, to provide control and monitoring via smartphone or Web browser .This will be effective to the normal persons. But physically challenged persons may face problem with this. To overcome its problem voice recognition is added to control the home appliances.

Cite this article as: M.Shabnam Banu, M.S.Fathima Bevi, K.Thilagavathi. "AN INTERACTIVE IMPLEMENTATION ON A SMART PHONE FOR DISABLED PERSONS TO ACCESS HOME APPLICATIONS." *International Conference on Information Engineering, Management and Security (2015)*: 338-342. Print.

2.SYSTEM DESIGN AND IMPLEMENTATION

2.1 INTRODUCTION ABOUT MSP430

The Texas InstrumentsMSP430 family of ultra-low power microcontrollers consists of several devices featuring different sets of peripherals targeted for various applications. The architecture, combined with five low-power modes, is optimized to achieve extended battery life in portable measurement applications. The device features a powerful 16-bit RISC CPU, 16-bit registers, and constant generators that contribute to maximum code efficiency. The calibrated digitally controlled oscillator (DCO) allows wake-up from low-power modes to active mode in less than 1 μs.

The MSP430F23x/24x (1)/2410 series are microcontroller configurations with two built-in 16-bit timers, a fast 12-bit A/D converter (not MSP430F24x1), a comparator, four (two in MSP430F23x) universal serial communication interface (USCI) modules, and up to 48 I/O pins. The MSP430F24x1 devices are identical to the MSP430F24x devices, with the exception that the ADC12 module is not implemented. The MSP430F23x devices are identical to theMSP430F24x devices with the exception that a reduced Timer B, one USCI module, and less RAM is integrated. Typical applications include sensor systems, industrial control applications, hand-held meters, etc.

2.2 BLOCK DIAGRAM

The figure 3.2 MSP430G2231 shows the internal block diagram that explains various peripherals in the architecture of the microcontroller.

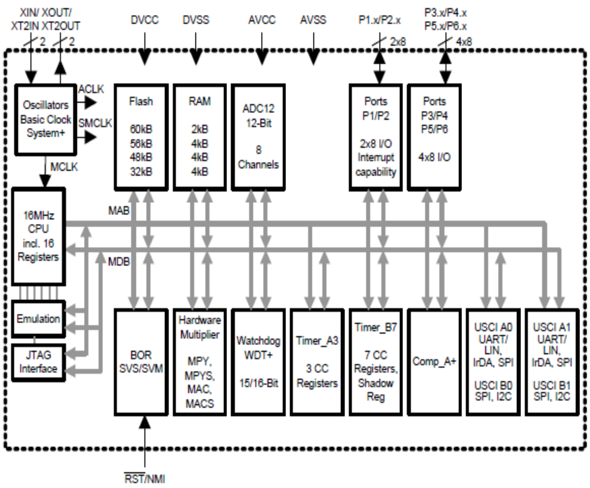


Figure 2.1 Block Diagram of MSP430G2231

2.3 SPEECH RECOGNITION SYSTEM

Speech recognition will become the method of choice for controlling appliances, toys, tools and computers. At its most basic level, speech controlled appliances and tools allow the user to perform parallel tasks (i.e. hands and eyes are busy elsewhere) while working with the tool or appliance .The heart of the circuit is the HM2007 speech recognition IC. The IC can recognize 20 words, each word a length of 1.92 seconds.

The speech recognition system is a completely assembled and easy to use programmable speech recognition circuit. Programmable, in the sense that you train the words (or vocal utterances) you want the circuit to recognize. This board allows you to experiment with many facets of speech recognition technology. It has 8 bit data out which can be interfaced with any microcontroller for further development. Some of interfacing applications which can be made are controlling home appliances, robotics movements, Speech Assisted technologies, Speech to text translation, and many more. Speech recognition will become the method of choice for controlling appliances, toys, tools and computers. At its most basic level, speech controlled appliances and tools allow the user to perform parallel tasks (i.e. hands and eyes are busy elsewhere) while working with the tool or appliance. The heart of the circuit is the HM2007 speech recognition IC. The IC can recognize 20 words, each word a length of 1.92 seconds.

2.4 TRAINING WORDS FOR RECOGNITION

Press “1” (display will show “01” and the LED will turn off) on the keypad, then press the TRAIN key (the LED will turn on) to place circuit in training mode, for word one. Say the target word into the onboard microphone (near LED) clearly. The circuit signals

Cite this article as: M.Shabnam Banu, M.S.Fathima Bevi, K.Thilagavathi. “AN INTERACTIVE IMPLEMENTATION ON A SMART PHONE FOR DISABLED PERSONS TO ACCESS HOME APPLICATIONS.” *International Conference on Information Engineering, Management and Security (2015)*: 338-342. Print.

acceptance of the voice input by blinking the LED off then on. The word (or utterance) is now identified as the “01” word. If the LED did not flash, start over by pressing “1” and then “TRAIN” key.
You may continue training new words in the circuit. Press “2” then TRN to train the second word and so on. The circuit will accept and recognize up to 20 words (numbers 1 through 20). It is not necessary to train all word spaces. If you only require 10 target words that are all you need to train.

2.5 CHANGING & ERASING WORDS

Trained words can easily be changed by overwriting the original word. For instances suppose word six was the word “Capital” and you want to change it to the word “State”. Simply retrain the word space by pressing “6” then the TRAIN key and saying the word “State” into the microphone. If one wishes to erase the word without replacing it with another word press the word number (in this case six) then press the CLR key .Word six is now erased.

2.6 SIMULATED INDEPENDENT RECOGNITION

The speech recognition system is speaker dependent, meaning that the voice that trained the system has the highest recognition accuracy. But you can simulate independent speech recognition.
To make the recognition system simulate speaker independence one uses more than one word space for each target word. Now we use four word spaces per target word. Therefore we obtain four different enunciations of each target word (Speaker independent). The words spaces 01, 02, 03 and 04 are allocated to the first target word. We continue do this for the remaining word space. For instance, the second target word will use the word spaces 05, 06, 07 and 08. We continue in this manner until all the words are programmed. If you are experimenting with speaker independence use different people when training a target word. This will enable the system to recognize different voices, inflections and enunciations of the target word. The more system resources that are allocated for independent recognition the more robust the circuit will become. If you are experimenting with designing the most robust and accurate system possible, train target words using one voice with different inflections and enunciation's of the target word.

2.7 THE VOICE WITH STRESS & EXCITEMENT

Stress and excitement alters ones voice. This affects the accuracy of the circuit’s recognition. For instance assume you are sitting at your workbench and you program the target words like fire, left, right, forward, etc., into the circuit. Then you use the circuit to control a flight simulator game, Doom or Duke Nuked. Well, when you’re playing the game you’ll likely be yelling “FIRE! …Fire! …FIRE!!…LEFT …go RIGHT!”. In the heat of the action you’re voice will sound much different than when you were sitting down relaxed and programming the circuit. To achieve higher accuracy word recognition one needs to mimic the excitement in one’s voice when programming the circuit. These factors should be kept in mind to achieve the high accuracy possible from the circuit. This becomes increasingly important when the speech recognition circuit is taken out of the lab and put to work in the outside world.

2.8 VOICE SECURITY SYSTEM

This circuit isn’t designed for a voice security system in a commercial application, but that should not prevent anyone from experimenting with it for that purpose. A common approach is to use three or four keywords that must be spoken and recognized in sequence in order to open a lock or allow entry.
The ability to listen to one person speak among several at a party is beyond the capabilities of today’s speech recognition systems. Speech recognition systems cannot (as of yet) separate and filter out what should be considered extraneous noise. Speech recognition does not understand speech. Understanding the meaning of words is a higher intellectual function. Because a circuit can respond to a vocal command doesn’t mean it understands the command spoken. In the future 4.1, voice recognition systems may have the ability to distinguish nuances of speech and meanings of words, to “Do what I mean, not what I say!”.
Speech recognition is divided into two broad processing categories; speaker dependent and speaker independent. Speaker dependent systems are trained by the individual who will be using the system. These systems are capable of achieving a high command count and better than 95% accuracy for word recognition.
The drawback to this approach is that the system only responds accurately only to the individual who trained the system. This is the most common approach employed in software for personal computers. Speaker independent is a system trained to respond to a word regardless of who speaks. Therefore the system must respond to a large variety of speech patterns, inflections and enunciation's of the target word. The command word count is usually lower than the speaker dependent however high accuracy can still be maintain within processing limits. Industrial applications more often require speaker independent voice recognition systems.

Cite this article as: M.Shabnam Banu, M.S.Fathima Bevi, K.Thilagavathi. “AN INTERACTIVE IMPLEMENTATION ON A SMART PHONE FOR DISABLED PERSONS TO ACCESS HOME APPLICATIONS.” *International Conference on Information Engineering, Management and Security (2015)*: 338-342. Print.

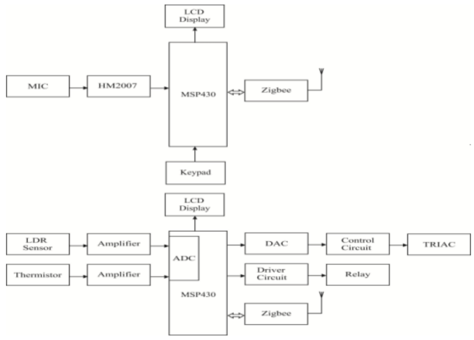


Figure 2.2 block diagram

The first module contains the mike through which voice input is given that is then stored in the hm2007 voice recognition system. The voice recognition system transfers the voice command to the next module with the help of Zigbee which acts as the transmitter during transfer.

In the second module receives the voice signals along with the keyboard text input. Here the Zigbee acts as the receiver and the signals are converted from analog to digital and vice versa using respective convertors. The light on and off options depends upon the temperature of the room.

2.8 LIGHT DEPENDENT RESISTOR – LDR

Figure 2.3 Light Dependent Resistors

Two cadmium supplied photoconductive cells with spectral responses similar to that of the human eye. The cell resistance falls with increasing light intensity. Applications include smoke detection, automatic lighting control, and batch counting and burglar alarm systems. Figure 2.3 is light dependent resistors. Intensity of the light is varies depends on temperature value.

3. RESULTS AND DISCUSSIONS

3.1 Communication between MSP430 and hm2007

The voice command send trough the voice recognition of hm2007 to msp430 microcontroller UART communication unit from there the data's are send to the voice through zigbee module. MSP430 provides the main part of the project. Both receiver and sender side contains MSP430. HM2007 is used. The voice received from the HM2007 is stored in MSP430. The voice signal is then transferred to receiver side through zigbee.

This project consists of keyboard. Keyboard consists 5 butoons that are used for giving different commands. The various instructions are relayon, light on, light off, bright, dim. HM2007 is used to recognize the vioce signal that is stored in MSP430.

4. CONCLUSION

The proposed system proposes methodology that is more efficient than the existing system that uses design display. The existing system has the disadvantages dealing with the software. The proposed system defines a effective speech recognition system efficiently store the voice input and has extended memory that increases the performance of the system. The proposed system would be more helpful for the disabled people to more than that of the existing System since touching design would lead to more software error problem. The proposed system overcomes all the disadvantages of the existing system and helps the disabled people more effectively.

5. FUTURE WORK

Single home appliance is used to implement the project. Many home appliances with advanced microcontroller will be implemented in the future work. Blind people can be used to implement the project as the future work. many different voices will be stored in the speech recognition. Voice quality accuracy is achieved.

Cite this article as: M.Shabnam Banu, M.S.Fathima Bevi, K.Thilagavathi. "AN INTERACTIVE IMPLEMENTATION ON A SMART PHONE FOR DISABLED PERSONS TO ACCESS HOME APPLICATIONS." *International Conference on Information Engineering, Management and Security (2015)*: 338-342. Print.

accurately track the time scale. Each output sample from the QRS discriminator is compared against a set threshold to detect the presence of a beat. Pulse period is incremented by one during every sample period. Because each sample occurs every 1/512 second, it is easy to track the time scale based

REFERENCES

[1] Alexandrov, A., (2010) "Characteristics of single-item measures in Likert scale format", The Electron Journal of Business Research Methods. Vol.No. 8, pp.1 -12.

[2] Costa, L. C. P., Almeida N. S., Correa, A. G. D., Lopes, R. D. and Zuffo, M. K., (2012), "Accessible Display Design to Control Home Area Networks", IEEE Transactions on Consumer Electronics, Vol. 59, No. 2.

[3] Costa, L.C.P., Ficheman, I. K., Correa, A. G. D., Lopes, R. D. and Zuffo, M. K., (2012), "Accessibility in digital television: designing remote controls", IEEE Transaction on Consumer Electronics, Vol. 58, No.2, pp.605-611.

[4] Casey, D. E., (1992) "Descriptive research: techniques and procedures." In: R. Colwell (Ed.). Handbook of research on music teaching. Schirmer: New York, pp. 115-123.

[5] Do, J., Jang, H., Jung, S. H., Jung, J. and Bien, Z., (2005), "Soft remote control system in the intelligent sweet home", IEEE International Conf. on Intelligent Robots and Systems, pp. 3984-3989.

[6] Ghorbel, M., Arab, F., Monhtari, M., (2008) "Assistive housing: case study in a residence for elderly people", IEEE Second International Conf. on Pervasive Computing Technologies for Healthcare, pp.140-143.

[7] Urokohara, H., Tanaka, K., Furuta, K., Kurosu, M., (2000),"NEM: Novice Expert ratio Method. A usability evaluation method to generate a new performance measure", Extended Abstracts of ACM Conference on Human Factors in Computing Systems, pp. 185–186.

[8] Valles, M., Manso, F., Arredondo, M. T. and Del Pozo, F., (1996), "Multimodal environmental control system for elderly and disabled people" 18th Annual International Conf. of the IEEE in Eng. in Medicine and Biology Society, pp. 516-517.

[9] Vogel, B., Kurti, A., Milrad, M., Kerren, A., (2010) "An Interactive web-based visualization tool in action: user testing and usability aspects", IEEE International Conf. on Computer and Information Technology, pp. 403-408.

[10] Xianmei, W., Lingyan L.,Ti, D., Zhiliang, W.,(2010) "Smart home control system for the disabled using the head and the mouth movement". IEEE 29th Chinese Control Conf., pp. 3527-3530

[11] Zhu, J., Gao, X., Yang, Y., Li, H. and Cui, Z. A. X., (2010), "Developing a voice control system for zigbee-based home automation networks", IEEE 2nd International Conf. on Network Infrastructure and Digital Content, pp. 737-741.

Cite this article as: M.Shabnam Banu, M.S.Fathima Bevi, K.Thilagavathi. "AN INTERACTIVE IMPLEMENTATION ON A SMART PHONE FOR DISABLED PERSONS TO ACCESS HOME APPLICATIONS." *International Conference on Information Engineering, Management and Security (2015)*: 338-342. Print.