

2nd Annual International Conference on

Innovative Techno Management Solutions for Social Sector

IEMCON 2012

17th-18th January, 2012.

Venue: Science City, Kolkata.



Conference Proceedings

It has become very necessary to make modern management and technology available to all sections of the Society to reap maximum benefit of emerging avenues. Technology has grown leaps and bounds but the fruits have reached only the well to do sections of the Society. The gadgets used by the neo-rich, the cars used by the affluent, the communication devices used by the powerful; all caters to the people of the top echelons of the Society.

It is high time that academicians, bureaucrats, entrepreneurs and government authorities come together to put in place a system wherein the benefit of technological advancement and growth of managerial capacity reaches every section of the society and culminate in inclusive growth and development, especially the rural population.

Imagine a situation where a rickshaw puller has to exert bare minimum effort to make his rickshaw move. If our women folk can draw water from a deep tube well with minimum effort there cannot be a better service to our rural siblings.

Carrying of water is definitely a problem of serious nature which our rural woman faces. If a mechanized system for ferrying water house to house can be adopted, a very notable percentage of woman can be saved from getting affected with spondylitis and osteoporoses. We envisage a day when the bullock cart is replaced with a mechanised form of transport. Notable portion of ripe fruits fall off for want of timely plucking and lack of preservation. Can we not find out a way to avoid this criminal waste of our natural resource?

We wish to bring modern management and latest technology to the doors of each citizen in general but to the downtrodden in particular. If the life of the rural people can be made even a bit comfortable by adopting any of the latest available know-how, our mission will be successful.

Keeping the above in mind we are organizing the subject conference so that all stake holders can exchange ideas and prepare a road map for future.

ACKNOWLEDGEMENT

IEMCON - 2012 is thankful to have Tata Consultancy Services
as Publication Partner



WE ARE VERY GRATEFUL TO OUR SPONSORS

INSTITUTE OF ENGINEERING & MANAGEMENT

IEEE

TATA CONSULTANCY SERVICES

AIRCEL

INFOSYS

NATIONAL ENTREPRENEURSHIP NETWORK

COGNIZANT

ACCLARIS

IBM

SUNPLANT

RN CONSULTANCY

ZEE AKASH PVT. LTD.

&

TELEVISION PARTNER

24 GHANTA

RADIO PARTNER

POWER FM

NEWS PARTNER

TIMES OF INDIA

PRINT & PUBLICATION PARTNER

TATA CONSULTANCY SERVICES

PROGRAM COMMITTEE

CHEIF PATRON

T.C.Dutt, IAS

Chairman, IEM Governing Body

PATRONS

| | |
|-----------------------------------|-----------------------------------|
| Prof. Sabyasachi Sengupta | VC, WBUT. |
| Prof. Pradip Narayan Ghosh | VC, Jadavpur University. |
| Prof. Ajay Ray | VC, BESU. |
| Prof. Suranjan Das | VC, Calcutta University. |
| Prof. Ashok Ranjan Thakur | VC, West Bengal State University. |
| Prof. Satyajit Chakrabarti | Director, IEM, Kolkata. |
| Prof. B.B.Ghosh | IEM, Kolkata. |

ADVISORY COMMITTEE

Chairman

Prof. Manish Kumar Mukherjee IEM, Kolkata.

Co-Chairman

Prof. Dipak Chatterjee IEM, Kolkata.

Prof. P.K.Mishra IEM, Kolkata.

Prof. Subrata Basak IEM, Kolkata.

Prof. Satyajit Chakrabarti IEM, Kolkata.

Prof. Biswajoy Chatterjee IEM, Kolkata.

Members

Prof. R. E. Critoph Warwick University, UK

Prof. Harpreet Singh Wayne State University, USA

Prof. Partha Dasgupta Arizona State University, USA

Prof. Sivaji Chakravorti Jadavpur University, India

Prof. Bhaskar Gupta Jadavpur University, India

Prof. Hiranmay Saha BESU, India

Prof. S.P.Gon Chaudhury BESU, India

Prof. Ajit Pal IIT, Kharagpur, India

Prof. Prabir Kr. Biswas IIT, Kharagpur, India

Prof. Jayanta Mukhopadhyay IIT, Kharagpur, India

Mr. Kushal Banerjee TCS

Mr. Abhijit Bhattacharya TCS

Prof. S. Nayek IEM, Kolkata

TECHNICAL COMMITTEE

Chairman

Prof. Asim Kar IEM, Kolkata

Co-Chairman

Prof. Alok Kumar Das IEM, Kolkata

Prof. P.K.Sinha Roy IEM, Kolkata

Prof. Debika Bhattacharya IEM, Kolkata

Prof. Mohuya Chakraborty IEM, Kolkata

Members

Prof. Goutam Saha IIT, Kharagpur, India

Prof. D. Mukherjee BESU, India

Prof. Nabendu Chaki Calcutta University

Prof. Amlan Chakraborty Calcutta University

Prof. Somnath Sarkar Calcutta University

Prof. Chandan Kr. Sarkar Jadavpur University

Prof. C.R.Sinha Jadavpur University

Prof. Saktipada Nanda IEM, Kolkata

Prof. S.K.Das IEM, Kolkata

ORGANIZING COMMITTEE

Chairman

Prof. Ranjan Ghosh IIM, Kolkata

Co-Chairman

Prof. K.K.Ghosh IEM, Kolkata

Members

Mrs. Gopa Goswami IEM, Kolkata

Prof. Arun Kumar Bar IEM, Kolkata

Prof. Abhishek Bhattacharya IEM, Kolkata

Prof. Indraneel Mukhopadhyay IEM, Kolkata

Prof. Sukalyan Goswami IEM, Kolkata

Prof. Ratna Chakraborty IEM, Kolkata

Prof. Moutushi Biswas Singh IEM, Kolkata

Prof. Susmita Mukherjee IEM, Kolkata

Prof. Arindam Chakraborty IEM, Kolkata

Prof. Biswadip Basu Mallik IEM, Kolkata

Prof. Koushik Banerjee IEM, Kolkata

Prof. Ranjan Dasgupta IEM, Kolkata

Co-Ordinators

Prof. Himadri Nath Saha IEM, Kolkata

Prof. Malay Ganguly IEM, Kolkata

Prof. Sanghamitra Poddar IEM, Kolkata

Prof. Prabir Kumar Das IEM, Kolkata

CONTENTS

| Name Of the Session | Page no. |
|---|----------|
| MEDICAL ELECTRONICS | 1 - 45 |
| COMPUTER & COMMUNICATION SYSTEMS | 46 - 76 |
| NEURAL &ARTIFICIAL NETWORKS | 77 - 114 |
| ANALOG & DIGITAL COMMUNICATION | 115- 153 |
| SOCIAL ENGINEERING & SOCIAL SCIENCE | 154-197 |
| APPLICATION FOR IT IN INDUSTRY | 198-250 |
| PATTERN RECOGNITION IN ELECTRICAL AND OPTICAL DOMAIN | 251-315 |
| FINANCE AND MANAGEMENT | 316-370 |
| SIGNAL PROCESSING & MOBILE COMMUNICATION | 371-400 |
| ANTENNA & PROPAGATION | 401-437 |

N.B. Please use the Bookmarks provided to find the respective Papers

MEDICAL ELECTRONICS

Date : 17th January,2012

Chairman : Prof. Monisha Chakraborty

Co-Chairman : Prof. Kamal Kanti Ghosh

Rapporter : Prof. Arunava Mukherjee

| Paper ID | Paper Title | Authors |
|----------|--|--|
| 1 | Total Quality Management in Health Sector | Abhijit Chakraborty,Sudip Kumar Deb and Ranjan Bhattacharya |
| 24 | Extraction of ECG Fiducial Points Using Discret Wavelet Transform | Swati Banerjee and Madhuchhanda Mitra |
| 25 | Digital Filtering of Amino Acid Sequence For Prediction of Cancer Cell | Suman Saha and Soma Barmanmandal |
| 30 | Detection Of Myocardial Infarction by Fourier Co-Efficient Analysis of ECG | Deboleena Sadhukhan and Madhuchhanda Mitra |
| 60 | A Comparative Study on the Accuracy of ECG Detection for Arrythmia By Different Wavelet Families | Anandarup Mukherjee, Sanjukta Chatterjee, Kajari Chattpadhyay, K K Ghosh |
| 78 | Design of Tele-healthcare Interactivity:A Semi Formal Approach | Imon Banerjee, Shrutilipi Bhattacharjee |
| 88 | Electrocardiografic Data Compression Using Wavelet | Aritra Basu,Amit Shankar Sinha, Arighna Deb and Tejendra Nath Sen |
| 80 | A Mathematical Analysis Of Blood Flow Through Stenosed Arteries: A Non-Newtonian Model | B.Basu Mallik , Saktipada Nanda |

Total Quality Management in Health Sector

Abhijit Chakraborty \$ Ranjan Bhattacharya @ Sudip Kr. Deb

Research Fellow, Dept. of Production Engineering, Jadavpur University, Kolkata-32,

Email: chakrabortyabhijit100@gmail.com,

\$ Professor, Production Engineering Dept., Jadavpur University, Kolkata-32, India

@ Professor, M.E. Dept., Assam Engineering College, Guwahati-13, India.

Abstract: Total Quality Management (TQM) is a powerful system to guarantee organization's survival in world-class competition. TQM is the art of managing the whole to achieve excellence. Health sectors in the country, which have a tradition over the years, are today faced with stiff competition not only from local units in the region but also from globally renowned hospitals. It is imperative that the TQM model, which has been successful in transforming business and making them competitive, be adapted in health sectors also while maintaining the societal responsibility as a core value in the health sector. The theory of TQM has to be applied to the health sector process in order to sustain the competitive edge. The integration of the developmental ideas into a total quality management framework, hitherto considered will be attempted in this paper. It also describes the application of theory of TQM to health sector and also provides guidelines for the betterment of the health sector services.

Keywords: Total quality management, health sector, system, hospital management.

1.0 INTRODUCTION:

Total Quality Management (TQM) is a framework that provides the capabilities to integrate the social needs along with the business compulsions of the health system. TQM is an evolving system of practices, tools, and training methods for managing companies to provide customer satisfaction in a rapidly changing world. TQM improves the performance of companies in several areas; eliminating

product defects, enhancing attractiveness of product design, speed in service delivery, and reducing costs among others. TQM is an approach to management for total quality, for effectiveness and competitiveness, involving each and every activity and person at different levels in the organisation. TQM provides an environment where fear is eliminated, where all the employees take pride in their work,

where they feel respected and accepted, where they feel part of the same team, and where they strive not only for own interests, but for the whole organisation. This needs the establishment the following three fundamental characteristics: 1. Commitment to never ending quality improvement and innovation, 2. Scientific knowledge of the proper tools and techniques, 3. Involvement - all in one team. Total Quality Management (TQM) has proved its significance and applicability in all industries. TQM, conjointly with HRD interventions, can be a possible solution to reduce the high rate of attrition and preclude or defer the managerial obsolescence. The motives, personality traits, employee satisfaction survey and finally the job behaviours (including organisational commitments) need periodic measurement, review, and thus initiate remedial actions.

In this paper, basic principle of total quality management (TQM) philosophy, considerations while implementing and maintaining TQM, quality system as a system, basic props of quality and golden ten rules of TQM have been briefly discussed. Thereafter, few major problems of health sector have been

highlighted followed by key attrition factors.

2.0 TQM PHILOSOPHY -TEN GOLDEN RULES:

TQM philosophy: The basic principles for the TQM philosophy of doing business are to satisfy the customers, satisfy the suppliers, and continuously improve the business processes. These are achieved through a structured management approach centered on quality, based on participation of all of its members and aiming at long-term success through customer satisfaction, and benefits to all members of the organisation and to society. The employees are to be treated as internal customers. Quality management has been defined in ISO 9000:2000 as "co-ordinated activities to direct and control an organisation with regard to quality". While implementing and maintaining TQM, one should remember that it is to be started at the top and allow it to cascade down throughout the organization.

2.1 Quality System:

To get the best return of a quality system, we should take a system approach. Quality System can be considered as

having four characteristic elements viz. Content, Structure, Communication and Control. These elements are common to all organisations irrespective of their sizes or processes/products.

2.2 Three Basic Props of Quality:

From the foregoing paragraphs we can view the quality system as an entity that translates the quality policy of an organisation to its quality practice. This quality policy emanates from the philosophy, which the organisation adopts. While everyone in the organisation is involved in the quality practices, possibly senior executives play a greater role in evolving a quality policy and this is to be generated and guided by the quality philosophy of the top most executives.

Philosophizing on quality, the three basic props of a quality system are:

Vision - Quality philosophy
Mission - Quality policy
Realisation - Quality practice

2.3 The Golden ten rules of TQM:

Thus, in line of above discussion the golden ten rules of Total Quality Management can be written as:

- Be reliable -Be responsible
- Be competent -Be accessible
- Be courteous - Be credible
- Be communicative -Be secure
- Be emphatic -Be attractive

3.0 HEALTH SECTOR PROBLEMS:

Health sector problems are many. Precisely there are four areas where the problems generate a) Availability of Proper Infrastructure b) Lack of Expert physician for critical diseases c) Skilled manpower at various levels d) Awareness of Environment, Health and Safety. It is evident that first two points are addressed in the policy of any state and depend on development of overall economy, socio-economic structure. But point number three and four can be handled with various tools and techniques. In these studies, TQM is considered as an effective tool for effective management of Hospital and Health care.

4.0 SIGNIFICANCE OF HOSPITAL MANAGEMENT:

Science of management applies not only to industry and commerce but also to all fields of employment. Management fundamentals and principles are universal. They apply to all areas of human activities, though the techniques and procedures of their application may

differ, depending on the nature to be performed. Thus, good and professional management is essential for all the fields of human activity and thus, the hospitals are no exception.

But in fact, the use of modern management techniques for the optimum utilization of scarce resources is widely accepted in the industry of our country; while its use in the area of public health, especially in hospital administration, where the system suffers from paucity of resources has not had the same reception.

Hospital administration in a developing country has got a lot of relevance in the present day development of medical science in India. We consider a developing country and in hospitals as very nicely described by Late Dr. V. Ramalingawamy, we have continued to follow the same administration, which was left behind when we became independent. Hardly any change or developments have been made. The concept of health has changed, people's expectations have also increased a great deal but we are still following the technique and the method of administering which is so archaic and out of date.

Now hospitals are fast becoming not only centers of healing but also centers of health and prevention of disease-concept of community health. The hospitals, therefore, have not only increased in complexity busy-with activity of team of experts, sophisticated diagnostic apparatus and equipment, laboratories, specialization and super-specialization, but also in size and the patients that they can look after. This has added more and more difficulties rather than solving difficulties in hospital administration. Thus, in a changing society, hospital management in its right perspective is very important. Moreover, hospitals are very complex organizations with a variety of jobs to be performed by various personnel – specialized as well as otherwise. Interactions and intra-action of several heterogeneous groups constitute a great challenge to hospital managers and behavioral scientists. Besides this, the hierarchy in which the staff members have to work is very sensitive and constant tension exists. Therefore, the proper training is required at every level, where persons have to deal with other persons.

Medical care is the major function of hospital organization. Medical care may be defined in different ways in different contexts, but as the WHO definition has it, it may symbolize personal health care. It encompasses preventive, curative and rehabilitative measures. This ‘medical care’, today, is no more a one –man show. It involves the services of a large number of persons, specialized in different areas. It also requires supporting services of paramedical persons and clerical personnel. Thus, it is a teamwork. There are different areas of work ranging from simple health check-up to detect disease, out-patient examinations, conducting in-patient wards, operation theatre, intensive care-unit, laboratory, X-ray unit, C.T scan etc. Over and above this, hospitals have to arrange for good food services, housekeeping, engineering, fire fighting and security, etc. Hence, management of hospitals needs the same care and consideration, which is essential for running a business or industrial enterprise. It is important to note that good doctors may not always be good administrators. The management techniques are highly technical. So, utilisation of expertise available from

other fields is essential in some case in the initial period. This is the only way to develop managerial competence amongst medical administrators. The problems of management are more acute and difficult in developing countries where the needs are disproportionately higher than the available resources. This is more true for our hospitals and the problems can be dealt with by adopting modern management techniques in hospital organizations.

5.0 GENERAL CONSIDERATION

PERFORMANCE SYSTEM:

5.1.1 General Consideration:

In order to consistently provide the services of desired quality at optimum cost for providing effective and efficient patient-care services, a hospital management should develop, implement and continuously improve the appropriate quality system. Any such system consisting of organizational structure, procedures, processes and resources needed to implement the quality management.

The characteristics of hospital services in terms of infrastructure and systems form the starting point. The other set relates to quality considerations, which should flow thereafter. These may be

quantitative or qualitative. Both these types of characteristics should be evaluated by hospital authorities in terms of the following attributes:

- a) **Availability-** The services should be available to the patient wherever needed.
- b) **Reliability-** The services should be accurate and reliable; for example, if a blood sample is tested in a pathological laboratory, the values obtained should be accurate and free from errors.
- c) **Completeness-** The services rendered should be complete in all respects. If a laboratory test is done for an in-patient but report not sent to the ward, it is not a satisfactory service because it is not complete.
- d) **Timeliness** – Rendering the services in time is a very important consideration, which otherwise, may even cost a patient's life. The medical service, in order to be called as good quality or satisfactory, should adhere to the time specified.

- e) **Courtesy** – A hospital service is a highly personalized one and should be rendered with due courtesy and pleasant behaviour to the customer. A patient, while suffering from a disease leaves a sigh of relief through courteous behaviour.
 - f) **Economy**- A hospital service should be provided to the patient at reasonable cost, which is generally affordable.
 - g) **Conscientiousness**- A strong service to redress grievances and make good losses arising out of lapses in services is also an attribute to quality of services rendered.
 - h) **Other considerations**- These may include confidentiality, hygienic considerations, comfort, aesthetics of environment and communication.
- 5.1.2 Customer (i.e. patients) Considerations:**
- Communication and interactions with customers:**
- One of the major indicators of quality of health care services is the feedback, which percolates down the line to every employee in the hospital. Another major element in the service area is the concept of consideration. Usually, price is a very

important concept in any industry and service. But in health care service area, price has been replaced by the element of consideration. Consideration is price plus loss of dignity, waiting time, anxiety, etc; the latest concept in management is that of customer delight. It has now become essential not only to ensure customer satisfaction but also to aspire for customer delight. Customer delight, for example, means that a receptionist is not merely courteous, but is warm, emphatic and well trained in counselling. Since health service is mainly people-oriented, prevention and reduction of human errors at all levels are of vital importance. This means that due priority would be given to human resource development in respective area of service.

Some common examples of waste, rework, negligence and error in patient care in hospitals are:

- X-ray / investigation ordered but not done
- Prolonged stay before and after the operation
- Consultation indicated and not carried out
- Wrong label on samples
- Wrongly numbered reports

- Missing requisitions, reports, etc.
- Inferior quality of drugs used
- Delayed reports
- Prescriptions not explained
- Medicine prescribed but not given by the nurse

- Incorrect patient identification
 Communication with customers involves listening to them and keeping them informed. Difficulties in communication or interaction with customers should be given prompt attention. These difficulties provide important information on areas for improvements in the service delivery process. The customers' perception of service quality is acquired often through communication with the service organizations' personnel and facilities.

5.1.3 Performance requirements:

Parameters have to be identified which describe the performance of each activity of the work phase. These parameters should relate to the quality of patient care activity and should be associated with the appropriate type of service required. In order to define standards of acceptance performance, the value of each parameter needs to be specified in terms of both the acceptable level and allowable tolerances.

Customer's needs and expectations from various activities of hospital services should be kept in mind while laying down various quality parameters. Some of these customer satisfiers are given below:

- a) Patients should get cured,
- b) Curing should be done in the shortest possible time and at the lowest cost,
- c) Cure should be lasting,
- d) Psychologically the patients should feel that they are being treated well,
- e) The hospital environment should be clean and inviting,
- f) The environment should prevent the spread of diseases to others,
- g) Emergencies should be attended to promptly,
- h) Measures should be taken in minimum time.

6.0 BUSINESS AND SOCIAL RESPONSIBILITY IN HEALTH CARE SECTOR:

India Spends 5.2 % of its Gross Domestic Product(GDP) for the growth of Health sector. Growth of this area has started from metro city to developed cities and slowly spreading to urban areas. Owing huge population growth and increased

criticality of diseases, it is inevitable to integrate Total Quality Management in these Hospitals and Health care units for quick disposal and effective service for life care. There are four key areas for assessing the effectiveness of study.

- i) Incident free Work Place,
- ii) Regulatory Compliance and process for continuous improvements,
- iii) Regular assessment of environment, Health and Safety,
- iv) Enhancement of sustainability of business

7.0 CONCLUSION:

The health sector can be improved to a large extent by the application of TQM. For it, the golden rules of TQM are to be strictly followed with the three basic props of quality system i.e., vision, mission and realization. When a health sector is going to look after its patients with general considerations of the performance systems such as economy, availability, reliability, timeliness, courtesy, conscientiousness, etc. along with customers i.e., patients' considerations by addressing all their queries and trying to give proper and prompt quality service, the level of

service in health sector is going to improve definitely.

REFERENCES:

- [1] Davies B., Wilson D., TQM - Organising for success, processing's of the Third International Conference TQM-3 (London UK), IFS Ltd / Springer - Verlag, 1990, PP 105-112.
- [2] Driscol, Robert and Eckstein, Daniel. G. (1882), Pfeiffer and Goodstein, 1982 Manual, University Associates, Inc. pp. 102-107.
- [3] Ersoz, Clara Jean, M D., St. Clair Hospital, "TQM: Healthcare's Roadmap to the 21st Century", presented at North Coast Quality Week, New Paradigms in Health, Erie, PA October 2, 1992.
- [4] Ghosh, Sadhan. K. "Total Quality Management", Oxford Publishing House, Kolkata.
- [5] Haldar, U.K. (1995) Project Work (Unpublished) entitled "Occupational values and Managerial styles: Study in a Central Government Departmental Enterprise" submitted to Indira Gandhi National Open University in partial fulfilment of Master degree in Business Administration.
- [6] Harris, John W. and Baggett, J Mark, Editor Quality Quest in the Academic Process Published by Stanford University, Birmingham, Alabama 35229, and by GOAL / QPC (1992).
- [7] Price R C Gaskil G P TQM in Research and Practice, Proceedings of 3rd International Conference TQM-3 (London, UK), IFS Ltd / Springer Verlag 1990, pp. 77-87.
- [8] Rao, T. V. (2006) "Old pillars of people management", Indian Management, Volume 45, Issue 6, June 2006, pp.74-81.
- [9] Raychoudhuri, Suvro, "Manpower Planning and Employee Attrition Analytic - A Markov Analysis Attempt for Attrition-Rate Prediction and Stabilisation".
- [10] Reddin, W.J., (1987) "How to make your management style more effective", Mc GRAW HILL Book Company (U.K) Limited, Maidenhead, Berkshire, England.
- [11] Rosenberg, Morris. (1957) "Occupations and Values". The Free Press Journal, Glencoe, Illinois, USA.
- [12] Tribus, Myron, "Deming's Way Quality First NSPE Publication #1459, 4* Edition, National Institute for

2nd Annual International Conference IEMCON 2012

Engineering Management and Systems,
1420 King Street, Alexandria, VA
22314.

Extraction of ECG Fiducial Points using Discrete Wavelet Transform

Swati Banerjee¹, M. Mitra¹

¹ Department of Applied Physics, Faculty of Technology, University of Calcutta
Kolkata-700009, India

Abstract— A feature extraction technique in the QT segment of digitized Electrocardiograph recordings based on discrete wavelets transform (DWT) is proposed in this paper. A multiresolution approach along with an adaptive thresholding is used for the detection of R - peaks. Then Q, S peak, QRS onset and offset points are identified. Finally, the T wave is detected. By detecting the baseline of the ECG data, height of R, Q, S and T wave are calculated. For R-peak detection, proposed algorithm yields sensitivity and positive predictivity of 99.8% and 99.6% respectively with MIT BIH Arrhythmia database, 99.84% and 99.98% respectively with PTB diagnostic ECG database. For time plane features, an average coefficient of variation of 3.21 is obtained over 150 leads tested from PTB data, each with 10000 samples.

Keywords- Multiresolution, QRS vector, Coefficient of variation, Fiducial points.

I. INTRODUCTION

The Cardiovascular disease remains the main cause of mortality in many region of the world, and several studies demonstrate the importance of reducing the time delay to treatment for improving the clinical outcome of the patients. From electrical point of view, the heart is situated at the center of the electrical field it generates. ECG is a well structured signal that can be decomposed into five different component waves, namely P, Q, R, S and T wave followed by a conditional U wave. These waves along with the morphologies are of clinical significance because most of the information in the ECG is found in the intervals and amplitudes defined by peaks and boundaries of the characteristic waves. [1]

Good performance of an automatic ECG analysing system depends upon the reliable and accurate detection of the QRS complex and P and T waves along with the measurement of the QT segment. The automatic delineation of the ECG has been widely studied and algorithms are developed for QRS detection and wave delineation [2]-[4]. A real time QRS detection algorithm implemented in assembly language is developed by Pan et. al. [5]. Some old detectors are presented in [6]-[8]. Denoising the signal becomes an essential part of the analysis process. The noise sensitivity of nine different QRS complex detectors are measure and analysed by G. M. Friesen et al [8].

In an attempt to develop more accurate algorithms for extraction of features different approaches are used. Some algorithms are based on mathematical models [9]-[10]. A simple, mathematical based method along with the concept of data structure is used to obtain the complex [11]. Some other approaches like matched filters [12], ECG slope

criteria [13], [14], second order derivatives [15], wavelet transforms [16], [17] are also studied. In [16], a multi scale QRS detector including a method for monophasic P and T waves was proposed. Method based on artificial neural network [18], was developed for wave delineation. ECG beats are classified by neuro-fuzzy networks where predefined feature vectors are used [19]. An association rule mining based classification is proposed in [20]. An unsupervised classification technique for classification of extrasystoles using bounded clustering algorithm is explored in [21].

In this paper, a discrete wavelet transform (DWT) based ECG feature extraction technique in the QT region is proposed. Wavelet transform is a well-suited tool for non-stationary signals like ECG. After decomposing the signal using DWT, well localized frequency domain features are obtained at different levels of decomposition. QRS complex and T wave have separate frequencies and those frequency bands are identified along with the frequency levels for different noises. The signal is denoised by discarding those frequency bands corresponding to noise levels. A multiresolution approach along with thresholding is used for the detection of R - Peaks in each cardiac beats. Hence, the heart rate is calculated. Then other fiducial points (Q and S) are detected by differentiation and slope criteria QRS onset and offset points are detected. Finally, the T wave peak is detected and QT interval is measured. Baseline is also detected in the TP segment and height of R, Q, S and T waves are calculated

II. THEORY

A. Wavelet Transform

Wavelet transform is a linear transform, which decomposes a signal into components that appears at different scales (or resolution). Time localization of spectral components can be obtained by multiresolution wavelet analysis, as this provides the time-frequency representation of the signal.

The DWT analyses the signal at different resolution (hence, multiresolution) through the decomposition of the signal into several successive frequency bands. The DWT utilizes two set of functions $\phi(t)$ and $\psi(t)$, each associated with the low pass and the high pass filters respectively [22,24]. These functions have a property that they can be obtained as the weighted sum of the scaled (dilated) and shifted version of the scaling function itself:

$$\phi(t) = \sum_n h[n] \phi(2t-n) \quad (1)$$

$$\psi(t) = \sum_n g[n] \phi(2t - n) \quad (2)$$

Here, $h[n]$ and $g[n]$ is the halfband lowpass filter and highpass filter respectively.

Conversely, a scaling function $\phi_{j,k}(t)$ or wavelet function $\psi_{j,k}(t)$ that is discretized at scale j and translation k can be obtained from the original (prototype) function $\phi(t) = \phi_{0,0}(t)$ or $\psi(t) = \psi_{0,0}(t)$.

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \quad (3)$$

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k) \quad (4)$$

Where: j controls the dilation or translation and k denotes the position of the wavelet function.

Different scale and translation of these functions allows us to obtain different frequency and time localization of the signal. Decomposition of the signal into different frequency band is therefore accomplished by successive lowpass and highpass filtering of the time domain signal. The original time domain signal $x(t)$ sampled at 1000 samples/sec forms a discrete time signal $x[n]$, which is first passed through a halfband highpass filter $g[n]$, and a lowpass filter $h[n]$ along with down sampling by a factor of 2.

Filtering followed by subsampling constitutes one level of decomposition, and it can be expressed as follows:

$$d1[k] = y_{\text{high}}[k] = \sum_n x[n] \cdot g[2k - n]. \quad (5)$$

$$a1[k] = y_{\text{low}}[k] = \sum_n x[n] \cdot h[2k - n]. \quad (6)$$

Where, $y_{\text{high}}[k]$ and $y_{\text{low}}[k]$ are the outputs of the highpass and lowpass filters, respectively.

III. PROPOSED METHOD

The proposed algorithm for feature extraction is described in Figure 1. This module subjects the signal to DWT decomposition and then goes through various stages to obtain the time plane features of the ECG signal.

In the ptb-db data files, sampling frequency of the ECG data is 1 kHz. The selection of the scale (frequency band) is based on sampling frequency of the current data file along with a prior knowledge of the clinical bandwidth of the signal and various noise levels. The noise frequencies and ECG bandwidth is already known [8].

A. Denoising the Signal

ECG signal gets corrupted due to various types of noises like power line interference, electrode contact noise, motion artifacts, baseline drift and electrosurgical noise which often hinders proper feature extraction. The drift of the baseline is caused due to respiration and the baseline variation frequency is 0.15~0.8 Hz. Motion artifacts are transient baseline changes within the frequency range of baseline drift. These are eliminated by removal of the lowest frequency component (A10), after decomposition of the ECG signal as shown in Fig. 2. Electrosurgical and Muscle Contraction noises are eliminated by discarding the details D1, D2. After removal of these noises the remaining ECG signal ranges from 0 -125 Hz.

Some noises still prevails in this range which are actually noise frequency overlapping with that of the original ECG. These are Power line interference and Electrode contact noise.

These two noises can be eliminated by the dyadic scale DWT technique by identification of the frequency band containing the frequency range of 59.5~60.5 Hz. Then we keep on decomposing the components containing this frequency range and finally discarding it.

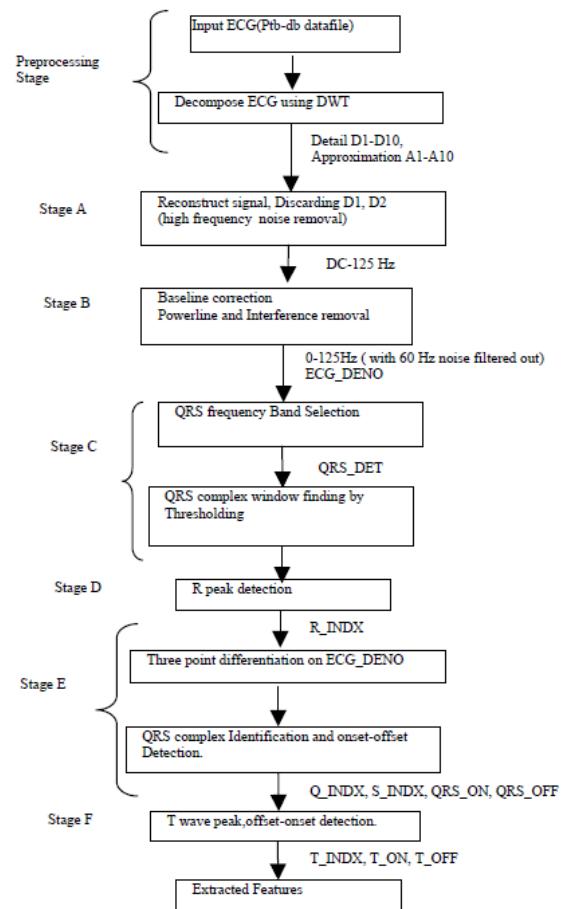


Fig1: Feature Extraction module

B. QRS Complex Band selection and finding the QRS window by thresholding

In this stage the objective is to find out the QRS region from the detail coefficients of the DWT decomposition and mapping of this zone in the denoised array ECG_DENO.

From Figure 2, it is observe that the QRS complex regions are more prominent in details at scale D3, D4 and D5, among which most of the energy of the QRS complex is concentrated[14],[18]. An array QRS_DET is formed by addition of components D4 and D5. Figure 3, shows the QRS_DET plot against ECG_DENO. To find the QRS complexes from QRS_DET, an empirically determined threshold value is set which is equal to the 15 % of the mean amplitude value of QRS_DET array.

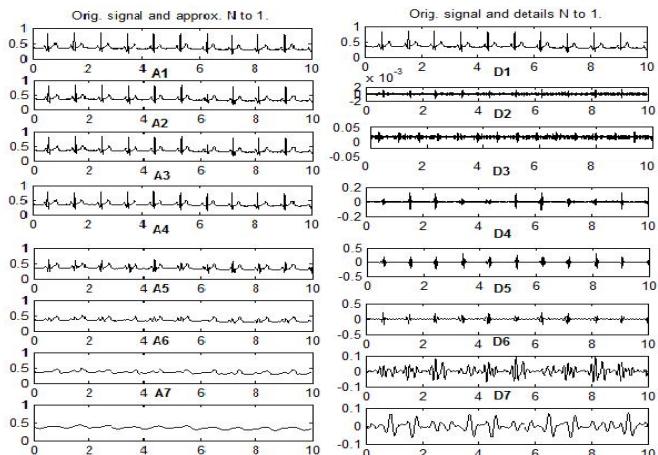


Fig 2: The DWT decomposition of ECG signal.

To detect the R peaks accurately, the indexes which exceed the threshold level are marked in the QRS_DET array. The region of the QRS complex starts from the first point where the absolute QRS_DET value becomes greater than the threshold. Since the maximum width of the QRS complex for any patient is not more than 160 ms [1], a fixed window of same width is searched in QRS_DET to detect the indexes where the threshold condition is satisfied. Between two consecutive searches a blanking period of 200 ms is offered [5].

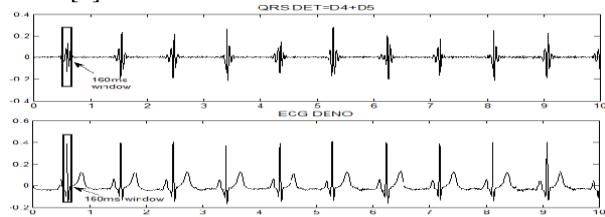


Fig 3: QRS complex located in Original signal from QRS_DET.

C. R peak detection (Q or S in case of inverted QRS)

The R (Q or S) peaks are the maximum (minimum) amplitude values within the QRS windows, set over ECG_DENO. In stage D, mapping each of the QRS starting index to the denoised signal ECG_DENO, the actual R peaks (Q or S peak) can be obtained by searching the local maximum (minimum) value within each QRS_START: QRS_START + 160 ms window. The amplitude of each R - index is again checked from ECG_DENO. If this is positive (negative), an R peak (an elongated or pathological Q or S) is identified. So, by this method all the peaks of the QRS complex are identified. All the R (Q or S) indexes are stored in an array R_INDX.

D. Q and S point detection, with QRS onset and offset detection and baseline detection

Once the R peak is detected accurately, the Q and S fiducial points are detected to find the complete QRS complex. Slope inversion method is used to find the Q and S points. Three point differentiation on ECG_DENO is taken using the formula stated below. Where h is the time division.

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} \quad (9)$$

Moving on either side of the detected R peak the points when the slope sign is inverted gives the Q and S points respectively. In case a detected peak is of negative

amplitude, the algorithm decides whether it is an elongated (pathological) Q or S wave depending on its nearby vicinity from the left or right side of the window. It is labeled as Q when it is near the left side of the 160 ms window and S when near the right side of the window. To find the onset and offset points of S and Q accurately, a tolerance value equal to $0.2 * \text{maximum slope of S wave}$ (Q wave is set). Whenever, the value becomes less than this tolerance limit, S wave ends and the offset is detected (Q wave ends and onset is detected). Baseline is the isoelectric line in the ECG signal. Most methods are based on the assumption that the isoelectric level of the signal lies on the area about 80 msec left of the detected R-peak, where the first derivative becomes equal to zero.

F. T wave Detection

After the offset of S wave is identified, the search for T wave starts. The R-R interval is calculated, and this distance is divided in a ratio of 2:1, corresponding to the T wave: P wave region [1]. Where, T wave is associated with the first cardiac beat and P wave with that of the second cardiac beat. Now the T wave detection window is estimated and the T wave amplitude is obtained by searching the absolute maximum value in this window.

IV. TESTING AND RESULTS

All the input data for this method has been randomly selected from Physionet, PTB diagnostic database [23]. Test results for R peak detection are shown in Table 1 and 2. For R-peak detection, proposed algorithm yields sensitivity and positive predictivity of 99.8% and 99.7% respectively with MIT BIH Arrhythmia database, 99.84% and 99.98% respectively with PTB diagnostic ECG database.

Where, Sensitivity (Se) = TP / (TP + FN).

Positive Predictivity(PP)=TP/(TP+FP)

TP = True positive (actual R peaks correctly detected as peaks)

FN = False negative (actual peaks not detected as peaks).

FP = False Positive (Non-peaks detected as peaks).

The denoising of the signal and baseline correction as explained in stage A and stage B is shown in figure 4.

Table 1: Beat Detection on PTB DB

| Patient id | Total beats | Detected beats | TP | FN | FP | Se (%) |
|--------------|--------------|----------------|--------------|-----------|----------|--------------|
| Patient024 | 7692 | 7662 | 7662 | 30 | 0 | 99.6 |
| Patient117 | 8040 | 8040 | 8040 | 0 | 0 | 100 |
| Patient007 | 9360 | 9341 | 9339 | 21 | 2 | 99.8 |
| Patien041 | 7680 | 7689 | 7680 | 0 | 0 | 100 |
| Patient027 | 10080 | 10060 | 10058 | 22 | 6 | 99.8 |
| Total | 42852 | 42832 | 42779 | 73 | 8 | 99.84 |

Table 2:Beat Detection on MITDB

| Total Beats | Detected Beats | TP | FN | FP | Se (%) |
|-------------|----------------|-------|----|----|--------|
| 19098 | 19062 | 19022 | 76 | 40 | 99.8 |

Some of the extracted features are tabulated in table 3.

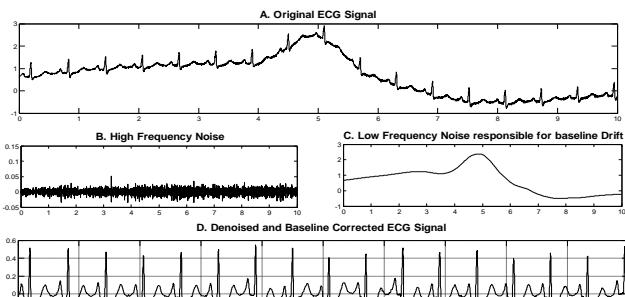


Fig 4: A. Original ECG signal, B. High Frequency Noise, C. Low Frequency noise, D. Denoised and Baseline Corrected ECG signal

VI. CONCLUSION

The proposed method uses multiresolution feature extraction using DWT, which can be used for feature extraction from ECG data. Additional features like T wave direction, pathological Q etc. can be used for characterization of the ECG wave. This paper also explores multiresolution analysis for identification of various frequencies present in an ECG signal. The noise frequencies are also identified and eliminated. The proposed method yields a sensitivity of 99.84 % and Predictivity of 99.92 % with ptb-db. With mit-db files, these figures are 99.6% and 99.5% respectively. Further, in future an ECG classifier can be proposed, for categorization of various kinds of abnormalities.

Table 3: Extracted Features

| Patient ID and record No. from Physionet | Lead | Q height | | | R height | | | S height | | | T height | | | |
|--|------|----------|------|---------|----------|------|---------|----------|------|---------|----------|------|------|-----------|
| | | Mean | S.D | C.V (%) | Mean | S.D | C.V (%) | Mean | S.D | C.V (%) | Mean | S.D | C.V | Direction |
| P004/S0020lre (MI) | V1 | 1.032 | .066 | 6.4 | I* | I | I | I | I | I | 0.356 | .036 | 10.2 | U* |
| | V2 | 1.129 | .080 | 7.1 | I | I | I | I | I | I | 0.473 | .029 | 6.2 | U |
| | V3 | 0.012 | 0 | 6.3 | 0.321 | .01 | 6.1 | 0.969 | .06 | 7.1 | 0.649 | .033 | 5.2 | U |
| | V4 | 0.785 | .050 | 6.4 | 0.591 | .033 | 5.7 | 0.736 | .02 | 3.2 | 0.549 | .068 | 12.4 | U |
| P013/s0045lre (MI) | V1 | 0.766 | .024 | 3.2 | I | I | I | I | I | I | 0.654 | .032 | 4.9 | U |
| | V2 | 1.318 | .054 | 4.1 | I | I | I | I | I | I | 0.254 | .018 | 7.1 | U |
| | V3 | 0.002 | I | I | 0.09 | .004 | 5.1 | 0.85 | .05 | 5.9 | 0.214 | .006 | 2.9 | U |
| | V4 | 0.069 | .003 | 4.9 | 0.39 | .020 | 5.2 | 0.574 | .024 | 4.3 | 0.137 | .005 | 3.8 | U |
| P177/s0291lre (normal) | V1 | 0.005 | I | I | 0.399 | .021 | 6.3 | 0.638 | .039 | 6.2 | 0.025 | .001 | 6.3 | In* |
| | V2 | 0.002 | I | I | 0.799 | .039 | 4.9 | 0.301 | .009 | 3.1 | 0.432 | .047 | 11.1 | U |
| | V3 | 0.009 | I | I | 0.776 | .055 | 7.1 | 1.610 | .07 | 4.9 | 0.695 | .056 | 8.2 | U |
| | V4 | 0.003 | I | I | 1.502 | .103 | 6.9 | 0.920 | .02 | 2.2 | 0.722 | .06 | 9.1 | U |
| P104/s0306lre (normal) | V1 | 0.010 | I | I | 0.125 | .006 | 4.9 | 0.973 | .07 | 7.9 | 0.035 | .001 | 4.2 | In |
| | V2 | 0.002 | I | I | 0.371 | .013 | 3.6 | 1.195 | .14 | 12.2 | 0.375 | .028 | 7.5 | U |
| | V3 | 0.003 | I | I | 0.996 | .040 | 4.1 | 1.629 | .08 | 5.5 | 0.341 | .017 | 5.2 | U |
| | V4 | 0.002 | I | I | 2.638 | .134 | 5.1 | 0.785 | .01 | 2.2 | 0.367 | .04 | 10.9 | U |

References

- [1] An introduction to Electro Cardiography by Leo Schamroth.
- [2] B.-U. Kohler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection," *IEEE Eng. Med. Biol. Mag.*, vol. 21, no. 1, pp. 42–57, Jan.–Feb. 2002.
- [3] C. Saritha, V. Sukanya, Y. Narasimha Murthy "ECG Analysis using wavelets", *Bulg. J. Phys.* , 35(2008) 68-77.
- [4] P.Ranjith et al., "ECG analysis using wavelet transform: Application to myocardial ischemia detection", *ITBM-RBM* 24 (2003) 44-47.
- [5] Pan et al. , "A real time QRS detection algorithm", *IEEE Trans. On Biomedical Engg.* Vol -32, No. -3, 1985.
- [6] O. Pahlm and L. Sörnmo, "Software QRS detection in ambulatory monitoring– A review," *Med. Biol. Eng. Comp.*, vol. 22, pp. 289–297, 1984.
- [7] F. Gritzali, "Toward a generalized scheme for QRS detection in ECG waveforms," *Signal Processing*, vol. 15, pp. 183–192, 1988.
- [8] G. M. Friesen et al., "A comparison of the noise sensitivity of nine QRS detection algorithms," *IEEE Trans. Biomed. Eng.*, vol. 37, pp. 85–98, Jan. 1990.
- [9] I. Murthy and G. D. Prasad, "Analysis of ECG from pole-zero models," *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 741–751, July 1992.
- [10] J. Vila, Y. Gang, J. Presedo, M. Fernandez-Delgado, and M. Malik, "A new approach for TU complex characterization," *IEEE Trans. Biomed. Eng.* Vol. 47. pp. 764-772, June 2000.
- [11] John Darrington, "Towards real time QRS detection: A fast method using minimal pre-processing", *Biomedical Signal Processing and Control* 1 (2006) 169–176.
- [12] A. S. M. Koelman, H. H. Ros, and T. J. van den Akker, "Beat-to-beat interval measurement in the electrocardiogram," *Med. Biol. Eng. Comput.*, vol. 23, pp. 213–219, 1985.
- [13] A. Algra and H. L. B. C. Zeelenberg, "An algorithm for computer measurement of QT intervals in the 24 hour ECG," in *Computers in Cardiology*. Los Alamitos, CA: IEEE Computer Society Press, 1987, pp. 117–119.
- [14] I. K. Daskalov, I. A. Dotsinsky, and I. I. Christov, "Developments in ECG acquisition, preprocessing, parameter measurement and recording," *IEEE Eng. Med. Biol. Mag.*, vol. 17, pp. 50–58, 1998.
- [15] J. G. C. Kemmelings, A. C. Linnenbank, S. L. C. Muilwijk, A. Sippens-Groenewegen, A. Peper, and C. A. Grimbergen, "Automatic QRS onsetand offset detection for body surface QRS integral mapping of ventricular tachycardia," *IEEE Trans. Biomed. Eng.*, vol. 41, pp. 830–836, Sept. 1994.
- [16] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, pp.21–28, Jan. 1995.
- [17] J. S. Sahambi, S. Tandon, and R. K. P. Bhatt, "Using wavelet transform for ECG characterization," *IEEE Eng. Med. Biol.*, vol. 16, no. 1, pp. 77–83, 1997.
- [18] Z. Dokur, T. Olmez, E. Yazgan, and O. Ersoy, "Detection of ECG wave forms by neural networks," *Med. Eng. Phys.*, vol. 19, no. 8, pp. 738–741, 1997.
- [19] Mehmet Engin, "ECG beat classification using neuro-fuzzy network", *Pattern Recognition Letters*, 25 (2004) 1715–172.
- [20] Themis. P. Exarchos et al, "Association rule-mining based methodology for automated detection of Ischemic ECG beats", *IEEE Trans. Of Biomedical Engineering* Vol 53, No 8, August 2006.
- [21] David Cuesta-Frau, Marcelo O. Biagetti, Ricardo A. Quinteiro, Pau Mico' -Tormos, Mateo Aboy "Unsupervised classification of ventricular extrasystoles using bounded clustering algorithms and morphology matching", *Med Bio Eng Comput*. vol.45,pp. 229–239,2007.
- [22] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis", *IEEE Trans. Inform. Theory*, 961–1005, 1990.
- [23] <http://physionet.org>.
- [24] Amara Grap, "An Introduction to wavelets", *IEEE Trans. Comp. Sc. And Eng*, Vol. 2, No. 2, 1995.

Digital filtering of Amino acid sequence for prediction of cancer cell

Suman Saha¹Soma Barman(Mandal)²

Institute of Radiophysics and Electronics, University of Calcutta, 92, A.P.C. Road, Kolkata-700009

s.saha@zoho.com¹, barman_s@email.com²

Abstract—Power Spectrum Density (PSD) of 20 Amino acids of various Homo sapiens cell is estimated using Discrete Fourier Transform (DFT). A digital IIR filtering model is realized to predict cancer disease. The response filtered signal of all the amino acids is 1000% less in case of cancer cells compare to normal cells.

Keywords- *Amino Acid, Deoxyribo Nucleic Acid, Discrete Fourier Transform, Digital Filter, Power Spectral Density*

I. INTRODUCTION

Genomic signal processing (GSP) is a quickly evolving interdisciplinary field that blends biosciences, medicine and engineering. One of the important area of GSP is DNA detection and analysis which is generally used to reveal some hidden structure, to distinguish coding regions from non coding regions in DNA sequence. Disease diagnosis is another important area of GSP to find out abnormalities present in DNA, because almost all human genetic diseases such as cancer and developmental abnormalities are characterized by the presence of genetic variation. Over the past decade, significant discoveries have been made that provide a better understanding of genetic basis of cancer disease. It has been understood from various medical reports that 20 types of amino acids in DNA plays an important role in the study of cancer disease. We believe that the technique of GSP can be used to process and interpret genomic signal for better understanding of the underlying biological mechanisms of cancers and to assist in the development of optimal strategies for cancer therapy. Controlled Amino Acid Therapy (CAAT) is an efficient medical treatment used to impair the development of cancer cell. Proteins drive most the biological processes in living organism. The protein's biological function is encrypted within the protein's primary structure, i.e. the sequence of amino acids. Amino acids that must be obtained from the diet are called "Essential Amino Acids", other amino acids that the body can manufacture from the other sources are called "Nonessential Amino Acid". The basic requirement to impair the development of cancer cells is controlling these Amino Acids.

Controlled Amino Acid Therapy (CAAT) used in cancer treatment to reduce certain amino acids in the daily diet of cancer patient.

Dr. Otto discovered that all cancer cells produce inordinate amount of lactic acid. Dr. Lee et al showed the deficiency of glucose can commit suicide of cancer cells. R.Sing et al showed

Arginine activity in case of breast cancer. George C et al stated that Aspartic Acid (D), Glutamic Acid (E), Glycine (G), Serine (S), Alanine (A) and Cysteine (C) generated through the synthesis of Glucose. Considering all above the literature surveyed it has been understood that the Amino Acids are highly related to cancer disease. Karlin et al presented analysis of multiple amino acids runs, charge and hydrophobic runs for five eukaryotic genome. They found that human proteins with multiple long runs are often associated with diseases and long glutamine runs induce neurological disorder and various cancers. After elaborate literature survey it has been inferred that Amino Acids are highly related to cancer disease.

The foregoing research works motivated the analysis of amino acids using digital signal processing. We estimated the spectral content of amino acids using DFT and the transformed signal is digitally filtered for better prediction of cancer genes. The technique is tested on various Homo sapiens databases downloaded from NCBI homepage.

This paper is structured as follows: Introduction, Overview of DNA, Methodology, Results and Conclusion.

II. OVERVIEW OF DNA

Genomic Signal Processing is primarily the processing of DNA sequence, RNA sequence and proteins. DNA (Deoxyribo Nucleic Acid) is double helix structure, containing the genetic information of living organisms. DNA is made up of a number of units called nucleotides consisting of pentose sugar, phosphoric acid and nitrogen bases such as Adenine (A), Guanine (G), Cytosine (C) and Thymine (T). The RNA (Ribo Nucleic Acid) molecule is closely related to the DNA. It is made up of four bases but a molecule called Uracil (U) is present instead of Thymine and U always pair with A. The structure of DNA sequence is shown in figure 1.

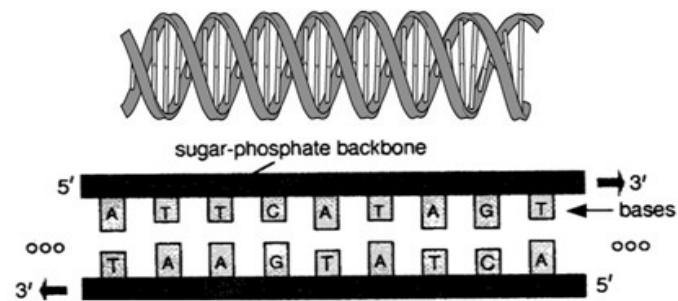


FIG 1

STRUCTURE OF DNA SEQUENCE

A DNA sequence can be separated into two regions, genes and intergenic spaces. Genes contain the information for generation of proteins. Genes has two types of sub-regions called exons and introns. The gene is first copied into a single stranded chain called the messenger RNA or mRNA molecule. The introns are then removed from the mRNA by a process called splicing. The spliced mRNA is divided into groups of three adjacent bases. Each triplet is called a codon. There are 64 possible codons. Each codon instructs the cell machinery to synthesize an amino acid. The codon sequence therefore uniquely identifies an amino acid sequence which defines a protein. This mapping is called the genetic code and is shown in Figure 2.

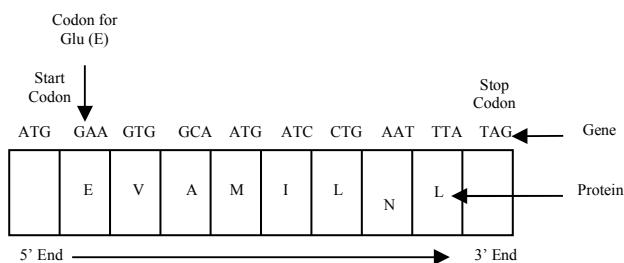


FIG. 2 CODON TO AMINO ACID MAPPING

The 64 possible codons generate 20 amino acids that are responsible for forming proteins. Each protein is formed by a chain of amino acids linked together through peptide bonds.

Amino acids are the ‘building blocks’ of the body out of which eight are essential, rests are non-essential. Eight essential amino acids human body cannot synthesize from other compound at the level needed for normal growth, so they must be obtained from food. Besides building cells and repairing tissue, amino acids form antibodies to combat invading bacteria and viruses; they are part of the enzyme and hormonal system; they carry oxygen throughout the body and are part of all muscular activity.

According to various medical reports, some amino acids are essential for the growth of tumor cells and restricting them or inhibiting them may be beneficial for curing cancer patients.

A long chain of amino acid in human proteins often are associated with various cancers leukemias and neurological disorder.

The authors in this article proposed a digital filtering model for prediction of cancer cells. The model is tested on various Homo sapiens genes consisting of BRCA1, BRCA2, Oncogenes and normal gens.

III. METHODOLOGY

The spectral content of 20 amino acids in a DNA sequence is studied using Discrete Fourier Transform and a model has been realized to filtered the transform signal. A matlab simulink environment is used for this purpose.

A. EIIP Mapping

Digital processing techniques can easily be applied to amino acids sequence as they can be viewed as character sequence with a definite alphabet consisting of twenty letters **A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y**. For that, the sequence should be digitized first, by mapping the character to suitable numerical values. Here we choose EIIP mapping technique to map twenty amino acids of human genome because this mapping technique reduced the computational overhead by 75% and provides relatively better result. With each amino acid molecule in a protein it is possible to associate a unique nonnegative number called the EIIP (average electron-ion interaction potential). This number ranges from 0.0 to 0.1263. Table 1 shows the EIIP values of 20 Amino acids.

TABLE I. EIIP VALUE OF TWENTY AMINO ACIDS

| Sl. No | Abbreviation | Amino Acid | EIIP Value |
|--------|--------------|------------------|------------|
| 1 | A | Alanine | 0.0373 |
| 2 | C | Cysteine (has S) | 0.0829 |
| 3 | D | Aspartic Acid | 0.1263 |
| 4 | E | Glutamic Acid | 0.0058 |
| 5 | F | Phenylalanine | 0.0946 |
| 6 | G | Glycine | 0.005 |
| 7 | H | Histidine | 0.0242 |
| 8 | I | Isoleucine | 0 |
| 9 | K | Lysine | 0.0371 |
| 10 | L | Leucine | 0 |
| 11 | M | Methionine | 0.0823 |
| 12 | N | Asparagine | 0.0036 |
| 13 | P | Proline | 0.0198 |
| 14 | Q | Glutamine | 0.0761 |
| 15 | R | Arginine | 0.0959 |
| 16 | S | Serine | 0.0829 |
| 17 | T | Threonine | 0.0941 |
| 18 | V | Valine | 0.0057 |
| 19 | W | Tryptophan | 0.0548 |
| 20 | Y | Tyrosine | 0.0516 |

For example, an amino acid chain of a DNA sequence of length N is

$$x[n] = M \ P \ I \ G \ S \ K \ E \ R \ P \ T \ F \ F \ E \ I \ F \ K \ T \ R \ C \ N \ K \ A$$

After EIIP mapping

$$\begin{aligned} x(n) = & 0.0823 \ 0.0198 \ 0.0000 \ 0.0050 \ 0.0829 \ 0.0371 \ 0.0058 \\ & 0.0959 \ 0.0198 \ 0.0941 \ 0.0946 \ 0.0946 \ 0.0058 \ 0.0000 \ 0.0946 \\ & 0.0371 \ 0.0941 \ 0.0959 \ 0.0829 \ 0.0036 \ 0.0371 \ 0.0373 \end{aligned}$$

After this conversion DSP methods can be applied effectively to predict the abnormalities of the DNA sequence.

B. Power Spectrum Estimation of Amino Acid Sequence using DFT method

Fourier technique is used to analysis the given sequence of amino acid.

The DFT of the binary indicator sequence, obtained by the mapping techniques is given by

$$X_s[k] = \sum_n X_s[n] e^{-j2\pi nk/n}$$

The power spectral content of the sequence is given by the following expression

$$P_s[k] = \sum_k |X_s[k]|^2$$

Where

$X_s[k]$ = DFT of sequence

N = Length of DNA sequence

$k = 0, 1, 2, \dots, N-1$

$n = 0, 1, 2, \dots, N-1$

s = A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y

C. Digital Filtering of Amino acid sequence

The transform signal obtained from DFT method is filtered using a digital IIR filter with the following specifications:

| | | |
|---------------------------------|------------------------|---------------------|
| Response Type | Band Pass | |
| Design Method | IIR - Chebyshev Type 1 | |
| Filter Order | 10 | |
| Frequency Specifications | Units | Normalized (0 to 1) |
| | Wpass 1 | 0.5625 |
| | Wpass 2 | 0.601 |
| Magnitude Response | Units | dB |
| | Apass | 1 |

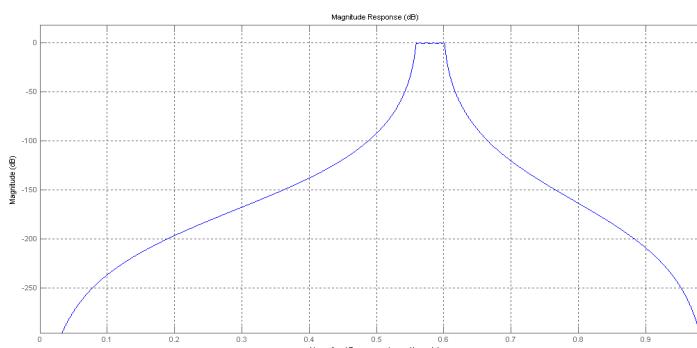


FIG. 2 FILTER RESPONSE

A Matlab Simulink environment has been realized to filter the signal. The proposed model is shown in figure 3.

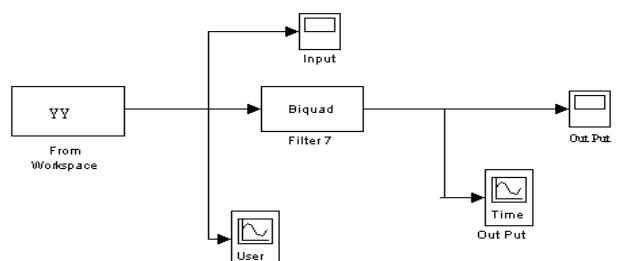


FIG.3 SIMULINK MODEL

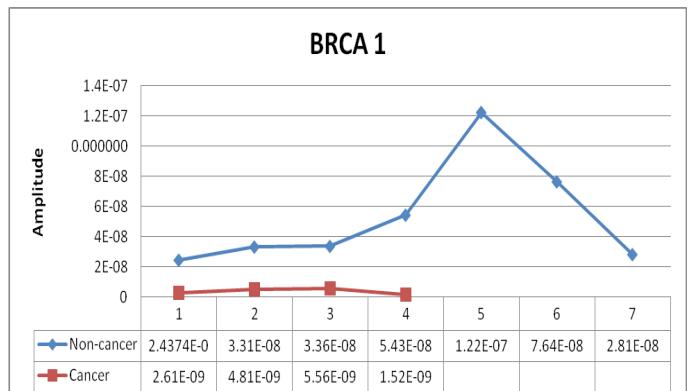
IV. RESULT

The proposed model is tested for Homo sapiens genes consists of BRCA1, BRCA2, Oncogenes and Normal cells downloaded from NCBI homepage shown in Table 2.

TABLE II DATABASES WITH ACCESSION NO.

| Type of Cancer Cell | Accession No |
|-------------------------|--|
| BRCA 1 (Ovarian Cancer) | NM_000465, NM_004656, NM_006768, NM_032043 |
| BRCA 2 (Breast Cancer) | AAG32681, AF309413, AF348515, AL137247, NM_000059, NM_016567, NM_024675, NM_078468, P38398, Q9P287, U43746 |
| Oncogene | AAG17262, CAA45108, CAA45109, CAA45110, CAA45111 |
| Non Cancer Cells | AF007546, AF186607, AF186613, AF540397, AY013301, L48213, NM_012400 |

The simulated results shows the Amplitude Modulated (AM) waveform both for cancer and noncancer cells. It has been observed that peak amplitude of the filtered signal for normal cell is always 10 times higher than cancer cells (BRCA1, BRCA2, Oncogenes). The simulated results are shown in Fig 4.



V. CONCLUSIONS

A digital filtering model has been proposed to predict abnormalities present in amino acid sequence. The performance of the model is successfully tested for various Homo sapiens genes. Digital filtered response of the transformed signal shows higher strength in normal cells compared to cells associated with cancer disease. Though EIIP mapping is used here for numerical conversion of amino acid sequence, other types of mapping may be explored for better prediction of cells using the proposed model.

REFERENCES

- [1] J. Chen, and S. T. C. Wong, "Nanotechnology for Genomic Signal Processing in Cancer Research," IEEE Signal processing magazine, January 2007, Vol 24, Issue 1, pp. 111-121.
- [2] www.apjohnancerinstitute.org
- [3] Rajan Singh, Shehla Pervin, Ardesir Karimi, Stephen Cederbaum, and Gautam Chaudhuri, "Arginase Activity in Human Breast Cancer Cell Lines: Nv-Hydroxy-L-arginine Selectively Inhibits Cell Proliferation and Induces Apoptosis in MDA-MB-468 Cells", CANCER RESEARCH 60, 3305-3312, June 15, 2000.
- [4] George C. Rock and Kendall W. King, "Amino Acid Synthesis from Glucose-U-¹⁴C in Argyrotaenia velutinana (Lepidoptera: Tortricidae) Larvae", The Journal of Nutrition, 95,369-373.
- [5] Samuel Karlein, Luciano Brocchieri, Aviv Bergman, Jan Mrazek and Andrew J. Gentles, "Amino acid runs in eukaryotic proteomes and disease associations", PNAS, January 8,2002, Vol 99, Vo 1, Pg 333-338.
- [6] P.P Vaidyanathan, " Genomics and Proteomics: a signal processor's tour." IEEE circuit and system magazine, March 2004, Vol 4, Issue 4.
- [7] R.F.Voss, "Evolution of long-range fractal correlations and 1/f noise in DNA base sequences.", Phy.Rev.Lett.,Vol.68 , No.25 , pp.3805-3808, 1992.
- [8] M.Rani and C.K Mitra, "Correlation analysis of the distribution of amino acids in protein sequence", Journal of Bio Science Vol 20, pp7-16, 1995.
- [9] National Centre for Biotechnology Information (NCBI). [Online]. Available: <http://www.ncbi.nlm.nih.gov/>

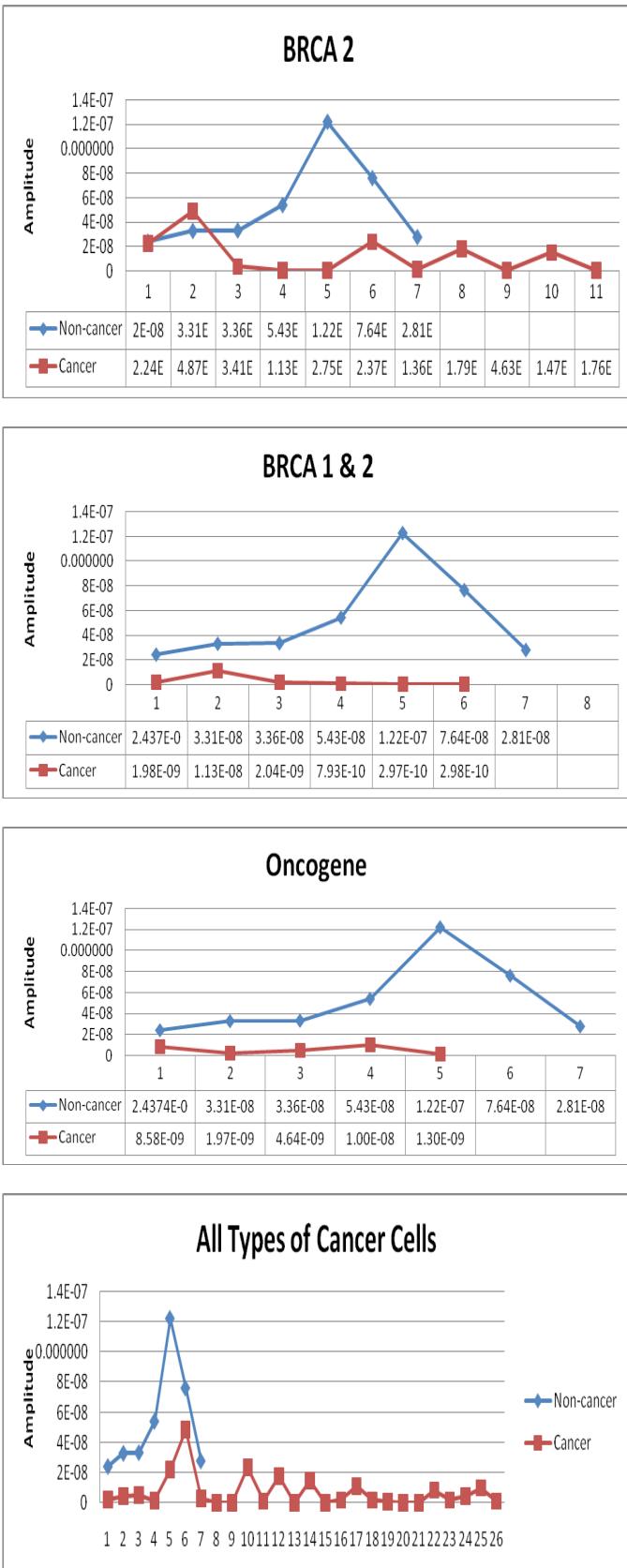


FIG. 4 SIMULATED RESULTS

Detection of Myocardial Infarction by Fourier Coefficient Analysis of ECG

Deboleena Sadhukhan, Madhuchhanda Mitra

Department of Applied Physics

University of Calcutta

Kolkata-700009, India

madhuchhanda94@rediffmail.com, deboleena.rainbow@gmail.com

Abstract— Myocardial infarction (MI) has been a major cause of death all over the world. Accuracy of the MI detection techniques based on the time plane ECG features are affected by variations in body composition of different subjects, presence of different noise and other diseases which alters the amplitude dependent features. The frequency domain and time-frequency domain analysis techniques involve complex feature sets. In this paper we aimed to detect MI by identifying the changes in the Fourier coefficients of the ECG caused due to presence of MI. Fourier spectra of the QT segment was considered because most detectable changes in the ECG caused by MI occurs in the QT segment. A computationally simple R-peak detection algorithm was developed based on sorting and thresholding of the squared double difference signal of the digital ECG data and RR interval processing, which showed a detection sensitivity of 99.8% as tested on 12-lead ECG data of PTB diagnostic ECG database. The QT segment of the ECG data was extracted by detecting slope values above particular thresholds and slope inversion points of selected windows on both sides of the detected R peaks. Discrete Fourier Transform was performed to obtain frequency spectra of the QT segment of ECG data of chest lead V4. The first twenty Fourier co-efficient for healthy and infarction data showed detectable difference. Analysis of the amplitude change of these co-efficients was done to determine a parameter which was significantly different for normal and infarction data and hence can be used as an easy identification tool for detecting MI.

Keywords- *Electrocardiogram; double difference; QT segment; Discrete Fourier Transform; Fourier coefficient;*

I. INTRODUCTION

The electrocardiogram (ECG) is the recording of the electrical activity of the myocardium of the heart during one cardiac cycle recorded by placing electrodes on the body surface [1]. The normal ECG is characterized by a recurrent sequence of P, QRS, T and a conditional U wave which represents the rhythmic depolarization and repolarization of the myocardium associated with the contractions of the atria and ventricles over each cardiac cycle as shown in Fig. 1. The clinical bandwidth of the ECG signal varies from 0.05-100 Hz and has low voltage ranges. Usually a 12-lead system is used for ECG recording of which standard leads I, II and III are bipolar leads, aVR, aVL, aVF are uni-polar limb leads and V1-V6 uni-polar pre-cordial chest leads to get an overall view of the heart's activity.

ECG is an important diagnostic tool for various cardiac diseases. The cardiac abnormalities cause change in the heart's

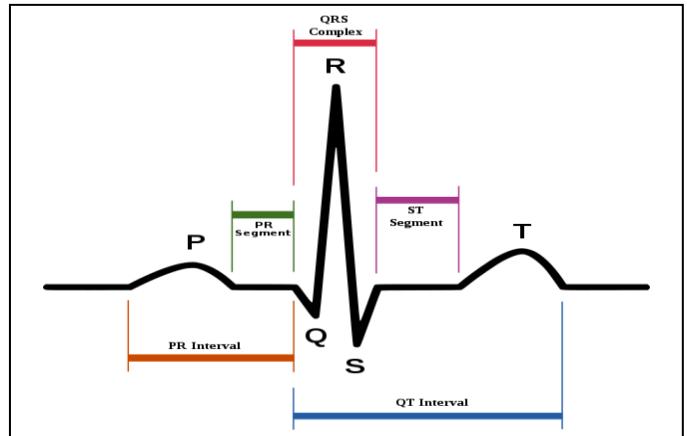


Figure 1. Schematic representation of ECG wave

electrical activity and hence the ECG. Cardiologists infer about the cardiac condition of a patient based ECG features like the PR interval, QRS interval, QT interval, ST interval, PR segment, and ST segment as shown in Fig. 1 [2-4]. The different computer aided analysis techniques aims at extracting these time plane features from digitized ECG data. Detection of QRS complexes and the R-peak is the most important step for automatic analysis of the ECG as it acts as the starting point of the other ECG feature extraction like the P, T waves ST segment [5]. It is important for heart rate variability (H.R.V) analysis [3] and for diagnosing other cardiac diseases [4]. Different QRS detection algorithms are available in literature [6-15]. The derivative based approaches [6-8] use first, second and other modified derivative operators. Other approaches are based on use of digital filters [9], mathematical models [10], and wavelet based techniques [11]. Statistical models like Hidden Markov Model [12] and pattern recognition techniques [13] are also used. Use of ANN or SVM as classifiers to detect the QRS complex [14-15] has also been reported. Due to ease of implementation and computational simplicity derivative based approaches are widely used.

Myocardial infarction (MI) is a cardiac disorder which has been the leading cause of mortality all over the world. Hence early detection of MI is very important. It is mainly caused by blockage of the coronary arteries supplying blood to the heart muscles leading to death of some heart muscle cells [16]. This causes changes in ECG signal like ST segment changes, T wave inversion, and reduced R wave amplitude [1]. Many

automatic MI detection algorithms are based on analysis of these features but had low detection sensitivity [2, 17-18]. To increase detection sensitivity other complex features were formulated based on the ST segment information and different tools were used as classifier [19-20]. But the accuracy of all these approaches was affected by body composition variations in different subjects, presence of noise superimposed on the ECG which alters the ECG signal amplitudes and presence of other diseases which simultaneously affects the selected ECG features. In these cases frequency plane analysis proved efficient. Spectrogram analysis [21] and time-frequency plane analysis techniques based on wavelet transforms showed better results [22]. Other frequency domain features were also used for detection of MI [23-24]. But most of them involved complex feature set which involved high computational complexity.

In this paper we aimed to detect MI by identifying the changes in the Fourier co-efficients of the ECG signal affected with MI. We analyzed the Fourier spectra of the QT segment only because most detectable changes in the ECG caused by MI occur in the QT segment [1]. A computationally simple R-peak detection algorithm was developed based on sorting and thresholding of the squared double difference signal of the ECG data and processing the RR intervals. The proposed algorithm was tested on 10s ECG data from the PTB diagnostic ECG database and a detection sensitivity of 99.8% was reported. QT segments of the ECG data were extracted by searching slope values above particular thresholds and slope inversion points on selected windows on both sides of the detected R-peaks. Discrete Fourier Transform was performed on QT segments of the ECG data of lead V4 for normal patients and patients affected with MI (anterior, antero-lateral, antero-septal). There was detectable difference in the first twenty Fourier co-efficient for healthy and infarction data. Analysis of the amplitude change of these co-efficients was done to determine a parameter which was significantly different for normal and infarction data and hence can be used as a tool to classify MI.

II. METHODOLOGY

The proposed methodology consists of three basic steps. The digital ECG data was first filtered to eliminate noise. Time plane feature extraction was done to extract the QT segment from the ECG data. First, the R-peaks were detected by a derivative based approach and then the other fiducial points (Q, T) are detected by slope comparison and searching for slope inversion points. The QT region of the ECG is extracted as it undergoes maximum alteration due to presence of MI. Discrete Fourier transform was then done on the extracted QT segments to determine the Fourier co-efficients. All 12 ECG leads are not necessarily used for diagnostic purpose due to information redundancy. For our algorithm we selected only lead V4 which mainly covers the anterior, antero-septal and antero-lateral parts of the heart. Analysis of the Fourier co-efficients for both healthy and infarcted data was then done to identify changes caused by MI.

The processing steps are described as follows.

A. Initial filtering of ECG data

The ECG signals gets corrupted due to various types of noises like power line interference, electrode contact noise, Electromyography noise, motion artifacts and others. In the present work the R- peak detection is based on the differencing technique. Presence of high frequency noise will result in unwanted difference peaks causing false peak detections. So the ECG data were first filtered to eliminate high frequency noises.

B. R-peak detection

The proposed algorithm for R-peak detection is based on the derivative based approach and operates on digitized ECG data from a single lead. The digital ECG data from a single lead is read as a 2-d array of the time instants and the sample points. The method involves 3 different stages for the accurate detection of the R peaks:

1) Detection of QRS regions:

Owing to the high frequency content of the QRS region [5-15 Hz] the derivative of these regions of ECG have higher amplitudes. As the sampling instants of digital ECG data remains constant the amplitude differences are proportional to the derivatives. Double differencing and squaring intensifies the magnitudes of the difference signal in the QRS regions which aids in the localization of the QRS regions.

- From the ECG data array the squared double differences are calculated at all points to yield the difference array.
- The difference array is sorted in descending order of magnitude and the difference peaks above a constant threshold value of 3% of the maximum are selected.
- All peaks obtained within an interval of 75ms on either side of the selected peaks are eliminated to eliminate possibility of detection of several peaks in the same QRS region.
- As the approximate duration of the QRS region is 150 ms the QRS regions are identified within a window of 75ms on either side of the difference peaks.

2) Detection of R-peaks:

The R peaks are the positive peaks of the QRS regions. These are detected by relative magnitude comparison in each QRS regions. A search for maximum amplitude was done on the relative magnitudes for each window to eliminate errors due to baseline wander.

- For each QRS window mean of the maximum and minimum values of the window are calculated and subtracted from all data points of the window to get the relative magnitudes.
- The position of the maximum of the relative magnitudes is taken as the R- peak location of the corresponding QRS window.

3) Processing of RR intervals:

R peaks obtained may not be accurate. False detections can occur due to noise which causes spurious R peak detections. Or there can be some missed peaks. To ensure accurate detection, the RR intervals are processed according to certain criteria.

- It is considered that the minimum difference between two successive R peaks can be 200ms. Any peaks detected within 200ms of the first is considered as noise peaks and eliminated.
- The average RR interval for three successive R peaks, one on either side of the R peak corresponding to the highest difference peak is calculated.
- If the RR interval between any two detected peaks is less than 70% of the average RR interval then the 2nd peak is eliminated and if it is more than 180% a search for another R peak in that interval is initiated with decreased threshold for the difference signal.

Fig. 2 illustrates the processing steps for detection of the R peaks.

C. Extraction of QT segment

Once the R-peaks are detected, the Q and T waves are detected by searching for high slope values and slope inversion points on both sides of the R peaks for extraction of the QT segment. For our analysis only one QT segment was extracted from each ECG data. The 2nd R-peak location of the ECG was identified.

1) Detection of Q point

The first slope inversion point on the left of the R peak where the amplitude difference changes sign is noted. A search window of ± 20 data points around the slope inversion point is chosen. Position of the absolute maximum of the window was identified to be the Q point.

2) Detection of T wave

Then search for slope values exceeding a threshold value and slope inversion points was performed on the right of the R peak leaving a window of 75ms to avoid detection of S point. The point where slope exceeds the threshold value corresponds to approximate T onset. The slope inversion point corresponds to the T wave peak. The point where the slope value decreases below the threshold value was detected to be the end of T wave.

Hence the QT segment was extracted from the ECG data. The extracted QT segment was visually inspected also. In cases where the QT detection algorithm failed, exact QT segment was detected by visual inspection.

Fig. 3 shows the processing steps for extraction of QT segment from the ECG data.

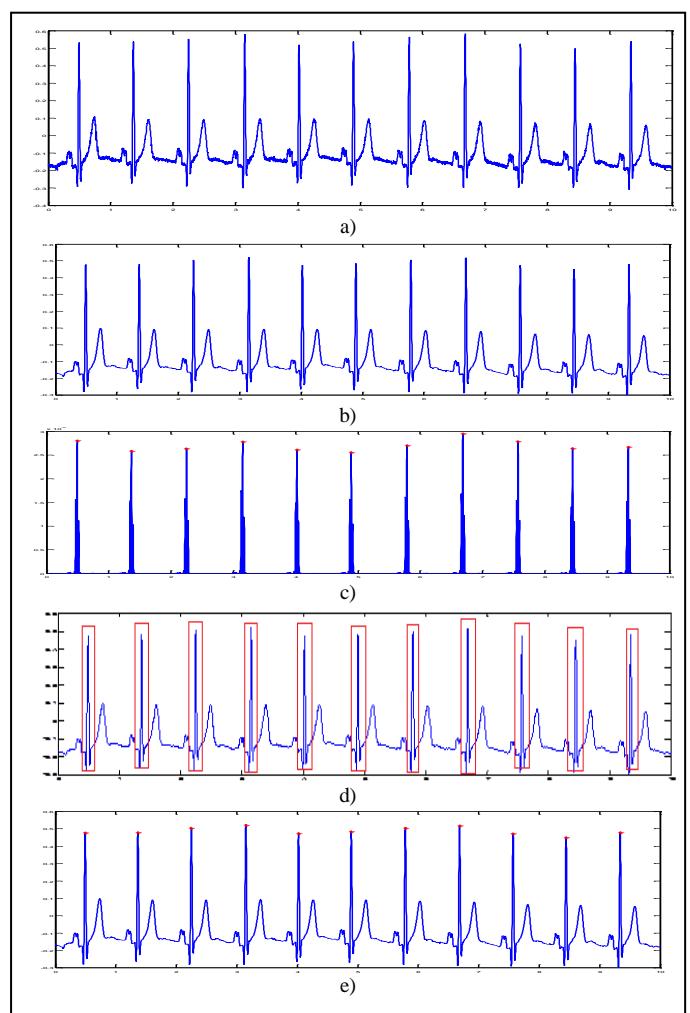


Figure 2. Processing steps for detection of R-peaks a)ECG signal b) denoised ECG signal c) selection of the squared difference peaks marked with “*” d) selection of QRS windows on the ECG signal e) Detection of R peaks

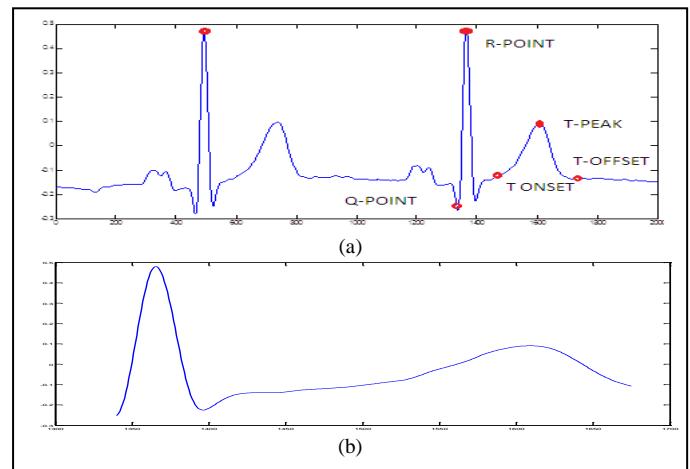


Figure 3. Extraction of QRS segment a) Detection OF Q and T waves b) Extracted QRS segment

D. Discrete Fourier Transform

Discrete Fourier Transform is a powerful computational tool for frequency domain analysis of discrete-time signals. The Fourier co-efficients of the QT segments was calculated using the Fourier transform equation.

$$C_k = (1/N) \sum_{n=0}^{N-1} x_p(n) e^{-j2\pi kn/N} \quad (1)$$

where $k=0, 1, \dots, N-1$, $x_p(n)$ is the QT segment data points with total N no of points. The first 20 co-efficients were extracted for frequency analysis of the normal and infarction ECG data.

III. RESULTS

The PTB diagnostic ECG database of digital ECG data of sampling rate of 1000 Hz available under the Physionet website [25] was used to test the performance of the proposed algorithm.

A. Detection of R-peaks

It was tested on 10s ECG data for all 12 leads of both healthy patients and patients affected with different diseases including MI, hypertrophy, bundle branch block and others.

The figure of merit for R peak detection algorithm is measured by the detection sensitivity S_e which is defined as

$$S_e = TP / (TP+FN) \quad (2)$$

Where TP = True positive (actual R peaks correctly detected as peaks); FN = False negative (actual peaks not detected as peaks).

Table I and II tabulates the test results for healthy and diseased patients and also calculates the detection sensitivity for each lead.

An overall detection sensitivity of 99.8% was detected which is a fairly good performance.

TABLE I. TEST RESULTS FOR SOME NORMAL PATIENTS

| LEAD NOS | TOTAL R | DETECTED R | TP | FN | S_e (%) |
|----------|---------|------------|-----|----|-----------|
| I | 109 | 109 | 109 | 0 | 100 |
| II | 109 | 109 | 109 | 0 | 100 |
| III | 109 | 109 | 108 | 1 | 99 |
| aVR | 109 | 109 | 109 | 0 | 100 |
| aVL | 109 | 109 | 109 | 0 | 100 |
| aVF | 109 | 109 | 109 | 0 | 100 |
| V1 | 109 | 109 | 109 | 0 | 100 |
| V2 | 109 | 109 | 109 | 0 | 100 |
| V3 | 109 | 109 | 109 | 0 | 100 |
| V4 | 109 | 109 | 109 | 0 | 100 |
| V5 | 109 | 109 | 109 | 0 | 100 |
| V6 | 109 | 109 | 109 | 0 | 100 |

TABLE II. TEST RESULTS FOR SOME DISEASED PATIENTS

| LEAD NOS | TOTAL R | DETECTED R | TP | FN | S_e (%) |
|----------|---------|------------|-----|----|-----------|
| I | 507 | 507 | 507 | 0 | 100 |
| II | 507 | 505 | 505 | 2 | 99.6 |
| III | 507 | 506 | 506 | 1 | 99.8 |
| aVR | 507 | 506 | 506 | 1 | 99.8 |
| aVL | 507 | 504 | 504 | 3 | 99.4 |
| aVF | 507 | 506 | 506 | 1 | 99.8 |
| V1 | 507 | 506 | 506 | 1 | 99.8 |
| V2 | 507 | 505 | 505 | 2 | 99.6 |
| V3 | 507 | 506 | 506 | 1 | 99.8 |
| V4 | 507 | 506 | 506 | 1 | 99.8 |
| V5 | 507 | 506 | 506 | 2 | 99.6 |
| V6 | 507 | 506 | 506 | 1 | 99.8 |

B. Fourier co-efficients

Fourier co-efficients were found for the QT segments of the ECG data only because most of the detectable changes caused by MI occur in this region [16-18]. The Fourier analysis was performed on ECG data of lead V4 which gets affected due to presence of Anterior, Antero-lateral, or Antero-septal MI. Hence frequency analysis was performed on healthy patients and on patients affected with the above three kinds of MI. The data was randomly selected from the PTB diagnostic ECG database.

The Fourier spectra of the first twenty co-efficients of each QT segments showed significant changes for normal ECG data and infarction data.

Fig. 4 and 5 shows the first 20 DFT co-efficients of QT segments for healthy and infarction ECG respectively.

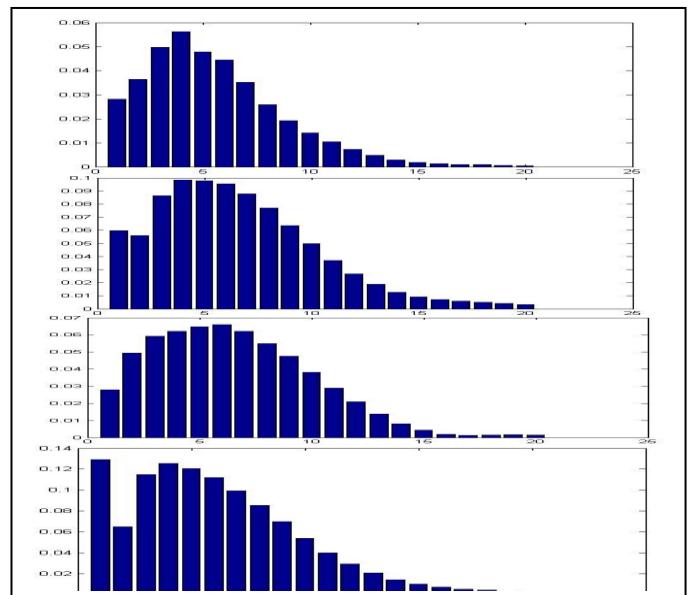
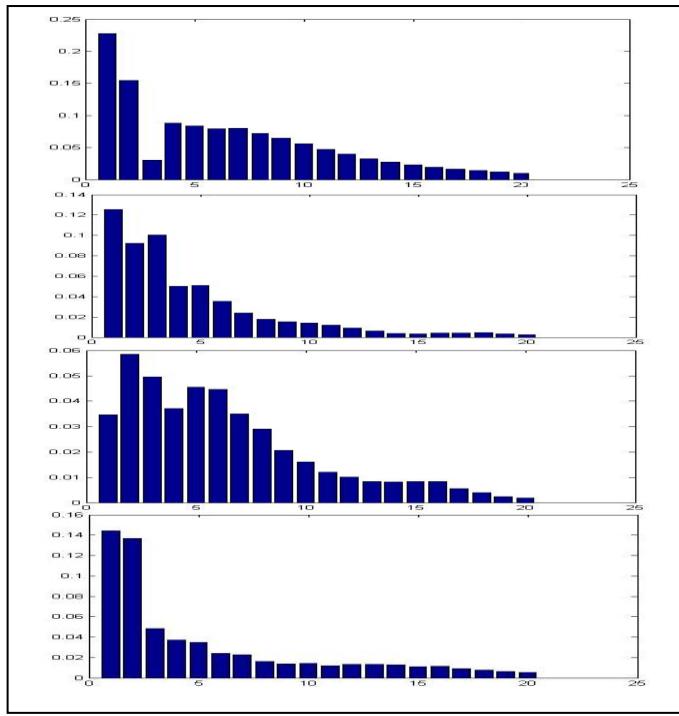


Figure 4. DFT spectra for healthy data



C. Analysis of Fourier co-efficients for classification

The Fourier spectra of the QT segment show significant changes for normal and infarction data. A close study of the Fourier spectra revealed that the fall in amplitude level of the 7th to 14th co-efficients was rapid in case of healthy data.

The change in amplitude between two successive co-efficients (c) was calculated to give the amplitude difference (ad) and sum of the 7th to 14th amplitude differences (s) was calculated for each data.

$$ad(i) = c(i+1) - c(i) \quad \text{where } i=1,2,\dots,19 \quad (3)$$

$$s = ad7 + ad8 + ad9 + ad10 + ad11 + ad12 + ad13 + ad14 \quad (4)$$

$$S = -(s * 100) \quad (5)$$

The parameter S was calculated for ECG data of lead V4 for twenty normal and twenty infarction patient records. Table III tabulates the values for 10 data of each kind

TABLE III. S VALUES FOR NORMAL AND INFARCTION DATA

| NORMAL RECORD NO. (LEAD V4) | S | INFARCTION RECORD NO. (LEAD V4) | S |
|-----------------------------|------|---------------------------------|------|
| s0287 | 5.75 | s0015 | 2.45 |
| s0291 | 7.47 | s0036 | 2.4 |
| s0306 | 8.95 | s0046 | 2.57 |
| s0464 | 5.41 | s0253 | 2.67 |
| s0466 | 5.73 | s0507 | 2.04 |
| s0471 | 7.03 | s0557 | 3.71 |
| s0473 | 6.41 | s0083 | 1.81 |
| s0312 | 7.64 | s0107 | 2.88 |

| NORMAL RECORD NO. (LEAD V4) | S | INFARCTION RECORD NO. (LEAD V4) | S |
|-----------------------------|------|---------------------------------|------|
| s0479 | 6.09 | s0117 | 2.23 |
| s0481 | 4.89 | s0140 | 1.67 |

Based on the observations a threshold value of 4 was selected for the parameter S and simple rule for classification of the normal and infarction ECG data for lead V4.

CLASSIFICATION RULE:

If the value of parameter S is less than 4 then it is infarction data.

The rule was applied to classify twenty normal and twenty infarction data and the results and classification sensitivity is calculated in table IV.

TABLE IV. CLASSIFICATION RESULTS

| TYPE | NO. OF RECORDS | TP | FN | S _e (%) |
|------------|----------------|----|----|--------------------|
| Normal | 20 | 19 | 1 | 92 |
| Infarction | 20 | 18 | 2 | |
| Total | 40 | 37 | 3 | |

IV. DISCUSSIONS

The R peak detection algorithms available in literature though reports high detection sensitivity but either are computationally complex [9, 10] or requires suitable mother wavelets and scale values [11]. The proposed R-peak detection algorithm developed was computationally simple and hence needed less execution time. But as it is a derivative based approach it is highly sensitive to HF noise.

The algorithm applied for detection of Q and T waves depending on slope values and slope inversion points is a crude method for detection. In case pathological ECG with altered features this algorithm proved insufficient and QT segments were extracted by visual inspection.

Detection of MI based on time plane feature of the ECG [17-20] are subject to alterations due to body composition variations of different subjects or presence of noise which alters the amplitude related features of the ECG. In frequency plane these variations are eliminated by selecting the required frequency domain of interest. The various frequency domain and time frequency domain analysis reported in literature uses complex feature set for classification purpose [20-24]. The proposed methodology aims at observing the significant changes in the Fourier spectra of normal and infarcted ECG. This can be aid in the analysis of ECG.

V. CONCLUSION AND FUTURE WORK

This paper reports that the Fourier co-efficients of the ECG can be an easy tool for detection of MI. After thorough observation of the first twenty co-efficient values of both normal and diseased ECG data we formulated a parameter to

2nd Annual International Conference IEMCON 2012

classify normal and infarcted ECG for only lead V4. Development of classification rules for all other leads is under process. The main advantage of this method was the reduced effect of superimposed noise on ECG and body composition variations which alters the time domain features.

Further studies are being carried on the analysis of the Fourier co-efficients of the ECG to use them as a reliable tool for detection of other cardiac diseases. Detection of a particular trend pattern in the Fourier co-efficients of normal ECG and diseased ECG can aid in generation of diseased ECG signal by using Inverse Discrete Fourier Transforms on the generated co-efficients which would be of severe importance for research purposes.

REFERENCES

- [1] L. Schamroth, "An introduction to electrocardiography," Wiley-Blackwell, 1991.
- [2] M.C. MacLachlan, B.F. Nielsen, M. Lysaker, A. Tveito, "Computing the size and location of Myocardial Ischemia using measurements of ST-segment shift," IEEE Trans. on Bio Med. Eng., vol. 53 (6), pp 1024-1031,2006.
- [3] H.-L. Chan, W.-S. Chou, S.-W. Chen, S.C. Fang, C.S. Liou, Y.S. Hwang, "Continuous and online analysis of heart rate variability," J . Med. Eng. Technol., vol. 29(5), pp 227-34, Sep-Oct, 2005.
- [4] P. Gomis, D.L. Jones, P. Caminal, E.G. Berbari, P. Lander, "Analysis of abnormal signals within the QRS complex of the high resolution Electrocardiogram," IEEE Trans. on Biomed. Eng., vol. 44 (8), pp 681-693, 1997.
- [5] F. Gritzali, G. Frangakis, G. Papakonstantinou, "Detection of the P and T waves in ECG," Comp. Biomed. Res., vol. 22, pp. 83-91,1989.
- [6] W.P. Holsinger, K.M. Kempner, and M.H. Miller, "A QRS pre processor based on digital differentiation," IEEE Trans. on Biomed. Eng., vol. BME-18, no. 3, pp. 212-217, 1971.
- [7] J. Pan , W.J. Tompkins,"A real-time QRS detection algorithm," IEEE Trans on Biomed. Eng., vol. 32(3), pp. 230-236,1985.
- [8] Y.C. Yeha, W.J. Wang, "QRS complexes detection for ECG signal: The difference operation method," Comp. Methods and programs in biomedicine, vol. 91, pp. 245-254, 2005.
- [9] M. Okada, "A digital filter for the QRS complex detection," IEEE Transactions on BME, Vol.-30, pp. 651-657,1983 .
- [10] T. Murthy and G. D. Prasad, "Analysis of ECG from pole-zero models," IEEE Trans. Biomed. Eng., vol. 39, pp. 741-751, July 1992.
- [11] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," IEEE Trans. Biomed. Eng., vol. 42, pp.21-28, Jan. 1995.
- [12] D.A. Coast, R.M. Stem, G.G. Cano and S.A. Briller, "An approach to cardiac arrhythmia analysis using hidden Markov models", IEEE Trans. Biomed. Eng., BME-37, 826 – 836, 1990.
- [13] P.E. Trahanias and E. Skordalakis, "Syntactic pattern recognition of the ECG," IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, pp. 648 –657, 1990.
- [14] Y.Suzuki, "Self organising QRS-wave recognition in ECG using neural net-works," IEEE Trans. on Neural Networks, vol. 6(6), pp. 1469–1477, 1995.
- [15] S. S. Mehta, N. S. Lingayat, "Detection of QRS complexes in electrocardiogram using support vector machine," J. of Med. Eng. & Tech. , Vol. 32, No. 3, pp. 206 – 215,2008.
- [16] Kristian Thygesen, Joseph S. Alpert, and Harvey D. White , "Universal definition of myocardial infarction," Journal of the American College of Cardiology, Vol. 50, No. 22, 2007.
- [17] N. Herring and D.J. Paterson, "ECG diagnosis of acute ischemia and infarction: past, present and future," *QJM*, vol. 99, no. 4, pp.219-230, 2006.
- [18] S. Yusuf, M. Pearson, H. Sterry, S. Parish, D. Ramsdale, P. Rossi, P. Sleight, "The entry ECG in the early diagnosis and prognostic stratification of patients with suspected acute myocardial infarction," *Eur Heart J.* , vol. 5 (9), pp. 690-696, 1984.
- [19] McSharry, "Advanced methods and tools For ECG data analysis," Artech House Publishers, Oct. 2006.
- [20] H. Haraldsson, L. Edenbrandt and M. Ohlsson "Detecting acute myocardial infarction in the 12-lead ECG using Hermite expansions and neural networks," *Artif. Intell. Med.*, vol. 32, pp. 127-136, 2004.
- [21] A.K.M Fazlul Haque, Md. H. Ali, M.A. Kiber, "Improved spectrogram analysis for ECG signal in emergency medical applications," *International Journal of Advanced Computer Science and Applications*, Vol. 1, No. 3, September 2010.
- [22] H. Dickhaus, L. Khadra and J. Brachmann, "Quantification of ECG late potentials by wavelet transformation", *Comput. Methods Programs Biomed.*, vol. 43, pp. 185 – 192,1994.
- [23] G. Bakul , U.S. Tiwary,"Automated risk identification of myocardial infarction using relative frequency band coefficient (RFBC) features from ECG", *The Open Biomedical Engineering Journal*, vol.4, pp. 217-222, 2010.
- [24] J.W. Abdulsada and H.J. Abbas, "Automatic diagnosis from electrocardiograms," *Int. J. Syst. Sci. (UK)*, vol. 19, pp. 2157 – 2162,1988.
- [25] <http://www.physionet.org/physiobank/database/mitdb/>

A Comparative Study on the accuracy of ECG Detection for Arrhythmia by different wavelet families

Anandarup Mukherjee, Sanjukta Chatterjee, Kajari Chattopadhyay, K. K. Ghosh

Department of Electronics and Communications Engineering,

Institute of Engineering and Management (IEM), Kolkata, 700 091, India

E-mail: anandarupmukherjee@gmail.com, sanjukta86chatterjee@gmail.com

Abstract - In this paper we present a comparative study on the effects of different continuous wavelet transform families and varying scale factor for classification of ECG signals using a Gaussian Classifier. The classifier distinguishes two different ECG patterns, one is the normal sinusoidal rhythm and the other is arrhythmic rhythm. Three different and commonly used wavelet families are taken for this purpose, namely, Daubechies Wavelet, Symlet Wavelet and Coiflet Wavelet. The scale factor was also varied for each wavelet family and the accuracy of the classifier was analyzed against these changes. We check which wavelet family and scale factor gives the best accuracy while classifying these two signals.

Keywords – *Arrhythmia, CWT, ECG classification, Daubechies, Symlet, Coiflet, Scale factor*

I. INTRODUCTION

Heart disease is one of the leading causes of death worldwide, immediate detection of abnormal functions of heart is, therefore, important. The widely used tool for detection of the health of the heart is ECG. It provides a graphical tracing of depolarization and repolarization of the heart. This tracing is in the time domain viz. it is a plot of cardiac potential in millivolts (mV) with respect to the time of occurrence of that potential.

The time domain representation of the ECG signal does not give any information about the variation of frequency components of the ECG signal. Instead of using Frequency domain analysis such as Fourier Transform which only gives information about the frequency components present in the signal, the wavelet transform method is used which provides localization of the signal components in both frequency and time domain [1].

In wavelet transform, a linear operation transforms the signal to decompose it into various scales. For this purpose, the signal is passed through a series of highpass and lowpass filters in order to analyze both high as well as low frequency components. There is

no defined procedure for selection of a wavelet for analytical operations on the signal. The wavelets closely resembling the signal are generally selected for the analysis [2, 3]. There are several wavelet families such as Haar, Daubechies, Coiflets, Symlets, Biorthogonal, Mexican Hat, Morlet and several other Real and Complex wavelets [4]. We have taken Daubechies, Symlet and Coiflet for our study because these wavelets closely resemble the QRS complexes and their energy spectrum is centered on low frequencies [5].

The chosen wavelets were used for the study of change in the accuracy of a Gaussian classifier. A Gaussian classifier; using continuous wavelet transform distinguishes between Arrhythmia and Normal Sinusoidal Rhythm by changing various parameters of the signal during feature extraction [6]. In the paper , the accuracy level of three different wavelet families together with variation of the scale factor from 1:8 to 1:128 were determined in order to show which sub-family at what scale factor attains highest accuracy and hence proves to be most useful in detection of physiological disorders.

a. Cardiac Arrhythmia and Its Detection by Electrocardiogram (ECG)

Cardiac arrhythmia is the condition of abnormal electrical activity in the heart that leads to irregularity in heart beat. Heart beat originates as electrical impulse from sino-atrial node or SA node causes the both atria to contract. The impulse then activates the atrio-ventricular or AV node and spreads through both ventricles via the Bundle of His and the Purkinje fibers generating a synchronized contraction of the ventricular muscles. Normal heart rate of adults is 60-80 beats per minute. Cardiac arrhythmia is characterized by slower (bradycardia) or faster (tachycardia) and regular and irregular heartbeats. Arrhythmia can be caused by problem in the cardiac impulse formation, impulse conduction or both [7]. Abnormality in the discharge rate from the normal pacemaker (e.g., sinus node) or discharge from ectopic pacemakers leads to the disorders of impulse

formation. On the other hand, disorders of impulse conduction can be caused either by delay in the conduction or re-entry of electrical impulse in cardiac muscle. Delay in conduction can lead to both bradycardia and tachycardia. Re-entry tachycardia occurs when a small group of inactivated cells discharge impulses during the refractory period of the rest of the cardiac cells, hence causing re-entry of localized electrical impulse. When multiple micro-reentry circuits are generated throughout one heart chamber causing it to beat abnormally it is called fibrillation [8].

Cardiac arrhythmias can be life-threatening in many cases and therefore their diagnosis is very important from clinical point of view. Electrocardiogram (ECG) is commonly used as diagnostic test for cardiac arrhythmias [9]. Electrocardiogram (ECG) involves measurement and interpretation of the electrical activity of the heart detected by electrodes attached to the outer surface of the skin in different parts of the body. The ECG device detects and amplifies the small electrical changes on the skin that are caused by cardiac polarization and depolarization during each heart beat [10]. Electrical voltages are picked up through different electrodes and system of pins (called "lead") from various parts of the body and are recorded in time scale in the form of typical graph on an ECG paper. A typical ECG tracing consists of different waves – P wave, QRS complex, T wave and U wave. A normal ECG plot is shown in *Figure 1*.

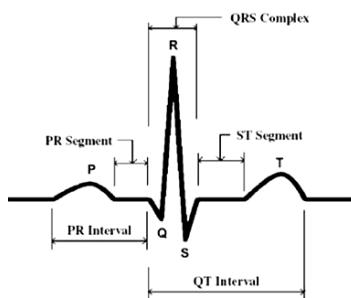


Figure 1: The Time-Domain Representation of an ECG Signal

b. The Need for Using Continuous Wavelet Transform as a Detection Technique

Nowadays automated systems are used in the analytical computations of medical data. Modern techniques outsource a large volume of data which helps in better understanding of physiological behavior. A physician's decision for the diagnosis of the patient depends on this set of data which not always ambiguous. Hence, a system is required which will help in minimization of errors for better

diagnosis of the diseases. In modern technologies the electrical activity of the heart is measured by using simple but useful digital signal processing algorithms in electrocardiographs.

These algorithms can be divided into 3 classes based on:

- Time
- Frequency
- Both time and frequency

The first two classes were successful in many diagnostic and analytical description of ECG; for example, detection of the QRS complex, deviation of the ST segment, detection of heart beat variability and many more. Results detected only on the basis of time gives poor sensitivity since it causes a small change in amplitude. Results based on frequency are more sensitive but it cannot detect in which part of the heart cycle the variation has occurred. Modern techniques uses algorithms based both on time and frequency for much better results, they give frequency analysis together with the range of time. Based on the type of application, a specific type of wavelet transform can be implemented [11, 12].

A continuous wavelet transform has been implemented in order to detect normal heart beat, abnormality in PVC (premature ventricular contraction) and APC (atrial premature contraction of heart beats). Here, the wavelet transform will be used to decompose the ECG signal into wavelet coefficients.

c. Wavelet Transform

An efficient method of providing localization of a signal in both time and frequency domain is given by Wavelet Transform. Wavelet theory is a very versatile concept and is applicable to several subjects such as Image Processing, Signal Processing, Feature Extraction, Bio-medical Acoustics, etc [13]. All wavelet transforms may be considered forms of time-frequency representation for continuous-time signals and so are related to harmonic analysis. Almost all practically useful discrete wavelet transforms use discrete-time filter banks. These filter banks are called the wavelet and scaling coefficients in wavelets nomenclature. Wavelets are defined by the wavelet function $\psi(t)$ (i.e. the mother wavelet) and scaling function $\phi(t)$ (also called father wavelet) in the time domain [14, 15].

Wavelet transforms are broadly divided into three classes: continuous, discrete and multi-resolution based. In continuous wavelet transform, a given

signal of finite energy is projected on a continuous family of frequency bands [16, 17].

The subspace of scale a or frequency band [$1/a$, $2/a$] is generated by the functions (sometimes called *child wavelets*)

$$\psi(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (1)$$

Where, a is positive and defines the scale and b is any real number and defines the shift. The projection of a function x onto the subspace of scale a then has the form

$$x_a(t) = \int_{\mathfrak{R}} WT_{\psi} \{x\}(a,b) \cdot \psi_{a,b}(t) db \quad (2)$$

With *wavelet coefficients*

$$WT_{\psi} \{x\}(a,b) = \langle x, \psi_{a,b} \rangle = \int_{\mathfrak{R}} x(t) \psi_{a,b}(t) dt \quad (3)$$

In the present work, we have made use of three different wavelet families e.g. Daubechies, Symlet and Coiflet.

i. Daubechies Wavelets

These wavelets have no explicit expression except for db1, which is the Haar wavelet. However, the square modulus of the transfer function h_k is explicit and fairly simple. Let,

$$P(y) = \sum_{k=0}^{N-1} C_k^{N-1+k} y_k \quad (4)$$

Where N is the order and C_k^{N-1+k} denotes the binomial coefficients. Then

$$|m_0(\omega)|^2 = (\cos^2(\frac{\omega}{2}))^N P(\sin^2(\frac{\omega}{2})) \quad (5)$$

$$\text{Where } m_0(\omega) = \frac{1}{\sqrt{2}} \sum_{k=0}^{2N-1} h_k e^{-ik\omega} \quad (6)$$

Where, $|m_0(\omega)|^2$ is the square modulus of h_k , h_k is the transfer function of the wavelet and k denotes the location of the transfer function with respect to the

wave and is known as the location index of the transfer function h_k . The supported length of ψ and φ is $2N - 1$. The number of vanishing moments of ψ is N . Most dbN are not symmetrical. For some, the asymmetry is very pronounced. The analysis is orthogonal [18].

ii. Symlet Wavelets

In Symlet, N is the order. The Symlets are nearly symmetrical, orthogonal and biorthogonal wavelets proposed by Daubechies as modifications to the db family. The properties of the two wavelet families are similar. So to construct Symlet wavelet from Daubechies wavelet $m_0(\omega)$ is reused. For filtering purposes the $|m_0(\omega)|^2$ is represented as a function W of $z = e^{i\omega}$. Then W can be factored in several different ways in the form of

$$W(z) = U(z) \overline{U(\frac{1}{z})} \quad (7)$$

because, the roots of W with modulus not equal to 1 go in pairs. If one of the roots is z_1 , then $1/z_1$ is also a root. By selecting U such that, the modulus of all its roots is strictly less than 1, Daubechies wavelets dbN is built. The U filter is a "minimum phase filter."

iii. Coiflet Wavelets:

The Coiflet ψ and φ are much more symmetrical than the dbNs. With respect to the support length, coifN has to be compared to db3N or sym3N. With respect to the number of vanishing moments of ψ , Coiflet has to be compared to db2N or sym2N .

d. Gaussian Classifier

The Gaussian classifier uses maximum likelihood estimation to calculate the means and variances of each class. The class conditional probability density function $p(x/c)$ is then calculated by using multivariate Gaussian distribution, which is given by,

$$p(x) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right] \quad (8)$$

Where, $p(x)$ is the likelihood of an observation, d is the dimension of the data, x is d - component column

vector, μ is d- component mean vector, Σ is d-by-d covariance matrix . Finally, posterior probabilities are calculated using Bayes' rule, which can be given as,

$$P\left(\frac{\omega_j}{x}\right) = \frac{p\left(\frac{x}{\omega_j}\right)P(\omega_j)}{\sum p\left(\frac{x}{\omega_j}\right)P(\omega_j)} \quad (9)$$

Where, $P\left(\frac{\omega_j}{x}\right)$ is the posterior probability, $p\left(\frac{x}{\omega_j}\right)$ is the likelihood and $P(\omega_j)$ is the prior probability. An observation is assigned to the class with the highest posterior probability [19, 20].

II. METHODOLOGY:

The ECG recordings, both Normal Sinus Rhythm and Arrhythmia, were taken from freely available MIT-BIH Normal Sinus Rhythm and MIT-BIH Arrhythmia database respectively [21].

The data was selected from both databases separately and adjusted for wavelet transform, without changing any values and the nature of the recordings. The data were subjected to continuous wavelet transform by changing the wavelet families and the scale factor to generate wavelet coefficients. Relevant flow chart is shown in *Figure.2*. These coefficients, both from the Normal Sinus Rhythm Database and The Arrhythmia Database were used to train a Gaussian Classifier and the accuracy of classification between Normal Sinus Rhythm and Arrhythmia was observed for different wavelet families at different values of scale factor.

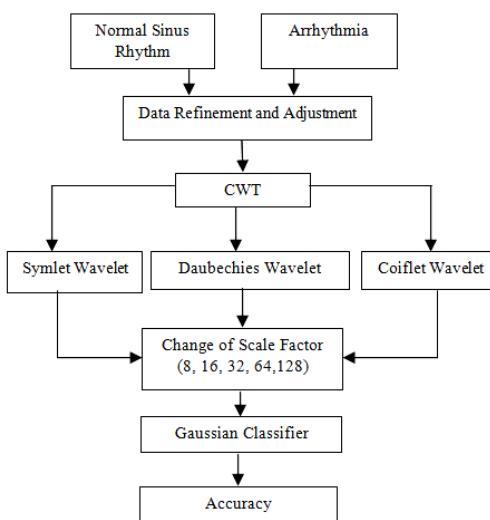


Figure 2: Flowchart for Classification of Normal Sinus Rhythm and Arrhythmias using a Gaussian Classifier

III. RESULT AND CONCLUSION

Table 1: Accuracy of symlet family with respect to different scales

| Scales: | 8 | 16 | 32 | 64 | 128 |
|-------------|----------|-----------|-----------|----------------|------------|
| sym2 | 26.2295 | 27.4454 | 34.8395 | 42.4044 | 26.2295 |
| sym3 | 26.2295 | 29.5492 | 42.8893 | 57.9918 | 26.2295 |
| sym4 | 26.2295 | 28.2104 | 39.6551 | 59.1189 | 26.2295 |
| sym5 | 26.2295 | 28.89 | 47.9816 | 79.5833 | 26.2295 |
| sym6 | 26.2295 | 29.2042 | 41.7555 | 76.3388 | 26.2295 |
| sym7 | 26.2295 | 30.2596 | 60.4747 | 95.8641 | 26.2295 |
| sym8 | 26.2295 | 31.4037 | 53.0908 | 94.0471 | 26.2295 |

Table 2: Accuracy of coiflet family with respect to different scales

| Scales: | 8 | 16 | 32 | 64 | 128 |
|--------------|----------|-----------|-----------|----------------|------------|
| coif1 | 26.2295 | 27.4283 | 32.0902 | 39.679 | 26.2295 |
| coif2 | 26.2295 | 28.1523 | 39.474 | 59.2862 | 26.2295 |
| coif3 | 26.2295 | 30.4645 | 41.9262 | 67.9781 | 26.2295 |
| coif4 | 26.2295 | 31.5642 | 45.2459 | 87.2541 | 26.2295 |
| coif5 | 26.2295 | 31.9604 | 60.7104 | 96.7145 | 26.2295 |

Table 3: Accuracy of db family with respect to different scales

| Scales | 8 | 16 | 32 | 64 | 128 |
|-------------|----------|-----------|-----------|----------------|------------|
| db2 | 26.2295 | 27.4454 | 34.8395 | 42.4044 | 26.2295 |
| db3 | 26.2295 | 29.5492 | 42.8893 | 57.9918 | 26.2295 |
| db4 | 26.2295 | 30.3689 | 44.7917 | 80.2152 | 26.2295 |
| db5 | 26.2295 | 32.2199 | 51.5847 | 90.6455 | 26.2295 |
| db6 | 26.2295 | 32.5205 | 58.0977 | 99.0847 | 26.2295 |
| db7 | 26.2295 | 34.4228 | 63.3402 | 99.9932 | 26.2295 |
| db8 | 26.2295 | 34.0266 | 72.2165 | 99.9863 | 26.2295 |
| db9 | 26.2295 | 35.7582 | 78.3504 | 99.9761 | 26.2295 |
| db10 | 26.2295 | 35.7104 | 90.625 | 99.9829 | 26.2295 |

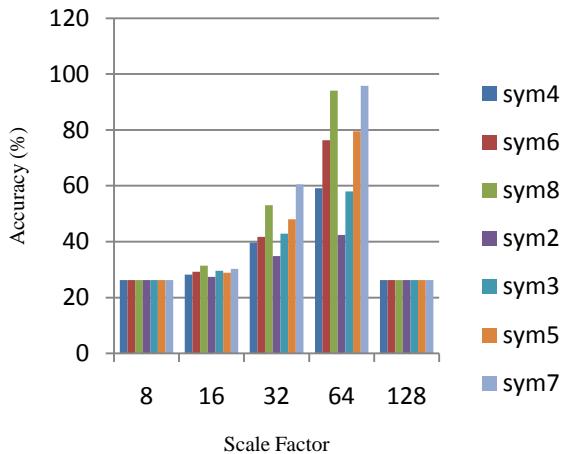


Figure 3: Accuracy vs. scale graph for symlet family

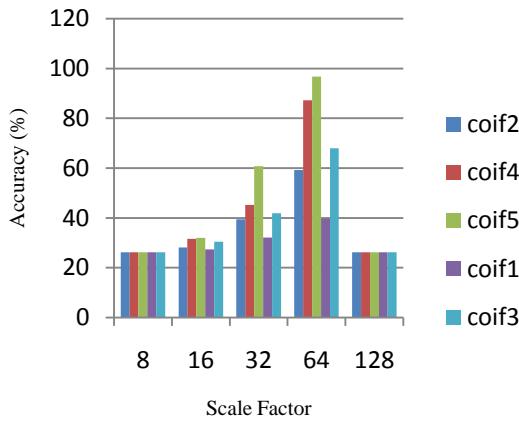


Figure 4: Accuracy vs. scale graph for coiflet family

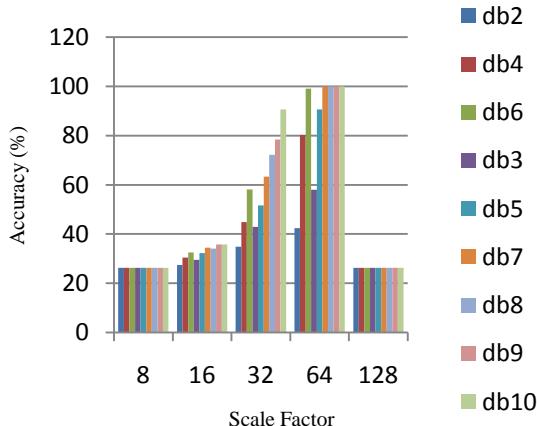


Figure 5: Accuracy vs. scale graph for db family

The effects of change of wavelet family and the corresponding change in scale factor observed for each family are given in *Tables 1, 2 and 3*. The corresponding graphical comparison is given in *Figures 3, 4 and 5*.

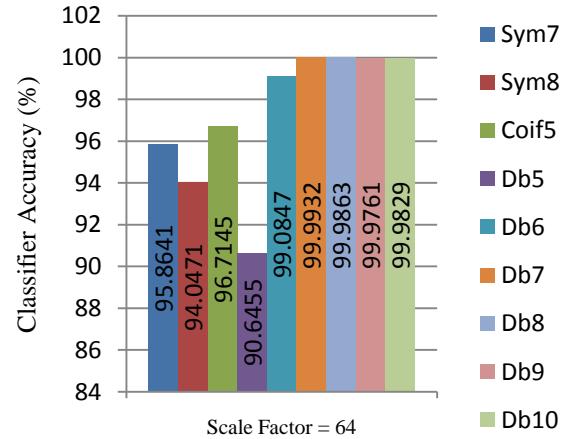


Figure 6: Comparison of wavelet families with accuracies > 90%

The performance of different wavelet families are depicted in *Figure 6*. If a Gaussian classifier be used to classify normal sinus rhythm and arrhythmia, it is observed that if wavelets “Symlet-7”, “Symlet-8”, “Coiflet-5”, “Daubechies-5”, “Daubechies-6”, “Daubechies-7”, “Daubechies-8”, “Daubechies-9” and “Daubechies-10” are used at a scale factor of 64, the Gaussian Classifier gives the optimum result. However, the best accuracies are given by “Daubechies-7”, “Daubechies-8”, “Daubechies-9” and “Daubechies-10”. The Optimum scale factor is 64, with the rest having less or no significant effect on accuracy of the classifier. Increase in scale factor does not necessarily mean increase in classifier accuracy, as is clear from *Tables 1, 2 and 3*.

Thus, for achieving best results while classifying between Normal Sinus Rhythm and Arrhythmia, the best choice of wavelet family is “Daubechies-7” at a scale factor of 64.

IV. REFERENCES

- [1] C. Saritha, V. Sukanya, Y. Narasimha Murthy, “ECG Signal Analysis using Wavelet Transform”, PACS number:87.85.J; 02.30,Nw Bulg.J.Phys.35(2008)68-77.
- [2] Vanishree K, J Singaraju, “Detection of R-R Interval using DWT”, International Journal on Computer Science and Engineering

- (IJCSE), ISSN:0975-3397, Vol 3, No. 4, April 2011
- [3] P. Ghorbanian, A. Ghaffari, A. Jalali, C. Nataraj, "Heart Arrhythmia Detection Using Continuous Wavelet Transform and Principal Component Analysis with Neural Network Classifier". Computing in Cardiology 2010; 37:669- 672, ISSN: 0276-6574.
- [4] Yuliyan Velchev, Ognian Boumbarov, "Wavelet Transform Based ECG Characteristic Points Detector", International Scientific Conference , Computer Science'2008, 49
- [5] S.Z. Mahmoodabadi, A. Ahmadian, M.D. Abolhasni, M. Eslami, J.H. Bidgoli, "ECG Feature Extraction based on Multiresolution Wavelet Transform", Proceedings of the 2005 IEEE, Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, September 1-4, 2005.
- [6] H. Khorrami, "A Comparative Study of DWT, CWT and DCT Transformations in ECG Arrhythmias Classification", Expert Systems with Applications, Volume-37, Issue-8, August 2010, Pages 5751-5757
- [7] Brian F. Hoffman (1999). "Cardiac Arrhythmias: What Do We Need to Know about Basic Mechanisms?" J Cardiovasc Electrophysiol, Vol. 10, pp. 414-416.
- [8] Douglas P. Zipes (2003). "Mechanisms of Clinical Arrhythmias". J Cardiovasc Electrophysiol, Vol. 14, pp. 902-912.
- [9] Van Mieghem, C; Sabbe, M; Knockaert, D (2004). "The clinical value of the EKG in noncardiac conditions." Chest 125 (4): 1561-76.
- [10] Cardiovascular Physiology
Electrocardiogram, Richard E.Klabunde
2007.
- [11] Dina Kicmerova, Ivo Provaznik (2008). "Wavelet contour based detector for arrhythmia analysis". European journal for biomedical informatics –vol :4 .
- [12] F. Zhang, "Novel QRS detection by CWT for ECG sensors", IEEE Xplore: Biomedical circuits and systems, 2007.
- [13] Robert X. Gao, Ruqiang Yan, "Wavelet: Theory and applications for manufacturing", ISBN-10: 1441915443, 2010
- [14] Robi Polikar, "The Wavelet Tutorial", 2nd Edition, Part-1.
- [15] Paul S. Addison, "Wavelet Transforms and the ECG: a review", Physical Meas. 26 (2005) R155- R199, Institute of Physics Publishing.
- [16] D. Lee Fugal, "Wavelets beyond comparison", World Satellite Magazine, May 2009, 35
- [17] Chun-Lin Liu, "A Tutorial of the Wavelet Transform", February 23, 2010
- [18] Kristian Sandberg, "The Daubechies Wavelet Transform", University of Colorado.
- [19] Nuno Vasconcelos , "The Gaussian Classifier", www.svcl.ucsd.edu/courses, ECE Department, UCSD.
- [20] Hongwei Guo , "A simple algorithm for fitting a Gaussian Function" ,IEEE Signal Processing Magazine 28(9):134-137 (2011)
- [21] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220; 2000 (June 13).

Design of Tele-healthcare Interactivity: A Semi Formal Approach

Imon Banerjee¹, Shrutilipi Bhattacharjee²

¹ Department of Information Technology, Institute Of Engineering & Management, WB
India

² Department of Information Technology, National Institute Of Technology, Durgapur
¹ imonban@gmail.com, ² shrutilipi.2007@gmail.com

Abstract—In this paper, we are supposed to manage the workflow of a web-based multiple participants in distance diagnosis and medical help referred as Telehealthcare Organization. The presented approach conspicuously helps the developer to comprehend the interactive relationship of different entities of Telehealthcare system but also to easily construct a shared and trusted medical environment. The interaction designing is done by Dynamic Petri Net model. This presented design can help to comprehend the interaction relationship among Administration-Client-Doctor, as well as with each query-diagnosis process of the organization.

Keywords— Distance diagnosis, Telehealthcare, Interaction, Dynamic Petri Net, Formal Design.

INTRODUCTION

Distance Diagnosis Organization is an evolving paradigm of diagnosis and healthcare that attempts to overcome both distance and time constraints found in traditional medical system. Telehealth system is a field of medical science and technology that aims to solve query as well as to provide medical help by doctors, hospitals to those clients/patients who are not physically present on the site. The main goal of telehealth system is to impart distance independent medical help with the use of information and communication technology. Rather than taking appointment physically, hospitals, doctors and clients may communicate at a predefined time schedule by exchanging electronic media, or through technology that allows them to communicate in real time and through other online ways. This enhance the diagnosis and any medical process to be location independent removing the distance constraint by networked PCs.

Being an interactive system, there is a number of interactions between each entity and the organizational processes itself. So rigorous formalization of each Distance Diagnosis Organization interactions for the effective implementation of the system is needed here. Petri nets, introduced by C. A Petri in 1962 [1], provide an elegant and useful mathematical formalism for modeling concurrent systems and their behaviors as well as interactions. In many applications, however, modeling by itself is of limited practical use if one cannot analyze the modeled system.. In this paper we use a new extension to the Petri Net model to design those interactions to get the best support in implementation level. The Dynamic Petri Net (DPN) is a powerful extension to the Petri Net family. DPN adds several new elements such as dynamic places, control places, control

functions, control output arcs, and control variables. The extensions provide for the creation and destruction of conventional Petri-Net Places and Transitions inducing dynamic behavior of Petri-Net operation.

A. Formal Definition of DPN

A Dynamic Petri Net structure, S, is a 10-tuple.
 $S = (P, T, I, O, \tau, P_d\{F\}, N, F, P\{F\}, O_c\{F\})$.

- 1) $P = \{P_1, P_2, P_3, \dots, P_X\}$ where $X \geq 0$, is a finite set of places.
 - 2) $T = \{t_1, t_2, t_3, \dots, t_Y\}$ where $Y \geq 0$, is a finite set of Transitions., where $P \cap T = \emptyset$ i.e. the set of the places and transitions are disjoint.
 - 3) $I: T \rightarrow P^\infty$ is the Input Arc, a mapping from places to bags of transitions.
 - 4) $O: T \rightarrow P^\infty$ is the Output Arc, a mapping from transitions to bags of places.
 - 5) $\tau_{1,1}, \tau_{1,2}, \tau_{1,3}, \dots, \tau_{a,1}, \tau_a \} = \tau_a$, where $a \geq 0$, is a finite set of time intervals representing playback time intervals. This is derived from OCPN.
 - 6) $N = \{n_1, n_2, n_3, \dots, n_b\}$, where $b \geq 0$, is a finite set of persistent control variables. These variables are persistent through every marking of the net.
 - 7) $F = \{f_1, f_2, f_3, \dots, f_c\}$, where $c \geq 0$, is a finite set of control functions that perform functions based on any control variable N.
 - 8) $P\{F\}: P$ is a finite set of static control places (a subset of P) that executes any control functions F.
 - 9) $O^c\{F\}: O^c$ is a finite set of static control output arcs that may be disabled or enabled according to the control functions F.
 - 10) $P^d\{F\}: P^d$ is finite set of dynamic places (a subset of P) that takes their value from some control function F.
- Token= { token₁, token₂, token₃, ..., token_z }, $z \geq 0$, is a finite set of dynamic marking on places.

RELATED WORK

This section presents an overview of approaches in the research literature associated with the research issues of this work. The subject of Telehealthcare has engaged researches all over the world. Many solutions have been proposed for distance healthcare and collaboration over the Internet.

Various methods have been proposed for synchronous healthcare with an on-line facilitator [1][2].

We here use Dynamic petri net model to design the system interactions. Similarly, several properties can be verified using Dynamic Petri Nets. Dynamic Petri Net were developed to model to interactive distributed multimedia environment. Naveed et al. [3] introduce OCPN for solving intra stream and inter stream synchronization problem in real time multimedia communication . Charls et al.[4] and Roy et al.[5] have introduce the concept of Dynamic Petri Net. In [5] the authors described a modeling technique for distributed sensor network. DPN has been used in designing of a multimedia orchestration tool[5].

MODELS OF TELEHEALTHCARE INTERACTIONS

We used The Dynamic Petri Net model to establish the activities in Telehealthcare Organization. It includes the following five stages:

A. Admin Interaction:

This stage prevent the loss of effective traditional medical techniques, thus providing the means for all instructors to continually improve existing curriculum in Distance Diagnosis Organization . Thus before selecting Doctors and hospitals in order to deliver best healthcare and diagnosis, the administrators(Admin) of the organization collects the feedback form outside world as well as form the hospitals and doctors.

First of all, the Administrator login with granted authorization for selecting specific doctors and efficient hospitals is illustrated in Fig 1 . The subsystem should provide the independency for the Administrator to choose doctors and hospitals based on the Market review. Administrator could also change previously selected doctors according to their performance in diagnosis. They first take into account the Market review and existing doctors interaction for selecting a new doctors and hospitals. Here the doctors are chosen according to their specialization but when selecting a hospital, its infrastructure perspective and location constraints are also taken into account. In both the cases Admin have to pay them for their services.

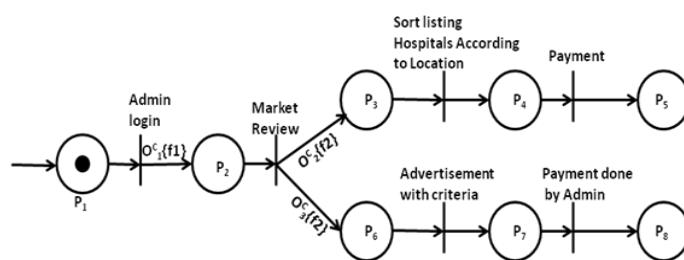


Fig. 1 Dynamic Petri Net model for Admin Interaction

P_1 : Start.

P_2 : Admin Home Page.

- P_3 : Hospital Selection.
- P_4 : Collaboration and Agreement.
- P_5 : Hospital registration.
- P_6 : Doctor Selection.
- P_7 : Agreement with sorted Doctor.
- P_8 : Doctor Registration.

```

f1: admin_login( )
{
    if(admin_id & password = true)
    {
        O^c_1 enabled;
    }
}
  
```

```

f2: market_review( )
{
    if(market_review .need= Doctor)
    {
        O^c_2 disabled;
        O^c_3 enabled;
    }
    if(market_review .need= Hospital)
    {
        O^c_2 enabled;
        O^c_3 disabled;
    }
}
  
```

B. Client Registration:

After selecting doctors and hospitals, the next step is to register the clients, who want to submit their query to their site. Here registration is authenticated by administration. So the Administrator login with granted authorization for client registration process is mandatory. At first when a client want to submit a query, the admin first checks its clients database whether this client is already registered or a new client. If he/she is already registered in the database then his/her membership validity is checked. If a valid member then only query can be submitted else he/she has to renew the membership by yearly payment before submitting the query. For the new clients the first step done by admin is to authenticate the client. Once authentication is successful and payment is done, the client becomes a member and able to post any query. This total process is done in a location independent manner.

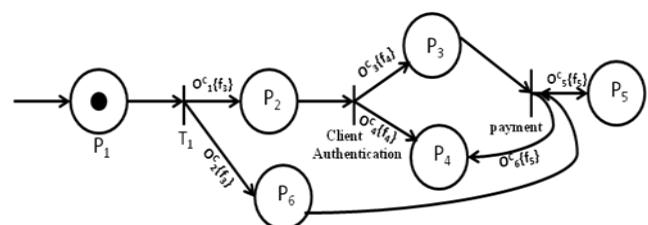


Fig. 2 Dynamic Petri Net model for Client Registration

P_1 : Start.

P₂: Registration Page.
 P₃: Agreement.
 P₄: Client Discarded.
 P₅: Login information provided to clients.
 P₆: Registration renewal page.
 T₁: Navigation to home page.

```
f3:check_client()
{
    if(admin_id & password = true && new client)
    {
        Oc1 enabled;
        Oc2 disabled;
    }
    if(admin_id & password = true &&  $\neg i$  new client)
    {
        Oc1 disabled;
        Oc2 enabled;
    }
}
```

```
f4: client_authentication ()
{
    if(client authentication =successful)
    {
        Oc3 enabled;
        Oc4 disabled;
    }
    if(client authentication =  $\neg i$  successful)
    {
        Oc4 enabled;
        Oc3 disabled;
    }
}
```

```
f5: payment ()
{
    if(payment=successful)
    {
        Oc5 enabled;
        Oc6 disabled;
    }
    else
    {
        Oc6 enabled;
        Oc5 disabled;
    }
}
```

C. Appointment:

Once the client is authenticated and registered he/she is able to submit their medical queries to the site which is a location and time independent process. In order to submit a query client has to login with the user-id and password

provided by the admin followed by authentication process done by admin. Now once the client submitted the query in, admin would process the query. Depending on the specification of registered doctors and particulars of registered hospitals, client would take the decision where to send the query to provide the diagnosis as best as possible. If the clients query requires any doctors help and treatment or any suggestion, then immediately it is forwarded to the particular doctor of that specialization and appointment details sent to the client. If the query need any direct interaction with any local hospital for some sort of test or any formal check up, the query is forwarded to the nearest and best suited hospital and client receives the appointment details sent by the admin. If admin is not able to provide any sort of help to the clients query, it is discarded immediately without any dilemma.

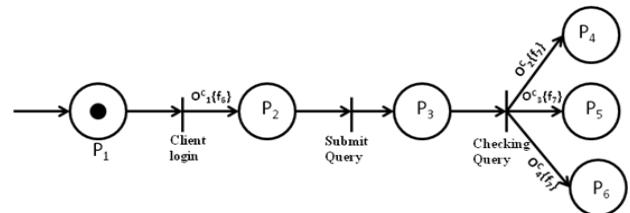


Fig. 3 Dynamic Petri Net model for Appointment

P₁: Start.
 P₂: Home Page.
 P₃: Process Query.
 P₄: Respective Doctors Appointment according to doctors Specialization.
 P₅: Appointment of Hospital.
 P₆: Discarded Query.

```
f6:client_login()
{
    if(client_id & password = true)
    {
        Oc1 enabled;
    }
}

f7: checking_query ()
{
    if(Doctor's Consultancy needed)
    {
        Oc2 enabled;
        Oc3 disabled;
        Oc4 disabled;
    }
    if(Hospital's Consultancy needed)
    {
        Oc2 disabled;
        Oc3 enabled;
        Oc4 disabled;
    }
    if(Invalid query)
    {

```

```

Oc2 disabled;
Oc3 disabled;
Oc4 enabled;
}
}
}

```

D. Doctor Client Interaction:

In order to get the solution submitted previously doctor client interaction is needed. This may operate in two modes, synchronous and asynchronous. In the synchronous mode, on the particular time of the appointment (provided by admin earlier) both doctor and client have to be present to interact online. Client submits its query to the doctor and interacts with him through a number of dialogue interchange. This process lasts until the client gets solution as best as possible or the appointment time up. If appointment time is not true or doctor is not present, client may also submit query to the doctor in offline mode. Doctor also can provide solution to the client in this manner. This is asynchronous type of interaction between doctors and clients.

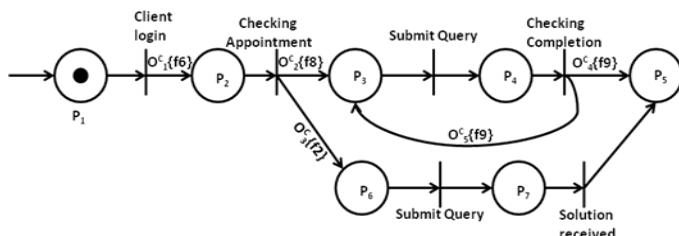


Fig. 4 Dynamic Petri Net model for Doctor Client Interaction

P₁: Start.

P₂: Home page.

P₃ : Initialization of Interaction with Doctor in synchronous mode.

P₄: Processing Query.

P₅: End Session.

P₆: Interaction with Doctor in asynchronous mode.

P₇: Analyzing of Query.

f6: client_login()

f8: checking appointment()

```

{
    if(appointment.time = true && Doctor=present)
    {
        Oc2 enabled;
        Oc3 disabled;
    }
    if(appointment.time =  $\neg i$  true || Doctor=  $\neg i$  present )
    {
        Oc3 enabled;
        Oc2 disabled;
    }
}

```

```

}
f9: checking completion( )
{
    if(interaction=completed )
    {
        Oc4 enabled;
        Oc5 disabled;
    }
    else
    {
        Oc5 enabled;
        Oc4 enabled;
    }
}

```

E. Hospital Client Interaction:

If the query is not solvable by doctors interaction and need some direct assessment then hospital client interaction is necessary. Here in this interaction the client first has to login with his particular user id and password. Then after login hospital will check appointment of that client done in admin client interaction. If appointment on that time is true then process further else discard the query. Now client would submit the query directly to the hospital person. If the query is not perfectly understood by hospital person, then client has to submit the query again. This loop will continue until and unless hospital is satisfied to make out what the problem is. After completely understanding that respective hospital will give any solution to the client like check up appointment or call for any testing.

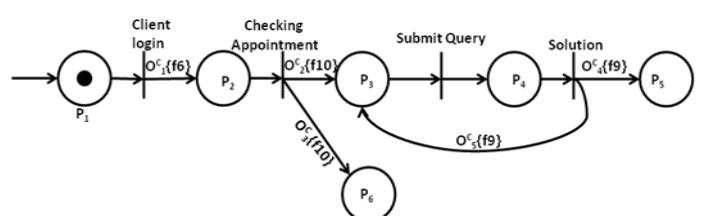


Fig. 5 Dynamic Petri Net model for Hospital Client Interaction

P₁: Start.

P₂: Home page.

P₃: Virtual Interaction Session.

P₄: Query Processing.

P₅: End Session and solution provided.

P₆: Discard Request.

f6: client_login()

f10: checking Hospital.appointment ()

```

{
    if(Hospital.appointment=true)
    {
        Oc2 enabled;
        Oc3 disabled;
    }
}

```

```

else
{
    O3c enabled;
    O2c disabled;
}
}

```

f9: checking completion()

CONCLUSIONS

In this paper, we proposed a trust development framework and approach to the understanding of the Distance Diagnosis by Dynamic Petri Net. This is the important comprehensive and coherent framework to guide our understanding of distance diagnosis in medical environment and society. This paper can be used as a basic research framework and tool to design and understand the characteristics of distance diagnosis and to explore its optimal location independent medical application.

ACKNOWLEDGMENT

This work was supported by Dept of Information Technology of National Institute of Technology (Durgapur),WB.

REFERENCES

- [1] Ken Boddy,Dimitrios Sotiriou Quality in Telemedicine Services: The HERMES Approach. 4th International Conference on Medical Aspects of Telemedicine Jerusalem, June 6-11, 1999
- [2] Telemedicine for the benefit of patients, healthcare systems and society, COMMISSION STAFF WORKING PAPER SEC(2009)943 final June 2009 .
- [3] Naveed U. Qazi, Miae Woo, Arif Ghafoor, "A and Communication Model for Distributed Multimedia Objects".
- [4] Charles J. Graff , Charles Giardina : "Dynamic petri-nets: A new modeling technique for sensor networks and distributed concurrent systems".
- [5] Roy Tan and Sheng-Uei Guan "A Dynamic Petri Net Model for Iterative and Interactive Distributed Multimedia Presentation", IEEE Transactions on multimedia, vol. 7, no. 5, october 2005.

ELECTROCARDIOGRAPHIC DATA COMPRESSION USING WAVELET

Amit Shankar Sinha¹, Aritra Basu¹

¹Department of Electronics and Communication Engineering
Institute of Engineering & Management
Salt Lake Electronics Complex, Kolkata-700091
West Bengal, India

amitshankar_sinha@yahoo.co.in , aritrabasu@live.com

Arighna Deb², Tejendra Nath Sen²

²Department of Electronics and communication Engineering
Institute of Engineering & Management
Salt Lake Electronics Complex, Kolkata-700091
West Bengal, India

arighna87@rediffmail.com , tejendra.sen@gmail.com

Abstract - Digital storage media is not expensive and computational power has exponentially increased in past few years, the possibility of electrocardiogram (ECG) compression still attracts the attention, due to the huge amount of data that has to be stored and transmitted; the amount that grows (depending upon the sampling rate, quantization levels and number of sensors) at the rate of 7.5-540 KB per minute per patient, depending upon the time and amplitude, sampling rate and number of sensors. Besides the increased storage capacity for archival purposes, ECG compression allows real-time transmission over telephone networks, economic off-line transmission to remote interpretation sites. A wide range of compression techniques based on different transformation techniques like DCT, FFT were evaluated to find an optimal compression strategy for ECG data compression. Wavelet compression techniques were found to be optimal in terms of compression.

Index Terms— ECG, Compression, DWT, CR

I. INTRODUCTION

High performance compression methodologies for the transmission of medical data and images play an important part in High-speed, cost-effective, and efficient networked telemedicine. Wavelet-based compression techniques and tools have received significant attention, especially for different biomedical signal-processing applications [1] [2]. The techniques proposed in this paper account for exploring efficient electrocardiogram (ECG) data compression methodologies that are suitable for mobile telecardiology applications that are computationally efficient and can reduce transmission times at the Mobile transmitter end and simultaneously provide clinically acceptable ECG data at the receiver end.

The main goal of the methods presented here is to achieve sufficiently high compression ratios (CRs) without

affecting the diagnostic characteristics of the ECG signals. Previous ECG data affecting the diagnostic characteristics of the ECG signals. Previous ECG data compression techniques have been reported in the literature and include direct time-domain techniques, transform-domain techniques [3] [4], average beat subtraction, and different modeling methods [5] [6].

II. ELECTROCARDIOGRAM (ECG)

The ECG is nothing but the recording of the heart's electrical activity. The deviations in the normal electrical patterns indicate various cardiac disorders. The "Electrocardiogram" (ECG) is an invaluable tool for diagnosis of heart diseases. The volume of ECG data produced by monitoring systems can be quite large over a long period of time and ECG data compression is often needed for efficient storage of such data. Similarly, when ECG data need to be transmitted for telemedicine applications, data compression needs to be utilized for efficient transmission. ECG systems are used by paramedics responding to accident scenes in emergency vehicles. They are also used by clinicians at remote sites. Certain military and/or space missions also employ ECG.

The objective of this project is to compress a complex ECG that easily to transmit from remote areas to the databank of the clinical chamber. Different compression techniques are used in ECG waveform application. A decomposer (filter) and reconstructor (filter) based on a new emerging transformation technique; called as '**Wavelet Transformation**' coding is chosen and found to give a better compression on real world offline ECG signals tested while having a low complexity and fast random access to arbitrary part of the ECG signal, have much better compression by selecting the threshold in filtering techniques while retaining the essential characteristics of the ECG waveform. The ECG as shown in Figure 1 records the electrical activity of the heart, where each heart beat is displayed as a series of electrical waves characterized by peaks and valleys. Any ECG gives two kinds of information.

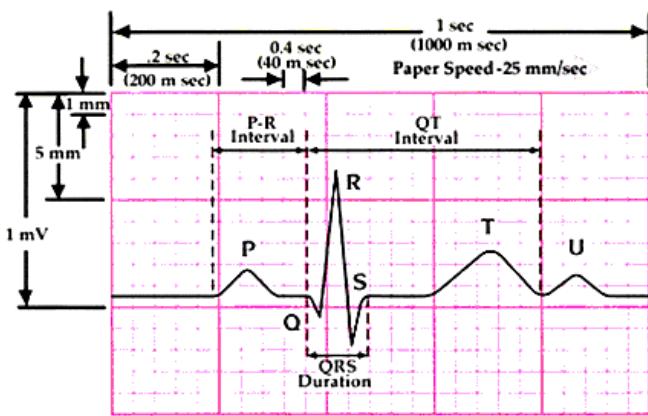


Figure 1. A typical cardiac waveform

One, the duration of the electrical wave crossing the heart which in turn decides whether the electrical activity is normal or slow or irregular and the second is the amount of electrical activity passing through the heart muscle which enables to find whether the parts of the heart are too large or overworked. Normally, the frequency range of an ECG signal is of 0.05–100 Hz and its dynamic range of 1–10 mV. The ECG signal is characterized by five peaks and valleys labeled by the letters P, Q, R, S, T. In some cases we also use another peak called U. The performance of ECG analyzing system depends mainly on the accurate and reliable detection of the QRS complex, as well as T- and P waves.

The P-wave represents the activation of the upper chambers of the heart, the atria, while the QRS complex and T-wave represent the excitation of the ventricles or the lower chamber of the heart. The QRS complex is the most striking waveform within the ECG. Since it reflects the electrical activity within the heart during the ventricular contraction, the time of its occurrence as well as its shape provide much information about the current state of the heart. Due to its characteristic shape it serves as the basis for the automated determination of the heart rate, as an entry point for classification schemes of the cardiac cycle, and often it is also used in ECG data compression algorithms. In that sense, QRS detection provides the fundamentals for almost all automated ECG analysis algorithms. Once the QRS complex has been identified a more detailed examination of ECG signal including the heart rate, the ST segment etc. can be performed.

III. WAVELET TRANSFORM

Mathematically speaking, the wavelet transform is a convolution of the wavelet function $\psi(t)$ with the signal $x(t)$. Orthonormal dyadic discrete wavelets are associated with scaling functions $\varphi(t)$. The scaling function can be convolved with the signal to produce approximation coefficients S . The discrete wavelet transform (DWT) can be written as:

$$T_{m,n} = \int_{-\infty}^{\infty} x(t) \psi_{m,n}(t) dt$$

By choosing an orthonormal wavelet basis, $\psi_{m,n}(t)$, and we can reconstruct the original. The approximation coefficient of the signal at the scale m and location n can be presented by:

$$S_{m,n} = \int_{-\infty}^{\infty} x(t) \varphi_{m,n}(t) dt \quad (1)$$

In practice our discrete input signal $S_{0,n}$ is of finite length N , which is an integer power of 2: $N = 2^M$. Thus the range of scales that can be investigated is $0 < m < M$. A discrete approximation of the signal can be shown as

$$X_0(t) = x_m(t) + \sum_{m=1}^M d_m(t) \quad (2)$$

Where the mean signal approximation at scale M is

$$x_M(t) = S_{m,n} \varphi_{m,n}(t) \quad (3)$$

And the detail signal approximation corresponding to scale m is defined for a finite length signal as

$$d_m(t) = \sum_{n=0}^{2^{M-m}-1} T_{m,n} \varphi_{m,n}(t) \quad (4)$$

Adding the approximation of the signal at scale index M to the sum of all detail signal components across scales gives the approximation of the original signal at scale index 0. The signal approximation at a specific scale was a combination of the approximation and detail at the next lower scale.

$$x_m(t) = x_{m-1}(t) - d_m(t) \quad (5)$$

If scale $m = 3$ was chosen, it can be shown that the signal approximation is given by

$$X_3(t) = x_0(t) - d_1(t) - d_2(t) - d_3(t) \quad (6)$$

Corresponding to the successive stripping of high frequency information (contained within the $d_m(t)$) from the original signal at each step [4]. This is referred to as multiresolution analysis of a signal using wavelet transform, and is the basic of our procedure.

In discrete case, the wavelet transform is modified to a filter bank tree using the Decomposition/ reconstruction given in Figure 2.

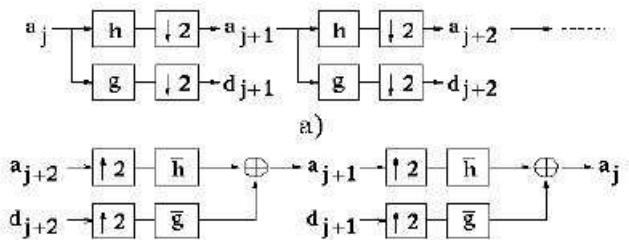


Figure:-2 (Filter bank tree of a) Decomposition and b) Reconstruction)

In the present work, Symlet wavelet is chosen although the Symlet algorithm is conceptually more complex and has a slightly complicated computations, yet this algorithm picks up minute detail that is missed by other wavelet

algorithms, like Haar wavelet algorithm. Even if a signal is not represented well by one member of the Symlet [Dau92] family, it may still be efficiently represented by another.

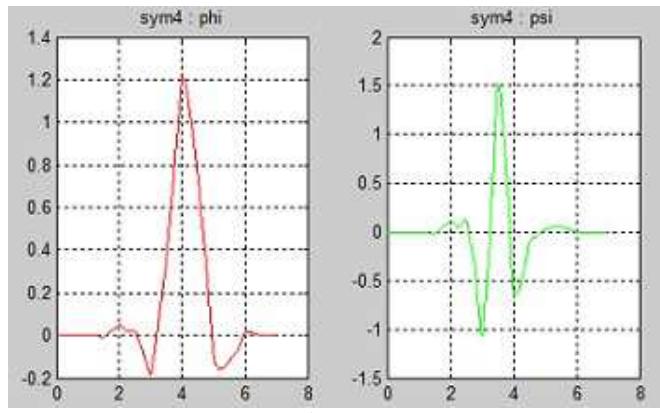
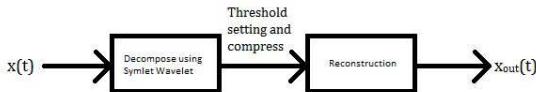


Figure 3

The phi part and the psi part of sym4 which are used in our experiment that is shown in Fig:-3.

IV. APPLYING WAVELET IN ECG

Here we use Symlet Wavelet family. At first the sym4 is applied on the ECG wave. Here we choose the level 7 multi level decomposition. The output after decomposition we gate level 7 approximations where the low frequency part is present and level 1 to level 7 detail part which we gate from high pass filter where the higher frequency part is present in decreasing order.



In case of ECG, the most of the portions are P-Q-R-S-T-U peaks are maximum in the higher frequency part. For that reason when we decompose the ECG signal, the maximum data remains on the detail part. We see that detail 1, detail 2 and detail 3 contain very low information part of ECG signal, though in this part, maximum frequency are present. In detail 4, detail 5 and detail 6, in ‘sym4’ decomposition we find the maximum part of ECG. But detail 7 and approximate 7 we may say very low frequency part, contain low information part. That’s why when we select the threshold; we have to be careful to set the threshold on detail 4, detail 5 and detail 6. In detail 7 we set the threshold in a very low scale.

After selecting the threshold of the decomposition part when we compress and reconstruct the ECG signal then the outcomes of the sym 4, some portion which has been eliminated by selecting the threshold, has been filled by null value by default in wavelet.

V. Result and Discussion

In this paper, we have presented an effective ECG Data Compression algorithm to be implemented in MATLAB based on a newly developed mathematical tool i.e. Wavelet Transformation. There exist many algorithm schemes based on different transformation techniques. So for reliability and existence of our threshold based compression algorithm we have compared with global threshold which has been present in wavelet compression. To reach our goals we followed certain steps which include observing and analyzing the different wave forms and selecting the threshold and compare with the global threshold compression.

The original signal which has been taken is shown below fig4.

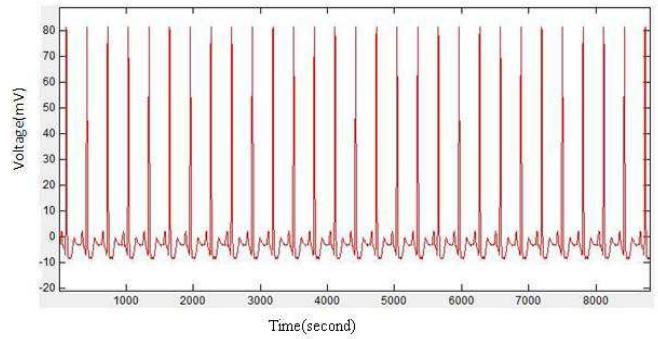


Figure:-4 (Original Signal)

After decomposition of level 7 using Sym 4 we get the seven detail parts and approximation part is shown in Fig5

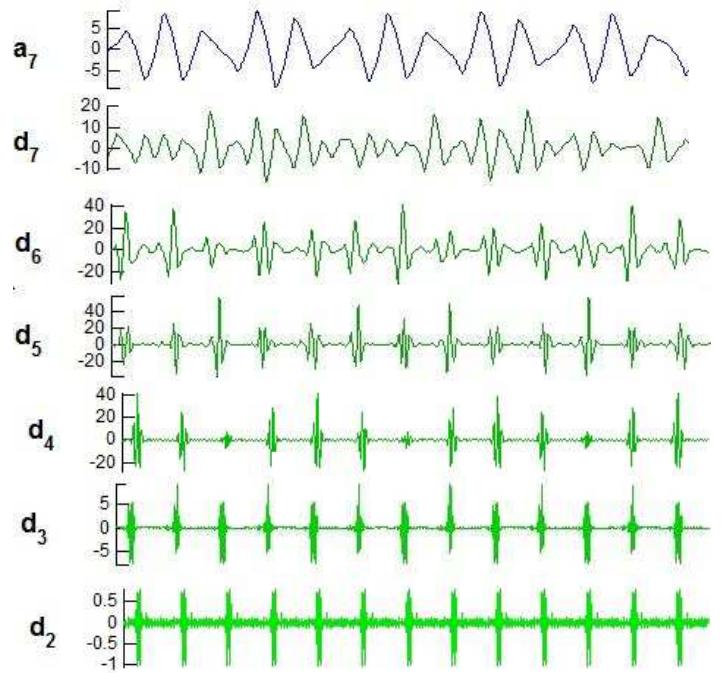




Figure:-5 (approximation and detail part)

After selecting the threshold of the detail part of the decomposed ECG signal, then reconstruct the decomposed signal we found the compressed signal which is shown in figure 6

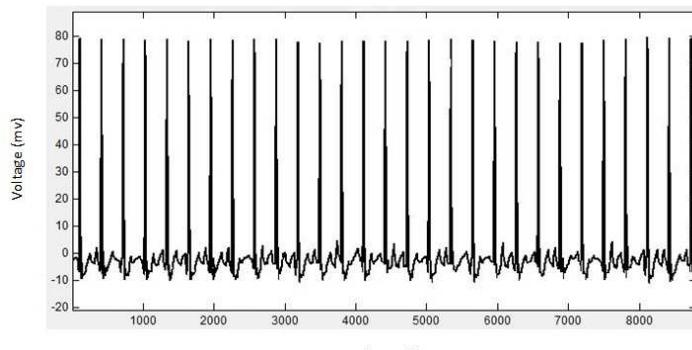


Figure:-6(compressed signal)

Here we see that the retained energy is 97.9% and zeros in that signal is 96.53%

VI. CONCLUSION

In this experiment, we have presented an effective ECG Data Compression to be implemented in MATLAB based on a newly developed mathematical tool i.e. Wavelet Transformation. There exist many algorithm schemes based on different transformation techniques. For reliability & existence of our presented technique to reach our goal, we followed certain steps which include observing & analyzing the different waveforms and determining various performance evaluation parameters obtained in MATLAB.

The wavelet compression technique is very effective compared to other compression method. The method that we have used in wavelet transform technique seems to be the best choice for ECG data compression.

V. Reference

- [1] M. Akay, *Time Frequency and Wavelets in Biomedical Signal Processing*. Piscataway, NJ: IEEE Press, 1998, sec. Biomed. Eng...
- [2] M. L. Hilton, "Wavelet and wavelet packet compression of electrocardiograms," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 394–402, May 1997.
- [3] B. R. S. Reddy and I. S. N. Murthy, "ECG data compression using Fourier descriptors," *IEEE Trans. Biomed. Eng.*, vol. BME-33, pp. 428–434, 1986.
- [4] W. Philips, "ECG data compression with time-warped polynomials," *IEEE Trans. Biomed. Eng.*, vol. 40, pp. 1097–1100, Nov. 1993.
- [5] P. S. Hamilton and W. J. Tompkins, "Compression of ambulatory ECG by average beat subtraction and residual differencing," *IEEE Trans. Biomed. Eng.*, vol. 38, pp. 253–259, 1991.
- [6] M. L. Hilton, "Wavelet and wavelet packet compression of electrocardiograms," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 394–402, May 1997.
- [7] S. Jalaleddin, C. Hutchens, R. Stratton, and W. Coberly, "ECG data compression techniques – a unified approach", *IEEE Trans. On Biomedical Engineering*, vol. 37, pp. 329-343, 1990.
- [8] Z. Lu, D. Y. Kim, and W. A. Pearlman, "Wavelet compression of ECG signals by the set partitioning in hierarchical trees algorithm", *IEEE Trans. on Biomedical Engineering*, vol. 47, pp. 849-856, July 2000.
- [9] Chen J., Itoh S., and Hashimoto T., (1993). ECG data compression by using wavelet transform. *IEICE Trans. Inform. Syst.*, vol. E76-D, 1454–1461, Dec. 1993.
- [10] M. L. Hilton, "Wavelet and wavelet packet compression of electrocardiograms," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 394–402, May 1997.

A Mathematical Analysis Of Blood Flow Through Stenosed Arteries:A Non-Newtonian Model

¹B.Basu Mallik & ²Saktipada Nanda

Department of Basic Science & Humanities

Institute of Engineering & Management, Salt Lake Electronics Complex

Kolkata - 700091, West Bengal, India.

e-mail: ¹b.basumallik@gmail.com ²saktipada_nanda@yahoo.com

Abstract— A mathematical model is developed in this paper for studying blood flow through a narrow artery with multiple stenoses. Consideration of Non-Newtonian character of blood is taken into account by taking constitutive equation of blood described by Herschel- Bulkey equation. An improved shape of the stenosis is described with boundary specified by Krogh model. The two important characteristics viz. the presence of yield stress and the dependence of viscosity with respect to the shear rate have been taken into account in the constitutive equation. The analysis bears the potential to explore a variety of information regarding some phenomenological aspects of this physiological problem. Analytical expressions for fluid velocity, plug flow velocity, flow rate and the shear stress are derived by using the model. The derived expressions are computed numerically and graphical representations of the variation of the physical quantities with change in the value of the parameters are computed to give an insight on stenotic diseases. The results for Newtonian model of blood are obtained as special case from the model. A medical treatment is suggested to address the problem realistically by careful use of catheter.

Keywords- *Stenosis, Herschel-Bulkley model, shearing stress, viscosity, fluid index.*

I. INTRODUCTION

The medical term ‘Stenosis or Arteriosclerosis’ means narrowing of any body passage, tube or orifice, and is a frequently occurring cardio-vascular disease in human arteries. Arteries are narrowed by abnormal and unusual deposition of cholesterol and some other substances resulting in the development of stenosis in the arteries. One of its most serious consequences is the increased resistance to the blood flow bringing about significant alterations in pressure distribution, wall shear stress and the flow resistance (impedance). The altered haemodynamic system (the flow of blood) may further influence the development of the disease and arterial deformity and change the regional blood rheology. Then there is significant fluid retention in the body, including the heart-muscle, leading to multi-organ failure. Also the deposition of atherosclerotic plaques in the coronary arteries gradually

diminishes the flow of blood through the arteries into the heart leading to cardiac ischemia. Although the exact mechanisms responsible for the initiation of this phenomena are not clearly known it is believed that the fluid dynamical parameters, particularly the wall shear stress play an important role in the genesis of the disease.

A survey conducted on the previous literatures indicates that the studies are based on Newtonian behavior of blood and single symmetric or non-symmetric stenosis. But the stenosis may develop in series (multiple stenoses) or may be of irregular shape or overlapping (composite). Also the assumption of Newtonian behavior of blood (single phase homogeneous viscous fluid) is acceptable for high shear rate in case of a flow through larger arteries. Many analytical as well as experimental studies confirmed that blood being a suspension of erythrocytes (red cells) in plasma, exhibits remarkable non-Newtonian behavior when it flows through narrow arteries at low shear rates, particularly, in diseased state. Herschel-Bulkley (H-B) fluid model and Casson fluid models are used in the theoretical studies of blood flow (non-Newtonian nature) through narrow arteries. Investigators have mentioned that blood obeys H-B equation at low shear rates when flowing through a tube of diameter 0.095 mm or less and represents fairly closely occurring flow of blood in arteries.

A review of the literature on the effects of stenosis in the arterial flow of blood indicate the contribution of Young [1] Shukla et al.[2] Chaturany and Sami [3] , Misra and Chakraborty [4]. These studies are based on the assumption that blood behaves like a Newtonian fluid. Srivastava et al [5] considered the blood flow through a composite stenosis in catheterized arteries assuming that the flowing blood behaves like a Newtonian fluid. The mathematical analysis presented by Sankar [6] discusses the pulsatile flow of blood through stenosed narrow artery with body acceleration, treating blood as H-B fluid. Jain et al. [7] examined the effects of mild stenosis on blood flow, in an irregular axi- symmetric artery with oscillating pressure gradient.

In this study, we have mathematical analysis of blood flow over stenosis in narrow constricted human artery. Theoretical analysis is carried out for situations where the non-Newtonian characteristics of blood are depicted by Herschel-Bulkey fluid. An extensive qualitative analysis is carried out by performing numerical computations of the desired quantities having more physiological significance to explore the effects of viscosity, shear stress and velocity over the stenosis. Finally, comparisons are made with the other existing results to substantiate the applicability of the present model under study.

II. MODEL DESCRIPTION

Let us consider the axi-symmetric flow of blood through an artery of circular cross section with a mild stenosis specified at the position shown in Fig.1. The blood is modeled as Herschel-Bulkey fluid. It is assumed that the flow is laminar, viscous and incompressible. The shape of the tube is under zero pressure and it is further assumed that the tube wall have no axial motion so that no slipping takes place between the fluid and the wall. The pressure gradient is prescribed oscillatory which is compatible with a pumping heart motion. The length of the artery is assumed to be large enough when compared to its diameter, so that entrance, end and special wall effects may be neglected.

The geometry of the stenosis is assumed to be manifested in the arterial segment given by (Kapoor, et al [8])

$$\frac{R}{R_0} = 1 - \frac{\delta}{2R_0} \left\{ 1 + \cos \frac{2\pi}{L_0} (z - d - \frac{L_0}{2}) \right\}, d \leq z \leq L_0 + d \\ = 1, \text{ otherwise} \quad (1)$$

Where R_0 is the radius of unobstructed tube and R is the radius of obstructed tube. L_0 is the length of the stenosis and d is the location of the stenosis. The maximum height of stenotic growth is taken as δ . The schematic diagram is shown in Figure 1.

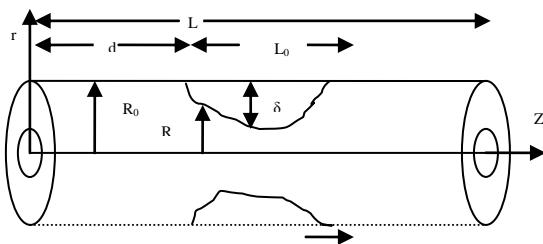


Fig1.Geometry of the stenosed artery

The momentum equation of motion is

$$\nabla \cdot V = 0 \quad (2)$$

$$\rho \frac{dv}{dt} = -\nabla p - \nabla \cdot \tau \quad (3)$$

Where ρ is the density, p is the pressure, and τ is the shearing stress tensor.

Herschel-Bulkley law to model the fluid behavior of blood flow, taking into account two characteristic features, which has emerged from the experimental data namely:

- The presence of a yield stress,
- The dependence of the viscosity with respect to the shear rate. (Kapoor, et al [8])

Let τ_0 be the yield stress, the coefficient of viscosity is μ and γ' be the strain rate.

Then constitutive equation in one dimensional form for Herschel-Bulkley pulsatile fluid with the shearing stress τ , is given by

$$\tau = \mu(\gamma')^n + \tau_0, \tau \geq \tau_0 \\ \gamma' = 0, \tau < \tau_0 \quad (4)$$

(Tu et al [9])

The governing equation of motion for steady incompressible blood flow with pressure gradient through the mild stenosis in an artery reduces to the following form:

$$P = \frac{1}{r} \frac{\partial}{\partial r} (r\tau) \quad (5)$$

where

$$\frac{\partial p}{\partial z} = -P \quad (6)$$

P being a constant. Integrating Equation (5) with respect to r which is the radial co-ordinate, we have,

$$\tau = P \frac{r}{2} \quad (7)$$

From Equations 4 and 7, we have

$$\frac{Pr}{2\mu} - \frac{\tau_0}{\mu} = (\gamma')^n \quad (8)$$

For the Herschel-Bulkley fluid in circular tube, we have $\gamma' = 0$ when $\tau \leq \tau_0$ and there is a core region which flows as a plug.

Let the radius of this plug region be r_p . At the surface of this plug, the stress is τ_0 , so that considering the force on the plug, we get

$$P \times \pi r_p^2 = \tau_0 \times 2\pi r_p$$

$$\text{Or } P \times r_p = 2\tau_0 \quad (9)$$

Then Equation (8) changes to

$$(\gamma')^n = \frac{\Pr}{2\mu} - \frac{\Pr_p}{2\mu} \quad (10)$$

$$\text{We know that } \frac{dv}{dr} = \gamma' \quad (11)$$

Then equation (10) changes to

$$\frac{dv}{dr} = \left(\frac{P}{2\mu}\right)^{\frac{1}{n}} (r - r_p)^{\frac{1}{n}} \quad (12)$$

The boundary conditions are

$$v = 0 \text{ at } r = R \text{ and } R_0 \quad (13)$$

Integrating (10) and using conditions (13), we get

$$v = \frac{n}{n+1} \left(\frac{P}{2\mu}\right)^{\frac{1}{n}} (R)^{\frac{1+n}{n}} \left\{ \left(\frac{r}{R} - \beta\right)^{\frac{1+n}{n}} - (1-\beta)^{\frac{1+n}{n}} \right\} \quad (14)$$

where

$$\frac{r_p}{R} = \beta \quad (15)$$

Plug flow exists whenever the shear stress does not exceed yield stress. The velocity of the plug flow can be obtained by putting

$$r = \beta R \quad (16)$$

Then we get

$$v_p = -\frac{n}{n+1} \left(\frac{P}{2\mu}\right)^{\frac{1}{n}} \left(\frac{r}{\beta}\right)^{\frac{n+1}{n}} (1-\beta)^{\frac{n+1}{n}} \quad (17)$$

Flow rate Q is obtained as follows:

$$Q = \frac{n\pi}{3n+1} \left(\frac{P}{2\mu}\right)^{\frac{1}{n}} c^3 \left(1 - \frac{3n+1}{n(2n+1)} \beta\right) \quad (18)$$

where

$$c = (-R)^{\frac{1+n}{n}} \quad (19)$$

The shear stress τ_w is obtained as follows:

$$\tau_w = \left(\frac{(3n+1)Q\mu^n}{n\pi c^3}\right)^{\frac{1}{n}} \left(1 + 3n \frac{\delta}{c}\right) \left(1 + \frac{3n+1}{2n+1} \beta\right) \quad (20)$$

III. NUMERICAL RESULTS & DISCUSSION

To discuss the results of the study quantitatively, the values of the different constants and rheological parameters from standard literatures are taken. The values of the power law index 'n' for blood flow problems are generally taken to lie between 0.9 and 1.1(Sankar, D.S.[6]). In this analysis, we have used the value 0.95 for n<1 and 1.05 for n>1. In order to validate our theoretical procedure, the results obtained on the basis of the present study for n = 1 (Newtonian model) were compared with those reported in Mishra et al. [4]

The change in the flow pattern, the shear stress in an artery are estimated if blood is modeled as non- Newtonian (H-B type). The variation of the fluid velocity (v) with radius of obstructed tube(r) for different values of the coefficient of viscosity (μ) and taking n = 0.95 are exhibited in Fig.2. For a value of μ , the velocity increases with r. Naturally in the absence of any stenosis in the artery, the core radius will be maximum and the flow will be free indicating the increase in velocity. Also as μ (the coefficient of viscosity) increases the velocity decreases for a fixed r which is in full agreement with the physical nature of blood (Blood is more viscous than water). Actually viscosity describes a fluid's internal resistance to flow and may be thought of as a measure of fluid friction.

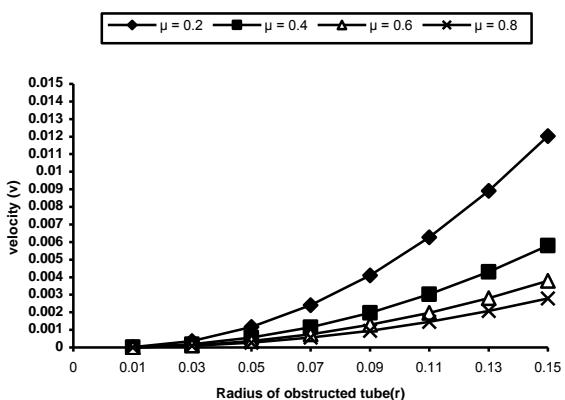


Fig.2 Velocity profile vs. radius of obstructed tube(r) for different values of μ and $n = 0.95$.

Fig.3 and Fig.4 depicts the same variation taking $n = 1$ & 1.05.

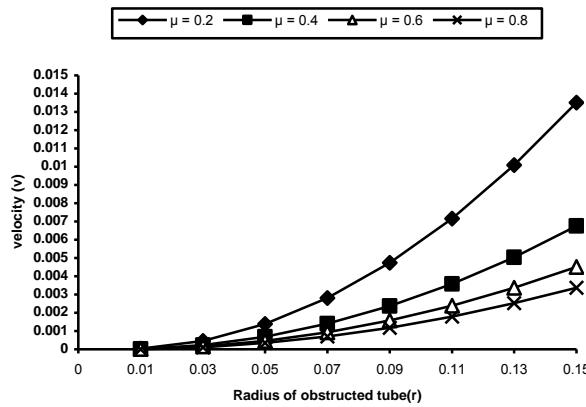


Fig.3 Velocity profile vs. radius of obstructed tube(r) for different values of μ and $n = 1$

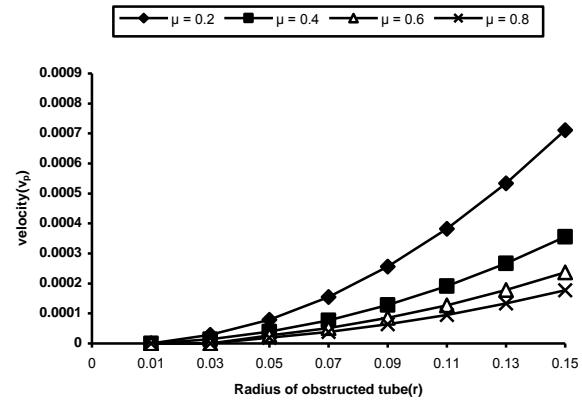


Fig.6 Velocity profile of plug region vs. radius of obstructed tube(r) for different values of μ and $n = 1$

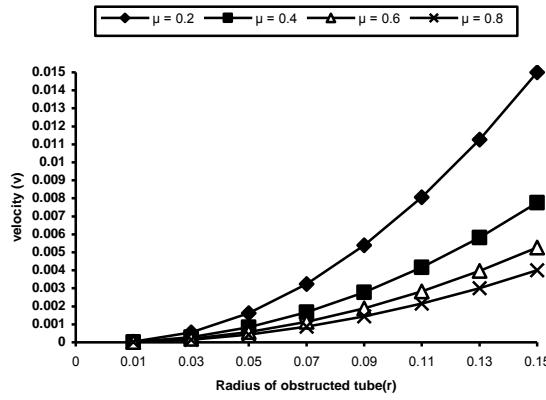


Fig.4 Velocity profile vs. radius of obstructed tube(r) for different values of μ and $n = 1.05$

The variation of plug core velocity (v_p) with r and $n = 0.95$, 1 and 1.05 are shown in Fig.5, Fig.6 and Fig.7 respectively. It is found that the velocity decreases with increase of μ for a fixed r .

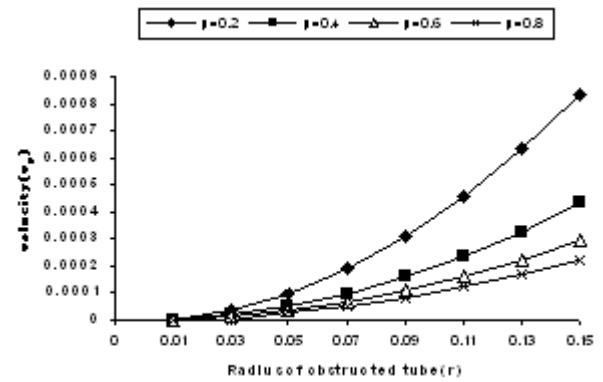


Fig.7 Velocity profile of plug region vs. radius of obstructed tube(r) for different values of μ and $n = 1.05$

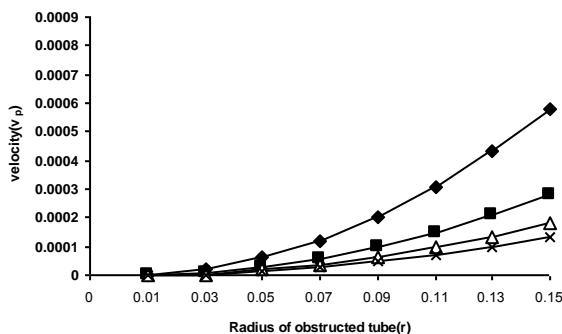


Fig.5 Velocity profile of plug region vs. radius of obstructed tube(r) for different values of μ and $n = 0.95$

It is observed that the shear stress increases with decrease in r . $r = 0.03$ shows a rapid fall in the value of the shear stress (τ_w). The shear stress become higher if the Newtonian model of the blood is considered i.e. non-existence of erythrocytes in the plasma.

These are the quantities of physiological relevance

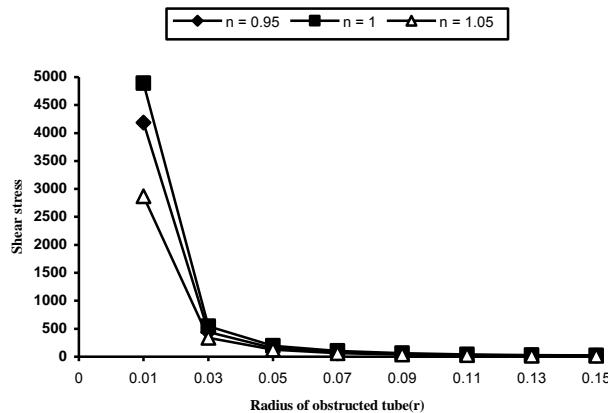


Fig.8 Variation of shear stress (τ_w) vs. radius of obstructed tube(r) for different values of n .

IV. CONCLUSION

The use of catheters may be of immense importance and a standard tool for diagnosis and treatment of stenosis. Transducers attached to catheters are of large usage in clinical works. This technique may be used for detection and cure of the disease. The catheter is carefully guided to the location of stenosis and the balloon is inflated to fracture the fatty deposits and widen the narrowed portion of the artery. It requires to estimate the flow characteristics and the shape of stenosis before artery catheterization. The artery length should be large enough as compared to its radius so that the entrance, end and special wall effects may be neglected. Further careful investigations are suggested to address the problem more realistically and to overcome the restrictions imposed on the present work.

REFERENCES

- [1] D.F. Young, "Effects of a time-dependent stenosis on flow through a tube," *J. Eng. Ind., Trans ASME*, vol.90, pp. 248-254, 1968.
- [2] J.B. Shukla,R.S. Parihar and S.P.Gupta, " Effects of stenosis on non-Newtonian flow of the blood in an artery, *Bull Math Biol.*, vol. 42., pp. 283-294,1980.
- [3] P. Chaturani and R. P. Samy, "Pulsatile flow of Casson's fluid through stenosed arteries with application to blood flow," *Biorheology*, vol.23, pp. 499 – 511, 1986.
- [4] J.C. Misra and S. Chakravarty, "Flow in arteries in the presence of stenosis," *J.Biomech.*, vol. 19 (11), pp. 907-918 , 1986.
- [5] V.P.Srivastava, R.Vishnoi, S.Mishra, and P. Sinha, "Blood flow through a composite stenosis in catheterized arteries," *E- journal of science & technology*, vol. 4(5), pp. 55-64, July 2010.
- [6] D.S.Sankar, "Mathematical analysis of blood flow through stenosed arteries with body acceleration," *Proceedings of the National Conference on Applied Mathematics(NCAM)*, pp. 15-22, 2010.
- [7] M.Jain,G.C. Sharma and S.K. Sharma, " A Mathematical model for blood flow through narrow vessels with mild-stenosis" , *IJE Transactions B:Applications*, vol. 22(1), pp.99–106, April 2009.
- [8] J.N. Kapoor , "Some Mathematical Models in Medical Sciences," *Internat. J. Math. Edn. Sc. Tech*, vol 115 .pp. 587-600, 1984.
- [9] C. Tu and M. Deville, "Pulsatile flow of Non-Newtonian fluids through arterial stenoses," *J. Biomech*, vol. 29(7), pp. 899 – 908, 1996.

COMPUTER COMMUNICATION & SYSTEMS

Date : 17th January,2012

Chairman : Prof. Anup Kumar Bandyopadhyay

Co-Chairman : Prof.S.Chakrabarti

Rapporter : Prof. Sukanya Mukherjee

| Paper ID | Paper Title | Authors |
|----------|--|--|
| 5 | Processing Delay Analysis of MLFFA based on Queuing Model | Boudhayan Bhattacharya,Sharmistha Bhattacharjee and Dwijesh Dutta Majumder |
| 37 | Speech De-noising by Adaptive Filter and Wavelet Transform | Abhinaba Banerjee and Dulal Mandal |
| 49 | Fear and Risk Issues In Cloud Computing | Sarasij Majumdar, Hemanta Dey, Soumen Santra and Rakhi Bhattacharjee |
| 64 | Design and Implementation of Real Time AES-128 on Real Time Operating System for Multiple FPGA Communication | Rourab Paul, Sangeet Saha, Suman Sau, Amlan Chakrabarti |
| 92 | Microstrip Antenna For Hispeed WLANs | Priyanarayan Misra and Debasis Mishra |
| 95 | Energy Efficient Administrator Based Secure Routing For Manet | Aniruddha Bhattacharyya, Arnab Banerjee and Dipayan Bose |

Speech De-noising by Adaptive Filter and Wavelet Transform

Abhinaba Banerjee

Department of Electronics and Communication
Heritage Institute of Technology
Kolkata-700107, India
abhi2652254@gmail.com

Abstract— In many application of noise reduction, the changes in signal characteristics could be quite fast and random. This requires the utilization of adaptive algorithms. Least Mean Squares (LMS) and Normalized Least Mean Squares (NLMS) adaptive filters are being used in a wide range of signal processing application because of their simplicity in computation and implementation. Now LMS is computationally not much complex but is slow so a different method is tried along with wavelet transforms. In this paper adaptive filtering is combined with wavelet transform was presented, and the comparison is made with conventional methods of Spectral subtraction, Thresholding and Wiener Filtering. The proposed method is efficient and yields satisfying results.

Keywords- Adaptive Filter ,Normalized Least Mean Square(NLMS), Least Mean Square(LMS),Speech enhancement, Wavelet Transform.

I. INTRODUCTION

Speech is a form of communication in the daily life of every human being. People who cannot communicate by speech face a lot of problems in their daily life. Since human civilization started to communicate by speech, technological advancements were being carried out and for better speech enhancements. As applications for cellular and satellite systems are being enhanced day-by-day, so speech processing need to be enhanced for better results. There are four generic attributes of speech coding. They are complexity, quality, bit rate and delay. These attributes affect the applications of speech. There are other issues like robustness to transmission errors, multistage encoding/decoding, and accommodation of non-voice signals such as in-band signaling and voice band modem data play an important role in speech coding as well.

There are a number of technologies which can be used for speech enhancement. All provide quite satisfying results but for some problems, specific technologies are effective. So, here we are using Adaptive Filter algorithms Normalized Least Mean Square(NLMS) and Least Mean Square(LMS) [1] along with wavelet transform. Adaptive algorithms self-adjusts its transfer function according to an optimization algorithm driven by an error signal. Adaptive algorithms use a

Dulal Mandal

Department of Electronics and Communication
Heritage Institute of Technology
Kolkata-700107, India
dulal.mandal@heritageit.edu

feedback signal in the form error signal for noise cancellation, signal prediction, echo cancellation etc. NLMS is Normalized LMS has less computational complexity compared to that of LMS. But LMS shows better results than NLMS.

But by using only LMS better performance might not be achieved as the SNR value might be below threshold. So, Wavelet Transform can be used along with LMS filter for better results. Wavelet Transforms[13] provides frequency component information at all times. As Wavelet transform analyses both high and low frequency regions, it is very flexible. This property makes wavelet transform to be particularly suitable for speech signal processing where better time information is required at high frequencies to detect rapidly changing transients of the signal, while better frequency resolution is needed at low frequency. The other advantage of Wavelet Transform is its simplicity for hardware implementation. The computational complexity of fast WT is on the order of N while that of the FFT is $N \log_2 N$ where N is the data length.

The paper discusses Adaptive Filtering with wavelet transform. Before Wavelet Transform[14] was introduced by Donoho[7], Spectral Subtraction[5] and Wiener Filtering[4] were used for speech enhancement. But here we are using all these methods along with Adaptive Filtering with wavelet transform which proves to be an efficient method from the other two methods.

The sections II provides the basics of Adaptive Filtering, III and IV shows Wiener Filtering and Spectral Subtraction respectively. V shows the Wavelet Transform method and steps involved. Section VI shows the simulation results. Section VII concludes the paper.

II. OVERVIEW ON NLMS AND LMS ALGORITHMS

In Fig. 1, we show an adaptive filter where $Z(n)$ is the noise reference, $Y(n)$ is the noisy speech and $E(n)$ is the enhanced speech.

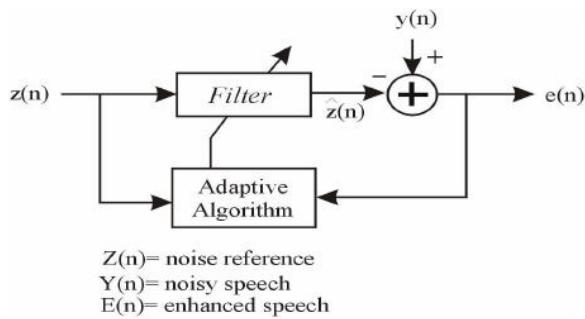


Fig.1: Adaptive system

It is well known that the filter vector update equation for the LMS algorithm[11] is given by :

$$w(n+1)=w(n)+\mu *Z(n)E(n) \quad (1)$$

where

$$Z(n)=w^T(n-1)u(n), E(n)=Y(n)-Z(n) \quad (2)$$

$u(n)$ is the vector of buffered input samples at step n, $w(n)$ is the vector of filter weights estimates at step n and μ is the step-size that determines the convergence speed and steady-state mean-square error (MSE). The LMS algorithm calculates filter weights as:

$$Z(n)=w^T(n-1)u(n) \quad (3)$$

$$E(n)=Y(n)-Z(n) \quad (4)$$

$$w(n+1)=w(n)+f(u(n),E(n),\mu) \quad (5)$$

The weight update function for the LMS adaptive filter algorithm is defined as:

$$f(u(n),E(n),t)=\mu *E(n)u^*(n) \quad (6)$$

$u^*(n)$:The complex conjugate of the vector of buffered input samples at step n.

III. WIENER FILTERING

The Wiener filter[8,10,12] is a filter whose purpose is to reduce the amount of noise present in a signal by comparison with an estimation of the desired noiseless signal. In order to achieve noise reduction with a highly reduced musical noise effect, the concept of a priori SNR estimate is being introduced into the classical speech enhancement, including Wiener filtering.

$$g_K=\xi_K/(\xi_K+1) \quad (7)$$

where K denotes the Kth frequency bin

The prior SNR ξ_K is estimated using the decision-directed method. g_K is the gain function in Wiener filtering. At frame m, $\xi_K(m)$ is estimated as:

$$\xi_K(m)=\alpha |\dot{X}_K(m-1)|^2 / |D_K(m-1)|^2 + (1-\alpha) * \max(|Y_K(m)|^2 / |D_K(m)|^2 - 1, 0) \quad (8)$$

α is the smoothing constant, $\dot{X}_K(m=1)$ denotes the enhanced signal spectrum at frame m=1, $Y_K(m)$ and $D_K(m)$ are the noisy speech and estimated noise spectra respectively.

$$\dot{X}_K(m)=Y_K(m)*g_K(m) \quad (9)$$

$g_K(m)$ is the gain function at frame m.

IV. SPECTRAL SUBTRACTION

Spectral subtraction[9] is a technique that exploits the idea that the human hearing system is insensitive to phase information Boll and Lim et al .proposed that the short-time spectral magnitude be used to estimate the speech spectrum . While a number of methods exist to extract this estimate of the noise spectrum and subtract it from the contaminated speech signal, the basic technique uses an estimate derived from:

$$|\dot{s}(\omega)|=\max\{|s(\omega)|^2-\beta|N(\omega)|^\alpha, \gamma|N(\omega)|\} \quad (10)$$

$s(\omega)$, $N(\omega)$ and $\dot{s}(\omega)$ are the magnitude spectrum of the noisy speech, noise estimation, and the enhanced speech, respectively. β is the over subtraction coefficient and γ is a parameter that determines the remaining noise for frames. This algorithm is useful for frames in which subtraction has a negative value.

V. WAVELET TRANSFORM METHOD

The performance of the adaptive system, in real situation, when the noise reference has a complicated relation SNR is low, resulting in poor performance. Thus a method along with the wavelet transform is being applied.

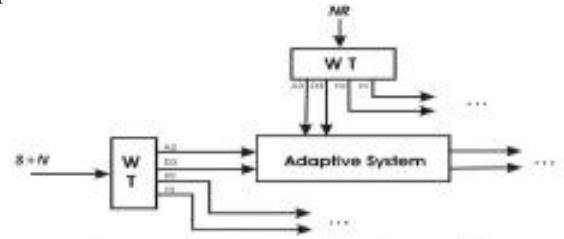


Fig.2: Adaptive system with wavelet transform

The essential energy of the wavelet transform[2,3,6] of Speech signals is concentrated in the first coefficients (lower scales). As the scale is increased, the coefficients are attenuated effectively. But for noise signals such as white noise, energy is spread in the entire coefficients uniformly. Hence the SNR value in the lower scales is more than the higher scales. By using the above adaptive filter in the lower scales, better SNR is attained, even, when the higher scales (last coefficients) are kept without any change. The wavelet transform causes the spectrogram of the noise reference and the noisy speech to be whiter, especially for the lower scales.

Thus the eigenvalues become more concentrated and the rate of convergence improves. Furthermore, because of more similarity between the noise of the noisy speech and the noise reference in each scale, the SNR improves. Figure 4.a and 4.b demonstrate that the adaptive system in the first scale has better SNR than the last scale.

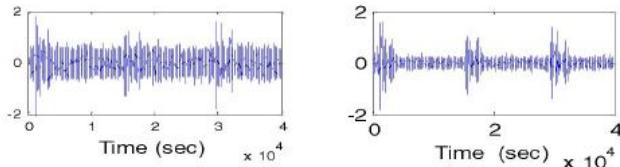


Fig. 3a: The noisy speech and the enhanced speech in the first scale(A3)

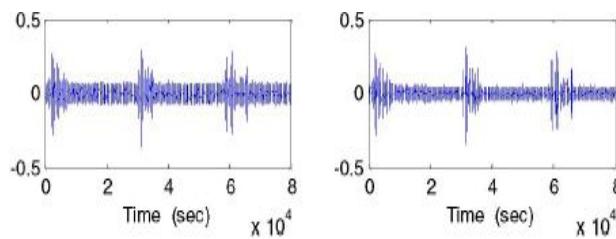


Fig.3b: The noisy speech and the enhanced speech in the last scale(D1)

Steps involved:

The enhancement algorithm of Adaptive filters along with wavelet transform may be summarized as follows:

- Transforming the noisy speech and noise reference by applying discrete wavelet transform. For this purpose, Daubechies filter of order 1 is employed (dbl) and the signal is decomposed into 4 levels of decomposition.
- Extract noise feature from the noise reference for Spectral Subtracting or Wiener filtering.
- Apply LMS algorithm to the main wavelet coefficients(Fig.3).
- Apply one of the following three methods for the rest of the coefficients.
 - Make them zero
 - Use Spectral Subtraction method
 - Apply Wiener filter using one of the following two approaches.
 - Maximum likelihood
 - Decision directed approach.
- Transform the obtained coefficients back to the time Domain.

VI. SIMULATION RESULTS

For our simulation, a noisy speech and noise reference which were sampled at 16 kHz and quantized to 16 bits. The observed speech and noise is divided into overlapping frames of length 512. The Overlap is nearly 5000. By this frame length the speech is nearly stationary in each frame. Furthermore, enough samples exist for adaptation process in each scale. The speech and noise are decomposed into four levels by using Daubechies filter of order 1. For adaptive filtering, LMS algorithm is applied. The filter length is set to 150. Simulation results show that t=.005 is a suitable value for our method. For comparison between all these methods, the differential SNR is calculated. It is assumed that the clean speech is not available. So for estimating the power of the clean speech, the power of the some of the silent frames is calculated and then is subtracted from the power of active frames.

$$\text{SNR} = 10 \log_{10} (\text{mean}(\mathbf{y}_{\text{active}})^2 - \text{mean}(\mathbf{y}_{\text{silent}})^2) / \text{mean}(\mathbf{y}_{\text{active}})^2 \quad (11)$$

TABLE I. Different SNR by using LMS Algorithm in the Time and Wavelet Domains

| SNR(dB)→METHOD↓ | -5 | 0 | 5 |
|------------------|------|------|------|
| Time- Adaption | 2.2 | 2.84 | 3.31 |
| Wavelet Adaption | 4.95 | 6.12 | 6.74 |

TABLE II. The Comparison of Improved SNR for different Suggested Methods in the Wavelet Domain

| Approach↓ SNR(dB)→ | -5 | 0 | 5 |
|--|------|------|------|
| Thresholding | 6.04 | 7.34 | 7.86 |
| Spectral Subtraction | 6.91 | 7.88 | 8.21 |
| Max Likelihood(Wiener Filter) | 6.96 | 8.04 | 8.59 |
| Decision Directed(Wiener Filter) | 8.48 | 9.22 | 9.91 |

From Table II it shows that Decision Directed approach gives better SNR than the maximum likelihood. The results of the maximum likelihood approach are similar to the spectral subtraction method. It must be mentioned that Wiener filtering has more distortion than the other methods; but it has the best

SNR. Figure 4 demonstrates the differential SNR for different approaches versus the initial SNR. Curves from bottom to up depict 'Time-adap', 'Wave-Thr', 'Wave-SS', 'Wiener-max' and 'Wiener-DD', respectively. 'Timeadap' is the same as traditional LMS algorithm. 'Wave-Thr' employs LMS in the lower scales (first coefficient of the DWT), whereas uses thresholding in the rest scales. 'Wav-SS' combines LMS to Spectral subtraction. For this purpose it uses LMS in the lower scales and Spectral Subtraction in the higher one. The 'Wiener-Max' method represents the hybrid of the LMS method in wavelet and the Wiener filter with the maximum likelihood approach. 'Wiener-DD' which has the best performance is the same as previous approach except that uses the wiener filter with decision directed approach in the last scales.

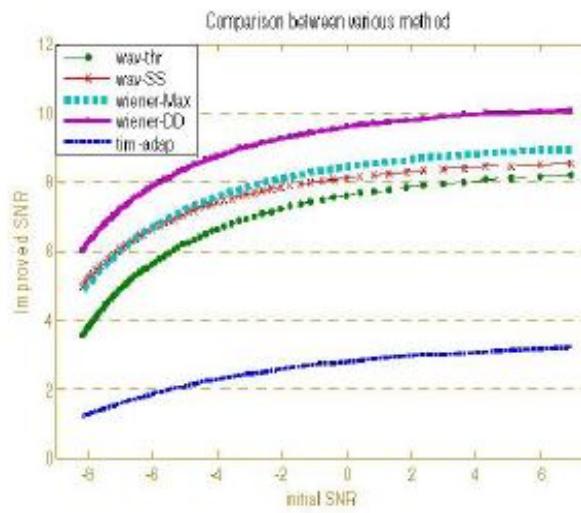


Fig.4: The comparison of Time adaption and proposed methods

VII. CONCLUSIONS

A very efficient adaptive system based on wavelet transforms for speech enhancement is proposed in this paper. The main idea is to use adaptation process in the lower scales. The main components of speech exist in these lower scales and the value of SNR is higher than the higher scales. Thus the adaptation process has a better performance. Furthermore, in the wavelet transform spectrum is divided into several scales. In the lower scales speech frames are present whereas in the higher frames, silent and non-voiced frames are present which have higher probability of being corrupted by noise . In order to enhance the speech signal in the last scales, three methods are used. The wavelet transform method shows better results than the other three methods. These approaches are based on using Thresholding method, Spectral Subtraction and Wiener

filtering that has been used for speech enhancement in the time domain. Experimental results confirm that this algorithm is computationally efficient with excellent results.

REFERENCES

- [1] S. Haykin, Adaptive Filter Theory, 3th Edition, Prentice Hall Inc.,1996.
- [2] C.S. Burrus , "Introduction to wavelet and wavelet transforms," Prentice Hall Inc., 1998.
- [3] Mohammad A Akhaee, Ali Ameri, Farokh A Marvasti, "Speech enhancement by adaptive noise cancellation in the wavelet domain", ICICS, pp.719-723,2005.
- [4] Fei Chen and Philipos C. Loizou, "Speech enhancement using a frequency-specific composite wiener function", ICASSP,pp.4726-4729,2010.
- [5] Boll SF: "A spectral subtraction algorithm for suppression of acoustic noise in speech," . Proceedings of IEEE ICASSP, pp.200–203, 1979.
- [6] H. Sheikhzadeh, and H. R. Abutalebi, "An improved wavelet-based speech enhancement system," Proc. of the Eurospeech, 2001.
- [7] D. L Donoho, and I .M Johnstone, " Ideal spatial adaptation via wavelet shrinkage, " Biometrika, vol. 81, pp.425-455, 1992.
- [8] A. Amehraye, D. Pastor, and A. Tamtaoui, "Perceptual improvement of wiener filtering," in Proc. ICASSP, Las Vegas, NV, pp. 2081-2084, 2009.
- [9] Boh Lim Sim, Yit Chow Tong, Joseph S. Chang, and Chin Tuan TanA, "A parametric formulation of the generalized spectral subtraction method," in IEEE Transactions On Speech and Audio Processing, Vol. 6, No.4, July 1998.
- [10] L. Lin, Kff. Holmes and E. Ambihirajnh, "Subband noise Estimation for Speech Enhancement using a Perceptual Wiener Filter," Proceedings of IEEE ICASSP,pp.80-83, 2003.
- [11] Douglas M. Chabries, Richard W . Christiansen, Robert H. Brey, Martin S.Robinette and Richard W. Harris, "Application of adaptive digital signal processing to speech enhancement for the hearing impaired," Journal of Rehabilitation Research and Development Vol . 24 No . 4 pp. 65-74, Fall 1987.
- [12] Jingdong Chen, Jacob Benesty, Yiteng (Arden) Huang, and Simon Doclo, " New Insights Into the Noise Reduction Wiener Filter," IEEE Transactions On Audio, Speech, And Language Processing, Vol. 14, No. 4, July 2006.
- [13] S.Manikandan, "Speech Enhancement Based On Wavelet Denoising," Academic Open Internet Journal, Vol.17,2006.
- [14] Amir Reza Momen, Hossein Ahmadi-Noubari, Alireza Mirzaee, "A New deniosing method in spread spectrum systems based on wavelet transform," CCECE/CCGEI, Saskatoon,May2005.

Fear and Risk Issues in Cloud Computing

Mr. Sarasij Majumdar,
Assistant Professor,
Techno India College of
Technology,
WBUT, W.B, INDIA
sarasisj.majumdar@gmail.com

Mr. Hemanta Dey,
Assistant Professor,
Techno India College of
Technology, WBUT,
W.B, INDIA
hemantadey13@gmail.com

Mr. Soumen Santra,
Student,
Kalyani Government
Engineering College,
W.B, INDIA
soumen70@gmail.com

Ms. Rakhi
Bhattacherjee., Lecturer,
Techno India College of
Technology, WBUT,
W.B, INDIA
say2rakhi@gmail.com.

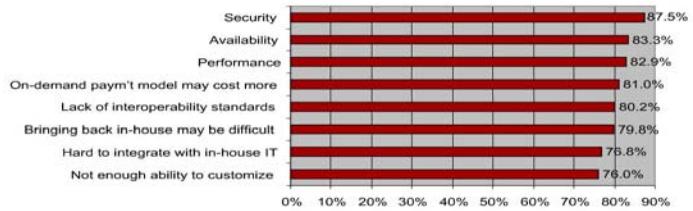
ABSTRACT - Cloud computing can be pointed as the next generation evolution in the world of computer after world wide web(www) and internet, which is actually as a on demand software solution. Cloud computing provides a platform for delivering software services and other applications through remote file servers rather than having an IT infrastructure on users own premises. In spite of these potentials the Cloud Computing has for revolutionizing many aspects of the software industry, there are significant shortcomings in the area of security and risk assessment. The most important significance of cloud computing is that rather than storing and accessing information in users own computers, data, information, and software are placed on remote servers and are accessible for you always, everywhere provided users have internet connections. The above scenario indicates it is firmly related security aspect in this field. The security challenges of cloud computing don't lie entirely in the technology itself but also on a lack of transparency, a loss of control over data assets, and unclear security assurances. We need to keep a view on different security domains and understand the processes that need to be followed in overall cloud security operation.

Keywords – Network Security; Virtual Machine; Data Leakage; Data security; Channel Attack;

INTRODUCTION

In today's world cloud computing is one of the popular internet based computing which provides a range of computing services to the enterprises as well as to the institutions. It dynamically delivers a complete package of services on user demand. The services provided through this technology is vast: like from network to operating system, from storage to hardware and software, resources everywhere. These services broadly classified into three types: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Cloud computing provides services in different platforms, yet there are some serious concerns regarding the security of the cloud. Some researchers already make some useful survey on the biggest challenges found in cloud computing security like "survey of IDC Enterprise Panel" [1] shows.

Q: Rate the challenges/issues of the 'cloud'/on-demand model



(Scale: 1 = Not at all concerned 5 = Very concerned)

Source: IDC Enterprise Panel, 3Q09, n = 263, September 2009

Proper security measures play an important role in preventing service failures and cultivating trust in cloud computing. In particular, cloud service providers need to secure the virtual environment, which enables them to run services for multiple clients and offer separate services for different clients. Some serious risk factors in cloud computing discuss over here. The Cloud Security Alliance's report [2] points out 13 different domains of risks. Here 18 different domains of security risks and preventive measures discuss to make cloud more secure. Some of the concerns are elaborately described below:

I. IDENTITY MANAGEMENT

Clouds deployment models are broadly classified in public cloud, private cloud, hybrid cloud and community cloud. In case of a public cloud, the responsibility for application security in terms of identity management and data protection depends only on the cloud provider. To improve security and user satisfaction, some public cloud providers offer identity management features such as SSO (Single Sign-On) and limited user provisioning. However, the majority of security and identity controls are with the public cloud provider. On the other hand a private cloud is an extension of the enterprise protected by a firewall. Such clouds are primarily integrated with client's identity management systems for different purposes such as SSO, user authentication, authorization, audit, provisioning, role management, and compliance. Furthermore, private clouds offer dedicated storage or virtualized layering for data isolation and application partitioning - reducing the risk of data breach.

2nd Annual International Conference IEMCON 2012

II. AUTHENTICATION AND AUTHORIZATION

Security authentication and authorization is a process for assessing the security of a system. It identifies if one has the access to enter in the cloud or not. It is an application to identify the risks. In authentication and authorization matters three types of security controls must be allocated.

- a) **System Specific Controls** – Controls implemented with an information system.
- b) **Common Controls** – Controls inherited by an information system.
- c) **Hybrid Controls** – Controls that have both system specific and common characteristics.

III. DATA LEAKAGE

Data leakage is a major concern for an organization as critical & sensitive data are kept in a cloud. It is mainly caused by multiple tenants sharing physical resources. It may happen sometime the users of the same cloud share the same data storage and data processing facilities; as a result they are exposed to the risk of information leakage, along with accidental or intentional disclosure of information.

The type of leaking or breaching of the data is by:

- Lost or Stolen
- Accidental
- Disposal
- Hacked
- Malicious(Viruses/Worm)
- Negligence

IV. VIRTUAL MACHINE (VM) PROTECTION

Virtual machines contain critical applications and sensitive data. VMs take these data off premise to public and shared cloud environments.

Virtual machines are dynamic in nature. They can quickly be reverted to previous instances, paused and restarted, relatively easily. They can also be readily cloned. In this way it can be seamlessly moved between physical servers. This dynamic nature and potential for VM sprawl makes it difficult to achieve and maintain consistent security. Henceforth, it is difficult to maintain an auditable record of the security state of a virtual machine at any given point of time. In cloud computing environments, it will be of utmost necessary to be able to prove the security state of a system, regardless of its location or proximity to other, potentially insecure virtual machines.

V. PHISHERS IN THE CLOUD

In the arena of cloud computing, phishing is a non-legal, sometimes a criminal activity using social engineering techniques. Phishers uses different fraud techniques to acquire sensitive information, such as personal details, passwords and credit card details, by entering the scenario as a trustworthy person or business in an electronic communication [3].

VI. PERSISTENT CLIENT-DATA SECURITY

In traditional models of software application, organizational data resides within the boundary of the organization. In this model data is subject to its physical, logical and personnel security and access control policies. In SaaS mode data is stored along with the application at the SaaS vendor end, so in this case proper safeguards need to be adopted for the data security. There also need to be strict safety to prevent leakage of sensitive data due to application vulnerabilities or malicious privileged users at the SaaS provider end.

VII. LACK OF TRANSPARENCY

By the way of accessing the cloud, some consumer's perception is cloud is generally less secure than an in-house system. Only better transparency could help addressing this issue. Data stored in a cloud provider's devices isn't located on a single machine in a single location or country. Rather, the data is stored and processed across the entire virtual layer. There are two issues involved in transparency: one is the physical location of the storage and processing sites, and the other is the security profiles of these sites.

VIII. NETWORK ATTACK

As the cloud computing is different from traditional softwares, it doesn't have a firewall normally. So the service provider and user should take special care of the infrastructure used for connecting and accessing the cloud.

IX. FORENSICS USED IN THE CLOUD

We know in the cloud, data processing is mainly of the decentralized nature. So traditional approaches like evidence collection and recovery is no longer practical. Unknown physical location of the companies' assets in the cloud further complicates the situation. Normally, if a security problem occurs, the corporate security team wants to be able to perform their own investigation without dependency on third parties. In the Cloud, this is not possible anymore. The Cloud Service Provider obtains all the power over the Cloud environment mainly biasing the way an investigation may be processed.

X. VULNERABILITIES IN THE CLOUD

Vulnerability scanning is a well-known practice of the Information Security community. Enterprises invest considerable time and money developing vulnerability management programs to help assess IT security risk across applications and infrastructure. Specifically, vulnerability scanners help identify potential security weaknesses at scale; e.g. missing patches, default passwords, coding or configuration weaknesses. A case of fixing of 3 major cloud vulnerabilities is shown in [15]. Vulnerability scanning is front of mind for Internet

2nd Annual International Conference IEMCON 2012

exposed or partner connected infrastructure. However, when said infrastructure is owned and/or operated by a service provider, some of the existing challenges associated with vulnerability scanning are magnified:

- Scans can cause outages. This can happen if the scanning policy includes Denial of Service checks or the scanning engine is configured with "aggressive" settings;
- Identifying unauthorized scans.
- Scanning may trigger automated or manual actions by the provider. A common automated response from a provider is to apply traffic shaping to slow down the scan, or simply block the client IP address via an ACL update.

XI. CROSS-VM SIDE-CHANNEL ATTACKS

Here an attacker could place a malicious virtual machine (VM) in close proximity to a target server in a shared, "cloud" environment. From there, it would be possible to launch a "cross-VM attack" using a variety of different hacking strategies. Such strategies might be employed by an attacker looking to access a specific target or server, or they might be used by hackers to cast a wide net and searching for any vulnerable server. The risk might be such that the automated methods used to place VMs in a cloud environment might be exploited to allow an attacker to intentionally create VMs and place them near a likely target server. The malicious VM then could use "side channels" to learn more about the target server, eventually collecting enough "leaked" data to allow penetration.

XII. AVAILABILITY OF SERVICE

In cloud computing service availability is the most important part of the services. Cloud Service Providers make sure the business organizations and a customer for all the time data is available for access. The application-layer distributed denial-of-service (DDoS) threat actually amplifies the risk to data center operators. That's because IPS devices and firewalls become more vulnerable to the increased state demands of this emerging attack vector – making the devices themselves more susceptible to the attacks. The remedies are pointed as: If one CSP is not functioning well, then one should use Multiple Cloud Providers to provide Business Continuity.

XIII. UPTIME

Uptime can be defined as the amount of time that a server has stayed up and running. This is usually listed as a percentage, like "99.9% uptime." Uptime is a perfect measure of how good a Web hosting provider is at keeping their systems up and running. If a hosting provider has a high uptime percentage, then we can say that their servers stay up and running and so any site you host with them should stay up and running too. As we all

know the web pages can't keep customers if they are down, so uptime is very important. [9].

XIV. SINGLE POINT OF FAILURES

No matter how well a network is designed single point of failures (SPOF) can always take place in the arena of cloud computing. Every engineering solution is it load-balanced or round-robin or off-site or cold-backup – everything is dependent on a single database query to provide their functionality. Again there is a human angle in every computing paradigm. So SPOFs can occur anytime anywhere. When you're living in the cloud, there is always a chance that a third party can make decisions about your data. Even the third party uses platform in ways never seen before in computing. It is known as Third Party Data Control. All the threats related to the third party are not understood well. The main reason behind it is that when a third party takes control of one's data, it is hard to decisions on which type of failure it is. In the above cases user himself/herself fails to understand the situation properly. Here he/she can just point out what problem such user is facing in the cloud.

XV. AUDITABILITY

To fully ensure the data security and save the cloud users' computation resources, it is of critical importance to enable public auditability for cloud data storage. This helps the users to go to a third party auditor (TPA), who has expertise and capabilities that the users do not, to audit the outsourced data when needed. Based on the audit result, TPA could release an audit report, which would not only help users to evaluate the risk of their subscribed cloud data services, but also be beneficial for the cloud service provider to improve their cloud based service platform [8]. In a word, enabling public risk auditing protocols will play an important role for this nascent cloud economy to become fully established; where users will need ways to assess risk and gain trust in Cloud.

XVI. DATA LOCK-IN

So far, there is no single specific cloud user interface to access cloud. Different cloud providers provide different data access method using different format. Many customers are afraid that by choosing a particular cloud service provider, they will enter a "lock-in" situation with that particular vendor. Due to the lack of standardization, most of the current cloud vendors have implemented and deployed proprietary solutions that lack interoperability with each other. Some cloud related standard organizations, like the Open Cloud Consortium, are working on interoperability standards for cloud computing and frameworks for interoperateing between clouds. Standardize APIs is a solution of the above problem. But that is not being done at today's cloud scenario.

2nd Annual International Conference IEMCON 2012

XVII. TRANSITIVE NATURE

A transitive dependency relationship means that if Domain A depends on Domain B and Domain B depends on Domain C, then an implicit dependency relationship exists automatically between Domain A and Domain C; For cloud computation the matter is almost same. Here a customer has a cloud provider. This cloud provider depends on other contractors. If these contractor(s) fail service to the customer will be affected.

XVIII. ACCESS CONTROLS

Different access control mechanisms have been evolving in the past thirty years and various techniques [10, 13] have been developed to effectively implement fine-grained access control, which allows flexibility in specifying differential access rights of individual users. Traditional access control architectures usually assume the data owner and the servers storing the data are in the same trusted domain, where the servers are fully entrusted as an omniscient reference monitor [14] responsible for defining and enforcing access control policies. This assumption however no longer holds in cloud computing since the data owner and cloud servers are very likely to be in two different domains. On one hand, cloud servers are not entitled to access the outsourced data content for data confidentiality; on the other hand, the data resources are not physically under the full control of the owner.

To keep an eye on security following are the best practice for companies in the cloud inquire about exception monitoring systems

- Be vigilant around updates and making sure that staffs don't suddenly gain access privileges they're not supposed to.
- Ask where the data is kept and inquire as to the details of data protection laws in the relevant jurisdictions.
- Seek an independent security audit of the host.
- Find out which third parties the company deals with and whether they are able to access your data.
- Be careful to develop good policies around passwords; how they are created, protected and changed.
- Look into availability guarantees and penalties.
- Find out whether the cloud provider will accommodate your own security policies.

CONCLUSION

In today's computing scenario cloud computing is one of the most happening things. With the evolution of cloud computing a new generation of computing arena is on the verge of starting. But the fear factor in the security in cloud computing is more important issues. Some of the big giant companies already hacked by cloud-based

system. JP Morgan and Citi banks were among those affected companies [22]. If service providers can provide these securities to customers and customers also behave responsibly, the numerous facilities of the cloud computing will be available to each one of us.

ACKNOWLEDGEMENT

The authors of this paper Mr. Sarasij Majumdar, Mr. Hemanta Dey, Mrs. Rakhi Bhattacharjee are thankful to the Research and Development Cell of Techno India College of Technology and Mr. Soumen Santra is thankful to Kalyani Government Engineering College, Kolkata, for necessary supports during the research work.

REFERENCES

1. Cloud Computing 2010: An IDC Update by IDC Enterprise Panel. <http://www.slideshare.net/JorFigOr/cloud-computing-2010-an-idc-update>.
2. Security Guidance for Critical Areas of Focus in Cloud Computing. <http://www.cloudsecurityalliance.org/guidance/csaguide.pdf>.
3. Phishers in cloud computing <http://www.mxpolice.com>. Amazon Elastic Compute Cloud (Amazon EC2) <http://aws.amazon.com/ec2/#details>.
4. Amazon Simple Storage Service (Amazon S3) <http://aws.amazon.com/s3/>.
5. Managing Amazon EC2 - SSH login and protecting your instances.<http://blog.taggesell.de/index.php?/archives/73-Managing-Amazon-EC2-SSH-login-and-protecting-your-instances.html>.
6. Sony PlayStation Network attack shows Amazon EC2 a hackers' paradise. <http://www.ibtimes.com/articles/146224/20110516/sony-playstation-network-qriocity-amazon-ec2-elastic-compute-iaas-cloud-computing-hackers-outage-wik.htm>.
7. M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents," Cryptology ePrint Archive, Report 2008/186, 2008, <http://eprint.iacr.org/>.
8. Latest cloud storage hiccups prompts data security questions. http://www.computerworld.com/s/article/9130682/Latest_cloud_storage_hiccups_prompts_data_security_questions.
9. Recent Breaches Spur New Thinking On Cloud Security. <http://www.darkreading.com/cloud-security/167901092/security/attacks-breaches/229402577/recent-breaches-spur-new-thinking-on-cloud-security.html>.
10. P. D. McDaniel and A. Prakash, "Methods and limitations of security policy reconciliation," in *Proc. of SP'02*, 2002.
11. T. Yu and M. Winslett, "A unified scheme for resource protection in automated trust negotiation," in *Proc. of SP'03*, 2003.
12. J. Li, N. Li, and W. H. Winsborough, "Automated trust negotiation using cryptographic credentials," in *Proc. of CCS'05*, 2005.
13. J. Anderson, "Computer Security Technology Planning Study," Air Force Electronic Systems Division, Report ESD-TR-73-51, 1972, <http://seclab.cs.ucdavis.edu/projects/history/> Format string attacks. <http://pferrie.tripod.com/papers/vtrootkits.pdf>.
14. Cloud Cartography & Side Channel Attacks <http://cloudsecurity.org/blog/2009/08/31/cloud-cartography-side-channel-attacks.html>.
15. Time to lay down the cloud computing law for uptime <http://searchcio.techtarget.com/news/1518323/Time-to-lay-down-the-cloud-computing-law-for-upptime>.

Processing Delay Analysis of MLFFA based on Queuing Model

Boudhayan Bhattacharya
 Assistant Professor
 Dept. of Computer Application
 Brainware Group of Institutions
 Kolkata, India
 mail4boudhayan@rediffmail.com

Sharmistha Bhattacharjee
 Scientist "C"
 Education & Training
 DOEACC Society
 Kolkata, India
 sharmi9870@yahoo.com

Dwijesh Dutta Majumder
 Professor Emeritus, ISI
 Emeritus Scientist, CSIR
 Secretary & Director, ICSIT
 Kolkata, India
 ddm@isical.ac.in

Abstract- This paper compares the Processing delay of our own data fusion architecture called MLFFA (Multi-level Federated Filter Architecture) with other architectures available in the literature as Centralised, Cascaded, Federated, and Distributed in order to improve the filtration process of signaling sensors for a reference sensor (RS) within its fusion domain. The processing delay analysis of a signaling sensor is done by calculating the difference between the time of receipt of a packet by a sensor and getting it processed by that particular sensor from its own fusion queue. This piece of writing shows improvement with respect to the existing architectures in term of processing delay.

Keywords- Master Fusion Filter, Reference Sensor, Local Filter, Centralised, Cascaded, Federated, Distributed MLFFA, Queue and Processing Delay.

I. INTRODUCTION

Data fusion is a methodology which combines the data from different sensors sources to cooperate each other to take the total effect in higher level than the independent sum of effects. Data fusion helps in finding out the state of a system from correlated data originating from different sources. In the recent years, the data fusion had gone through rapid change. The information that needs to be fused can vary greatly based on the architectures it is using to the local filters and master fusion filter. Each sensor is a separate data source that produces records with several fields such as the identification and location of the sensor that generated the reading, a time stamp, the sensor type, and the value of the reading.

Queues are situation where the claim for a service capability surpasses the capacity of that facility for which the service is not delivered immediately upon request but must linger or wait. The various architectures of data fusion can suffer from the problem of data queue in the intermediate sensors used for effective transmission of fused data. As a result the waiting time of the signals ready to be processed can get increased resulting in a delay at various stages of fusion. At the same time, it is true that the diminution of waiting time requires extra investment in the technology. This is why we require a data fusion model and approach to evaluate such circumstances. This is where our model works and the MLFFA have been found to be more

efficient than the available literatures. This paper deals with processing delay of our architecture MLFFA.

II. MULTISENSOR DATA FUSION

The MultiSensor Data Fusion (MSDF) was first coined for military applications including battlefield surveillance, automatic multi-target tracking and recognition [1][2] and then are applied to civil industries [3][4][5] including guidance and control of autonomous vehicles and robotic systems. MSDF is also applied to estimation and identification theory, control engineering, statistics and decision theory, signal processing and pattern recognition, artificial intelligence and knowledge engineering. So the multidisciplinary nature of MSDF can be applied to wide range of applications. MultiSensor network consist of very large number of devices, each of which is a data source, measuring some quantity of the object's location and requiring some data fusion model for their function.

III. CHARACTERISTICS OF DATA

The general characteristics of the data can be one-dimensional (signal), two-dimensional (image) or three-dimensional (robotics). The quality of the data to be processed and its heterogeneity are often more important. This causes a further level of difficulty, which has to be counted not only in the modeling but also in the algorithms. The data is mostly objective (provided by sensors), which differentiates them from subjective data such as what can be provided by individuals. However, they preserve a portion of subjectivity (for example, in the choice of the sensors or the sources of information, or also of the acquisition parameters). There is also some subjectivity in how the objectives are expressed. The quality of objective data is usually tarnished either because of flaws in the acquisition systems or because of the processes to which it is subjected.

A first characteristic involves the type of information we wish to fuse. It can consist of direct observations, results obtained after processing these observations, more generic knowledge expressed in the form of rules or opinions of experts.

Information can also be either factual or generic. Factual information is directly related to the observations of phenomenon. Generic information can be a replica of the observed fact, common rules and integrity constraints.

Generic information is more general and serves as a “default” and is considered if the specific information is not available or reliable. Again, information can be static or dynamic for which there are different ways of modelling and describing it. In a fusion process, first important point consists of the elements of information we wish to fuse together and secondly the extra information used to direct or help the combination. Imperfection is one of the essential qualities of information in fusion which is always there (otherwise fusion would not be necessary) and can take different forms.

IV. DATA FUSION FILTER

The purpose of data fusion is to produce an enhanced model or estimation of a system from a set of independent data sources. There are various multisensor data fusion approaches, of which Kalman filtering is one of the most noteworthy. The Kalman filter was developed for applications in aircraft navigation, control and guidance since the 1970s. After that, various filter architectures and filtering algorithms based on Kalman were proposed for different data fusion methods to fuse multiple navigation sensor data to achieve the desired performance. The main job of a sensors network is to provide information about a process variable in the situation by taking measurements, and because these measurements are noisy and are taken at discrete points in time, it is necessary to fuse multiple measurements to rebuild the parameter of interest. In general, given an observation vector corresponding to time, we want to guess a process state vector. This is where the Filter works. The actual state of a process entity shall be estimated by using actual measurement from different sensors and may even be compared with information gained from previous measurements. Usually, filtering is performed in real time. The filter architectures can be broadly classified as: Centralised, Cascaded, Federated and Distributed.

A. Centralised Filter Architecture

The Centralised Filter architecture, measures the data from all navigation sensors and are processed in a central data fusion filter to obtain the accurate estimates of the navigation states. [6][7][8].With the increasing number of sensor systems, the algorithm used in the filter can be very complicated and its computation can be very time-consuming one. In order to overcome this problem, the other filter architectures have been proposed.

B. Cascaded Filter Architecture

In Cascaded Filter architecture, the output of one filter is used as input to a later filter. The filter outputs include the estimates of the states of the system and their error covariance. [9][10][11][12]

C. Federated Filter Architecture

The Federated Filter architecture is a two-stage filtering architecture. Here all parallel local filters combine their own sensor system with a common reference system for data propagation with estimation of the local system states. [13][14]. Then these local estimates are fused in master filter to get global estimates. All parallel filters have a common state vector as they all share common reference system. Comparing federated and centralized filters will show that federated architecture provides much

enhancement in detection of failures, isolation and recovery and fault tolerance over centralized filter. It is used in many multi-sensor navigation systems.

D. Distributed Filter Architecture:

No standard model exists for Distributed Filter architecture. From the viewpoint of information use, there are two main data fusion approaches to the design of distributed filters - Measurement Fusion and State Fusion. [15][16]. In case of state fusion, the local filter estimates the local states which are then fused in a central filter to get global estimates. In measurement fusion, different subsets of all the sensor measurements are fused by means of a pool of Kalman filters to get multiple state estimation versions of the global system states. These are then compared or weighted to attain the more correct global state estimation and to detect sensor or system failures. In a fully distributed multisensor data fusion system, there may be no central data fusion. In the design of multisensor navigation systems, the distributed filter architecture presents the most flexible method.

V. QUEUING, QUEUING THEORY, QUEUING MODEL

Queue is a line of people or things waiting to be executed, usually in sequential order commencing at the beginning of the line or sequence. In computer technology, a queue is a series of work objects that are waiting to be processed. The possible aspects, arrangements, and processes related to queues is known as queuing theory. Thus, Queuing theory [17] is the mathematical study of queues and deals with problems which involve queuing (or waiting).. The theory enables mathematical analysis of several related processes including arriving at the (back of the) queue, waiting in the queue (essentially a storage process), and being served at the front of the queue. The theory permits the origin and calculation of several performance measures including the average waiting time in the queue or the system, the expected number waiting or receiving service, and the probability of encountering the system in certain states, such as empty, full, having an available server or having to wait a certain time to be served.

A queuing model is used to estimate a real queuing situation or system, so the queuing behavior can be investigated mathematically. Queuing models permits a number of practical steady state performance measures to be determined, including: the average number in the queue, or the system, the average time spent in the queue, or the system, the statistical distribution of those numbers or times, the probability the queue is full, or empty, and the probability of finding the system in a particular state.

A. Components of a Queuing System

A queuing system [18] is characterized by three components: Arrival process, Service mechanism and Queue discipline.

1) Arrival Process

Arrivals may initiate from one or several sources referred to as the calling population. The calling population can be limited or 'unlimited'. Request Arrival Rate (a)-

2nd Annual International Conference IEMCON 2012

Service requests arrive according to one of four patterns: steady, irregular, regular, or random. Service Distribution Rate (s) - The mean number of requests processed within a time period. Utilization (u) - It is the traffic intensity which means the request arrival rate divided by the service rate.

2) Service Mechanism

The service mechanism of a queuing system is specified by the number of servers (denoted by s), each server having its own queue or a common queue and the probability distribution of customer's service time

3) Queue Discipline

Discipline of a queuing system means the rule that a server uses to choose the next customer from the queue (if any) when the server completes the service of the current customer. Commonly used queue disciplines are: FIFO - Customers are served on a first-in first-out basis. LIFO - Customers are served in a last-in first-out manner. SIRO : Service-In-Random-Order, etc Priority - Customers are served in order of their importance on the basis of their service requirements.

B. Communication Delays

We know that when a system gets congested, the service delay in the system increases. A good understanding of the association between congestion and delay is essential for scheming effective congestion control algorithms. There are various components of delay [19] in a messaging system. The total delay experienced by messages can be classified into the following categories:

1) Processing Delay

This is the delay between the times of receipt of a packet for transmission to the point of putting it into the transmission queue. On the receiving end, it is the delay between the time of reception of a packet in the receive queue to the point of actual processing of the message. This delay depends on the Data Filter (in our case) speed and Data Filter load in the system.

2) Queuing Delay

This is the delay between the point of entry of a packet in the transmit queue to the actual point of transmission of the message. This delay depends on the load on the communication link.

3) Transmission Delay

This is the delay between the transmissions of first bit of the packet to the transmission of the last bit. This delay depends on the speed of the communication link.

4) Propagation Delay

This is the delay between the points of transmission of the last bit of the packet to the point of reception of last bit of the packet at the other end. This delay depends on the physical characteristics of the communication link.

5) Retransmission Delay

This is the delay that results when a packet is lost and has to be retransmitted. This delay depends on the error rate on the link and the protocol used for retransmissions.

C. Little's Theorem

Little's theorem states that: The average number of customers (N) can be determined from the following equation:

$$N = \lambda T$$

Here λ is the average customer arrival rate and T is the average service time for a customer. The most important characteristics of a queuing system are explained below:

1) Arrival Process

The probability density distribution that determines the customer arrivals in the system. In a messaging system, this refers to the message arrival probability distribution.

2) Service Process

The probability density distribution that determines the customer service times in the system. In a messaging system, this refers to the message transmission time distribution. Since message transmission is directly proportional to the length of the message, this parameter indirectly refers to the message length distribution.

3) Number of Servers

Number of servers available to service the customers. In a messaging system, this refers to the number of links between the source and destination nodes.

D. Queuing notation

It is common to use to use the symbols [20]: λ to be the mean (or average) number of arrivals per time period, i.e. the mean arrival rate μ to be the mean (or average) number of customers served per time period, i.e. the mean service rate. There is a standard notation system to classify queuing systems as A/B/C/D/E, where:

A represents the probability distribution for the arrival process

B represents the probability distribution for the service process

C represents the number of channels (servers)

D represents the maximum number of customers allowed in the queuing system (either being served or waiting for service)

E represents the maximum number of customers in total

In the Poisson probability distribution, the observer records the number of events that occur in a time interval of fixed length. In the (negative) exponential probability distribution, the observer records the length of the time interval between consecutive events

Common options for A and B are:

M for a Poisson arrival distribution or a exponential service time distribution

D for a deterministic or constant value

G for a general distribution (but with a known mean and variance)

If D and E are not specified then it is assumed that they are infinite.

Examples of queuing systems that can be defined with this convention are:

For example the M/M/1 queuing system, the simplest queuing system, has a Poisson arrival distribution, an exponential service time distribution and a single channel (one server).

M/D/n: Here the arrival process is Poisson and the service time distribution is deterministic. The system has n servers. (e.g. a ticket booking counter with n cashiers.) Here the service time can be assumed to be same for all customers)

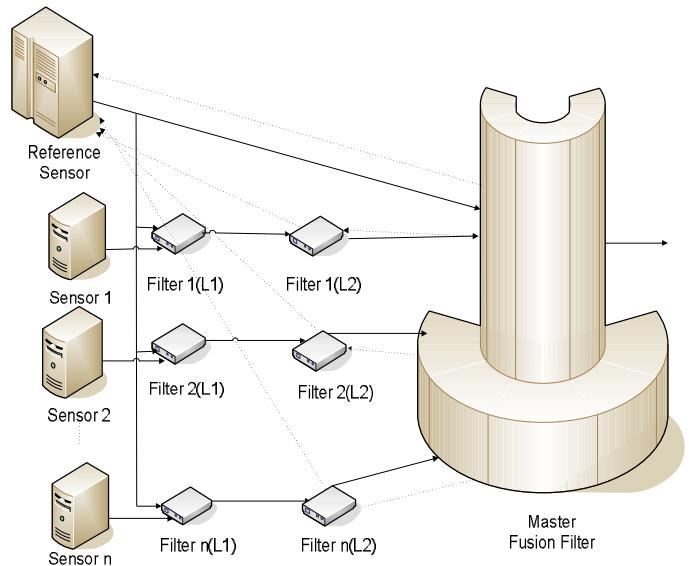
G/G/n: This is the most general queuing system where the arrival and service time processes are both arbitrary. The system has n servers. No analytical solution is known for this queuing system.

VI. MULTI-LEVEL FEDERATED FILTER ARCHITECTURE (MLFFA)

Our Multi-level Federated Local filter scheme [21] (Fig 1) having adjacent and different levels can ease out the burden from MFF. It has got multiple levels of filter for each sensor. This multiple levels have been done for more refinement of the collected data from the sensors. Since different filtering algorithms are executed at different levels of filter, fine-tuning of the data increases. Moreover, it has a direct communication with the reference filter to compare the fused data. At each level, all parallel filters merge their individual sensor data with a common reference data and calculates local system states. Then these local estimates are sent to the parallel filters of the next level which repeats the same process as the parallel filters in the previous level and it continues. After crossing all the levels, the local estimates from the last level are finally fused in master filter to get global estimates. Parallel filters have a common state vector as they all share common reference system. Now, when the signals are arriving to the filters at various levels, they are to be processed and transmitted.

A filter can process only one signal at a time. If signals arrive faster than the filter can process them (such as in a burst transmission) the filter puts them into the queue (also called the buffer) until it can get around to transmitting them. The maximum queuing delay is proportional to buffer size. The longer the line of signal waiting to be transmitted, the longer the average waiting time is. However, this is much preferable to a shorter buffer, which would result in ignored ("dropped") signals, which in turn would result in much longer overall transmission times. During network congestion, queuing delays can be considered infinite when the signal is dropped. The retransmission of such signals causes significant overall delay because all forms of delay will be incurred more than once. If the network congestion continues, the packet may be dropped many times.

FIGURE 1: MLFFA



VII. DELAY IN PROCESSING TIME

We deduced the mathematical model to calculate processing delay for the data fusion using the Queueing Model and applied this to our own architecture named MLFFA. The model calculates the amount of time taken by a local filter for the initiation of a signal with respect to its time of entry into the filter.

The processing time is based on clock pulse for multi-fusion sensor to change state. Let the processing time be T_f . Therefore

$$T_f = n * d * \rho * \beta * \log(t_i + \delta) * (1 - \alpha) + \log(t_o + \delta) * (1 - z) \text{ where}$$

n = Number of elements

d = Distance between two filters

ρ = Filter density

β = Average time complexity for 1d, 2d, 3d data

t_i = Clock pulse generated by incoming data stream

δ = Time delay for the queue process

α = Ratio of active to inactive filter operations

t_o = Clock pulse for the out going data stream

z = Cache hit to miss ratio

VIII. PERFORMANCE ANALYSIS

We have formulated all the four architectures and our new Multi-level Federated filter architectures to bring up the

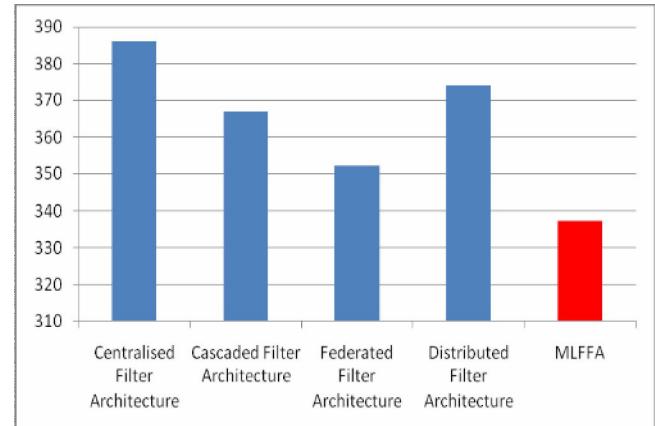
comparison between them. The values of the standard parameters used here are from the literature [22].

| Parameters | Value |
|---|-----------------------------|
| Bias uncertainty (o/h) | 10-40 |
| Scale factor stability (ppm) | 100-500 |
| Alignment (arcs) | 200 |
| Random noise (o/h/ $\sqrt{\text{Hz}}$) | 1-5 |
| Flattening (f) | 1/298.257223563 |
| R_w | 8 |
| R_c | 0.5 |
| R_l | 0.5 |
| d_{LIL2} | 16 |
| d_{LILn} | \sqrt{n} |
| V | 28.9 m/s |
| P | 0.0002 users/m ² |
| A | 5 |
| λ_a | 0.0008/s |
| λ_d | 0.0008/s |
| ρ | 1 |
| β | 0.001 |
| δ | 0.005 |
| α | 1/5 |
| z | 1/10 |
| γ_a | 1 |
| β_a | Log n |

Considering the output from the above formulation, our architecture has been proved to be more effective than the other classical architectures. The following figure shows the processing delay for a single filter in different architectures – Centralised, Cascaded, Federated, Distributed Filter architectures and our MLFFA.

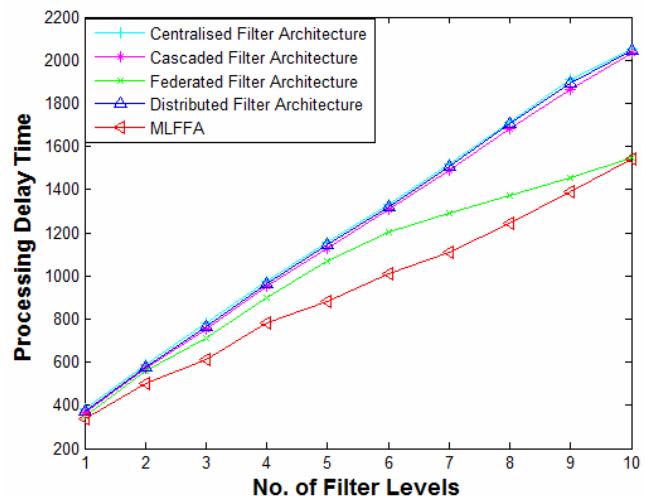
Considering the number of LF levels as 1, our formulation yields 386, 367, 352, 374 and 337 ms respectively for Centralised, Cascaded, Federated, Distributed and Multi-level Federated filter architectures as shown in the figure below.

FIGURE 2: PROCESSING DELAY FOR SINGLE FILTERS ACROSS ARCHITECTURE



Considering the number of LF levels as 1 to 10, here follows the pictorial representation of our simulation result using the MATLAB software:

FIGURE 3: COMPARISON OF DELIVERY TIME DELAY



IX. CONCLUSION

We designed a mathematical model based on queuing theory for reducing processing delay of data fusion with the help of our architecture called Multi-level Federated Filter Architecture. It is significant to prevent processing delay so that signals in the signalling sensors waiting at the queue does not get bottlenecked as voice, video, and data traffic are all delivered through the same sensor network. Considering economic restraint and limited sensor ability, the processing time of a sensor is typically a challenging factor, where each sensor has to relay the traffic from one sensor to another. Although all architecture has its own merits and demerits, considering performance evaluation based on Processing delay our new MLFFA has been proposed to be more efficient compared to other architectures.

2nd Annual International Conference IEMCON 2012

X. REFERENCES

- [1] Blackman, S.S., Multiple Targets Tracking with Radar Applications. Artech House Inc. 1986
- [2] Hall, D. L. and Llinas, J., An Introduction to Multisensor Data Fusion, Proceedings of the IEEE Vol. 85, No. 1, Jan. 1997, 6 -23
- [3] <http://www.data-fusion.org/article.php>
- [4] Luo, R. C. and Kay, M. G., Multisensor Integration and Fusion in Intelligent System, IEEE Trans. On System, Man, and Cybernetics, Vol. 19, No. 5, Sept/Oct 1989, 901-931.
- [5] Kokar, M. and Kim, K., Review of Multisensor Data Fusion: Architecture and Techniques, Proceedings of The International Symposium on Intelligent Control, Chicago, Illinois, USA, Aug. 1993, 261-266
- [6] Diesel, J. W., Integration of GPS/INS for Maximum Velocity Accuracy, Journal of The Institute of Navigation, Vol. 34, No. 3, Fall 1987.
- [7] Diesel, J.W. GPS/INS Integration for Civil Aviation, Proceedings of National Telesystems Conference, 1991, 223 -228
- [8] Hyslop, G.; Gerth, D.; Kraemer, J. GPS/INS integration on the Standoff Land Attack Missile (SLAM), IEEE Aerospace and Electronics Systems Magazine, Vol. 5, No. 7, July 1990, 29 -34.
- [9] Schlee, F.H.; Toda, N.F.; Islam, M.A. and Standish, C.J., Use of an external cascade Kalman filter to improve the performance of a Global Positioning System (GPS) inertial navigator, Proceedings of the IEEE NAECON 1988, 345-350.
- [10] Wade, M. and Grewal, M. S., Analysis of A Cascaded INS Calibration Filter, Proceedings of the IEEE 1988, 366-373.
- [11] Karatsinides, S. E., Enhancing Filter Robustness in Cascaded GPS-INS Integrations, IEEE Trans. On Aerospace and Electronic Systems, Vol. 30, No. 4, Oct., 1994, 1001-1008.
- [12] Bell, W.B; Gorre, R.G; and Cockrell, L.D, Cascading Filtered DTS Data into a Loosely Coupled GPS/INS System, Proceedings of IEEE PLANS'98, 1998, 586-593.
- [13] Carlson, N.A., Federated Filter for Fault-Tolerant Integrated Systems, Proceedings of 1988 IEEE PLANS, 1988, 110-119.
- [14] Felter, S.C., An overview of decentralized Kalman filter techniques, Proceedings of IEEE Southern Tier Technical Conference, 1990, 79-87.
- [15] Bar-Shalom, Y. and Fortmann, T. E., Tracking and Data Association, Academic Press, New York, 1988.
- [16] Liggins, M. E., II; Chee-Yee Chong; Kadar, I.; Alford, M. G.; Vannicola, V. and Thomopoulos, S., Distributed Fusion Architectures and Algorithms for Target Tracking, Proceedings of the IEEE, Vol. 85, No. 1, Jan. 1997, 95-107.
- [17] http://en.wikipedia.org/wiki/Queueing_theory
- [18] <http://www.cs.utexas.edu/~browne/cs380ns2003/Papers/SimpleQueuingModels.pdf>
- [19] http://www.eventhelix.com/realtimemantra/congestioncontrol/queueing_theory.htm
- [20] <http://people.brunel.ac.uk/~mastjjb/jeb/or/queue.html>
- [21] Sharmistha Bhattacharjee, Boudhayan Bhattacharya, Dwijesh Dutta Majumder "MLFFA- A Novel Filter Architecture for Data Fusion Management Scheme" ,NCIE 2011, WBUT, Kolkata, February 17-18
- [22] Spitzer, C.R, The Avionics Handbook, CRC Press LLC, 2001

DESIGN AND IMPLEMENTATION OF REAL TIME AES-128 ON REAL TIME OPERATING SYSTEM FOR MULTIPLE FPGA COMMUNICATION

Rourab Paul¹
 Dept. of Electronic
 Science¹
 Kolkata, India
 rourabpaul@gmail.com
 University of Calcutta

Sangeet Saha²
 Dept. of Computer
 Science and Engineering²
 Kolkata, India
 sangeet.saha87@gmail.c
 om
 University of Calcutta

Suman Sau³
 A. K. Choudhury School
 of Information
 Technology³
 Kolkata, India
 sumansau@gmail.com
 University of Calcutta

Amlan Chakrabarti⁴
 A. K. Choudhury School
 of Information
 Technology⁴
 Kolkata, India
 acakcs@caluniv.ac.in
 University of Calcutta

Abstract- Security is the most important part in data communication system, where more randomization in secret keys increases the security as well as complexity of the cryptography algorithms. As a result in recent dates these algorithms are compensating with enormous memory spaces and large execution time on hardware platform. Field programmable gate arrays (FPGAs), provide one of the major alternative in hardware platform scenario due to its reconfiguration nature, low price and marketing speed. In FPGA based embedded system we can use embedded processor to execute particular algorithm with the inclusion of a real time operating System (RTOS), where threads may reduce resource utilization and time consumption. A process in the runtime is separated in different smaller tasks which are executed by the scheduler to meet the real time dead line using RTOS. In this paper we demonstrate the design and implementation of a 128-bit Advanced Encryption Standard (AES) both symmetric key encryption and decryption algorithm by developing suitable hardware and software design on Xilinx Spartan- 3E (XC3S500E-FG320) device using an Xilkernel RTOS, the implementation has been tested successfully. The system is optimized in terms of execution speed and hardware utilization.

Keywords: - Reconfigurable architecture, RTOS, Real Time Communication, AES, Security, XPS, EDK

I. INTRODUCTION

Data security is an essential objective for the military and diplomatic services which have many commercial uses and applications such as electronic banking, electronic mail, internet network service, messaging networks etc. As an efficient and cost-effective cryptographic algorithm AES [1] algorithm has broad applications, including smart cards and cellular phones, WWW servers and automated teller machines (ATMs). Establishing reliable communication between

multiple FPGA systems/cards is an essential component for developing complex real time systems used for applications like real time data acquisition and processing [2, 3]. Though related work exists for FPGA based AES implementation [4,5,6,7,8,9,10,11,12]. But there exist no implementation based on real time data encryption and decryption over a communication interface between multiple FPGA systems using an RTOS, where each thread is capable of running the encryption and decryption individually.

Our proposed work is an FPGA based design and implementation of the AES-128 algorithm on Real Time Operating System. We have successfully established a secured link between two FPGA systems through RS232 link and have achieved good throughput with a minimum number of resource utilization, compared to the other existing works [13, 14, 15, 16, 17]. The total system functions like a complete system where data are taken from key board in real time by the FPGA, the encryption is done using a thread running on the RTOS. Through the RS232 communication link the cipher data is being transmitted to another FPGA, where the decryption is done by another thread of the RTOS. The decrypted data is visualized on the HyperTerminal for verification. The algorithm and the real time scenario of the system has been analyzed and its hardware utilization proves to be better compared to the related works [13,14,15]. The development platform of our work is Xilinx EDK 11.1 and has an RTOS we have chosen Xilinx owned Xilkernel [18].

The organization of the paper is as follows, Section II describes the design-flow of the AES algorithm for both encryption and decryption process. Section III details out the Real time Operating System on Hardware, Section IV is discussed on proposed hardware architectural design. Section V briefs on the implementation, results and comparison with existing works and the concluding remarks are presented in Section VI.

II DESIGN OVERVIEW OF AES

The algorithm is composed of three main parts: Cipher, Inverse Cipher and Key Expansion. Cipher converts data, commonly known as plaintext, to an unintelligible form called cipher. Key Expansion generates a key schedule that is used in the Cipher and the Inverse Cipher procedure. Cipher and Inverse Cipher are composed of specific number of rounds (Table 1). For the AES algorithm, the number of rounds to be performed during the execution of the algorithm is dependent on the key length [1].

AES operates on a 4×4 array of bytes (referred to as “state”). The algorithm consists of four different simple operations. These operations are:

- Sub Bytes
- Shift Rows
- Mix Columns
- Add Round Key

TABLE I: FEATURES OF AES FOR DIFFERENT KEY LENGTHS

| | Block size N_b words | Key length N_k words | Number of rounds N_r |
|------------------------------|---------------------------|---------------------------|---------------------------|
| AES- 128_bits key | 4 | 4 | 10 |
| AES-192_bits key | 4 | 6 | 12 |
| AES-256_bits key | 4 | 8 | 14 |

A. Encryption Process:

The Encryption and decryption process consists of a number of different transformations applied consecutively over the data block bits, in a fixed number of iterations, called rounds. The number of rounds depends on the length of the key used for the encryption process. For key length of 128 bits, the number of iteration required are 10 ($N_r = 10$). As shown in Fig. 1(a), each of the first N_{r-1} rounds consists of 4 Transformations: *SubBytes()*, *ShiftRows()*, *MixColumns()* and *AddRoundKey()*.

1) *Sub Bytes Transformation*: It is a non-linear substitution of bytes that operates independently on each byte of the state using a substitution table (*S* box). This invertible *S*-box is constructed by first taking the multiplicative inverse in the finite field GF (28) with irreducible polynomial $m(x) = x^8 + x^4 + x^3 + x + 1$. The element {00} is mapped to itself. Then affine transformation is applied (over GF (2)).

2) *Shift Rows Transformation*: Cyclically shifts the rows of the state over different offsets. The operation is almost the same in the decryption process except for the fact that the shifting offsets have different values.

3) *Mix Columns Transformation*: This transformation operates on the state column-by-column, treating each column as a four-term polynomial. The columns are considered as polynomials over GF (28) and multiplied by modulo $x^4 + 1$

with a fixed polynomial $a(x) = \{03\} x^3 + \{01\} x^2 + \{01\} x + \{02\}$.

4) *Add Round Key Transformation*: In this transformation, a Round Key is added to the state by a simple bitwise XOR operation. Each Round Key consists of N_b words from the key expansion. Those N_b words are each added into the columns of the state. Key Addition is the same for the decryption process.

5) *Key Expansion*: Each round key is a 4-word (128-bit) array generated as a product of the previous round key, a constant that changes each round, and a series of *S-Box* lookups for each 32-bit word of the key. The Key schedule Expansion generates a total of $N_b * (N_r + 1)$ words.

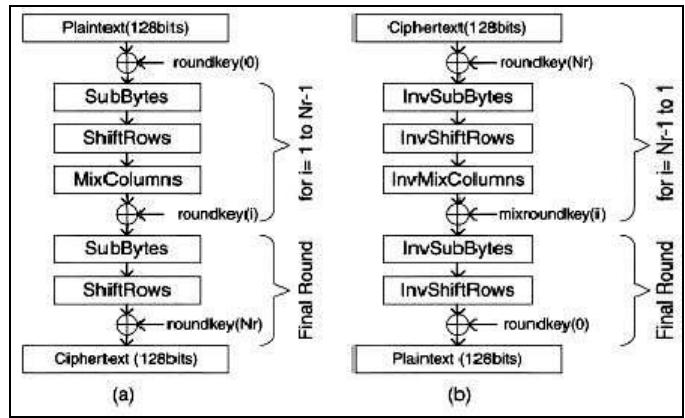


Fig. 1: AES algorithm (a)Encryption structure(b) Equivalent Decryption Structure

B. Decryption Process:

For decryption, the same process occurs simply in reverse order, taking the 128-bit block of cipher text and converting it to plaintext by the application of the inverse of the four operations. *AddRoundKey* is the same for both encryption and decryption. However the three other functions have inverses used in the decryption process: *Inverse SubBytes*, *Inverse ShiftRows*, and *Inverse MixColumns*. This process is direct inverse of the encryption process. All the transformations applied in the encryption process are inversely applied to this process. Hence the last round values of both the data and key are first round inputs for the decryption process and follows in decreasing order as shown Fig. 1(b).

This is the brief description of the AES algorithm for details see the reference [1].

III. REAL TIME OPERATING SYSTEM

REAL – time embedded systems are typically designed for various purposes such as to control or to process data. Characteristics of real-time system include meeting certain deadlines at the right time. To achieve this purpose, real-time operating systems (RTOS) are often used, more specifically an RTOS is a piece of software with a set of APIs for users to develop applications [19] which requires deadline based execution. RTOSes are typically differentiated from generic OSes regarding the following criteria i.e. preemptive or

priority-based scheduling, predictability in task synchronization, deterministic behaviors [19]. There are various RTOSes available for microcontroller as well as for FPGA based design i.e. VxWorks, QNX, eCos, LynxOS, and RTLinux [20].

Here we have chosen Xilkernel [18] as our RTOS, which is provided by Xilinx. Xilkernel is a small, robust, and modular kernel. It is highly integrated with the Platform Studio Frame [21]. The advantages for the usage of Xilkerenel as RTOS here, are

- Xilkernel has very low memory footprint, it uses 7-16 kb of BRAM in a multi threaded program [18], which is much smaller than the RTOS used in microcontroller [10].
- In the context of Xilkernel, a **thread** is the unit of execution and is analogous to a process. Threads are coded like functions.
- Xilkernel is structured as a library. The user application source files must link with Xilkernel to access Xilkernel functionality.
- The MicroBlaze kernel requires an external timer to generate periodic interrupts and this is the only hardware requirement that the kernel places.
- If an interrupt controller is present in the system, Xilkernel exports an interrupt handler registering mechanism and invokes the handlers after it pre-processes each hardware interrupt.[18]

Below we shown an outline of the code to implement the Xilkernel on RTOS

```
int main(void)
{
    xilkernel_main(); /* Start the kernel */

    /* Control does not reach here */

}

void* main_thread(void) /* Statically created first thread */
{
    //Here the AES encryption and decryption runs in each board
    // Child thread can be created by calling creation routine }

Void* child thread()
{
    // again another thread can be created
}
```

As soon as the Xilkernel library is being called, the scheduler takes the responsibility to execute the thread in a sequential manner depending on type of scheduling imposed, control doesn't return back to main function for specific function call.

As far as the scheduling among the threads is concerned it can be either of the two types that Xilkernel supports, one is Round robin type and another of priority based [18].

IV. HARDWARE ARCHITECTURAL DESIGN

The proposed work is implemented using the Xilinx EDK 11.1 (version) and Xilinx Spartan 3E FPGA prototyping board has been used for the hardware implementation and testing. Using the Xilinx platform studio from EDK (Embedded Development Kit) the hardware portion of the embedded system has been developed. A soft core 32-bit RISC processor Micro Blaze has been used as a CPU for this embedded computing unit and all the required soft core peripherals are UART 1 (used for RS232 DCE (Data Circuit-Terminal Equipment) port), UART 2 (used for RS232 DTE (Data Terminal Equipment) port). The blocks used to build up the FPGA based embedded computing unit is shown in Fig. 2.

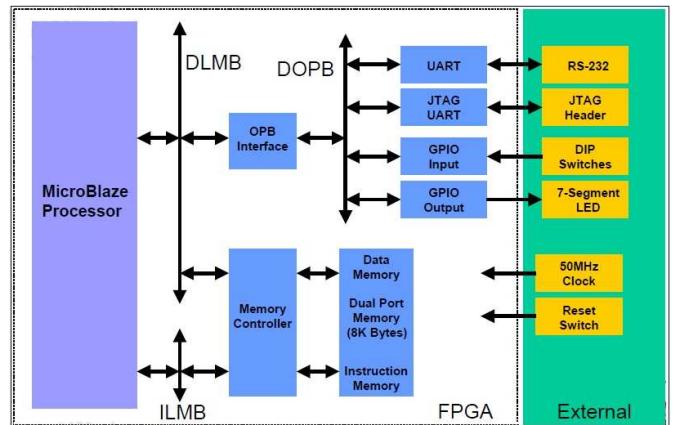


Fig. 2: Internal Architecture Block of FPGA Systems

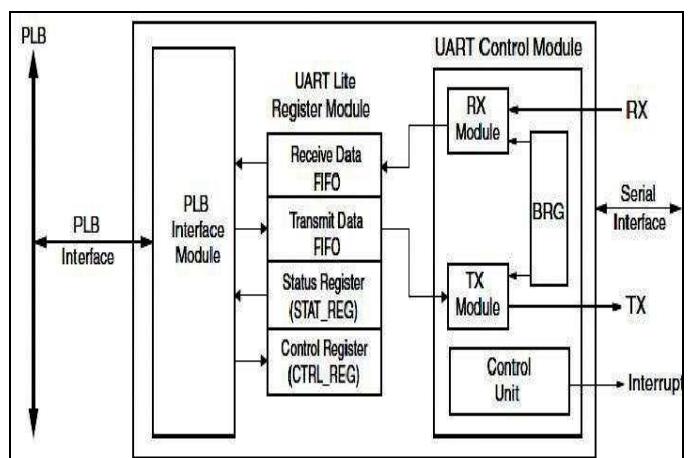


Fig. 3: Universal Asynchronous Receiver Transmitter (UART) system module attached with the Processor Local Bus

V. SERIAL COMMUNICATION

The component that is used for establishing the serial communication between the multiple FPGA systems is the UART (Universal Asynchronous Receiver Transmitter). The Block diagram of the UART is shown in Fig. 3. Here “BRG” stands for “Baud Rate Generator” which controls the speed of the data communication in RS232 channel. Both receiver and sender side must work in the same band ratio otherwise data will be lost. BRG controls the received data store initially at received FIFO and the *transmit data FIFO* transfer the data through the transmitter Module (TX Module). Status and Control registers are used to check status of the FIFO whether it is full or empty and the control of RX (Receiver) and TX modules respectively.

VI IMPLEMENTATION

The proposed architecture was synthesized using Xilinx ISE 11.1 [22] and was implemented on XC3S500e Spartan 3E FPGA Board [22]. The necessary software for this design is written using the feature-rich C/C++ code editor and compilation environment provided within the SDK (Xilinx Software Development Kit). The SDK provides an environment for creating software platforms and applications targeted for Xilinx embedded processor (MicroBlaze). SDK works with hardware designs created with Xilinx Platform Studio (XPS) [22]. We have tested the real time execution of the program in the hardware in RTOS environment. The Fig. 4 gives the work flow of the RTOS.

Xilkernel (RTOS) can be configured by the software platform settings of the EDK. At the time of configuration we can also set the scheduling method, here we chose Round-Robin scheduling with specific time slice. The kernel starts with the execution of a static thread. The real time data are taken from the key board into the board 1 through RS232 port where the AES encryption process is going on. The encrypted plain text called cipher text is being sent to the board 2 using RS232 port. After receiving the cipher text, board 2 decrypts the encrypted text and converts to the plaintext again. This plaintext is being sent to the hyper terminal of a PC also connected with the board 2, through the RS232 interface. Fig. 5 shows the architectural picture of the proposed work.

A. Results

In our test case we have taken 16 byte data by the key board
 $\text{Input} = [36\ 46\ e6\ a8\ 88\ 5a\ 30\ 8c\ 28\ 31\ 98\ a2\ e0\ 37\ 07\ 34]$.
 $\text{Cipher Key} = [2b\ 7e\ 15\ 16\ 28\ ae\ d2\ a6\ ab\ f7\ 15\ 88\ 09\ cf\ 4f\ 3c]$.

Then after ten rounds of the AES the cipher text will appear as

$\text{Cipher text} = [27\ 37\ c1\ 82\ 83\ 29\ a4\ f1\ 43\ 93\ 9c\ 5e\ a6\ b0\ 7c\ e1]$ as shown in Fig. 7. For decryption we use cipher text as input and use the same cipher key for decryption algorithm and find original data= $[36\ 46\ e6\ a8\ 88\ 5a\ 30\ 8c\ 28\ 31\ 98\ a2\ e0\ 37\ 07\ 34]$.

This decrypted data is sent to the PC through an RS232 port and been verified using the hyper terminal interface (Fig. 8). Table II shows the resource utilization of the FPGA in our

design and Table III shows the execution time of our algorithm and the throughput of encryption and decryption from the equation given below:

$$\text{Throughput} = (\text{bits processed for encryption or decryption}) / \text{second}$$

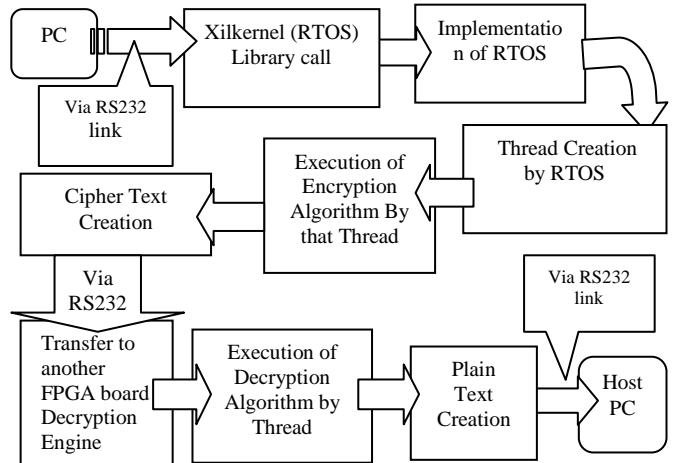


Fig. 4: Work flow using RTOS

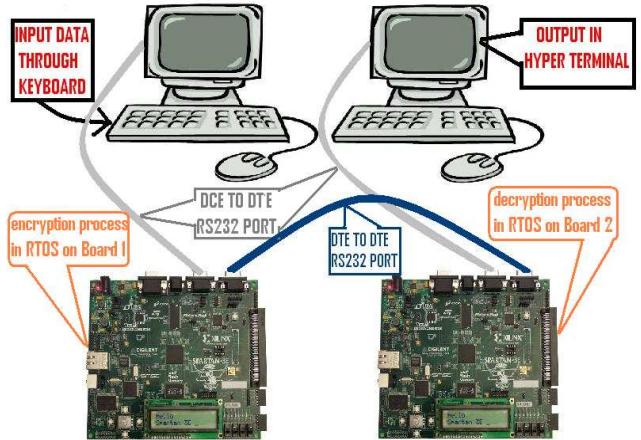


Fig. 5: Architectural Picture of the Experiment.

B. Comparison with existing works

Table IV, shows the comparison of our work with the existing work and we can claim that our design proposes an efficient solution for the implementation of AES using the FPGA devices.

TABLE II: RESOURCES UTILIZATION

| Logic Utilization | Used | Available | Utilization |
|--|-------|-----------|-------------|
| Number of Slice Flip Flops | 2,621 | 9,312 | 28% |
| Number of 4 input LUTs | 2,871 | 9,312 | 30% |
| Number of occupied Slices | 2,495 | 4,656 | 53% |
| Number of Slices containing only related logic | 2,495 | 2,495 | 100% |
| Number of Slices containing unrelated logic | 0 | 2,495 | 0% |

2nd Annual International Conference IEMCON 2012

TABLE III: ENCRYPTION AND DECRYPTION TIME

| Process | Clock frequency (MHz) | Time (ms) | Throughput (byte/sec) |
|------------|-----------------------|-----------|-----------------------|
| Encryption | 50 | 4.0274 | 3972.2 |
| Decryption | 50 | 4.1524 | 30825.1 |

TABLE IV: COMPARISION WITH EXSISTING WORKS

| Design | Device | Frequency MHz | Slices | BRAMS |
|------------------------|---------------------------|---------------|--------------|------------|
| Elbirt et al[13] | XCV1000-4 | 31.8 | 10992 | 0 |
| M. McLoone et al[14] | XCV812e-8 | 93.9 | 2000 | 244 |
| K.U.Jarvinen et al[15] | XCV1000e-8 | 129.2 | 11719 | 0 |
| G.P.Sagges [16] | XCV2000e-8 | 158 | 5810 | 0 |
| F. Standaert [17] | XCV3200e-8 | 154 | 15112 | 0 |
| Proposed | Spartan3e Xc3s500e | 50 | 2,495 | 320 |

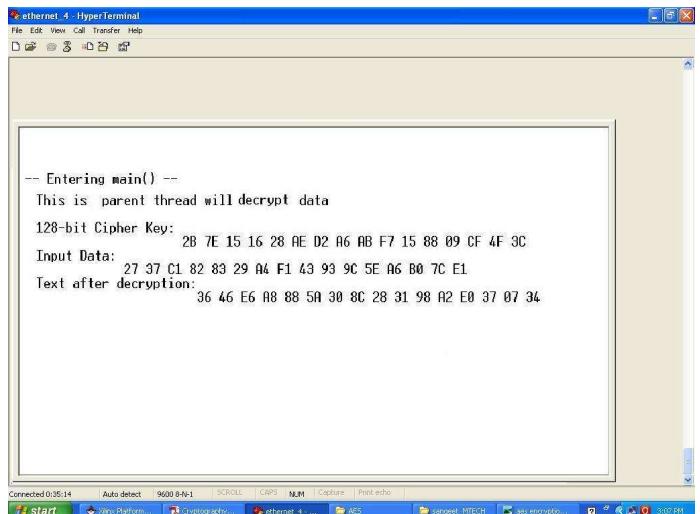


Fig. 8: Decrypted text on Hyper Terminal

VI. CONCLUSION

The aim of this proposed design is to perform a real time data communication exhibiting a significant level of security and providing a faster processing time where necessary. The bit length of the key used in our experiment is 128 bit and the result is obtained successfully. The key width can be varied with a little modification in the algorithm. Though we have performed the real time encryption and decryption of data for RS232 serial communication the technique remains the same for Ethernet data communication using the EMAC core. In future we will try to perform the encryption and decryption of data where inputs will be audio, image, video data coming from different multimedia applications performed over FPGA. Usage of single FPGA with dual processor implementation, where one processor will execute the algorithm while other one will be responsible for input data acquisition, so that the executing processor can handle the algorithm without any interruption, will also be a good step in the world of hardware design.

REFERENCES:

- [1] "Advanced Encryption Standard (AES)" Federal Information Processing Standards Publication 197, Nov. 2001
- [2] http://www.xilinx.com/appnotes/FPGA_NSREC98.pdf.
- [3] W. Kühn et. al.,FPGA based Compute Nodes for High Level Triggering in PANDA,Journal of Physics: Conference Series 119 (2008) 022027.
- [4] James H. Wiebe,AES-128 implementation on a Virtex-4 FPGA Proc. 2007 IEEE International Symposium on Signal Processing and Information Technology.
- [5] B.Singh,H.Kaur2, and H.Monga FPGA Implementation of AES Co-processor in Counter Mode proc pp. 491–496, 2010. © Springer-Verlag Berlin Heidelberg 2010
- [6]"Analysis of AES Hardware Implementations" Song J. Park Department of Electrical & Computer Engineering Oregon State University, Corvallis,
- [7] "Efficient Hardware Design and Implementation of AES Cryptosystem" PRAVIN B. GHEWARI, MRS. JAYMALA, K. PATIL AMIT B.

Fig 7: Cipher text on Hyper Terminal

2nd Annual International Conference IEMCON 2012

CHOUGULE, International Journal of Engineering Science and Technology
Vol. 2(3), 2010, 213-219.

[8] "AES on FPGA from the Fastest to the Smallest" Tim Good and Mohammed Benaissa.

[9] "HARDWARE IMPLEMENTATION OF THE IMPROVED WEP AND RC4 ENCRYPTION ALGORITHMS FOR WIRELESS TERMINALS" Panu Hämäläinen, Marko Hännikäinen, Timo Hämäläinen, and Jukka Saarinen.

[10] "Architectures and VLSI Implementations of the AES-Proposal" Rijndael. N. Sklavos and O. Koufopavlou, IEEE TRANSACTIONS ON COMPUTERS, VOL. 51, NO. 12, DECEMBER 2002.

[11] "Utilizing hard cores of modern FPGA devices for high-performance cryptography" Tim Güneysu. J Cryptogr Eng (2011) 1:37–55 DOI 10.1007/s13389-011-0002-2.

[12] A. Mazzeo, L. Romano, G. P. Saggese and N. Mazzocca. 2003. FPGABased Implementation of a Serial RSA Processor. Design. Proceedings of the conference on Design, Automation and Test in Europe - Volume 1. ISBN:0-7695-1870-2 .

[13] A. J. Elbirt, W. Yip, B. Chetwynd, and C. Paar. An FPGA implementation and performance evaluation of the AES block cipher candidate algorithm finalist. presented at Proc. 3rd AES Conf. (AES3). [Online].Available:<http://csrc.nist.gov/encryption/aes/round2/conf3/aes3paper.s.html>.

[14] M. McLoone and J. V. McCanny, "Rijndael FPGA implementation utilizing look-up tables," in IEEEWorkshop on Signal Processing Systems, Sept. 2001, pp. 349–360.

[15] K. U. Jarvinen, M. T. Tommiska, and J. O. Skytta, "A fully pipelined memoryless 17.8 Gbps AES-128 encryptor," in Proc. Int. Symp. Field-Programmable Gate Arrays (FPGA 2003), Monterey, CA, Feb. 2003, pp. 207–215.

[16] G. P. Saggese, A. Mazzeo, N. Mazzocca, and A. G. M. Strollo, "An FPGA based performance analysis of the unrolling, tiling and pipelining of the AES algorithm," in Proc. FPL 2003, Portugal, Sept. 2003.

[17] F. Standaert, G. Rovroy, J. Quisquater, and J. Legat, "Efficient implementation of Rijndael encryption in reconfigurable hardware: Improvements & design tradeoffs," in Proc. CHES 2003, Cologne, Germany, Sept. 2003.

[18] xilkernel_v3.00.pdf on www.xilinx.com.

[19] Tran Nguyen Bao Anh*†, Su-Lim Tan†"Survey and performance evaluation of real-time operating systems (RTOS) for small microcontrollers.

[20] "Operating system for Xilinx embedded processor" at <http://www.em.avnet.com>.

[21] Sarat Yoowattana, Chinnapat Nantajiwakornchai, Manas Sangworasil "A Design of Embedded DMX512 Controller using FPGA and XILKernel" ,2009 IEEE Symposium on Industrial Electronics and Applications (ISIEA 2009), October 4-6, 2009, Kuala Lumpur, Malaysia.

[22] <http://www.xilinx.com>.

Microstrip Antenna For Hispeed WLANs

Priyanarayan Misra

Dept. of Electronics and Communication
EASTERN ACADEMY OF SC. & TECH
Bhubaneswar, India
pnmels@gmail.com

Debasis Mishra

Dept. of Electronics and Communication
VSSUT
Burla, India
debasis_uce@rediffmail.com

Abstract— This paper presents a wideband microstrip antenna for high speed WLANs operating in the 5–6 GHz range. A simple rectangular patch antenna is designed at frequency 5.5 GHz and matched at impedance with 50 ohms of transmission line. The technique to achieve wide bandwidth to cover both the high speed WLAN frequency bands (5.15 - 5.35 GHz and 5.725 - 5.80 GHz) is proposed by etches out slot on the ground plane and patch geometry configuration. The configuration of patch and the size of slot on the ground plane are important to enhanced bandwidth, and inset feed is used for good impedance matching in wide frequency range. The EM Simulator CST Microwave Studio is used for the antenna design.

Keywords- Wideband antenna, patch, slot, inset.

I. INTRODUCTION

In recent years the demand for multiband & broad-band antennas has increased for use in multi-frequency & high frequency and high speed data communication. Printed antenna's are economical and can be accommodated in the device package. The microstrip antennas are printed circuits for VHF electronics and microwave which are referred as patch antenna [1]. Simply a microstrip structure consists of a thin metal conductor plate on one side of dielectric substrate, and other side covered with metal conductor completely, ground plane. Basically, the radiating patch may be square or rectangular configuration. The popular microstrip line feed with a conducting strip is easy to fabricate and matching can be achieved by controlling the inset position [2]. Generally, the patch antenna are having very narrow bandwidth, this bandwidth can enhanced by using some technique such as slotting from the patch or/and ground plane. Mostly, the slot on the ground plane is introduced in several shaped for radiating field and adjusted some parameters for good matching. In this paper, the proposed patch antenna comprises a C-shaped wide slot on the ground plane. The vertical and horizontal inset techniques on the patch antenna are also introduced to achieve wideband. The first design of microstrip antenna configuration is simple rectangular patch antenna. Secondly, etch out C-shaped wide slot from the ground plane and patch modification. The designed patch antenna can cover both the high speed WLAN frequency bands (5.15 - 5.35 GHz and 5.725 - 5.80 GHz). The slots are taken on the ground and patch and simulated with EM Simulation Software CST Microwave Studio.

II. PATCH ANTENNA STRUCTURE DESIGN

The first design step is to choose a suitable dielectric substrate of appropriate thickness (h). In this paper for the proposed antenna the most common substrate material FR4 having dielectric constant 4.4 and loss tangent=0.019 is used and the height (h) of the substrate is taken as 1.6 mm. The ground plane ($W_g \times L_g$) of the patch antenna having dimension 38 mm x 25.5 mm and a microstrip feed line of width 3 mm is used.

A. Rectangular Patch Antenna

The FR4 substrate is chosen in order to get acceptable low cost antenna. A rectangular patch with microstrip line feed is designed according to design formulas based on the transmission line model [2], [3]. The size of the rectangular patch antenna can radiate at a frequency 5.5 GHz. The geometry of designed patch antenna at 5.5 GHz is shown in figure 1.

Dimensions of the antenna are given as:

$L_g = 25.5$ mm
 $W_g = 38$ mm
 $L_p = 12.128$ mm
 $W_p = 16.44$ mm
 $y_0 = 4.568$ mm

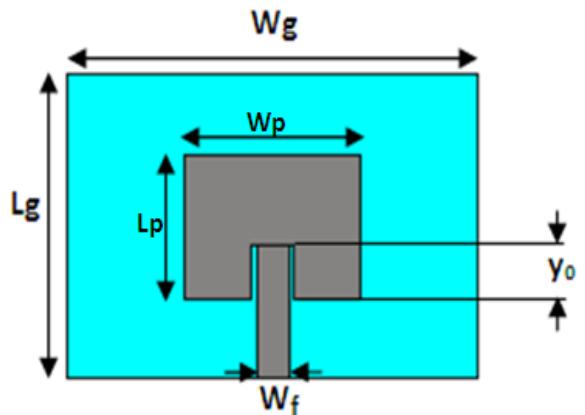


Figure 1. Patch antenna geometry at 5.5 GHz

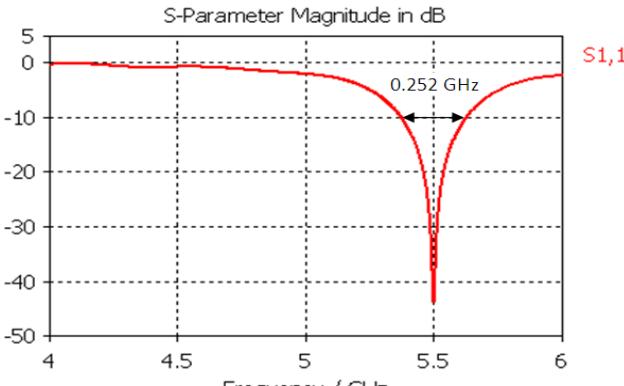


Figure 2. Simulated return loss of the patch antenna

From the simulation result in fig. 2, the plot in terms of return loss shows that the center frequency is at 5.5 GHz with return loss -44 dB. The antenna bandwidth at -10 dB of return loss is found to be from 5.376- 5.628 GHz (BW = 0.252 GHz). This antenna can be reconfigured to achieve wideband by slot loading technique on the ground plane and radiating patch.

B. Ground Plane With Rectangular Slot

A rectangular slot is etched out from the ground plane as shown in figure 3 with the slot dimensions as $LS1=6.5$ mm, $WS1=37$ mm at $LG1=18$ mm. The simulated return loss with the rectangular slot on the ground plane is shown in fig. 4. The rectangular slot on the ground plane results the resonant frequency from 5.5 GHz to 5.81 GHz which meets with the required higher band frequency range 5.72 GHz to 5.80 GHz.

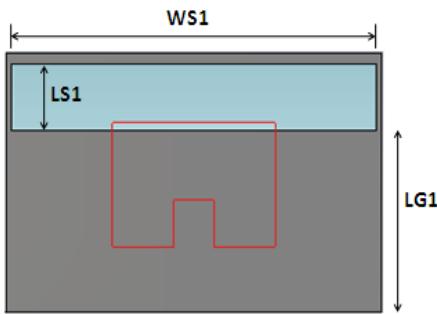


Figure 3. Rectangular slotted ground plane

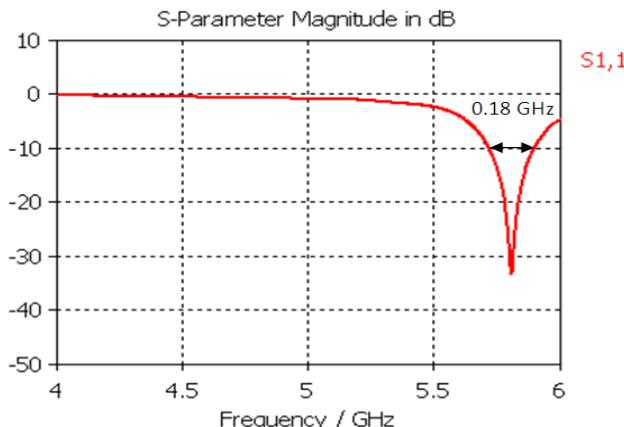
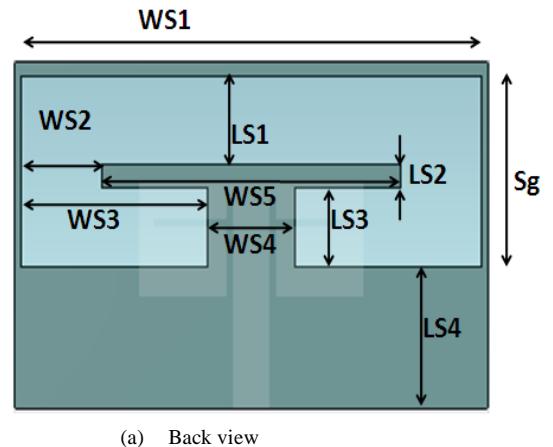


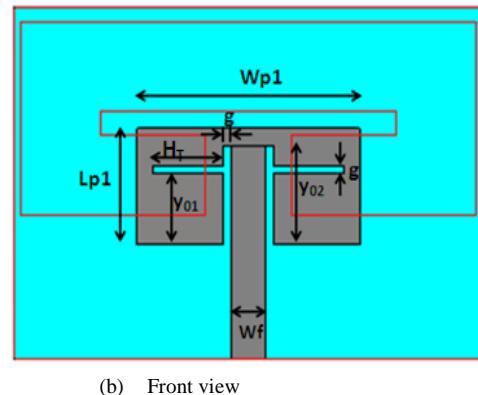
Figure 4. Simulated return loss with rectangular slotted ground plane

C. Ground Plane With C-Slot and Modified Patch

The symmetrical C-shaped wide slot is etched out from the ground plane or the back of patch antenna is shown in fig. 5 (a). Starting from the rectangular slot taken before two more L shaped slots are added again which becomes a C slot on the ground plane. The radiator shape is also modified to match with proper impedance. The vertical inset feed can enhance bandwidth but the first edge frequency is higher than desirable frequency [4], [5]. The horizontal inset on the patch is introduced along with the vertical inset to achieve desired wide bandwidth is shown in fig. 5 (b). The parameter ‘g’ is fixed to 0.5 mm just above the centered horizontal plane but the ‘ H_T ’ parameter is varied from 2 to 6 mm and finally best result is obtained at 5.8 mm from notch edge between patch and feed line with the inset feed distance y_{01} . The impedance matching is achieved by increasing the vertical inset distance upto y_{02} where H_T remains at y_{01} distance from the edge of patch. The return loss at -10 dB is obtained with a wide frequency range from 5.05-5.85 GHz. The simulated return loss graph is shown in fig. 6.



(a) Back view



(b) Front view

Figure 5. C-Slotted ground plane and modified patch

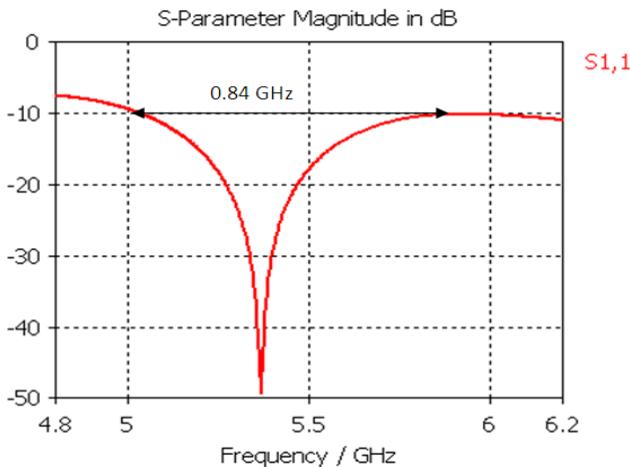


Figure 6. Simulated return loss with C-slotted ground plane and modified patch

The dimensions of C-slot components on the ground plane along with the vertical and horizontal inset distances on the patch are varied along different directions to achieve the best result and fixed at the values which are given in table 1.

TABLE 1. SLOTTED GROUND PLANE AND MODIFIED PATCH PARAMETER DIMENSIONS

| | |
|------------------|---------|
| WS1 | 37 mm |
| WS2 | 6.5 mm |
| WS3 | 15 mm |
| WS4 | 9 mm |
| WS5 | 24 mm |
| LS1 | 6.5 mm |
| LS2 | 1.75 mm |
| LS3 | 5.85 mm |
| LS4 | 10.9 mm |
| Sg=(LS1+LS2+LS3) | 14.1 mm |
| Lp1 | 8.5 mm |
| Wp1 | 18 mm |
| Wf | 3 mm |
| g | 0.5 mm |
| H _T | 5.8 mm |
| y ₀₁ | 3.8 mm |
| y ₀₂ | 6.7 mm |

The final geometry of the proposed antenna with wideband characteristics comprises of a C-slotted ground plane and patch with both vertical and horizontal inset. Also a very good agreement is achieved with impedance matching in wideband frequency range 5.05-5.89 GHz. The bandwidth of the antenna

is increased upto more than three times from the previous antenna

III. RADIATION PATTERNS

The far field radiation patterns in terms of polar plot for the proposed antenna are shown in figures 7, 8 and 9 at frequencies 5.15 GHz, 5.5 GHz and 5.8 GHz respectively.

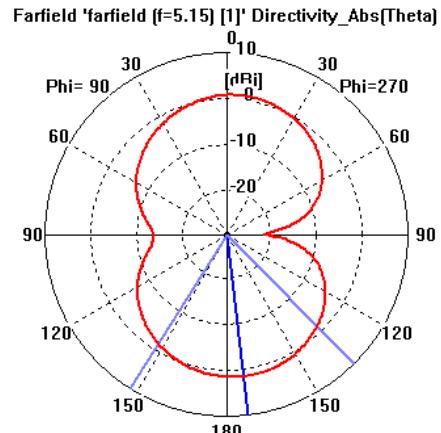


Figure 7. Radiation pattern at 5.15GHz

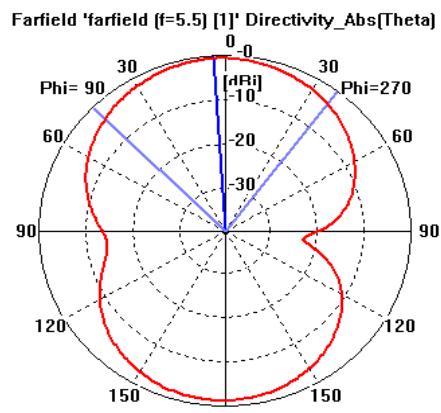


Figure 8. Radiation pattern at 5.5GHz

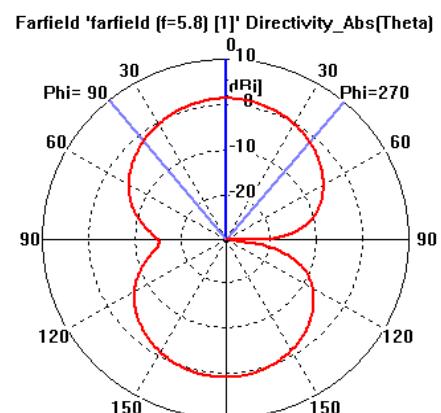


Figure 9. Radiation pattern at 5.8 GHz

2nd Annual International Conference IEMCON 2012

IV. CONCLUSION

The C-shaped wide slot and inset techniques are investigated to achieve wideband operating frequency from 5.05 GHz to 5.85 GHz with good impedance matching. The designed antenna can be suitable for the high speed WLAN frequency bands (5.15-5.35 GHz and 5.725-5.80 GHz).

ACKNOWLEDGMENT

The authors wish to express their appreciation to Prof. R.K. Mishra because the research reported here was initially inspired by him.

REFERENCES

- [1] Fred Gardiol, Microstrip Circuits, A Wiley-Interscience Publication, John Wiley & Sons, Inc., 1994.
- [2] Constantine A. Balanis, Antenna Theory and Design, John Wiley & Sons, Inc., 1997.
- [3] R. Garg, P. Bhartia, I. Bahl, and A. Ittipiboon, Microstrip Antenna Design Handbook, Boston, Artech House, 2001.
- [4] S. Satthamsakul, N. Anantrasirichai, C. Benjangkaprasert and T. Wakabayashi, "Rectangular Patch Antenna with Inset Feed and Modified Ground- Plane for Wideband Antenna," SICE Annual Conference 2008, August 20-22, 2008.
- [5] T. Huynh and K.F.Lee, "Single-layer single patch wideband microstrip antenna," Electron Letter, vol.31, no.16, pp. 1310– 1312, August 1995.
- [6] Vedaprabhu, B. and K. J. Vinoy, "A double U-slot patch antenna with dual wideband characteristics," Proc. National Conference on Communications, 361-365, Jan. 2010.
- [7] Li, R., B. Pan, J. Laskar, and M. M. Tentzeris, "A novel low-profile broadband dual-frequency planar antenna for wireless handsets," IEEE Trans. Antennas Propagat., Vol. 56, 1155-1162, 2008.

ENERGY EFFICIENT ADMINISTRATOR BASED SECURE ROUTING FOR MANET

Aniruddha Bhattacharyya

Department Of Computer Science & Engineering
Institute Of Engineering & Management, Salt Lake
Kolkata, India
aniruddha.aot@gmail.com

Arnab Banerjee

Department Of Computer Science & Engineering
Institute Of Engineering & Management, Salt Lake
Kolkata, India
arnab.saheb@gmail.com

Dipayan Bose

Department Of Computer Science & Engineering
Institute Of Engineering & Management, Salt Lake
Kolkata, India
connect2dipayan@gmail.com

Abstract— Energy efficiency is a key issue in wireless ad-hoc network. Energy-efficient MANET routing protocol OSLR is widely used for routing in ad hoc networks. Residual energy of nodes plays a vital role in route discovery in MANET. To support energy-efficient routing, accurate state information about energy levels of nodes should be available. Several directions have been explored to maximize network lifetime using various energy efficient routing. In this paper we have proposed new algorithms which has made routing in MANET more energy efficient. Simulation results show that such protocol can remarkably increase the life-span of network with lower energy consumption when compared to the existing routing schemes.

Keywords- *Manet , OLSR , energy efficient routing,willingness function*

I. INTRODUCTION

Mobile Ad hoc network (MANET) can operate in a self-organized and non predefined infrastructure. The nodes in such network can communicate with each other through direct wireless links or multi-hop routing. It has been used in a wide range of applications ranging from a battlefield to the user's living room. However, due to the limit battery energy of mobile nodes, method of prolonging the lifetime of nodes as well as network becomes the key challenge in MANET. The performance of MANET depends on the routing scheme employed, and the traditional routing protocols do not work efficiently in a MANET. Developing routing protocols for MANETs has been an extensive research area in recent years, and many proactive, reactive and hybrid protocols have

been proposed from a variety of perspectives [7]. These protocols try to satisfy various properties, such as : battery capacity, coverage area and reliability.

II. WORKING METHODOLOGY OF OLSR

The Optimized Link State Routing protocol (OLSR) [1] is a routing protocol that is optimised for MANET and also used on other wireless ad-hoc networks. It is a proactive link-state routing protocol that floods a topology table of it's neighbors to all nodes in the network which then compute optimal forwarding paths locally. AODV(Ad hoc On-Demand Distance Vector) and DSR(Dynamic Source Routing) are other popular protocols used in MANET [2][3].

In OLSR protocol, there are four main steps for creating a routing table: neighbor sensing, ADMIN selection, ADMIN information declaration and route table calculation.

During the neighbor sensing, each node periodically broadcasts the HELLO message containing information about its one-hop neighbors and their link status. All the one-hop neighbors on receiving the HELLO messages doesn't transmit it to further nodes, it updates its knowledge of its one hop neighbors and two-hop neighbors, which gets recorded in a neighbor table on basis of which each node performs the selection of its ADMIN set.

During the ADMIN selection, each node independently selects its ADMIN set according to the ADMIN selection scheme. As a result, all of the two-hop neighbors of each node are contained in the union of the neighbor sets of its ADMINs.

Then, each node declares its ADMINs in the subsequent HELLO messages. From the HELLO messages, which contain the ADMINs, each node can inform its ADMIN Selectors and construct its ADMIN Selector table.

During the ADMIN information declaration, each node broadcasts specific control messages called Topology Control messages to declare its ADMIN Selector set. The TC messages are forwarded through ADMIN nodes and transmitted to all nodes in the MANET. According to the ADMIN selectors and the information in TC messages, a node maintains a network topology table to record the ADMINs of other nodes. The topology table is a base of calculating the route table.

During the route table calculation, a node calculates the route table based on the information contained in the neighbor table and the topology table. To find a path from a source (S) to a destination (D), an intermediate node (I1) one hop away to D has to be found and the connection pair [I1, D] obtained; then, a node (I2) one hop away to I1 has to be found and the connection pair [I2, I1] obtained; and so forth, until a node In is found in the ADMIN sets of S. Based on this process, the route table is built by tracking the connection pairs included in the topology entries in a topology table.

III. RELATED WORKS

Each node of ad hoc network contributes in maintaining the properly functioning network. So node energy is a crucial parameter in ad-hoc network [4]. Due to mobility, it might be possible that a single node becomes a connecting node between two network partitions. If the node battery gets depleted; the network can again become partitioned. In OLSR protocol, the selection of ADMIN depends on willingness value of each node which depends on a node's willingness to forward packets on behalf of other nodes. In case of tie with equal coverage, node with highest willingness value is taken as ADMIN. It depends mainly on battery power of the node [15]. OLSR has other parameters also that can be tuned to improve performance [5]. Willingness value for each node is sent as part of HELLO messages sent to announce its presence. This value is stored internally by each receiving node in its one-hop and two-hop neighbor list. A node chooses its ADMIN based upon the list entries by taking note of their current willingness value to be chosen as ADMIN. Nodes having willingness value as WILL_NEVER suggest that node is unwilling to relay packets as it is sever resource constrained and is never selected as ADMIN by other nodes. OLSR defines willingness value as follows:

| Willingness | Actual value |
|--------------|--------------|
| WILL_NEVER | 0 |
| WILL_LOW | 1 |
| WILL_DEFAULT | 3 |
| WILL_HIGH | 6 |
| WILL_ALWAYS | 7 |

Table 1: OLSR defined willingness values

There has been few other works [6][7][8][9] regarding energy efficient selection of ADMIN based on current battery power of each node and residual energy. But relying solely on battery power might unnecessarily increase number of ADMINs. Without taking care of the chosen node's coverage, reliability etc., an unreliable node may get selected as ADMIN. As increase in number of ADMINs significantly degrade the network performance owing to increase of topology, routing message exchange in network and radio collisions.

IV. PROPOSED ALGORITHMS

In this paper, we are proposing a new mechanism for calculating willingness of a node by taking care of more exhaustive parameters so that numbers of admin nodes are kept to minimum (as per basic OLSR). Also we are incorporating WATCHNODE concept to improve security. Only when it is absolutely necessary, admin node can switch from one node to another – offloading their job to other one increasing network runtime.

A. Admin Node selection:

First we will discuss about Admin node selection algorithm. The value of willingness will be derived from next algorithm given in part B of section IV.

Few definitions:

- **ADMIN(x):** Admin set of node x which is running this algorithm.
- **N(x):** One hop neighbor set of node x (symmetric neighbors)
- **N2(x):** Two hop neighbor set of node x [symmetric neighbors of nodes in N(x)].The two hop neighbor set N2(x) of node x does not contain any one hop neighbor N(x) of node x.
- **D(x,y) :** Degree of one hop neighbor node y (where y is a member of N(x) -- means y belongs to N(x)), is defined as the number of symmetric one hop neighbors of node y EXCLUDING the node x and all the symmetric one hop neighbors of node x, i.e.,

$$D(x, y) = N(y) - \{x\} - N(x) \quad (1)$$

- **V** = Willingness value of the node.

Algorithm:

- Step 1.** Start with an empty ADMIN(x) set.
- Step 2.** Calculate D(x,y), where y is a member of N(x), for all nodes in N(x) (put for all +ve sign)
- Step 3.** First select as ADMINs those nodes in N(x) which provide the "only path" to reach some nodes in N2(x)

Step 4. For each node in $N(x)$ calculate a [] = V

{

4.1. IF (WILL_VALUE is MAX && WILL_VALUE not 0)

{

 SELECT as a ADMIN

}

ELSE IF the WILL_VALUE is equal to more than one node in $ADMIN(x)$

{

 Take node $ADMIN(x)$ whose coverage is MAX.

}

4.2. While there still exist some nodes in $N_2(x)$ that is not covered by $ADMIN(x)$:

{

 For each node in $N(x)$, calculate the no. of nodes in $N_2(x)$ which are not yet covered by $ADMIN(x)$ and are reachable through this one hop neighbor of x.

}

4.3. Select as a ADMIN that node of $N(x)$ which reaches the maximum number of uncovered nodes in $N_2(x)$.

4.4. If a tie occurs, select that node as ADMIN whose $D(x,y)$ is greater.

}

Step 5. To optimize, process each node y in $ADMIN(x)$, one at a time, if $ADMIN(x) - \{y\}$ still covers all nodes in $N_2(x)$ then remove y from $ADMIN(x)$.

Step 6. After that Convert the link between node x and ADMIN as SYM_LINK to ADMIN_LINK

Step 7. Exit

B. Dynamic Willingness calculation:

Our algorithm takes a weighted sum of battery power of a node, coverage area and reliability of the node while calculating willingness value.

Willingness (P, C, R) =

$$(0.75 * P) + (0.15 * C) + (0.1 * R) \quad (2)$$

Where, P: power available for that node (in %)

C: coverage (in %)

R: reliability of the node (in %)

Power (P) is defined as:

$$P = (\text{current node power/rated capacity of the node}) * 100 \quad (3)$$

Coverage (C) is defined as:

$$C = (\text{no of 1-hop neighbors of that node / no of 2-hop neighbors of nodes that want to select this node as its ADMIN}) * 100 \quad (4)$$

Reliability (R) is calculated from various sensor inputs regarding outside environment condition. Value of R ranges from 0% to 100% depending upon where the node is situated. Nodes deployed in hazardous area such as high radiation, high temperature and zones where shock and vibration is common will be less reliable and will likely to fail more often than nodes that are situated outside of such area. Choosing such unreliable nodes as ADMIN will make network vulnerable to network outage and network being partitioned.

$$R = \{0\% \dots 100\% \} \quad (5)$$

Each node sends HELLO message after certain time interval to let know its neighbors that it still exist in network. Each entry in 1-hop and 2-hop neighbor is stored with a timeout value, upon expiration of that the entry is deleted. The entry is renewed upon receiving a HELLO messages from that node. Therefore the modified willingness value will be updated in each of its 1-hop neighbors and later to 2-hop neighbors after certain interval (generally within few seconds).

C. WatchNode calculation:

This algorithm helps the network from external threats. Any node whose behavior is malicious can be identified that WATCH NODE and this message broadcasted inside the network as a warning message. A node whose behavior is malicious can never be accepted as a trusted node in future.

Step 1. Initialize the value of Turn, T=0;

count_admin(T) → no. of ADMINS in the ADMIN set at

 level 'T+1' ;
 source_node(T) → 1st node ;
 Admin (count_admin (T)) → immediate ADMINS of
 source_node(t) ;

Step 2. Initial source node [source_node(T)] broadcasts control message to the network;

T++;

Step 3. the intermediate ADMINS [Admin(count_admin (t))] receives the broadcasted message of source_node(T) ;
 Set c=0 ;

Step 4. s=T-1 ;

Step 5. prev_source(T) ← source_node(s);
 source_node(T) ← Admin(c) ;

Step 6. source_node(T) broadcasts the message received from source_node(s) also known as prev_source(T) ;

Step 7. IF the broadcasted message sent by source_node(T)

reaches the prev_source(t) within a threshold time

$$[2 * (\text{Hop Delay}) + x * v]$$

where v= random waiting interval, x = 0,1,2,3.

THEN

Step 7a. the source_node(T) is a normal node.

ELSE

Step 7b. the source_node(T) is a malicious node
and an alarm is set that source_node(T)
is a malicious node.

[The prev_source broadcasts the message to its neighbor nodes and the neighbor nodes further forwards it to its further neighbor nodes]

watch_node \leftarrow source_node(s);

Step 8. c++ ;

Step 9. repeat steps 5 to 8 until c<= count_admin (T)

Step 10. T++ ;

```
if count_admin (T) <> 0
then
goto Step 3
```

Step 11. End.

V. PERFORMANCE EVALUATION

A. Simulation Environment

We used OLSR protocol implementation from Niigata University for Glomosim [12].

Table 2: Simulation parameters

| Parameter | Value |
|---------------------------------|---------------------|
| Terrain Dimension | (600x500) sq. meter |
| Simulation Time | 500 minutes |
| Channel | Noisy |
| Noise Figure | 10 dB |
| Radio Frequency | 2.4 Ghz |
| Radio Receive Threshold | -65.046 dBm |
| Radio Transmit Power | 22.5 dBm |
| Node Placement | Random |
| Mobility Speed | 0-10 m/s |
| MAC Protocol | 802.11 |
| MAC Propagation Delay | 1000 ns |
| Bandwidth | 11 Mbps |
| Routing Protocol | OLSR, EEABSR |
| Number of Interface per node | 2 |
| Rated Battery Power (each node) | 1500 mAh |
| Data Packet Type | FTP, CBR |
| Data Packet Size | 2044 byte |

To simulate the proposed algorithm we used Glomosim 2.03

network simulator [10]. Glomosim can simulate both wired and wireless network with layered TCP/IP stack with model based on noisy & noiseless channel with MAC protocol 802.11/CSMA/MACA/TSMA and various network, transport & application layer protocols. GlomoSim is written using PARSEC language [11], a C derivative for large scale parallel simulation.

B. Energy Consumption Model

We are using IEEE 802.11b (DSSS modulation) as MAC protocol. The transceiver uses energy both to transmit and to listen for incoming packet. It also consumes energy in idle state. Let, the energy needed to transmit a packet E_t for duration t_t and to receive a packet E_r for duration t_r . Also assume it waits for t_i consuming energy E_i . Then total energy consumed by that node will be approximately:

$$E_c = E_t * t_t + E_r * t_r + E_i * t_i \quad (6)$$

We assumed each node will use 5V DC battery with rated capacity of 1500 mAh. Transmission energy consumed will depend on radio signal strength of transmission; here we assumed 22.5 dBm; which approximately translate into 177.83 mW.

$$A = \frac{W}{V} \quad (7)$$

From equation (7) we get, $A = 35.57$ mA for $V=5V$ DC. If we draw the same amount of current, using 1500 mAh battery, we'll get approximately 42 hour of runtime before the battery dies. Adding Idle and receiver power we'll get less than that. We used value for Idle and receiver power based on typical low power sensor network currently used, with appropriate overages added to make it more realistic. [13]

C. Simulation Results

We ran the simulation with varying number of nodes using default OLSR and our modified version using dynamic willingness keeping other parameters same. The result was dependant on actual data being sent and frequency of data transfer. First we evaluate number of admin in the network by both protocol variant as a function of number of nodes. Maximum numbers of nodes were set to 50. Simulation results are illustrated in following figures:

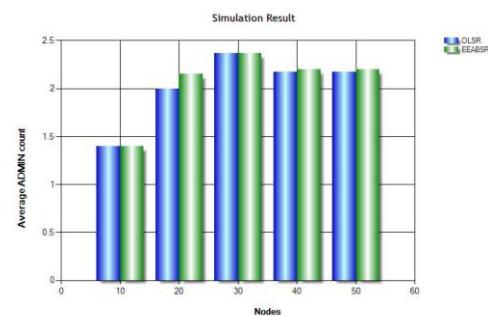


Figure 1: ADMIN Count

Here we can't see much difference in average ADMIN count over basic OLSR protocol. We can also see that number

2nd Annual International Conference IEMCON 2012

of ADMIN count increased slightly when numbers of nodes were 20, 40 & 50. This increase in ADMIN count is due to shift of responsibility as the node's willingness changes with time. Having significantly increased ADMIN will adversely decrease network performance. But here the count has increased only slightly. Figure 2 depicts average HOP counts with variation of node density.

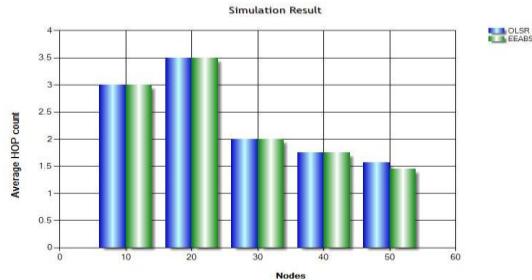


Figure 2: Average HOP Count

As the network became denser, so does the decrease of node hopping, because nodes came closer together decreasing requirement of intermediate HOP nodes. But with increase in ADMIN will affect in radio layer packet collision, as depicted in figure 3:

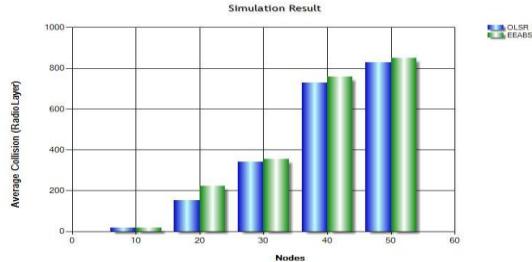


Figure 3: Average Collision

We can see the collision in fact has increased, but only slightly as the increase of ADMIN count was not so drastic. This increase was due to reselection of ADMIN and subsequent topology message being broadcasted internally. Collision increases with network density as more and more nodes are trying to compete for radio frequency. Using 802.11b reduced collision due to deliberate use of collision avoidance scheme (such as RTS/CTS) built into radio layer protocol itself. Unfortunately we saw no change in throughput what so ever in both protocol, as depicted in following figure:

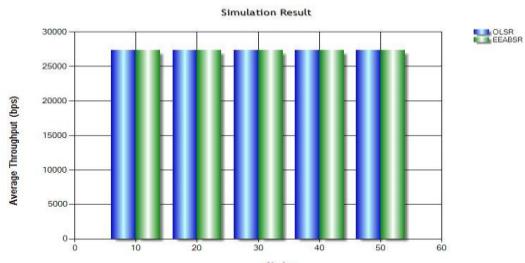


Figure 4: Average Throughput

With 11 Mbps network bandwidth and multiple FTP and CBR data transfer, we saw average throughput stayed around

27 kbps. This we suspect due to multiple HOPs required and other control message being broadcasted internally consuming precious network bandwidth. We also saw improvement in end-to-end delay with our proposed protocol:

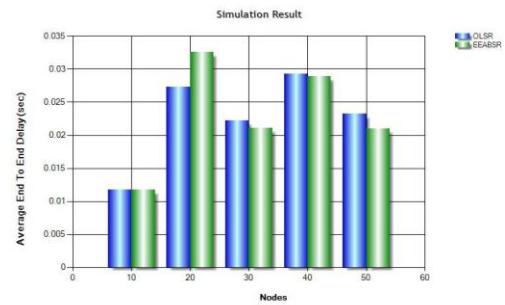


Figure 5: Average End-To-End Delay

End-To-End delay first increased with node number then it suddenly dropped around 30 nodes and then it again rose. As the plot shows, our protocol can decrease end to end delay slightly. Whereas, packet delivery rate hovers around 1:

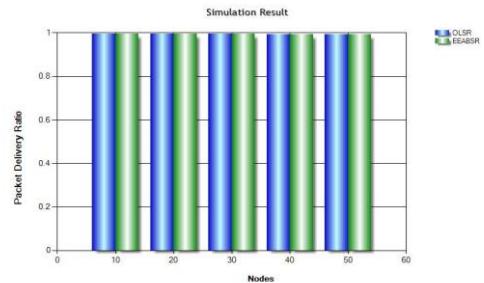


Figure 6: Packet Delivery Ratio

Our protocol's main goal was to reduce average energy consumption per node. It is achieved by delegating roles to other nodes when node faced either energy crisis, coverage changes or node's movement towards unfavorable environment. In next few figures we will be discussing about energy consumption side of the simulation. As we have said earlier, we have tried to simulate energy consumption as close as possible with real world. But, real world scenario might be dependent on many other parameters such as ambient temperature, embedded microcontroller & network subsystem power requirement and other unforeseen effects (like battery aging, battery type- Li-ion versus Alkaline versus Ni-MH etc). So there might be some margin of error in results.

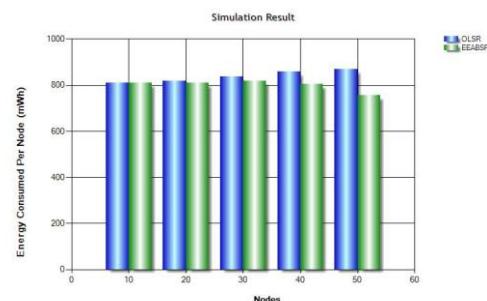


Figure 7: Average Energy Consumed Per Node

Energy efficiency is a critical for increasing the lifetime of the individual nodes as well as the overall network [14]. We can see, energy consumed per node increase in OLSR but our protocol it saved energy a little as load was distributed among nodes. Saving will be more prominent when running for full runtime that is possible with currently assumed rated battery. In the following figure we are plotting average energy remained per node with varying number of nodes:

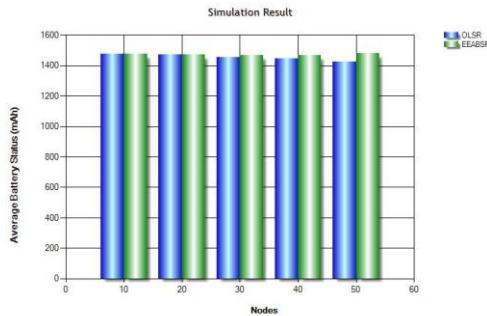


Figure 8: Average Battery Status

From figure 8, it is clear that our protocol is saving precious node energy, thereby increasing network runtime. The effect will be more prominent with more running time of network.

VI. FUTURE WORK

We have already implemented three parameters of willingness functions (power, coverage area, reliability). For future work we'll be adding more parameters into willingness function such as: link type (symmetric / asymmetric), link stability and bandwidth. After incorporating three parameters the tuning is required for the optimization. Also these parameters help us to reduce the number of ADMIN NODE. As we know that less number of ADMIN NODES reduces drastically collision & packet dropping factor. So the algorithm will be much more efficient. We are also going to incorporate the security factor into the newly designed routing protocol by implementing the WATCH NODE concept. The selection of the Watch Node is dynamic and whose job is to look after the intermediate nodes that only forward the destined packets. If any node performs its job in a correct way then that node will be rewarded (increasing Trust value) else the Trust value will be deducted. If the trust value is less than a threshold value, then that particular node will be treated as a malicious node. This message will be broadcasted to the entire network.

VII. CONCLUSION

In this paper Energy efficient administrator based secure routing has been proposed. This novel feature allows us to select ADMIN NODES in an efficient way. The parameters inside the WILLINGNESS function made the selection procedure dynamic. The performance of this routing algorithm in comparison to OLSR has been improved. More parameters, further tuning and security implementation (by Watch Node) will make the algorithm more compact and concur. The hop count between source and destination also minimized. The security implementation also protects the network from internal and external threats.

REFERENCES

- [1] T. Clausen, P. Jacquet, "Optimized Link State Routing Protocol (OLSR)", RFC 3626, <http://tools.ietf.org/html/rfc3626>
- [2] Shaista Anju, Sarita Singh Bhaduria, "TCP and UDP based analysis of AODV and OLSR in Mobile Ad-Hoc Networks", International Conference on Communication Systems and Network Technologies, 2011
- [3] Floriano De Rango, Juan-Carlos Cano, Marco Fotino, Carlos Calafate, Pietro Manzoni, Salvatore Marano, "OLSR vs DSR: A comparative analysis of proactive and reactive mechanisms from an energetic point of view in wireless ad hoc networks", Computer Communications, October 2008
- [4] Saoucene Mahfoudh, Pascale Minet, "Survey of energy efficient strategies in wireless ad hoc and sensor networks", International Conference on Networking, April 2008
- [5] C. Gomez, D. Garcia, J. Paradells, "Improving Performance of a Real Ad-hoc Network by Tuning OLSR Parameters", IEEE Symposium on Computers and Communications, June 2005
- [6] Floriano De Rango, Marco Fotino, Salvatore Marano, "EE-OLSR: Energy Efficient OLSR routing protocol for mobile adhoc networks", D.E.I.S. Department, University of Calabria, Rende, Italy , 2008
- [7] Saoucene Mahfoudh, Pascale Minet, "An energy efficient routing based on OLSR in wireless adhoc and sensor networks", 22nd International Conference on Advanced Information Networking and Applications - Workshops, 2008.
- [8] Floriano De Rango, Marco Fotino, "Energy efficient OLSR performance evaluation under energy aware metrics", SPECTS'09, July 2009
- [9] Li Zhiyuan, Hu Jinhong, "Simulation and Analysis of Optimized OLSR", International Conference on Multimedia Information Networking and Securit, 2010
- [10] Xiang Zeng, Rajive Bagrodia, Mario Gerla, "GloMoSim: A Library for Parallel Simulation of Large-scale Wireless Networks", Workshop on Parallel and Distributed Simulation, May 1998
- [11] Rajive Bagrodia, Richard Meyer, Mineo Takai, Yu-an Chen, Xiang Zeng, Jay Martin, Ha Yoon, "Parsec: A Parallel Simulation Environment for Complex Systems", October 1998
- [12] Niigata University, Information & Communication Networks laboratory, "OLSR_Niigata", <http://www2.net.ie.niigata-u.ac.jp/olsr-e.php>.
- [13] Thomas Kunz, "Accurately Predicting Residual Energy levels in MANETs", IEEE International Conference on Wireless & Mobile Computing, Networking & Communication
- [14] Zhang, Xiaoying; Kunz, Thomas; Li, Li; Yang, Oliver, "An Energy-Efficient Broadcast Protocol in MANETs: Design and Evaluations", Communication Networks and Services Research Conference (CNSR), 2010 Eighth Annual, 11-14 May 2010, pp. 199 – 206.
- [15] Kunz, T. "Energy-Efficient Variations of OLSR", Wireless Communications and Mobile Computing Conference, 2008. IWCMC '08. International, pp. 517 - 522

NEURAL & ARTIFICIAL NETWORKS

Date : 18th January,2012

Chairman : Prof. B.N.Roy

Co-Chairman : Prof. Dipak Chatterjee

Rapporter : Prof. Sourav Saha

| Paper ID | Paper Title | Authors |
|----------|---|---|
| 19 | Benefit Based Artificial Intelligence Regulated Stability Governanee of a Practical Power Network | Saptarshi Bhattacharya, Rumna Kuthial, Partha Pratim Das and Sawan Sen |
| 41 | Artificial Neural Network Modeling of Instrusion Detection and Prevention System | Tanusree Chatterjee, Indraneel Mukhopadhyay and Mohuya Chakraborty |
| 42 | Intrusion Detection In Wireless Sensor Networks Using Fuzzy Logic | Moutushi Singh, Abdur Rahaman Sardar and Subir Kumar Sarkar |
| 44 | Artificial Neural Network Based Automatic Mobile Radio Signal Classifier | Swagata Roy Chaterjee, Indranil Banerjee, Soma Bhattacharyya and Mohua Chakrobory |
| 72 | Study The Nonlinear Behaviour of PFC Boost Converter | Arnab Ghosh |
| 98 | A Review of Secure Routing in Mobile ad hoc Network | Abari Bhattacharya, Himadri Nath Saha |
| 103 | Leader-Follower Formation Control Using Artificial Potential Functions | Kishorekumar H Kowdiki, Ranjit Kumar Barai, Samar M Battacharya. |

Leader-Follower Formation Control Using Artificial Potential Functions

Kishorekumar H Kowdiki,
 Department of Electrical Engg.,
 Jadavpur University,
 Kolkata-700032
 Email: kkowdiki@yahoo.com

Ranjit Kumar Barai,
 Department of Electrical Engg.,
 Jadavpur University,
 Kolkata-700032
 Email: ranjit.k.barai@gmail.com

Samar M Bhattacharya.
 Department of Electrical Engg.,
 Jadavpur University,
 Kolkata-700032

Abstract- This paper presents a novel hybrid formation control technique for a group of differential-drive mobile robots employing both artificial potential field based navigation and leader-follower formation control scheme. In the proposed method, the leader robot of the group determines its path of navigation by an artificial potential field. Then the other robots in the group follow the leader maintaining a particular formation employing the leader-follower formation control scheme. As the leader robot navigates itself by artificial potential field, its locomotion control is stable and robust against collision while reaching to the goal position. The follower robots adapt its formation by suitably controlling the desired separation distance l_{12}^d and the bearing angle Ψ_{12}^d . Thus, the original formation can be regained even if the formation is lost due to passage through narrow opening / path. Therefore, the overall formation control scheme results into a robust and adaptive formation control for a group of autonomous differential-drive mobile robots. The effectiveness performance of the proposed method has been verified in simulation.

Keywords- Artificial Potential Field based Navigation, Formation Control, Leader-Follower approach, Tracking Control.

I. INTRODUCTION

Formation control means to control a group of robots to form up and to move in specified geometric shapes. Formation control of multi-robot is a methodology of keeping multi-robot system in the scheduled formation and adapting to the environment while moving to the target. Formation control is an important issue in coordinated control for a group of unmanned autonomous vehicles/robots. In many applications, a group of autonomous vehicles are required to follow a predefined trajectory while maintaining a desired spatial (geometric) pattern. Various types of shapes have been employed in formation control. The more common formation shapes are column, line, wedge, triangle, and circle. Formation control of multi-robot is very important in military, aerospace, industry and so on, and is widely used in these fields.

Robot motion design can be approached at three levels: a) A geometric level, where the trajectory in the configuration space is designed in the presence of obstacles; b) A kinematic level, where the velocity profiles are designed; and c) a dynamics level, where robot forces are designed as

well. A good robot motion planning method must lead to a robot trajectory with desirable geometrical features (e.g. robot moves to target along a short path while keeping a good safety distance from obstacles), desirable kinematical features (e.g. robot maintains a reasonably uniform and brisk speed while traveling but slows down in tight spots), and desirable dynamic features (e.g. robot forces are reasonable and easy to compute).

The artificial potential field (APF) approach is a widely adopted approach to mobile robot navigation and control and which claims to address all three levels either directly or indirectly. APF is easy to implement with low computational complexity and good performance.

Among all the approaches to formation control reported in the literature, formation control using APF by considering inter-agent interaction has been adopted by some researchers [8] [9], while some researchers has adopted the leader-follower formation control by the graph theoretic methods[1] [11][12]. In this method, each robot takes another neighboring robot as a reference point to determine its motion. The referenced robot is called a leader, and the robot following it called a follower

In our approach, we are allowing the leader robot to move in the APF i.e. the path planning of the leader robot is done by the APF and the followers will track the leader's path by suitably controlling the desired separation distance l_{12}^d and the bearing angle Ψ_{12}^d . Thus, the formation control problem becomes a tracking control problem.

This paper is organized as follows. In section II the kinematic model is given, section III will describe the artificial potential functions for leader's path planning, section IV deals with the tracking control. In section V simulation results are presented and finally conclusion and the future work to be carried out are shown in section VI.

II. KINEMATIC MODEL

We are considering the kinematic model described by [2], in which two scenarios for feedback control within the formation are described. In the first scenario, one robot follows another by controlling the relative distance and orientation between the two ($l - \psi$) and in another scenario, a robot maintains its position in the formation by maintaining a

specified distance from two robots, or from one robot and an obstacle in the environment ($l - l$) .

We have considered the $l - \psi$ formation scheme. In the $l - \psi$ control of the two mobile robots, the aim is to maintain a desired length (separation distance), l_{12}^d and a desired relative angle (bearing angle) ψ_{12}^d between the two robots. The kinematic equations for the system of two mobile robots shown in Fig. 1 is given by

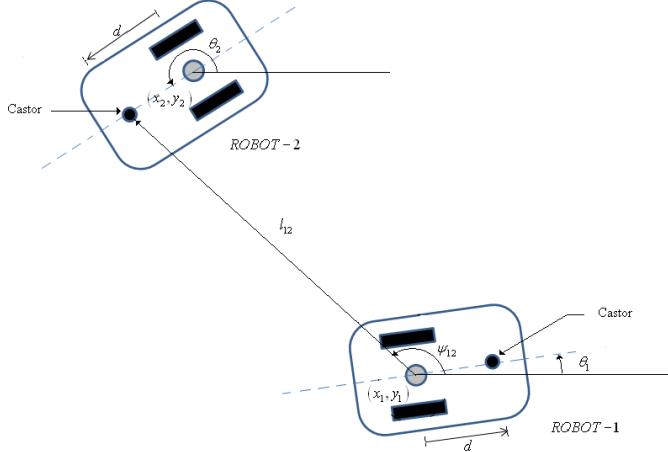


Fig. 1: Notation for $l - \psi$ control

$$\begin{aligned}\dot{x}_i &= v_i \cos \theta_i \\ \dot{y}_i &= v_i \sin \theta_i \\ \dot{\theta}_i &= \omega_i\end{aligned}\quad (1)$$

The model for leader-follower formation using $l - \psi$ formation scheme is given as:

$$\begin{aligned}\dot{l}_{12} &= v_2 \cos \gamma_1 - v_1 \cos \psi_{12} + d \omega_2 \sin \gamma_1 \\ \dot{\psi}_{12} &= \frac{1}{l_{12}} \{v_1 \sin \psi_{12} - v_2 \sin \gamma_1 + d \omega_2 \cos \gamma_1 - l_{12} \omega_1\} \\ \dot{\theta}_2 &= \omega_2\end{aligned}\quad (2)$$

where, $\gamma_1 = \theta_1 + \psi_{12} - \theta_2$ and $v_i, \omega_i (i = 1, 2)$, are the linear and angular velocities at the center of the axle of each robot. In order to avoid collisions between robots, we will require that $l_{12} > d$. where d is the distance between the castor wheel and the centre of rear wheels.

III. ARTIFICIAL POTENTIAL FUNCTION FOR PATH PLANNING OF THE LEADER ROBOT

We consider the simple artificial potential function described by [3]. In this method, a robot is modeled as a moving particle inside an artificial potential field that is generated by superposing an attractive potential that pulls the robot to a goal configuration and a repulsive potential that pushes robot away from obstacles. The negative gradient of the generated global potential field is interpreted as an artificial force acting on the robot and dictating its motion.

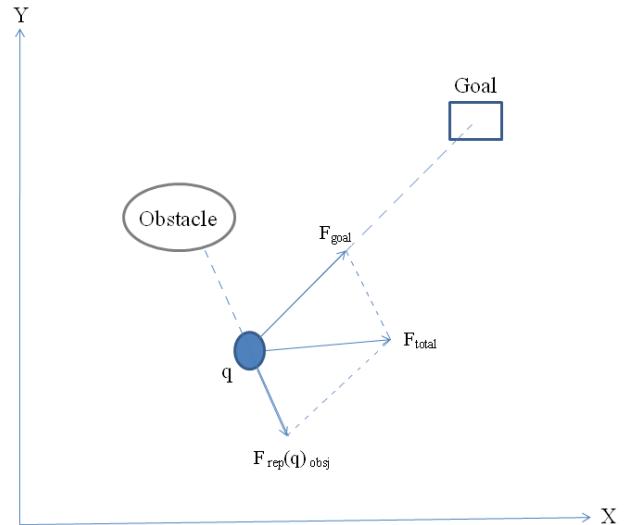


Fig. 2: Moving direction of robot in Artificial Potential Field

Artificial potential field (APF) two has two kinds of potential sources: gravitation pole and repulsion pole. The target is the gravitation pole, and the obstacle is the repulsion pole. They jointly produce the artificial potential field. As shown in Fig. 2, the negative gradient of the APF is the moving direction of the robot in the system.

The target gravitation and obstacle repulsion in APF are defined as:

Let q be the position of the robot, $\rho(q, g)$ be the distance between the robot and the target g , the gravitational potential field U_g and gravitation F_g at robot q are defined as:

$$U_g(q) = \frac{1}{2} \xi \rho^2(q, g) \quad (3)$$

$$F_g(q) = \xi \rho(g, q) \quad (4)$$

Let q_{obsj} ($j=1, \dots, m$) be the position of the j^{th} obstacle, $\rho(q, q_{obsj})$ be the distance between robot q and the q_{obsj} , the repulsive potential field $U_{rep}(q)$ and gravitation $F_{rep}(q_{obs})$ at the robot q are defined as:

$$U_{rep}(q) = \begin{cases} \frac{1}{2} \zeta \left(\frac{1}{\rho(q, q_{obsj})} - \frac{1}{\rho_s} \right)^2 & \text{if } \rho(q, q_{obsj}) \leq \rho_s \\ 0 & \text{if } \rho(q, q_{obsj}) > \rho_s \end{cases} \quad (5)$$

$$F_{rep}(q_{obs}) = \begin{cases} \zeta \left(\frac{1}{\rho(q, q_{obsj})} - \frac{1}{\rho_s} \right)^2 \frac{1}{\rho^2(q, q_{obsj})} \nabla \rho(q, q_{obsj}) & \\ 0 & \end{cases}$$

$$\begin{aligned} & \text{if } \rho(q, q_{obj}) \leq \rho_s \\ & \text{if } \rho(q, q_{obj}) > \rho_s \end{aligned} \quad (6)$$

So the resultant force of robot q in the APF is:

$$F_{total}(q) = F_g(q) + \sum_{j=1}^m F_{rep}(q_{obj}) \quad (7)$$

A fundamental problem in the application of potential field method is how to deal with the local minima that may occur in a potential field environment [4]. We have not considered the problem of local minima.

IV. TRACKING CONTROL

The robots are considered as a point mass. The leader's path is defined by the artificial potential functions, depending upon location of static obstacles and the goal position. The follower robot will follow the leader by keeping the separation distance l_{ij} and bearing angle Ψ_{ij} .

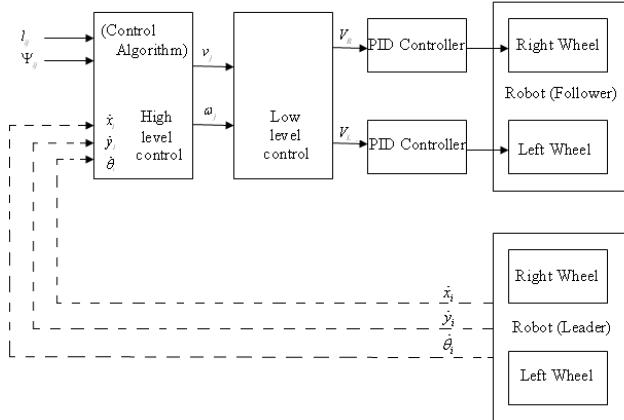


Fig. 3: Block Schematic for Formation Control.

Fig. 3 shows the block schematic for formation control. Since, next position and orientation of the leader robot are dictated by the artificial potential field, we have \dot{x}_i , \dot{y}_i and $\dot{\theta}_i$. l_{ij} and Ψ_{ij} i.e. the separation distance and the bearing angle respectively are the command inputs, which the follower robot has to maintain while following the leader robot. We will have to develop the control algorithm for deriving the velocities of the left and right wheels of the follower robot.

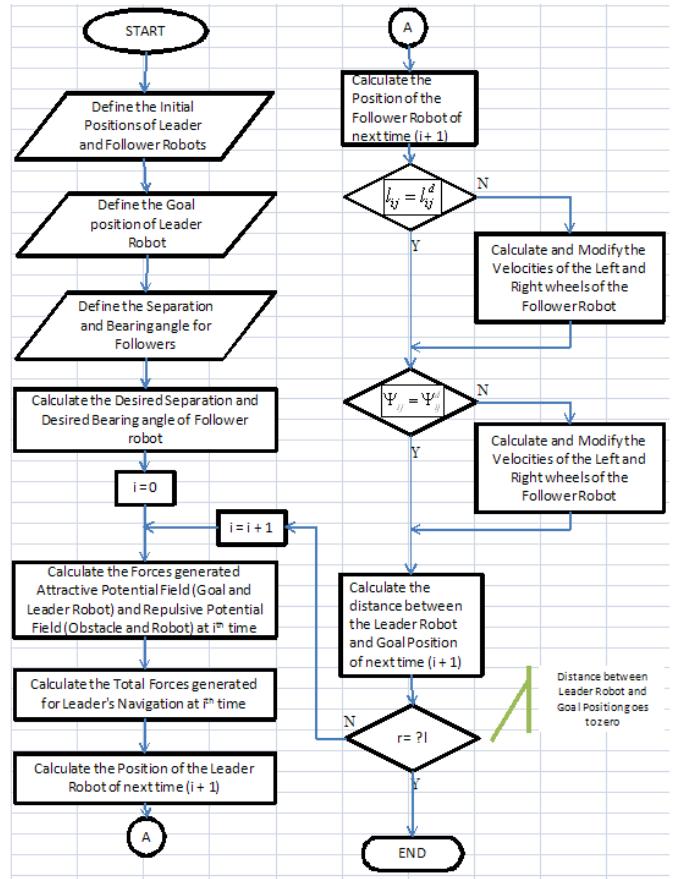


Fig. 4: Flow Chart for Formation Control.

The formation control scheme is described in the flow chart as shown in Fig. 4.

To avoid collisions, separation distances are measured from the back of the leader to the front of the follower, and the kinematic equations for the front of the j^{th} follower robot can be written as:

$$\dot{q} = \begin{bmatrix} \dot{x}_j \\ \dot{y}_j \\ \dot{\theta}_j \end{bmatrix} = S_j(q_j)v_j = \begin{bmatrix} \cos\theta_j & -d\sin\theta_j \\ \sin\theta_j & -d\cos\theta_j \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_j \\ \omega_j \end{bmatrix} \quad (8)$$

Where d is the distance from the rear axle to the front of the robot,

Consider the tracking controller error system presented in [5] used to control a single robot as

$$\begin{bmatrix} e_{j1} \\ e_{j2} \\ e_{j3} \end{bmatrix} = \begin{bmatrix} \cos\theta_j & \sin\theta_j & 0 \\ -\sin\theta_j & \cos\theta_j & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{jr} - x_j \\ y_{jr} - y_j \\ \theta_{jr} - \theta_j \end{bmatrix} \quad (9)$$

$$\begin{aligned} \dot{x}_{jr} &= v_{jr} \cos\theta_{jr}, \dot{y}_{jr} = v_{jr} \sin\theta_{jr}, \dot{\theta}_{jr} = \omega_{jr}, \\ \dot{q}_{jr} &= \begin{bmatrix} \dot{x}_{jr} & \dot{y}_{jr} & \dot{\theta}_{jr} \end{bmatrix}^T \end{aligned} \quad (10)$$

where x_j , y_j , and θ_j are actual position and orientation of the robot, and x_{jr} , y_{jr} , and θ_{jr} are the positions and orientation of a virtual reference cart robot j seeks to follow [5].

The basic tracking control problems can be extended to formation control as follows. The virtual reference cart is replaced with a physical mobile robot acting as the leader i , and x_{jr} and y_{jr} are defined as points at a distance l_{ijd} and a desired angle Ψ_{ijd} from the lead robot. Now basic navigation problems can be introduced for leader-follower formation control as.

Let there be a leader i for follower j such that

$$\begin{bmatrix} \dot{x}_i \\ \dot{y}_i \\ \dot{\theta}_i \end{bmatrix} = \begin{bmatrix} \cos \theta_i & -d \sin \theta_i \\ \sin \theta_i & d \cos \theta_i \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_i \\ w_i \end{bmatrix} \quad (11)$$

$$x_{jr} = x_i - d \cos \theta_i + l_{ij}^d \cos(\Psi_{ij}^d + \theta_i)$$

$$y_{jr} = y_i - d \sin \theta_i + l_{ij}^d \sin(\Psi_{ij}^d + \theta_i)$$

$$\theta_{jr} = \theta_i$$

And

$$v_{jr} = [v_j \quad \omega_j]^T \quad (13)$$

Then the actual position and orientation of the follower j with respect to leader i can be defined as

$$\begin{aligned} x_j &= x_i - d \cos \theta_i + l_{ij} \cos(\Psi_{ij} + \theta_i) \\ y_j &= y_i - d \sin \theta_i + l_{ij} \sin(\Psi_{ij} + \theta_i) \\ \theta_j &= \theta_i \end{aligned} \quad (14)$$

Where l_{ij} and Ψ_{ij} is the actual separation and bearing of the follower j .

Using (12), (14) and simple trigonometric identities, the error system (9) for follower robot j with respect to leader robot i can be rewritten as

$$\begin{bmatrix} e_{j1} \\ e_{j2} \\ e_{j3} \end{bmatrix} = \begin{bmatrix} \cos \theta_j & \sin \theta_j & 0 \\ -\sin \theta_j & \cos \theta_j & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} l_{ij}^d \cos(\Psi_{ij}^d + \theta_i) - l_{ij} \cos(\Psi_{ij} + \theta_i) \\ l_{ij}^d \sin(\Psi_{ij}^d + \theta_i) - l_{ij} \sin(\Psi_{ij} + \theta_i) \\ \theta_i - \theta_j \end{bmatrix} \quad (15)$$

And after further simplification (14) becomes

$$e_j = \begin{bmatrix} e_{j1} \\ e_{j2} \\ e_{j3} \end{bmatrix} = \begin{bmatrix} l_{ij}^d \cos(\Psi_{ij}^d + e_{j3}) - l_{ij} \cos(\Psi_{ij} + e_{j3}) \\ l_{ij}^d \sin(\Psi_{ij}^d + e_{j3}) - l_{ij} \sin(\Psi_{ij} + e_{j3}) \\ \theta_i - \theta_j \end{bmatrix}$$

$$= \begin{bmatrix} l_{ij}^d \cos(\Psi_{ij}^d + e_{j3}) - l_{ij} \cos(\Psi_{ij} + e_{j3}) \\ l_{ij}^d \sin(\Psi_{ij}^d + e_{j3}) - l_{ij} \sin(\Psi_{ij} + e_{j3}) \\ \theta_i - \theta_j \end{bmatrix} \quad (16)$$

This transformed error system now acts as a formation tracking controller which not only seeks to remain at a fixed desired distance l_{ij}^d with a desired angle Ψ_{ij}^d relative to the leader robot i , but also achieves the same orientation as the leader robot which is desirable when $\omega_i = 0$

In order to calculate the dynamics of the error system (16), it is necessary to calculate the derivatives of l_{ij} and Ψ_{ij} , and it is considered that the desired separation distance l_{ij}^d and the desired bearing angle Ψ_{ij}^d are constant. Consider the two robot formation as shown in fig. 1. The x and y components of l_{ij} can be defined as

$$\begin{aligned} l_{ijx} &= x_{irear} - x_{jfront} = x_i - d \cos \theta_i - x_j \\ l_{ijy} &= y_{irear} - y_{jfront} = y_i - d \sin \theta_i - y_j \end{aligned} \quad (17)$$

And the derivatives of the x and y components of l_{ij} can be found as

$$\begin{aligned} \dot{l}_{ijx} &= v_i \cos \theta_i - v_j \cos \theta_j + d \omega_j \sin \theta_j \\ \dot{l}_{ijy} &= v_i \sin \theta_i - v_j \sin \theta_j + d \omega_j \cos \theta_j \end{aligned} \quad (18)$$

Noting that $l_{ij} = \sqrt{l_{ijx}^2 + l_{ijy}^2}$ and

$$\Psi_{ij} = \arctan\left(\frac{l_{ijy}}{l_{ijx}}\right) - \theta_i + \pi, \text{ It can be shown that the}$$

derivatives of separation and bearing are similar as the kinematic equation described in (2).

$$\begin{aligned} \dot{l}_{ij} &= v_j \cos \gamma_j - v_i \cos \Psi_{ij} + d \omega_j \sin \gamma_j \\ \dot{\Psi}_{ij} &= \frac{1}{l_{ij}}(v_i \sin \Psi_{ij} - v_j \sin \gamma_j + d \omega_j \cos \gamma_j - l_{ij} \omega_i) \end{aligned} \quad (19)$$

Where $\gamma_j = \Psi_{ij} + e_{j3}$.

Now, using derivative of (16), equation (19) and applying simple trigonometric identities, the error dynamics becomes

$$\begin{bmatrix} \dot{e}_{j1} \\ \dot{e}_{j2} \\ \dot{e}_{j3} \end{bmatrix} = \begin{bmatrix} -v_j + v_i \cos e_{j3} + \omega_j e_{j2} - \omega_i l_{ij}^d \sin(\Psi_{ij}^d + e_{j3}) \\ -\omega_j e_{j1} + v_i \sin e_{j3} - d \omega_j + \omega_i l_{ij}^d \cos(\Psi_{ij}^d + e_{j3}) \\ \omega_i - \omega_j \end{bmatrix} \quad (20)$$

Examining (20) and the error dynamics of a tracking controller for a single robot in [5], it can be seen that dynamics of a single follower with a leader is similar to [5], except additional terms are introduced as a result of (8) and (19).

We have considered only $(l - \Psi)$ control of leader-follower formation schemes, this can be also implemented for $(l - l)$ control.

V. SIMULATION RESULTS

A triangular formation of three identical mobile robots is considered where the leader's path /navigation is dictated by the artificial potential field and are considered as the desired formation trajectory, and simulations are been carried out in MATLAB with various cases. Simulation cases of formation control with one, two and three obstacles in the environment.

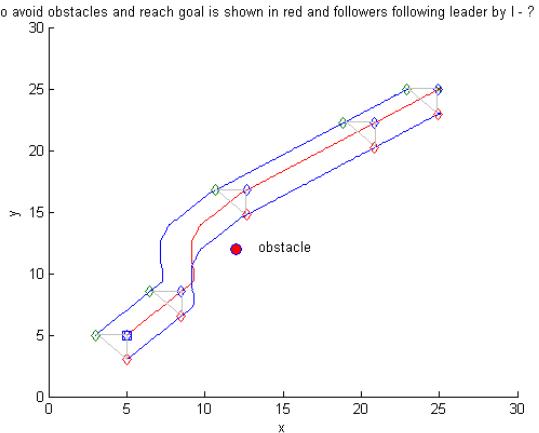


Fig. 5: Simulation of triangular formation with one obstacle.

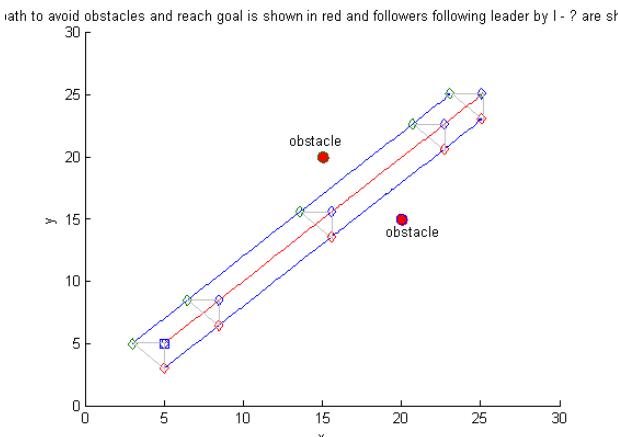


Fig. 6: Simulation of triangular formation with two obstacles.

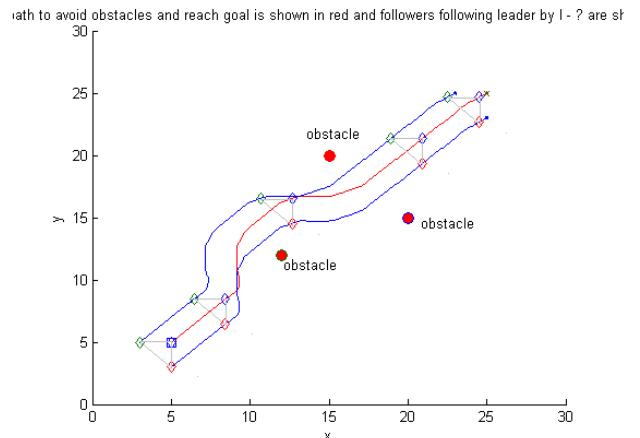


Figure 7: Simulation of triangular formation with three obstacles.

Figure 5, 6 and 7 shows the triangular formation with one, two and three obstacles in the environment respectively. The initial positions of the leader and two followers are defined. The goal/target position for the leader is defined. The positions of the obstacles are defined. The leader's path is dictated by the artificial potential fields as in section III shown in red line. The followers (robots) follow the leader as described in section IV and are shown in blue lines. Robots positions after regular time intervals are also shown.

VI. CONCLUSION AND FUTURE WORK

Based on the simulation results, it can be seen that as the leader robot navigates itself by artificial potential field, its locomotion control is stable and robust against collision while reaching to the goal position and the followers are following the leader's path effectively.

Thus, from the simulation results, we one can see that the desired formation control using leader-follower scheme is effective. In future work, a suitable controller will be selected to show that the error dynamics described in this paper will tend to zero as time tends to infinity. And in this work we have considered only the kinematics of the differential-drive mobile robot, so in future we will include the dynamics as it is well known that due to non-holonomic constraint of the differential-drive mobile robot, the perfect velocity tracking will not hold, we will have to consider the torque as well.

REFERENCES

- [1] A. Das, R. Fierro, V. Kumar, J. Ostrowski, J. Spletzer and C. Taylor, "A Vision-Based Formation Control Framework", *IEEE Transactions on Robotics and Automation*, Vol. 18, No. 5, October 2002.
- [2] J. P. Desai, J. P. Ostrowski, and V. Kumar, "Controlling formations of multiple mobile robots," in *Proceedings 1998 IEEE International Conference on Robotics and Automation*, Leuven, Belgium, May 1998, pp. 2864–2869.
- [3] O. Khatib, "Real-Time Obstacle Avoidance for Manipulators and Mobile Robots", *In Proceedings 1985 IEEE International Conference on Robotics and Automation*, March 1985, pp. 500-505.
- [4] Y. Koren and J. Borenstein, "Potential field methods and their inherent limitations for mobile robot navigation", *IEEE International Conference*

2nd Annual International Conference IEMCON 2012

- on Robotics and Automation*, Sacramento, CA, Apr. 1991, pp. 1398–1404..
- [5] R. Fierro, and F.L. Lewis, "Control of a Nonholonomic Mobile Robot: Backstepping Kinematics Into Dynamics", *Proc. IEEE Conference on Decision Control.*, Kobe, Japan, 1996, pp. 1722-1727.
 - [6] T. Dierks and S. Jagannathan, "Control of Nonholonomic Mobile Robot Formations: Backstepping Kinematics into Dynamics", *IEEE International Conference on Control Applications*, Singapore, October 2007, pp. 94-99.
 - [7] M. Zang, Y. Shen and Q. Wang, "Dynamic artificial potential field based multi-robot formation control", *IEEE International conference Instrumentation and Measurement Technology*, Austin, Tx, May 2010, pp. 1530-1534.
 - [8] M. Egerstedt and X. Hu, "Formation Constrained Multi-Agent Control", *IEEE Transactions on Robotics and Automation*, Vol. 17, No. 6, December 2001.
 - [9] V. Gazi and R. Ordóñez, "Target Tracking using Artificial Potentials and Sliding Mode Control", *International Journal of Control*, Vol. 80, No. 10, October 2007, pp. 1626-1635.
 - [10] Y. Bin-qiang, Z. Ming-fu and W. Yi, "Research of path planning method for mobile robot based on artificial potential field", *International Conference on Multimedia Technology*, Hangzhou, July 2011, pp. 3192-3195. .
 - [11] C. Ke-Cai, G. Xiang and H. Yang, "Formation Control of Multiple Nonholonomic Mobile Robots", *International Conference on Information Science and Technology*, Jiangsu, China, March 2011, pp. 1038-1042.
 - [12] E. G. Hernández-Martínez and E. Aranda-Bricaire, "Convergence and Collision Avoidance in Formation Control: A Survey of the Artificial Potential Functions Approach", *Multi-Agent Systems - Modeling, Control, Programming, Simulations and Applications*, ISBN 978-953-307-174-9, April 2011, pp. 103-126.
 - [13] X. Li and J. Xiao, "Robot Formation Control in Leader-Follower Motion Using Direct Lyapunov Method", *International Journal of intelligent Control and Systems*, Vol. 10, No. 3, September 2005,244-250.
 - [14] Instruction manuals of MATLAB 7.5 from www.mathworks.com .

Benefit Based Artificial Intelligence Regulated Stability Governance of a Practical Power Network

R. Kuthial, S. Bhattacharya, P. P. Das

EE Department, Techno India
Kolkata, India

rumna90@gmail.com, sappy.sls@gmail.com

S. Sen, S.Chanda

EE Department, Techno India
Kolkata, India

sen_sawan@rediffmail.com, sandipee1978@gmail.com

Abstract—FACTS devices are of prime importance in the spheres of stability governance which is essential to realize the complete utilization of the full potential of the transmission network. SVC, one such FACTS device can be systematically used for constant sustenance of system voltage profile. In the present scenario due to existing socio-economic reasons the power networks operate under optimum stress. Under this condition a further increase in stress can cause imminent dipping of system voltage profile. This paper utilizes a standard technique for the recognition of the weakest bus in a practical system and employs the SVC device for the improvement of the voltage profile of the same. The work depicted in this paper enlightens the fact that irrespective of stress in the network, SVC regulates the voltage profile within its linear operating zone, thereby accentuating the flexibility of operation. The main objective of the paper is to restore the bus voltage profile with the union of SVC by classical method and its subsequent corroboration with another methodology based on ANN model (Elman Network). Here an Artificial Neural Network (ANN) regulated SVC module insulates the system from voltage instability and precisely regulates the firing angles of the SVC employed. The promising results not only establish the necessity of the proposed methodology ahead of classical method but also improve the flexibility of the system in stressed condition. In this regard, a practical system (Eastern Grid of India) has been used as a test-bed to demonstrate the applicability of the proposed algorithm with prospective results.

Keywords-Artificial Neural Network; loading stress; Static Var Compensator; voltage profile

I. INTRODUCTION

Due to uncontrollable growth of electrical power, the existing power networks are operating at optimum stress. This excessive loading is accountable for dipping the voltage profile of almost every electrical network, which further trims down the maximum power transfer capability of lines, stability margin of the network and efficiency of transmission. For optimum utilization of the power network, minimization of this unwanted sag of voltage profile is of utmost importance [1].

FACTS controllers are able to change the network parameters in a fast and effective way in order to achieve better system performance [2-5]. Most of these FACTS devices modify the line reactance to alter power flow in such a way that the disturbances incurring in the network can be avoided by achieving a new stable operating point. FACTS devices were used to enhance the Optimal Power Flow [6] by redistributing

real and reactive power using the customized Security Constrained Optimal Power Flow (SCOPF) and Interline Power Flow Controller (IPFC) [7]. However it did not guarantee the voltage stability and security of the existing network. There is also environmental concern regarding the electrical and magnetic fields surrounding the overhead lines and AC cables. These limitations can be eliminated if direct current (DC) transmission is used, which concurrently improves transmission capabilities, has lower losses and the transmission lengths are practically unlimited due to the elimination of capacitive currents [8], [9]. Moreover, during contingency or fault, HVDC can insulate two interconnected systems from each other, thereby improving reliability of the entire system. Though HVDC has numerous technical advantages over FACTS but its application is limited owing to the high investment cost associated with it. But, at present, deregulation tendency is driving conventional utilities to be more sensitive to new investment schemes like FACTS which will harness better real time operating condition at a cost certainly lesser than HVDC. Although one of the main purposes of deregulation is to promote efficient use of cheap generation, there are often constraints on the power transfer through transmission system, causing subsequent overloading or congestion. It can also be caused by different level of contingencies and in this situation, the expenses rise and the system runs in a sub-optimal mode in terms of power system economics [10-12]. This problem can be dealt through generation rescheduling and load curtailment but these lead to excess loss revenue and unnecessary rise of rate of electricity in the power market.

Increased demand for power not only increases stress but also gives rise to line congestion in the power system. FACTS devices were used for the purpose of eliminating the possibility of congestion [13]. Stability Constrained OPF auction model [14] allowed for the inclusion of FACTS devices and their dynamic effects on transmission system congestion management and pricing. FACTS controllers have been recognized as one of the cost effective solutions for network congestion management [15]. However, the voltage profile improvement of the lines was not inspected and line flow didn't improve significantly.

Moreover, the laborious and time consuming calculations of the classical methods have paved way for a less cumbersome and hassle free technique, such as any soft computing based

artificial intelligence algorithms to evaluate the similar parameters as enunciated by the classical analysis. Artificial Neural Networks (ANN) have attracted great deal of attention because of their pattern recognition capability and their ability to handle corrupt data. They have been effectively applied in certain power engineering problems [16], [17]. ANN has the ability to solve problems with high non-linearity such as system security and stability assessment [18], [19]. Once the ANN based model was programmed with training set of input-output patterns, from which, it learnt the relationship between them, it was able to calculate the laborious and time-consuming calculations in a very smooth way. But, its ability to perform is well affected by the chosen training data as well as network topology and training scheme.

Hence, it can be inferred that to utilize the complete potential of any transmission system the FACTS devices must be incorporated into the system with utmost accuracy to manage congestion as well as improve line flow and voltage profile. In this paper, a methodology has been proposed which employs SVC, one of the prominent FACTS devices to mitigate unwanted voltage drops incurring in various points of a practical multi-bus power system network, both using classical method and a trained ANN model. Determination of weakest bus has been done using standard techniques and the proposed program changes the control parameter of SVC to decrease congestion by sustaining the voltage profile. Throughout the course of this work, a practical 203 bus power network (Eastern Grid of India) has been used to highlight the efficacy of the proposed methodology.

II. THEORY

The proposed methodology rests on proper formulation of objective function for the improvement of voltage profile using SVC. The deterministic algorithm has been framed to calculate voltage profile of the concerned network both in presence and absence of SVC in the weakest bus. The objective function for the concerned problem has been formulated as below and the equality, inequality and security constraints have been specified.

A. Problem Statement

The objective function for the proposed algorithm is given as:

$$\text{Minimize: } \sum_{i=1}^n (V_i - V_{ij}) \quad (1)$$

where, n = No. of load buses in the system, V_{ij} = Bus voltage magnitude in stressed condition, V_i = Bus voltage in normal condition

1. Equality or power balance constraints:

$$P_{Gi} - P_{Di} - V_i \sum_{j=1}^n V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) = 0 \quad (2)$$

$$Q_{Gi} - Q_{Di} - V_i \sum_{j=1}^n V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) = 0 \quad (3)$$

P_{Gi}, Q_{Gi} = Active and reactive power injected in bus i , P_{Di}, Q_{Di} = Active and reactive power demand on bus i , G_{ij} = Conductance of transmission line from bus i to j , B_{ij} = Susceptance of transmission line from bus i to j .

2. Inequality or generator output constraints:

$$P_{gi}^{\min} \leq P_{gi} \leq P_{gi}^{\max} \quad (4)$$

$$Q_{gi}^{\min} \leq Q_{gi} \leq Q_{gi}^{\max} \quad (5)$$

P_{gi}, Q_{gi} = Active and reactive power of generator i respectively, $P_{gi}^{\min}, Q_{gi}^{\min}$ = Lower limit of active and reactive power of the generators, $P_{gi}^{\max}, Q_{gi}^{\max}$ = Upper limit of active and reactive power of the generators

3. Voltage constraint:

$$V_i^{\min} \leq V_i \leq V_i^{\max} \quad (6)$$

V_i^{\max}, V_i^{\min} = Upper and lower limits of V_i

4. Transmission constraint:

$$P_{ij\min} \leq P_{ij} \leq P_{ij\max} \quad (7)$$

$P_{ij\max}, P_{ij\min}$ = Maximum and minimum line flow limits of P_{ij}

5. Cost constraint:

$$C_i = AP_{gi}^2 + BP_{gi} + C \quad (8)$$

A, B, C = Cost co-efficient of generators, P_{gi} = Generation of i^{th} generator in MW

B. SVC Modeling

Due to the limitations imposed by constraints on the operation of ac power transmission lines, they may not be able to deliver the power demand. FACTS is a modern technology and its key role is to enhance controllability and power transfer capability in an ac system. SVC is a FACTS device and it is extensively accepted for reactive power control in transmission and distribution system. SVC has principal role in voltage control of the buses. A FC-TCR type Static Var Compensator consists of one fixed capacitor bank in parallel with one thyristor-controlled reactor in each phase. The SVC structure shown in Fig. 1 is used to derive a SVC model that considers the TCR firing angle α as state variable.

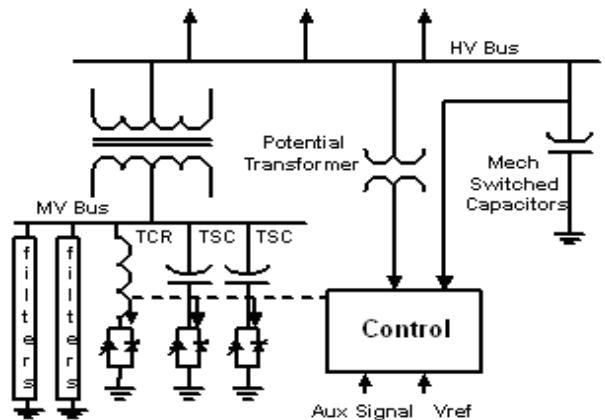


Figure 1. Layout of a typical Static Var System

Depending on the nature of equivalent SVC's reactance (capacitive or inductive) the SVC draws either capacitive or inductive current. For a FC-TCR type SVC, the variable TCR equivalent reactance $X_{L_{eq}}$, at fundamental frequency is given by,

$$X_{L_{eq}} = X_L \frac{\pi}{2(\pi - \alpha) + \sin(2\alpha)} \quad (9)$$

where α is the thyristor firing angle.

The SVC reactance effective reactance X_{eq} is determined by the parallel combination of X_C and $X_{L_{eq}}$, as given by (10),

$$X_{eq} = \frac{X_C X_L}{\frac{X_C}{\pi}(2\pi - 2\alpha + \sin 2\alpha) + X_L} \quad (10)$$

The SVC is represented by the following equation with the firing angle considered as state variable.

$$\begin{bmatrix} \Delta P_k \\ \Delta Q_k \end{bmatrix}^i = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\partial Q_k}{\partial \alpha} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \frac{\partial Q_k}{\partial \alpha} \end{bmatrix}^i \quad (11)$$

$$\text{where, } \frac{\partial Q_k}{\partial \alpha} = \frac{2V_k^2}{X_L} (\cos(2\alpha) - 1) \quad (12)$$

At the end of the iteration i , the variable firing angle α is updated according to the following equation,

$$\alpha^{i+1} = \alpha^i + \Delta\alpha^i \quad (13)$$

Hence the admittance matrix $[Y_{bus}]$ of the whole network may be customized using (11), (12) and (13) and accordingly voltages of buses will be modified.

With a view to support the novelty of the classical analysis with another technique that judges voltage dip in a practical environment and also subsequently rectifies it, an ANN model has been proposed. With its fast response characteristics the ANN model scores largely over classical analysis by being far less cumbersome. Associated theory related to the proposed methodology is given as follows.

C. ANN Model

In recent years, substantial effort has been devoted for development of a real time technique, for voltage stability assessment with high speed and precision. For voltage stability analysis, a number of ANN algorithms have been projected in different research papers [20], [21]. An ANN is composed of several processing units, also identified as artificial neurons, and interconnected in an encoded mode to accomplish a desired pattern recognition task. A neural network can execute a particular function by adjusting the values of connections (weights) between processing elements. Conventional computer programs have large number of instructions, and they operate in a sequential mode to execute one instruction after another on a computer. In comparison, ANN can perform massive parallel operations and nonlinear computations in short duration. ANN has found many applications in power

systems, especially in the areas where pattern classification is necessary, based on historical examples. Once trained, the neural network is able to provide sufficiently accurate result for on-line application. A supervised artificial neural network has been used in this paper, which requires a number of input-target pairs to train the network. The Elman networks are two layer back propagation networks, with the addition of feedback connection from output of hidden layer to its input layer. This recurrent connection permits the network to detect and generate time varying patterns. The amalgamation with two layer (hidden layer and output layer) networks and transfer functions can approximate any function (with a finite number of discontinuities) with arbitrary accuracy. The delay in the recurrent connection stores values from the previous time step. Thus, even if two Elman networks, with same weights and biases are given identical inputs at a prearranged time step, their outputs can be dissimilar due to different feedback states, because the network can amass information for future orientation. The basic model of Elman network is shown in following figure.

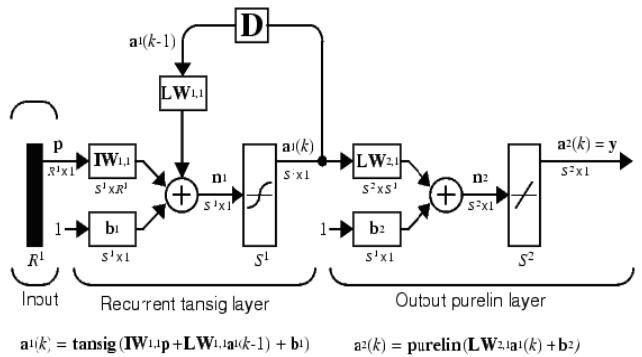


Figure 2. Basic model of Elman network

III. SIMULATION

In this paper, the efficacy of the proposed methodology has been effectively demonstrated on a 203 bus-265 lines practical power network (Eastern Grid of India). Table I is the system description table which portrays the various parameters of the concerned network.

TABLE I. SYSTEM DESCRIPTION

| | |
|---------------------|---------|
| No. of buses | 203 |
| No. of lines | 265 |
| Total demand (MW) | 7619 |
| Total demand (MVAR) | 4808.35 |
| No. of generators | 24 |

A diagrammatic representation of the scheme of operation employed in this regard is given as follows.

TABLE II.

EFFECTIVENESS OF SVC IN RESTORING THE BUS VOLTAGE PROFILE UNDER STRESSED CONDITION FOR PRACTICAL SYSTEM

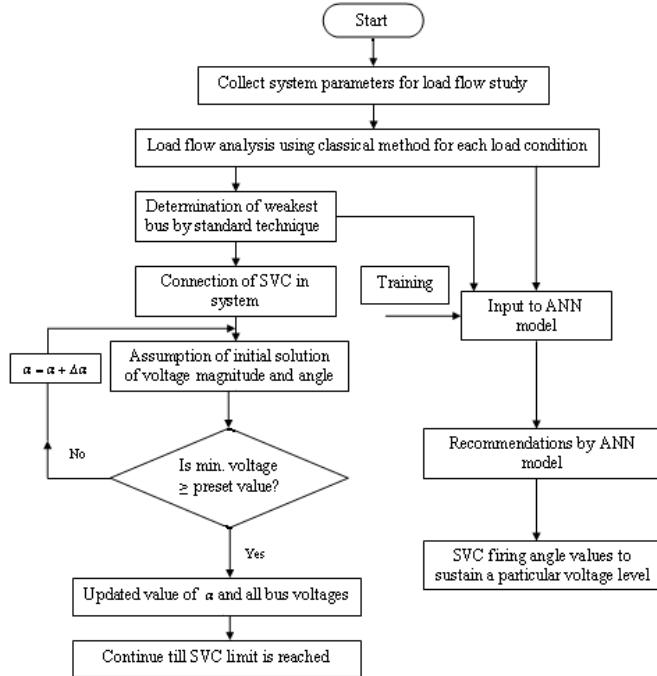


Figure 3. Proposed Methodology

On commencement of the methodology, the foremost step determines the weakest bus of the system using the assistance of a reliable technique. From Figure 4, bus number 168 has been found as the weakest bus followed by bus 169 and 173. Thus SVC incorporation has been done on these three buses for the increase in demand.

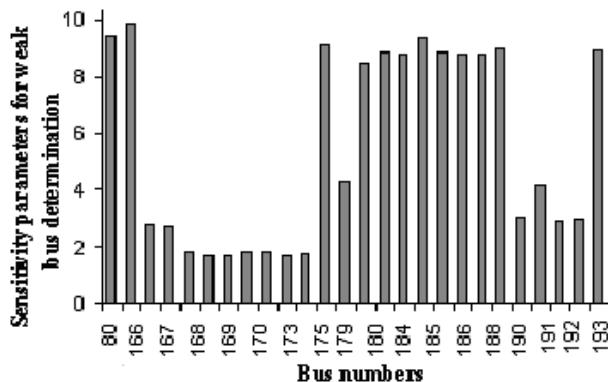


Figure 4. Weak bus determination

Consequent connection of the SVC to the weakest bus of the system causes recognizable improvement in the voltage profile of the network, as has been depicted in the following Table II. The analysis of these tables successfully highlights the self healing propensity that sets in within the network on connection of SVC, as long as the operational conditions are constrained within a bounded zone of two predetermined extremes.

As evident from observations enlisted, marked improvement is obtained in the degree of sustenance of the voltage profile which appreciably ameliorates the dynamic stability of the system. This in turn progressively leads to increased efficiency of operation in the consumer bank to which the considered system is delivering power thus addressing all major concerns in the spectrum of discussion.

In the next step of the simulation, the ANN based model has been trained with a large number of data sets of input training vector to meet convergence criterion. The convergence criterion for the application has been achieved in 5000 iterations respectively with a training goal of 0.0001. All the other parameters of the proposed model have been tabulated in Table III.

TABLE III. DESIGN PARAMETER OF NEURAL NETWORK

| Parameter | ANN application |
|-------------------------|-----------------|
| Training Vector | 200 |
| Testing Vector | 50 |
| Input Neurons | 2 |
| Hidden Layers | 1 |
| Output Neurons | 4 |
| Neurons in Hidden Layer | 100 |

Results obtained after employment of this Artificial Intelligence technique was tabulated and compared to those obtained from the previous classical analysis. As a subsequent step, Table IV correlates the same network analysis when the SVC remains connected to the weakest bus.

2nd Annual International Conference IEMCON 2012

TABLE IV. VERIFICATION USING ANN

| Classical Calculation | | | | ANN | |
|-----------------------|-------------------------------|----------------------------------|----------------------------|----------------------------------|----------------------------|
| Bus No. | Loading Stress applied (p.u.) | Min. bus voltage with SVC (p.u.) | Firing angle of SVC (rad.) | Min. bus voltage with SVC (p.u.) | Firing angle of SVC (rad.) |
| 168 | 0.200 | 0.9400 | 2.1508 | 0.9321 | 2.143 |
| | 0.380 | 0.9400 | 2.2128 | 0.9392 | 2.212 |
| | 0.395 | 0.9401 | 2.2302 | 0.9371 | 2.230 |
| | 0.410 | 0.9400 | 2.2311 | 0.9390 | 2.251 |
| | 0.500 | 0.9400 | 2.2514 | 0.9390 | 2.211 |
| 169 | 0.183 | 0.9400 | 2.1508 | 0.9370 | 2.130 |
| | 0.200 | 0.9400 | 2.1652 | 0.9380 | 2.315 |
| | 0.230 | 0.9429 | 2.4302 | 0.9389 | 2.460 |
| | 0.245 | 0.9463 | 2.5451 | 0.9373 | 2.495 |
| | 0.305 | 0.9537 | 2.6786 | 0.9487 | 2.568 |
| 173 | 0.150 | 0.9400 | 2.1508 | 0.9380 | 2.170 |
| | 0.200 | 0.9400 | 2.1799 | 0.9390 | 2.169 |
| | 0.240 | 0.9463 | 2.5452 | 0.9420 | 2.545 |
| | 0.320 | 0.9537 | 2.6787 | 0.9497 | 2.578 |
| | 0.325 | 0.9537 | 2.6787 | 0.9497 | 2.578 |

Here also it is evident from the table that the classical analysis and the neural network analysis agree to a high degree with a few minor ignorable departures from the acceptable values as procured from the realms of classical analysis. However, the spectrum of advantages of the proposed ANN significantly scores over the negligible inaccuracy in its performance. The successful agreement of this aforesaid neural network methodology with classical results envisages the obvious practicability of the neural network in shunning out cumbersome classical methods to the periphery and bringing about relatively simpler techniques for network analysis to the fore.

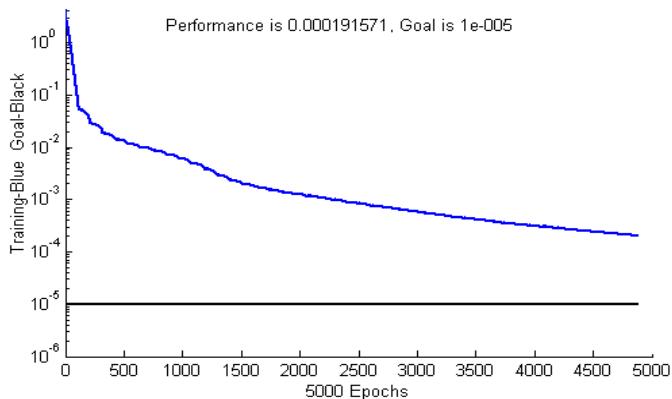


Figure 5. Training performance of proposed ANN based model

IV. CONCLUSION

For upholding the proficient operational condition, voltage profile of the system network should remain unaltered during excess loading. The work presented in this paper has employed a FACTS device to regulate the voltage profile in inadvertent states of the system. This facilitation of improved power flow through the transmission corridors of the network paves the way for management of line congestion, alleviation of undesirable power losses, increased efficiency of operation and therefore, a marked improvement in the system security. Also the novelty of using an artificially intelligent voltage restoration algorithm in carrying out all encompassing system surveillance is very effectively highlighted in this paper. Thus the ability of a system to heal itself in a stressed environment using the concept of reactive power balancing by an artificially intelligent SVC module is the high watermark of this paper. Even though both classical method and the ANN model based methodology serves the same purpose, the artificially intelligent voltage profile improvement algorithm has its novelty ingrained in its comparatively faster response characteristics and its high extent of accuracy of output result. If this methodology is adopted by the system operator to control the system parameters, the enhancement of operating conditions in terms of stability margin can also be brought about by the assurance of constant voltage profile.

REFERENCES

- [1] Sawan Sen, S. Sengupta, S. Chanda, A. Chakrabarti, "Voltage Profile and Loss Assessment of a Power Network under Stressed and Contingent Condition using TCSC and HVDC Link in a Deregulated Environment," International Journal of Engineering Research and Technology, vol. 3, no. 3, pp. 467-480, 2010.
- [2] Hingorani, N.G. and Gyugyi, L, Understanding FACTS - Concept and Technology of flexible ac transmission systems, 1999 IEEE press.
- [3] H. Ambriz-Perez, E. Ache, C.R. Fuerte Esquivel, " Advanced SVC model for N-R Load Flow and Newton Optimal Power Flow Studies, " IEEE Transaction on Power Systems, vol. 15, no. 1, pp. 129-136 2000.
- [4] M. Moghavemi, M. O. Faruque, "Effects of FACTS devices on static voltage stability, " in the Proceedings of TENCON, pp. 356-362, 2000.
- [5] A.Kazami, B. Badrzadeh, "Modeling and Simulation of SVC and TCSC to Study their Limits on Maximum Loadability point," Electrical Power and Energy System, vol. 26, pp 381-388, 2004.
- [6] A. Berizzi, M. Delfanti, P. Marannino, M. Savino Pasqua, A. Silvestrion, "Enhanced Security-Constrained OPF with FACTS Devices" IEEE Transaction on Power Systems, vol. 20, no. 3, 2005.
- [7] J. Zhang, A. Yokoyama, "Optimal Power Flow Control for Congestion Management by Interline Power Flow Controller (IPFC)," in the proceedings of International Conference on Power System Technology, 2006.
- [8] M. H Haque, "Determination of Steady State Voltage Stability Limit of a Power System in the Presence of SVC," Power Tech Proceedings, vol. 2, 2001.
- [9] A. Gustafsson, M. Jeroense and J. Karlstrand, "Energy Efficiency with HVDC Technology," ABB Power Systems, Raleigh, North Carolina.
- [10] M Shaaban, Yixin Ni, Felix F Wu, "Transfer capability computations in deregulated power system," Proceedings of the 33rd Hawaii International Conference on System Sciences, vol. 4, 2000.
- [11] "Transmission Power Flow Framework," PSERC Project Report.
- [12] Thomas Tarnowski, " FACTS Devices Optimizing an Electrical Energy Market Subject to Power Transmission Constraints," IEEE Transaction on Power Systems, vol. 15, no. 2, pp. 129-136 2003.
- [13] L. Yao, P. Cartwright, L. Schmitt, Xiao-Ping Zhang "Congestion Management of Transmission Systems using FACTS," IEEE/PES

2nd Annual International Conference IEMCON 2012

- Transmission and Distribution Conference & Exhibition: Asia and Pacific Dalian, China, 2005.
- [14] Claudio A. Cañizares, Sameh K. M. Kodsi, "Dynamic Versus Steady-State Modeling of FACTS Controllers in Transmission Congestion" Power Tech Proceedings, vol. 2, 2002.
- [15] X. P. Zhang, B. Chong, K. R. Godfrey, L. Yao, M. Bazargan, L. Schmitt, " Management of Congestion Costs Utilizing FACTS Controllers in a Bilateral Electricity Market Environment ;" in the proceedings of Power-Tech, 2007.
- [16] S.Chakrabarty, B. Jeyasurya, "Online Voltage Stability Monitoring using Artificial Neural Network," *Proceedings of Power Engineering, LESCOPE-04 Large Engineering System Conference*, pp-71-75, 2004.
- [17] C. A. Jimenez, C. A. Castro, "Voltage Stability Security Margin Assessment via Artificial Neural Network," *Proceedings of Power Tech, IEEE Russia*, pp-27-30, June 2005.
- [18] H. B. Wan, Y. H. Song, A. T. Johns, "Artificial Neural Networks for Power System Voltage Stability Assessment," *Proceedings of UPEC*, vol: 1, 1995.
- [19] B. Suthar, R. Balasubramanian, "Application of an ANN Based Voltage Stability Assessment Tool to Restructured Power Systems," *Proceedings of Conference on Power Bulk Power System Dynamics and Control – VII*, pp-1 – 8, Aug. 2007
- [20] Vidya Sagar, S. Vankayala and Nutakki D. Rao: Artificial Neural Networks and its Applications to power systems- a bibliographical survey- Electric Power System Research 28, 1993, pp- 67-79.
- [21] S. Kamalasadan, A. K. Srivastava, D. Thukaram, "Novel algorithm for online stability assesment based on feed forward neural network", Power Engineering Society General Meeting, 2006, IEEE.

Artificial Neural Network Modeling of Intrusion Detection and Prevention System

Tanusree Chatterjee

Dept. of Computer Science & Engg.
IEM, Kolkata
tnsr.chatterjee@gmail.com

Indraneel Mukhopadhyay

Dept. of Information Technology
IEM, Kolkata
imukhopadhyay@gmail.com

Mohuya Chakraborty

Dept. of information Technology
IEM, Kolkata
mohuyacb@yahoo.co.in

Abstract — With the rapid proliferation of computer networks during the past decade, security has become a crucial issue for computer systems. Different soft computing based methods have been proposed in recent years for the development of intrusion detection systems. However due to the vulnerabilities of the networks towards attacks and to keep our sensitive data safe from attacks, detection is not enough, rather we have to prevent them. In this paper, we present an artificial neural network based model of Intrusion Detection and Prevention System (IDPS), which not only detects different network attacks but also prevents them from being forwarded. We have used KDDCUP'99 dataset for classifying five categories of attacks and normal data. Our simulation test results in MATLAB 7.11.0.584(R2010b) and SIMULINK are quite satisfactory.

Keywords- Artificial Neural Network, Back Propagation, Intrusion Detection and Prevention, Network Attack, Simulink.

I INTRODUCTION

In the last three years, the networking revolution has finally come of age. More than ever before, we see that the Internet is changing computing, as we know it. The possibilities and opportunities are limitless; unfortunately, so too are the risks and chances of malicious intrusions. However, we can try to detect these intrusion attempts so that action may be taken to repair the damage later. Intrusion detection is the process of monitoring the events occurring in a computer system or network and analyzing them for signs of possible incidents, which are violations of computer security policies or imminent threats to the standard security practices.

Intrusion prevention is the process of performing intrusion detection and attempting to stop detected possible incidents. Intrusion detection and prevention systems (IDPS) [1] are primarily focused on identifying possible incidents, logging information about them, attempting to stop them, and reporting them to security administrators.

Earlier we have worked on Back Propagation ANN approach towards classification of different attacks and have found out the four main attack categories [2]:

- DOS (Denial of Service):** An attacker tries to prevent legitimate users from using a service e.g. TCP SYN Flood, Smurf etc.
- Probe:** An attacker tries to find information about the target host e.g. scanning the target host in order to get information about available resources.
- U2R (User to Root):** An attacker has local account on victim's host and tries to get the root's privileges.
- R2L (Remote to Local):** An attacker does not have local account on the victim's machine and tries to obtain it.
- Other:** Unknown attacks that are yet to be classified.

We have designed a two layer feed forward Artificial Neural Network (ANN) with sigmoid hidden and output neurons. The hidden layer contains 12 neurons and output layer contains 6 neurons. We have conducted a machine learning using the Neural Network Pattern Recognition Tool of MATLAB 7.11.0.584 (R2010b) to train our neural network (shown in Figure 1) by using back propagation scaled conjugate gradient training algorithm [3] where we have obtained a mean squared error of about 10^{-2} and best validation performance of 0.021587 at epoch 33 for the dataset (shown in Figure 2). After training the ANN we generated the basic Simulink model of our neural network based attack classifier. Finally we have designed and tested the model of IDPS in Simulink by using all the 41 features of KDDCUP'99 dataset of Intrusion Detection [4].

The organization of the paper is as follows. After the introduction in section I we discuss about the ANN attack classifier model in section II. Section III provides an overview of our designed model of IDPS using Simulink. In section IV the performance analysis of our designed model is discussed. Finally we conclude the paper in section V with some highlights on future work.

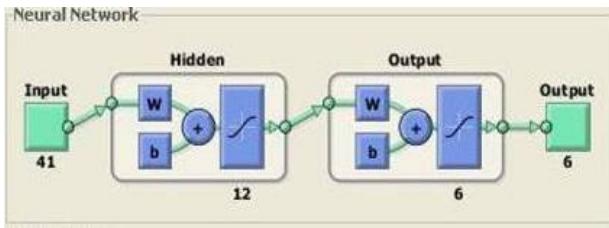


Figure 1 Artificial Neural Network Block Diagram

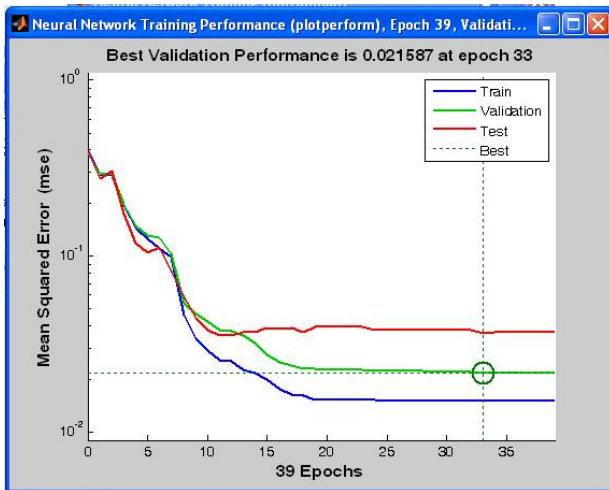


Figure 2 Least Mean Squared Error

II ANN MODEL FOR ATTACK CLASSIFIER

Simulink is an environment for multi-domain simulation and Model-Based Design for dynamic and embedded systems. It provides an interactive graphical environment and a customizable set of block libraries that allows to design, simulate, implement, and test a variety of time-varying systems, including communication, control, signal processing, video processing, image processing etc [5].

During training of the ANN we have specified some values to detect attacks e.g., Normal 0; DoS 1; Probe 2; U2R 3; R2L 4; Other 5.

After training our ANN with the given dataset we generate the Simulink ANN Attack classifier model (shown in Figure 3) to classify the attack and the normal data separately.

Here the input signature consists of the 41 attributes of our dataset and the IDPS block output is the classified attack or normal.

The ANN Attack Classifier Block can be further subdivided as shown in figure 3. The different blocks are Process Input, Layer 1, Layer 2 and Process Output. The output $\{a\}$ of hidden layer i.e. Layer 1 is fed to output layer i.e. Layer 2 as input. The internal functionalities that take place in the neurons of the layer 1 are shown in Figure 4.

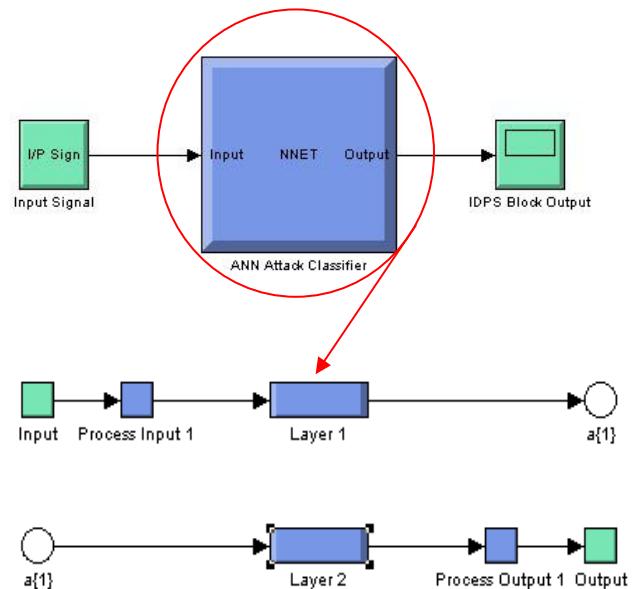


Figure 3 Simulink Model of ANN Attack Classifier

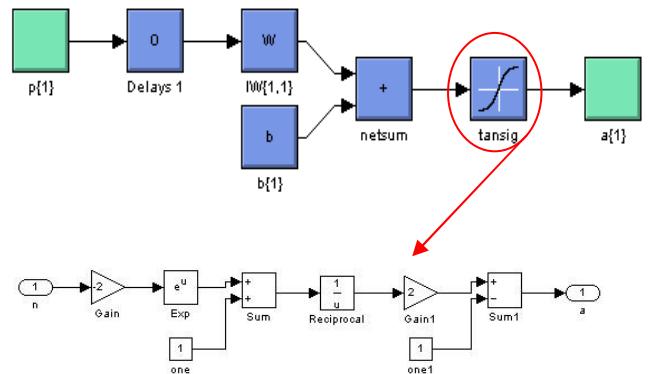


Figure 4 Layer 1 Function

Here the input vector $p[1]$ after being delayed is multiplied by the weights w of all the neurons of input layer to form the product that becomes the total weighted input IW . The neuron has a bias b , which is summed with the weighted inputs to form the net input $netsum$ as given by equation (1).

$$\begin{aligned} n &= w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b \\ &= W \cdot p + b \end{aligned} \quad (1)$$

The network contains delays, so the input to the network would normally be a sequence of input vectors that occur in a certain time order. Finally the net input is passed to the transfer function $tansig$, which produces the output vector a as given by equation (2).

$$a = f(Wp + b) \quad (2)$$

The sigmoid transfer function shown in the figure takes the input, which can have any value between plus and minus infinity, and squashes the output into the range 0 to 1. The output a will be the input to the layer 2 i.e. Output layer. There would be 41 weight blocks of Layer 1 (like the 6 blocks of Layer 2 shown in Figure 5). The inner functionalities of layer 2 are shown in Figure 5 that is self explanatory.

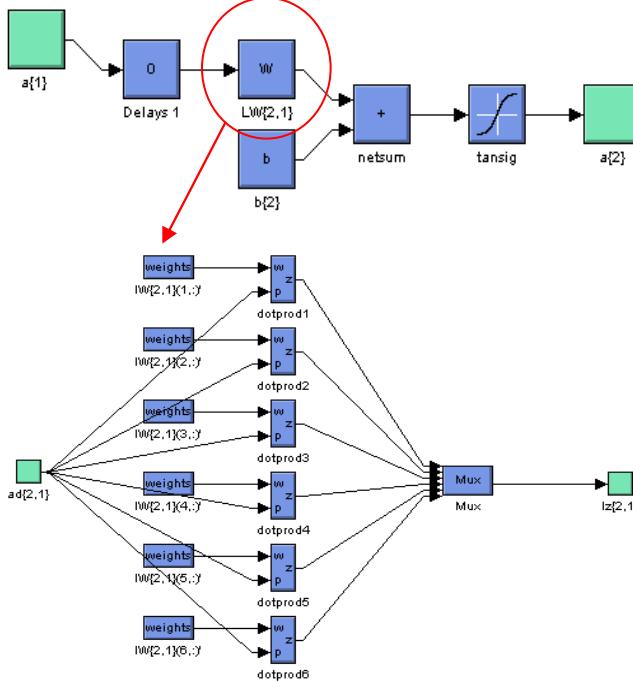


Figure 5 of Layer 2 Functions

III OVERVIEW OF SIMULINK MODEL OF IDPS

Now we proceed a step further in the Simulink model where we have used Embedded MATLAB function to allow only the normal data to propagate and will stop the attacks from propagating through the network to create the prevention mechanism shown in the Figure 6. The Intrusion Detection & Prevention Subsystem contains the embedded MATLAB code to detect the attack data as well as prevent them to be forwarded to the output port *out1* [6].

The IDS output block showing 1 in the second position with others 0 indicates that DoS attack has been detected as discussed in section II. If the input signature happens to be an attack the subsystem prevents them from being forwarded as shown by the IDPS output [7].

The same Simulink model was tested using a Normal data signature that allowed the input signature to be forwarded to the output as shown in Figure 7.

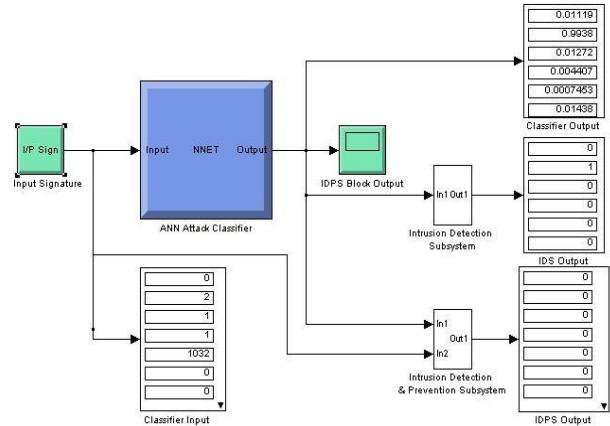


Figure 6 Simulink model of IDPS showing prevention of DoS attack

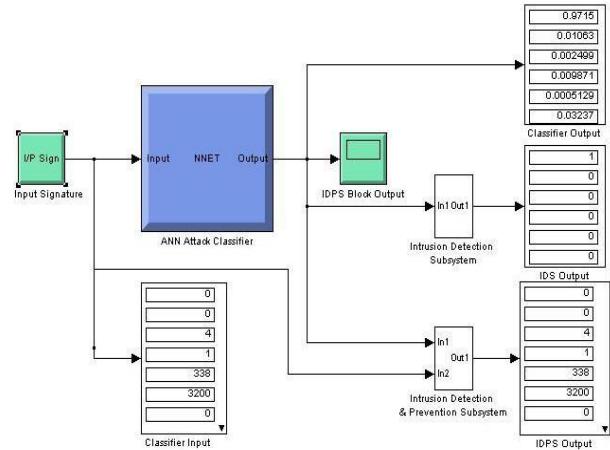


Figure 7 Simulink model of IDPS allowing normal data to pass

V PERFORMANCE ANALYSIS OF IDPS

The performance of the system was found to be satisfactory. We had tested the system taking 25% of the data from the dataset. Figure 8 shows the False Positive Graph taking into account number of Input Signature (Normal) vs. Number of False positive generated [8].

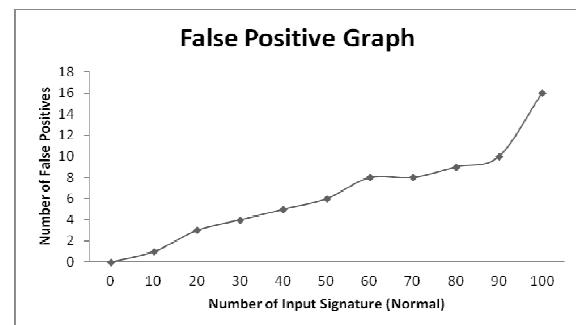


Figure 8 False Positive Graph

VI CONCLUSION

In this work we have designed an ANN based IDPS using MATLAB 7.11.0.584 (R2010b) and Simulink. The result of our tests is very encouraging. Our further objective is to implement the designed model on FPGA (Spartan 3E). This implementation aims to contribute in hardware integration solutions in the areas of monitoring, diagnosis and management of network intrusions. Since the Simulink library provided by Xilinx, has all the blocks that are necessary for the design of Artificial Neural Networks except a few functions such as sigmoid functions. In our work an approximation of the sigmoid functions will be proposed and which will be implemented on FPGA using Xilinx Library.

REFERENCES

- [1] "Guide to Intrusion and Prevention System", csrc.nist.gov/publications/nistpubs/800-94/SP800-94.pdf
- [2] I.Mukhopadhyay M.Chakrabarty, S.Chakrabarti T.Chatterjee, "Back Propagation Neural Network Approach to Intrusion Detection System", In Proc. IEEE Conference on Recent Trend in Information System (ReTIS - 2011).
- [3] "Neural Network ToolBox - MathWorks", www.mathworks.com/help/pdf_doc/nnet.pdf
- [4] Results of the DARPA 1998 Offline Intrusion Detection Evaluation, Richard P. Lippmann, Robert K. Cunningham, David J. Fried, Isaac Graf, Kris R. Kendall, Seth E. Webster, Marc A. Zissman.
- [5] "Getting Started with Simulink - MathWorks", www.mathworks.com/help/pdf_doc/simulink.pdf
- [6] Mehdi Moradi and Mohammad Zulkernine, "Neural Network Based System for Intrusion Detection and Classification of Attacks", 148-04.
- [7] I Mukhopadhyay et.al, "HawkEye Solution" ICWET '11 Proceedings of the International Conference & Workshop on Emerging Trends in Technology, pages 252-257, ISBN: 978-1-4503-0449-8
- [8] Ray-I Chang, Liang-Bin Lai, Wen-De S, Jen-Chieh Wang, Jen-Shiang Kouh "Intrusion Detection by Back propagation Neural Networks with Sample-Query and Attribute-Query", International Journal of Computational Intelligence Research, ISSN 0973-1873 Vol.3, No. 1 (2007), pp. 6-10.

INTRUSION DETECTION IN WIRELESS SENSOR NETWORKS USING FUZZY LOGIC

Moutushi Singh¹, Abdur Rahaman Sardar², Subir Kumar Sarkar³

¹ moutushi_singh@rediffmail.com Department of Information Technology, Institute of Engineering and Management

² abdur.sardar@gmail.com Department of Computer Science and Engineering, Institute of Technology and Marine Engineering

³ su_sircar@gmail.com Department of Electronics and Telecommunication, Jadavpur University

Abstract— Wireless Sensor Networks (WSNs) detect and report interesting events when they occur in the target region. These networks are vulnerable to security breach due to wireless communication and lack of infrastructure. In sinkhole attacks, an attacker attracts network traffic by forging or replaying routing messages through compromised nodes. Thus attracted traffic is used for selective forwarding, denial of service (DoS) and false report attacks. Adapting various sensor devices to communicate within sensor networks empowers us by providing range of possibilities. The sensors in sensor networks need to know their measurable belief of trust for efficient and safe communication. In this paper we have proposed the intrusion detection scheme using fuzzy logic for detecting and defending sinkhole and DOS attacks in sensor networks.

Index Terms— Denial of Service (DoS), Fuzzy Logic, Intrusion detection, Wireless Sensor Networks (WSN)

I. INTRODUCTION

In the last decade, wireless sensor networks (WSNs) have been growing rapidly in various applications. Significant effort has been made in both academia and industry to meet the vision of a sensor-rich world. Wireless sensor nodes equipped with sensing, computing, and communication capacities are now available. Typical examples include UC Berkeley's Telos and Mica family, CMU's FireFly, Intel's IMote2, Sun's SPOT, UCLA's Medusa, and MIT's μ AMPS. Commercial sensor node products and solutions are also offered by many vendors, e.g. Crossbow, Rockwell, MicroStrain, Ember, Sentilla, and Dust Networks. While their physical sizes continue to decrease, these sensor node products are becoming cheaper and more powerful than ever. The availability of these products makes it possible to deploy WSNs at a large scale and a low cost that were impractical or even unimaginable just a few years ago. WSNs are typically used for information gathering in applications like habitat monitoring, military surveillance, agriculture and environmental sensing, and health monitoring. The primary functionality of a WSN is to sense and monitor the state of the physical world. In most cases, they are unable to affect the physical environment. However, in many applications, observing the state of the physical system is not sufficient, it is also expected to respond to the sensed

events/data by performing corresponding actions on the system. This stimulates the emergence of wireless sensor/actuator networks (WSANs). Featuring coexistence of sensors and actuators, WSNs enable the application systems to sense, interact, and change the physical world. They can be deployed in lots of applications such as disaster relief, planet exploration, intelligent building, home automation, industrial control, smart spaces, pervasive computing systems, and cyber-physical systems.[1]

Real-world WSN applications have their requirements on the quality of service (QoS). Some QoS metrics may be used to measure the degree of satisfaction of these services. Technically, QoS can usually be characterized by, e.g., delay and jitter, packet loss, deadline miss ratio, and/or network utilization (or throughput) in the context of WSNs. Meeting QoS requirements in WSNs is difficult.[1]

Some major challenges are described as follows.

WSNs are normally resource constrained. Sensor nodes are usually low-cost, low-power, small devices equipped with limited data processing capability, transmission rate, energy, and memory. Due to the limitation in transmission power, the available bandwidth and the radio range of the wireless channel are also limited. While actuator nodes typically have stronger computation and communication capabilities and more energy budget relative to sensors, resource constraints apply to both sensors and actuators.

WSNs are highly dynamic in nature. The network topology may possibly change over time due to node mobility, node failure, node addition, and exhausted battery energy. The channel capacity may also change because of the dynamic adjustment of transmission powers of the sensor/actuator nodes. [1][2]

WSNs feature inherent node heterogeneity. Having different functionality, sensors and actuators do not share the same level of resource constraints. The coexistence of sensors and actuators makes WSNs and WSANs fundamentally distinct.

This paper deals with network security in WSNs. A fuzzy logic control based network security scheme will be developed to facilitate security support in resource-constrained WSNs operating in dynamic and unpredictable environments. This approach is by no means an almighty solution to all of the above challenges; it is, however, the first attempt to explicitly

address the impact of unpredictable variations in traffic load on the security of WSNs. The variability of traffic loads over wireless connections may be a natural result of network topology changes, ambient interferences, and/or system reconfiguration, just to mention a few.[2] [3]

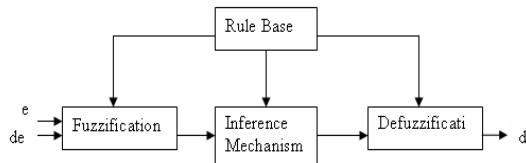


Figure 1. Inner structure of Fuzzy Logic Controller

A fuzzy logic controller is designed to dynamically adjust the sampling period of relevant sensor in a way that the intrusion is kept at a desired level. Taking advantage of the feedback control technology, the FLC-QM can provide QoS guarantees while achieving predictable application performance. This solution is generic, scalable, and easy to implement. It can simultaneously address multiple QoS problems such as delay, packet loss, and network utilization. Simulation results will be given to demonstrate the effectiveness of the proposed FLC-QM scheme.[4][5]

After the introduction in section I, we will describe wireless sensor network in section II & intrusion in section III. Fuzzy logic is discussed in section IV. Section V is on network implementation & section VI is on software implementation. Experimental results are analyzed in section VII.

II. INTRUSION

Sensor networks are particularly vulnerable to several key types of attacks. Attacks can be performed in a variety of ways, most notably as DoS attacks, but also through traffic analysis, privacy violation, physical attacks, and so on. DoS attacks on wireless sensor networks can range from simply jamming the sensor's communication channel to more sophisticated attacks designed to violate the 802.11 MAC protocol or any other layer of the wireless sensor network. Due to the potential asymmetry in power and computational constraints, guarding against a well orchestrated DoS attack on a WSN can be nearly impossible.[1][3] A more powerful node can easily jam a sensor node and effectively prevent the sensor network from performing its intended duty.[1][2][4][6]

A. Denial of Service Attack

The projected use of sensor networks in highly critical and sensitive applications. A sensor network designed to alert building occupants in the event of a fire could be highly susceptible to a denial of service attack. A denial of service attack on such a sensor network could prove very costly. Some of the major types of denial of service attacks are discussed below.

A standard attack on wireless sensor networks is simply to jam a node or set of nodes. Jamming is simply the transmission of a radio signal that interferes with the radio frequencies being used by the sensor network. No messages are able to be sent or received. Attacks can also be made on the link layer itself. One possibility is that an attacker may simply intentionally violate the communication protocol, e.g., ZigBee or IEEE 801.11b (Wi-Fi) protocol, and continually transmit messages in an attempt to generate collisions. Such collisions would require the retransmission of any packet affected by the collision. Using this technique it would be possible for an attacker to simply deplete a sensor node's power supply by forcing too many retransmissions. At the routing layer, a node may take advantage of a multihop network by simply refusing to route messages. This could be done intermittently or constantly with the net result being that any neighbor who routes through the malicious node will be unable to exchange messages with, at least, part of the network.[4] Extensions to this technique include intentionally routing messages to incorrect nodes (misdirection). The transport layer is also susceptible to attack. Flooding can be as simple as sending many connection requests to a susceptible node. Eventually a node's resources will be exhausted, thus rendering the node useless.[5][6][7]

B. The Sybil attack

The Sybil attack relates to WSN, which is defined as a "malicious device illegitimately taking on multiple identities". It was originally described as an attack able to defeat the redundancy mechanisms of distributed data storage systems in peer-to-peer networks. In addition to defeating distributed data storage systems, the Sybil attack is also effective against routing algorithms, data aggregation, voting, fair resource allocation and foiling misbehavior detection. Regardless of the target, the Sybil algorithm functions similarly. All of the techniques involve utilizing multiple identities. To attack the routing protocol, the Sybil attack would rely on a malicious node taking on the identity of multiple nodes, and thus routing multiple paths through a single malicious node.[5][6][8]

C. Traffic Analysis Attacks

Wireless sensor networks are typically composed of many low-power sensors communicating with a few relatively robust and powerful base stations. It is not unusual, for data to be gathered by the individual nodes where it is ultimately routed to the base station. Often, for an adversary to effectively render the network useless, the attacker can simply disable the base station. A *rate monitoring attack* simply makes use of the idea that nodes closest to the base station tend to forward more packets than those farther away from the base station. An attacker need only monitor which nodes are sending packets and follow those nodes that are sending the most packets. In a *time correlation attack*, an adversary simply generates events and monitors to whom a node sends its packets. To generate an event, the adversary could simply generate a physical event that would be monitored by the sensor(s) in the area (turning on a light, for instance).[7][9]

D. Node Replication Attacks

A node replication attack is quite simple. An attacker seeks to add a node to an existing sensor network by copying (replicating) the node ID of an existing sensor node. A node replicated in this fashion can severely disrupt a sensor network's performance: packets can be corrupted or even misrouted. This can result in a disconnected network, false sensor readings, etc. If an attacker can gain physical access to the entire network he can copy cryptographic keys to the replicated sensor and can also insert the replicated node into strategic points in the network. By inserting the replicated nodes at specific network points, the attacker could easily manipulate a specific segment of the network, perhaps by disconnecting it altogether.[7][8][9]

E. Attacks against Privacy

Sensor network technologies offer great benefits to users. Particularly relevant concerns are privacy problems, since sensor networks provide increased data collection capabilities. Adversaries can use even seemingly innocuous data to derive sensitive information if they know how to correlate multiple sensor inputs. The main privacy problem, however, is not that sensor networks enable the collection of information. In fact, much information from sensor networks could probably be collected through direct site surveillance. Rather, sensor networks aggravate the privacy problem because they make large volumes of information easily available through remote access. Hence, adversaries need not be physically present to maintain surveillance. They can gather information in a low-risk, anonymous manner. Remote access also allows a single adversary to monitor multiple sites simultaneously. Some of the more common attacks against sensor privacy are: *Monitor and Eavesdropping, Traffic Analysis, Camouflage*[1][7][9][10]

1) Monitor and Eavesdropping

This is the most obvious attack to privacy. By listening to the data, the adversary could easily discover the communication contents. When the traffic conveys the control information about the sensor network configuration, which contains potentially more detailed information than accessible through the location server, the eavesdropping can act effectively against the privacy protection.[11]

2) Traffic Analysis

Traffic analysis typically combines with monitoring and eavesdropping. An increase in the number of transmitted packets between certain nodes could signal that a specific sensor has registered activity. Through the analysis on the traffic, some sensors with special roles or activities can be effectively identified.

3) Camouflage

Adversaries can insert their node or compromise the nodes to hide in the sensor network. After that these nodes can masquerade as a normal node to attract the packets, then misroute the packets.

F. Physical Attacks

Sensor networks typically operate in hostile outdoor environments. In such environments, the small form factor of the sensors, coupled with the unattended and distributed nature of their deployment make them highly susceptible to physical attacks, i.e., threats due to physical node destructions. Unlike many other attacks mentioned above, physical attacks destroy sensors permanently, so the losses are irreversible. For instance, attackers can extract cryptographic secrets, tamper with the associated circuitry, modify programming in the sensors, or replace them with malicious sensors under the control of the attacker. Recent work has shown that standard sensor nodes, such as the MICA2 motes, can be compromised in less than one minute. While these results are not surprising given that the MICA2 lacks tamper resistant hardware protection, they provide a cautionary note about the speed of a well-trained attacker. If an adversary compromises a sensor node, then the code inside the physical node may be modified.[6][11][12]

III. NETWORK IMPLEMENTATION

We have implemented a sensor network that consists of basic three types of nodes: - Sensor, Monitor and Head (Base Station).

Sensor node consists of sensor equipments for measuring environmental data, radio resources for transmitting data to the base station and receiving packets. Every sensor node in the network has a unique id with which each sensor node will be identified. Every sensor node has a message processing queue with limited storage length.

The monitoring nodes are in a limited amount in the network with extra processing power and radio range for monitoring the sensor nodes by overhearing sensor transmission. Based on the location information, sensor nodes are assigned to the nearest monitoring node and each monitoring node only monitors or overhears the transmission from or to the sensor nodes assigned to it. Monitoring node further reports the information to base station for making intrusion related decision. Besides this, monitoring node also delivers the routing packet to the sensor nodes under its supervision. Every monitor node has a message processing queue with limited storage length.

The head node is the main node of the network which collects all the network information, make routing decisions and intrusion related decision. The Head node has much more power and radio resources than the other nodes. We have implemented the Fuzzy Logic Controller (FLC) to make the intrusion related decision.[13][14][15]

1) MONITORING PROCEDURE

1. Each monitoring node overhears the transmission from and to the sensor nodes under its supervision and collects the statistics about the number of received and sent message from each sensor node.

2. Based on the statistics it calculates the number of messages dropped and generated by the sensor node from the start till the present moment.
3. Each monitoring node sends the information about the number of dropped and generated messages to the base station via monitoring network.
4. On receiving this packet, base station extract and update the information for each sensor node and pass the data to the Fuzzy Logic Controller for decision making.

future no other packet is sent or received through the affected node.

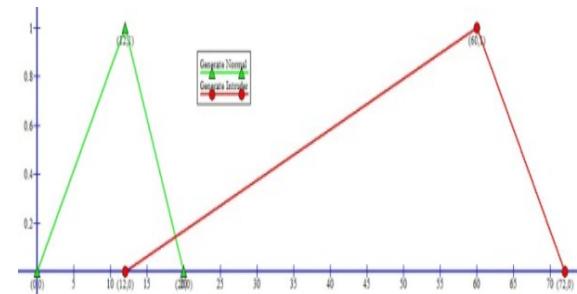


Figure 2. Membership function for Generated Msg.

Figure 1. Snap shot of the simulation network

5. The monitor repeats this monitoring procedure after a certain time interval which is given by **MONITOR_MSG_INTV**.

2) DETECTION PROCEDURE

1. When the packet is received by the base station, it extracts and updates the information for each sensor node and sends the dropped count and generated count to the Fuzzy Logic Controller.
2. The Fuzzy Logic Controller at the base station will decide whether any intruder is there or not. The decision will be made based on the following rule base:

If no_of_generated packet = normal and no_of_dropped packet= normal then Intruder_possibility=low

If no_of_generated packet = normal and no_of_dropped packet= high then Intruder_possibility=high

If no_of_generated packet = high and no_of_dropped packet= normal then Intruder_possibility=high

If no_of_generated packet = high and no_of_dropped packet= high then Intruder_possibility=high

3. Thus on making the decision, if any node is found affected then it will be excluded from the network by updating the routing table for each node so that in

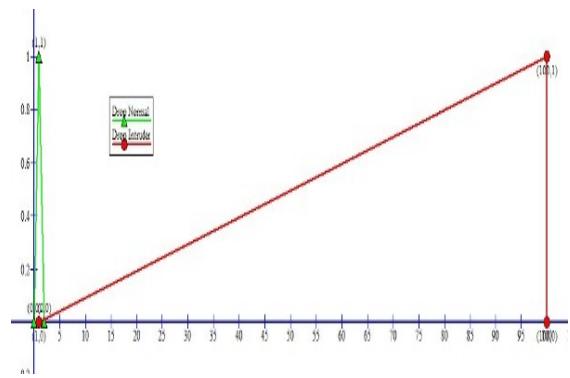


Figure 3. Membership function for Dropped Message

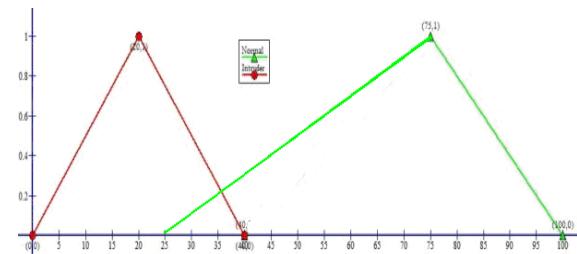


Figure 4. Membership function for Intruder possibility

IV. SIMULATION

We have carried out the simulation using a software simulation tool called OMNeT++ simulation library [16]. C++ was used for carrying out the coding of the various network components.

We have recorded the data from both the ends i.e. sensor and base station. Based on the result obtained we have checked whether the monitoring process has successfully executed and whether the actual intruders are detected or not.

For each sensor node we have recorded the number of dropped messages and generated message.

2nd Annual International Conference IEMCON 2012

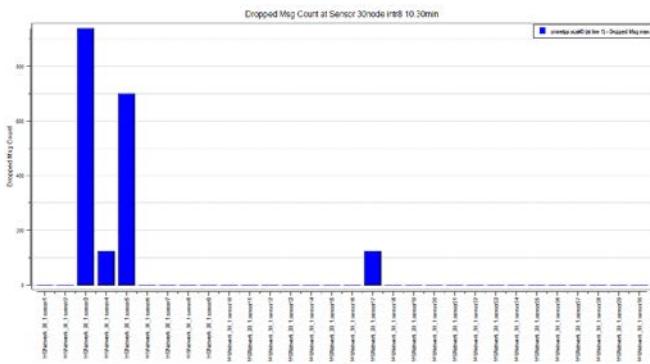


Figure 5. Overall dropped message count recorded at sensor.

At the sensor end the data being recorded is plotted. From the plot we observe that node 3,4,5 and 17 have dropped all the messages.

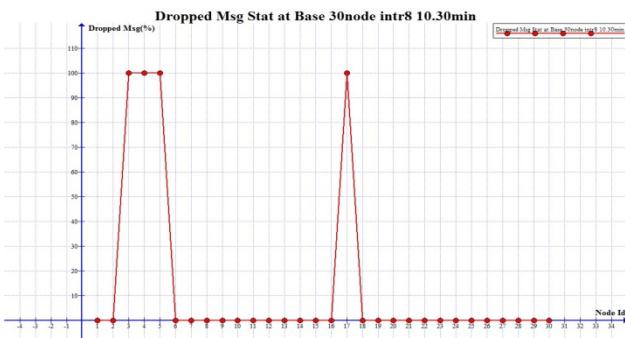


Figure 6. Dropped message count recorded at base.

Here we have plotted the percentage of dropped messages as reported by the monitoring node.

For each sensor node we have also recorded the number of generated messages.

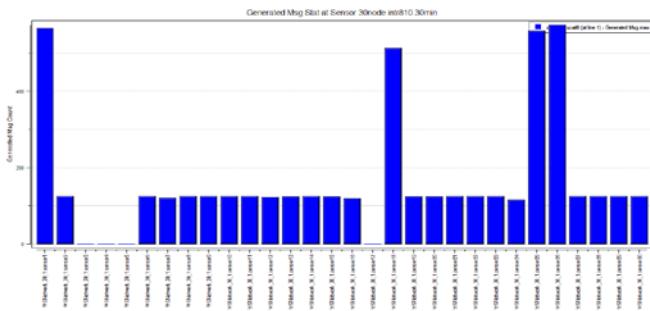


Figure 7. Generated message count recorded at sensor node.

Each of the shaded bars in the graph indicates the total message generated in each sensor node. For most of the sensor nodes the message generation rate is almost same which thereby indicates the normal message generation rate for the nodes in the network. But we can observe that for four nodes the message generation rate is very high compared to others. Also there are four nodes whose message generation rate is zero.

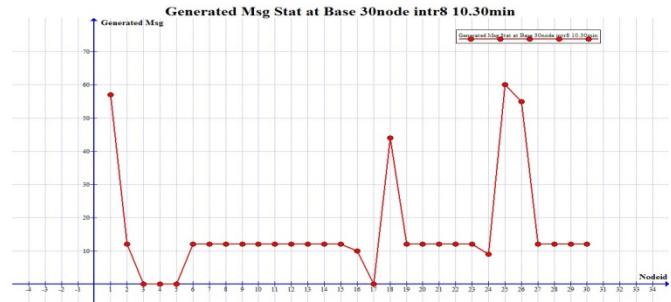


Figure 8. Generated message statistics recorded at base station.

From the data received at the base station, we obtain similar properties for the nodes suspected as an intruder in the previous graph.

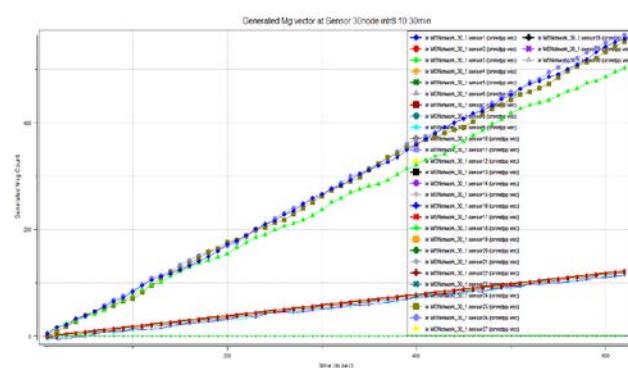


Figure 9. Generated message count with time recorded at sensor.

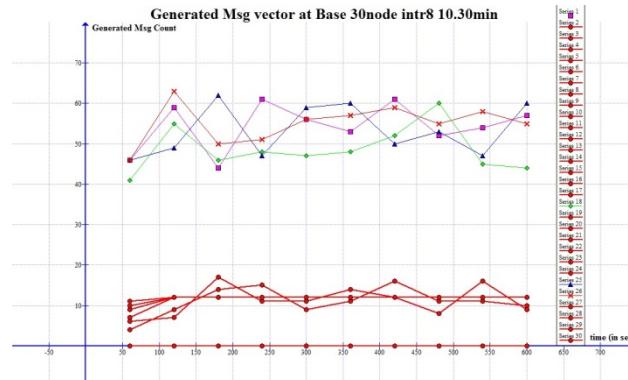


Figure 10. Generated message count with time recorded at the base station.

The above two graphs represent the variation of the results with time. From the results obtained we can conclude that after every reporting we can detect an intruder.

The data obtained from the simulated network are sent to the Fuzzy Logic Controller (FLC) which finally gives us the intruder state of the sensor nodes.

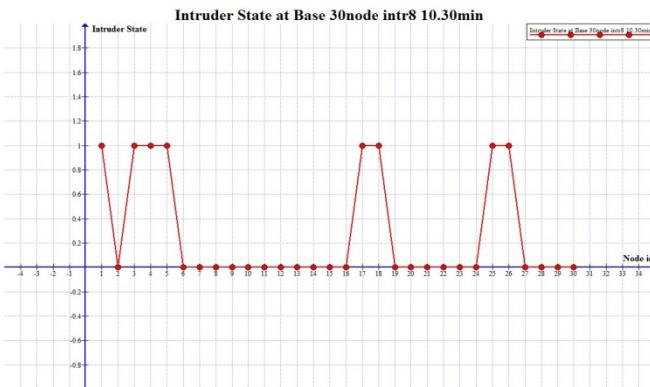


Figure 11. Intrusion Detection in FLC.

In this simulation we have simulated the network by varying the number of intruders and also by varying the type of attacks. And finally our fuzzy logic based scheme will detect the intruder correctly.

ACKNOWLEDGMENT

We would like to thank SUVADEEP SAHA & TANEYA TALUKDAR for helping us in the simulation work.

REFERENCES

- [1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," IEEE Communication Magazine, Aug. 2002.
- [2] John Paul Walters, Zhengqiang Liang, Weisong Shi, and Vipin Chaudhary, "Wireless Sensor Network Security: A Survey", Security in Distributed, Grid, and Pervasive Computing, 2006.
- [3] P. Albers and O. Camp. Security in ad hoc networks: A general intrusion detection architecture enhancing trust based approaches. In First International Workshop on Wireless Information Systems, 4th International Conference onEnterprise Information Systems, 2002.
- [4] C. Karlof, N. Sastry, and D. Wagner. Tinysec: A link layer security architecture for wireless sensor networks. In *Second ACM Conference on EmbeddedNetworked Sensor Systems (SensSys 2004)*, pages 162–175, November 2004.
- [5] J. Douceur. The sybil attack. In Proc. of the 1st International Workshop on Peer-to-Peer Systems (IPTPS'02), February 2002.
- [6] Y. Mun and C. Shin, "Secure routing in sensor networks: Security problem analysis and countermeasures," in International Conference on Computational Science and Its Applications - ICCSA 2005, May 9-12 2005, vol. 3480 of Lecture Notes in Computer Science, (Singapore), pp. 459–467, Springer Verlag, Heidelberg, D-69121, Germany, 2005.
- [7] Sophia Kaplantzis, "Security Models for Wireless Sensor Networks", March 20, 2006
- [8] A. Agah, S. K. Das, K. Basu, and M. Asadi, "Intrusion detection in sensor networks: A non-cooperative game approach," in Proceedings - Third IEEE International Symposium on Network Computing and Applications, NCA 2004, Aug 30- Sep 1 2004, Proceedings - Third IEEE International Symposium on Network Computing and Applications, NCA 2004, (Cambridge, MA, United States), pp. 343– 346, IEEE Computer Society, Los Alamitos, CA 90720-1314, United States, 2004.
- [9] C. E. Loo, M. Y. Ng, C. Leckie, and M. Palaniswami, "Intrusion detection for sensor networks," 2004.
- [10] O. Kachirski and R. Guha, "Effective intrusion detection using multiple sensors in wireless ad hoc networks," in System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on, p. 8 pp., 2003.
- [11] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler. Spins: security protocols for sensor networks. *Wireless Networking*, 8(5):521–534, 2002.
- [12] C. Shen, C. Srisatjapornphat, and C. Jaikaeo, "Sensor information networking architecture and applications," IEEE Pers. Communication, pp. 52–59, Aug. 2001.
- [13] Tiranuch Anatvatee and Jie Wu, "A Survey on Intrusion Detection in Mobile Ad Hoc Networks" Signals and Communication Technologies, ISSN 1860-4862 ,Wireless Network Security, Springer US 2007, pp. 159-180.
- [14] T. Park and K. Shin. LiSP: A Lightweight Security Protocol for Wireless Sensor Networks. ACM Transactions on Embedded Computing Systems, Vol 3, No. 3, Pages 634-660, August 2004.
- [15] "Fuzzy Logic and Control: Software and Hardware Applications", :M. Jamshidi, N. Vadiee, and T. J. Rose, PTR Prentice Hall, Englewood Cliffs, New Jersey 07632.
- [16] "Omnet++ community site," November 2005. <http://www.omnetpp.org/>.

Artificial Neural Network based Automatic Mobile Radio Signal Classifier

S. Roy Chatterjee^{#1}, I. Banerjee^{#2}, S. Bhattacharyya^{#3}, M. Chakraborty^{#4}

^{#1}*Netaji Subhash Engineering College*

rcswagata@gmail.com

^{#2,#3,#4}*Institute of Engineering & Management*

#2 nilbanerjee11@gmail.com

#3 somabhattacharyya86@gmail.com

#4 mohuyacb@yahoo.com

Abstract— Spectrum sensing performed by Cognitive Radio (CR) is a critical and challenging task. It must include detection and identification of communication signals in order to avoid interference with the primary user. In this paper an Automatic Mobile Radio Signal Classifier (AMRSC) is designed based on cyclostationary feature detection and artificial neural network (ANN) to distinguish different digital modulated signals viz., Phase Shift Keying (PSK), Frequency Shift Keying (FSK), Direct sequence Spread Spectrum (DSSS) and Gaussian Minimum Shift Keying (GMSK). Recent developments of ANN have made them a very powerful tool for pattern recognition and classification. By combining the above two techniques a more efficient and reliable classifier is developed where a significant amount of processing is performed. The performance of the classification scheme is investigated through simulations using MATLAB 7.11.0.584 (R2010b). High recognition rate of about (98.5%) is obtained. However, there are probabilities of misclassification of about (1.5%).

Index Terms- Cognitive radio, Artificial Neural network, Modulation Classification, Cyclostationary

I. INTRODUCTION

The most appealing property of the Cognitive Radio (CR) is its ability to sense and characterize its RF environment using spectrum sensing technique and adapting to it accordingly [1], [2], [3]. It has received international scientific attention since the last three decades up to date due to its several possible roles in both civilian and military applications such as signal confirmation, interference identification, spectrum management, and surveillance. Spectrum sensing must include detection and identification of mobile radio signals in order to avoid interference with the primary user. By classifying digital modulated signals it is possible to detect and classify different mobile radio signals exhibiting the corresponding modulation schemes. Efficient design of

AMRSC will be one of the most challenging tasks for the design and implementation of CR based wireless communication systems.

Modulation classification can be categorized into two major groups, i.e. decision-based (DB) methods and feature-based (FB) methods [4], [5]. The DB approach is based on the likelihood functions of received signals where pre-defined decision thresholds are determined by comparing the likelihood ratios. However the dynamic nature of the mobile radio environment for most wireless communication systems makes the likelihood ratio test ineffective. On the other hand, the FB approach is based on extracting some basic characteristics of the signals. Extracting a proper set of features for classification is a very critical task in real applications. Cyclostationary feature detection technique is favorable for spectrum sensing in low SNR scenarios due to its robustness against the feature of uncertainty in noise power [6], [7], [8].

Artificial neural network (ANN) ability to automatically learn from examples makes them attractive and exciting in designing feature classifier [9], [10]. Here a reliable and efficient Automatic Mobile Radio Signal Classifier (AMRSC) is designed based on cyclostationary feature detection and ANN.

The organization of the paper is as follows. After the introduction in section I, section II provides an outline of cyclostationary signal processing followed by spectral correlation density (SCD) functions of different mobile radio signals used for their cyclic spectral analysis [8]. Section III gives a brief overview of ANN used for signal classification. Framework of the mobile signal classifier is described in section IV. Performance analysis of the designed classifier is provided in section V. Section VI concludes the paper with some highlights on future works.

II. OUTLINE OF CYCLOSTATIONARY SIGNAL PROCESSING

Cyclic spectral analysis deals with second order transformation of a function and its spectral representation [6], [8], [11]. A continuous-time signal $x(t)$ is said to be cyclostationary (in wide sense), if it exhibits a periodic auto-correlation function which is given by

$$R_x(t, \tau) = E[x(t)x^*(t - \tau)] \quad (1)$$

where, $E[\cdot]$ represents statistical expectation operator. Since $R_x(t, \tau)$ is periodic, it has the Fourier series representation

$$R_x(t, \tau) = \sum_{\alpha} R_x^{\alpha}(\tau) e^{j2\pi\alpha t} \quad (2)$$

where sum is taken over integer multiple of fundamental cycle frequencies, α . The term $R_x^{\alpha}(\tau)$ in equation (2) is known as cyclic auto-correlation function, which is defined as:

$$R_x^{\alpha}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} R_x(t, \tau) e^{-j2\pi\alpha t} dt \quad (3)$$

The spectral correlation density (SCD) function is defined as the Fourier transform of cyclic auto-correlation function of $x(t)$. The SCD of a signal $x(t)$ is given by,

$$S_x^{\alpha}(f) = \int_{-\infty}^{\infty} R_x^{\alpha}(\tau) e^{-j2\pi f \tau} d\tau \quad (4)$$

Signals usually exhibit distinctive features in frequency domain that may not be present in the time domain. Those features act as a fingerprint of the signal to detect their presence. The FFT Accumulation Method (FAM) is a computationally efficient algorithm to measure the SCD function [8]. Cyclic spectral analysis is a very useful tool for signal classification due to the following reasons: [8]

- Different types of modulated signals (FSK, BPSK, GMSK, DS-CDMA) with overlapping power spectral densities have highly distinct SCD functions.
- There is no spectral correlation exhibited by stationary noise.
- The spectral correlation density function contains phase and frequency information related to timing parameters in modulated signals (carrier frequencies, pulse rates, chipping rates in spread spectrum signaling, etc.).

III. BRIEF OVERVIEW OF ARTIFICIAL NEURAL NETWORK

An ANN can be viewed as a data processing system consisting of large number of artificial neurons in a network structure. Artificial neurons are highly interconnected processing elements with many inputs and one output. An artificial neuron is well described mathematically by binary threshold unit according to computational model of McCulloch and Pitts [10] as shown in figure1. This computes a weighted sum of its n input signals, $x_j, j = 1, 2, \dots, n$, and generates an output of 1 if this sum is above a certain threshold ' u ', otherwise an output of zero results. Mathematically,

$$y = \theta \left\{ \sum_{j=1}^n (w_j * x_j) - u \right\} \quad (5)$$

where $\theta\{\}$ is a unit step function at 0, and w_j is the synapse weight associated with the j^{th} input.

The neuron has two modes of operation; the training mode with validation and the testing mode. In the training mode, the training data set is used to update the synaptic weights. The performance of the network is evaluated on the validation set after each iteration and the training is stopped if the performance of validation did not increase for more than 15 training iterations or the minimal gradient was reached [9]. In the testing mode, the testing data set is used to measure the performance of the network after it has been trained.

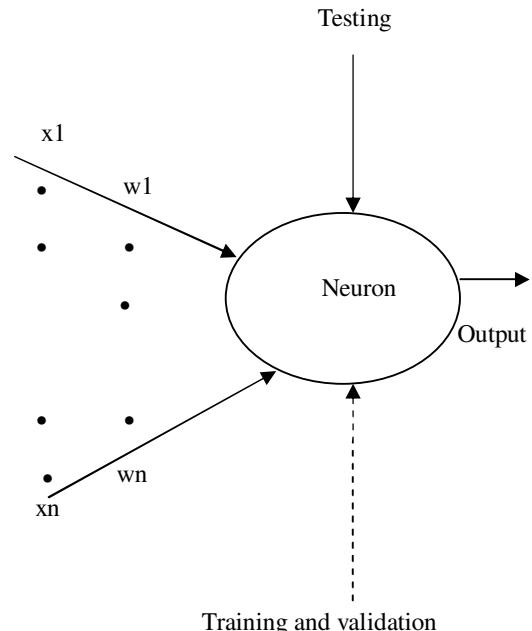


Figure1. McCulloch-Pitts model of a neuron

Positive weights correspond to excitatory synapses, while negative weights model inhibitory ones. Based on the connection pattern, ANN can be grouped into two categories: feed-forward networks and recurrent. In feed-forward networks, neurons are organized into layers that have unidirectional connections between them. It is also called multilayer perception in which graphs have no loop and recurrent (or feedback) networks, in which loops occur because of feedback connections.

The major advantage of neural networks over traditional expert systems is its ability to learn underlying rules (like input-output relationships) from the given collection of representative examples. A learning process in the ANN is used for updating network architecture and connection weights so that a network can efficiently perform a specific task.

IV. FRAMEWORK OF MOBILE RADIO SIGNAL CLASSIFIER

We have designed AMRSC combining spectral correlation analysis and ANN. The workflow of AMRSC consists of two stages.

- Feature extraction of different digital modulated signals used for mobile radio communication using SCD function.
- Classification of the signals using Neural Network Pattern Recognition (NPR) tool of MATLAB.

2nd Annual International Conference IEMCON 2012

The feature parameters that are almost insensitive to signal to noise ratio (as given in Table I) are selected as feature vectors for the signal classifier.

TABLE I
FEATURE PARAMETERS

| Features | Description |
|----------|--|
| x1 | Frequency of peak at α -domain |
| x2 | Frequency of peak at f-domain |
| x3 | Maximum value of peak at α -domain |
| x4 | Maximum value of peak at f-domain |
| x5 | No. of peaks ($f=0$) at α -domain |
| x6 | No. of peaks ($\alpha=0$) at f-domain |
| x7 | Energy of the signal at α -domain |
| x8 | Carrier frequency (fc) |

These features are calculated for different digital modulated signals and for the vacant bands (VC) at different carrier frequencies for different signal to noise levels. VC is assumed to be the absence of signal or presence of stationary noise. The set of values of eight feature vectors for SNR= -2dB are listed in Table II.

TABLE II
CALCULATED FEATURE PARAMETER VALUES

| | BPSK | FSK | DSSS | GMSK | VC1 | VC2 | VC3 | VC4 |
|----|---------|---------|---------|----------|---------|----------|-----------|---------|
| x1 | 7998 | 2002 | 4004 | 2158 | 0 | 0 | 0 | 0 |
| x2 | 3750 | 937.5 | 2500 | 1250 | 0 | 0 | 0 | 0 |
| x3 | 0.8562 | 0.3679 | 0.3876 | 0.4434 | 0 | 0 | 0 | 0 |
| x4 | 1 | 1 | 1 | 0.8486 | 0 | 0 | 0 | 0 |
| x5 | 4 | 15 | 38 | 45 | 0 | 0 | 0 | 0 |
| x6 | 1 | 2 | 2 | 4 | 0 | 0 | 0 | 0 |
| x7 | 31.8132 | 77.0573 | 88.1391 | 237.2844 | 449.637 | 664.3724 | 445.31178 | 816.644 |
| x8 | 4000 | 1000 | 2000 | 1000 | 4000 | 1000 | 2000 | 1000 |

These features are normalized to avoid numerical computational error by subtracting mean of each feature from the original feature and dividing the result by the standard deviation δxi of the same feature [11] as given by equation (6).

$$xi' = (xi - \bar{xi}) / \delta xi \quad i = 1, 2, 3, 4 \quad (6)$$

After normalization, the feature vector $x = (x1', x2, \dots, x8')$ are generated as the input to NNTK tool of MATLAB 7.11.0.584 (R2010b) to classify eight different signals with the following experimental criteria as given in Table III.

TABLE III
EXPERIMENTAL CRITERIA

| | |
|--------------------------------|--------------------------|
| Number of Input Layer Neurons | 8 |
| Number of Hidden Layers | 1 |
| Number of Hidden Layer neurons | 20 |
| Number of Output Layer Neurons | 8 |
| Training Algorithm | Scale Conjugate Gradient |
| Performance | Mean Square Error |

V. PERFORMANCE ANALYSIS

A. Simulation Environment:

To measure the performance efficiency of AMRSC, simulation was performed using MATLAB 7.11.0.584 (R2010b) with 1008 sample feature data. 70% of sample data was used for training, 10% for validation and the remaining 20% was used for testing. Figure 2 shows the experimental neural network setup for ANN signal classifier.

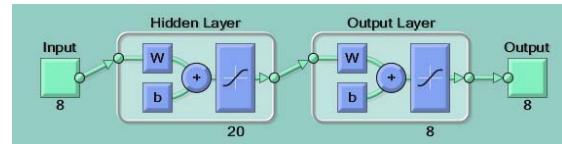


Figure 2 Experimental setup of ANN signal classifier

The input layer of the experimental setup has eight input neuron to describe eight attributes in the feature data set. For the hidden layer we have used ten neuron with bias ' b ' and weight ' w ' that give the optimal result. Since eight different types of signals are taken, we have used eight neurons in the output layer. It took 64 iterations to train the network in two seconds.

B. Simulation results

Figure 3 shows confusion matrices i.e Success rate vs Error rate for all stages like training, validation and testing. The results are given in Table IV. Green boxes indicate success (%) whereas red boxes indicate failure (%). Figure 4 shows the performance of the system taking into account training validation and testing. It shows that the best validation was 0.006121 at 58 epoch. Figure 5 shows the error histogram graph for training, validation and testing. The receiver operating characteristic (ROC) for testing as shown in figure 6 consist of true positives and no false positive that indicates a quite satisfactory result.

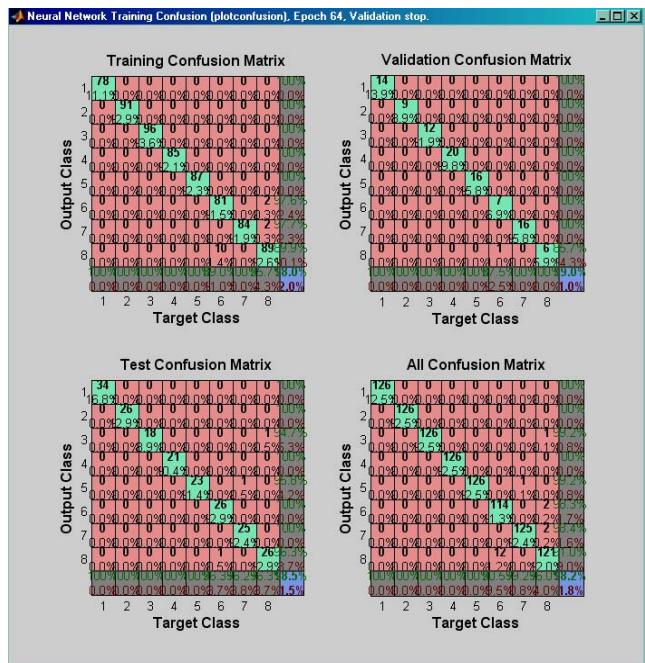


Figure 3 Confusion Matrices

TABLE IV

SUCCESS AND ERROR RATES FOR DIFFERENT EXPERIMENTAL PHASES

| Phase | Success Rate | Failure Rate |
|------------|--------------|--------------|
| Training | 98.0% | 2.0% |
| Validation | 99.0% | 1.0% |
| Testing | 98.5% | 1.5% |

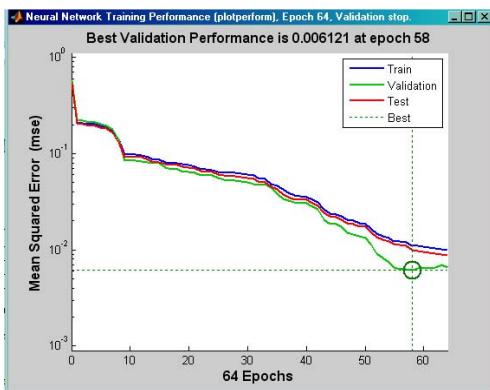


Figure 4 System Performance

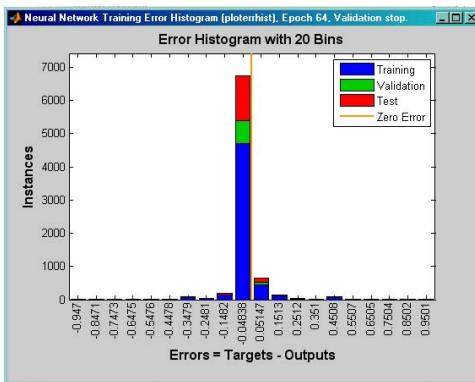


Figure 5 Error histogram

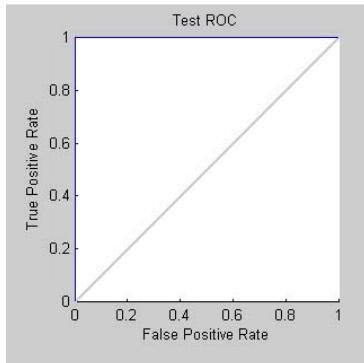


Figure 6 Receiver operating characteristics for testing phase

VI. CONCLUSION

We have designed a simple and reliable AMRSC which can classify different signal with high accuracy. In our future work we propose to implement the AMRSC in Spartan 3 FPGA for real time signal classification.

REFERENCES

- [1] Tevfik Yucek and Huseyin Arslan,"A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications", IEEE Communications Surveys & Tutorials, vol. 11, no. 1, First Quarter 2009 .
- [2] Shahzad A. Malik, Madad Ali Shah, Amir H. Dar, Anam Haq, Asad Ullah Khan, Tahir Javed, Shahid A. Khan"Comparative Analysis of Primary Transmitter Detection Based Spectrum Sensing Techniques" in CognitiveRadio Systems Australian Journal of Basic and Applied Sciences, 4(9): 4522-4531, 2010.
- [3] Ian F. Akyildiz, Won-Yeol Lee, Mehmet C. Vuran , Shantidev Mohanty," NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey " in Computer Networks 50 (2006) 2127–2159.
- [4] Octavia A. Dobre1, Ali Abdi, Yeheskel Bar-Ness and Wei Su" Survey of automatic modulation classification techniques: classical approaches and new trends", Communications, IET April 2007 Vol.1,pp.: 137 – 156.
- [5] Victor Iglesias, Jes'us Grajal and Omar Yeste-Ojeda," Automatic Modulation Classifier for Military Applications", in 19th European signal Processing Conference,29Aug-2Sep,2011.
- [6] Gardner, W.A. "An Introduction to Cyclostationary Signals, Chapter 1, Cyclostationarity in Communications and Signal Processing", IEEE Press,Piscataway,NJ,1993.
- [7] Anwer Al-Dulaimi, Nazar Radhi, H. S. Al-Raweshidy,"Cyclostationary Detection of Undefined Secondary Users",in Third International Conference on Next Generation Mobile Application,Services and Technologies,2009.
- [8] S. Roy Chatterjee, R. Hazra, A. Deb and M. Chakraborty, "Cyclostationary Spectral Analysis Approach to Spectrum Sensing for Mobile Radio Signals",in Proc. IEMCON 2011 organised by IEM in collaboration with IEEE on 5th & 6th of Jan,2011.
- [9] A. Bhavani Sankar, D. Kumar, K. Seethalakshmi,"Neural Network Based Respiratory Signal Classification Using Various Feed-Forward Back Propagation Training Algorithms" European Journal of Scientific Research , Vol.49 No.3 (2011), pp.468-483.
- [10] Anil K. Jain, Jianchang Mao, K.M. Mohiuddin,"Artificial Neural Network: A Tutorial"IEEE Computer, Vol. 29 (Mar 1996) pp. 31 – 44.
- [11] Hau Ho,"Cyclostationary Approaches to Signal Detection and Classification in Cognitive Radio", 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, (DySPAN) 2007.

Study The Nonlinear Behaviour of PFC Boost Converter

Arnab Ghosh

Department of Electrical Engineering
Jalpaiguri Govt. Engineering College
Jalpaiguri, India
e-mail: arnab.4u2010@gmail.com

Abhisek Pal

Department of Electrical Engineering
Jalpaiguri Govt. Engineering College
Jalpaiguri, India
e-mail: abhisekpal.ee@gmail.com

Dr. P. K. Saha

Prof. and Head of the Department of Electrical Engineering
Jalpaiguri Govt. Engineering College
Jalpaiguri, India
e-mail: p_ksaha@rediffmail.com

Dr. G. K. Panda

Professor of the Department of Electrical Engineering
Jalpaiguri Govt. Engineering College
Jalpaiguri, India
e-mail: g_panda@rediffmail.com

Abstract—With rapid development in power semiconductor devices, the usage of power electronic systems has expanded to new and wide application range that include residential, commercial, aerospace and many others. However, their nonlinear behavior puts a question mark on their high efficiency. This paper aims to develop a circuit for PFC boost converter to observe nonlinear behaviors. It is clear that the output storage capacitor is a main contributing parameter on the system stability, therefore, bifurcation maps are developed to determine the accurate minimum output capacitance value that assures the system stability under all operating conditions.

Keywords- *PFC boost converter; Phase Plane Trajectories ; Bifurcation diagrams.*

I. INTRODUCTION

The power electronic engineers observe some strange phenomena noise like oscillation. Actually power electronics system can exhibit a variety of nonlinear behaviours because of periodic switching of the circuits. This kind of nonlinearity is the main cause of harmonics generation i.e. degradation of input power factor. In the last decade, bifurcation and chaotic phenomena have been reported in some type of DC-DC converters [1][3][6]. Here we are discussing about some nonlinear phenomena of Power Factor Corrected (PFC) Boost Converter.

The operation of the boost PFC converter [2][4] has been analyzed in details by many researchers. In practical circuits, it is much more difficult to arrange pure DC source, as well as the setup is much more expensive. So we are considering rectified dc in spite of pure DC. This DC contains several kinds of harmonics. We consider this kind of converter circuit and want to maintain the input power factor high.

Recently, power-factor-correction (PFC) circuits are widely used in power electronics. One of the most common and more attractive circuits than other is the boost power factor correction circuit operating at the

continuous conduction mode (CCM)[3] because of the unity power factor and the reduced current stress. But it has been found that the PFC boost converter may operate in discontinuous conduction mode (DCM) near the crossover point of the line voltage, in particular, the PFC converter may operate in DCM during the whole line cycle under the light load condition. It has been found that the transition from CCM to DCM can cause bifurcation and chaos.

Most prior researches introduced some assumptions that force the time-varying PFC system to be linear. They linearised the system as their assumption. They assumed a very huge output capacitance (not acceptable in industry) and it resulted in the time-invariant feedback signal that neglected the time-varying effect. Also, they replaced the input voltage with its root mean square (r.m.s.) value, neglecting the effect of its amplitude variation. Then, they introduced a small-signal equivalent circuit and the stability was examined by this linear model.

The PFC converter is nonlinear system[4] due to a multiplier using and a large variation of duty cycle. There is also present nonlinear term in its state equations. Here we will observe the nonlinear phenomena and bifurcation of this converter.

II. PFC BOOST CONVERTER AND ITS PROPOSED MODEL

Modelling, simulation and circuit analysis are done by MATLAB respectively. These not only help in developing a deeper understanding of PFC converters but are also extremely important tools for design verification and performance evaluation. These techniques help in the evaluation of a system without risking the huge cost and effort of developing and testing an actual converter.

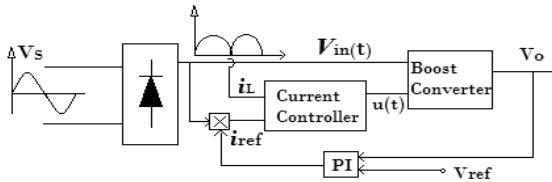


Fig. 1: Block Diagram of a boost PFC circuit

From the above Fig.1 the single phase sinusoidal voltage source v_s is rectified by diode bridge and the rippled DC voltage v_{in} is fed to the boost converter. The output voltage v_o (ripple is presented as the value of the capacitor is taken small) is obtained from load side. The output voltage v_o is compared with a reference voltage (DC) v_{ref} . We use an integral controller to get steady state value of error signal. i_{ref} is obtained after combining the the result of controller, v_{in} and inductor current (i_L). Now i_{ref} or i_L^* is compared with i_L . The duty cycle is maintained by the result of the comparator. The clock period and the value of the inductor are so chosen that the inductor current never falls to zero.

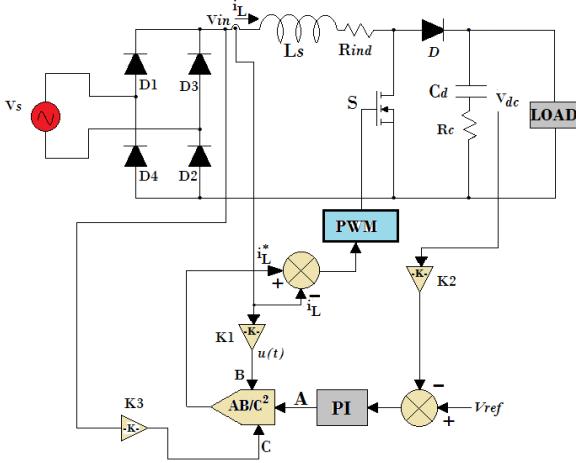


Fig. 2: Boost PFC ac-dc regulator under fixed-frequency Current Mode Control

Depending upon the block diagram we design the above model and derived the several expressions[2][4] which are given below.

Supply system:

Under normal operating conditions the supply system can be modelled as a sinusoidal voltage source of amplitude v_m and frequency f_s . The instantaneous voltage is:

$$v_s(t) = v_m \sin \omega t \quad (1)$$

where $\omega = 2\pi f_s$ electrical radians/second and t is instantaneous time.

In some topologies, the input is rectified line voltage $v_d(t)$ which can be given as:

$$v_d(t) = |v_s(t)| = |v_m \sin \omega t| \quad (2)$$

From the sensed supply voltage, an input-voltage template $u(t)$ is estimated for converter topologies with AC side inductor as:

$$u(t) = v_s(t)/v_m \quad (3)$$

The input-voltage template for converter topologies with a DC side inductor is obtained from:

$$u(t) = |v_s(t)|/v_m \quad (4)$$

Feedback controller:

PFC converters, like most power electronics systems, cannot function without feedback control. Fig.1 shows a block diagram of typical control scheme for PFC converters – the current mode control [3]. This control scheme ensures regulated DC output voltage at high input power factor. The output DC voltage regulator generates a current command, which is the amount of current required to regulate the output voltage to its reference value. The output of the DC voltage regulator is then multiplied with a template of input voltage to generate an input current reference. This current reference has the magnitude required to maintain the output DC voltage close to its reference value and has the shape and phase of the input voltage – an essential condition for high input power factor operation.

(i) Output voltage controller:

A proportional integral (PI) voltage controller is selected for zero steady-state error in DC voltage (rippled in nature) regulation. The o/p capacitor voltage v_{dc} (or v_o) is sensed and compared with the set reference voltage v_{ref} . The resulting voltage error $v_{e(n)}$ at the nth sampling instant is:

$$v_{e(n)} = v_{ref} - v_{dc(n)} \quad (5)$$

The output of the PI voltage regulator $v_o(n)$ at the nth sampling instant of the PI controller will be:

$$v_{o(n)} = v_{o(n-1)} + k_p \{ v_{e(n)} - v_{e(n-1)} \} + k_i v_{e(n)} \quad (6)$$

Here k_p and k_i are the proportional and integral gain constants, respectively. $v_{e(n-1)}$ is the error at the $(n - 1)$ th sampling instant. The output of the controller $v_{o(n)}$ after limiting to a safe permissible value is taken as the amplitude of the input current reference A (Fig. 2).

(ii) Reference current controller:

The input voltage template $u(t)$ obtained from the sensed supply voltage is multiplied by the amplitude of the input current reference A to generate a reference current. The instantaneous value of the reference current is given as:

$$i_L^* = AB / C^2 \quad (7)$$

where B is the input voltage template $u(t)$ and C is the input voltage feed forward component obtained by low-pass filtering the sensed input voltage signal.

Semiconductor switches:

Semiconductor switches, Mosfet S and Diode D are modelled as pure ON-OFF switches. No snubbers or non-idealities in the switches are modelled.

Load:

The converters are modelled as resistive loads having resistance R.

Power circuit:

The power circuit is modelled by first-order differential equations describing the circuit behaviour.

These modelling equations are obtained by application of Kirchoff's and Ohm's laws to the power circuit.

III. STATE EQUATIONS FOR MODELING OF BOOST CONVERTER

There are two states[1][3] of the circuit depending on whether the controlled switch is open or closed. When switch is closed, the current through the inductor rises and any clock pulse arriving during that period is ignored. The switch opens when reaches the reference current. When switch is open, the current falls. The switch closes again upon the arrival of the next clock pulse.

The State Equations during "ON" period

$$\frac{di_L}{dt} = V_{in}/L - (r_i * i)/L \quad (8)$$

$$\frac{dv_c}{dt} = -v_c/C(R + r_c) \quad (9)$$

The State Equations during "OFF" period

$$\frac{di_L}{dt} = V_{in}/L - i*(r_i + R*r_c/(R + r_c))/L - v_c*R/L(R + r_c) \quad (10)$$

$$\frac{dv_c}{dt} = (R*i - v_c)/C(R + r_c) \quad (11)$$

where,

V_{in} =Input Voltage

L = Inductor

C = Capacitor

i_L = Inductor Current

v_c = Capacitor Voltage,

r_i & r_c = Parasitic Elements

IV. SIMULATION OF PFC BOOST CONVERTER

Simulation of PFC Boost Converter is done by MATLAB 7.8R2009a. The model is totally designed by SimPowerSystem and Simulink blocks.

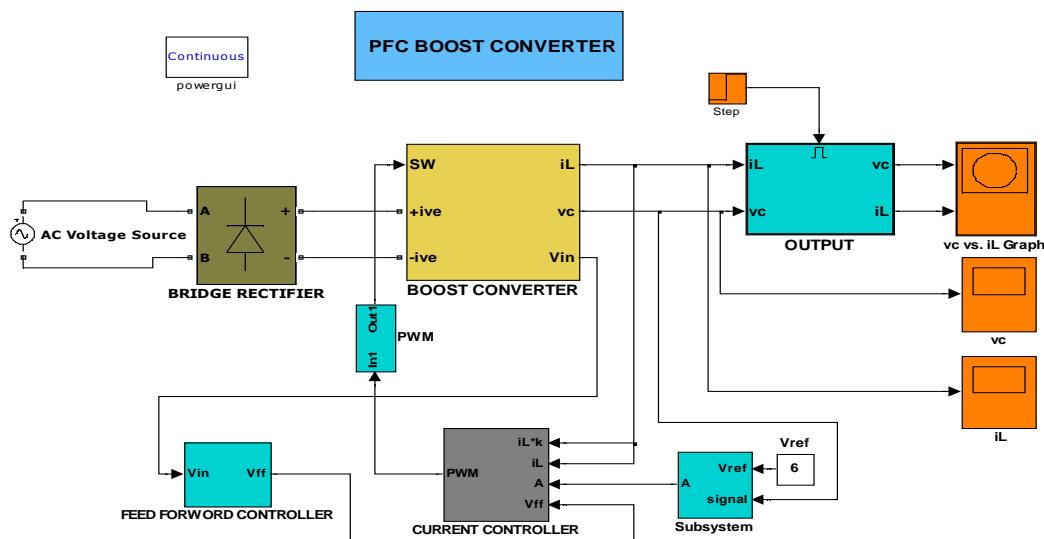


Fig. 3: Simulation of PFC Boost Converter

V. EXPERIMENTAL RESULTS

Here we are varying the value of Current Gain K1 (is placed in Current Controller Block, Fig.2) and we obtain the several periodic behavior of converter.

Case I (Period I Operation)

$V_s=220\sin w t$, $L=40\text{mH}$, $C=100\mu\text{F}$, $R=35\text{ohm}$, $K1=850$

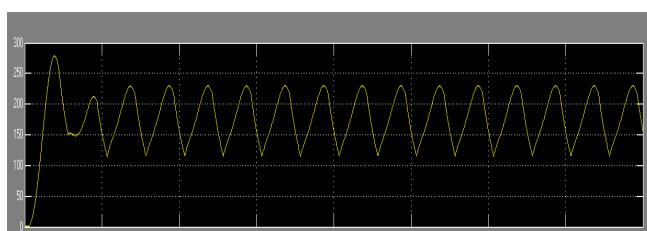


Fig.4(a) : O/P Voltage Waveform at Period I operation ($K = 850$)

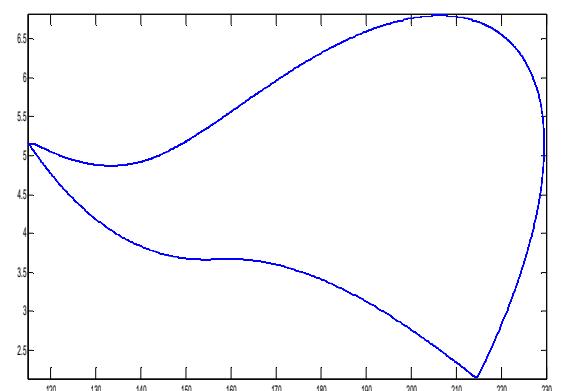


Fig.4(b) : Phase Plane Trajectory(Case I)
Capacitor Voltage vs Inductor Current (Period I)

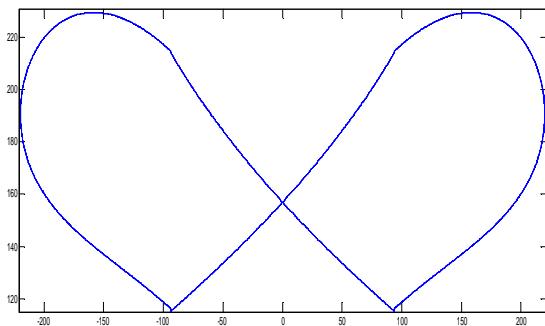


Fig.4(c) : V_{in} vs V_c (Case I)

Case II(Period II Operation)

$V_s=220\sin \omega t$, $L=40mH$, $C=100\mu F$, $R=35\Omega$, $K_1=450$

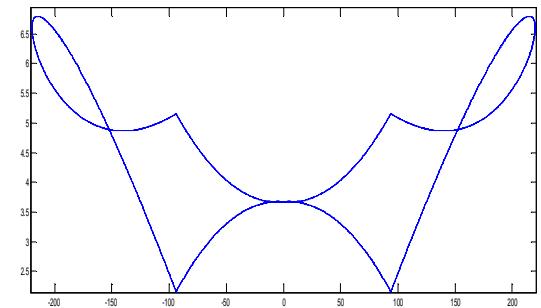


Fig.4(d) : V_{in} vs i_L (Case I)



Fig.5(a): O/P Voltage Waveform at Period II operation ($K = 450$)

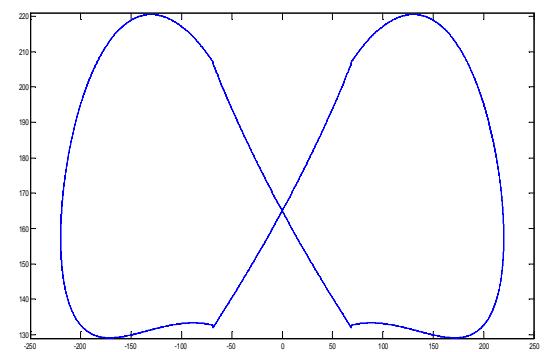


Fig.5(c) : V_{in} vs V_c (Case II)

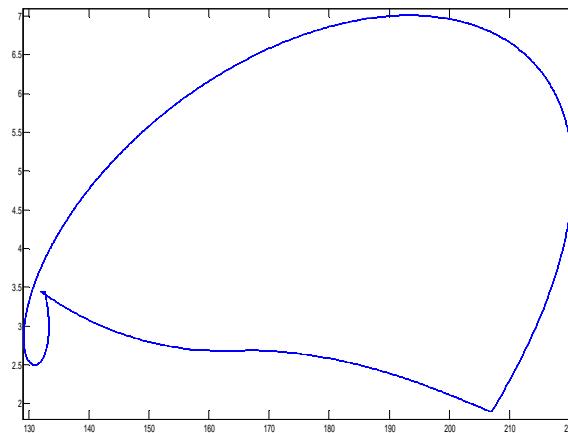


Fig.5(b) : Phase Plane Trajectory(Case II)
Capacitor Voltage vs Inductor Current (Period II)

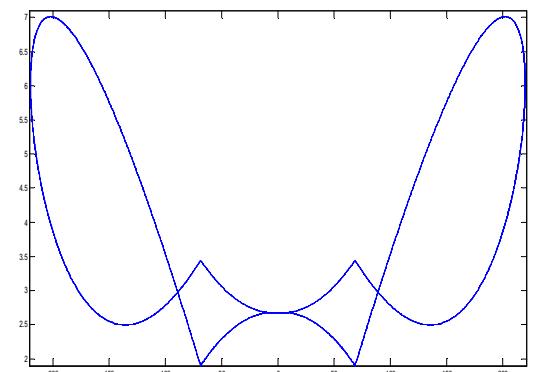


Fig.5(d) : V_{in} vs i_L (Case II)

Analysis of Experimental Results:

Here the state variables are inductor current (i_L) and capacitor voltage (V_c). From above results we see that case I(Fig.4b) is operating at period I condition[5] and case II(Fig.5b) is operating at period II condition[5]. The output voltage waveform in Fig.5(a) is much more ripple free than Fig.4(a). We get better result of output voltage at same value of inductor (L) and capacitor (C), only changing the value of current gain(K_1). The value of Capacitor(C) is chosen small just it operates as a boost converter. If we can decrease more values of current gain K_1 , the system will operate at chaotic region and we can get better ripple free output voltage. This is the main observation that we get

better output voltage profile at least value of capacitor. So, the investment is much more less than other conventional practical instruments.

VI. BIFURCATION DIAGRAMS

Bifurcation diagrams are obtained from FORTRAN and ORIGIN 5.0 software. The data files are obtained after executing the FORTRAN programme of State Equations (8), (9), (10), (11) of PFC Boost Converter and i_{ref} equ (7). This data files are plotted by ORIGIN 5.0. Results are given below.

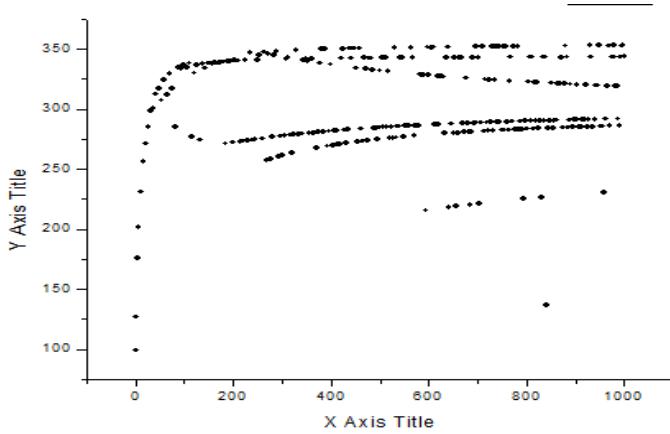


Fig.6(a):R vs Vc (R is varied 1 to 1000ohm with step of 0.5)

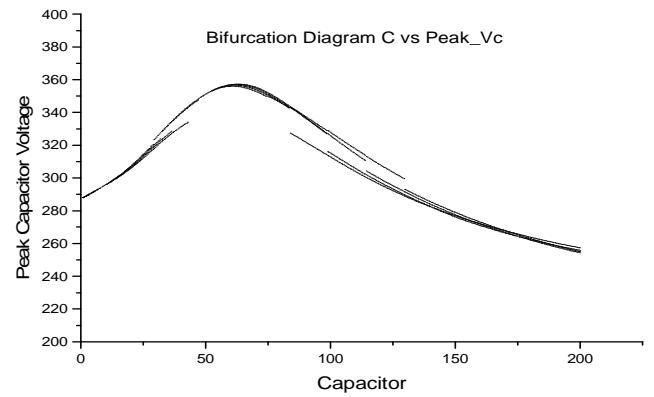


Fig.6(c):C vs Peak_Vc (C is varied 1 to 200ohm with step of 0.5)

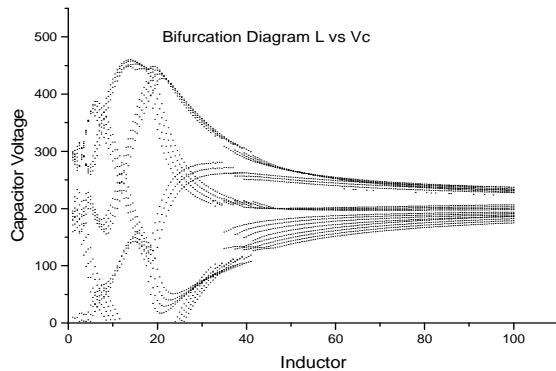


Fig.6(b):L vs Vc (L is varied 1 to 100ohm with step of 0.5)

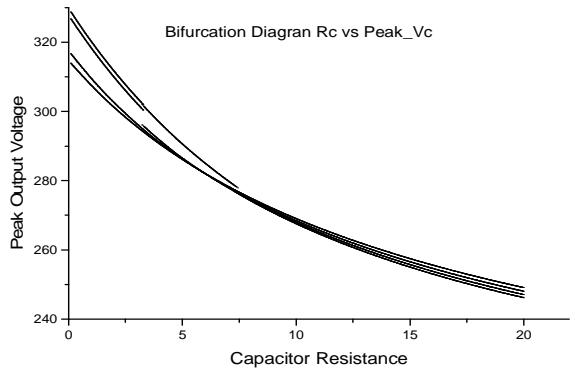


Fig.6(d):Rc vs Peak_Vc (Rc is varied 0.1 to 20ohm with step of 0.01)

Analysis of Bifurcation Diagrams:

The above bifurcation diagrams are much more differ from other conventional bifurcation diagrams. In conventional process the border [6] is fixed i.e. I_{ref} (or i_L^*) is constant. In our experiment I_{ref} (or i_L^*) is time varying nature i.e. order is time-varient. It is very difficult to analysis the bifurcation diagram properly. Actually the bifurcation is Period Doubling [5][6] in nature. The analysis is not given here. We are now working on analysis of didgrams.

VII. CONCLUSION

The boost PFC converter with current mode control has been examined. Results highlight that the proposed model of practical pfc converter, experimental results and bifurcation diagrams. The value of current gain is decreased; the output capacitor voltage waveform is going to period I to period II i.e. period doubling bifurcation is observed. But the main benefit is the output voltage ripple is going less than the previous. In a DC/DC converter system, the input voltage is constant and therefore the dynamical behavior is periodic with the switching frequency. On the other hand, the input voltage of the boost AC/DC PFC converter system is periodic with the

line frequency. The results highlight that the dynamical behavior is periodic with the line frequency not with the switching frequency and simulation results are also agree with our statements.

REFERENCES

- [1] Soumitro Banerjee and Krishnendu Chakrabarty, "Nonlinear Modeling and Bifurcations in the Boost Converter" in IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 13, NO. 2, MARCH 1998.
- [2] Ashish Pandey, Dwarka P. Kothari, Ashok K. Mukerjee and Bhim Singh, "Modelling and simulation of power factor corrected AC-DC converters" in International Journal of Electrical Engineering Education.
- [3] Ammar Nimer Natsheh, "ANALYSIS, SIMULATION AND CONTROL OF CHAOTIC BEHAVIOUR IN POWER ELECTRONIC CONVERTERS" in Doctoral Thesis Submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of Loughborough University.
- [4] Yi-Jing Ke, Yu-Fei Zhou and Jun-Ning Chen, "Control Bifurcation in PFC Boost Converter under Peak Current-Mode Control" in IPEMC 2006.
- [5] Soumitro Banerjee, "Dynamics for Engineers" (John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England).
- [6] Soumitro Banerjee and George C. Verghese, "NONLINEAR PHENOMENA IN POWER ELECTRONICS"(Published by John Wiley & Sons, Hoboken, NJ)

A Review of Secure Routing in Mobile Ad Hoc Network

Abari Bhattacharya

Programmer Analyst Trainee,
Cognizant Technology Solutions

Prof.Himadri Nath Saha

Computer Science Engineering Department
Institute of Engineering & Management(WBUT)

Abstract— The aim of this paper is the study of various routing protocols of Mobile Ad hoc NETwork (MANET) and the security attacks they are prone to. MANET being an infrastructure less network, providing routing security is a major issue. In this paper we present a study of two MANET routing protocols namely Ad hoc On demand Distance Vector Routing (AODV) protocol and Optimized Link State Routing (OLSR) protocol. We discuss the various security attacks they are prone to and the measures taken to counter these attacks.

Keywords-*Wireless Network , MANET, Routing Protocols, Routing Security , Authentication, Hash Function, Cryptography, Digital Signature, black hole attack, wormhole attack , Sybil attack, impersonation, spoofing*

I. INTRODUCTION

MANET is a wireless ad hoc network. It is basically a collection of mobile hosts (nodes). These nodes use wireless transmission for communication with other nodes in the network(fig.1) .MANET does not have any established infrastructure or centralized administration. Nodes are free to move randomly. Hence, MANET does not have any fixed topology and it can change randomly. Here, nodes act both as hosts (capable of sending and receiving messages) as well as routers (i.e., they can also forward data intended for other nodes. In this paper we consider routing in MANET. We all know that security is an important aspect of any network and secure routing of data is also indispensable. However wireless networks like MANET has several security issues. They are as follows:

- MANET does not have a fixed topology and it is dynamic in nature. As a result, malicious nodes can easily join the network and detection of these adversaries is very difficult.



Figure 1. Mobile Ad hoc Network

- MANET lacks centralized monitoring system. So it becomes easy for adversaries to attack the network and hence the security of the network is compromised.
- Being a wireless network it is more susceptible to eavesdropping than wired network.

In this paper we focus on two routing protocols- Ad hoc On demand Distance Vector Routing (AODV) protocol and Optimized Link State Routing (OLSR) protocols, the security attacks they are susceptible to and the countermeasures. But before going into this we would like to generally discuss the security requirements which must be fulfilled by the routing protocols in order to achieve secure routing against various security attacks that pose a threat to MANET (in section 3). After that (in section 4), we deal with the routing protocols, attacks and the countermeasures. In section 4, we conclude the paper.

II. SECURITY REQUIREMENTS

A. Authenticity

The receiver must be able to make sure that the message it has received has come from an authentic source but not from an adversary.

B. Confidentiality

The message which is sent across the network should remain confidential between the source and the destination. Any other node which does not have any permission to view the message should not be able to access it.

C. Integrity

The integrity of the message should not be compromised with, i.e., the message should not be altered by any means during its transmission. The receiver should be able to get the message in the exact condition in which it was sent.

D. Non Repudiation

The sender should not be able to claim that it has not sent any message to the receiver when it has actually sent it, i.e., the sender should not be able to disown a message it has sent to the receiver.

III. SECURITY ATTACKS ON MANET

Several attacks can compromise the security of mobile ad hoc network. They can be broadly classified as follows:

Internal attacks: In this category of attacks the attacker poses as one of the nodes and gains direct access to the network either by impersonation or by compromising a proper node and using it to do its malicious activities.

External attacks: Here, the attacker attacks from outside the network, creating congestion in the network traffic by propagating meaningless messages, thereby disturbing the entire communication of the network.

Let us get familiar to some of the security attacks which we will find later in the paper. These attacks pose a threat to the routing protocols of mobile ad hoc network.

A. Impersonation

This is one of the most severe attacks. Here the attacker can spoof as an innocent node and join the network. In this way, when several such nodes join the network, they gain the control of the network and conduct malicious behavior. They propagate fake routing information and they also gain access to confidential information. A network is vulnerable to such attacks if it does not employ a proper authentication mechanism.

B. Eavesdropping

In this type of attack the aim of the attacker is to get some confidential information, while it is being transmitted from one node to the other. This attack is very hard to detect and the secret information like private key, public key, password etc of the nodes can get compromised due to this attack.

C. Denial of Service

The aim of this attack is to make sure that a specific node is not available for service. The entire service of the network might be compromised due to this attack.

D. Wormhole attack

In this attack, the adversary connects two distant parts of the network and tunnels messages received in one part of the network to the other. Here a low latency link is used to pass the

messages. Wormholes are sometimes used to create black holes in the network.

E. Black hole attack

In this attack the adversary lures the traffic of the network towards a compromised node and thus a black hole is created with the adversary at the centre. Fig.2 shows Black hole attack. Here generally the attacker offers an attractive route to the surrounding nodes. This attack can also be coupled with other attacks like dropping packets, denial of service, replay of information, selective forwarding and so on.

F. Sybil attack

In Sybil attack, a node tries to have several different identities. This helps the malicious node to gain more information about the network. The effectiveness of fault tolerant schemes like distributed storage, multipath routing, topology maintenance etc has a notable decrease.

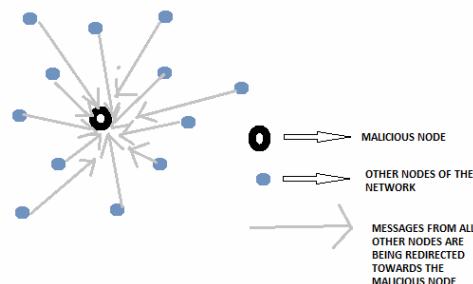


Figure 2. Black hole attack

IV. ROUTING IN MOBILE AD HOC NETWORK

In MANET, there is no fixed network topology; rather it is dynamic in nature. So nodes in such networks need to discover routes themselves. Hence, in MANET routing protocols, nodes need to announce their presence to other fellow nodes as well as know the presence of their fellow nodes in the network. Basically they learn about each other's presence and also discover routes to communicate with each other.

The following is a list of some MANET routing protocols :

- *Proactive (table driven) routing protocols*
- *Reactive (on demand) routing protocols*
- *Flow oriented routing protocols*
- *Adaptive (situation aware)routing protocols*
- *Hybrid (both proactive and reactive) routing protocol*

and so on.

Now, we will concentrate only on two of the most popular MANET routing protocols, viz., Reactive routing protocol and Proactive routing protocol, how they work, algorithms that follow these protocols (AODV and OLSR respectively) and

also the attacks they are prone to and how to secure them from these attacks.

A. Proactive or table driven routing

In this protocol, proper and updated routing information is maintained between every pair of node in the network. So, the whole network topology should be known to each and every node. This is done by propagating routing information among the nodes at a fixed interval of time and since this routing information is maintained in a table format, proactive routing is also known as ‘table driven routing’. As routing information is propagated among the nodes of the network, this protocol causes a routing traffic overhead, but as the nodes get to know the topology of the network, there is no initial delay in communication.

Some proactive protocols are:

- Destination Sequences Distance Vector (DSDV) routing
- Optimized Link State Routing (OLSR)
- Clustered Gateway Switched Routing (CGSR)
- Wireless Routing Protocol(WRP) etc.

Now, OLSR is elaborated.

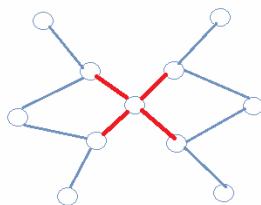


Figure 3. Multipoint relaying

• Optimized Link State Routing:

OLSR is nothing but a variation of the traditional Link State Routing , improved for application in mobile ad hoc networks. Unlike the Link State Routing , OLSR floods network information across the network in an optimized manner , so as to preserve bandwidth (due to power constraints in MANET). This optimization is based on something called MultiPoint Relaying (MPR). MPR(Figure 3.) helps to reduce the overhead caused when the network information is flooded.

For a node , its MPR set is the set of its neighbors that are at most two hops away from itself. Nodes get to know about their neighbors through HELLO messages which are exchanged at a fixed interval of time. This HELLO message sent by a node contains a list of neighbors and also an attribute which includes the link directionality of each neighbor. Once the nodes know their MPR set , routes within the network can be determined.

Topology Control (TC) messages are exchanged by the nodes in order to get the topology information of the network. A TC message for a particular node lists its neighbors whose MPR set has this node. This is called the multipoint relay sector of the node.

When a node receives a TC message from its fellow nodes , it can create new entries or modify the old entries of its routing table accordingly.

Attacks: OLSR is prone to several attacks.

A node may generate incorrect control messages. A misbehaving node M may send HELLO message on behalf of some other node X. (See fig.3). Now nodes A and B can announce their reachability to X through their control messages, even though actually they don't have reachability to X. This results in conflicting routes to X , which may also lead to connectivity loss.

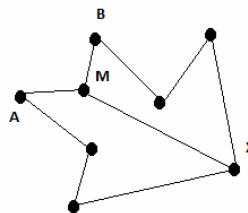


Figure 4. Malicious node sending Hello message to A and B on behalf of node X.

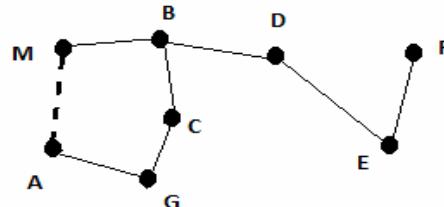


Figure 5. Malicious node M advertising a high quality route to node A thus performing link spoofing attack

Link Spoofing is also another issue to worry about in OLSR. In a control message, signalization of an incorrect set of links is called link spoofing. For example, a compromised node M can advertise a link to a node A. Now another node B will have a wrong MPR set because of this and the messages generated by node D will not reach A , but will reach the malicious node X (Figure 5.)

TC messages with spoofed originator address can also cause untrue neighbor relationship to be advertised in the network. However, for this attack to be successful the generated TC message must bear an Advertised Neighbor Sequence Number (ANSN) greater than the highest ANSN value referenced to the node on behalf of whom the adversary has sent a TC message.

TC messages with spoofed links can also perturb the topology of the network and in this process an adversary can get all the information meant for other nodes.

If the control messages are not properly relayed the network becomes prone to attacks like Black hole attack, Wormhole attack and Replay attack.

An attacker can perform Black hole attack by not relaying TC messages. In such a case , the network will suffer from connectivity problems as well as connectivity loss.

Replay attacks can also be performed if the attacker records old valid HELLO messages and resends them to the nodes. The nodes update their routing tables with wrong routing information which again results in the breakdown of the connectivity of the network.

Wormhole attack can be performed when an extraneous link say XY is created by an adversary A , by tunneling control messages between X and Y. However, for this attack to be successful , A has to wait until enough HELLO messages are exchanged through the wormhole. This will help to establish a symmetric link and TC messages are not processed if the relater node, i.e. the last hop is not a symmetric neighbor.

Securing OLSR: [8] proposes a method for securing OLSR. They propose an architecture where public key cryptography is used. Here both authenticated and non authenticated nodes are verified and besides verification ‘traceability’ is also taken into account so that if a trusted node misbehaves , it can be identified among other innocent nodes.

The architecture targets two different kinds of cryptography which are

- i. identity based, where public key of a node is derived from some identity of the node
- ii. traditional asymmetric technique, where a public key exists and it must be distributed possibly with a certificate.

The proposed architecture can be summarized as follows :

- An Authentication Authority (which is not necessarily in the ad hoc network) in charge of assigning keys to each node participating in the network. This key is denoted as the global key(Figure 6.).

- Key Distribution: Any node joining the network can diffuse its public key to the network with proper parameters, certificates and signatures. The key which is later used to sign messages is known as ‘local key’. It can be either the global key or a newly generated public/private key. (Figure 7.)

When the global key is distributed with traditional asymmetric cryptography, a Key Distribution message is originated by the node. This message includes the global public key, a certificate of the global public key, including a node identifier (originator IP address). If the local key is different from the global key, all the parameters of the local key are included. This Key Distribution message is signed with the global key. Now the message is distributed by flooding it across the network.

If the global key is identity based, no certificate is needed and the information about the local key is transmitted which is secured with a signature.

- Protocol message signing: When a node generates a control message, it is signed with the local key with a specific extension.

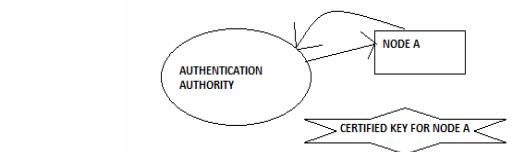


Figure 6.

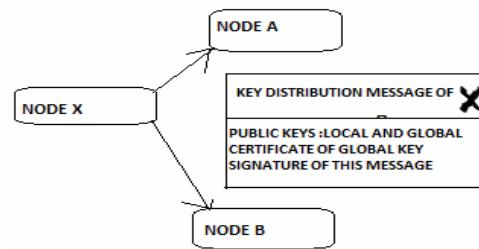


Figure 7

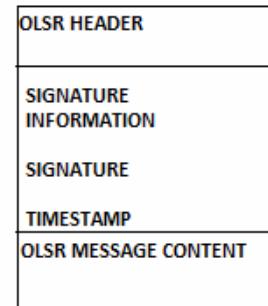


Figure 8. OLSR message signature and message

In order to prevent the Wormhole attack, a variant of the counting technique is used in this architecture. Here, each node periodically declares how many messages it has received from its neighbors. A counting mismatch is detected if a compromised link drops a packet.

For preventing mutation of packets by compromised links, each node advertises a set of hash-es of the packets received over each of last ‘k’ intervals. This allows the neighboring nodes to check whether packet loss has crossed a threshold. In case it does, the link is assumed to be compromised and no longer advertised.

Replay attacks are prevented by traditional use of timestamps. A message is rejected if the timestamp is too old.

B. Reactive or On Demand

In Reactive protocols routes are set up on demand , where the a node initiates demand for a route to another node with which it wants to communicate and it does not have a

route to reach the destination node. Once a route is established between two nodes , it is maintained until it is no longer used or the destination node has somehow become inaccessible.

Some Reactive routing protocols are as follows:

- Dynamic Source Routing (DSR)
- Temporarily Ordered Routing Algorithm(TORA)
- Ad hoc On demand Distance Vector (AODV) routing

Now we'll study AODV routing protocols, various attacks possible while using this protocol and relevant security measures.

When a node wants to communicate with another node, but it does not have a path to it, it starts the process of discovering paths to reach the destination node. It generates a Route REquest message which is broadcasted to its neighbors and so on. This is continued until the destination node is reached or some other node is reached which has a valid route to the destination .AODV uses Destination Sequence Number (DSN) to ensure that all routes are loop free and routing information is proper and valid. Each node in the network has its unique ID and sequence number. This sequence number is incremented while RREQ is forwarded from node to node. Every RREQ has broadcast id , sequence number, the source node's IP address and the destination node's recent sequence number.During the process of forwarding, the RREQ packets, the intermediate nodes record the address of the neighbor from whom the first copy of broadcast message is received in their routing tables. This helps to establish a reverse path.

When the RREQ message is finally received by the destination or the node having a valid route to the destination, then the destination or this other node responds by unicasting a route reply (RREP) packet back to the source. The path which this RREP packet follows is the path which was established by the intermediate nodes while relaying RREQ.

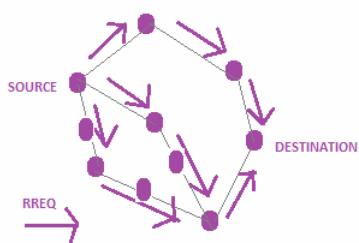


Figure 9.

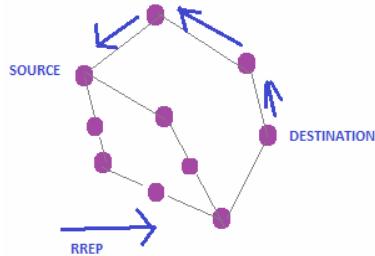


Figure 10.

If the newly established route is not used for a certain time or if there is any kind of link failure then the source node again has to find another route to the destination. Due to the on demand route establishment nature of the protocol there is an initial delay in communication.

Connectivity between the nodes in a particular route is maintained by periodic exchange of HELLO messages.

Attacks:

Attacks like Modification attack, Impersonation (spoofing) and attacks using fabrication (black hole attack, wormhole attack) etc poses threat to AODV routing protocol.

Attacks using modification: AODV uses the hop count parameter to determine the shortest path between two nodes. A malicious node can set false hop counts. It can also set wrong value of sequence number. This leads redirection of network traffic. DoS attack is also possible by source route modification.

Attacks using Impersonation: By impersonation, an adversary can perform several attacks e.g. a traffic that belongs to a compromised node can be redirected to the malicious node. Loops may also be formed. AODV is also prone to Sybil attack where a mailicious node takes up the identity of several nodes.

Attacks using fabrication: In this type of attack a malicious node generates false routing information. Some fabrication attacks are as follows -

- *Blackhole attack:* Here, an adversary pretending to be the destination node sends forged RREP packets to the source node(which initiated the RREQ message).The RREP message which is sent by the attacker has higher DSN than the DSN of the real destination. The source judging the packet by its DSN accepts the one with the greater value. So we see that the source becomes unable to communicate with the actual destination and inadvertently sends all the messages (meant for the actual destination) to the malicious node. In this way the malicious node lures the source to send all its messages to it and it continues to do the same with other nodes of he network, gradually creating a black hole in the network. All messages from various sources are directed toward this black hole.

- *Wormhole attack:* It is a kind of a replay attack and it is hard to defend against. It is even effective when all the routing information is encrypted, confidential and

authenticated. In wormhole attack the attacker directs the RREQ message through a tunnel straight to the destination node keeping the hop count unincremented. The attack generally involves two remote nodes, say, X and Y, which are connected via a wormhole link. They target to attack those nodes which have something to send to some other node in the network.

In Figure 11. we see that when source S broadcasts an RREQ message to find its way to the destination D, A and C receives it. Now when X receives RREQ(forwarded by A) it records and tunnels the RREQ to Y. Now Y forwards it to B and finally RREQ reaches D. Again RREQ reached D through another route S-C-P-E-G-F-D but the RREQ reaching

D through the other path reaches faster. So D ignores the message received through S-C-P-E-G-F-D route as it has a higher hop count than RREQ from S-A-X-Y-B-D route. Now D unicasts RREP through the route S-A-B-D and S also sends message to D through route S-A-B-D. Thus all the data passes through the wormhole between the malicious nodes X and Y.

These malicious nodes can also transmit the eavesdropped messages to some other channel available to the attacker. The wormhole attack can also be coupled with message dropping attack to prevent destination nodes fro receiving packets meant for them. As a result securing AODV against wormhole attack is a big challenge.

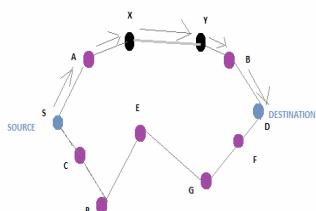


Figure 11.

*Securing AODV :*Secure AODV (SAODV)[3] is an improved version of AODV which helps in securing AODV from several attacks. SAODV is based on public key cryptography and all the routing messages are digitally signed so that integrity and authenticity of the messages are maintained.

In SAODV when a node generates any routing message it signs it with its private key and the nodes which receive the message verifies the signature with the sender's public key. As the sender cannot sign the hop count, SAODV uses a hash chain mechanism. Again RREP message is signed by the destination itself.

SAODV has a feature called 'double signature'. When a node, say node A generates a RREQ message it includes a second signature which is computed on a fictitious RREP message towards A. Now the intermediate nodes can store this second signature in their routing table along with other information related to node A. Now if one of the intermediate nodes receive RREQ towards A later sometime it can reply on behalf of A with a RREP (this is similar to what happens in AODV).

This RREP generated by the intermediate node includes the signature by node A which it stored along with other information related to node A in its routing table. It then signs the message with its own private key.

As mentioned earlier SAODV uses hash chains to authenticate the hop count of the routing messages. Every time a node wants to send a routing message , it generates a random number (seed) . The hash field in the signature extension is set to the seed. The Top Hash field is set to the seed hashed Max hop count times. Every time a node receives a routing message , it verifies the hop count by hashing the Max hop count times the Hash field and checking if it is same as Top Hash. If the check fails, then the node drops the packet.

In this way SAODV provides security to AODV.

V. CONCLUSION

Security is one of the biggest issues in mobile ad hoc network. In this paper we have discusses of two of the most popular MANET routing protocols , OLSR and AODV protocols. We briefly discussed some security problems in MANET such as eavesdropping. Impersonation attack , Black hole attack, Worm hole attack and so on. Finally we made a study of some secure versions of these two protocols. These secure architectures promise better and secure routing in mobile ad hoc networks , and also provides the basic security attributes such as authentication , confidentiality , integration and non repudiation.

REFERENCES

- [1] Satoshi Kurosawa1, Hidehisa Nakayama1, Nei Kato1, Abbas Jamalipour2, and Yoshiaki Nemoto1_(Corresponding author: Hidehisa Nakayama)___Graduate School of Information Sciences, Tohoku University1,Aoba 6-3-09, Aramaki, Aoba-ku, Sendai, Miyagi 980-8579, Japan. (Email: hidehisa@it.ecei.tohoku.ac.jp),School of Electrical and Information Engineering, The University of Sydney, Sydney NSW 2006,Australia2
- [2] MANET Routing Protocols and Wormhole Attack against AODV
Jatin D. Parmar ,Ashish D. Patel2 ,Rutvij H. Jhaveri1 , Bhavin I. Shah4
Department of Computer Engineering and Information Technology
S.V.M. Institute of Technology
Bharuch, India
- [3] Performance Comparisons of AODV, Secure AODV and Adaptive Secure AODV Routing Protocols in Free Attack, Simulation Environment Mohd Anuar Jaafar ,Zuriati Ahmad Zukarnain Dept. of Communication Technology and Networks Faculty of Computer Science and Information Technology University Putra.
- [4] Extension to the OLSR protocol
Andreas Hafslund, Andreas Tønnesen, Roar Bjørgum Rotvik, Jon Andersson and Øivind Kure
- [5] A Security-Aware Routing Protocol for Wireless Ad Hoc Networks Seung Yi, Prasad Nalburg, Robin Kravets fseungyi,nalburg,rhkg@cs.uiuc.edu
Dept. of Computer Science University of Illinois
- [6] Mobile Ad Hoc Networking: An Essential Technology for Pervasive Computing Jun-Zhao Sun MediaTeam, Machine Vision and Media Processing Unit, Infotech Oulu P.O.Box 4500, FIN-90014 University of Oulu, Finland
- [7] Chris Karlof , David Wagner Secure routing in wireless sensor networks:attacks and countermeasures , University of Berkeley CA 94720 USA
- [8] C'edric Adjih, Daniele Raffo, Paul M'uhlethaler
INRIA, Domaine de Voluceau, France1t.

ANALOG & DIGITAL COMMUNICATION

Date : 18th January,2012

Chairman : Prof. Tapan Kumar Chatterjee

Co-Chairman : Prof. Asim Kar

Rapporter : Prof. Moumita Chakraborty

| Paper ID | Paper Title | Authors |
|----------|--|---|
| 7 | Performance Analysis of Transistor Level Adder Circuits by Varrying Number of Transistors and Implementing Different Logic Style of Addition | Subhadeep Nandy and Rajib Pal |
| 15 | A Binary Adaptive Differential Evolutionary Approach for Multi_User Detection in Multi_Carrier Code Division Multiple Access | Jubin Hazra and Arindam Chakraborty |
| 21 | Design Of Sequential Circuits Using Single Electron Encoded Logic | Mili Sarkar |
| 34 | Implementing Security Measurement on qa Lifetime Enhancing Routing Protocol of WSN | Avik Modak and Anuran Roychowdhury |
| 45 | A Novel Multiplexer and Demultiplexer Design and Permissible Fault Analysis Using Quantum Cellular Automata | Ratna chakroborty, Kamal Kumar Ghosh and Debasis De |
| 58 | Hairpin model of CRLH Transmission Line and Floating Slot Approach for Wide -Band Bandpass Filter | Pratik Mondal,Bhupesh Mukherjee and Susanta Kumar Parui |
| 61 | Programmable Array Logic Design Using Quantum Dot Cellular Automata | Ratna Chakrabarty,Arnab Sarkar and Debasis De |

A Binary Adaptive Differential Evolutionary Approach for Multi-User Detection in Multi-Carrier Code Division Multiple Access

Jubin Hazra¹, Student Member IEEE, Arindam Chakraborty¹

¹ Electronics and Telecommunication Engineering, Institute of Engineering and Management
Kolkata, India

Email:jubinhazra.iem@gmail.com, arindam365@gmail.com

Abstract— Multi-Carrier Code Division Multiple Access (MC-CDMA) is an emerging wireless communication technology that incorporates the advantages of Orthogonal Frequency Division Multiplexing (OFDM) into the original Code Division Multiple Access (CDMA) technique. But it suffers from the inherent defect called Multiple Access Interference (MAI) due to inappropriate cross-correlation possessed by the different user codes. To reduce MAI, the multi-user detection (MUD) technique has already been proposed in which MAI is treated as noise. Due to high computational cost incorporated by the optimal MUD detector with increasing number of users, researchers are looking for sub-optimal MUD solutions. This paper proposes a new modified adaptive binary Differential Evolution algorithm with a novel crossover strategy (MBDE_pBX) for multi-user detection in a synchronous MC-CDMA system. Since MUD detection in MC-CDMA systems is a problem in binary domain a binary encoding rule is introduced which converts problem of binary domain of any number of dimension into 4-dimensional continuous domain problem. The BER performance of MBDE_pBX has been compared with the binary version of the Genetic algorithm (GA) and Particle Swarm Optimization (PSO). The simulation results show that this new binary DE variant can achieve superior bit error rate performance (BER) within much lower optimum solution detection time outperforming its competitors as well as achieving 99.62% reduction in computational complexity as compared to the MUD scheme using exhaustive search.

Index Terms Multi-Carrier CDMA (MC-CDMA), Multi User Detection (MUD), Differential Evolution, p-best Crossover, Parameter Adaptation.

I. INTRODUCTION

Code division multiple access (CDMA) [1] is a very popular broadband wireless transmission technique being a channel access method used by various radio communication technologies. CDMA employs spread spectrum technology and a special coding scheme (where each transmitter is assigned a code) to allow multiple users to be multiplexed over the same physical channel. In recent times, the concept of CDMA is extended to a multi-carrier code division multiple access (MC-CDMA) [2] where the idea of orthogonal frequency division multiplexing is incorporated in CDMA to improve the performance of CDMA as well as accomplishing high frequency diversity gain. Here the spreading sequence of different users are modulated by all the available sub-carriers over the same frequency band removing the multi-path

problem in CDMA and the users are able to compromise the same frequency band. However, for a perfect CDMA system, the spreading codes assigned to each specific user should be perfectly orthogonal possessing no cross-correlation between them. But if the cross-correlation values are finite, it can give rise to Multiple Access Interference (MAI) between the different users. In second generation CDMA systems, specifically, in the Qualcomm IS-95 CDMA standard, the multiple access interference (MAI) is treated as noise. To mitigate MAI, an idea regarding joint detection of all users in a CDMA system has been put forward where MAI has been estimated as a noise. This idea is regarded as multi-user detection (MUD) [3].

The Differential Evolution (DE) [4] algorithm has emerged as a very competitive form of evolutionary computing more than a decade ago which appeared as a technical report by Rainer Storn and Kenneth V. Price in 1995 [4]. In this article we propose an innovative binary Differential Evolution algorithm which has been shown to serve as a suitable MUD detector with huge reduction in computational complexity. In this paper, an encoding rule is presented to generate a bit vector in the binary domain to compute the value for the objective function for the MUD detector from the continuous domain DE operation. This rule basically converts the binary domain problem into a 4-dimensional continuous domain problem. Now, several modifications have been proposed to improve the performance of the Differential Evolution (DE) in the 4-dimensional continuous domain. To increase the convergence speed a novel exploitative crossover mechanism called *p*-best crossover has been incorporated. Novel and simple adaptation rules have been proposed to update the scale factor (*F*) and crossover probability (*Cr*) values utilizing the fitness value in order to increase the robustness of MBDE_pBX. The results obtained with the proposed algorithm MBDE_pBX as a MUD detector for the MC-CDMA system has been shown in two ways. At first, BER performance has been demonstrated for the MBDE_pBX against the binary versions for two of the state-of-the-art evolutionary algorithms namely Genetic algorithm (GA) [5] and Particle Swarm Optimization (PSO) [6]. Secondly, it has been shown that MBDE_pBX is successful to achieve huge reduction in computational overhead as compared to the traditional exhaustive search technique. This superior convergence performance can be attributed to the proposed modifications in the algorithmic

components of MBDE_pBX ensuring massive reduction in computational cost (99.62%) in comparison with the exhaustive search algorithm. The algorithmic components of MBDE_pBX do not incur any additional computational overhead as they are implemented based on simple DE operators.

II. MULTI CARRIER CDMA (MC-CDMA)

Orthogonal frequency division multiplexing (OFDM) [7] is a parallel data transmission scheme in which high data rates can be achieved by transmitting several orthogonal subcarriers. OFDM systems reduce the inter symbol interference (ISI) as well as the inter channel interference (ICI) by the insertion of guard intervals. Recently CDMA systems based on the combination of CDMA schemes and OFDM signaling, which are referred to as multi-carrier CDMA systems, have attracted much attention in the wireless digital communication [8]. This is mainly due to the fact that there is a need to enhance high data rate services in wireless environment.

Based on signal spreading model multicarrier can be categorized mainly into two types. In the first class of schemes the serial data stream is first spread by a spreading code and then converted into parallel chip sequences with each chip modulating a different subcarrier. The number of subcarrier equals the number of chips per data symbol. Spreading operation in this type of multicarrier CDMA arrangements occurs in the frequency domain. This type of systems combines robustness of orthogonal modulation with the flexibility of the CDMA schemes [9]. In the second type of multicarrier CDMA system the original data stream is first serial to parallel converted into sub-streams. Then each sub-streams. Then each sub-stream is spread using a given spread in the time domain and finally modulates a different subcarrier with each of the data stream.

III. PROBLEM FORMULATION FOR MC-CDMA

In this section, we consider a bit-synchronous MC-CDMA system where K users simultaneously transmit their bits over an idealized AWGN channel. The number of chips is equal to the number of sub-carriers where each user bit is first spread in the frequency spectrum by a specific pseudo-noise sequence and then modulated by "M" available sub-carriers by means of orthogonal frequency division multiplexing (OFDM). The k^{th} user's transmitted signal is given by,

$$s_k(t) = \sqrt{\frac{2E_k}{M}} \sum_{m=1}^M c_{k,m}(t) \cdot b_k(t) e^{j\omega_m t} \quad (1)$$

Here, E_k : The transmit power of the k^{th} user ($k = 1, \dots, K$)

M : Number of sub-carriers,

T_b : Data bit duration,

$b_k(t)$: The transmitted bit sequence of the k^{th} user.

$c_{k,m}(t)$: Time-domain representation of the k^{th} user's spreading sequence on the m^{th} sub-carrier.

$c_k(t)$: Time-domain representation of the k^{th} user's spreading code over all sub-carriers.

$\omega_m = 2\pi f_m$, where $f_m (m = 1, \dots, M)$ are the sub carrier frequencies and $f_m = f_1 + (m-1)\Delta f$, where frequency spacing is $\Delta f = \frac{1}{T_b}$.

The Figure 1 shows a simple MC-CDMA receiver followed by a binary DE based MUD. The demodulated data on the sub carrier frequency f_1 yields a composite part of the spreading signals of all users, contributed by the 1^{st} chips of their respective spreading sequences. Similarly, for the sub-carriers f_2 till f_M , the contribution is from the set of the 2^{nd} chips till the last set of chips of the pseudo-random sequences from all users respectively.

The received signal on the m^{th} sub-carrier is,

$$r_m(t) = \sum_{k=1}^K \left[\sqrt{\frac{2E_k}{M}} c_{k,m}(t) \cdot b_k(t) \right] + n(t) \quad (2)$$

Here $n(t)$ is the AWGN noise added to the signal with two-sided power spectrum density $\frac{N_0}{2}$. Now, the joint optimum decision rule for a K -user discrete-time synchronous CDMA model can be easily extended to a K -user multi-carrier CDMA (MC-CDMA) system model where the data is transmitted over M number of orthogonal sub carriers. So, the contribution of all the likelihood functions will yield the optimum vector \hat{b} .

$$\Omega(b) = \sum_{m=1}^M \Omega_m(b) = \sum_{m=1}^M \left\{ 2\Re \left[b^T \xi C^* Z_m \right] - b^T \xi C R_m C^* \xi b \right\}, \quad (3)$$

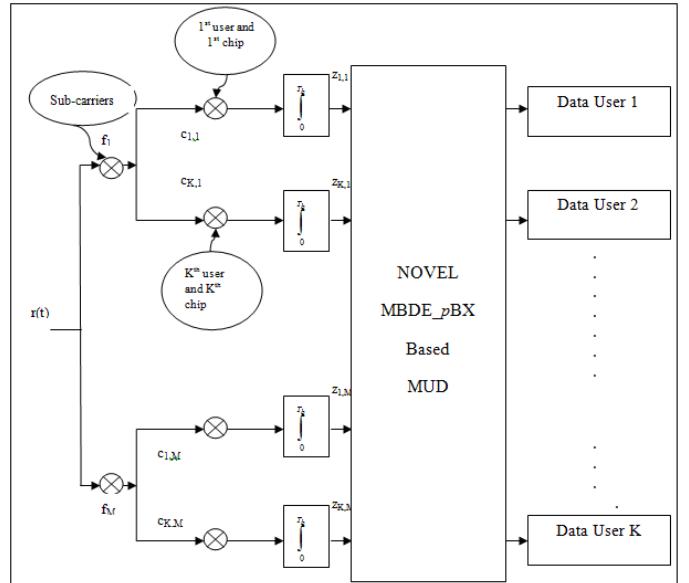


Figure 1. A typical MC-CDMA Reciever Model

where $\Omega_m(b)$ denotes the discrete-time correlation metric on the m^{th} sub-carrier. Our aim is to search for the specific bit sequence b out of 2^K possible combinations that maximizes the correlation metric of equation (3).

IV. CLASSICAL DIFFERENTIAL EVOLUTION

DE is a simple real parameter optimization algorithm. We explain each stage separately in the following subsections.

A. Initialization of the Parameter Vectors

DE searches for a global optimum point in a D -dimensional real parameter space \Re^D . It begins with a randomly initiated population of N_p D dimensional real-valued parameter vectors. Each vector, also known as *genome/chromosome*, forms a candidate solution to the multi-dimensional optimization problem. We shall denote subsequent generations in DE by $G = 0, 1, \dots, G_{\max}$. We adopt the following notation for representing the i -th vector of the population at the current generation:

$$\vec{X}_{i,G} = [x_{1,i,G}, x_{2,i,G}, x_{3,i,G}, \dots, x_{D,i,G}]. \quad (4)$$

The initial population (at $G = 0$) should cover the range as much as possible by uniformly randomizing individuals within the search space constrained by certain minimum and maximum bounds:

$$\vec{X}_{\min} = \{x_{1,\min}, x_{2,\min}, \dots, x_{D,\min}\} \text{ and}$$

$$\vec{X}_{\max} = \{x_{1,\max}, x_{2,\max}, \dots, x_{D,\max}\}.$$

Hence we may initialize the j -th component of the i -th vector as:

$$x_{j,i,0} = x_{j,\min} + \text{rand}_{i,j}[0,1] \cdot (x_{j,\max} - x_{j,\min}), \quad (5)$$

where $\text{rand}_{i,j}[0,1]$ is a uniformly distributed random number lying between 0 and 1 (actually $0 \leq \text{rand}_{i,j}[0,1] \leq 1$).

B. Mutation with Difference Vectors

After initialization, DE creates a *donor* vector $\vec{V}_{i,G}$ corresponding to each population member or *target* vector $\vec{X}_{i,G}$ in the current generation through mutation. It is the method of creating this donor vector, which differentiates between the various DE schemes. Five most frequently referred mutation strategies implemented in the public-domain DE codes available online at <http://www.icsi.berkeley.edu/~storn/code.html> are listed below:

$$\text{"DE/rand/1": } \vec{V}_{i,G} = \vec{X}_{r_1^i,G} + F \cdot (\vec{X}_{r_2^i,G} - \vec{X}_{r_3^i,G}). \quad (6a)$$

$$\text{"DE/best/1": } \vec{V}_{i,G} = \vec{X}_{best,G} + F \cdot (\vec{X}_{r_1^i,G} - \vec{X}_{r_2^i,G}). \quad (6b)$$

"DE/target-to-best/1":

$$\vec{V}_{i,G} = \vec{X}_{i,G} + F \cdot (\vec{X}_{best,G} - \vec{X}_{i,G}) + F \cdot (\vec{X}_{r_1^i,G} - \vec{X}_{r_2^i,G}). \quad (6c)$$

"DE/best/2":

$$\vec{V}_{i,G} = \vec{X}_{best,G} + F \cdot (\vec{X}_{r_1^i,G} - \vec{X}_{r_2^i,G}) + F \cdot (\vec{X}_{r_3^i,G} - \vec{X}_{r_4^i,G}). \quad (6d)$$

"DE/rand/2":

$$\vec{V}_{i,G} = \vec{X}_{r_1^i,G} + F \cdot (\vec{X}_{r_2^i,G} - \vec{X}_{r_3^i,G}) + F \cdot (\vec{X}_{r_4^i,G} - \vec{X}_{r_5^i,G}). \quad (6e)$$

The indices $r_1^i, r_2^i, r_3^i, r_4^i$, and r_5^i are mutually exclusive integers randomly chosen from the range $[1, N_p]$, and all are different from the index i . These indices are randomly generated once for each donor vector. The scaling factor F is a positive control parameter for scaling the difference vectors. $\vec{X}_{best,G}$ is the best individual vector with the best fitness (i.e. lowest objective function value for minimization problem) in the population at generation G .

C. Crossover

To enhance the potential diversity of the population, a crossover operation comes into play after generating the donor vector through mutation. The donor vector exchanges its components with the target vector $\vec{X}_{i,G}$ under this operation to form the trial vector $\vec{U}_{i,G} = [u_{1,i,G}, u_{2,i,G}, u_{3,i,G}, \dots, u_{D,i,G}]$.

On the other hand, binomial crossover is performed on each of the D variables whenever a randomly generated number between 0 and 1 is less than or equal to the Cr value. In this case, the number of parameters inherited from the donor has a (nearly) binomial distribution. The scheme may be outlined as:

$$u_{j,i,G} = \begin{cases} v_{j,i,G}, & \text{if } (\text{rand}_{i,j}[0,1] \leq Cr \text{ or} \\ & x_{j,i,G}, \text{ otherwise,} \end{cases} \quad (7)$$

where, as before, $\text{rand}_{i,j}[0,1]$ is a uniformly distributed random number, which is called anew for each j -th component of the i -th parameter vector. $j_{rand} \in [1, 2, \dots, D]$ is a randomly chosen index, which ensures that $\vec{U}_{i,G}$ gets at least one component from $\vec{V}_{i,G}$. It is instantiated once for each vector per generation. We note that for this additional demand, Cr is only approximating the true probability p_{Cr} of the event that a component of the trial vector will be inherited from the donor.

D. Selection

To keep the population size constant over subsequent generations, the next step of the algorithm calls for *selection* to determine whether the target or the trial vector survives to the next generation i.e. at $G = G + 1$. The selection operation is described as:

$$\begin{aligned} \vec{X}_{i,G+1} &= \vec{U}_{i,G}, & \text{if } f(\vec{U}_{i,G}) \leq f(\vec{X}_{i,G}) \\ &= \vec{X}_{i,G}, & \text{if } f(\vec{U}_{i,G}) > f(\vec{X}_{i,G}), \end{aligned} \quad (8)$$

where $f(\vec{X})$ is the objective function to be minimized. Therefore, if the new trial vector yields an equal or lower value of the objective function, it replaces the corresponding target vector in the next generation; otherwise the target is retained in the population. Hence, the population either gets better (with respect to the minimization of the objective function) or remains the same in fitness status, but never deteriorates. Note that in eqn (8) the target vector is replaced by the trial vector even if both yields the same value of the objective function – a feature that enables DE-vectors to move over flat fitness landscapes with generations.

V. MODIFIED BINARY DIFFERENTIAL EVOLUTION WITH P-BEST CROSSOVER (MBDE_PBX)

Differential Evolution has already emerged as a powerful tool for optimisation of continuous space functions. The multi-user detection (MUD) mechanism in a synchronous multi-carrier CDMA system is a combinatorial maximization problem. But the continuous nature of DE does not permit the algorithm to apply to combinatorial optimization problems. In this article a novel binary version of a modified Differential Evolution named as MBDE_pBX is designed to solve the optimum K -user MUD detection problem in MC-CDMA system. MUD detection demands fast convergence towards near-optimal solutions without incorporating large amount of computational complexity.

A. Initial Population

The appropriate selection of initial population is necessary for any evolutionary algorithm to approach optimal fitness. To create an initial population we devise a method by which each component of the population vector is uniformly initialized within the search bound [0, 1]. Here, we present the excellent concept of orthogonal array which are used in experimental simulation methods. If, in certain design purpose there are m factors and each factor have n levels then to find appropriate setting of each factor's level, $m \times n$ experiments must be done. So, it is required to generate a brief but representative sample of combinations for testing. The orthogonal arrays were developed for this purpose. For an experiment having m factors and each having n levels, an orthogonal array is an array with m rows and n columns which is a representative sample of some testing experiments that satisfies the condition that for the factors in any column, every level occurs the same number of times. For each column of $[SOA]_{m \times n}$, a random permutation of 1, ..., m is generated and denoted as sequence Z. Then the elements in Z are picked sequentially one by one and filled randomly in the column. When all elements in Z were picked, the process starts once again from the beginning of sequence. So in every column of $[SOA]_{m \times n}$, each of m elements will appear the same number of times. In the beginning, the proposed algorithm generates initial solutions by using the simulated orthogonal array which ensures uniform initialization within the search range [0, 1].

B. Scale Factor Adaptation

In this article we aim at elevating F whenever a particle is situated far away from the favourable region where the suspected optima lies, i.e. the fitness value of the particle differs much from the best solution value. On the other hand we should reduce F whenever the objective function value of the particle nears the best solution. These particles are subjected to lesser perturbation so that they can finely search the surroundings of some suspected optima. The scheme may be outlined as:

$$F_i = F_{\min} + (F_{\max} - F_{\min}) \cdot \frac{\Delta f_i}{1 + \Delta f_i} \quad (9)$$

where F_i represents the value of the scale factor for the i^{th} target vector and $\Delta f_i = |f(\vec{X}_i) - f(\vec{X}_{best})|$.

F_{\min} and F_{\max} are the lower and upper bounds of F set to 0.2 and 0.9 respectively. As clear from equation (9), F_i depends

on the factor $\frac{\Delta f_i}{1 + \Delta f_i}$. The factor $\frac{\Delta f_i}{1 + \Delta f_i}$ can be modified to $\frac{1}{1 + \frac{1}{\Delta f_i}}$. For a target vector if Δf_i is large $\frac{1}{\Delta f_i}$ decreases,

so the factor $\frac{1}{1 + \frac{1}{\Delta f_i}}$ increases, so F_i gets enhanced and the

particle will be subjected to larger perturbation so that it can jump to a favorable region in the landscape. For the best individual $\Delta f_i = 0$, so F_i will be equal to F_{\min} (0.2) which is evident because the best vector is required to undergo a lesser perturbation so that it can perform a fine search within a small neighborhood of the suspected optima. Thus the particles which are distributed away from the current fittest individual have large F_i values and keeps on exploring the search space, maintaining sufficient population diversity.

C. The p-Best Crossover

The crossover operation named as p -best crossover where for each donor vector, a vector is randomly selected from the $p\%$ top-ranking vectors (according to their objective function values) in the current population and then normal binomial crossover is performed as per eqn. (7) between the donor vector and the randomly selected p -best vector to generate the trial vector at the same index. This novel crossover scheme promotes exploitation and improves the convergence performance of MBDE_pBX.

D. Crossover Probability Adaptation (Cr)

The adaptation of Cr is also devised based on the difference between the objective function value of the target and the best particle. The concept follows that if the solution of the target

vector is better Cr should be smaller, otherwise Cr should be assigned a larger value. So the solution of the target vector is utilized to make a decision whether Cr is larger or smaller. To assist the convergence of the target vector, the value of Cr controls the flow of genetic information to the trial vector. Suppose that a particle is situated in an adverse region of the fitness-space. During the crossover operation the donor vector created by perturbing the target particle should inject information to the trial vector to a greater extent so that it can switch to a favorable area in the fitness landscape. So the value of Cr should be relatively high in accordance with the binomial crossover scheme ensuring more contribution of genetic information from the donor vector. The adaptation scheme is given below:

$$Cr_i = Cr_{\min} + (Cr_{\max} - Cr_{\min}) \cdot (1 - e^{-\Delta f_i}) \quad (10)$$

where Cr_i represents the crossover probability value for the i^{th} target vector and $\Delta f_i = |f(\vec{X}_i) - f(\vec{X}_{best})|$

Cr_{\min} and Cr_{\max} are the lower and upper limits of Cr_i set to 0.1 and 0.94 respectively. As evident from equation (10), Cr_i depends on the factor $(1 - e^{-\Delta f_i})$. If Δf_i is large, the factor $(1 - e^{-\Delta f_i})$ attains a higher value resulting in larger crossover probability values.

VI. SIMULATION AND RESULTS

The simulation results are taken based on the three main performance criteria. They are Bit Error Rate (BER), optimum solution detection time and the associated computational burden.

A. Comparison of BER Performance with Binary GE and Binary PSO

In this section the BER performance of MBDE_pBX is compared with Binary DE referred to as BDE, Binary GA and Binary PSO for a 10-user MC-CDMA system under 9 dB SNR. For Binary DE (BDE), the mutation scheme applied is DE/current-to-best/1, normal binomial crossover and scale factor (F) and crossover probability (Cr) values kept at 0.9 and 0.9 respectively. For Binary GA and PSO, we employ the best-suited parametric set up chosen with guidelines from their respective literatures.

A close scrutiny of figure 2 reveals that MBDE_pBX can converge to the optimum within only 4 generations for a 10-user MC-CDMA system whereas BDE, Binary PSO and Binary GA converges to the optimal solution within 12, 14 and 16 generations under the same conditions. This superior convergence performance of MBDE_pBX can be attributed to the three modifications proposed in MBDE_pBX, especially the novel p -best crossover scheme ensuring good convergence speed. In figure 3, we have investigated the BER performance of all the algorithms with increasing SNR values. There is a problem associated with increasing SNR that the Multiple Access Interference gets increased. But even under such adverse condition the performance of MBDE_pBX is not

affected in a worst way as in the case of Binary DE (BDE), Binary GA and Binary PSO.

In Figure 4, we have shown how MBDE_pBX performs with the increasing number of users in the MC-CDMA system. MBDE_pBX is robust with the increasing number of users in a MC-CDMA system. As can be seen from the Figure 4, up to near about 14 users it can converge fully to the optimal solution, but when the number of users is increased further MBDE_pBX is able to converge to the near-optimal solution. This is due to the reason that as the number of users increases the computation complexity of the system involved also increases. So for large number of users, it basically becomes impossible to converge to the actual optima. As can be seen from figure 4, the performance of Binary DE (BDE), Binary GA and PSO has failed to be robust with increasing number of users and their BER performance have started deteriorating from much lower user level.

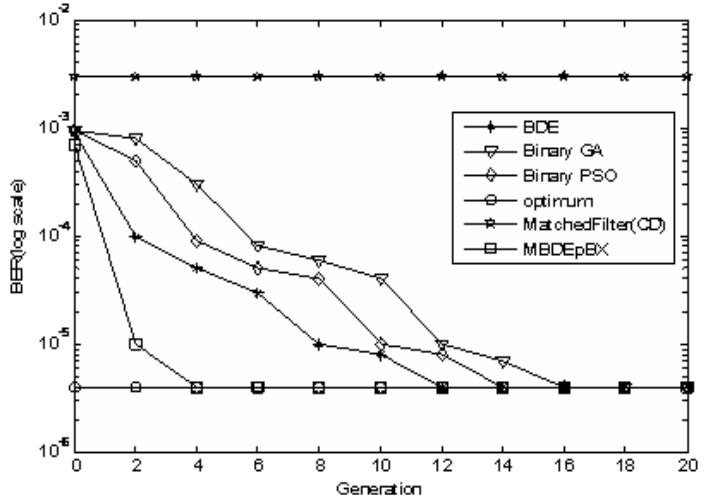


Figure 2. Convergence comparison of MBDE_pBX with Binary DE, Binary GA, Binary PSO and Matched Filter Conventional Detector for a 10-user MC-CDMA system under 9 dB SNR

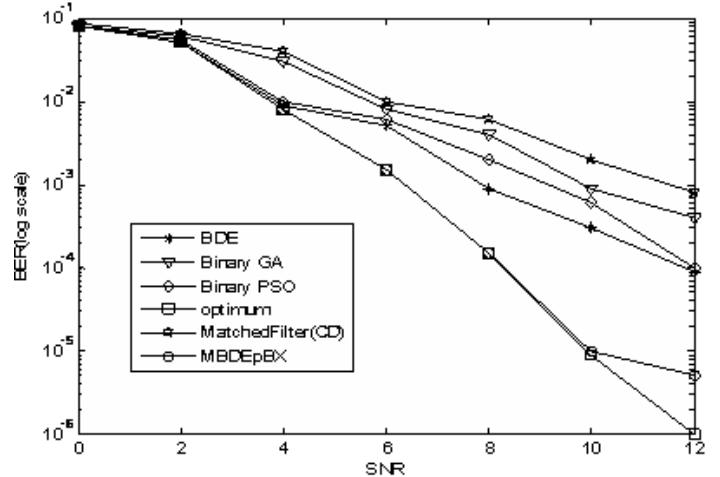


Figure 3. BER Comparison of MBDE_pBX vs. BDE, Binary GA, and Binary PSO for a 15-user MC-CDMA system with increasing SNR

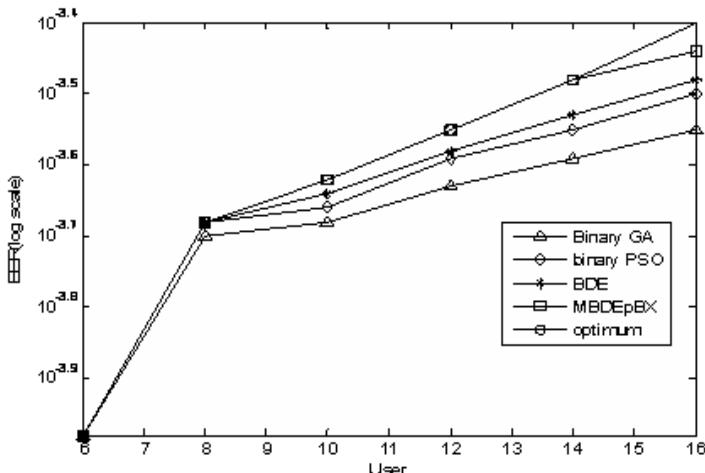


Figure 4. BER performance of MBDE_pBX for increasing number of users in a MC-CDMA System

B. Reduction in Computational Complexity achieved by proposed MBDE_pBX

In this article we have defined the computational complexity as the number of the cost function evaluation needed to reach the optimal solution. The exhaustive search method will require 2^K number of function evaluations to search for the optimal solution where K is the number of users. On the other hand, the number of cost function evaluations taken by MBDE_pBX is the product of the population size and the number of generations required to give the optimum solution. So the percentage of the complexity involved

$$\text{is } \left(\frac{\text{pop_size} \times \text{Gen}}{2^K} \right) \times 100\% \quad \text{where}$$

pop_size represents the population size and *Gen* is the number of generations involved. Now for $K = 15$ i.e. for a 15-user system MBDE_pBX can approach the near-optimum fitness within only 5 generations. So the number of cost functions involved is $25 \times 5 = 125$ as population size used is 25. So, percentage in computational complexity involved achieved by MBDE_pBX as compared to the traditional

exhaustive search is given by $\frac{125}{2^{15}} \times 100\% = 0.38\%$. So,

MBDE_pBX can attain 99.62% significant reduction in complexity as compared to the conventional exhaustive search.

VII. CONCLUSION

MC-CDMA has emerged as a popular wireless broadband transmission technique combining the advantages of both OFDM and CDMA. Nowadays, the numbers of users are gradually looking up and the complexity of the optimal MUD detector in MC-CDMA system also increases as a function of the number of users. So, researchers are now quite focused on developing suitable sub-optimal MUDs. The proposed algorithm MBDE_pBX has successfully emerged itself as a

potential sub-optimal MUD by carrying out a balanced ratio of exploration and exploitation throughout the evolutionary stages. The p -best crossover is exploitative and improves the convergence performance of MBDE_pBX overcoming the problems of slow convergence. At the same time, to avoid premature convergence a less exploitative mutation scheme and simple adaptation schemes of F and Cr based on the fitness information has been incorporated to increase the robustness of the algorithm. It is to be noted that these algorithmic components do not incur any additional computational complexity because they are realized based on simple DE operators. A close scrutiny of simulation and results reveals that MBDE_pBX has completely outperformed the binary version of the DE (BDE), GA and PSO in terms of BER performance as well as the optimum solution detection time measured in terms of number of generations taken for complete convergence. Hence, MBDE_pBX has been able to ensure huge reduction in computational complexity (99.62%).

REFERENCES

- [1] R. Rick and L. Milstein, "Parallel acquisition in mobile DS-CDMA systems," *IEEE Transactions on Communications*, vol. 45, pp. 1466–1476, November 1997.
- [2] S. L. Miller and B. J. Rambolt, "MMSE Detection of Multi-carrier CDMA," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 2356-2362, Nov 2000.
- [3] H. Wei, and L. Hanzo, "Reduced-Complexity Near-Optimum Genetic Assisted Multiuser Detection for Synchronous Multicarrier CDMA", *Proc. of 2004 IEEE VTC-Spring Conf.*, Milan (I), May 17-19 2004, Vol. 3, pp. 1717-1721.
- [4] R. Storn and K. V. Price, "Differential Evolution - a simple and efficient adaptive scheme for global optimization over continuous spaces", *Technical Report TR-95-012, ICSI*, <http://http.icsi.berkeley.edu/~storn/litera.html>, 1995.
- [5] Witold Kosiński, Stefan Kotowski, and Jolanta Socala, "On Asymptotic Behaviour of a Binary Genetic Algorithm", XXI Autumn Meeting of Polish Information Processing Society Conference Proceedings pp. 139–144.
- [6] J. Kennedy and R.C. Eberhart, "A discrete binary version of the particle swarm algorithm." In *Proceedings of the World Multiconference on Systems, Cybernetics and Informatics*, pages 4104-4109, 1997.
- [7] S. B. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Transactions on Communication Technology*, vol. 19, pp. 628–634, October 1971.
- [8] M. Kavehrad and B. Ramamurthy, "Direct-sequence spread spectrum with DPSK modulation and diversity for indoor wireless communications," *IEEE Transactions on Communications*, vol. COM-35, pp. 224–236, February 1987.
- [9] M. K. Varanasi and B. Aazhang, "Near-optimum detection in synchronous code-division multiple-access systems," *IEEE Transactions on Communications*, vol. 39, pp. 725–736, May 1991.

Design of Sequential Circuits using Single Electron Encoded Logic

Mili Sarkar
milisarkar77@yahoo.co.in
Institute of Engineering and Management

Sunit Das
sunitdas09@gmail.com
Institute of Engineering and Management

Upasana Roy Chowdhury
roychowdhuryupasana@gmail.com
Institute of Engineering and Management

Dibyajyoti Das
dibyajyotisrkr@gmail.com
Institute of Engineering and Management

Abstract - A new realization of sequential circuits is presented in this paper. Memory elements like RS flip-flop, D flip-flop, JK flip-flop have been designed using Threshold Logic Gates. The Threshold Logic Gates are implemented using Single Electron Encoded Logic (SEEL) in which the Boolean values are encoded as zero or one electron charges. SEEL is based on Single Electron Tunneling(SET) technology. The proposed designs have been verified by means of simulation using VHDL.

Keywords – Threshold Logic Gates, Single Electron Tunneling(SET),SEEL,Flip-flops.

I. INTRODUCTION

Gradual decrease of circuit feature sizes in MOS based circuits will eventually come to an end due to physical restrictions [13]. Single Electron Tunneling (SET) [1] [8] is an emerging technology which offers greater scaling potential than MOS. This new SET technology is different from MOS technology as in SET based circuits the controlled transport of individual electrons through tunnel junctions is possible. This paper focuses on the implementation of logic gates and memory elements using SET technology.

Currently there are two design styles available for designing SET based logic circuits. The first style approach the tunnel junction as a switch using it to implement the SET equivalent of a MOS transistor[5][6]. In this process, charge is transported through an “open” SET transistor. By this approach the existing MOS based designs can easily be ported to SET technology. The main disadvantage of this approach is that the current transport through an “open” transistor still comprises of a large number of individual electrons “dripping” through the tunnel junctions. This results in increased delay and power consumption.

The second approach is to design SET logic gates that operate according to the Single Electron Encoded Logic (SEEL) paradigm, i.e., charge transport due to switching activity is limited to a single electron[3]. Given that this results in the transport of fewer electrons, it is expected that SEEL circuits have both reduced delay and reduced energy consumption.

Earlier investigations have revealed that (buffered) SEEL linear threshold logic gates can be constructed based on the Coulomb blockade of SET tunnel junctions[2][4] and that such gates operate correctly in larger networks. Threshold logic (TL) networks are fundamentally more powerful than networks of standard Boolean gates[10][12], e.g. TL gate based implementations of Boolean functions potentially require a smaller number of gates and less gate levels.

The remainder of this paper is organized as follows : Section II introduces the SET theory and briefly describes the generic SET Linear Threshold gate and the SET Buffer which are combined as a generic SEEL buffered Threshold gate. Section III investigates the SEEL threshold gate based implementation of R-S flip-flop, D flip-flop and J-K flip-flop. Then some final remarks are presented in Section IV.

II.BACKGROUND OF SET AND GENERIC THRESHOLD GATE

SET circuits are centered around tunnel junction through which single electrons can be transported in a controlled manner. A tunnel junction can be viewed as a leaky capacitor. *Tunneling* is referred to as the transport of charge through this tunnel junction while the transport of single electron through a tunnel junction is termed as *tunnel event*. Electrons tunnel through the tunnel junction strictly one after the other.

The voltage required to enable a tunnel event through a tunnel junction is called the voltage threshold. Let this critical voltage across the tunnel junction be represented as V_c . The capacitance of the tunnel junction is assumed as C_j . The remainder of the circuit as viewed from the tunnel junction’s perspective has an equivalence capacitance C_e . Following the approach presented in [7] the threshold voltage is calculated as

$$V_c = q_e / 2(C_e + C_j) \quad (1)$$

The charge of the electron is considered as $q_e = 1.602 \times 10^{-19}$ C in the calculation of V_c as well as in the rest of the discussions. Strictly speaking this is wrong since the charge of electron is negative. However it is more intuitive to consider the electron charge as a positive constant in the formulas that will determine whether a tunnel event will occur or not.

Defining the voltage across a junction as $V_{j,a}$ tunnel event will occur across this junction if and only if

$$|V_j| \geq V_c \quad (2)$$

If $|V_j| < V_c$ for all junctions in a circuit,tunnel events cannot occur in any of the circuit's tunnel junctions and the circuit is said to be in a stable state.In our paper,we focus on circuits where a limited number of tunnel events may occur resulting in a stable circuit.

Threshold Logic Gates are devices that are able to compute any linearly separable Boolean function given by

$$Y = \text{sgn}\{\mathbb{F}(X)\} = \begin{cases} 0, & \text{if } \mathbb{F}(X) < 0 \\ 1, & \text{if } \mathbb{F}(X) \geq 0 \end{cases} \quad (3)$$

$$\mathbb{F}(X) = \sum_{i=1}^n \omega_i x_i - \psi \quad (4)$$

where x_i are the n Boolean inputs and ω_i are the corresponding n integer weights.The LTG compares the weighted sum of inputs and the threshold value ψ .If the weighted sum of inputs are greater than or equal to the threshold,the gate produces a logic 1.Otherwise,the output is logic 0.

Figure 1(a) depicts a generic SEEL threshold gate already been proposed earlier in [2].The threshold gate can be implemented using both positive and negative weights.Due to the passive nature of the threshold gate buffers are required for the gates to operate correctly in networks[11].SET transistors can be utilized for buffers[9] as it requires active components.Figure 1(b) displays a non-inverting static buffer.It can also be modified to form an inverting static buffer.

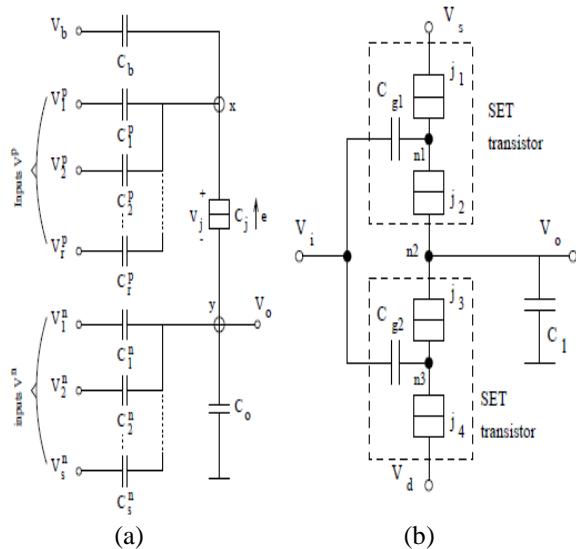


Fig. 1 (a)n-input LTG (b)Non-Inverting Buffer

The SET threshold gate combined with either inverting or non-inverting static buffer forms the basic building

block for any threshold gate network.In our investigation we have used this basic building block to design various Boolean logic functions that forms the fundamental blocks for any digital circuit.The basic Boolean logic functions AND,OR,NAND,NOR and NOT can be represented in the form of (3) and (4) as follows

$$\text{AND}(a,b) = \text{sgn}\{a+b-1.5\} \quad (5)$$

$$\text{OR}(a,b) = \text{sgn}\{a+b-0.5\} \quad (6)$$

$$\text{NAND}(a,b) = \text{sgn}\{-a-b+1.5\} \quad (7)$$

$$\text{NOR}(a,b) = \text{sgn}\{-a-b+0.5\} \quad (8)$$

$$\text{NOT}(a) = \text{sgn}\{-a+0.5\} \quad (9)$$

We can implement these basic logic functions with generic threshold gate circuits as shown in figure 1(a) augmented with a buffer given in figure 1(b).

III.TRESHOLD GATE-BASED MEMORY ELEMENTS

A. S-R LATCH IMPLEMENTATION :

The S-R latch has two inputs; namely SET (S) and RESET (R) ; and two outputs Q and \bar{Q} . The two outputs are compliment to each other.In the Boolean logic , the logic function of the S-R latch can be expressed as :

$$Q_{n+1} = \overline{RS} + \overline{R}Q_n \quad (10)$$

$$\overline{Q}_{n+1} = \overline{SR} + S\overline{Q}_n \quad (11)$$

TABLE I : Functional Truth Table for S-R Latch

| R | S | Q_{n+1} | \overline{Q}_{n+1} | Functions |
|---|---|-----------|----------------------|---|
| 0 | 0 | Q_n | \overline{Q}_n | Hold current output values Q_n and \overline{Q}_n |
| 0 | 1 | 1 | 0 | Set output Q to 1 and \overline{Q} to 0 |
| 1 | 0 | 0 | 1 | Reset output Q to 0 and \overline{Q} to 1 |
| 1 | 1 | ? | ? | Unspecified input combination |

In Boolean gate-based S-R latch implementation, it is observed that for the input combination $R=S=1$ the output becomes unstable.This problem can however be resolved in threshold gate-based implementation without the use of any additional gate.The S-R latch can be implemented at the cost of one threshold gate only as

$$Q_{n+1} = \overline{RS} + \overline{R}Q_n + SQ_n \quad (12)$$

$$\overline{Q}_{n+1} = \text{NOT}(Q_{n+1}) \quad (13)$$

If we verify we will find that when $S=R=1$, $Q_{n+1} = Q_n$ (Hold).A threshold gate-based implementation of this logic equation would require only one three input threshold gate implementing

$$Q_{n+1} = \text{sgn}\{S + \bar{R} + Q_n - 2\} \quad (14)$$

Now this equation can be written as

$$Q_{n+1} = \text{sgn}\{S - R + Q_n - 1\} \quad (15)$$

As stated earlier the threshold gates derived from generic threshold gate scheme require an output buffer for proper functioning,we therefore propose a S-R Latch implementation as displayed in fig. 2.

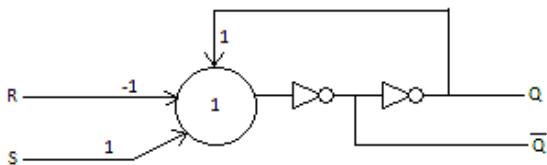


Fig. 2. Threshold gate-based S-R latch implementation

B. Negative edge triggered S-R Flip-flop

An edge-triggered S-R flip-flop is a memory element with three inputs (S,R,CLK) and two outputs (Q , \bar{Q}). The operation of the basic S-R latch can be modified by providing one additional control input(CLK) that determines, when the state of the is to be changed.The latch with the additional control input is called flip-flop. Flip-flop is capable of storing one bit.

In Boolean logic , the logic function of the Boolean gate-based S-R flip-flop can be expressed as :

$$Q_{n+1} = \overline{\text{CLK}}Q_n + \text{CLK}(\bar{R} + \bar{Q}_n + SQ_n) \quad (16)$$

Implementation of a negative edge triggered S-R flip-flop consists of two cascaded S-R latches. In the first latch the CLK signal is connected directly and to the second latch the CLK signal is fed through a NOT gate ($\overline{\text{CLK}}$).The outputs of the first latch serve as inputs of the second latch.The behaviour of a negative edge triggered S-R Flip-flop is given as follows:

Table II :Truth table for negative edge triggered S-R Flip-flop

| Clock(CLK) | S | R | Q_{n+1} |
|------------|---|---|--------------|
| 0 | d | d | Q_n (hold) |
| 1→0 | 0 | 0 | Q_n (hold) |
| 1→0 | 0 | 1 | 0 |
| 1→0 | 1 | 0 | 1 |
| 1→0 | 1 | 1 | Q_n (hold) |

As stated earlier in equation (16),

$$Q_{n+1} = \overline{\text{CLK}}Q_n + \text{CLK}(\bar{R} + \bar{Q}_n + SQ_n)$$

Now let us suppose

$$A = \bar{R} + \bar{Q}_n + SQ_n \quad (17)$$

Thus we can write equation (16) as

$$Q_{n+1} = \overline{\text{CLK}}Q_n + CLKA \quad (18)$$

$$\text{Let } B = \overline{\text{CLK}}Q_n \quad (19)$$

Therefore the final expression becomes

$$Q_{n+1} = B + CLKA \quad (20)$$

This can be implemented using threshold gate as

$$Q_{n+1} = \text{sgn}(2B + A + \text{CLK} - 1.5) \quad (21)$$

A can be designed by designing \overline{A} and then using a buffer or inverter to obtain A .This will also meet the requirement of an output buffer for the proper functioning of threshold gates derived from threshold generic scheme.Using the same concept the design of B can also be done.Thus we need to design the following

$$\overline{A} = \text{sgn}(-S + R - Q_n + 1) \quad (22)$$

$$\overline{B} = \text{sgn}(-Q_n + \text{CLK} + 0.5) \quad (23)$$

Thus we propose the threshold logic gate based negative edge triggered S-R flip-flop implementation as depicted in fig. 3.

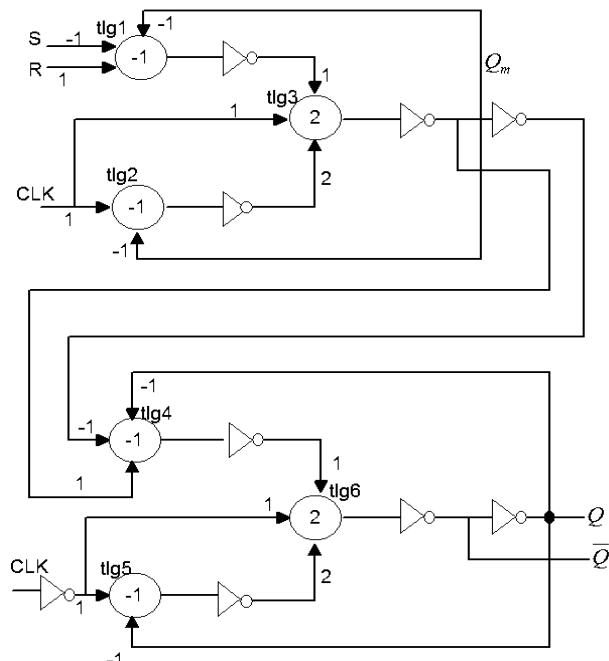


Fig.3. Negative edge triggered S-R flip-flop implementation

The proposed design is verified by VHDL simulator. Simulation results obtained are shown below.

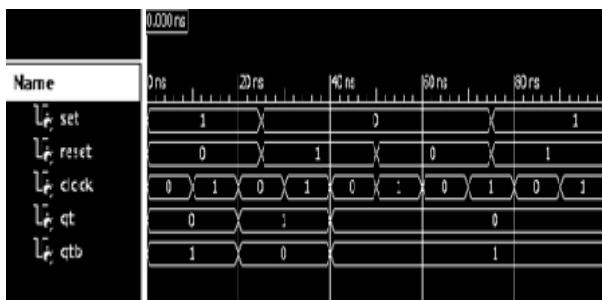


Fig.4 Simulation results of threshold gate based negative edge triggered S-R flip flop

C. D Latch

The D latch is a memory element with two inputs, D and L and two outputs Q and \bar{Q} . The behaviour of the D latch, is summarized in the following table

Table III : Truth table for D Latch

| L | D | Q | \bar{Q} |
|---|---|--------|----------------|
| 1 | 0 | 0 | 1 |
| 1 | 1 | 1 | 0 |
| 0 | d | last Q | Last \bar{Q} |

d=don't care

Thus it can be seen that, when L=0, the D latch holds the current output values, when L=1, the output Q follows the input D.

The logic function of the Boolean gate based D- Latch implementation is given as

$$Q_{n+1} = LD + \bar{L}Q_n \quad (24)$$

$$\bar{Q}_{n+1} = NOT(Q_{n+1}) \quad (25)$$

Verification of the above Latch equations can be shown as

If L=0, $Q^{n+1} = Q^n$ (hold)

If L=1, $Q^{n+1} = D$ (transparent).

However, in all cases, the output \bar{Q} is the complement of Q.

But this equation cannot be implemented by a single threshold gate. At least two threshold gates are required to implement such types of logic functions.

Splitting the equation into two parts we get,

$$A = LD \quad (26)$$

$$Q_{n+1} = A + \bar{L}Q_n \quad (27)$$

And hence the implementation of each of the equations as a single threshold gate is possible, as shown :-

$$A = \text{sgn}(L + D - 2) \quad (28)$$

$$Q_{n+1} = \text{sgn}(2A - L + Q_n - 2) \quad (29)$$

Equations (28) and (29) will be used to design D Flip Flop. Equation(28) results only in A=1, when both L and D are equal to 1 and thus represents a Boolean logic AND. Similarly, second equation gives $Q^{n+1}=1$ when

A=1 or when both \bar{L} and Q^n are 1. The threshold gates derived from the generic threshold gate scheme require an output buffer for correct operation in a network structure. The applied buffer inverts its input signal and hence, the threshold equation of A is modified such that it calculates $\bar{A} = \text{sgn}(-L - D - 1)$. Thus the combined result of the threshold gate and its buffer/inverter is a buffered gate that calculates A.

D.Negative edge triggered D-Flip-flop

An edge triggered D flip flop is a memory element with two inputs D and L, and two outputs Q and \bar{Q} .

The behaviour of the negative edge triggered D flipflop is as follows.

Table IV:Truth Table for negative edge triggered D flip-flop

| CLK | D | Q | \bar{Q} |
|-------|---|--------|----------------|
| 1 → 0 | 0 | 0 | 1 |
| 1 → 0 | 1 | 1 | 0 |
| 0 | d | last Q | Last \bar{Q} |
| 1 | d | last Q | Last \bar{Q} |

The transition of the clock from 1 to 0(a falling or negative edge on the time graph) results in the copy of the value of D to the output Q. But for all other input combinations, the circuit holds its current output values.

In the negative edge triggered D flipflop consists of two cascaded D Latches in which the output Q of the first D latch behaves as the input D of the second latch and the CLK is connected directly to the input L of the first latch, while the input L of the second clock is connected to NOT (CLK).

Adjustments of the Boolean function equations specified by equation (24) to $Q^{n+1} = \bar{L}D + LQ^n$ will require one

less inverter block .The resulting threshold gate-based equations are

$$A = \text{sgn}(L + D - 1) \quad (30)$$

$$Q_{n+1} = \text{sgn}(2A + L + Q_n - 1) \quad (31)$$

Again, as the applied buffer inverts its input signal, the threshold equation is modified to $\bar{A} = \text{sgn}\{L-D\}$. In this way, the combined result of the threshold gate and its buffer/inverter is a buffered gate that calculates A.The resulting threshold logic gate based negative edge-triggered D flipflop implementation is proposed , as shown in the fig. 5. below.

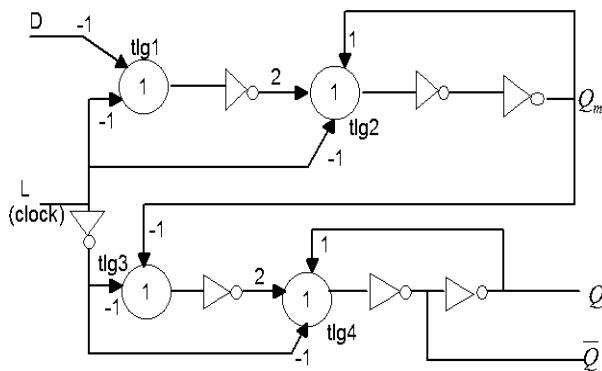


Fig. 5. Negative edge triggered D Flip-Flop implementation

We have verified the proposed D flipflop implementation by using the VHDL simulator.The simulation results thus obtained are shown below in fig.6.

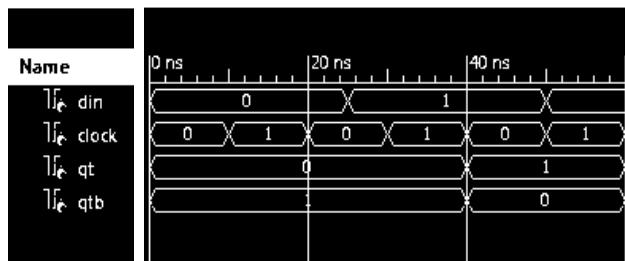


Fig. 6. Simulation results of threshold gate based negative edge triggered D flip flop

E. Negative edge triggered J-K Flip-flop

A J-K flip-flop has a characteristic similar to that of an S-R flip-flop.In addition the indeterminate condition of the S-R flip-flop is permitted in it.Inputs J and K behave like inputs S and R to set and reset the flip-flop, respectively.When $J=K=1$,the flip-flop output toggles, i.e. switches to its complement state;if $Q=0$,it switches to $Q=1$ and vice versa.

Thus the behaviour of the negative edge triggered J-K flip-flop can be summarized as follows.

Table V:Truth Table for negative edge triggered J-K flip-flop

| CLK | J | K | Q_{n+1} |
|-------------------|---|---|----------------------|
| 0 | d | d | Q_n (hold) |
| $1 \rightarrow 0$ | 0 | 0 | Q_n (hold) |
| $1 \rightarrow 0$ | 0 | 1 | 0 |
| $1 \rightarrow 0$ | 1 | 0 | 1 |
| $1 \rightarrow 0$ | 1 | 1 | \bar{Q}_n (toggle) |

A negative edge triggered J-K flip-flop can be obtained from the negative edge triggered S-R flip-flop by augmenting two AND gates as shown below in fig.7.The AND gate is implemented using threshold gate following equation (5).Instead of giving the threshold value as 1.5, the threshold value used here is 2 which results in the same output.The data input J and the output \bar{Q} are applied to the first AND gate and its output ($J \bar{Q}$) is applied to the S input of S-R flip-flop.Similarly the data input K and the output Q are connected to the second AND gate and its ouput ($K Q$) is applied to the R input of S-R flip-flop.

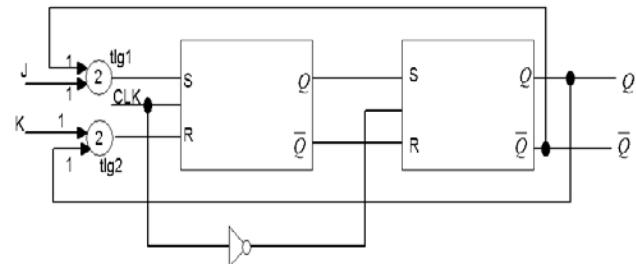


Fig. 7. J-K flip-flop using S-R flip-flop

Here the two clocked S-R latch used is implemented using threshold gate based design as shown in fig.3.

The condition $J = K = 1$ can be easily verified.Under these conditions and considering the previous state is a SET state, then $S = J \bar{Q}_n = 1 \times 0 = 0$ and $R = K Q_n = 1 \times 1 = 1$.Since $S = 0$ and $R = 1$,the flip-flop RESETS on the negative edge of the clock pulse,i.e. the flip-flop toggles from SET to RESET state.Similarly if the previous state is a RESET state then S becomes 1 and R becomes 0 and the flip-flop on the application of the negative edge of a clock pulse will toggle to the SET state.

The above proposed design is verified by using the VHDL simulator. The simulation result is shown below.

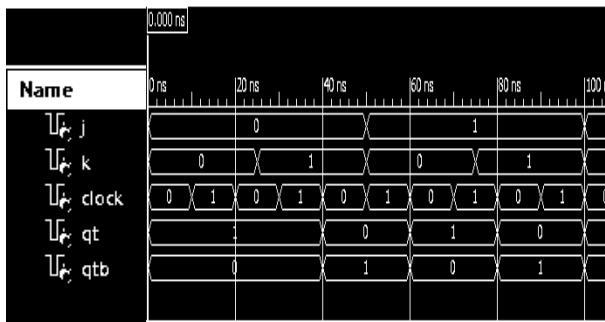


Fig. 8. Simulation results of threshold gate based negative edge triggered J-K flip flop

IV. CONCLUSIONS

If we compare the Boolean and threshold gate-based implementations of the above memory elements we can see that the threshold gate based designs has at most the same number of logic gates and the same network depth. In fact we have been successful in reducing the circuit elements using threshold gates. The successful design of latches and flip-flops will lead to the designing of binary counters and registers using SEEL based threshold gates. Thus various sequential circuits can be designed using lesser number of circuit elements. This technology can be very useful for designing of RAM cells. This technology is a promising alternative to Boolean gate-based implementations. Thus in future Threshold Logic Technology has a great prospect of being a forerunner in the development and inventions in Electronics engineering.

REFERENCES

- [1] A. Korotkov, "Single-electron logic and memory devices," *Int. J. Electron.*, vol. 86, no. 5, pp. 511–547, 1999.
- [2] C. Lageweg, S. Cotfan'a, and S. Vassiliadis, "A linear threshold gate implementation in single electron technology," in *IEEE Computer Society VLSI Workshop*, Apr. 2001, pp. 93–98.
- [3] Y. N. Nazarov and S. V. Vyshenskii, "SET circuits for digital applications," in *Single-Electron Tunneling and Mesoscopic Devices*. ser. Electron. Photon., H. Koch and H. Lubbig, Eds. Berlin, Germany: Springer-Verlag, 1992, vol. 31, pp. 61–66.
- [4] C. Lageweg, S. Cotfan'a, and S. Vassiliadis, "Static buffered SET based logic gates," in *2nd IEEE Nanotechnology Conf.*, Aug. 2002, pp. 491–494.

- [5] A. Korotkov, R. Chen, and K. Likharev, "Possible performance of capacitively coupled single-electron transistors in digital circuits," *J. Appl. Phys.*, vol. 78, pp. 2520–2530, Aug. 1995.
- [6] J. R. Tucker, "Complementary digital logic based on the "Coulomb blockade"," *J. Appl. Phys.*, vol. 72, no. 9, pp. 4399–4413, Nov. 1992.
- [7] C. Wasshuber, "About single-electron devices and circuits," Ph.D. dissertation, Elect. Eng. Dept., Tech. Univ. Vienna, Vienna, Austria, 1998.
- [8] K. Likharev, "Single-electron devices and their applications," *Proc.IEEE*, vol. 87, pp. 606–632, Apr. 1999.
- [9] R.H.Chen, A.N.Korotkov, and K.K.Likharev. Singleelectron Transistor Logic. *Applied Physics Letters*, Vol. 68(No. 14):pp. 1954–1956, April 1996.
- [10] S. Muroga, *Threshold Logic and Its Applications*. New York: Wiley, 1971.
- [11] C.Lageweg and S.Cotofana and S.Vassiliadis. Achieving Fanout Capabilities in Single Electron Encoded Logic Networks. In *6th International Conference on Solid-State and IC Technology (ICSICT)*, pages 1383–1386, October 2001.
- [12] Biplob Roy, and Biswarup Mukherjee, "Design of Coffee Vending Machine using Single Electron Devices," 2010 International Symposium on Electronic System Design.
- [13] Y.Taur, D.A.Buchanan, W.Chen, D.Frank, K.Ismail, S.Lo,G.Sai-Halasz, R.Viswanathan, H.Wann, S.Wind, and H.Wong. CMOS Scaling into the Nanometer Regime. *Proceeding of the IEEE*, Vol. 85(No. 4):pp. 486–504, 1997.

Implementing Security Measurement on a Lifetime Enhancing Routing Protocol of WSN

AVIK MODAK

Dept. of ECE
MCKV Institute of Engineering
Lilua , Howrah
avikmodak@gmail.com

ANURAN ROY CHOWDHURY

Dept. of ECE
MCKV Institute of Engineering
Lilua , Howrah
anuranforu@gmail.com

Abstract: We are aware that the power limitation is a critical issue for the sensor nodes, so different scheme or algorithm was proposed to implement the routing protocols to expand the minimum power utilize the maximum work [1,2].But it is unfortunate that most of them are not security concerned. As we know that the wsn is used for military purpose mainly so fulfilling privacy and security requirements in an appropriate architecture for WSNs offering pervasive services is essential for acceptance. Though prolonged lifetime may help the sensor network to work for several year without change the battery of the nodes but still it is the security that is most needed. If an attacker may hack the information of the network then prolonged lifetime have no importance to the network as longer the lifetime of the network more information will be hacked. In this paper we will discuss some internal and external security aspects that the routing protocols have faced and different solutions to make the protocols secure as well as prolonged life time concerns.

Key Words:

Secure Routing, Lifetime, Attacks, Privacy, Cryptography.

1. INTRODUCTION

We are concentrating to provide adequate security to the sensor networks, which are required both commercially and technically. The sensor nodes of the sensor networks are small, cheap, have limited battery capability and transmission range to perform limited monitoring and sensing functions. Sensor nodes are distributed over a potentially vast geographical area to form a static, multi-hop, self-organizing network. However, also mobile WSNs and mobility within WSN are conceivable. A major benefit of WSN is that they perform in-network processing to reduce large streams of raw data into useful aggregated information.

The Key constraints are providing importance when we design an wireless sensor network is security, lifetime and the transmission delay of the network. Proper design of each of these features depends the system architecture design of a WSN.

Security is typically an important issue in sensor networks including sensor networks , where the communication medium is broadcast in nature and, hence, an adversary can overhear all messages sent by any user. For this reason, a sender must authenticate the receiver and

encrypt any messages it sends. In WSN the channel is assumed to be insecure and the end-points cannot in general be trusted. An attacker may physically pick up sensor nodes and extract sensitive information .Making the sensor nodes “tamper-resistant may be a good solution as it impacts security but it makes the node expensive. Also the limited computing and storage capabilities make modular arithmetic with large numbers difficult and thus asymmetric (public key) cryptography unsuitable. Security is sometimes viewed as a standalone component of a system’s architecture, where a separate module provides security.

To achieve a secure system, security must be attached with every design parameters, since parameters designed without security can become a point of attack. Consequently, security must pervade every aspect of system design. While we designing a scheme for a sensor network, security security should be incorporated to the mechanisms after design has completed.

An ultimate limitation of building a multi-hop routing topology around a fixed set of base stations is that those nodes within one or two hops of the base stations are particularly attractive for compromise. After a significant number of these nodes have been compromised, all is lost. This indicates that clustering protocols where cluster-heads communicate directly with a base station may ultimately yield the most secure solutions against node compromise and insider attacks.

In this paper we will discuss countermeasures and design considerations for secure routing protocols in sensor networks. For that we first discuss a proposed routing scheme which is modification of the PEGASIS [1],to prolonged the lifetime of the network and then discuss the threats that the scheme may face and trying to resolve a solution to make it secure.

The paper is oriented as follows: In section II we will discuss about literature preview,In section III we will concentrate on our proposed scheme and its experimental results and simulation graphs ,In next section we will focus on the attacks against the scheme ,In section V we are trying to resolve the attacks to make it secure and at last we will draw some conclusion and the future research works that will come as a challenge to the scientists.

2. LITERATURE REVIEW

In this section we will discuss some previous work on the routing protocols and some security management that is already taken for wsn.

Massive experiments and research works were done regarding the lifetime time enhancement routing protocols. Hierarchical or cluster-based routing, originally proposed in wire line networks, are well-known techniques with special advantages related to scalability and efficient communication. As such, the concept of hierarchical routing is also utilized to perform energy-efficient routing in WSNs. Heinzelman *et al.* in [2] developed a cluster-based routing scheme called Low-Energy Adaptive Clustering Hierarchy (LEACH), where in each cluster, member nodes adopt a Time Division Multiple Access (TDMA) protocol to transmit their data packets to the cluster head. After receiving data packets from all its neighboring sensor nodes, a cluster head performs data aggregation and sends the final aggregated packet to the Base Station under the Carrier Sense Multiple Access (CSMA) protocol. LEACH utilizes a more accurate energy model and offers much better performance in terms of energy efficiency and network lifetime. The Power Efficient Gathering in Sensor Information Systems (PEGASIS) scheme proposed in [1] is based on a greedy chain, which starts from the farthest node from the Base Station. By connecting the last node on the chain to its closest unvisited neighbor, PEGASIS greatly reduces the total communication distance and achieves much better energy and lifetime performance than LEACH for different network sizes and topologies. The PEGASIS scheme depends upon a greedy chain formation whereas the LEACH scheme randomizes the leader selection in the network. While the greedy chain cannot always guarantee minimal energy consumption, the randomized leader selection does not take into account the node's capability in being the leader, in terms of its energy content and transmit distance.

Security in sensor networks has been well enumerated in the literature [4]. Sensor network security has been studied in recent years in a number of proposals. Kulkarni et al. [5] analyzes on the problem of assigning initial secrets to users in ad-hoc sensor networks to ensure authentication and privacy during their communication and points out possible ways of sharing the secrets.

Secure routing protocols for ad-hoc networks based on symmetric key cryptography have been proposed [7] and for public cryptography is proposed in [4]. But the protocols are too expensive in terms of node state and packet overhead and are designed to find and establish routes between *any* pair of nodes. In [6] Karlof et al. thoroughly discussed the problem of secure data transmission for different routing protocols and they conclude that Many sensor network routing protocols have been proposed, but none of them have been designed with security as a goal. They suggested the security goals required for routing in sensor networks.

3. PROPOSED SCHEME

In this section we modify the PEGASIS algorithm to increase the lifetime of the network here we will first

concentrate about the implementation of the scheme followed by the simulation results.

A. Problem Formulation and Implementation

In this paper we have implemented a allocation scheme for the leader selection of the WSN. PEGASIS is power efficient data gathering protocol for wireless sensor network where routing occurs as a greedy chain formation technique. In this work we have compared simultaneously the greedy chain formation with the data gathering using leader allocation strategy. We have assumed that each sensor node in the network bears a initial energy. The sensor nodes are deployed randomly in the network. For a particular node we have allocated initially with certain rounds after selection of that node as a leader. We have simulated the network life time and mean energy of the network using chain formation with leader assigned for fixed rounds. Here we keep track of the leaders i.e. which nodes are selected as leaders and no of rounds they are assigned as leaders. From this statistical analysis we keep track of the min. no. of rounds that a particular node may be posted as leader. Suppose the minimum no. of round is N, in the methodology for leader allocation it set as a restriction for every leader.

We observe that a single node communicates with the base station hence the possibility of collision between the signals may be avoided by leader allocation strategy. Disadvantage of the LEACH protocol [2] where several nodes communicate with the base station either in TDMA or CDMA which we assigned it as leader [8]. So in greedy chain formation we count the possibility of the number of nodes as leader using MATLAB. We have restricted the no of times a nodes being allowed to be leader. We allocate the least number of counts for a node acting as a leaders shown in table 1

For measurement of dissipation of energy different radio model are discussed[3] According to first order radio model the energy dissipated in transmitting a k-bit message over a distance d is given by :

and the amount of energy lost due to receiving the k -bit packet is:

where e_t is the energy dissipated per bit in the transmitter circuitry and $e_d * d^n$ is the energy dissipated for transmission of a single bit over a distance d , n being the path loss exponent .

Table 1 Count of the Nodes as Cluster head

| Node ID | Count of the Nodes as a Cluster Head. | Mean of the nodes to be a Cluster Head |
|---------|---------------------------------------|--|
| 1 | 359 | 145 |
| 2 | 284 | |
| 3 | 119 | |
| 4 | 238 | |
| 5 | 91 | |
| 6 | 97 | |
| 7 | 151 | |
| 8 | 113 | |
| 9 | 104 | |
| 10 | 86 | |
| 11 | 221 | |
| 12 | 126 | |
| 13 | 87 | |
| 14 | 220 | |
| 15 | 95 | |
| 16 | 69 | |
| 17 | 85 | |
| 18 | 248 | |
| 19 | 64 | |
| 20 | 48 | |

B. Simulation Results

We have simulated the clustering scheme of data gathering using MATLAB .Consider the BS is located at (100, 100) in a 100m x 1000m field. We have simulated in C to

determine the number of rounds of communication when 10%, 20%, 50% and 100% of the nodes die using direct transmission, LEACH, PEGASIS and Proposed algorithm with each node having the same initial energy level (0.1mJ). Once a node dies it is considered dead for the rest of the simulation. Our simulations shows that proposed algorithm achieves approximately three times the number of rounds compared to PEGASIS and five times the number of rounds compared to LEACH when 10%, 20%, 50%, and 100% nodes die for a 100m x 100m network.

The Experiment table is shown in the table II as follows

Table 2 : Comparative of the Network Lifetime with Node density and initial energy of the Network.
FND: First Node Dies, LND: Last Node Dies

| Node density of the Network | Initial Energy (mJ) | Network Lifetime (No of rounds) | | | |
|--|---------------------|----------------------------------|-----|-----------------|------|
| | | PEGASIS | | PROPOSED SCHEME | |
| | | FND | LND | FND | LND |
| 100 sensor Deployed in a 50* 50 Square field | 100 | 76 | 167 | 78 | 206 |
| | 250 | 200 | 416 | 206 | 465 |
| | 500 | 313 | 780 | 313 | 1029 |
| 200 sensor Deployed in a 50* 50 Square field | 100 | 69 | 160 | 69 | 213 |
| | 250 | 153 | 440 | 153 | 380 |
| | 500 | 357 | 771 | 357 | 1028 |

The Network Lifetime enhancement is shown in the comparative figure shown below

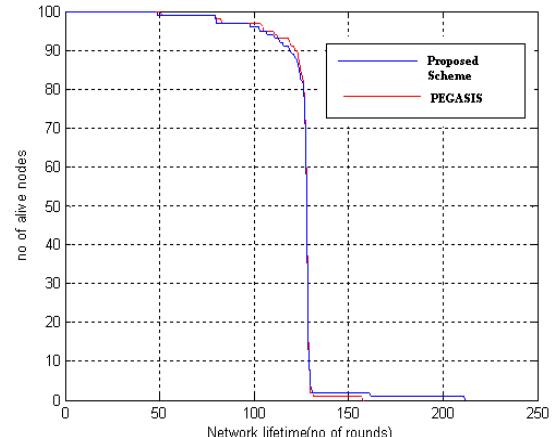


Fig 1: Comparative plot with 100 no. of nodes over a square area of 50*50 with initial energy 100J

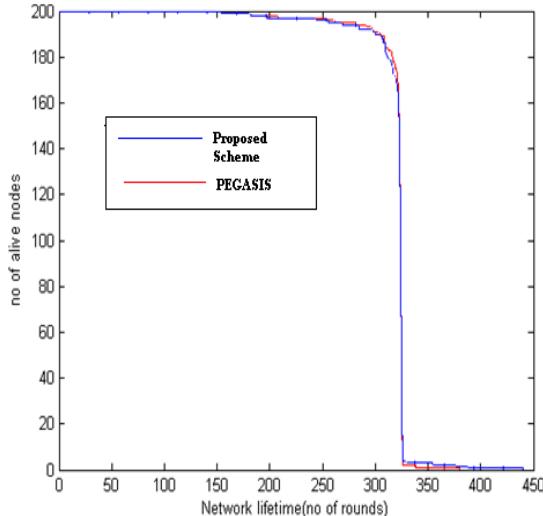


Fig2: Comparative plot of 200 no. of nodes over a square area of 50*50 with initial energy 250 J.

Comparative plot of the two schemes are shown with increased number of nodes and initial battery energy which shows significant improvement of network lifetime .as tabulated in table 3 and compare in chart in figure 3.

Table 3 : Comparison of Network Lifetime in (100m x 100m) square field with BS located at (100m, 100m)

| PROTOCOL | FND : FIRST NODE DIES | LND : LAST NODE DIES |
|-------------------|-----------------------|----------------------|
| DIRECT | 48 | 102 |
| LEACH | 302 | 663 |
| PEGASIS | 388 | 1002 |
| PROPOSED ALGOITHM | 441 | 1115 |

Our performance produces a significant enhancement of Network Lifetime over existing scheme. But it really not save from various attacks so security impact of the scheme is not efficient .

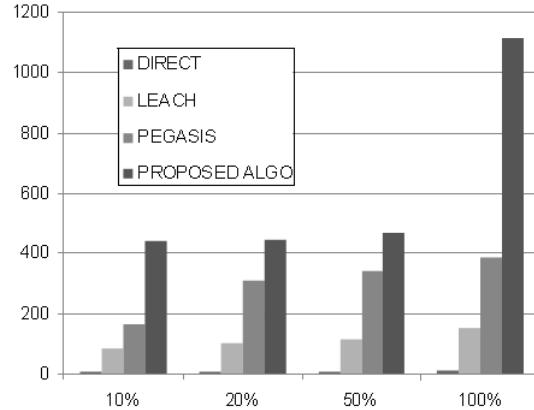


Fig 3.: Percentage of node death and performance results for a 100m x 100m network with initial energy of 0.1J/node

4. ATTACKS ON PROPOSED SENSOR NETWORK SCHEME

Before focusing on different attacks on the proposed algorithm we will first discuss some assumption underlying about the network scheme we have proposed in the previous section.

A. configuration of sensor nodes:

As in wsn all the data are conveying wirelessly so attackers may attack to the radio link between the nodes to inject some hacking data to penetrate the information of the network. And also we also assume that the nodes are not tamper resistance as it tends to add significant per-unit cost, and sensor nodes are intended to be very inexpensive.

B. Trust Requirements:

The base station of the network is considered as loyal i.e. it must behave correctly and can be trusted as it interface a sensor network to the outer world.

C. Threat Models:

Several types of attacker are present viz. outsider & insider attackers, mote-class & laptop-class attackers. The outsider attackers are unauthorized person though the insider are authorized person and they may be mounted from either compromised sensor nodes running malicious code or adversaries who have stolen the key material, code, and data from legitimate nodes. The laptop attackers are very massive attackers as they can jam the whole radio links of the sensor network , have a high bandwidth, low-latency communications channel not available to ordinary sensor nodes, allowing such attackers to coordinate their efforts.

D. Security Goals:

Security provides authenticity, integrity, privacy to a sensor network to covey messages properly in the network and also in the real world. Here we consider eavesdropping which is produced by cloning or rerouting of a data flow as the primary security goal and routing protocols should prevent it.

The effectiveness of a routing protocol in achieving the above goals should degrade no faster than a rate approximately proportional to the ratio of compromised nodes to total nodes in the network.

Now we are elaborate on some shortlisted attacks that our proposed protocol is faced off. The attacks that we will discuss in this section are

- Selective forwarding
- HELLO flood attacks

In this description below we will show how the attacks try to manipulate user data directly and affect our proposed scheme.

A. Selective forwarding

When we implement the scheme we always considered that every nodes in the network passed the message to its neighbor faithfully. If a node refuse to forward the message to its neighbor then the balance of the network is pretended. This type of attacks known as Selective forwarding attacks. The worst case occurs when a malicious node drops every message that it receive. In such case the neighbor nodes may report that it has failed and decide to seek another route. A more precise form of this attack is when an adversary selectively forwards packets. If the attacker implement this attacks in path of data flow then it becomes precise and very effective. Though, it is considerable that adversary ignoring a flow passing through neighboring nodes might be able to emulate selective forwarding by jamming or causing a collision on each forwarded packet of interest but such an effort are tricky at best, and may border on impossible. Thus, we believe an adversary launching a selective forwarding attack will likely follow the path of least resistance and attempt to include herself on the actual path of the data flow. In the next section we will focus on the HELLO flood attacks.

B. HELLO flood attack

HELLO flood attack is introduced in [6]. Here we assume that a node may broadcast HELLO packets to assign themselves to their neighbors, and a node receiving such a packet may assume that it is within (normal) radio range of the sender. This assumption may be false: a laptop-class attacker broadcasting routing or other information with large enough transmission power could convince every node in the network that the adversary is its neighbor. Under this circumstances the network is left in a state of confusion. So conveying and propagating the packet between neighboring nodes for topology maintenance or flow control are also subject to this attack. It may be thought of as one-way, broadcast wormholes[9].

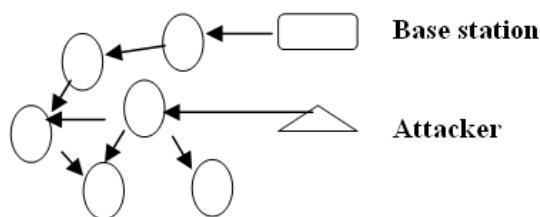


Fig4: an attacker broadcasting hello packets with more transmission power than a base station

Sensor networks pose some unique challenges regarding the security of the networks as the traditional security

management that is implement on traditional networks cannot be applied to the sensor networks because of unlike traditional networks, sensor nodes are often deployed in accessible areas, presenting the added risk of physical attack. existing security mechanisms are inadequate, and new ideas are needed.

5. COUNTER MEASURE AGAINST SECURITY THREATS

In order to meet the application level security requirements, the individual nodes must be capable of performing ,*authentication, privacy, Key establishment and trust setup thus overall a secure routing scheme*. Here we will discuss some counter measure against the security challenges that our proposed scheme faced.

A. Secure routing:

Data conveying in proper route is an essential requirement for enabling the communication in the sensor networks. In the previous section we discuss that our scheme faces a major challenges against the Selective forwarding, HELLO flood attacks. Here we will try to resolve this attack to make our protocol secure.

a. Selective forwarding

As a solution of the selective forwarding attacks multiple path routing is a good option . Messages routed over multipath whose nodes may have common but link are completely distinguished are completely protected against selective forwarding attacks involving at multiple nodes and still offer some probabilistic protection when over these nodes are compromised .This multipath that may have common nodes but all links are uncommon are known as Braided paths[10]. probabilistic protection may be taken against the selective routing attacks using Braided paths. Here nodes are allowed to choose a packet's next hop probabilistically from a set of possible candidates randomly can reduce the control of attackers against the data flow.

b. HELLO flood attacks

We assign a particular key to encrypt each request message that a node receives To defend against attack . In this way, any node's reachable neighbors can decrypt and verify the REQ message while the attacker will not know the key and will be prevented from launching the attack. But This defense gets less effectiveness when an attacker has a highly sensitive receiver as well as a powerful transmitter. Thus a different way of reliable exchange of messages among nodes and base stations is required so that when any particular node has different route to send data, this problem will be cured. If we assume that, there are a number of base stations in the network who have control over specific number of nodes and also, there are common means of communications among base stations and follows that steps below.

step 1:each nodes uses its new key to exchange messages among them.

step2:Transmission of request massage from base stations to the nearest nodes follows this format :



HCN is the base station's one-way hash chain number. Receiving node verifies that the REQ comes from the base

station, then it forwards the REQ to its neighbor node in same format using new key.

step3: *Step 3:* When any ordinary node receives this REQ message, it checks the sender ID to verify neighbor, *then it* decrypts and authenticates the sender with computed new key. If the message sender is valid, it replaces the HCN with the new value and encrypts the REQ message with its new key and broadcasts the newly encrypted message. The whole process is described it figure 5 considering four base stations with their communication range and sensor nodes with their communication range.

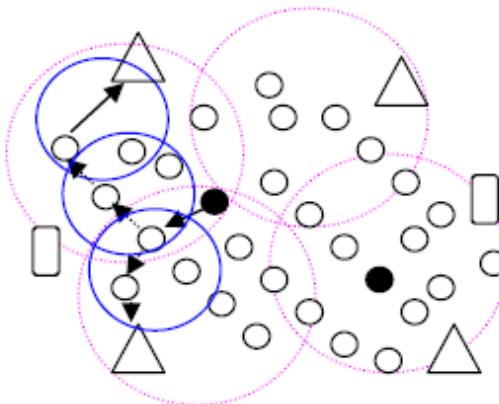


Fig 5: Ordinary node gets REQ message from compromised node but does not forward message to it, rather it sends message to its verified neighbor by alternative routes

B. Authentication and Privacy:

Here we focus our attention on the problem of assigning initial secrets to users in sensor network so that they can use those secrets to ensure authentication and privacy during their communication. As in sensor network all the information are exchanging wirelessly so the security of the information may be easily penetrated so a sender must authenticate the receiver. The sender and receiver may be processed to a secure communication by using a common shared key. But it is not an efficient solution. Clearly, if we require that the secret shared between two users is not known to any other user in the network, then each user must maintain $n-1$ secrets where n is the number of users. To reduce the number of secrets, the user are allowed to share a collection of secrets, and require that no other user in the network knows all the secrets in that collection. Clearly, in this situation, it would be possible for the users to use a combination of these secrets to ensure privacy and authentication.

C. Key establishment and trust setup:

Key establishment and trust setup is one of the primary requirement to establish a sensor network setting up. Network shared key is the simplest but the inefficient solution regarding this challenge. Our approach is to preconfigured the network with a shared unique symmetric key between each pair of nodes. Here we consider n nodes, each node needs to store $n-1$ keys, and $n * (n - 1)/2$ keys need to be established in the network.

Bootstrapping keys is used for trust setup station is where each node needs to share only a single key with the base station and set up keys with other nodes through the

base station. In this protocol the network may incorporate tamper-resistant packaging for the base station, ameliorating the threat of physical attack.

6. CONCLUSIONS

So from the compact discussion it is clear that it the security that must be considered as the primary constraint for designing a genuine wsn as Security plays a crucial role in the proper functioning of wireless sensor networks. Though Many other problems also need further research. One is how to secure wireless communication links against eavesdropping, tampering, traffic analysis, and denial of service and also to managing propagation delay of the network but we are optimistic that much progress will be made on all of them.

7. REFERENCES

- [1] Lindsey, S., Raghavendra, C.S. : PEGASIS: Power Efficient Gathering in Sensor Information Systems, In Proceedings of IEEE ICC 2001 (2001) 1125-1130
- [2] Wendi Heinzelman, Anantha Chandrakasan and Hari Balakrishna, “Energy-Efficient Communication Protocol for Wireless Microsensor Networks”, In Proceedings of 33rd Hawaii International Conference on System Sciences, Jan. 2000, pp. 1-10
- [3] Ayan Acharya, Anand Seetharam, Abhishek Bhattacharyya, Mrinal KantiNaskar, “Balancing Energy Dissipation in Data Gathering Wireless SensorNetworks Using Ant Colony optimization”,10th International Conference Distributed Computing and Networking-ICDCN 2009, pp437-443, Jan 3-6,2009.
- [4] L. Zhou and Z. Haas, “Securing ad hoc networks,” IEEE Network Magazine, vol. 13, no. 6, November/December 1999.
- [5] S. S. Kulkarni, M. G. Gouda, and A. Arora, “Secret instantiation in adhoc networks,” Special Issue of Elsevier Journal of Computer Communications on Dependable Wireless Sensor Networks, pp. 1–15, May2005.
- [6] C. Karlof and D.Wagner, “Secure routing in wireless sensor networks: Attacks and countermeasures,” Elsevier’s AdHoc Networks Journal, Special Issue on Sensor Network Applications and Protocols, 1(2-3):293–315, September 2003
- [7] Y.-C. Hu, A. Perrig, and D. B. Johnson, “Ariadne: A secure on-demand routing protocol for ad hoc networks,” Department of Computer Science, Rice University, Tech. Rep. TR01-383, December 2001.
- [8] Ayon Chakraborty, Swarup Kumar Mitra, and M.K. Naskar, “AnEfficient Hybrid Data Gathering Scheme in Wireless Sensor Networks”, ICDCIT 2010, LNCS 5966, pp. 98–103, 2010.
- [9] Y.-C. Hu, A. Perrig, and D. B. Johnson, “Wormhole detection in wireless Tech. Rep. TR01-384, June 2002.
- [10] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin, “Highly-resilient, energy-efficient multipath routing in wireless sensor networks,” Mobile Computing and Communications Review, vol. 4, no. 5, October 2001

A Novel Multiplexer & Demultiplexer Design and Permissible Fault Analysis Using Quantum Cellular Automata

Ratna Chakrabarty¹, K.K.Ghosh¹

Department of Electronics Engineering
Institute of Engineering & Management
Salt Lake, Sector-V, Kolkata, India
e-mail-ratna_chakrabarty@yahoo.co.in

Debashis De²

Department of Computer Science & Engineering
West Bengal University of Technology
Salt Lake, Kolkata, India.
School of Physics
University of Western Australia, Perth, Australia
Dr.debashis.de@gmail.com

Abstract—Quantum Dot Cellular Automata (QCA) represents an emerging area in the field of nano-technology with its promising extra low power consumption, high speed performance and for their extremely small feature size. Construction of logical circuits with QCA is cost effective and can be useful component for the design of many logical circuits. QCA is based on single electron effect in quantum dots and molecules. A number of conventional digital circuits hitherto developed by CMOS can equally be realized by proper configuring the basic building blocks of quantum cells. This paper presents QCA based 2:1 multiplexer and demultiplexer design and their various cell displacement fault analysis which can occur during fabrication. The proposed multiplexer and demultiplexer has been compared to few designs and a comparative study has been done by simulation

Keywords: *QCA; Majority Voter; Electron beam lithography; Multiplexer; Demultiplexer.*

I. INTRODUCTION

Owing to the physical limitations in down-sizing CMOS and its associated chips there has been growing interest in diversified field of nanotechnologies challenging the short comings of traditional CMOS circuits and devices with a replacement of a new technology. Though there has been significant improvement in the performance of CMOS VLSI devices in terms of size, power consumption and speed but the major obstacles are physical limitations of lithography and ever increasing cost of foundry which prevent the long term continuation process. This has motivated the exploration of a new technology capable of offering significant improvement in terms of speed, power consumption and integration level over VLSI based technology. The QCA technology has emerged by itself as a very important innovative technology in nano-scale. QCA represents wireless computing paradigm. Using QCA we introduce here the novel design of multiplexer and demultiplexer circuits. Multiplexer and demultiplexer has numerous applications in digital systems such as data selection,

data routing, parallel to serial conversion, waveform generation and many other logic function generation. It can be used as a part of control sequencer circuit. QCA is a nano level computing mechanism by which we can easily design a basic 2:1 multiplexer and demultiplexer. CMOS also provides microscale computing with low power, very large scale integration technique; but it has many problems due to its high leakage current, high lithography cost and limitation of speed which is in GHz range, where QCA works in THz frequency range. QCA is a better alternative of CMOS technology which had been proposed in [1, 2, 3, 4]. The device part of QCA is a simple quantum cell consisting of four (or sometime five) quantum dots confined in one cell of size 10-9 cm. Each cell contains two extra electrons at the two diagonally opposite corners. In QCA two stable states are considered, one and zero state corresponding to the position of the two electrons polarity. The individual cells and the two neighboring cells work on the principle of Coulomb's electrostatic force law and as such to maintain a stable configuration of the cell structure, the cellular electrons occupy diagonally opposite dots. This definite polarization status of the dot electrons contains definite digital information. Transmission of digital information by interconnecting a number of quantum cells instead of using interconnecting wires makes the QCA circuit a novel one [5] In QCA no interconnecting wires are needed in between the cells and no current flows to carry the digital information; instead the QCA transfers the information by propagating its polarization state using a clock. The clock not only synchronizes and controls the flow of information but also it supplies the necessary power to the cells which run under no external power source

The main contribution in this work are reconstruction of 2:1 multiplexer and demultiplexer, its area calculation and compare the result with the design proposed by Sara Hashemi et al [6] and M. Askari et al [7]. The area calculation of the multiplexer is half the previous one [6]. A two bit demultiplexer circuit also has been proposed in this paper.

Here in this paper some faults are introduced which may occur during fabrication and the output is shown by simulation.

The paper is organized as follows, section II introduces the brief concept of working principle of QCA Technology, section III, illustrates 2-bit multiplexer and demultiplexer design, comparison the circuit with the other previously designed circuits described in [6,7,8,9]. Section IV describes the fabrication defect and permissible defect tolerance of the proposed multiplexer and demultiplexer circuits.

Section V ends with the conclusion. All simulations carried out using QCAD designer tool [10].

II. BASIC PRINCIPLES UNDERPINNING QCA

A. Physics of quantum cell

QCA Cell was first introduced by Prof C. S. Lent et al [11] at the University of Notre Dame where processing of information is based on Columbic interactions between the electrons in the cell. Generally a cell is created with 4 quantum dots positioned at the vertices of a square. Two mobile electrons are there which tend to occupy the diagonals due to Coulomb electrostatic repulsion. The electrons are trapped inside the dot of low potential region surrounded by a region of high potential. A very common way to implement such dots is to use metal and especially the nano-scale dots are constructed by aluminum using the technology of electron beam lithography. Due to high potential barrier, tunneling is not possible outside a cell; however electrons are able to tunnel between the dots. Two possible polarizations (+1 or -1) according to the electron positions in a cell are showing in Fig 1(a). A QCA binary wire is shown in Fig1 (b) where cells are able to transmit information in coded form from one cell to another without any current flow. Polarization of each cell solely depends upon the polarization of its previous neighboring cell. This phenomenon is useful in nanotechnology which offers high resolution fast electronic circuits.

B. Elementary building blocks in QCA

The fundamental QCA logic circuits are the majority gate and the inverter. The majority gate has three inputs and one output. The center cell is the device cell. If A, B, C are the inputs, the majority gate computes the function $M=AB+BC+CA$. By fixing the input polarization of one input +1 or -1 i.e., logic 1 or logic 0, OR and AND logical operations can be performed respectively. Fig 2(a) shows the majority OR gate, Fig 2(b) shows the majority AND gate and Fig 2(c) shows the Inverter gate realized with quantum cells. In the inverter due to presence of 45 degree cell the signal alternates between the input value and its logic complement as it propagates the chain towards the output. Inverter consumes a substantial area in QCA. In the proposed design of 2 bit multiplexer a very small area inverter is used, thus number of cells are reduced in the design. A QCA circuit consists of an array of QCA cells arranged in a Cartesian plane. The input cells of the QCA array are in fixed polarization, and the entire array is then allowed to relax to its ground state. The output is read by sensing the state

of the output cells and cells are allowed to switch to whatever polarization that achieves from system at ground state. When the array settles to its ground state, QCA computes then correctly. A kink occurs when the system is in metastable (no true energy ground state) state. The kink energy requires exciting the system from the ground state to the first excited state



Figure 1. (a) Polarized cell (b) QCA binary wire

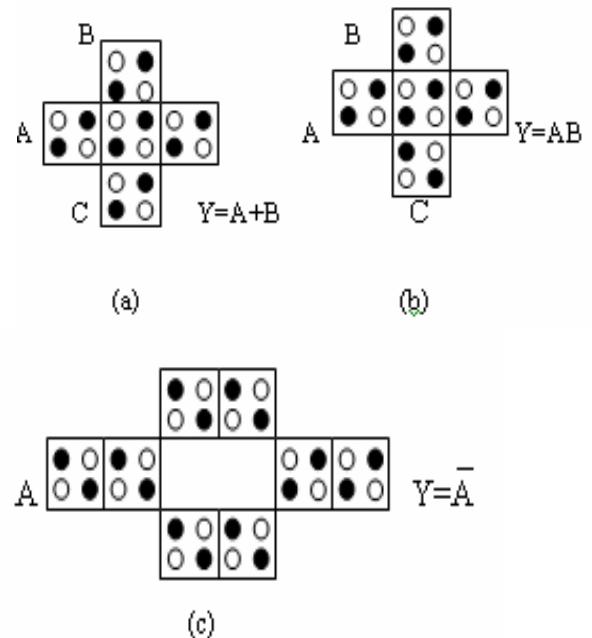


Figure 2. (a) QCA OR Gate (b) QCA AND Gate(c) QCA Inverter circuit.

C. QCA Clocking Scheme

Multi phased clocking mechanism is used in the QCA circuit to flow the information. Four distinct and periodic clock signals are required for QCA circuits to operate and these signals are shifted by 90 degree of relative phase difference with each other. Clock signals are generated through an electric field which is applied to the cells to control the tunneling barrier height between the dots occurs from a mono-stable to a bi-stable state. The four clocking zones can be divided as switch, hold, release and relax state as shown in Fig 3. In switch phase the actual computation occurs according to the input. Here in this phase, inter dot potential barriers are low first and then barriers are raised and QCA cell become polarized to the state of its input driver. By the end of this clock phase, barriers are high enough and there is no possibility of electron tunneling and hence cell states are fixed. The second phase of clocking at where barriers are still high and does not allow the electron tunneling and hence output of this

state can be used as the inputs to the next. In the third phase i.e. in release phase potential barrier is lowered and electron positions are configured according to the original state. The cells are in relaxed or unpolarized state which is the fourth clocking phase. So the overall polarization of the QCA cell is determined in the switch and hold phase where cells are in polarized state depending upon the neighboring cell polarizations. In release and relax phases the states are unpolarized.

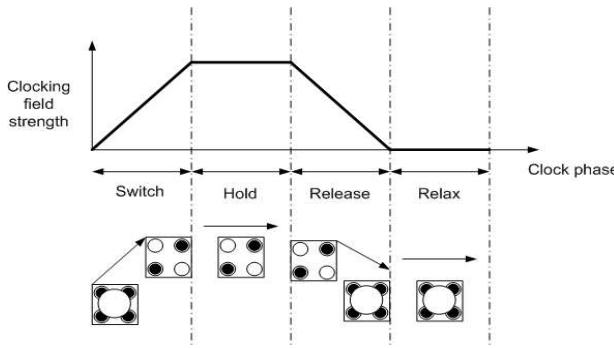
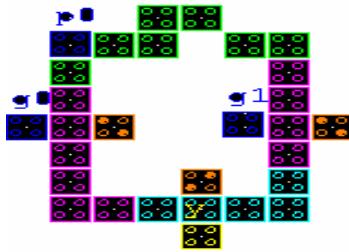


Figure 3. QCA clocking with four clocking zones.

III. PROPOSED DESIGN OF MULTIPLEXER AND DEMULTIPLEXER

In this section 2: 1 multiplexer and demultiplexer designs are proposed. The proposed designs are composed of 29 and 23 cells illustrated in Fig 4(a) and 5(a), which are the first designs of 2 bit multiplexer and demultiplexer of such a less number of cells. The output $Y = \text{maj}(\text{maj}(g_1, P_0 \text{ bar}, 0), \text{maj}(g_0, P_0, 0), 1)$. Where g_1 and g_0 are the inputs and P_0 is the select line. For the two inputs multiplexer will give just one output. For the demultiplexer just for one data two outputs are there. The first output $Y_0 = dS_1 \text{ bar}$. And for the second output $Y_1 = dS_1$, where d is the data line. When select line S_1 is zero output $Y_0 = d$ and when select line S_1 is one output $Y_1 = d$. Fig. 4(b) shows the simulated output for the multiplexer for different inputs. When select line P_0 is high, output follows g_0 , when P_0 low output follows g_1 at clock 2. Fig 5(a) shows the circuit diagram of 2 bit demultiplexer and Fig 5(b) shows the simulated output. For $S_1 = 1$ output Y_1 follows the data d . For $S_1=0$ output Y_0 follows d . Fig 6(a) and Fig 6(b) shows the previous designs by Sara Hashemi et al [6], L Sousa et al [8], Fig 7(a) shows the design proposed by K.Kim et al [9]. Fig 7(b) shows the design proposed by M. Askari [7]. In [12] M. Amiri et al proposed multiplexer based CLB design where they designed the 2 bit multiplexer with 34 cells as shown in Fig 7(c). In [19] Debarko et al proposed the design as shown in fig 7(d) with 49 cells. Compared to these designs the proposed design consumes less area and have lesser complexity. Sec IV shows the robustness of the new design.



(a)



(b)

Figure 4. (a) Two bit QCA Multiplexer. (b) Simulated output

Table 1 shows the comparison of the proposed method with the previously designed methods. Table 2 shows the improvement of proposed design compared to other previously designed methods. The proposed design is more robust and reliable compared to other previous designs. Sec IV shows the robustness of the new design. By checking every possibility of the above mentioned design the proposed design can be consider the best method of design of a 2 bit multiplexer and demultiplexer.

TABLE I. COMPARISON OF SOME 2:1 MULTIPLEXER DESIGNS WITH THE PROPOSED DESIGN

| Improvement | mux presented in [6]. Fig(6a) | mux presented in [8]. Fig(6b) | mux presented in [9]. Fig(7a) | mux presented in[7]. Fig(7b) |
|-------------------------|-------------------------------|-------------------------------|-------------------------------|------------------------------|
| Complexity (cell count) | 19% | 67% | 36% | 14.7% |
| Area | 50% | 78.57% | 62.5% | 25% |

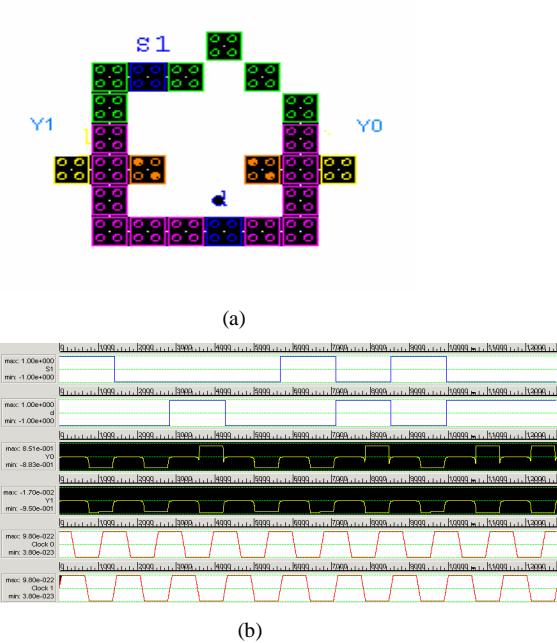
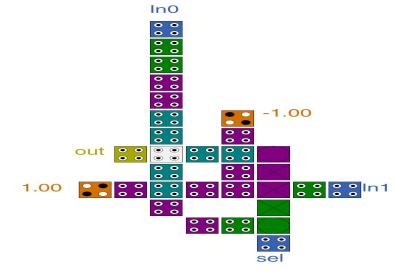


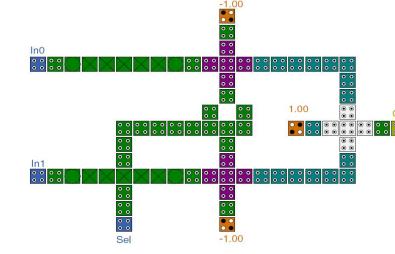
Figure 5.(a) Two bit demultiplexer (b) Simulated output

TABLE II. IMPROVEMENT OF PROPOSED MULTIPLEXER

| 2:1 Multiplexer | Complexity(cell count) | Area (μm^2) |
|---------------------------------------|------------------------|--------------------------|
| Multiplexer presented in [8] Fig(6b) | 88 | 0.14 |
| Multiplexer presented in [9] Fig(7a) | 46 | 0.08 |
| Multiplexer presented in [6].Fig(6a) | 36 | 0.06 |
| Multiplexer presented in [19].Fig(7d) | 49 | 0.06 |
| Multiplexer presented in [7].Fig(7b) | 34 | 0.04 |
| Proposed Multiplexer. Fig(4a) | 29 | 0.03 |



(a)



(b)

Figure 6. (a) Multiplexer proposed in [6] (b) proposed in [8]

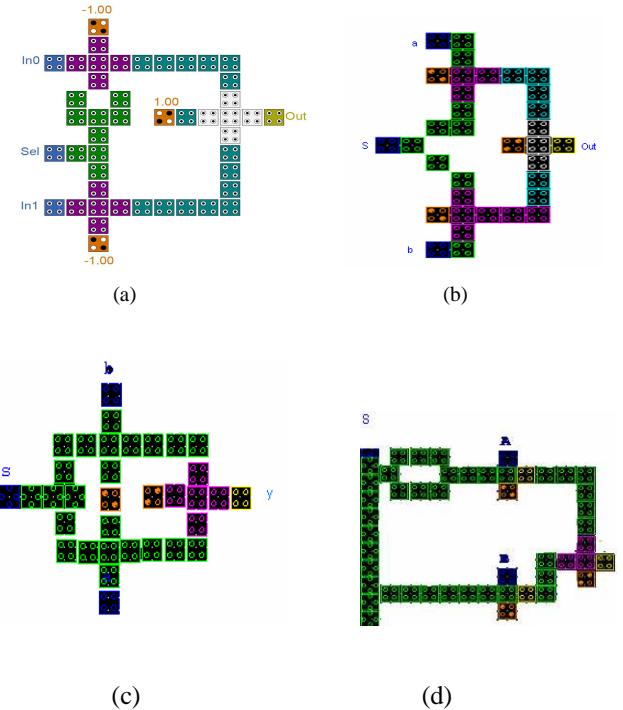


Figure 7. (a) Multiplexer proposed in [9] (b) Multiplexer proposed in [7] (c) Multiplexer proposed in [12] (d) Multiplexer proposed in [19]

IV. FAULT ANALYSIS

In Quantum cellular automata fault may arise due to manufacturing variations and device defects such as missing of one cell, misplacement of cell, orientation due to angular displacement of cell etc. QCA array is analogous to a two dimensional crystal lattice. The defects affect the properties of crystal lattice as well as performance of the QCA array defined in [13, 14]. External energy defined as Ekink can cause a cell in a wire or a system to switch into a mistake state. Suppose the cell in one extremity of the wire is externally forced to change its polarization, and suddenly a “kink” in polarizations appear between the first and the second cell of the array. This “kink” must then propagate through the array until it reaches the other extremity. Thus as the distance between cells increases the kink energy will decrease indicating that a smaller amount of external energy could excite a cell into a mistake state which is undesirable. If the angle of off centeredness between cells increases the kink energy also decrease. This energy has a minimum value when the two cells have the same polarization, and has the maximum value when they have opposite polarizations. The difference between these maximum and minimum values is called the “kink” energy. The electrostatic energy of a system composed by two cells (cell a and cell b, with respective polarizations p_a and p_b , side by side. The total energy of the two cells is calculated by the sum of the electrostatic energy between each of the four quantum-dots of cell a, (with charge q_{ai} and location r_{ai}) and each of the four quantum-dots of cell b, (with charge q_{bj} and location r_{bi}); both i and j range from 1 to 4, as there are 4 quantum dots in each cell. The kink energy is the amount of external energy that will excite a cell into a mistake state or create a kink in the transmission. That is why we get a binary ‘1’ instead of ‘0’.

$$\text{Kink energy is defined as } E_{\text{kink}} = E_{p_a \neq p_b}^{a,b} - E_{p_a = p_b}^{a,b}$$

In this section we introduce some faults in multiplexer and demultiplexer circuits in the majority gate, inverter and the output sections by misplacement of cells which are the main parts of the design. We present the simulation results due to these errors such as misalignment of majority gate which is very sensitive for input as well as output, inverter, input defect (cell displacement) in assembly of array fabrication. We also demonstrate the possibility of designing fault tolerant QCA circuit of multiplexer and demultiplexer. In Fig 8(a) we introduce some fault in input gate g0 of the multiplexer and test the circuit for several inputs and we observe that displacement of input cell g0 is permissible for less than 8 nm. Simulated output waveform is shown in Fig 8(b). After 8nm the output is not correct one. We also misplace the majority gate up to 6 nm of the multiplexer as shown in Fig 9(a), it is still possible to design circuits that performed to the desired level of output despite of their faulty nature, after this the output is not correct one. The output waveforms are shown in Fig 9(b). Outputs are taken in clock 2 for every case. We also showed another fault in the output section by 4 nm, after which the output will be erroneous.

For select input P0 is zero output always follows input g1. The corresponding circuit diagram and simulated results are shown in Fig 10 (a) and in Fig 10(b). Fig 11(a) shows the circuit diagram of misplacement fault of 6nm is permissible in inverter gate, after that distance the error occurs in the output. The error less simulated output is shown in Fig 11(b). We introduce some faults in demultiplexer circuits also as shown in Fig 12 (a) where one of the majority gates of the output section is misplaced by 5 nm; the corresponding simulated output is shown in Fig. 12(b). In this paper we propose the 2 bit multiplexer and demultiplexer design that offer remarkable robustness with respect of input and output cell misplacement. The simulation shows that the collective cellular behavior exhibits a very high degree of robustness. The output has been checked for different input patterns.

The proposed design is the most suitable design from the previous designs. Various faults of QCA based designs are analyzed in [15, 16, 17, 18].

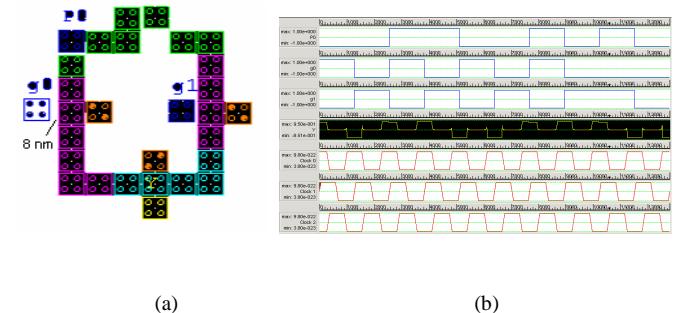


Figure 8. Layout of misplacement of input cell by 8 nm. (b) Simulated output

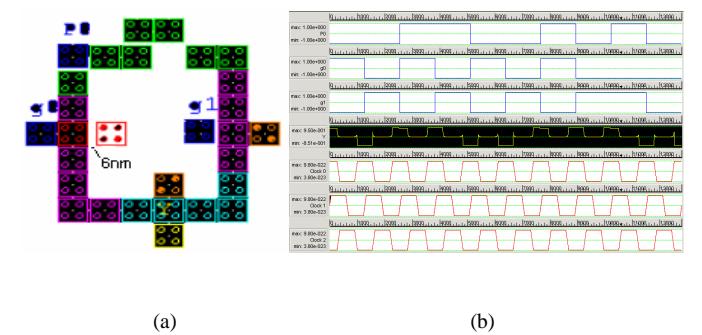


Figure 9. Layout for misplacement of majority gate of input section by 6nm of multiplexer. (b) Simulated output after displacement

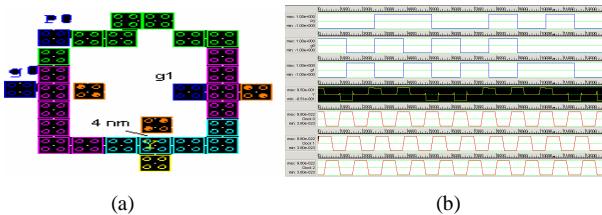


Figure 10. (a) Fault for misplacement of majority gate by 4nm of multiplexer in Output section. (b) Simulated output after displacement

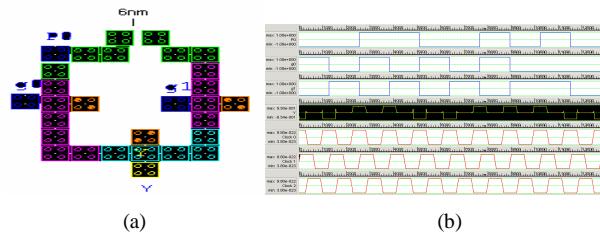


Figure 11. (a) Fault for misplacement of inverter gate by 6 nm of multiplexer (b) Simulated output after displacement

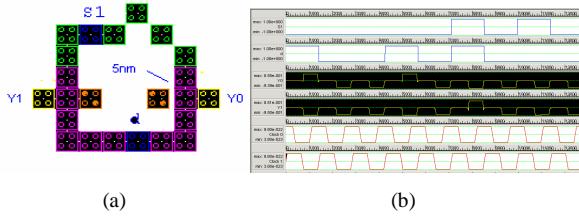


Figure 12. (a) Fault for misplacement of majority gate of Y0 output by 5nm of demultiplexer (b) Simulation due to fault by 5 nm of majority gate of Y0 output of demultiplexer

V CONCLUSION

The paper presents the design and analysis of a 2 bit multiplexer and demultiplexer. Multiplexers are very important digital device, it selects one input among so many inputs and steers the information to the output. The multiplexer and demultiplexer circuits using QCA which has been discussed here present some challenges and the new proposed design of two bit multiplexer compared with some established designs. As the number of cells of the structure is reduced the relevant energy scales increases even in the higher temperature environment. This scalability down to atomic dimensions is the key feature of nanotechnology. From the proposed design of 2:1 multiplexer other larger input multiplexers (such as 4:1, 8:1 etc.) can be easily designed. The proposed designs analyze the outputs inspite of some fabrication error such as misplacement of input cell and cell misplacement in majority gate and check the output upto maximum displacement; after that the error in output has occurred. The design schemes does not fail as the size is reduced rather it becomes more robust. From the analysis it can be definitely conclude that the proposed design is less complex and reliable among all the previous designed multiplexers and can be composed as the building block for more complex device structures and hierarchical design.

ACKNOWLEDGEMENT

Authors are grateful to Department of Science and Technology (DST) for sanctioning a research Project under Fast Track Young Scientist scheme reference no.: SERC/ET-0213/2011 under which this paper has been completed.

REFERENCES

- [1] Lent,C.S, Tougaw, P.D., W.Porod and G.H. Bernstein, *Quantum Cellular Automata* , Nanotechnology, 4(1993), pp. 49-57.
- [2] Touggaw, P.Douglas, Lent, Craig S, *Logical devices implemented using quantum cellular automata*, Journal of Applied Physics 75(3) (1994), pp. 1818-1825.
- [3] Tougaw, P.Douglas, Lent,Craig S, *Logical Devices implemented using quantum cellular automata*, J of App phy 75(3) (1994) pp. 1818- 1825
- [4] J. Timler , C.S. Lent, *power gain and dissipation in quantum dot cellular automata* , J of App phy, (1) (2002), pp. 823-831
- [5] W.J.Townsend , J.A . Abraham, 4th. IEEE conference Nano technology, (2004), pp.625.
- [6] Sara Haesmi, M.R. Azghadi, *A Novel Multiplexer design*, International Symposium on Telecommunications, 2008..
- [7] M.Askari, M. Taghizadeh, K.Fardad, *Digital Design Using QuantumDot Cellular Automata (A Nanotechnology Method)*, Proceedings of the International Comference on Computer and Communication Engineering 2008. Malaysia
- [8] T. Teodosio and L. Sousa, *QCA-LG: A tool for the automatic layout generation of QCA combinational Circuits*, In 25th IEEE Norchip Conference, IEEE, Denmark, November 2007.
- [9] K. Kim, K. Wu, and R. Karri, *The Robust QCA Adder Designs Using Composable QCA Building Blocks*, IEEE Transactions on Computer-Aided Design Of Integrated Circuits and Systems, 26(1), 2007.
- [10] <http://www.qcaddesign.ca>
- [11] C.S.Lent, P.D. Tougaw, W.Porod ans G.H. Bernstein, *Quantum Cellular Automata*, Nanotechnology, .4(1993), pp.49-57.
- [12] M. Amin Amiri, M Mahadvi, *QCA Implementation of a Mux based FPGA CLB* , ICONN 2008.
- [13] R.B.Philips, *Crystal defects and microstructure: modeling across scales* New York, NY, University Press (2001).
- [14] J. weertman, J.R weertman, *Elementary dislocation theory*, New York NY, Oxford University Press (1992).
- [15] N.Jha, S. Gupta, *Testing of digital systems* , Cambridge University Press. Cambridge, 2003
- [16] A. Fijany , B . N Toomarian, *New design for Quantum Dot Cellular Automata to obtain fault tolerant logic gates*, J. Nanopart Res (3) (2001), pp. 27-37.
- [17] M.Thoori, J . Huang, M. Momenzadeh, F Lombardi, *Testing of quantum cell automata* , IEEE Transaction Nanotechnol, 3(4), (2004), pp. 432-442
- [18] K.Das, D.De , *QCA defect and fault analysis of diverse nanostructure for implementing logic gates* , Int J Recent Trends Eng Finl. 3(1) (2010), pp. 1-5.
- [19] D. Mukhopadhyay, S. Dinda, P.Dutta, *Designing and Implementation of Quantum Cellular Automata 2:1 Multiplexer circuit*, International Journal of Computer Applications. 25(1) (2011).

Hairpin model of CRLH Transmission Line and Floating Slot Approach for Wide-Band Bandpass Filter

Pratik Mondal¹, Bhupesh Mukherjee², Arabinda Roy³, Susanta Kumar Parui⁴

Department of Electronics and Telecommunication Engineering
Bengal Engineering and Science University, Shibpur, Howrah-711103
Email: pratik665@gmail.com; arkapv@yahoo.com

Abstract-This paper presents construction of a composite right/left-handed (CRLH) transmission line (TL) using Hairpin microstrip coupled lines. For getting high coupling in comparison to conventional edge coupled microstrip line, the floating slot on the ground plane is used. Two configuration named as SOS(short-open-short) the OSO(open-short-open) model are studied for realization of band pass filter for the Hairpin model. The floating ground equalizes the even and odd mode phase velocities and solves the difficulties arised in the fabrication process as it relaxes the required physical dimensions like coupled gaps between lines. The Hairpin structure makes the structure more compact in size. Also using Defected Microstrip Structure(DMS) on the feed line minimizes the harmonics and gives a good stopband. Finally a bandpass filter having Fractional bandwidth of 70.83% for OSO model and 44.05% for SOS model are obtained.

Keyword: Composite right/left handed transmission line(CRLH); Open-short-open; Short-open-short; Floating slot; bandpass filter; Hairpin structure.

I. INTRODUCTION

Composite right/left-handed (CRLH) transmission lines (TLs) are TL based on metamaterials, which have recently led to numerous novel concepts and applications [1]. The CRLH transmission line developed a new class of highly functional microwave devices such as power dividers, couplers and filters [2]. For conventional transmission line i.e.; for right handed transmission line, a series inductor and a shunt capacitor generally is used. If the position of the inductor and the capacitor interchanges, the obtained structure is referred as left-handed transmission line. The models of a purely right-handed and left-handed transmission line shown in Fig.1. The left-handed transmission line is commonly known as metamaterials. Left handed transmission lines do not exist naturally, only they can be designed artificially using lumped elements and periodically loading a normal transmission line. The CRLH TL is generally implemented by loading a host TL periodically with series capacitances and shunt inductances, and exhibits both the left-handed (LH) and right-handed (RH) propagation bands. As the right-handed, parasitic effects cannot be avoided and when the realization of the left-handed TL is done, Composite right/left-handed (CRLH) transmission lines represent a practical left-handed TL [4]. The CRLH TL concept was introduced because of purely LH TL cannot be able to exist due to parasitic.

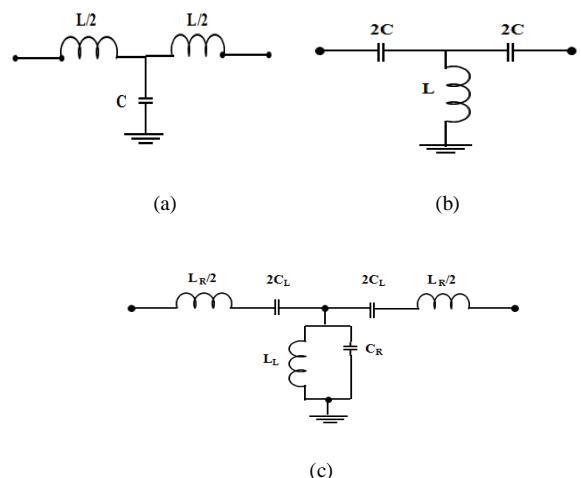


Fig 1. Schematic of the (a) Right handed TL (b) Left handed TL (c) CRLH TL

The broadside coupled coplanar waveguides was chosen to achieve tight coupling with practical dimensions, which reflects an enhancement in the bandwidth [3-4]. Also, filter using CRLH TL is studied[5].Now, for compact design of the bandpass filter hairpin model is used here. In this paper, a new type of hairpin bandpass filter has been realized using CRLH structure and floating slot on ground plane. It is shown that by proper design of the slotted ground, the coupling between the two microstrip lines can be increased dramatically keeping practical dimensions for the coupled line width and spacing.

II. COMPOSITE RIGHT/LEFT HANDED TRANSMISSION LINE

The normal-modes of symmetrical coupled lines are normal & coupled mode. Normal modes are classified as even and odd mode. The coupling between symmetrical lines can be determined in terms of characteristic impedance and phase velocities and for asymmetrical lines it can be designed as ‘C’ and ‘π’ parameters. The coupling between two microstrip lines will be described using homogeneous dielectric medium equations, where the electrical lengths are the same for both the modes .The current and voltage of the four port network as shown in “Fig. 2” can be related as , $V=[Z][I]$.

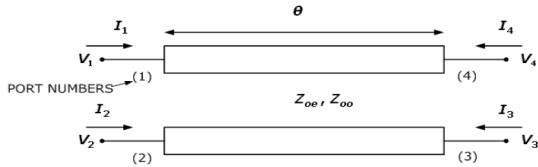


Fig.2: Schematic diagram of pair of coupled lines

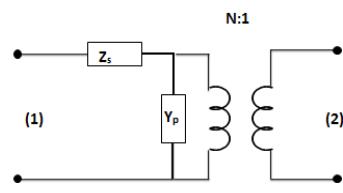
The port 3 and port 4 terminated with short ($V_3=0$) and open ($I_4=0$) circuits; respectively to realize the unit cell approach and the network becomes a two-port network. By eliminating I_3 from the equations we can get the ABCD transmission matrix –

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \frac{Z_{oe} - Z_{oo}}{Z_{oe} + Z_{oo}} (1 - \frac{4Z_{oe}Z_{oo}(\cot \theta)^2}{(Z_{oe} - Z_{oo})^2}) & \frac{-2jZ_{oe}Z_{oo}\cot \theta}{Z_{oe} - Z_{oo}} \\ \frac{-2j\cot \theta}{Z_{oe} - Z_{oo}} & \frac{Z_{oe} + Z_{oo}}{Z_{oe} - Z_{oo}} \end{bmatrix}$$

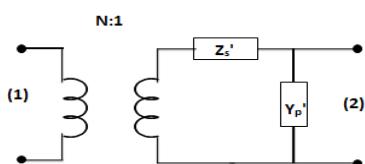
It has been observed from the above equation that, CRLH cell can be represented as series impedance and parallel admittance and can be represented by either two configurations depending on the transformer position as shown in Fig. 3. The equivalent circuit parameters in Fig. 3 are given by,

$$Z_s = \frac{1}{jG\mathcal{C}_s} = \frac{-2jZ_{oe}Z_{oo}\cot \theta}{Z_{oe} + Z_{oo}} \quad Y_p = \frac{1}{jG\mathcal{D}L_p} = \frac{-2j(Z_{oe} + Z_{oo})\cot \theta}{(Z_{oe} - Z_{oo})^2}$$

$$N = \frac{Z_{oe} - Z_{oo}}{Z_{oe} + Z_{oo}} \quad Z_s' = N^2 Z_s \quad Y_p' = \frac{Y_p}{N^2}$$



(a)



(b)

Fig 3. (a) Realization1 (b) Realization2

The parallel admittances Y_p and Y_p' acts as inductive susceptance for the coupled line electrical length $\theta < \pi/2$ and capacitive for $\theta > \pi/2$. Similarly for Z_s and Z_s' . To cancel the presence of the ideal transformer ,the CRLH unit cell may be constructed by cascading the above said circuit with its mirrored image. Here we found two configurations, either the unit cell shown in "Fig. 4(a)" and its equivalent circuit is

shown in "Fig. 4(b)" as T-network, or the circuit of "Fig. 5(a)" and its equivalent circuit is shown in Fig. 5(b) as II network. The configuration in "Fig. 4(a)" is known as OSO (open-short- open ports) and the configuration of "Fig. 5(a)" is SOS(short-open-short port).The design equations for the OSO and the SOS terminating impedance (i.e. $Z_o = 50$ ohm) are

$$Z_o^2 = Z_{oe}Z_{oo} \left(\frac{Z_{oe} - Z_{oo}}{Z_{oe} + Z_{oo}} \right)^2 \text{ for OSO} \quad Z_{oe}Z_{oo} \left(\frac{Z_{oe} + Z_{oo}}{Z_{oe} - Z_{oo}} \right)^2 \text{ for SOS}$$

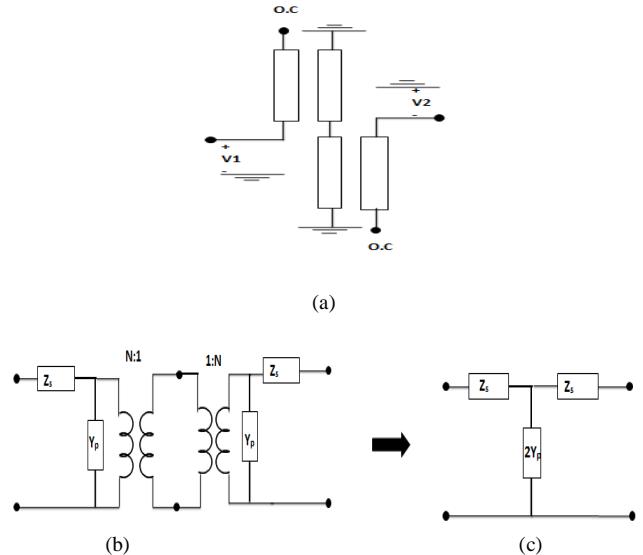
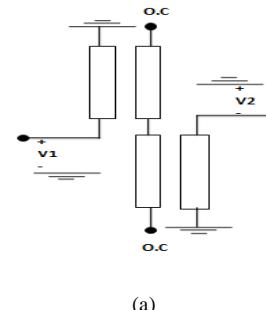
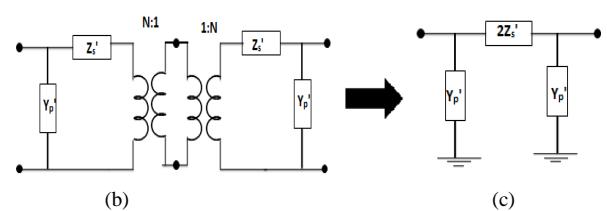


Figure 4. (a) OSO Coupled line configuration. (b) cascaded OSO configuration (c) equivalent T -Network



(a)



(b)

Figure 5. (a) SOS Coupled line configuration. (b) cascaded SOS configuration (c) equivalent π -Network

The floating slot at the ground plane solves the difficulties arising in the design procedure and the fabrication process as it relaxes the requirements on physical dimensions of the lines. In this design, adding the floating slot in the ground conductor as tends to equalize the even and the odd phase velocities and also increase the coupling dramatically. Also, by introducing the hairpin structure the compactness in size occurs.

III. IMPLEMENTATION OF HAIRPIN CRLH STRUCTURE

For getting High coupling the floating slot is used and for compact structure hairpin model is used. As, the CRLH transmission line network is a combination of the right and left-handed network, it is a bandpass filter with wide bandwidth for both the models.

In "Fig. 6" as illustrated, a substrate of dielectric constant, $\epsilon_r = 4.4$ and thickness, $h = 1.59$ mm and loss tangent = 0.02 is used. Here, the gap between two coupled lines $S_g=0.5$ mm is considered. To realize $Z_{in} = 184.46 \Omega$ and $Z_{out} = 73.78 \Omega$, the values of $W = 0.68$ mm, $S_g=0.5$ mm and the gap between slots $S_s=5$ mm are obtained. The length of the coupled line of the unit cell is taken as 26.25 mm and length of the floating slot as 25.75 mm for both the SOS and OSO model.

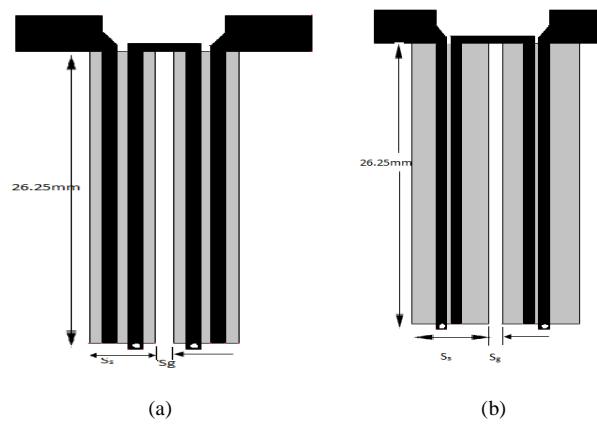


Figure 6. Schematic diagram of the (a) proposed OSO Coupled hairpin bandpass filter (b) SOS coupled CRLH hairpin bandpass filter

The Hairpin CRLH structure is implemented by using edge coupled microstrip lines with a floating ground is already shown in "Fig. 6" and the simulated S-parameters are shown in "Fig. 9". Using MOM based IE3D software. "Fig. 9(a)" shows for OSO model, after the pass band there is a stop band of 1.5 GHz at 30dB and the center frequency for the first harmonics is 4.7GHz. Here, to remove the harmonics as per the requirements of the proposed band pass filter we have used the DMS (Defective Magnetic Structures) in the 50 ohm lines.

In the present design as illustrated in "Fig. 7", all the parameters are taken same instead of a Defected Microstrip structure is added on the feed line. The width of the DMS slot is taken $g=0.2$ mm. The length of the the DMS structure are different for the need of suppression of the different harmonics. Here for SOS configuration –

DMS structure($a=7.8$ mm, $b_1=2$ mm, $b_2=2$ mm, $c_1=3$ mm, $c_2=3$ mm, $d_1=0.95$ mm, $l=2.72$ mm) for both the port .Also for OSO model – (a)DMS for left Port($a=5.8$ mm, $b_1=1.4$ mm, $b_2=1.2$ mm, $c_1=2.6$ mm, $d_1=0.8$ mm, $l=1.72$ mm) (b) (a)DMS for right Port($a=5.8$ mm, $b_1=1.4$ mm, $b_2=1.2$ mm, $c_1=2.6$ mm, $d_1=0.8$ mm, $e_1=0.8$ mm, $l=1.72$ mm)

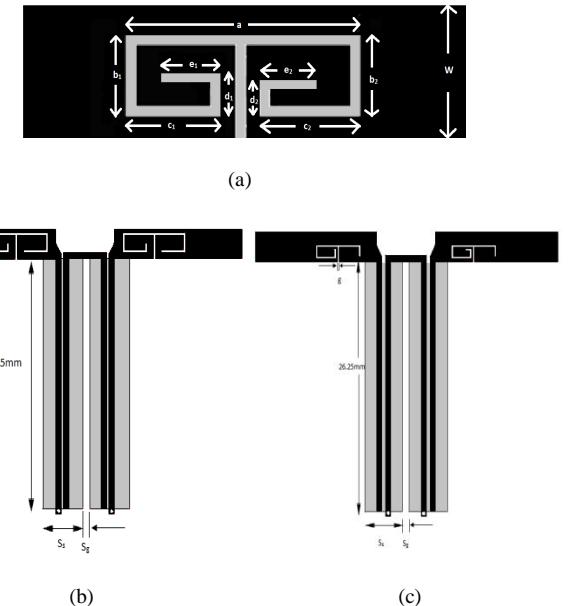
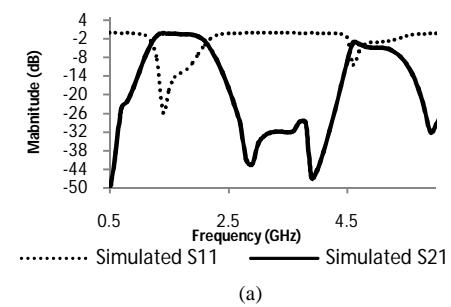


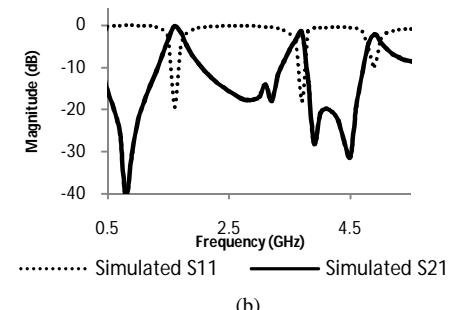
Figure 7. Schematic diagram of the (a) DMS structure (b) proposed OSO Coupled hairpin bandpass filter with DMS (c) SOS coupled CRLH hairpin bandpass filter with DMS

IV. RESULT AND DISCUSSION

The hairpin CRLH structure of the both OSO and SOS model has been simulated using MoM based IE3D software and the simulated S-parameters are shown in "Fig. 9". The center frequency at 1.78 GHz and 3dB bandwidth of 1.16 GHz with insertion loss of -0.373 dB is observed in OSO model. Also the Selectivity 52.94 dB/GHz for rising edge and 43.54 dB/GHz for falling edge is observed. The center frequency at 1.61 GHz and 3dB bandwidth of 0.25GHz is observed in SOS model. The selectivity for SOS model is 17.69 dB/GHz for rising edge and 35.411 dB/GHz for falling edge. The simulated results for the SOS model shows insertion loss of -0.56dB.



(a)



(b)

Fig 9. simulated S-parameters of the Hairpin CRLH-TL (a) OSO model and (b) SOS model

It is found that OSO model is suitable wide band application whereas the SOS model may be used for the narrow band application. Also by using the hairpin structure instead of normal CRLH structure as mentioned previous section, coupling increases which makes wider 3dB bandwidth for both the models and also the structure becomes more compact. It is also observed that higher order harmonics are present in the design which is minimised in the next section using DMS on the feed line.

The hairpin CRLH band pass filter with DMS of both OSO and SOS model has been simulated using MoM based IE3D software and the simulated S-parameters are shown in "Fig. 10". The center frequency at 1.56 GHz and 3dB bandwidth of 0.97 GHz with insertion loss of -0.17dB is observed in OSO model. Also the Selectivity 57.44 dB/GHz for rising edge and 37.5 dB/GHz for falling edge is observed. The center frequency at 1.79 GHz and 3dB bandwidth of 0.72GHz is observed in SOS model. The selectivity for SOS model is 60.01 dB/GHz for rising edge and 49.9 dB/GHz for falling edge.

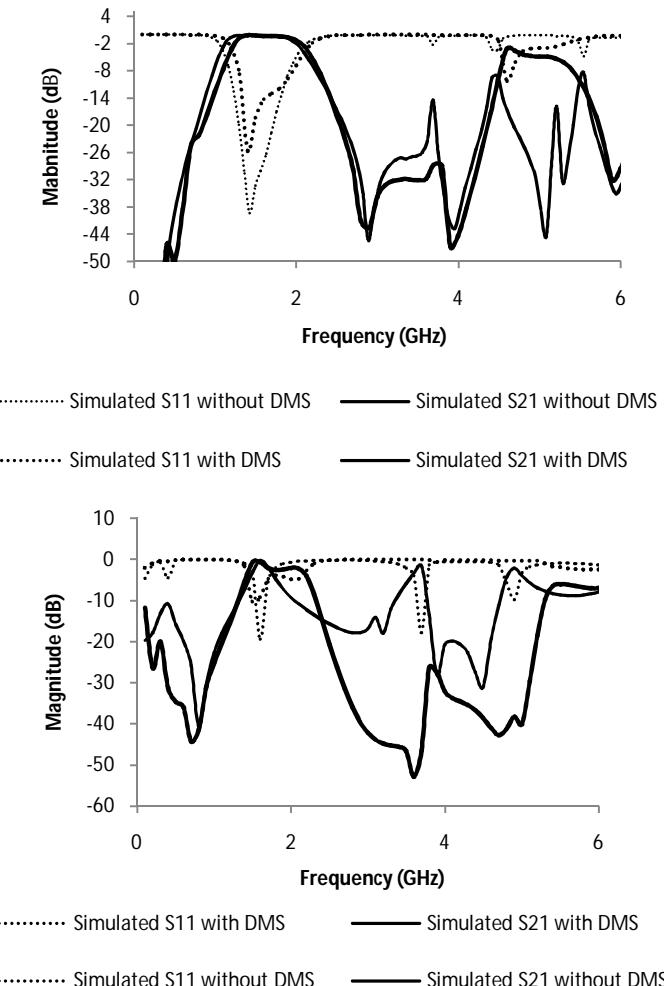


Fig 10. Comparative study between simulated S-parameters of the Hairpin CRLH filter with DMS and without DMS (a) OSO and (b) SOS model

It is observed that for OSO model all the harmonics are removed by the use of Defective magnetic structures on the both 50 ohm lines. Also, SOS model may be used for the narrow band application where by using DMS a wide stop band of around 3 GHz is obtained after the pass band. Also

by using the hairpin structure instead of normal CRLH structure miniaturisation occurs.

Prototype filter for both the model has been fabricated on FR-4 substrate which has dielectric constant of 4.4 and thickness of 1.59mm as shown in "Fig. 8(a)" and "Fig. 8(b)".⁸ The fabricated unit has been measured using Agilent make vector network analyzer (model N5230A). The measured results in "Fig. 10" shows center frequency at 1.44 GHz , 3dB bandwidth of 1.02GHz and fractional bandwidth of 70.83 % for OSO model and center frequency at 1.782 GHz, 3dB bandwidth of 0.765GHz and fractional bandwidth of 44.05% for SOS model. The EM-simulated result of Hairpin structure with DMS and without DMS are compared in "Fig. 10".

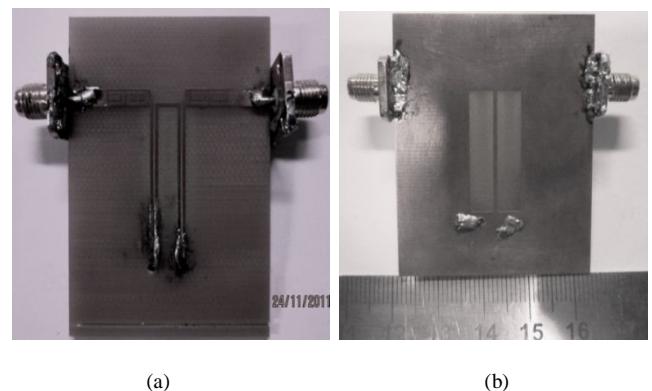


Fig 8.(a): Photograph of the proposed hairpin CRLH bandpass filter with DMS (OSO model) (a)Top Layer (b) BottomLayer

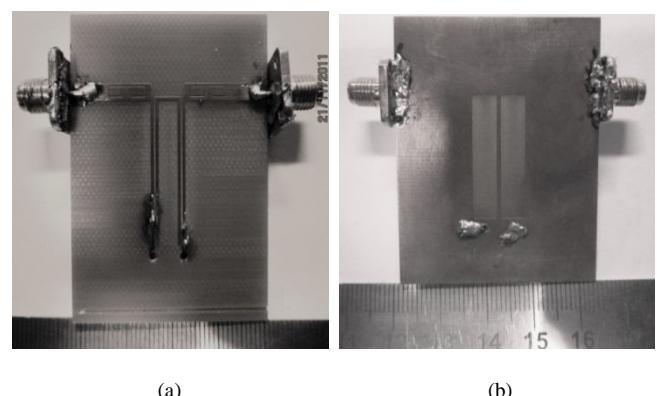
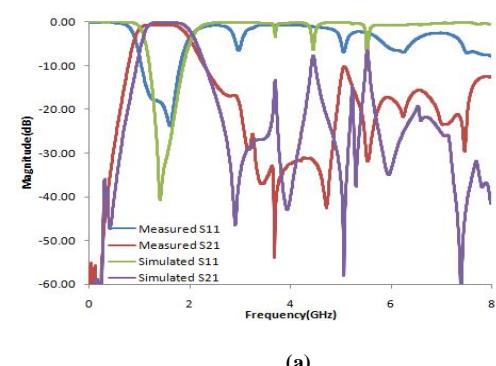
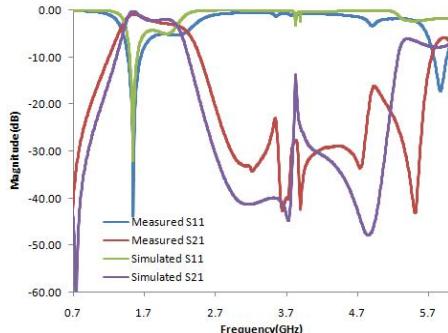


Fig 8.(b): Photograph of the proposed hairpin CRLH bandpass filter with DMS (SOS model) (a)Top Layer (b) BottomLayer

Also, the EM-simulated and experimentally measured results of hairpin bandpass filter with DMS are compared in "Fig. 11".





(b)

Fig 11.Comparative study between measured & simulated s-parameters of hairpin CRLH filter with DMS (a)for OSO model and (b)for SOS model

V. CONCLUSION

We propose to design a hairpin bandpass filter using composite right/left-handed (CRLH) transmission lines (TL) and floating slot in ground plane. We verified this idea through implementation on coupled microstrip lines with slotted ground. It is found that OSO model is for wide band application whereas the SOS model may be used for the narrow band application. The vias for both the modelling are almost identical except the positions are different. Also by using the hairpin structure instead of normal CRLH structure, coupling increases which makes wider bandwidth for both the models and also the structure becomes more compact. It is also observed that higher order harmonics are present in the design which is minimised using DMS on the feed lines. It offers low pass band insertion loss, wide bandwidth, high selectivity and high Fractional Bandwidth (FBW). Such CRLH filters may be found suitable in various applications.

ACKNOWLEDGMENT

This work is supported by CSIR (Council of Scientific and Industrial Research), Govt. of India.

REFERENCES

- [1] F. Bongard, J. Perruisseau-Carrier, J.R. Mosig, "Enhanced CRLH Transmission Line Performances Using a Lattice Network Unit Cell," IEEE Microwave and Wireless Components Letters, vol. 19, pp. 431–433, 2009.
- [2] Sun Mingming, Qin Weiping, Wang Haimeng, Chen Dong, "Resonator and Bandpass Filter Using CRLH Transmission Line Based on Microstrip-Coplanar-Waveguide Structure," International Conference on Microwave and Millimeter Wave Technology, vol. 3, pp. 1530–1592, 2008.
- [3] Aya F. Abdelaziz, Tamer M. Abuelfadl, and Osman L. Elsayed, "Realization of Composite Right/Left-Handed Transmission Line Using Broadside Coupled Coplanar Waveguides," IEEE Antennas and Propagation Society International Symposium, pp. 1-4,2009.
- [4] A. F. Abdelaziz, T. M. Abuelfadl, and O. L. Elsayed, "Realization of composite right/left-handed transmission line using coupled lines," Progress In Electromagnetics Research, PIER 92, pp. 299–315, 2009.
- [5] S. Pallaviram, "Microstrip coupled line filter elements as CRLH transmission line unit cells," Applied Electromagnetics Conference (AEMC),pp. 1 - 4 ,2009.

Programmable Array Logic Design using Quantum Dot Cellular Automata

Ratna Chakrabarty¹, Arnab Sarkar¹

Department of Electronics Engineering
Institute of Engineering & Management
Salt Lake, Sector-V, Kolkata, India
[e-mail- ratna_chakrabarty@yahoo.co.in](mailto:ratna_chakrabarty@yahoo.co.in)
arnabcricke@gmail.com

Debashis De²

Department of Computer Science & Engineering
West Bengal University of Technology
Salt Lake, Kolkata, India.
School of Physics
University of Western, Perth, Australia
Dr.debashis.de@gmail.com

Abstract— Quantum-dot Cellular Automata (QCA) has been widely considered as a new paradigm in the field of Nanotechnology. The basic feature of QCA require extremely low power and with a potential for high density and regular package system. These characteristics make QCA an attractive technology for manufacturing memories in which the paradigm of memory-in-motion can be fully exploited. This paper proposes Programmable Array Logic (PAL) architecture for QCA implementation. This architecture is based on utilizing new building blocks (referred to as tiles) in the storage and input/output circuitry of the memory. An extensive comparison of the proposed architecture and previous QCA serial memories is pursued in terms of latency, timing, clocking requirements, and hardware complexity.

Keywords: *QCA; Majority Voter; Macro Modelling; PAL.*

INTRODUCTION

Quantum Cellular Automata is a novel nanotechnology and a suitable alternative of traditional CMOS technology with respect to its size and power consumption [1], [2]. The CMOS technology is a useful one when it is used as a current switch to send the digital information in electrical form. To send the binary information by just switching on and off the current switch is a novel one, but it has some serious drawback when the size of the device is reduced as then the quantization of charge in both the doping level and in channel level become very significant one according to Lent et al [3]. Physical limits of using the CMOS devices have been successfully overcome in nanoscale level from the concept proposed in [4] which offers a solution at nanoscale level with a new method of computation and transformation of information that transfer the charge potential without the electron motion but the electron configuration within a quantum cell. Both individual QCA cell (semi-conductor and metallic) and multiple QCA arrangement

have been fabricated and tested [5],[6],[7]. Such a device is composed of four metal dots, connected with tunnel junctions and capacitors. Experiments have confirmed that the switching of a single electron in a double-dot cell can control the position of a single electron in another double-dot cell. The basic logic behavior with these cells has been demonstrated in [8], using its basic block as a majority voter (MV). It has been reported that room temperature operation requires QCA cells to be fabricated in the range of 1-5 nm in size. Lieberman et al. in [9], [10], [11] have proposed some possible realizations of molecular QCA. It describes the progress toward making QCA molecules at an extremely small cell size at room temperature. In this paper a Programmable Array Logic (PAL) is developed using QCA. Programmable logic devices can be built using the PAL circuit which can be a useful part for developing Microcomputer. QCA modelling tools available for such designs have been at the layout level. There are several approximate simulators available at the layout level, such as the bistable simulation engine and the nonlinear approximation methods [12], [13], [14].

The paper is organized as follows, section I describes the basic principles of quantum cell, section II describes the PAL in digital circuits, section III describes PAL using QCA, section IV describes simulations and result discussion and section V concludes the paper.

I. BASIC PRINCIPLES OF QUANTUM CELL

A QCA cell consists of four metallic dots positioned in four corners of a square quantum cell. Each cell has two mobile electrons, which carry the charge information from one cell to another according to the position of electrons in the dots due to mutual Columbic repulsion. The information in QCA flows without the electron movement from one cell to another cell. How the device cell reaches its lowest energy state is explained in [15]. Fig1 shows the polarization of quantum cell according

to the electron position in the quantum dots of that particular cell. According to the position of electrons in the dots of a quantum cell the polarization is given by the equation

$$P = \frac{(p_1+p_3)-(p_2+p_4)}{p_1+p_2+p_3+p_4}$$

The mutual coulomb repulsion between the electrons thus results in bistability between the $P = +1$ and $P = -1$ states.

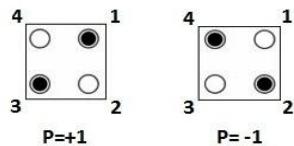


Figure 1. Quantum cells with four dots and polarization of $P = +1$ and $P = -1$.

A. The basic QCA logical device

The fundamental QCA logical circuit is the three-input majority gate that appears in Figure 2. The Majority gate can be realized by only 5 QCA cells [16]. The majority gate is used to build the fundamental logic AND gate and OR gate according to the polarization of one of the input gate. Computation is performed with the majority gate by driving the device cell (cell 4 in the figure 2.) to its lowest energy state. This happens when it assumes the polarization of the majority of the three input cells. An input cell simply as one that is changed by a signal that is propagating in a direction that is toward the device cell. The device cell will always assume the majority polarization because it is this polarization where electron repulsion between the electrons in the three input cells and the device cell will be at a minimum.

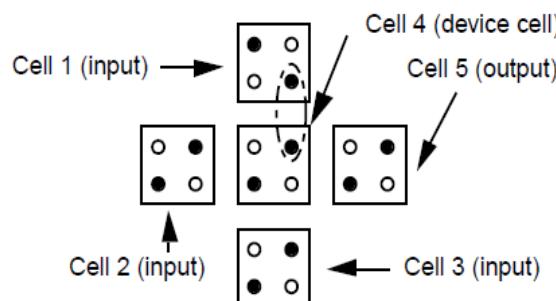


Figure 2. The fundamental QCA logic gate (majority gate)

If A, B and C are the three inputs and Y the output, the logical function for the majority gate is given by

$$Y = AB + BC + AC \quad (1)$$

The AND function can be implemented by setting one value (A, B, or C) in equation 1 to a logical 0. Similarly, the

OR function can be implemented by setting one value (A, B, or C) in equation 1 to a logical 1. This results in the equations:

$$\text{AND} = AB + B(0) + A(0) = AB$$

$$\text{OR} = AB + B(1) + A(1) = A + B$$

With this property i.e. the ability to generate the AND and OR functions, any logical circuit can be generated with QCA devices. Fig 3 shows the majority OR and AND gates.

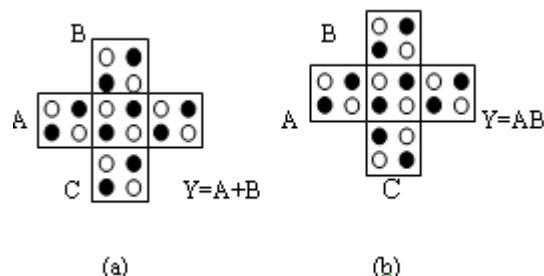


Figure 3. (a) Majority OR gate with $P = +1$ in C input (b) Majority AND gate with $P = -1$ in C input

B. QCA Wire

Fig 4 illustrates how a binary value propagates down the length of a QCA "wire" [17,18]. In this figure, the wire is a horizontal row of QCA cells. The binary signal propagates from left-to-right because of the Columbic interactions between cells.

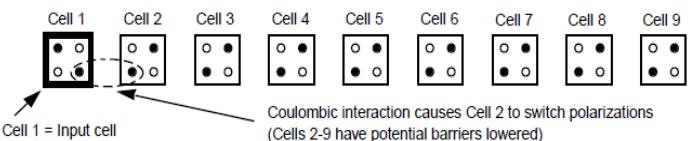


Figure 4 . A QCA Binary wire.

In the fig 4, a binary 0 (from polarization $P = -1$) will propagate down the length of the wire because of the Columbic interactions between cells. Initially, the electron repulsion caused by columbic interaction between cell 1 and cell 2 will cause cell 2 to change polarizations. Then, the electron repulsion between cell 2 and cell 3 will cause cell 3 to change polarizations. This process will continue down the length of the QCA "wire".

C. QCA Inverter

In logic circuits, an inverter flips the sense of a logic signal. A low signal input through an inverter results in a high output and vice versa. By arranging array of cells, as shown in fig 5, an inverter can be constructed. If the polarization of the one end is fixed, the polarization of the other end will be opposite.

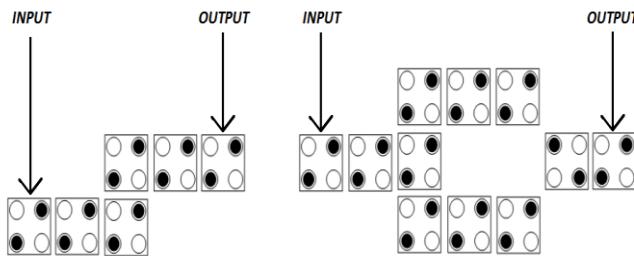


Figure 5. An Inverter circuit in QCA

D. QCA Clocking Phenomenon

Each QCA cell is synchronously latched and unlatched with the changing of the clock signal and therefore the information is distributed throughout the circuits [19], [20], [21]. QCA does not need any power to operate only four phased clock signals as shown in fig.6 are required for the operation of the circuit. The four clocking zones can be divided as switch, hold, release and relax state. In switch phase inter dot potential barriers are low and QCA cell become polarized according to the state of its input driver. In the second phase of clocking at where barriers are still high and does not allow the electron tunneling and output of this state is used as the inputs to the next. Potential barrier is lowered in the release phase and electron positions are configured according to the original state. The cells are relaxed or unpolarized in the fourth clocking phase. The overall polarization of the QCA cell is determined in the switch and hold phase where cells are in polarized state. The cells are in unpolarized state in the release and relax phases.

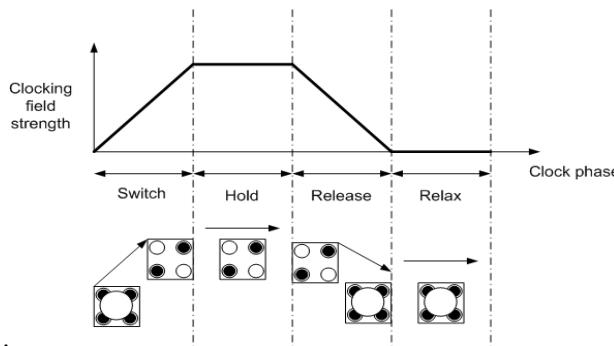


Figure 6. Four clocking phases in the QCA cell.

II. PAL IN DIGITAL CIRCUITS

PAL is an array of logic gates in a single chip with the AND- OR configuration. It has a programmable AND array and a fixed OR array as shown in Fig7.

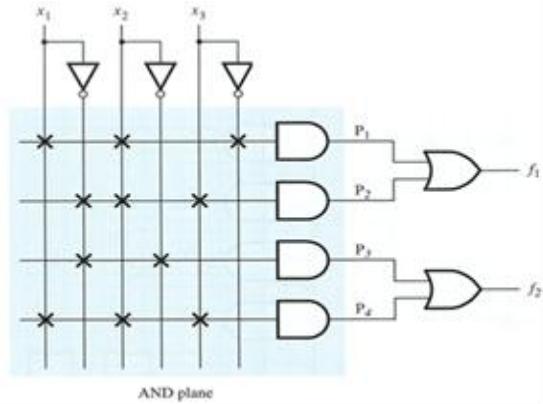


Figure 7. An example of PAL in Digital Electronics.

III. PAL IN QUANTUM CELLULAR AUTOMATA

In QCA we design the PAL Similar to digital electronics where there are more than one single layer. We have to fuse the interconnects as per our necessity. There are many technologies in doing this process, for example fuse and anti-fuse technology. In PAL we fuse the line of input signals with the wires which are connected to the AND gates, thus programmable AND logic is there. Then the output from the AND are connected to the fixed OR, thus non programmable OR logic is there. In the picture shown below we have made the equivalent PAL of the above circuit in QCA. Here the four structures are shown in two dimensions whereas in the real life case it will be a three dimensional structure in which the four individual block (each working as an AND function of multiple inputs) will be one over another. Here we have first shown a single layer in fig 8 and then a 3 layered (fig 9) quantum cell for propagation of signals or charge in proper way without affecting the charge of any other cells in the circuit.

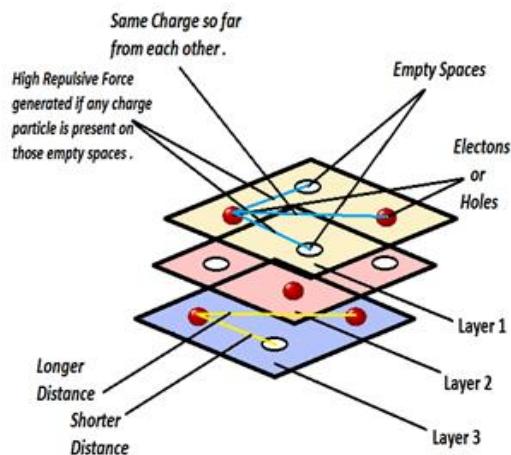


Figure 8. Single layered quantum cells.

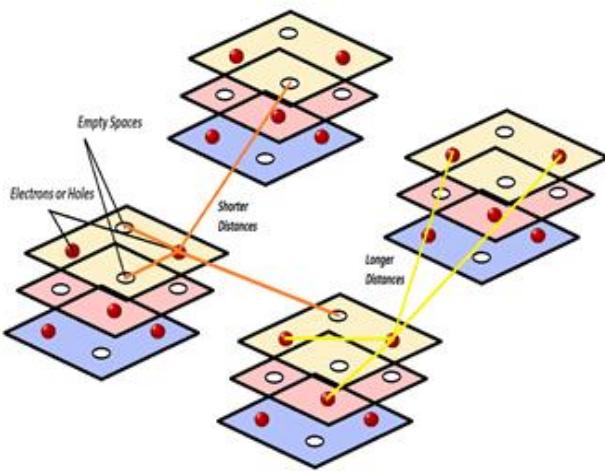


Figure 9. Three layered quantum cells.

IV. SIMULATION AND RESULT DISCUSSION

Following Fig 9 shows the PAL design using QCAD designer tool [22] in QCA.

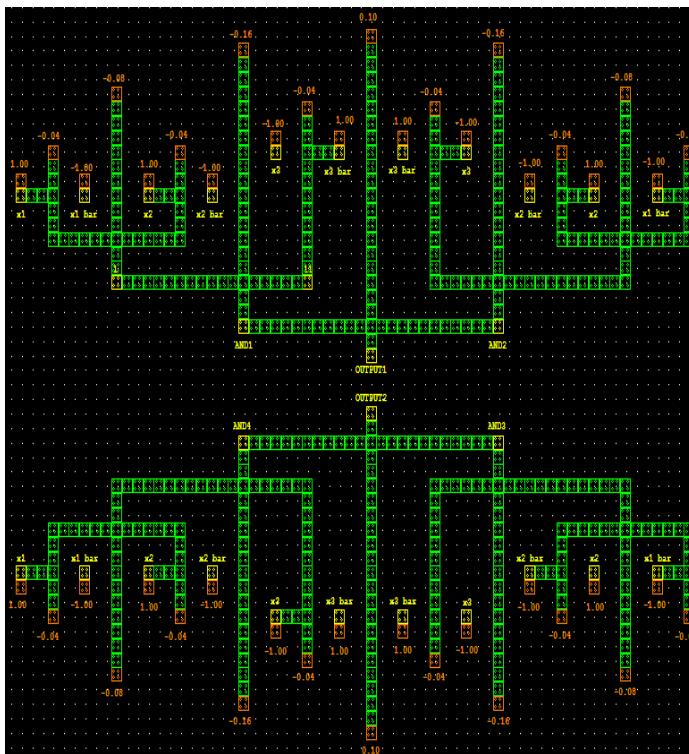


Figure10: The replica of the above PAL in fig7 in QCA.

Here,

$x_1 = 1$, $x_2 = 1$ and $x_3 = -1$.

$AND1 = P1$, $AND2 = P2$, $AND3 = P3$, $AND4 = P4$

$OUTPUT1 = f1$ and $OUTPUT2 = f2$

(refer to fig 7 for $P1, P2, P3, P4, f1$ and $f2$)

Simulated output of the PAL circuit using QCAD designer tool is shown in fig 11 . From the simulation we get the following operations such as

- The 1st AND gate gives an output ‘1’ as all the i/p’s to the AND gate were ‘1’
- The 2nd AND gate gives an output ‘-1’ as all the i/p’s to the AND gate were ‘-1’
- The 1st OR gate gives an output ‘1’ as the i/p’s to the OR gate were ‘1’ and ‘-1’.
- The 3rd AND gate gives an output ‘-1’ as the i/p’s to the AND gate were ‘0’ and ‘-1’.
- The 4th AND gate gives an output ‘-1’ as the i/p’s to the AND gate were ‘1’ and ‘-1’.
- The 5th OR gate gives an output ‘1’ as both the i/p’s to the OR gate were ‘-1’.



Figure11. Simulated output using QCAD designer tool.

2nd Annual International Conference IEMCON 2012

V. CONCLUSION

In this paper, QCA PAL circuit is proposed; this model is based on experimental demonstrations of QCA cells. Programmable array logic is used for designing circuits easily. For example large function which has large number of variables can be easily implemented by using programmable array logic.

This can further be used in complex circuitry combining it with other circuits, which will lead to the development of more superior or advanced architectures that is by bottom-up design process similar to that of the traditional CMOS design flow.

ACKNOWLEDGEMENT

Authors are grateful to Department of Science and Technology (DST) for sanctioning a research Project under Fast Track Young Scientist scheme reference no.: SERC/ET-0213/2011 under which this paper has been completed.

REFERENCES

- [1] Craig S Lent, P. Douglas Tougaw, "A Device Architecture for Computing with Quantum Dots," Vol 84,no-4,1997
- [2] S. Srivastava, S. Sarkar and S. Bhanja, "Power Dissipation Bounds and Models for Quantum-dot Cellular Automata Circuits", in *IEEE Transactions on Nanotechnology*, vol. 8, no. 1, pp. 116-127, 2009.
- [3] Craig S Lent, P. Douglas Tougaw, "A Device Architecture for Computing with Quantum Dots," Vol 84,no-4,1997
- [4] R. Compano, L. Molenkamp, and D.J. Paul, "Technology Roadmap for Nanoelectronics," European Commission IST Programme, Future and Emerging Technologies, 2000.
- [5] A.O. Orlov, I. Amlani, G.H. Bernstein, C.S. Lent, and G.L. Snider, "Realization of a Functional Cell for Quantum-Dot Cellular Automata," *Science*, vol. 277, pp. 928-930, 1997.
- [6] A. Orlov, R. Kummamuru, R. Ramasubramaniam, C. Lent, G. Bernstein, and G. Snider, "Clocked quantum-dot cellular automata shift register," *Surface Science*, vol. 532, pp. 1193–1198, 2003.
- [7] R. Kummamuru, A. Orlov, R. Ramasubramaniam, C. Lent, G. Bernstein, and G. Snider, "Operation of a quantum-dot cellular automata (qca) shift register and analysis of errors," *IEEE Transactions on Applied Physics*, vol. 50, pp. 1906–1913, September 2003.
- [8] I. Amlani, A.O. Orlov, G. Toth, C.S. Lent, G.H. Bernstein, and G.L.Snider, "Digital Logic Gate Using Quantum-Dot Cellular Automata," *Science*, vol. 284, no. 5412, pp. 289-291.
- [9] M. Lieberman, S. Chellamma, B. Varughese, Y. Wang, C.S. Lent,G.H. Bernstein, G. Snider, and F. Peiris, "Quantum-Dot Cellular Automata at a Molecular Scale," *Annals of the New York Academy of Sciences*, vol. 960, pp. 225-239, 2002.
- [10] C. Lent, B. Isaksen, and M. Lieberman, "Molecular quantum-dot cellular automata," *Journal of American Chemical Society*, vol. 125, pp. 1056–1063, 2003. March 8, 2006—9 : 06 pm DRAFT 33
- [11] W. Hu, K. Sarveswaran, M. Lieberman, and G. H. Bernstein, "High-resolution electron beam lithography and dna nanopatterning for molecular qca," *IEEE Transactions on Nanotechnology*, vol. 4, pp. 312–316, May 2005.
- [12] G. Toth and C. Lent, "Role of correlation in the operation of quantum-dot cellular automata," *Journal of Applied Physics*, vol. 89, pp. 7943–7953, June 2001.
- [13] I. Amlani, A. Orlov, G. Snider, C. Lent, W. Porod, and G. Bernstein, "Experimental demonstration of electron switching in a quantum-dot cellular automata (qca) cell," *Superlattices and Microstructures*, vol. 25, no. 1/2, pp. 273–278, 1999.
- [14] K. Walus, T. Dysart, G. Jullien, and R. Budiman, "QCADesigner: A rapid design and simulation tool for quantum-dot cellular automata," *IEEE Trans. on Nanotechnology*, vol. 3, pp. 26– 29, March 2004.
- [15] Craig S Lent, P Douglas Tougaw, Wolfgang Porod and Gary H Bernstein Department of Electrical Engineering, University of Notre Dame, December 1992.
- [16] Tougaw, P. D. & Lent, C. S. (1994). Logical Devices Implemented Using Quantum Cellular Automata. *Journal of Applied Physics*, Vol. 75, No. 3, (1994) page numbers 1818-1825.
- [17] Michael Thaddeus Niemier, B.S.Department of Computer Science and Engineering Notre Dame, Indiana .January 2004
- [18] P.D. Tougaw and C.S. Lent. Logical devices implemented using quantum cellular automata. *Journal of Applied Physics*, vol,75,pp-1818, 1994.
- [19] Hennessy, K. & Lent, C. S. (2001). Clocking of Molecular Quantum-dot Cellular Automata. *Journal of Vacuum Science and Technology B*, Vol. 19, No. 5, (Sep. 2001) page numbers 1752-1755.
- [20] Amiri, M. A.; Mahdavi, M. & Mirzakuchaki, S. (2008). QCA Implementation of a Mux-Based FPGA CLB, *Proceedings of International Conference On Nanoscience and Nanotechnology*, pp. 141-144, Australia, Feb. 2008, Melbourne.
- [21] Vankamamidi, V.; Ottavi, M. & Lombardi, F. (2008). Two-Dimensional Schemes for Clocking/Timing of QCA Circuits. *IEEE Transaction on Computer Aided Design of Integrated Circuits and Systems*, Vol. 27, No. 1, (2008) page numbers 34-44.
- [22] <http://www.qcaddesign.ca>

PERFORMANCE ANALYSIS OF TRANSISTOR LEVEL ADDER CIRCUITS BY VARRYING NUMBER OF TRANSISTORS AND IMPLEMENTING DIFFERENT LOGIC STYLE OF ADDITION

Subhadeep Nandy and Rajib Pal

Department of Electronics and Communication Engineering
Camellia Institute of Technology
Kolkata, India

Abstract - In this paper, a structured approach for analyzing the adder design is introduced. Analysis is based on some simulation parameters like power, delay and power- delay- product (PDP). Simulation is done by HSPICE in $0.18\mu\text{m}$ CMOS Technology at 1.8 V supply voltage. Three input pulses at different frequencies and sum and carry signals are taken from the COSMOS SCOPE software. The different circuit designs are studied and evaluated extensively. Simulations of several designs give new information which are applicable for different requirements. Each of these circuits cell exhibits different power consumption, delay and power-delay- product in different VLSI technology. This paper can be said as a library of different full adder circuits that will be beneficial for the circuit designers to pick the full adder cell that satisfied their specific application.

Keyword – Arithmatic circuit, Full adder, Power, Delay, Power-Delay-Product, HSPICE, COSMOS SCOPE.

I. INTRODUCTION

The Explosive growth in laptop, portable system, and cellular networks has intensified the research efforts in low power microelectronics. Today we find number of portable applications requiring low power and high throughput circuits. Addition is one of the fundamental arithmetic operations. It is used extensively in many VLSI systems such as application-specific DSP architectures and microprocessors. In addition to its main task, which is adding two binary numbers, it is the nucleus of many other useful operations such as subtraction, multiplication, division, address calculation, etc. In most of these systems the adder is part of the critical path that determines the overall performance of the system. That is why enhancing the performance of

the 1-bit full-adder cell (the building block of the binary adder) is a significant goal. Recently, building low-power VLSI systems has emerged as highly in demand because of the fast growing technologies in mobile communication and computation. The battery technology doesn't advance at the same rate as the microelectronics technology. There is a limited amount of power available for the mobile systems. So designers are faced with more constraints: high speed, high throughput, small silicon area, and at the same time, low-power consumption [1]. So building low-power, high - performance adder cells is of great interest. Designing systems aiming for low power is not a straightforward task, as it is involved in all the IC design stages beginning with the system behavioral description and ending with the fabrication and packaging processes. In some of these stages there are guidelines that are clear and there are steps to follow that reduce power consumption, such as decreasing the power-supply voltage. While in other stages there are no clear steps to follow, so statistical or probabilistic heuristic methods are used to estimate the power consumption of a given design[2], [3].

II. OVERVIEW OF POWER DISSIPATION

Power dissipation in CMOS digital circuits is categorized into two types: peak power and time-averaged power consumption. Peak power is a reliability issue that determines both the chip lifetime and performance. The voltage drop effects caused by the excessive instantaneous current flowing through the resistive power network affects the performance of a design due to the increased gate and interconnect delay. This large power consumption causes the

device to overheat which reduces the reliability and lifetime of the circuit. Also noise margins are reduced, increasing the chance of chip failure due to crosstalk. The time-averaged power consumption in conventional CMOS digital circuits occurs in two forms: dynamic and static. Dynamic power dissipation occurs in the logic gates that are in the process of switching from one state to another. During this process, any internal and external capacitance associated with the gate's transistors has to be charged, thereby consuming power.

Static power dissipation is associated with inactive logic gates (i.e., not currently switching from one state to another). Dynamic power is important during normal operation, especially at high operating frequencies, whereas static power is more important during standby, especially for battery-powered devices. Dynamic dissipation has historically been far greater than static power when the system are active, and hence static power is often ignored, although this will change as great and sub-threshold leakage increases.

III. REVIEW OF FULL ADDER CELLS

Various static CMOS logic styles have been used to implement low-power and high-performance 1-bit Full Adder cells. The 1-bit conventional CMOS full adder cell [4] has 28 transistors is shown in Fig.1. Different logic styles can be investigated from different points of view. Evidently, they tend to favour one performance aspect at the expense of others. In other words, it is different design constraints imposed by the application that each logic style has its place in the cell library development. The CMOS structure combines PMOS pull-up and NMOS pull-down networks to produce considered outputs. In this style all transistors (either PMOS or NMOS) are arranged in completely separate branches, each may consist of several sub-branches. Mutually exclusiveness of pull-up and pull-down networks is of a great concern.

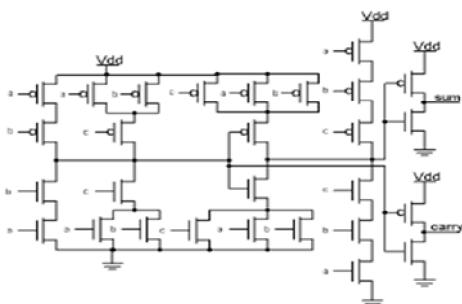


Fig. 1 Conventional CMOS full Adder

Fig.2 shows the Complementary Pass-Transistor Logic (CPL) Full adder [5]. Using pass transistor logic with CMOS output inverters, CPL adder features complementary inputs and outputs. These circuits generate many intermediate nodes and their complements in order to generate the final signals (Sum and C_{out}). Having a signal and complement produces high rate of switching activities. Therefore CPL full adder cell is not a suitable for low power applications.

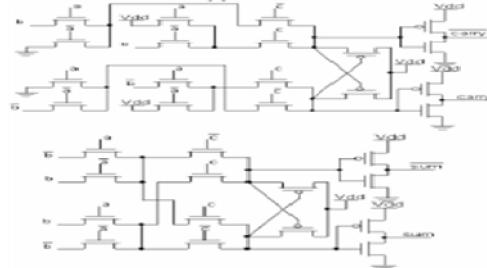


Fig. 2 The Complementary Pass-transistor Logic (CPL) Full Adder

Transmission function adder (TFA) [6] has simpler schematic than the conventional CMOS full adder shown in Fig.3. The TFA needs fewer MOS transistors in comparison with the previous one. The transistor count of TFA is 16, while the conventional CMOS full adder requires 28 transistors. TFA provides buffered outputs of the proper polarity with equal delay for sum and carry.

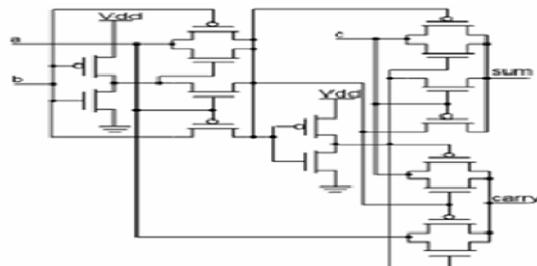


Fig. 3 Transmission function adder (TFA)

Figure 4 shows the schematic of 1-bit full adder MTCMOS circuit. A sleep control scheme is introduced for efficient power management. During active mode, sleep transistor (SL) is set low and sleep control transistors (MP and MN) are turned on. Since their on-resistances are small, the virtual supply voltage (VDDV and VSSV) almost function as real power lines. In the standby mode, SL is set to high and MN and MP will be turned off thus the leakage current is low.

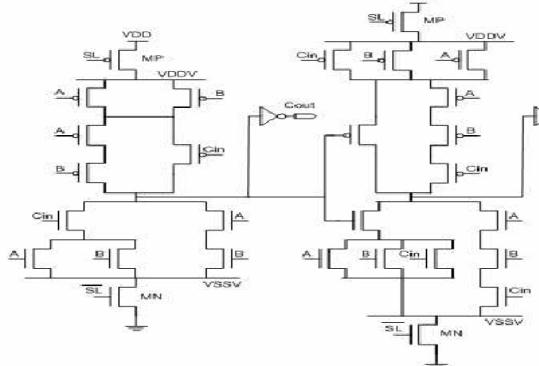


Fig. 4 1-bit full adder MTCMOS circuit

N-CELL1 has 14 transistors and utilizes the low power XOR/XNOR circuit and a pass transistors network to produce a non full swing Sum signal and uses four transistors to generate a full swing C_{out} signal, which do not provide enough driving power [6] is shown in Fig. 5.

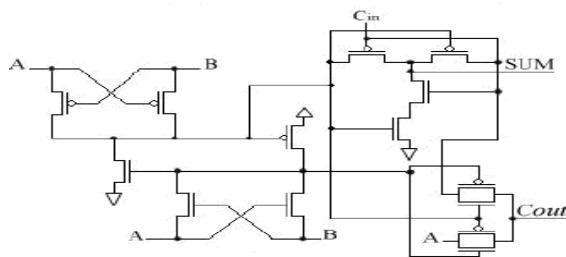


Fig. 5 N-Cell1

The 1-bit full adder circuit is a very important component in the design of application specific integrated circuits[5].This paper presents a novel low-power multiplexer-based 1-bit full adder that uses 12 transistors (MBA 12T) in Fig. 6. In addition to reduced transition activity and charge recycling capability, this circuit has no direct connections to the power-supply nodes, leading to a noticeable reduction in short-current power consumption.

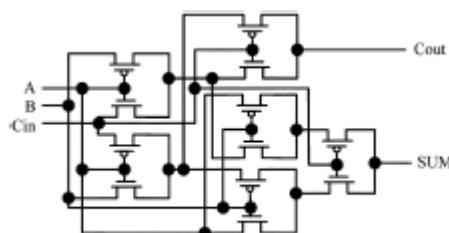


Fig. 6 MBT 12T

In comparison to ADDR10T13A, Fig. 7, the new MBA-12T adder has the ability of reclaiming all the node charges during the discharging cycle of those nodes. Since the new MBA-12T uses 12 transistors, one can expect a modest increase of the internal node Capacitances compared to ADDR10T13A adders. However, the switching activity of the MBA-12T adder will be considerably lower than that of previously proposed 10-transistors adders. That's because all of those adders (SERF and 10Ts) use an internally generated signal (XNOR) to control the output transistor gates (four gates in total).

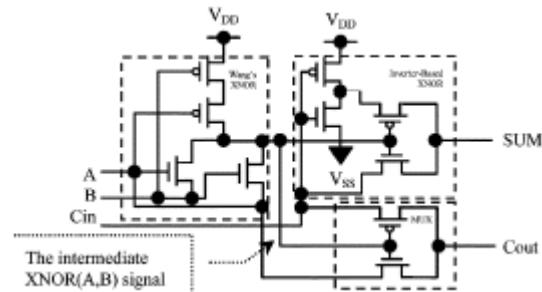


Fig. 7 ADDR10T13A

IV. FUNCTIONALITY TEST



Fig. 8 Functionality test of the MBT12T full adder circuit

V. SIMULATION RESULTS

The simulation result shows that in MBT12T full adder circuit the three inputs are taken in different frequencies. The three inputs of different frequencies are represented by the first three curves in Fig. 9. The output of sum & carry is shown by the next two curves respectively in Fig. 9 . Simulation result (Input & Output pattern) of MBT12T full-adder circuit is taken by using *COSMOS SCOPE* software.



Fig. 9: Input and output waveform of MBT12T adder (COSMOS SCOPE View)

TABLE I COMPARISON OF FULL ADDER CELLS

| Cells | Avg.power (μw) | Delay (μs) | PDP (J) |
|-------------------|-----------------------------|-------------------------|---------|
| Conventional CMOS | 1.178 | 130.183 | 153.35 |
| CPL | 1.018 | 83.73 | 85.24 |
| TFA | 0.627 | 2.860 | 1.92 |
| MTCMOS | 0.587 | 2.801 | 1.64 |
| N-Cell | 0.521 | 2.736 | 1.43 |
| ADDR10T13A | 0.519 | 2.632 | 1.36 |
| MBT12T | 0.489 | 2.112 | 1.03 |

VI. CONCLUSION

In this paper, various one bit full adder cells design has been reviewed from the most recent published research work. The comparison of full adder cells with each other in term of power, delay and power-delay-product is done. Based on survey, it is conclude that the MBT12T have good signal level, consume less power and have high speed compare to all other designs at low supply voltage. This circuit is suitable for arithmetic circuits and other VLSI applications with very low power consumption and very high speed performance.

REFERENCES

- [1] N. Weste and K. Eshraghian , *Principles of CMOS VLSI Design, A System Perspective*. Reading, MA: Addison- Wesley, 1993.
- [2] G. M. Blair, "Designing low-power CMOS," Inst. Elect. Eng. Electron. Commun. Eng. J., vol.6, pp. 229–236, Oct. 1994.
- [3] S.Devadas and S. Malik, "A survey of Optimization techniques targeting low-power VLSI circuits," in Proc. 32nd ACM/IEEE Design Automation Conf., San Francisco, CA, June 1995, pp. 242–247.
- [4] Shubhajit Roy Choudhary, Aritra Banerjee, Anirudha Roy, Hiranmay saha, "A high speed 8 transistor full adder design using Novel 3 transistor XOR gates" in International journal of electronics, Circuits and systems 2;4 @www.waset.org Fall 2008.RECENT ADVANCES in NETWORKING, VLSI and SIGNAL PROCESSING.
- [5] A Novel Multiplexer-Based Low- Power Full Adder.Yingtao Jiang, Abdulkarim Al-Sheraidah, Yuke Wang, Edwin Sha, and Jin Gyun Chung.
- [6] A.M. Shams, T.K. Darwish, M.A. Bayoumi, "Performance analysis of low-power 1-bit CMOS full adder cells," IEEE Transactions on VLSI Systems, Vol. 10, pp. 20–29, Jan.2002

SOCIAL ENGINEERING & SOCIAL SCIENCE

Date : 18th January,2012

Chairman : Prof. Jyoti Banerjee

Co-Chairman : Prof. Mohuya Chakraborty

Rapporter : Prof. Arighna Deb

| Paper ID | Paper Title | Authors |
|----------|--|--|
| 4 | "Is Emerging Green Technology Poses Threat of Trade Wars-Some Solutions?-A Survey" | Subro S. Thakur,Somasree Bhadra and Sougata Dey |
| 14 | Computational Optimization of Speed in an Unplanned Lane Traffic | Prasun Ghosal, Arijit Chakraborty and Sabyasachee Banerjee |
| 32 | PREMASA MODEL:-An Intgrated Approach to Customer Relationship Management | Debabrata Das |
| 36 | Energy Conservation In Buildings | Avijit Bandopadhyay |
| 53 | Design Of a Universal Solar Power Driven Mobile Phone Charger For Rural Areas Of India | Priyanka Roy, Swati Dey and Dr. Madhurima Chattopadhyay |
| 79 | Energy Consumption Modeling for Maintenance Cost Evaluation in RFID Inventory Tracking | Tanvi Agrawal, Pradip K Biswas, A.D.Raoot |
| 104 | Soil Fertility: Can We Estimate ? Assessment of techniques of soil fertility measurement | Ritam Sanyal, N. G. Shah |

Soil Fertility: Can We Estimate ?

Assessment of techniques of soil fertility measurement

Ritam Sanyal, N. G. Shah

Centre for Technology Alternatives for Rural Areas
Indian Institute of Technology, Bombay
Mumbai, India
ritam_sanyal@iitb.ac.in

Abstract— Agriculture will continue to play a significant role as livelihood for a large portion of Indian population for many decades to come. Sustenance in agriculture necessitates maintenance of soil fertility to acceptable level to assure desired productivity. Ideally, there should be proper soil fertility management system in place which would ensure sustenance of proper soil health. Soil fertility management comprises two components. They are soil fertility enhancement and soil fertility testing. Since “pre-green revolution” era, scientific communities as well as farmers have been searching for a convenient method for measuring soil fertility. For the purpose of soil testing, conventional laboratory methods are mostly practiced in India. These methods are generally laborious, time taking and require soil samples to be brought to laboratory from field. The agricultural scientific community prefers detailed soil analysis that aims to evaluate status of sixteen soil nutrients, including both macro and micro elements, as well as soil moisture and soil physical attributes. Conventional soil testing makes analyzing all the sixteen elements a very tedious and difficult process. The work represented here attempts to explore possible methods of soil testing which overcomes the drawback of conventional soil testing methods. This article describes the important concept of soil fertility and soil fertility testing by sensor technology. Soil fertility sensors, which are primarily based upon electrical and electromagnetic, optoelectronic, and electrochemical principles, are discussed in this article.

Keywords- soil fertility; precision agriculture; on-the-go system;

I. INTRODUCTION

Soil is the topmost layer of the earth's surface and one of Earth's most important natural resources. It is the natural medium in which plants grow. It consists of a mixture of minute particles of disintegrated rocks, minerals, organic matter and bacteria. Proper growth of a plant depends on soil. Apart from physical housing of a plant, soil provides the plants with water supply and essential nutrients for growth. Traditionally, soil fertility refers to the amount of nutrients in the soil, which is sufficient to support plant life. Actually, soil fertility is a congenial condition of soil which is ideal for plantation [1]. Fertile soil has the following properties—

- It is rich in nutrients necessary for basic plant nutrition, including nitrogen, phosphorus and potassium.
- It contains sufficient minerals (trace elements) for plant nutrition, including boron, chlorine, cobalt, copper, iron, manganese, magnesium, molybdenum, sulphur, and zinc.
- It contains soil organic matter that improves soil structure and soil moisture retention.
- Soil pH is in the range 6.0 to 6.8.
- Good soil structure, creating well drained soil.
- Proper texture that suits needs of the crops to be planted.
- A range of microorganisms that support plant growth.
- It often contains large amounts of topsoil.

Fig. 1 shows that soil fertility broadly depends upon four major factors viz. physical property, chemical property, biological property and soil.

There are different sets of fertility requirements for different crops. A typical soil type may be fertile for crop A but the same might not be fertile for crop B. So measurement of soil fertility is extremely important before deciding on any soil

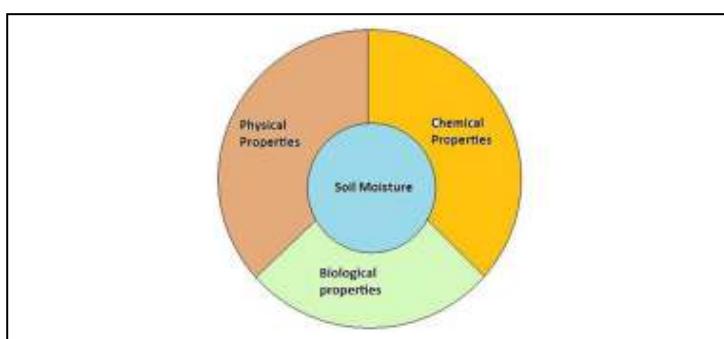


Figure 1. Contributing Factors of Soil Fertility

fertility enhancement activity. Measuring soil fertility has been a perennial issue, which has been addressed differently by people at different places at different time. There was a time when people used to decide the cropping pattern on a particular field just by looking at the colour of the soil. Efforts are made to use technicalities in order to remove the subjectivity. Conventional chemical methods of soil testing are widely practiced in India. These processes are not only time consuming, but also laborious and less accurate. The complete test for macro nutrients, pH level and soil electrical conductivity generally takes 1 day per sample. Sensor technology is one of the modern technologies which are used for the purpose of soil fertility measurement and it aims to overcome the limitations of conventional soil testing. It incorporates accuracy and rapidity in the process which is most desirable at a time when world agriculture scenario is slowly embracing precision agriculture.

II. SENSORS FOR SOIL FERTILITY MEASUREMENT

A *Sensor* is a device, which responds to an input quantity by generating a functionally related output usually in the form of an electrical or optical signal. Likewise in case of soil nutrient sensor, the sensors are supposed to do the same in response to the presence of elements like N, P, K etc. in soil. The challenge is to calibrate and then measure accurately that generally minute change properly according to the level of the nutrients.

There are a large variety of design concepts for the soil fertility sensors. But keeping the feasibility in mind, we will mainly concentrate on the following sensors—

- Electrical and electromagnetic sensors measure electrical resistivity/ conductivity and capacitance affected by composition of the tested soil.
- Optoelectronic sensors use electromagnetic waves to detect the level of energy absorbed/ reflected by soil particles.
- Electrochemical sensors use ion-selective membranes that produce a voltage output in response to the activity of selected ions. (H^+ , Na^+ , K^+ , NO_3^- , $P_2O_5^-$, NH_4^+ etc.)

A. Electrical and Electromagnetic Sensor

The field-scale application of apparent soil electrical conductivity to agriculture has its origin in the measurement of soil salinity but today, this parameter has evolved into a widely accepted means of establishing the spatial variability of several soil physico-chemical properties that influence the electrical conductivity measurement [2]. When current passes through a soil, the conductivity or resistivity of soil is affected by physical and chemical properties of soil. Using proper electrical circuitry, the conductivity or resistivity of the soil is measured. Electrical conductivity is strongly correlated with various soil properties [2].

Four types of electrical conductivity (EC) sensors are available: (i) Electrical Resistivity (ER) sensors, that utilise invasive electrodes (ii) non-invasive Electromagnetic Induction (EMI or EM) sensors (iii) Capacitive Sensor and

(iv) Time Domain Reflectometry (TDR) sensors. Invasive ER and non-invasive EM are the most popular sensors and they have been widely commercialized. The commercial development of a TDR sensor for use on a mobile apparatus is largely limited [3].

The purpose of ER surveys is to determine the resistivity (and thus conductivity) from a given soil volume. A pair of electrode is inserted into the soil. Current is applied at one electrode and the voltage drop (potential difference) across the two electrodes is measured as we can see in Fig. 2. The conductivity is then computed as the ratio of current to potential difference. The system can be stationary or mobile, and paired electrodes can be installed at different depths. Potential differences patterns provide information on the form of subsurface heterogeneities and their electrical properties.

EMI sensor technology uses electromagnetic energy to measure the apparent conductivity of the soil. In Fig. 3, it can be observed that the device is composed of a transmitter and a receiver coil installed usually 1.0 m (3.3 ft) apart in a non-conductive (wooden) bar in the opposite ends of the instrument. The transmitter coil is energized with an alternating current, generating a time-varying magnetic field in the earth. This magnetic field causes current to flow in the soil, and a secondary magnetic field is generated. The ratio of the secondary to the primary magnetic field is proportional to the ground conductivity of the soil [4].

Moisture content of soil can be measured using capacitive sensor. As the dielectric constant of water is greater than soil, the net capacitance of wet soil is greater than dry soil and capacitance of soil can be calibrated for the amount of moisture content in the soil. Depending upon this principle, Whalley et al. designed a sensor for soil moisture measurement and they reported that 84% of sensor variance could be explained through differences in moisture content [5]. Reference [5] mentioned that this type of sensor is capable to measure soil moisture content in the dynamic condition in the field although interferences from soil salinity, temperature and perhaps soil texture affected the moisture measurement. Through suitable instrumentation, the effect of temperature can be nullified.

Sensors based on electromagnetic properties have been most successfully applied to agriculture [5].

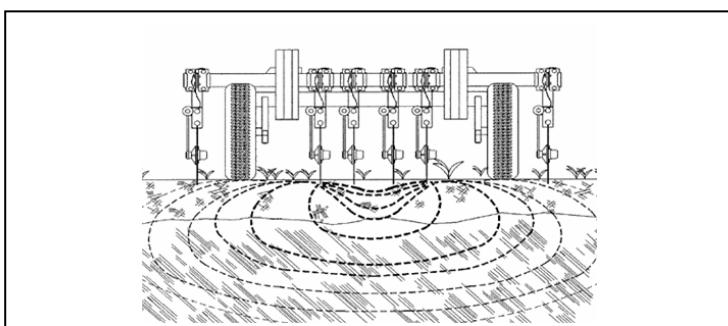


Figure 2. Sensor for Direct Fertility Measurement [4]

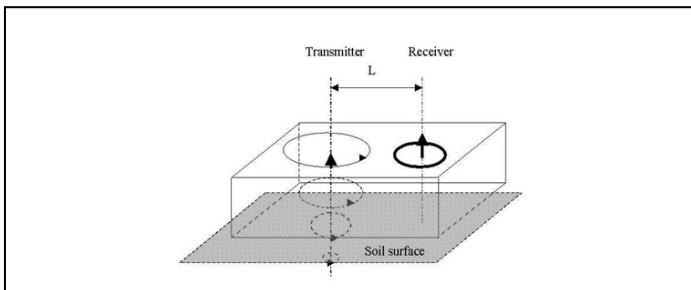


Figure 3. Electromagnetic Sensor for Measuring Apparent Conductivity of Soil [4]

It is evident from the discussion that both conductive and capacitive soil properties are affected by several factors inherent in the soil. It has been seen that soil types (mainly soil texture) influences the output of the commercially available ER sensors. Apparent soil electrical conductivity is influenced by a combination of physico-chemical properties including soluble salts, clay content and mineralogy, soil water content, bulk density, organic matter, and soil temperature. Capacitive sensor can measure the volumetric content of water in soil.

The main advantages of these sensors are their low cost of measurement per hectare, and their non-destructive nature. Both of them can map EC very fast. But EC measurement will reflect conjunctive effect of physical and chemical properties of the soil. The mapping can only be done during dry season to isolate the effect of water content [4].

B. Optoelectronic Sensor

Optoelectronics deals with the interaction of electronic processes with optical processes. Devices where such interaction takes place are called optoelectronic devices. Optoelectronic sensors in general measure the interaction pattern of different materials with light of different wavelengths. The system is composed of a light source (emitter) that emits light in a particular wavelength range, and a sensor (detector) for specific ranges of wavelength.

According to [5], there are numerous soil sensors which work on the basis of absorption, reflectance or transmittance characteristics of light for soil. Determination of the amount of energy reflected from the soil surface in a particular spectral range is the most popular approach. Like its electrical and electromagnetic counterpart, optical measurements are often affected by a combination of soil attributes. However different soil properties affect the response in the different parts of the spectral image. According to [6], soil reflectance is affected by moisture, organic matter, particle size, iron oxides, mineral composition, soluble salts, parent material and other attributes.

Soil colour is the visual portion of the spectra that is reflected from the soil surface. Historically, this has been one of the most obvious parameter used to characterize heterogeneity of soil in a given agricultural field. Aerial and satellite images using remote sensing are very helpful to analyze variation of soil reflectance at the landscape level but crop residue cover and other limitations of remote sensing restricts the application.

Reference [7] reported their earlier effort of a portable, on-the-go spectrophotometer designed to acquire NIR soil reflectance data at a number of narrow-band wavelengths. It successfully predicted soil organic carbon content across a range of soil types and moisture contents. Reference [5] reported a later version of this system which incorporated a spectrophotometer with a digital video camera, EC electrodes and a mechanical load sensor, and was used to study both spatial and temporal variability of SOM and nitrate content. Subsequent researchers targeted to improve the accuracy and repeatability of the sensors of this type.

Ehsani et al. used mid infrared (MIR) spectroscopy (in [8]) to reveal the nitrate ion that was not detected in NIR range. The MIR spectra results were compared to soil nitrate content over a range of nitrate concentration from 400 to 3000 mg/Kg of nitrate nitrogen. Based on this laboratory study, which included two soil types and two sources of mineral nitrogen added as a fertilizer, they concluded that MIR spectroscopy provides a viable technique to rapidly determine soil mineral nitrogen.

Reference [9] reported an optical sensor for analysis of soil nutrients which can measure the amount of ammonia nitrogen ($\text{NH}_4\text{-N}$), nitrate nitrogen ($\text{NO}_3\text{-N}$), available phosphorus (P_2O_5), available iron (Fe), exchangeable iron (Mn) and exchangeable calcium (CaO) in soil sample. Figure 4 shows the configuration of the developed sensor. The light sources consists of three LEDs green (G), red (R), infra red (IR). The wavelength of LED is chosen to fit the absorption peak of the colour developed solution. Different colour developed solutions are prepared from the extracted soil and the resultant solution is put into the transparent cell. The light from plastic optical fibre (POF) is transmitted into the cell and detected by a silicon photodiode. The amount of absorption of light can indicate the amount of a certain nutrient present in the soil if the photodiode output is properly processed through suitable electronic circuitry. This sensor shows high degree of agreement with the theoretical and conventional test method and this can be seen in Fig. 4. This method is also not as expensive as the sensors with spectrophotometer.

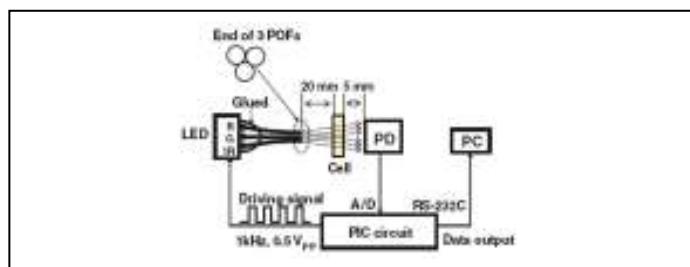


Figure 4. Configuration of Optoelectronic Sensor [9]

(POF- Plastic Optical Fibre, PD- Photodiode, PC- Personal Computer, PIC- Peripheral Interface Controller, A/D- Analog to Digital Converter)

Some of the major advantages of this type of sensors are high speed of analysis, simple sample preparation, non-destructive analysis of the sample, low cost and ease of implementation. Their main limitation is that the results may be sensitive to light changes in exposure; results are largely empirical, therefore, calibrations need to be performed for each medium being sensed.

C. Electrochemical Sensor

Most of the sensors described above have been used to measure some of the physical characteristics of soil. Though some of the optical sensor can measure the amount of some of the macronutrients concerned, their use is largely limited due to their slow response. Direct on-the-go measurement of soil characteristics such as pH or nutrient content has been the object of considerable research. Electrochemical methods have been successfully used to directly evaluate soil chemical properties. This is usually done by either an ion-selective electrode (ISE) (glass or polymer membrane), or an ion-selective field effect transistor (ISFET). In both cases, measured voltage (potential difference) between sensing and reference parts of the system is related to the concentration of specific ions (H^+ , K^+ , NO_3^- etc.). Ion selective electrodes have been historically used by commercial soil laboratories to conduct standard chemical soil tests, and they are widely used to measure soil pH [10].

A common ISE system consists of a membrane that is sensitive to specific ions and a reference electrode (pH electrode is an ISE for the Hydrogen ion.). The milli volt (mV) difference in the potential between the sensitive membrane and the reference is measured and converted to the activity of specific ions in the tested solutions. This activity usually correlates to the concentration of the ions of interest. The design of a combination ion-selective electrode allows both sensitive and reference parts to be assembled in one probe [11]. These systems measure ions of agricultural interest such as Cl^- , NO_3^- , Ca^{2+} , K^+ , Na^+ etc automatically and continuously but require previous sample extraction and pre-treatment. Clearly, this is primarily a laboratory method. Typical arrangement of ISE can be seen in Fig. 5.

According to [10], ISFETs are the result of the integration of two technologies: ISEs and microelectronics technology. The response mechanism of ISFETs is based on their semiconductor nature and on the electrochemical phenomena that occur in the chemically sensitive membrane placed on the gate of the transistor. ISFETs have a gate layer that is sensitive to pH variations. The materials making these layers include SiO_2 , Si_3N_4 and Al_2O_3 . Further progress has been introduced by the deposition of ion recognition membranes on top of the gate area of ISFETs in order to modify the selectivity of the device. Membranes based on poly vinyl chloride (PVC), similar to those developed for ISEs, have been deposited on ISFETs to produce devices sensitive to K^+ , Ca^{2+} , NH_4^+ and Na^+ [10]. A typical scheme of the ISFET mechanism can be observed in Fig. 6.

Current status of development of this sensor is as follows—

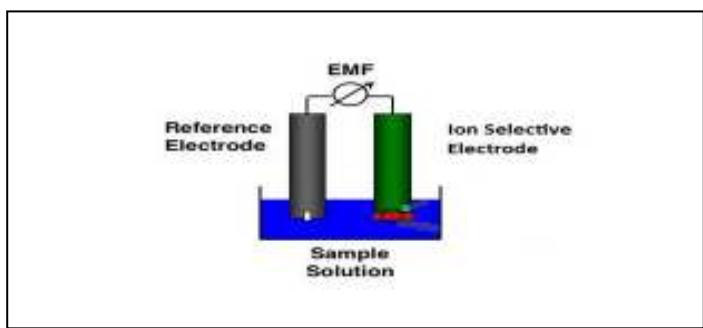


Figure 5. Example of a figure caption. (figure caption)

- Reference [8] reported some very promising development on detecting the presence of phosphate ion using electrochemical method. In order to detect phosphate ion concentration, they have developed an ISFET which uses platinum wire coated with PVC as membrane. The wire is connected to the FET and the drain current is proportional to phosphate concentration.
- Reference [10] developed a technology for evaluating concentration of potassium ion in soil samples. They have achieved some encouraging results which is highly similar to the result of conventional test. But this method is still being tested in laboratory condition [8].
- A successful automated system for on-the-go mapping of soil pH was developed and tested under field conditions [5]. While travelling across the field, a soil sampling mechanism located in a toolbar-mounted shank scooped a sample of soil from a depth of approximately 10 cm and brought it into firm contact with the sensitive membranes of two flat-surface ion-selective electrodes. After stabilization of the electrode output (typically 5–15 s), a new soil sample was obtained and the electrode surfaces were rinsed at the same time. This method was referred to by the authors as direct soil measurement (DSM). Every measurement was geo-referenced using a GPS receiver.

In order to obtain high resolution maps of soil nitrate and potassium levels at the time of pH mapping, intensive research has been going on. The major limitation lies in the fact that

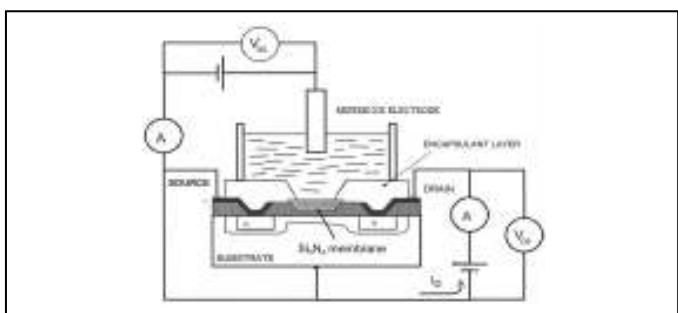


Figure 6. Example of a figure caption. (figure caption)

though it is feasible to use both NO_3^- and K^+ ISE to determine level of nitrate and potassium content in naturally moist soil in laboratory; the same cannot be done without compromising with accuracy in the conditions of field variability. Therefore, the potential for DSM using the ISE appears to decline in the order of $\text{pH} > \text{K}^+ > \text{NO}_3^-$. Additional research is required to develop sensor which will enable on-the-go measurement of the K^+ and NO_3^- ions in soil sample. According to preliminary field trials, integrated, automated, on-the-go mapping of pH, K^+ , and NO_3^- may be used to identify areas of fields with extremely low and high fertility levels and serve as a complimentary data layer for conventional soil sampling programs while significantly reducing the number of laboratory samples required [5].

Though commercial facilities are there for mapping soil pH, same is not available for the targeted ions such as potassium. Extensive research is needed before these sensors' output can be used for prescribing lime and fertilizer input to the land.

III. DISCUSSION

After reviewing some of the literature on most of the major sensor technologies for soil fertility measurement, it is evident that soil fertility sensing is very much an ever developing technology. Till the date, there is not a single sensor which can measure all the important macro nutrients, let alone other important aspects of soil fertility. In order to cater to the needs of precision agriculture, on-the-go soil fertility sensory systems have been developed. On-the-go system actually integrates some of the sensor for sensing different physical and chemical aspects of soil. The choices of sensors in that system are dictated by various factors like expectations of a particular crop from the field, cost, rapidity and required precision and accuracy.

All the sensors have their own merits and demerits. Most of the sensors are discussed above are largely limited in their practical use. Some of the sensors show appreciable accuracy and precision in laboratory testing but often cannot repeat the same in field condition. Those which are already working on field environment need further improvement in level of performance. For example, ISEs or ISFETs are capable of direct measurement of soil properties more accurately than any other method. But unlike previously described measurement concepts, electrochemical sensors require a significant amount of time for the sensing element to reach equilibrium with the measured soil or soil solution. In addition, tedious sample preparation is needed if the principal steps of conventional laboratory chemical tests are replicated in a field sensing system. Sensor prototypes capable of accomplishing this task are relatively complex and still under development [5].

Mainly due to the above two reasons, though various on-the-go soil sensors are under development, only electrical and electromagnetic sensors have been widely used in precision agriculture. Producers prefer sensors that provide direct inputs for existing prescription algorithms. Instead, commercially available sensors provide measurements, such as electrical

resistivity or conductivity that cannot be used directly since the absolute value depends on a number of factors as already mentioned. In contrast, electrical and electromagnetic sensors also give valuable information about soil differences and similarities which make it possible to divide the field into smaller and relatively homogeneous areas referred to as finite management elements (FMEs) or management zones. For example, such FME could be defined according to the various soil types found across a field.

IV. SUMMARY AND CONCLUSION

The sole aim of precision farming has been to manage soil and crops efficiently so that proper input to the soil brings out proper crop output. In order to do that, gathering proper knowledge and information about soil fertility of agricultural land with accuracy and rapidity is of utmost importance. To address this problem, soil fertility sensors become very important. But no sensor can measure all the factors of soil fertility. Physical characteristics of soil like EC, soil texture can be measured by electric and electromagnetic sensors etc. It has been mentioned earlier that measurement of EC gets affected by various inherent factors of soil so further study is needed to address this problem. Measurement of soil biomass can be done by optoelectronic sensor and radiometric sensor and research are going on to use improved GPR instrumentation for bringing more rapidity and accuracy to the process. As far as chemical properties of soil are concerned, optical techniques appear to have broad applicability across all ions of interest and offer intriguing potential to create an integrated sensing system in which shared optical components could be applied. Similarly, ISEs/ISFETs offer promise, but different membranes, extraction solutions, and methods complicate a multi-target system. Future work on these two paths is likely to drive the most immediate benefits for practical on-the-go analysis and some combination of the two may represent the shortest path to routine N-P-K field measurements. Beyond the implications of this analysis on specific sensors, it is clear that there is generally a need for more consistent presentation of sensor performance to more accurately document progress towards viable on-the-go units and direct research efforts towards the truly limiting aspects of any given sensing technology. Fully integrated on-the-go instruments require more research and considerable progress in that direction is required to serve the need of precision agriculture.

ACKNOWLEDGMENT

We are thankful to Dr. S. S. Prabhudesai, Soil Scientist, Regional Agriculture Research Station, Karjat, Maharashtra for helping us in understanding some of the key concepts of soil science. We would also like to thank Prof. A. W. Date, Professor Emeritus, Dept. of Mechanical Engineering, IIT Bombay for his support and valuable inputs towards this work.

REFERENCES

- [1] L. M. Thompson and F. R. Troeh, "Soil" in Soil and soil Fertility, 4 th Ed., New Delhi: McGraw-Hill, 1975, pp. 3-10.

2nd Annual International Conference IEMCON 2012

- [2] D. L. Corwin and S. M. Lesch , “Apparent soil electrical conductivity measurements in agriculture”, Computers and Electronics in Agriculture, vol 46, pp. 11-43, 2005.
- [3] B. Tisseyre and J. Taylor, “An overview of methodologies and technologies for implementing precision Agriculture in Viticulture”, pp. 45-54, 2006.
- [4] R. Sanyal, “Sensors for soil fertility management,” unpublished.
- [5] V. I. Adamchuk, J. W. Hummel, M. T. Morgan, and S. K. Upadhyaya, “On-the-go Soil Sensors for precision Agriculture”, Computers and Electronics in Agriculture, vol 44, pp. 71-91, 2004.
- [6] M. F. Baumgardner, L. F. Silva, L. L. Beihl, and E. R. Stoner, “Reflectance properties of soils”, Advances in agronomy, vol 38, pp.1-44, 1985.
- [7] J. W. Hummel, K. A. Sudduth, and S. E. Hollinger, “ Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor”, Computers and Electronics in Agriculture, vol 32, pp. 149-165, 2001.
- [8] J. V. Sinfield, D. Fagerman, and O. Colic, “Evaluation of sensing technologies for on- the-go detection of macro-nutrients in cultivated soils”, Computers and Electronics in Agriculture, vol 70, pp. 1-18 (2010).
- [9] M. Yokota, T. Okada, and I. Yamaguchi, “An optical sensor for analysis of soil nutrients by using LED light sources”, Measurement Science and Technology, vol 18, pp. 2197, 2007.
- [10] J. Artigas, A. Beltran, and C. Jiménez, “Application of ion sensitive field effect transistor based sensors to soil analysis”, Computers and Electronics in Agriculture, vol 31, pp. 281- 293, 2001.
- [11] V. I. Adamchuk, E. D. Lund, and B. Sethuramasamyraja, “Direct measurement of soil chemical properties on-the-go using ion-selective electrodes”, Computers and Electronics in Agriculture, vol 48, pp. 272-294, 2005.

Computational Optimization of Speed in an Unplanned Lane Traffic

Prasun Ghosal

Department of IT

Bengal Engg and Sc University, Shibpur Bengal Engg and Sc University, Shibpur
Howrah, WB, India
Email: prasun@ieee.org

Arijit Chakraborty

Purabi Das School of IT

Howrah, WB, India
Email: arijitchakraborty.besu@gmail.com

Sabyasachee Banerjee

Department of CSE

Heritage Institute of Technology
Kolkata, WB, India
Email: sabyasachee.banerjee@gmail.com

Abstract—Speed Optimization in an Unplanned Traffic (SOUT) is a very promising research problem. Searching for an efficient optimization algorithm to increase the degree of speed optimization and thereby increasing the traffic flow in an unplanned zone is a widely concerning issue. However, there has been a limited research effort on the optimization of the lane usage with speed optimization. This paper presents a novel SOUT technique to solve the problem optimally using the knowledge from analysis of speeds of vehicles, which, in turn will act as a guide for design of lanes optimally to provide better optimized traffic. The accident factors adjust the base model estimates for individual geometric design element dimensions and for traffic control features. Knowledge based analysis technique is applied to the proposed design and speed optimization plan. The experimental results are quite encouraging.

Index Terms—SOUT, Speed optimization, Unplanned traffic, Traffic engineering, Optimization Algorithm

I. INTRODUCTION

The challenges of speed optimization in an unplanned traffic design system are to move traffic safely and efficiently. Although highways and motor vehicles are designed to operate safely at speed. The purpose of our investigation was to create predictive models for optimizing the speed of the different types of vehicles in an unplanned zone, based on infrastructural design and traffic intensity. In this paper, the results for all transition points and lane transition of vehicles for speed optimization is discussed.

The first investigation was the identification of different zones on the lane. Three major zones on the lane were defined:

- Transition points.
- Entry zones and
- Exit zones.

A. Drawbacks of existing solutions

Many traditional speed optimizing algorithms for lanes were proposed earlier to optimize deterministic problems. But they cannot tackle the inherent randomness in the traffic systems. Therefore, to handle with such random realistic situation and generate some efficient solution, good computational models of the same problem as well as good heuristics are required. This requirement worked as a major driving factor towards our effort.

The solutions suggested by different traditional Non accidental algorithms for lanes cannot answer the questions like

What is the degree of robustness that the optimal solution will achieve optimum speed utilization of the vehicle? or What is the probability of failure that the optimal solution will fail if something unexpected occurs?

Hence, the above problems worked as motivating factors towards addressing the problem from a different angle. The purpose of this article is to address these three issues by introducing the approximation technique to be discussed throughout the paper for the optimal SOUT design.

The paper is primarily divided into two parts.

In the first part, we analyse the problem and determine the major factors for the determination of the optimum number of lanes and transition points.

In the second part, we conclude the possible lanes with optimum transits between the lanes for speed optimization of the vehicle.

The overall organization of the paper is given as follows. Important previous works are discussed in the following section. Section III discusses about the details of the proposed work with a very clear analysis of the proposed algorithm. Section IV presents the experimental results with simulated graphical analysis of the computational model. Section V concludes the paper with possible future directions of work.

II. PREVIOUS WORKS

The paper proposed by Jake Kononov, Barbara Bailey, and Bryan K. Allery [1], first explores the relationship between safety and congestion and then examines the relationship between safety and the number of lanes on urban free-ways. The relationship between safety and congestion on urban free-ways was explored with the use of safety performance functions [SPF] calibrated for multi-lane free-ways in Colorado, California, Texas. The Focus of most SPF modelling efforts to date has been on the statistical technique and the underlying probability distributions. The modelling process was informed by the consideration of the traffic operations parameters described by the Highway Capacity Manual [1].

In 2007, H Ludvigsen et al., has published Differentiated speed limits allowing higher speed at certain road sections whilst maintaining the safety standards are presently being applied in Denmark [2]. The typical odds that higher speed

limits will increase the number of accidents must thus be beaten by the project [2].

In another important work, C.J. Messer et al. [3] presented a new critical lane analysis as a guide for designing signalized intersections to serve rush-hour traffic demands. Physical design and signalization alternatives are identified, and methods for evaluation are provided. The procedures used to convert traffic volume data for the design year into equivalent turning movement volumes are described, and all volumes are then converted into equivalent through auto-mobile volumes. The critical lane analysis technique is applied to the proposed design and signalization plan. The resulting sum of the critical lane volumes is then checked against established maximum values for each level of service (A, B, C, D, E) to determine the acceptability of the design. In this work, the authors have provided guidelines, a sample problem, and operation performance characteristics to assist the engineer in determining satisfactory design alternatives for an intersection [3].

III. PROPOSED WORK

Speed Optimization in an Unplanned Traffic (SOUT) solutions finding methods in other literature are a family of optimization algorithms which incorporate level of traffic services in the algorithms. There are two major issues, in the first part, we have analysed the major issues residing in the latest practice of the accidental lane; and, in the last part, we have discussed the possible applications of this new technique and new algorithm.

A. Proposed Algorithm

The algorithm proposed for the solution of the problem of speed optimization in unplanned traffic is described below. The algorithm is designed under certain assumptions as described below.

Assumption

During the execution of the algorithm it is assumed that there will be no change in the current speed of the vehicle whenever the current speed of the vehicle is entered once.

B. Description of the Proposed Algorithm

The primary sections of the proposed algorithm and their major functionalities are described below.

- Step 1 is taking input from sensors, like the current speed of the vehicle, arrival time etc., and, counting the number vehicles the user has entered.
- Step 2 is categorizing the vehicles depending on their current speed.
- Step 3 is checking total how many numbers of lanes will be required for our sample data in an unplanned zone, and, which vehicle is moving in which lane.
- Step 4 is checking total number of transitions i.e. at which point of the lane and from which lane to where the transition will occur. Now these are the outputs of our simulation.

Algorithm_SOUT()

Input : Name of vehicle, maximum speed, arrival time.
Output: Type of vehicle, Number of lanes required,
Number of transitions.

```

Step 1.1: Set count = 1; /*Used to count the number of
vehicles.*/
Step 1.2: get_input(); /*Enter name of vehicle, current
speed, arrival time and store it into a record.*/
Step 1.3: Continue Step 1.1 until sensor stops to give
feedback and update count = count + 1
Step 2: for 1 ≤ i ≤ count for each vehicle
If current speed of the vehicle ( $V_i$ )  $0 < V_i \leq 10$  then
categorize vehicle (i) as type A
If  $11 \leq V_i \leq 30$  then categorize vehicle (i) as type B
If  $31 \leq V_i \leq 45$  then categorize vehicle (i) as type C
If  $46 \leq V_i \leq 50$  then categorize vehicle (i) as type D
If  $51 \leq V_i \leq 100$  then categorize vehicle (i) as type E
Step 3: Set counter: count1 := 1;
Set lane of first vehicle = 1;
for  $2 \leq i \leq count$  for each vehicle
for  $1 \leq j \leq count1$ 
Compare the type of (i)th vehicle with type of (j)th
vehicles present in the lane
if different update count1 = count1 + 1 and
laneof(i)thvehicle = count1;
else laneof(i)thvehicle = j;
end of loop;
end of loop;
Step 4: Set counter: count2 = count1;
for  $1 \leq i \leq count - 1$  for each vehicle
for  $2 \leq j \leq count$  for each vehicle
if (i)thvehicletype = (j)thvehicletype and  $V_i < V_j$  and
(i)thvehiclearrivaltime ≤ (j)thvehiclearrivaltime
set t =
(j)thvehiclearrivaltime - (i)thvehiclearrivaltime;
set t1 = 0;
begin loop
set t1 = t1 + 1;
set d =  $V_i \times (t + t_1)$ ;
set d1 =  $V_j \times t_1$ ;
if  $d_1 \leq d$  set count2 = count2 + 1;
if (j)thvehiclelane = 1 then transition will be to
2 - lane;
if (j)thvehiclelane = count1 then transition will be to
count1 - lane;
else transition will be to (j)thvehiclelane - 1 or
(j)thvehiclelane + 1;
end loop;
end loop;
end loop;
Step 5: Return Numberoflanesrequired =
count1, Numberofttransitionsrequired = count2;
Step 6: End

```

Algorithm 1: Proposed algorithm for speed optimization in an unplanned traffic (SOUT)

C. Analysis of the Proposed Algorithm

The salient points of the proposed algorithm are summarized below.

- The above algorithm is implemented on an open unplanned area.
- The objective will follow linear queue as long as speed/value/cost of preceding is greater than the immediate next.
- Transition/Cross over are used and they again follow appropriate data structure in order to maintain the preceding step rule.
- Here we assume the lanes are narrow enough to limit the bidirectional approach.
- Here we maintain optimize speed for each lane.
- Here we also maintain the transition points if speed/value/cost of a vehicle is found unable to maintain the normal movement and transition in all the calculated lanes.
- Transition points are recorded with their position and number and it follows appropriate data structure in order to maintain the record.

IV. EXPERIMENTAL RESULTS

Our work aimed to design /SOUT/ in an open unplanned area, so as to increase traffic movement in rush hours and to optimize the speed of the vehicles using the concept of transition points between adjacent Lanes.

In Figure 1, three vertical lanes are shown those are unidirectional, and $A = a_1, a_2, \dots, a_n, B = b_1, b_2, \dots, b_n, C = c_1, c_2, \dots, c_n$, with the property of the three lanes. I, II, III are the transition points through which vehicles can overtake its preceding vehicle with lesser speed and then immediately moves to its original lane. I is from lane A to B or B to A and II, III are from B to C or C to B. Here we assume that car speed in each and every lane is greater than 0 kmph. If the speed of any car is less than or equal to 0 kmph then we assume that there may be problem. The random distribution of entities in an open area to lanes is to be taken care of as far as possible.

A. Simulated Graphical Analysis

After the implementation of the proposed algorithm the whole scheme is simulated and fro the simulations the graphical results are generated which are shown in Figure 2 and Figure 3.

V. CONCLUSION AND FUTURE WORK

In this work, we introduce a novel design of a Speed Optimization Technique in an Unplanned Traffic. Our proposed design and speed optimization plan, SOUT, can robustly manage and can operate over any unplanned zone. As the Speed optimization with unplanned zone is widely concerning issue in rural development, therefore, our proposed method will optimize the lane requirements and speed of the vehicles, which indirectly helps also in energy saving and more economical growth of a country.

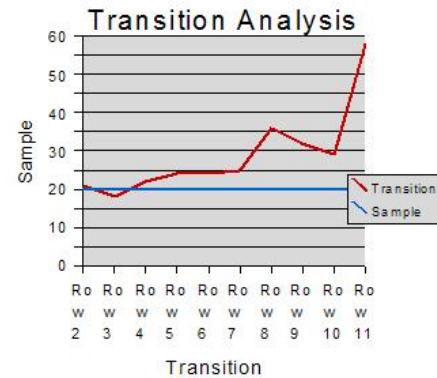


Fig. 2. Sample Size Vs Number of Transition points.

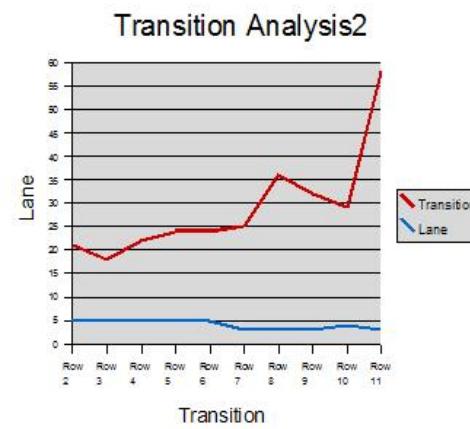


Fig. 3. Number of lanes Vs Number of Transition points.

The main limitation of the approach is that though in this paper lane requirement is optimized but lane usage is not optimized. For example at a certain time instance there may be a chance that only same type of vehicles are arriving and populating its designated lane but other lanes are not so populated, so lane usage optimization is required. Our future effort will certainly be on that direction. Here in this work, we have tried to optimize the speed of the vehicles and lanes and transitions between lanes in a secured non collision manner.

REFERENCES

- [1] Jake Kononov, Barbara Bailey, and Bryan K. Allery, *The Relationship Between Safety and Congestion*, Journal of the Transportation Research Board, No. 2083.
- [2] *Differentiated speed limits*, European Transport Conference Differentiated speed limits, 2007.
- [3] C.J. Messer and D.B. Fambro, *Critical Lane Analysis for Intersection Design*, Transportation Research Record No. 644; 1977, pp. 26-35.
- [4] Prasun Ghosal, Arijit Chakraborty, Amitava Das, Tai-Hoon Kim, Debnath Bhattacharyya, *Design of Non-accidental Lane*, In Advances in Computational Intelligence, Man-Machine Systems and Cybernetics, pp. 188-192, WSEAS Press, 2010.

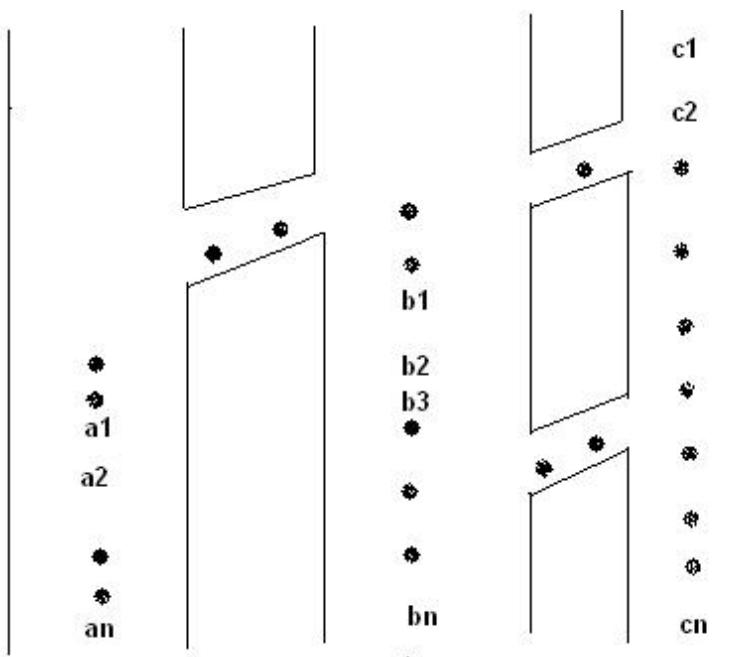


Fig. 1. Vertical lanes that are unidirectional and with the property of the three lane with transition points.

PREMASA MODEL: AN INTEGRATED APPROACH TO CUSTOMER RELATIONSHIP MANAGEMENT

Debabrata Das

Department of Business Administration.
Bengal School of Technology & Management.
Chinsurah, West Bengal, India.
E-mail: deb.brata.das@gmail.com

Abstract— CRM is a process or methodology which is used to learn more about customers' needs and behaviors in order to develop stronger and reliable relationships with them. A problem is that CRM means different things to different people. The meaning of CRM ranges from direct e-mails, mass customization of products or services to complicated technical jargon related to terms like OLAP (on-line analytical processing) and CICs (customer interaction centers). Though CRM has many technical components, but representing CRM in primarily technological terms is a mistake. With the ever increasing importance of service sector in global economy, the more useful way to think about CRM is as a business process that will help bring together lots of pieces of information about customers, sales, marketing effectiveness, responsiveness and market trends. This paper represents an empirical investigation involving internet service users in and around Kolkata whose opinions have been collected on some pre defined variables & multivariate statistical analysis has been used to identify key factors on which the customer marketer relationship depends and then it has been attempted to integrate the aspect of relationship marketing with consumer behaviour by incorporating the famous concept of cognitive dissonance to develop a CRM Model which should work not only as a theoretical framework but also as a base for future empirical research.

Keywords- *Customer Relationship Management, Relationship Marketing, Cognitive Dissonance, Trust, Satisfaction, Loyalty & Commitment, PREMASA Model.*

I. INTRODUCTION

Customer Relationship Management (CRM) is a business strategy to identify, acquire and retain customers who are the business' greatest assets. CRM helps to get the value out of customers by highly integrating sales, marketing and support efforts. By providing the means to manage and coordinate customer interactions, CRM process helps companies to maximize the value of every customer interaction and in turn drive improved corporate performance. With the rapid growth of service sector and more stress on profitability the research on CRM has increased manifold over the past few years and more and more companies in service sector are stressing on the need of relationship marketing to build a long term permanent

relations with the customers. The concept of cognitive dissonance has been reintroduced by the researchers and academicians in the field of relationship marketing to analyze & understand the inner facets of consumer behaviour. The future research on CRM should be oriented towards integrating the concept of relationship marketing with customer relationship management with the help of cognitive dissonance of the consumers and other aspects of consumer behaviour.

II. LITERATURE REVIEW

Research on CRM has confined itself on service quality enhancement and betterment of post purchase customer service etc but the effect of consumer behaviour on CRM has been ignored. Major research on CRM has only a single purpose and that is to provide a managerially useful, end-to-end view of the CRM process from a management perspective [1]. In other words, these models focus what the managers need to know about their customers and how that information should be used to develop a complete CRM process [2]. All these models are excellent tools for practicing marketing managers but it lacks the framework on the basis of which marketing analysts and researchers can further improve the CRM process because no such serious effort has been made to explore the relation between consumer behaviour and CRM and also the effect of consumer behaviour on CRM.

Reference [3] defines cognitive dissonance as the mental conflict that people experience when they are presented with evidence that their beliefs or assumptions are wrong. According to cognitive dissonance theory [4], there is a tendency for individuals to seek consistency among their beliefs and opinions. When there is an inconsistency between attitudes or behaviors (dissonance), something must change to eliminate the dissonance. Dissonance can be reduced by reducing the importance of the dissonant beliefs, or by adding more consonant beliefs that outweigh the dissonant beliefs, or by changing the dissonant beliefs so that they are no longer inconsistent. The theory of cognitive dissonance is highly relevant in explaining consumer behaviour as almost all major purchases result in cognitive dissonance or discomfort caused by post purchase conflict. After the purchase consumers are

satisfied with the benefits of the chosen brand and are glad to avoid the drawbacks of the brands not bought. At the same time consumers feel uneasy about acquiring the drawbacks of the chosen brand and about losing the benefits of the brands not purchased. Thus the consumers feel at least some post purchase dissonance for every purchase. To counter such dissonance, the marketer's after sale communications should provide evidence and support to help consumers feel good about their brand choices [5].

Reference [6] introduced the relationship marketing concept in the literature and the concept has since become very popular over the years. Relationship marketing can be defined as an umbrella concept which stresses the need to see exchanges from a long-term perspective rather than short-term [7]. According to reference [8], in recent times the trend to incorporate the concept of relationship marketing in formulating models and theories has increased among the academicians as well as the practicing managers and consultants. The fundamental assumption in relationship marketing is that retaining existing customers is considerably less expensive than acquiring new ones. Reference [9] claimed that only a 5% improvement in customer retention can lead to an increase in profitability between 25% and 85%, depending upon industry sector. The findings of Reference [9] regarding the economic value of loyal customers were certainly an important factor in this development. Research has also shown that growth in sales results from increased expenditure over time, positive word of mouth through loyal customers and the willingness of loyal customers to pay a price premium. As there is a lot of scope for greater research in this area, the importance of relationship marketing from different perspectives should be analyzed and more focus should be given to the relationship development process. Theoretical models along with empirical studies are the need of the hour in order to inject renewed energy into this subject [10].

III. RESEARCH OBJECTIVE

The primary objective of this study is to identify underlying dimensions or factors that explain customers' perceptions or attitude or simply what customers think about their ISPs or what makes them to stick to their existing ISPs or what can force them to switch over to some other ISPs or the relationship between the customers and the service providers. The identifiable variables that decide the nature & future of relationship between the consumers & the service providers are of great importance. The secondary objective is to use the identified key variables as critical success factors in developing a CRM Model incorporating the existing theories and concepts about CRM & Relationship Marketing available in the literature.

IV. RESEARCH METHODOLOGY

A. Research Design

As the primary objective of the study is to identify underlying dimensions or factors that decide the relationship between the customers and the service providers, primary data are of the main importance. A set of seven statements indicating different situations and conditions of various aspects of internet usage and ISP is formed on the basis of the

judgment of the author and all these statements are represented as seven unique variables for statistical analysis. As the objective is to identify underlying dimensions or factors that explain the correlations among the set of variables, Factor Analysis technique, which is an interdependence technique in which an entire set of interdependent relationships is examined, is used. As it is important that in case of Factor Analysis, the variables be appropriately measured on an interval or ratio scale, a 7 point Likert type balanced interval scale(1= strongly disagree, 7 = strongly agree) is used. Analyzing the result of the Factor Analysis, the extracted factors are identified and used to develop theoretical framework and conceptual model with the help of existing literature.

B. Methods of data collection

Data were collected in personal interviews with all respondents either in their work places or any nearby places near their workplace or at their homes. A total of one hundred respondents have been interviewed. Great care was taken in the fieldwork of this study in order to ensure a high response rate and a minimum amount of disruption to the respondents who participated. Fieldwork was completed thoroughly yet quickly, in just eight weeks.

C. Questionnaire Design & Interviewing

As this study is exploratory in nature, a combination of both open ended or unstructured & close ended or structured questionnaires are developed. At the first part mostly open ended questions with a very few close ended questions are used to collect general demographic information about the participants & their usage internet pattern, frequency & a choice of internet service providers. The second part of questionnaires has all structured questions. A seven point scale (1= strongly disagree, 7 = strongly agree) has been used & the respondents are asked to indicate their degree of agreement with seven different statements in terms of the 7 point scale. The following seven different statements represent seven general situations related to internet service & the internet service provider:

- V1- Though the service is not appropriate for me in terms of my need & preference; I shall stick to the existing ISP.
- V2 → If I get a new offer from my ISP on some values added services I believe these will be of good quality & I shall sincerely consider this proposal in future.
- V3→ If there is no change in prices I shall prefer my existing ISP over better quality ISP.
- V4→ Slight change in prices doesn't have much effect on my preference of ISP.
- V5→ The after sales services & customer care services are of much importance to me & I may switch to some other ISP over these issues.
- V6→ Massive change in prices may have some effect on my preference of ISP.
- V7→ The service related queries answered by the ISP personnel are true as per my knowledge & belief.

D. The pilot study

The author conducted pilot interviews with twenty five persons who are using different internet connections ranging from modem based broadband internet to wireless internet through USB stick in and around Kolkata. During these interviews, the questionnaires were tested and a personal approach, aimed at building a relationship of trust with the persons, was adopted. After these interviews minor changes were made to the questionnaire and plans were finalized for the actual survey to ensure an optimum response rate.

E. Actual Data collection

The author himself completed the interviews. The author first spoke to the persons known to him through personal connection & informed them of the purpose of the study & requested them to participate. Those in the snowball sample were contacted and asked to participate and their responses were collected.

V. ANALYSIS & INTERPRETATION

A. Examining the validity & applicability of Factor Analysis on the collected data

As the analytical process of Factor Analysis is based on a matrix of correlations between the variables, significance of the process depends on proper examination of this matrix [11]. Formal statistics like Bartlett's test of sphericity, Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and examination of the Anti Image Correlation matrix (AIC) can be used to examine whether the factor model is appropriate and can be used for further investigation or not. Bartlett's test of sphericity tests the null hypothesis that the variables are uncorrelated in the population or the population matrix is an identity matrix. In an identity matrix, all the diagonal terms are 1 and all off diagonal terms are 0. The test statistic for sphericity is based on a chi square (χ^2) transformation of the determinant of the correlation matrix. A large value of the test statistic will favour the rejection of the null hypothesis. If this hypothesis cannot be rejected then the appropriateness of the factor analysis should be questioned [12]. Another useful statistics is the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy. This index compares the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. Small values of KMO statistic indicate that the correlations between pairs of variables cannot be explained by other variables and that factor analysis may not be appropriate [13]. Generally a value larger than 0.5 is desirable which is called Kaiser's "middling" criteria. The Anti Image Correlation Matrix (AIC) represents the measures of sampling adequacy for each individual item on the diagonal of the anti image correlation matrix.

TABLE I. RESULTS OF BARTLETT'S TEST OF SPHERICITY & KMO STATISTIC

| Results Of Bartlett's test of sphericity | | Results Of Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy. |
|--|---------|---|
| Approx. Chi-Square(χ^2) | 235.802 | 0.641 |
| Degrees of Freedom(df) | 21 | |
| Significance (Sig.) | 0.000 | |

The negatives of the partial correlations between pairs of items are presented on the off diagonal. To be the correlation matrix factorable, the measures of sampling adequacy values on the diagonal of the anti image correlations matrix should be large and the values of the negatives of the partial correlations should be small [14].

As seen from Table I, the approximate chi square statistic with a certain degrees of freedom is significant at the 0.05 level as well as the value of KMO statistic is also large (>0.5).

TABLE II. THE ANTI IMAGE CORRELATION (AIC) MATRICES

| The Anti Image Correlation(AIC) Matrices | | | | | | | |
|--|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | V1 | V2 | V3 | V4 | V5 | V6 | V7 |
| V1 | 0.409 | 0.031 | -0.147 | 0.219 | 0.410 | -0.109 | -0.131 |
| V2 | 0.031 | 0.430 | 0.058 | 0.037 | 0.021 | -0.051 | -0.175 |
| V3 | -0.147 | 0.058 | 0.831 | -0.290 | -0.224 | -0.183 | -0.081 |
| V4 | 0.219 | 0.037 | -0.290 | 0.629 | 0.133 | -0.571 | -0.062 |
| V5 | 0.310 | 0.021 | -0.224 | 0.133 | 0.470 | -0.059 | -0.103 |
| V6 | -0.109 | -0.051 | -0.183 | -0.571 | -0.059 | 0.655 | 0.069 |
| V7 | -0.131 | -0.175 | -0.081 | -0.062 | -0.103 | 0.069 | 0.458 |

As seen from Table II, the measures of sampling adequacy values on the diagonal of the anti image correlations matrix are large and the values of the negatives of the partial correlations are small. So, it can be concluded that:

- The null hypothesis that the population correlation matrix is an identity matrix is rejected by Bartlett's test of sphericity.
- As the KMO statistic suggests that the sample size relative to the number of items in the scale is sufficient, factor analysis is an appropriate technique for analyzing the correlation matrix.
- The measures of sampling adequacy (MSA) statistics indicate that the correlations among the individual items are strong enough to suggest that the correlation matrix is factorable.
- So the factor analysis method can be applied to identify the underlying factors that are responsible in deciding the consumer marketer relationship in a service marketing scenario.

B. Determining the number of factors

In factor analysis, what number of factors should be extracted can be determined in three ways : by analyzing the Eigen values , by observing the Scree Plot and by considering the total variance explained by each factor. As an Eigen value represents the amount of variance associated with the factor, only factors with Eigen values greater than 1.0 or the variance greater than 1.0 should be retained.

A Scree plot is a plot of the Eigenvalues against the number of factors in order of extraction. The shape of the plot is used to determine the number of factors. Typically the plot has a distinct break between the steep slopes of factors. Experimental evidence indicates that the point at which the scree begins denotes the true number of factors [15].

TABLE III. THE INITIAL EIGEN VALUES OF ALL SEVEN COMPONENTS

| INITIAL EIGENVALUES | | | | | | | |
|-----------------------------------|-------------|-------------|-------------|------|------|------|------|
| Components | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Initial Eigen Values ^a | 2.58 | 1.32 | 1.17 | 0.87 | 0.61 | 0.34 | 0.12 |

a. Approximated up to two decimal points.

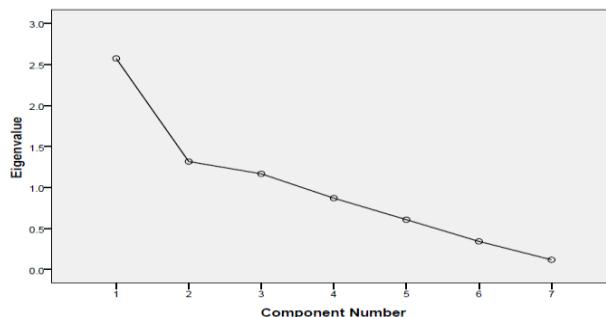


Figure 1. A Scree Plot depicting the variation of Eigen values across components

TABLE IV. TABLE TOTAL VARIANCE EXPLAINED BY THE FACTOR ANALYSIS

| Total Variance Explained | | |
|--------------------------|---------------|--------------|
| Component | % of Variance | Cumulative % |
| 1 | 36.782 | 36.782 |
| 2 | 18.821 | 55.602 |
| 3 | 16.679 | 72.281 |

From the Table III, it is clear that first three components have an Eigen value greater than 1. From Fig.1, it can be seen that the scree occurs at third component. Reference[14] recommends that the factors extracted should account for at least 60 percent of the variance and from Table IV, it is clear that the first three components explain 72% variance; so the result is satisfactory. So, from the factor analysis, three factors

have been extracted which can explain the variances in the population regarding the respondents' perception about internet service providers (ISPs).

C. Interpretation of Factors

Interpretation of factors is facilitated by identifying the variables that have large loadings on the same factor and that factor can then be interpreted in terms of the variables that load high on it and this can be done by analyzing the rotated component matrices.

The second method in interpretation is to plot the variables using the factor loadings as coordinates. Variables at the end of an axis are those that have high loadings on only that factor, and hence describe the factor. Variables near the origin have small loadings on both the factors. Variables that are not near any of the axes are related to both the factors [16].

TABLE V. THE ROTATED COMPONENT MATRIX

| Rotated Component Matrix | | | |
|--------------------------|--------------|---------------|--------------|
| | Component | | |
| | 1 | 2 | 3 |
| V1 | -0.033 | -0.837 | 0.111 |
| V2 | -0.077 | 0.029 | 0.737 |
| V3 | 0.851 | 0.083 | 0.053 |
| V4 | 0.935 | 0.078 | -0.009 |
| V5 | 0.110 | 0.818 | 0.108 |
| V6 | 0.933 | 0.041 | -0.005 |
| V7 | 0.110 | -0.035 | 0.778 |

In the rotated component matrix of Table V, factor I has high coefficients for variables V3, V4 and V6. Therefore, this factor may be labeled as customers' loyalty & commitment towards their ISPs factor. Similarly, factor II has high coefficients for variable V5 and a negative coefficient for V1. A negative coefficient for a positive variable V1 leads to a negative interpretation that the consumers will not stick to their existing ISPs if they don't get appropriate services from them i.e. appropriate service in terms of consumers' preferences will only make them satisfied otherwise the unsatisfied consumers will switch to other ISPs. So, Factor II may be labeled as consumers' satisfaction with their existing ISPs factor. Lastly, factor III has high coefficient for variables V2 and V7 and that's why Factor III may be labeled as consumers' trust towards their ISPs factor.

A plot of the factor loadings confirms this interpretation (Fig.2, Fig.3, and Fig.4). From the Fig.2, it can be seen that the variables V3, V4 and V6 are at the ends of the horizontal axis. (Factor I), with variables V1 and V5 are at the end of the vertical axis but in opposite direction (Factor II) and variables V2 and V7 are situated near the origin and so neither factor I nor factor II is associated with these variables.

From the Fig.3 and Fig.4, it can be seen that the variables V2 and V7 are at the end of the vertical axis. (Factor III). It can be summarized that the consumers think that there are two major reasons behind their choice of ISPs: their trust for the

ISP and their satisfaction for the service of the ISP and these two factors contribute to build their loyalty & commitment towards their ISPs.

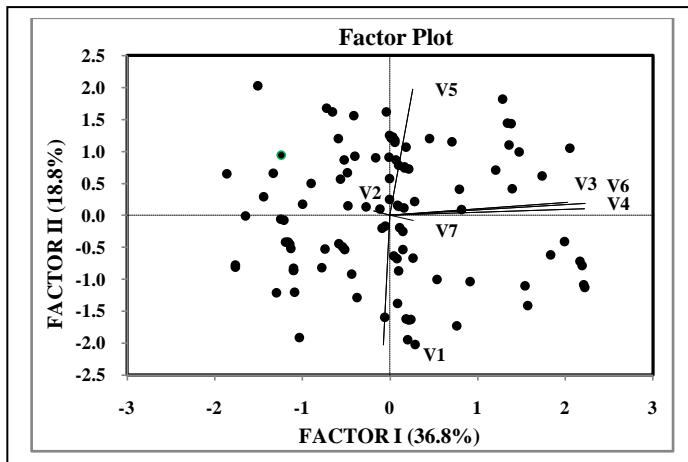


Figure 2. The Factor Plot between Factor I and Factor II.

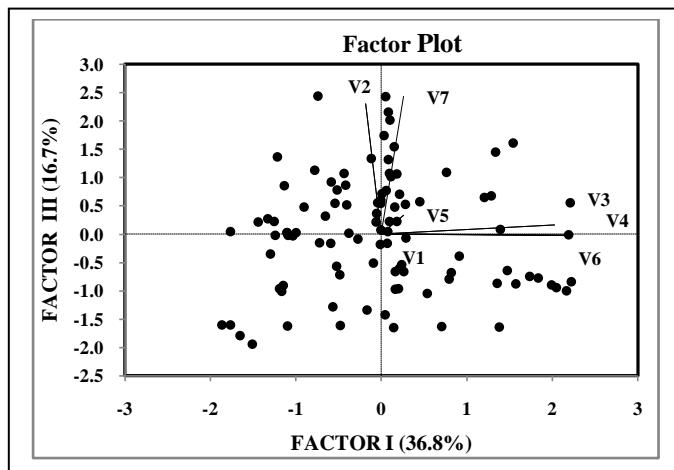


Figure 3. Example The Factor Plot between Factor I and Factor III.

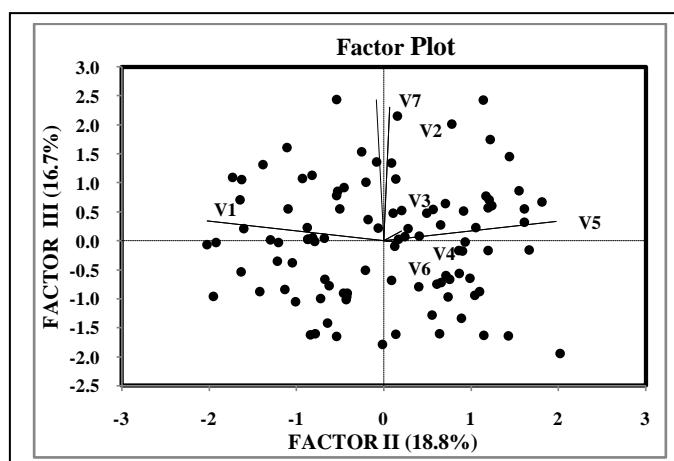


Figure 4. Example The Factor Plot between Factor II and Factor III.

VI. RECOMMENDATION AND SUGGESTION

A. Representation of Trust, Satisfaction, Loyalty & Commitment as Critical Success Factors in CRM program

From the results of the factor analysis it is clear that the three key factors in building and maintaining relationship between consumers & service providers are trust, satisfaction, loyalty and commitment. So, CRM should be used as a business approach that combines people, processes and technology to maximize the relation of an organization with all types of customers[17]. Concepts such as market orientation ([18],[19]), market focus [20] and market-based learning [21] should be revisited because these concepts emphasize the establishment of good information processes and capabilities within the firm to understand the needs and wants of customers, thus making firms more efficient and effective in managing customer relationships. In addition, the evolutionary journey from product, or brand management to customer management [22] and from product portfolio management to customer portfolio management [23] should continue so that the transitions, which are due in part to work in the area of brand equity, can help to recognize that equity resides in the minds of consumers [24] and thus can shift the focus of attention from brands and products to customers. CRM should help in understanding the customers better so that according to the needs of the customers, the organization can effectively customize their products and services in order to build the trust of the customers and also to increase customers' loyalty and satisfaction. Setting aside the traditional concept, CRM can be represented from consumer psychology point of view and thus the emerging concepts of relationship marketing can be integrated with the concepts of CRM so that CRM can be viewed as a process which has only three functions:-

- BUILDING TRUST
- CREATING SATISFACTION
- ENSURING LOYALTY AND COMMITMENT

In this new representation it is evident that building customer trust, creating customer satisfaction and thus ensuring customer loyalty and commitment are the Critical Success Factors (CSF) in any CRM program and that's why in this representation all other functions of CRM have been ignored. In this representation CRM works in three distinct time stage in the following manner:

- STAGE I Builds trust of the customers.
- STAGE II Creates customer satisfaction.
- STAGE III Ensures customer loyalty and commitment.

The three stages mentioned above are distinct but interrelated and sequential in nature and also each stage depends on the fulfillment of the previous stage. Customer satisfaction can only be ensured if the trust of the customers on the brand or the product or the company can be built and in the same way customer loyalty and commitment can be achieved

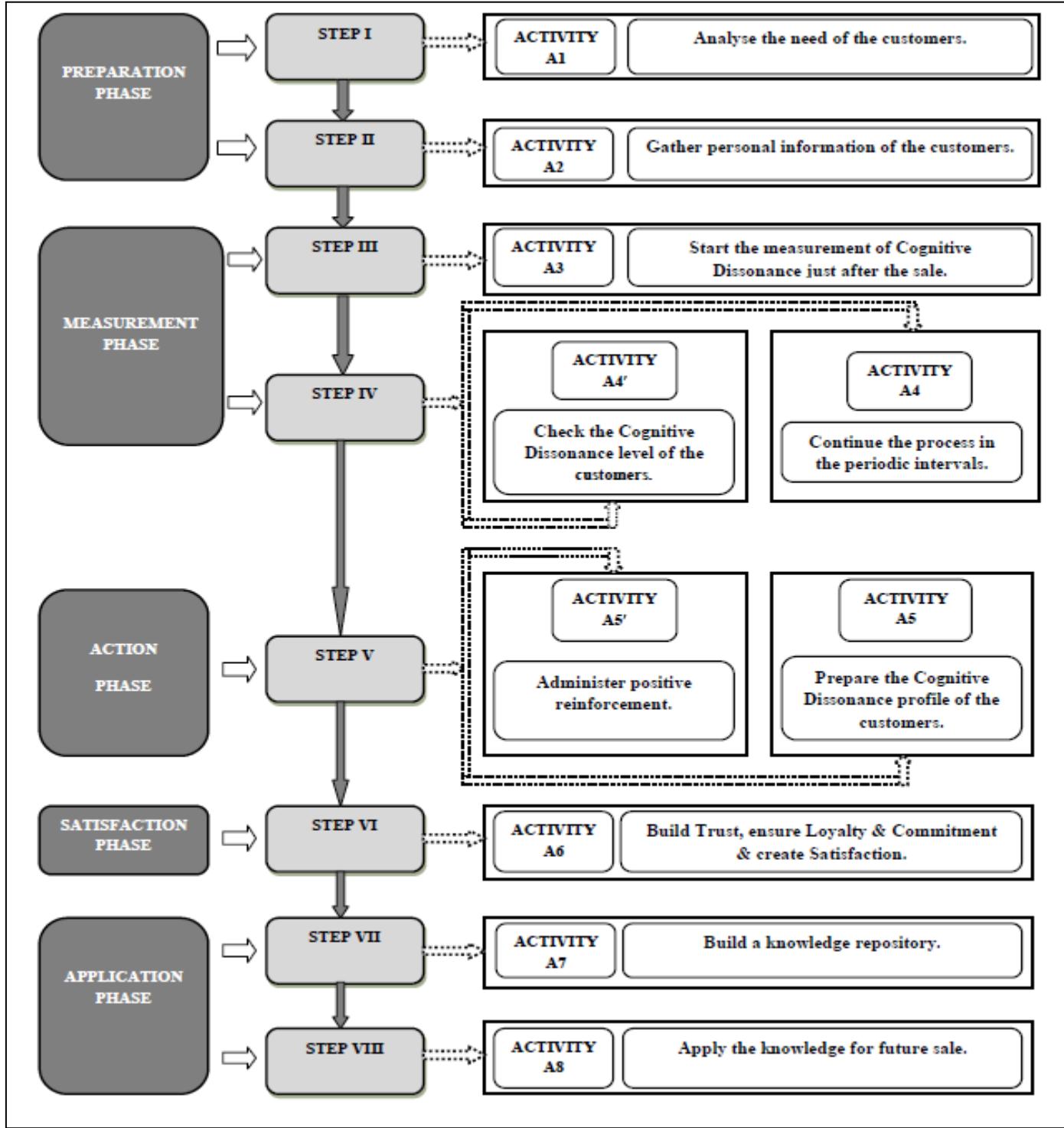


Figure 5. The interrelationship between the phases,activities & steps of the PREMASA MODEL.

only when the customers are satisfied with the product or the service of the company.

B. The Proposed CRM Model: The PREMASA Model

The model proposes a roadmap through which the three tasks of the CRM can be fulfilled. The model proposes total eight steps. Some of the steps have only one activity & some

steps have several activities. From Fig.5 it can be seen that there are total ten activities that the model executes through eight steps. Out of these ten activities some are needed to be performed simultaneously & some are needed to be performed sequentially. The model is multi dimensional in nature because the activities are simultaneous & sequential. The model integrates these sequential & simultaneous activities into eight steps. The steps 4 & 5 have two simultaneous activities each. Activities A4 & A4' as well as the activities A5 and A5' are

2nd Annual International Conference IEMCON 2012

needed to be performed simultaneously. All the ten activities have been arranged into five categories of phases & these are as follows:-

- Activity A1 & A2 fall into Preparation Phase.
- Activity A3, A4, & A4' fall into Measurement Phase.
- Activity A5 & A5' fall into Action Phase.
- Activity A6 falls into Satisfaction Phase.
- Activity A7 & A8 fall into Application Phase.

The categorization of the activities in five phases has been done on the basis of the tasks performed by the activities and that is why the name of the model has been given the PREMASA MODEL (Fig.6).Another important basis of categorization of activities into five unique phases is that activities are grouped together as per their time of execution. Each phase denotes a unique period of time when the activity /activities within that phase are needed to be performed in the following manner:-

- In the Preparation Phase the marketer needs to know the needs & requirements of the potential customers & also the marketer should collect as much personal information of the customers as possible. This can be done through activities A1 & A2.
- In the Measurement Phase the marketer should start measuring the cognitive dissonance level of the customers just after the sale & should continue the process in periodic intervals.
- In the Action Phase the marketer should introduce some positive reinforcement related to the existing services as added incentives or benefits to the customers so that the customers can be able to reduce their dissonance regarding their purchase. In this phase the marketer should create a cognitive dissonance (CD) profile of the customers on the basis of the measurement done in the previous phase.
- In the Satisfaction Phase the marketer should build the trust of the customers by satisfying them through the quality of the service product and through after sales service and also by engaging with the customers in every possible manner.
- In the Application Phase the marketer needs to build & maintain a customer database containing the cognitive dissonance (CD) profile of the customers along with their personal information & usage data. This database or knowledge repository should be used for future sales by predicting the behaviour of the existing or potential customers.

From the marketers' point of view across the phases from phase 2 to phase 5 cognitive dissonance of the customers should reduce gradually and customer trust, satisfaction,

loyalty and commitment should increase and that's why the five phases can be represented as a pyramid and the name of the pyramid has been given as the PREMASA PYRAMID(Fig.7).

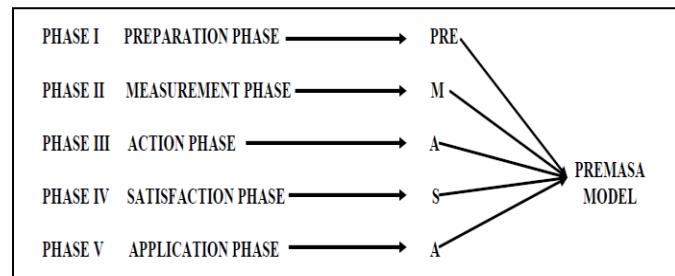


Figure 6. The PREMASA MODEL

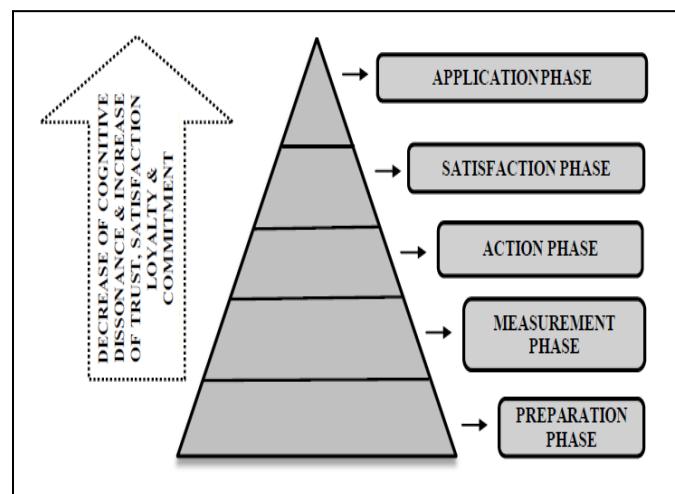


Figure 7. The PREMASA Pyramid

VII. LIMITATIONS & FUTURE DIRECTIONS

The time dependability of the identifiable factors are required to be found out and so a longitudinal study would be better to determine the variability of the key factors with time so that a more comprehensive model can be developed. During the factor analysis, the total variances explained by the three indefinable factors is 72.281 % which is enough to be suitable for empirical investigation but it is also true that in rest of 27.719 % variances there may exist some hidden factors which cannot be traced by this study and for that the number of participants as well as number of variables should be increased in future studies. Although the PREMASA model proposes to make significant contribution to CRM literature, some limitations of the proposed model should be recognized. Firstly, some individuals are comfortable with the disparity of their own experiences with information sources in the midst of decisions of high importance and may not seek to reduce their cognitive dissonance [25]. Secondly, this model does not consider the cultural aspects to how customers deal with

2nd Annual International Conference IEMCON 2012

cognitive dissonance [26]. Though reference [27] considers the marketer-dominated information sources (i.e. mass media advertisements) as insignificant in affecting trust; the PREMASA model should incorporate these sources.

So, future research can extend this model by incorporating the aspect of individual threshold for dissonance & also by including the above discussed limitations. The real challenge lies in the process of validating the reliability and universal applicability of the PREMASA Model in achieving its goal. A quantitative methodology for measuring the level of cognitive dissonance of the consumers is the next phase of development which will pave the way for empirical experiments and investigations to validate the proposed model. Future research should concentrate on the convergence of the theoretical model and experimental results which will provide a realistic assessment of the three critical success factors (Trust, Satisfaction, Loyalty & Commitment) for understanding & analyzing customer relationship and in this way the process of using the PREMASA Model as an integrated CRM framework will be accelerated.

VIII. CONCLUSION

The proposed PREMASA model on Customer Relationship Management revisits the early concepts of consumer behaviour viz the cognitive dissonance and attempts to establish a logical relationship between cognitive dissonance and cognitive, affective and behavioural variables like trust, satisfaction, loyalty and commitment. The proposed model also attempts to integrate the concept of relationship marketing with the traditional concept of CRM and thus makes an effort to strengthen the existing CRM framework to make it more effective & futuristic.

REFERENCES

- [1] E.W.T.Ngai, "Customer relationship management research (1992-2002): An academic literature review and classification," *Marketing Intelligence & Planning*, vol. 23, issue 6, pp.582 – 605, 2005.
- [2] R. S. Winer, "Customer Relationship Management: A Framework, Research Directions, and the Future," unpublished.
- [3] J. Montier, *Behavioural Finance: Insights into Irrational Minds and Markets*, 1st ed. Wiley, 2002.
- [4] L. Festinger, *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press, 1957.
- [5] L.Schiffman and L.Kanuk, *Consumer Behavior*, 10th ed., Prentice Hall, 2009.
- [6] L.L. Berry, "Relationship Marketing," in *Emerging Perspectives on Services Marketing*, L.L.Berry, G.L.Shostack and G.Upah, Eds. Chicago: AMA, 1983, pp 25-28.
- [7] A. Palmer, "The evolution of an idea: An environmental explanation of relationship marketing," *Journal of Relationship Marketing*, vol. 1, issue1, pp79-94, 2002.
- [8] S. Brown, *Postmodern Marketing Two: Telling Tales*. London: Thomson, 1997.
- [9] F.R. Reichheld and W.E.Sasser, Jr, "Zero Defections: Quality Comes to Services," *Harvard Business Review*, vol. 68, issue 5, pp 105-111, 1990.
- [10] A. Palmer, and R. Mayer, "A Conceptual Evaluation of the Multiple Dimensions of Relationship Marketing," *Journal of Strategic Marketing*, vol. 4,issue 4, pp 207-220,1996.
- [11] R. L. Gorsuch, *Factor Analysis*, 2nd ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1983.
- [12] J. Kim and C. W. Mueller, *Factor Analysis: Statistical Methods and Practical Issues*. Beverly Hills and London: Sage Publications, 1978.
- [13] H. H. Harman, *Modern Factor Analysis*, 3rd ed. Chicago: University of Chicago Press, 1976.
- [14] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 5th ed. Allyn and Bacon, 2006.
- [15] P. Kline, *An Easy Guide to Factor Analysis*. London: Routledge, 1994.
- [16] M. Norusis, *SPSS 13.0 Statistical Procedures Companion*. Upper Saddle-River, N.J.: Prentice Hall Inc, 2004.
- [17] A. Smith, "CRM and customer service: strategic asset or corporate overhead?," *Handbook of Business Strategy*,vol.7, issue 1, pp.87 – 93,2006.
- [18] A. K. Kohli and B.J. Jaworski, "Market Orientation: The Construct, Research Propositions, and Managerial Implications," *Journal of Marketing*, vol.5, pp.1–18, April 1990.
- [19] J.C. Narver and S.F.Slater, "The Effect of a Market Orientation on Business Profitability," *Journal of Marketing*, vol.20, pp.20–35, October 1990.
- [20] G.S.Day, "The Capabilities of Market-Driven Organizations," *Journal of Marketing*, vol. 58, pp. 37–52, October 1994.
- [21] D.W. Vorhies and N. A. Morgan, "Benchmarking Marketing Capabilities for Sustainable Competitive Advantage," *Journal of Marketing*, vol.69, pp. 80–94, January 2005.
- [22] J. Sheth, "The Benefits and Challenges of Shifting Strategies," in *Customer Management (MSI conference summary)*, Cambridge: Marketing Science Institute, 2005, pp.4–5.
- [23] M. D. Johnson and F.Selnes, "Customer Portfolio Management: Toward a Dynamic Theory of Exchange Relationships," *Journal of Marketing*, vol.68, pp. 1–17, April 2004.
- [24] K. L. Keller, "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, vol.57, pp.1–22, January 1993.
- [25] G.D. Bell, "The Automobile Buyer after Purchase," *Journal of Marketing*, vol.31, pp.12–16, July 1967.
- [26] S. Kitayama, A.C. Snibbe, H.R.Markus and T. Suzuki, "Is There Any "Free" Choice? Self and Dissonance in Two Cultures," *Psychological Science*, vol. 15, pp. 527-533, 2004.
- [27] R. Walczuch and H. Lundgren, "Psychological Antecedents of Institutional-Based Consumer Trust in E-Retailing," *Information and Management*, vol.42, pp. 159-177, 2004.

Energy Conservation in Buildings

A Bandopadhyay, B.E.E., M.E., M.B.A.
Associate Professor, Electrical Engineering
Institute of Engineering & Management
Salt Lake, Kolkata - INDIA

1.0.0 PREAMBLE

1.1.0 Energy is the prime mover of economic development of any developing nation. But the fact remains, that in pursuit of development, the consumption of energy is increasing at a rapid pace, while the available resources are limited. Out of the global requirement of primary energy, 80% comes from fossil fuels, particularly oil and gas. At current *Reserve/Production (R/P)* ratio, World oil and gas reserves are estimated to last just 42 years and 60 years respectively. Coal is likely to last for 122 years {Source : Bureau of Energy Efficiency].

1.2.0 Currently, India is the fourth largest oil-consuming country in the world. India imports about 75% of its crude oil needs, mainly from Gulf nations. The burden of import of crude oil, petroleum products and LNG, during the year 2008-09 was,

- Quantity (Million Metric Tonnes) ... 154.50
- Import Bill (Rs. Crores) ... 2,94,870

The country's demand for coal is also rising, particularly from the steel and the power sector which are expanding their capacities. About 15% of the total annual coal requirements of our country (both coking and non-coking coal) are met through import. Total quantity imported during the year 2008-09 was 59.00 Million Tonnes.

The aggregate installed capacity of Power Generating Stations in India stands at 1,61,351MW.

As far as electricity consumption is concerned, current level is 631 kWh per person per year, while the comparable figure for UK is 6206, for Japan 8,076, for USA 13,338 and for Canada is 17,179.

1.3.0 The use of electricity is growing rapidly in urban **Residential, Commercial and Services**. Out of the

total electricity demand in this sector, approximately 70% accounts for lighting purposes while the balance 30% is used for refrigeration, air-conditioning, pumping and other electrical appliances. The energy consumption especially for commercial and services activities is expected to grow rapidly due to growth rate in commercial establishments, hotels, shopping malls, IT parks and hospitality industries.

2.0.0 BUILDING DEFINITION

"Building", as per the definition in Energy Conservation (Amendment) Bill 2010, refers to any structure or erection or part of structure or erection, which is having a connected load of 100 kilo Watt (kW) or contract demand of 120 kilo Volt Ampere (kVA) and above and is used or intended to be used for commercial purposes.

3.0.0 ENERGY CONSERVATION BUILDING CODES (ECBC)

3.1.0 Energy Conservation Building Codes (ECBC) was developed to deal with rapidly increasing energy consumption in commercial buildings. ECBC sets minimum energy efficiency standards for design and construction of commercial buildings, taking into consideration the building's function, comfort, health and the productivity of the occupants, and at the same time ensuring minimum cost of construction and energy, over the entire life cycle of the building.

3.2.0 ECBC defines the norms of energy requirement per square meter of area and takes into consideration the climatic region of the country, where the building is located. The owners or occupants have to comply with the energy consumption norms and standards and / or to prepare and implement schemes for its efficient use and conservation.

3.3.0 In order to implement ECBC across the country in India, the country has been segmented into five (5) climatic zones as per weather conditions. These are,

- i) **Cold** : This includes Jammu & Kashmir, Northern part of Himachal Pradesh, Uttarakhand, Sikkim and part of Meghalaya, Arunachal Pradesh, Tamilnadu

- ii) **Composite:** This includes Punjab, Haryana, Delhi, Uttar Pradesh, Bihar, part of Rajasthan, Gujarat, Madhya Pradesh, Chhattisgarh, Jharkhand and the remaining areas of Himachal Pradesh, Uttarakhand
- iii) **Hot dry:** Remaining areas of Rajasthan, Madhya Pradesh and part of Gujarat, Maharashtra, Bihar, Orissa
- iv) **Moderate /Temperate:** Bangalore and neighboring areas of Karnataka
- v) **Warm-Humid:** Remaining areas of Gujarat, Maharashtra, Karnataka, Tamilnadu Orissa, Bihar, Meghalaya and Kerala, Andhra Pradesh, West Bengal, Assam, Tripura, Manipur, Mizoram, Nagaland

3.4.0 In order to effectively fulfill its objectives, ECBC sets guidelines in respect of the following aspects of building design:

- a) Building envelopes
- b) Mechanical system including heating, ventilation and air-conditioning
- c) Service hot water heating
- d) Interior and exterior lighting
- e) Electrical power (e.g., UPS) and motors (e.g., pumps, elevators/escalators etc.)

4.0.0 ECBC GUIDELINES

4.1.0 BUILDING ENVELOPE

The building envelope refers to the exterior facade, and comprises of wall, windows, roof, skylight, doors and other openings. From energy efficiency point of view, the envelope design must take into consideration both the external and internal heat loads, as well as day lighting benefits.

One of the key objectives of envelope design should be to introduce day lighting into the interior space of the building through windows and skylights, thereby reducing dependence on electric lighting. For that purpose, building orientation and sizing & placement of windows shall be given due consideration at design stage.

Other objective is to maintain thermal comfort and minimize internal cooling / heating loads. For that purpose, the building envelope needs to regulate and optimize heat transfer through roof, wall, windows, doors and other openings.

4.2.0 HEATING, VENTILATION & AIR CONDITIONING

Heating, ventilation and air conditioning (HVAC) refers to the equipment, distribution systems and terminals that constitute the overall system installed to meet the HVAC requirement of the building. The ECBC code sets important guidelines for designing / selection of equipment pertaining to the HVAC system:

- Natural ventilation : Maximum possible use of wind-induced natural ventilation.
- Equipment efficiency : Minimum equipment efficiencies are required to be met for all HVAC equipments. Standards and labeling as per star rating of B.E.E. or other applicable codes may be followed to ensure minimum power consumption for various sizes and capacities of HVAC equipments.
- Controls: In this context, the use of time switches, occupancy sensor etc. for temperature controls/thermostat and twin speed drive or VSDs are recommended.
- Pipe & duct work : For effective minimization of heat losses and gains, proper insulation of pipelines and duct work is important. Selection of insulation material and calculation of insulation thickness shall be done precisely based on the intended service conditions.
- System balancing : All HVAC systems shall be balanced with respect to generally accepted engineering standard. Balancing an air based or water based HVAC system of building provide improved comfort level and at the same time makes it truly energy efficient
- Condensers: Special care should be taken while locating the condenser so that the heat sink is free from interference from heat discharge from any other equipment located near by.
- Economizers : Each individual cooling fan system that has the design supply capacity over 2500 cfm and a total mechanical cooling capacity over 6.3 Tonnes shall include an economizer. Economizers allow the use of outdoor air to cool the building when the outside temperature is cooler than that inside.

- Hydronic systems : The code specifies the use of equipments of type and capabilities to reduce pump energy such as variable fluid flow, automatic isolation valves and variable speed drives.

4.3.0 SERVICE HOT WATER

- 4.3.1 Supplementary water heating facilities for hotels and hospitals must be given due consideration at planning stage. Solar water heating for at least 1/5th of the design capacity is very much feasible in most of the locations of our country. Maximum heat recovery from hot discharges like condensers of air conditioning units, exhaust of Diesel Generating sets shall also be incorporated in design features.
- 4.3.2 The entire hot water system including the storage tanks, pipelines shall be insulated conforming to the relevant IS standards.

4.4.0 LIGHTING

- 4.4.1 For proper design of lighting system, in a building, it is important to decide the required illuminance on different work places, vis-à-vis the type of lamp and luminaire ideally suited for each work place. For recommended illumination in different work places of residential, commercial buildings and services, reference may be made to *Illumination Engineers Society Recommended Handbook* and / or IS : 3646-1992. Decision regarding selection of type of lamp and luminaire most often is made as a function of both aesthetics and economics. However, it is important to examine the technical feature of each type of light source including polar curves furnished by the manufacturers, vis-à-vis the functional requirements of each work place.

- 4.4.2 The utility of using natural day lighting during the daytime is more preached than practiced. This aspect is being increasingly ignored especially in modern air-conditioned office spaces and commercial establishments. Innovative designs are possible which eliminates the glare of daylight and blend well with the interiors.

- 4.4.3 Code provides guidelines as to the lighting power required, in **lux/W/m²** for maintained illuminance in different workplaces / applications for Commercial lighting. If the actual lux per wall per square metre is largely at variance with the target figure recommended as per code, there is a scope for review.

- 4.4.4 Code guidelines are also applicable for lighting of building exterior features (e.g., façade, canopy) and outdoor pathways, lawn etc. power supply for which is catered from building's electrical service.

4.5.0 WATER PUMPING SYSTEM

The pumps used in the buildings are for chilled water circulation, cooling water circulation, domestic water, hot water and sewage water. Normally the pumps used in the buildings are of centrifugal type. The code requirement is selection of high efficiency pumps and motors, as well as suitable means for their efficient utilization, keeping in view their operational requirements. For example, installation of variable frequency drive for chiller motor and cooling water pump-motors, merit consideration

4.6.0 UPS

- 4.6.1 An Uninterruptible Power Supply (UPS) is a device that has an alternate source of energy, typically a battery back up that can maintain power supply to emergency loads, when the normal power source is temporarily disabled. A UPS comprises a converter (AC to DC), battery, inverter (DC to AC), monitor and control elements. Use of UPS is most common in modern intelligent buildings for supplying the loads of Building Management System, as well as computer installations, servers, telecom system etc. for office and Commercial establishments.

- 4.6.2 Energy efficiency of an UPS is important criterion for selection of its capacity. Most of the manufacturers specify their efficiency parameters and submit test reports with linear loads, while in reality most of the loads catered by an UPS, e.g., computers, servers, and other electronic equipment are non-linear in nature. Therefore actual efficiency of the same UPS is much lower when powering these non-linear loads.

- 4.6.3 UPS efficiency also depends largely on the power level at which it operates. Selection of capacity, ensuring minimum 60% loading at actual operating condition is necessary to avoid inefficient use.

4.7.0 ELEVATORS AND ESCALATORS

Losses normally encountered in elevators and escalator system of a Commercial building are, friction losses incurred while travelling and dynamic losses while starting and stopping. In general, the approach to energy efficiency in elevators and escalator equipment is to minimize these two loss factors by proper maintenance of the equipment, strictly as per manufacturer's recommendations.

5.0.0 **ASSESSMENT OF ENERGY PERFORMANCE IN BUILDINGS**

Any audit/independent study on energy performance of an existing Building should primarily focus on assessment of energy performance of HVAC System and Lighting System - which are the two main energy consuming areas in any Building for Commercial purposes & services. Other areas of energy use need to be addressed only when, significant deviations are recorded w.r.t. standard checklists for such items.

5.1.0 ENERGY SAVING OPPORTUNITIES - HVAC SYSTEM

- a) **Cold Insulation:** All cold lines / vessels shall be insulated properly using economic insulation thickness and appropriate insulating material
- b) **Building Envelope:** To optimize air conditioning volume by measures such as use of false ceiling, false flooring and segregation between critical and non-critical areas for air-conditioning by air curtains.
- c) **Building Heat Loads Minimization :** To minimize air-conditioning loads by measures such as roof cooling, roof painting, efficient lighting, sun film application, pre-cooling of fresh air etc.
- d) **Process Heat Loads Minimization:** For effective minimization of process heat loads in terms of TR capacity as well as refrigeration level, following measures are suggested:
 - Flow optimization
 - To increase heat transfer area
 - Compressor capacity control during partial load operation
 - Regular cleaning / de-scaling of heat exchangers

e) **At the Refrigeration A/C Plant :**

- To ensure regular maintenance of all parts of A/C Plant as per manufacturer's recommendation
- To ensure adequate quantity of chilled water and cooling water flows
- To avoid part load operations by matching loads and plant capacity

5.2.0 ENERGY SAVING OPPORTUNITIES - LIGHTING SYSTEM

- a) **Task Lighting**
- b) **Replacement by high efficiency lamps and luminaries**
- c) **De-lamping to reduce excess lighting**
- d) **Reduction of lighting feeder voltage :** A reduction in voltage as low as 380V reduces the lumen output of fixtures only marginally, by 5% - which is not even noticeable but the power consumption is reduced by 16%.
- e) **Use of Electronic ballasts:** The losses in electronic ballasts for tube lights are only about 1 watt, in place of 10 to 15 watt in standard electromagnetic chokes.
- f) **Occupancy Sensors:** Occupancy - linked control should be practiced in Office buildings using infra-red, acoustic or micro wave sensors.

5.3.0 ENERGY SAVING OPPORTUNITIES FOR OTHER AREAS

5.3.1 **Pumping System**

- a) To ensure proper selection of pumps by operating them near best efficiency points
- b) To replace old pumps by energy efficient pumps
- c) To stop running of multiple pumps
- d) To use booster pumps for small loads requiring higher pressures

5.3.2 **DG Sets**

- a) To ensure steady load conditions on the DG set. Parallel running of the sets shall be considered wherever feasible, to ensure improved loading and fuel economy.

- b) To improve air filtration
- c) To ensure compliance with maintenance checklist

5.3.3 Building Management System (BMS)

Now-a-days BMS can be developed by provision of suitable controllers as well as software, to include additional feature for effective control & monitoring of various services and functions (e.g., room temperature and humidity) of a building.

6.0.0 STAR RATING OF BUILDINGS

- 6.1.0 This programme under ECBC initiative aims to rate office buildings on a 1 – 5 star scale with 5 star labeled buildings being the most energy efficient. Energy Performance Index (EPI) in kWh per sq m per year is a measuring tool to evaluate the performance of the building in terms of the total energy consumption, for such rating. This shall include electricity generated from on-site diesel generating sets but not renewable sources like solar photovoltaic etc.
- 6.2.0 The bandwidths for Energy Performance Index for different climatic zones have been developed based on percentage air-conditioned space. For example a building in New Delhi having air conditioned area greater than 50% of the built up area, would qualify for 5 Star if its EPI falls below 90 kWh/sq m/year and 1 star if it is between 165-190 kWh/sq m/year.
- 6.3.0 The need of the hour is to formulate a proper incentive schemes for the occupiers of star rated buildings, depending on their EPI registered. The scope of National Energy Fund should be enlarged to include finance required for such incentives, vis-à-vis energy R&D initiatives – on a competitive basis. Energy conservation opportunities In Building sector must be given top most priority, since it can help our country to effectively reduce the energy requirement, without affecting the least, the economic development programmes of the country.

8.0.0 ENERGY AND ENVIRONMENT

- 8.1.0 Before concluding this Paper, it may not be out of place to mention that the combustion of hydrocarbon based fuels in industrial activity generates by-product materials, many of which are considered to be air pollutants. The worst polluting elements are particulate matters (dust), sulphur oxides, nitrogen oxides, hydrocarbons and carbon monoxide. Both SO_x and NO_x emissions have been identified all over the world as major air pollutants as they lead to acid rain.

- 8.2.0 Following are some of the key environmental issues faced by the mankind without regard to any particular country or region:

- Acid rain
- Ozone layer depletion
- Global warming and climate change
- Loss of biodiversity

These environmental issues have global significance and hence need to be addressed through International cooperation and initiatives.

REFERENCES

- [1] Energy Efficiency in Electrical Utilities – Guidebook for National Certification Examination for Energy Managers and Energy Auditors : Bureau of Energy Efficiency
- [2] ECBC Guide Book
- [3] Electrical Energy Audit – Back Ground Paper : National Productivity Council
- [4] <http://www.bee-india.nic.in>.

Is Emerging Green Technology Poses Threat of Trade Wars – some solutions? – A Survey

S. S. Thakur¹, S. Bhadra², Sougata Dey^{2*}

¹Department of Computer Science & Engineering, MCKV Institute of Engineering, Liluah, Howrah 711204
subroto_thakur@yahoo.com

²Department of IT, Seacom Engineering College, Dhulagarh

^{2*}Department of IT, MCKV Institute of Engineering, Liluah, Howrah 711204

Abstract Investors recognize that cheap oil won't last forever and emerging green technologies could revolutionize everyday business as much as computers did. As with any new technology, nations compete to perfect and produce new products for the world, making lots of money along the way, explains international economics columnist Bruce Stokes. But Stokes warns that the global trading system lacks adequate regulations on energy issues – and the result could be trade wars. Already, the US has blocked purchase of wind turbines from China to protect US jobs, and China has imposed quotas and export bans on needed materials to corner the market on some alternative energies. Deliberate domestic and global industrial policies – setting goals and nurturing demand – could speed development of these technologies. Failing that, ensuing trade wars could delay, even thwart, protection of the planet. In this paper some solutions for the same has been proposed, so trade wars can be minimized, along with useful applications.

Keywords- Green Computing, health hazards, carbon capture, threats etc.

I. INTRODUCTION

What's the biggest problem facing Earth's atmosphere today? The correct answer is increasing carbon dioxide levels. CO2 isn't the harshest greenhouse [1, 2] gas, but it's the one that is emitted the most and it's capable of causing quite a bit of damage. There's a finite amount of carbon dioxide in Earth's biogeochemical carbon cycle, the problem is we humans are releasing CO2 in droves, artificially accelerating the carbon cycle through means like removing fossil fuels and burning them.



Figure 1: Pollution in Environment

The best solution to this problem would be to simply stop releasing CO2, but see, we have this thing called the global economy and it runs on coal and oil. This is called

carbon capture (Figure 1). Instead, research has yielded a potentially equally acceptable solution, simply capturing the carbon dioxide before it makes its way into the air.

There are a couple of ways to capture CO2, and most are centered around coal, which even in clean form is exceptionally dirty. There's a single coal-fire power plant run by Southern Company here in Georgia that emits more CO2 than the entire electricity infrastructure of Brazil. One method is to remove carbon dioxide from coal before it's burned. Another is to catch it in the flues before it makes its way out in a puff of smoke. Researchers at Georgia Tech have come up with a synthetic, sand-like material that can catch CO2 and trap it. The stuff is called hyper branched amino silica, which is Latin for really, really branched amino silica. Now that you've got it captured, what to do with the CO2? Here's where the storage part comes in. Most CO2 storage going on today has to do with trapping the stuff below ground, another proposal is to keep it in liquid form in bags along the continental shelf. Either way, it's possible to find a way to regenerate captured carbon into a synthetic fossil fuel, forming a closed loop where fuel is burned, CO2 is captured and reused once more.

II. PROPOSED APPROACH

Our approach focuses on the following points discussed below.

- ÉEmerging Green Technology Poses Threat of Trade Wars
- ÉPossible solutions
- ÉImportance of Human Experience
- ÉGreen Building
- ÉImportance of Green Building
- ÉApplications
- ÉDiscussions and conclusions

EMERGING GREEN TECHNOLOGY POSES THREAT OF TRADE WARS:

After long time has been elapsed, now countries are waking up to the dangers of carbon-dioxide pollution [3, 4, 5] and a finite supply of fossil fuel. The bad news is that emergent green technology risks spawning new trade wars. A firestorm of public criticism emerged in the United States last year, with reports of taxpayer stimulus funds being used to buy 240 Chinese-made turbines for a new wind farm in west Texas. The world's pursuit of low-carbon sources of energy collided with the national need to create jobs. China, Germany, Spain and the US vie with others to reap

economic and environmental benefits of domestic green-energy sources while positioning themselves as market leaders in providing those technologies to the world. While environmentalists applaud this new focus on renewable energy, the inherent conflict between nations' climate needs and their competitiveness agendas threatens the global trading system, which lacks adequate rules and procedures on energy issues. Without a concerted effort to liberalize trade in green technologies, efforts to minimize climate change could rapidly devolve into new trade wars.

China intends to meet 20 percent of its energy needs from renewable energy sources by 2020. In so doing, it hopes to become the global production site for green technology. In 2009, China accounted for more than a third of the world's wind-capacity installations, more than doubling its cumulative installed capacity for the fourth year in a row. And the country has passed the US to become the world's largest wind-turbine market. By 2020, China aims to obtain up to a fifth of the nation's energy from renewable sources such as wind, solar and hydropower. China hopes to meet its goals with Chinese-made equipment. In 2004, four-fifths of Chinese purchases of wind power equipment came from abroad – overwhelmingly from Europe. This year less than a fifth of such technology will come from overseas. China has yet to become a major player in global wind-turbine markets, with exports to Europe and the US equaling less than 1 percent of total Chinese production. But the growth in domestic Chinese output of wind turbines now outpaces the country's own demand.

Chinese producers are poised to become big exporters. In 2008, China also emerged as the largest producer of solar panels in the world, accounting for roughly one-third of total solar shipments. Beijing hopes to increase domestic generation of electricity from solar panels from 3 gigawatts in 2010 to 20 gigawatts by 2020. But so far growth in Chinese production of solar panels has outstripped the growth of the installed solar capacity in China. The vast majority of solar panels produced in China are exported. Between 2007 and 2008, for example, the value of Chinese exports of solar panels to Europe more than doubled. Sales to the US during the first 11 months of 2009 were two-thirds higher than the value of such exports during the first 11 months of 2008.



Figure 2: Wind turbines as energy source

China's success in the renewable-energy technology market was accomplished through old fashioned industrial policy. Beijing muscled out American and European manufacturers seeking to gain a foothold in China's burgeoning market for renewables by establishing prohibitive quotas requiring purchase of homegrown solar and wind-turbine equipment, and by various schemes to disqualify bids from foreign companies to supply such technology. China also tried to corner the market for wind turbines by restricting the sale of rare-earth minerals, critical for manufacturing turbines, by blocking export of those metals. China has the world's largest known deposit of rare-earth minerals.

Europe also has high renewable-energy ambitions. Germany already gets 16 percent of its electricity from renewable sources such as solar and wind. A new McKinsey & Co. study concludes that "By 2050, Europe could achieve an economy-wide reduction of [greenhouse gas] emissions of at least 80 percent compared to 1990 levels." Achievement of this goal requires installation of about 1,900 square miles of solar panels, many on rooftops, and more than 2,000 new wind turbines per year, half of which might be at sea. The goal, while daunting, is about the same pace of installation achieved by the European wind sector over the last 10 years. The benefits for Europe would not be limited to slowing carbon emissions. By 2050, McKinsey expects the cost of energy per unit of GDP could actually be reduced by 30 percent in Europe, boosting competitiveness. Production of renewable technology could create tens of thousands of new jobs. As nations vie for global leadership in the wind, solar and other renewable energy fields, trade disputes are inevitable. In August 2009, two major German solar-technology firms filed a complaint with both the German government and the European Union about government subsidies allegedly given to Chinese competitors.

Such rows highlight the growing tension between international commitments to reduce national carbon footprints and existing global trade rules. The Texas wind farm case is only one example of this inherent conflict. Wind energy is abundant in west Texas and would help reduce US dependence on imported oil. But the purchase of wind turbines for the project was also intended to create jobs and spur economic recovery. In the end, the public demanded that turbines be produced in the US, not China. To free-trade advocates, this seems like a classic example of "Buy America," a practice criticized around the world as protectionist. But a recent study by the World Trade Institute in Geneva concluded: "current rules on government procurement do not systematically address the linkage to green procurement. There is therefore controversy as to what extent [countries] are entitled to condition government procurement in the light of goals set out in the Kyoto [climate change] Protocol."

POSSIBLE SOLUTIONS:

Built environments often define the fabric of our communities and play a central role in physical and

psychological health. Today, the majority of empirical data collection in built environments focuses on physical attributes and environmental performance, such as energy or water consumption. We are building increasingly sophisticated systems to collect, analyze, and use information on building energy consumption; information networks that soon will stretch from the power plant to a Smart Meter and, in some cases, to a Smart Phone. This creates unprecedented opportunities to manage energy use and improve energy efficiency.

While the volume of information about energy and, to a lesser degree water, is growing rapidly, information about the experience of people in and around built environments lags far behind. The relevant dimensions of human experience encompass traditional notions of occupant productivity, comfort, and satisfaction, as well as related concepts of walkability, well being, connectivity, community, and social capital (Dearly 2004). In an attempt to better understand these concepts, and develop a framework for the sustained collection of data on actual human experience within the built environment this paper explores the intersection between three important concepts:

ÉHuman experience;

ÉVolunteered Geographic Information (VGI);

ÉGreen building.

Our goal is to explore opportunities to test strategies with practice based experiments.

IMPORTANCE OF HUMAN EXPERIENCE:

Human experience is one of the most critical barometers of the success of a built environment. Traditionally, human experience in and around built environments [6, 7, 8, 9, 10] has been evaluated through surveys, interviews, and, in some cases, direct observations. These tried and true methods yield important insights, but they are not readily scalable or spatially extensible. Every observation requires substantial investments in time and energy and is difficult to generalize and iterate. We need new, scalable sources of information and systematic feedback processes to help advance consideration for occupant experience as a part of evidence based green building practice.

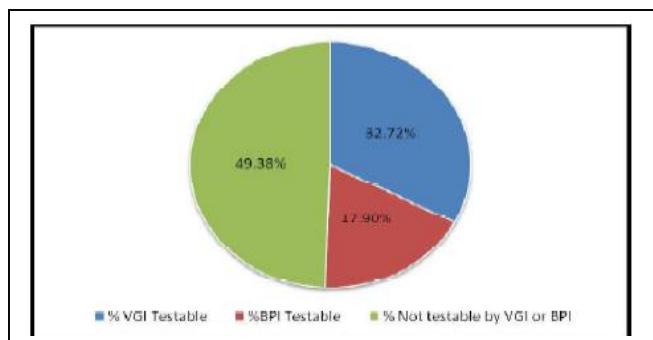


Figure 3: Percentage of Green Building practice

GREEN BUILDING:

Green building is a systematic effort to create, sustain, and accelerate changes in practice, technology, and behavior to reduce building-related environmental impacts while creating places that are healthier and more satisfying for people. In the United States, green building has largely emerged from communities of practitioners working to define beneficial processes and practices and create mechanisms to recognize and encourage their use. Today, the U.S. Green Building Council realizes this vision with a combination of processes and tools, including consensus based rating systems, rigorous third party review and certification, and educational resources. One of the most important tools is the Leadership in Energy and Environmental Design (LEED) green building certification program. LEED® rating systems include a combination of required prerequisites and elective credits. Here we refer to these elements collectively as green building strategies [11, 12, 13, 14].

Each strategy has a clearly defined intent that describes the desired outcome. Each intent is crafted by teams of volunteers and subjected to multiple rounds of public review and comment. Each credit provides one or more options for specific strategies that can achieve its intent, and each option is associated with specific required documentation. During the certification process, documentation is provided to an independent review body which ultimately determines whether each strategy meets LEED requirements. Requirements for LEED certification are not static. Rather, they reflect an explicit commitment to ñraising the barö toward an ultimate goal of regenerative built environments ó buildings and communities that actually improve conditions for people and the environment. Change is often manifested in the details of the rating system with new documentation requirements or so called alternative compliance paths. Credit intents represent the most stable aspect of the system; they reflect a high level aspiration that, in practice, may be fulfilled in many different ways. Current green building processes and practices have been successful in promoting the use of specific strategies during planning, design, construction, and operation of neighborhoods, new construction, and existing buildings.

This success is reflected in over 130,000 trained, accredited professionals and nearly 5,000 certified projects with another 20,000 in the process of pursuing certification. Each of these projects contains a distinct, sometimes unique, combination of green building strategies. Each strategy achieved by every certified project is recorded with USGBC and potentially available for analysis. Each project is also associated with a project team, typically including a LEED Accredited Professional. Each LEED AP® has demonstrated a familiarity with green building concepts and, at minimum, possesses a working vocabulary to describe important aspects of building performance and experience. Ultimately, LEED provides a number of

important elements to the confluence of human experience and VGI. LEED brings explicit aspirations or intents for individual green building strategies, verification of the implementation strategies [15, 16] on specific projects, and a cadre of trained professionals.

IMPORTANCE OF GREEN BUILDING:

Green building is a movement dedicated to the transformation of practice in the design, construction, and operation of built environments. The objective is to reduce the negative impacts of built environments while creating healthy, comfortable, and economically prosperous places for people to live, work, and play. The popular term “green building” encompasses the collection of processes, institutions, and individuals that serve to assess current practice, identify opportunities for improvement, develop and deploy tools, and provide independent review and recognition of results. The green building community has diversified from its origins in the architecture and engineering professions to encompass the full range of professionals involved in lifecycle of built environments.

III APPLICATIONS

The most impressive feature of the new Bahrain World Trade Center is, no doubt, the three massive wind turbines situated between the two towers comprising the main building. Each of these 80-foot turbines projects from a bridge between towers. The shape of the towers themselves channels and accelerates air moving between them which will help the building generate even more power. It is by far the largest wind-powered design incorporated into a massive building project to date.



Figure 4: Bahrain “World Trade Center”

The so-called Lilypad Project is perhaps the most fantastical of these green wonders and certainly the farthest from being built but is too amazing a concept not to mention. The idea is to create a series of floating self-sufficient ocean-going eco-city islands. Each one would be able to house 50,000 residents and would support a great deal of biodiversity. Collecting pools located in their centers would gather and filter water for use on board. These would be places for adventurers and refugees alike as water levels rise around the world and threaten many, particularly island, habitats.



Figure 5: Lilypad Project

The MagLev Wind Turbine is a big step forward in the world of wind power. By using magnetism to levitate the blades friction is eliminated and more power can be produced without any additional power expense (since the magnetics require no energy to run). The MagLev has a low threshold velocity for producing energy, could theoretically survive for centuries and can power up to 750,000 homes. Though the initial investment involved hundreds of millions of dollars the payoff is potentially huge.

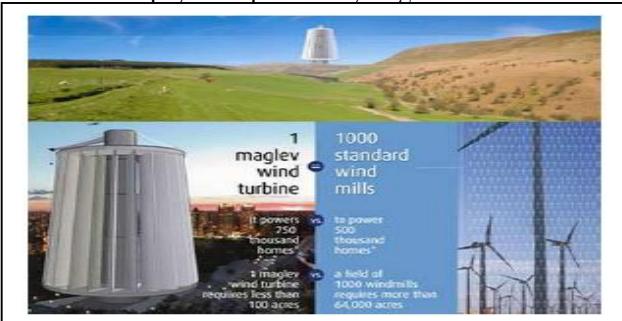


Figure 6: MagLev Wind Turbine

Urban skyscraper farms are still purely conceptual for now but are amazing theoretical propositions. They would provide locally grown food in densely packed urban centers. Some such designs incorporate elaborate rainwater reuse systems and other sustainable strategies intended to minimize their environmental impact and maximize their productivity [17, 18, 19, 20]. However, they are massive in scale and would cost a great deal to build. This huge initial outlay is part of what is keeping them out of production.



Figure 7: Urban skyscraper farms

IV DISCUSSIONS AND CONCLUSIONS

To avoid such controversies, the World Trade Institute advocates new energy negotiations under the auspices of the World Trade Organization to address all the pertinent problems, ranging from issues of classification of goods and services, to disciplines on subsidies, to issues of competition and state trading, as well as intellectual property rights and government procurement.⁶ International consensus for such a global pact does not yet exist and may require years of intensifying trade friction to convince governments of the desirability. Until then, the Obama administration and the EU are set to consider a less ambitious path: a deal among all nations to eliminate tariffs on renewable-energy technologies. Such an accord would liberalize rather than constrain trade, removing some potential future frictions between trade and climate goals. But even that may not be easy. China must be involved for any deal to be meaningful. Agreement on what constitutes a renewable-energy technology will not prove simple.

Nevertheless, the inherent tension between efforts to develop technologies that cut carbon emissions and longstanding commitments to trade cannot be ignored. Unless disciplined by new global rules of the road, competition for market share between China, Europe and the US will lead to abusive government actions and trade tensions that distract from the goal of protecting the planet.

We have an opportunity to shift green building from the implementation of Best Practices toward an evidence based practice based on practice-based evidence. This transition will require us to identify opportunities to adapt technology to better serve the purpose of understanding human experience in ways that create systematic information that can be combined with traditionally collected data and emerging sensor technology [21, 22]. Taken together, a sustained, integrative approach to understanding people, information, analytics, and practice can help drive market transformation in ways that increase the prevalence of practices that demonstrably improve human experience in built environments [23, 24, 25].

ACKNOWLEDGMENT

The authors ^{1,2*} are thankful to Principal, MCKVIE, Liluah for giving us the opportunity to do useful work for the International Conference. The author ² gives all the credit to her beloved husband for giving opportunity, and every support and guidance for the said work.

REFERENCES

- [1] Abbaszadeh, S., Zagreus, L., Lehrer, D., Huijzen, C. 2006. Occupant satisfaction with indoor environmental quality in green buildings. *Proceedings of Healthy Buildings, Lisbon*. Volume III: 365-370.
- [2] Arnold, J, Cochrane, L. 2009. Future opportunities for the energy ecosystem. *Power Engineering* July 2009. 2 pages.
- [3] Baum, M. 2007. Green Building Research Funding: An Assessment of Current Activity in the United States. U.S. Green Building Council. 37 pages.
- [4] Boulos, M.N.K., Wheeler, S. 2007. The emerging Web 2.0 social software: an enabling suite of sociable technologies in health and health care education. *Health Information and Libraries Journal*. 24(1):2-23.
- [5] Brager, G.S., Paliaga, G., and de Dear, R. 2004. Operable windows, personal control, and occupant comfort. *ASHRAE Transactions: Research*. RP-1161:17-35.
- [6] Browning, W.D., Romm, J.J. 1994. Greening the Building and the Bottom Line. 15 pages. California Department of General Services. 2002. Real Estate Services Division. Carnegie Mellon University, Center for Building Performance Diagnostics. eBIDS. <http://cbpd.arc.cmu.edu/ebids>.
- [7] Dannenberg, A.L., Jackson, R.J., Frumkin, H., Scieber, R.A., Pratt, M. Kochtitzky, C., Tilson, H.H. 2003. The impact of community design and land use choices on public health: a scientific research agenda. *American Journal of Public Health* 93(9):1500-1508.
- [8] Dearry, A. 2004. Impacts of our built environment on public health. *Environmental Health Perspectives* 112(11):A600-A601.
- [9] Ewing, R., Kreutzer, R. 2006. Understanding the Relationship between Public Health and the Built Environment. 137 pages.
- [10] Fisk, W.J. 2000. Review of Health and Productivity Gains From Better IEQ. 14 pages.
- [11] Goodchild, M.F. 2007. Citizens as Voluntary Sensors: Spatial data infrastructure in the World of Web 2.0. *International Journal of Spatial Data Infrastructures Research*. 2:24-32.
- [12] Green, L.W. 2006. Public health asks for systems science: to advance our evidence based practice, can you help us get more practice based evidence. *American Journal of Public Health* 96(3):406-409.
- [13] Kaplan, R. Kaplan, S. 1989. The experience of nature: A psychological perspective. Cambridge University Press.
- [14] Kats, G. 2003. The Costs and Financial Benefits of Green Buildings, A Report to California's Sustainable Building Task Force.
- [15] Kats, G. 2010. Greening our Built World: Costs, Benefits, and Strategies. Island Press. 258 pages.
- Kettel Kahn, L., et al. 2009. Recommended Community Strategies and Measurements to Prevent Obesity in the United States. Morbidity and Mortality Weekly Report 58(RR07):1-26.

- [16] McGinnis, J.M., Foege, W.H. 1993. Actual causes of death in the United States. *Journal of the American Medical Association* 270:2207-2012.
- [17] McGraw Hill Construction. 2006. Green Building SmartMarket Report. 42 pages. McGraw Hill Construction. 2008. The Green Home Consumer: Driving Demand for Green Homes. 33 pages.
- [18] McNeill, D.L., Wilkie, W.L. 1979. Public policy and consumer information: impact of new energy labels. *The Journal of Consumer Research* 6(1):1-11.
- [19] Milton, D. P. et. al. 2000. Risk of Sick Leave Associated with Outdoor Air Supply Rate, Humidification, and Occupant Complaints. *Indoor Air* 10(4): pp. 212-21.
- [20] Nelson, A., Rakau, O., Dorrenberg, P. 2010. Green buildings: a niche becomes mainstream. *Duetsche Bank Research*. April 12, 2010. 24 pages.
- [21] O'Reilly, T. 2007. What is Web 2.0: Design patterns and business models for the next generation of software. *Communications & Strategies* 1:17.
- [22] Peuquet, D.J. 1994. It's About Time: A conceptual framework for the representation of temporal dynamics in Geographic Information Systems. *Annals of the Association of American Geographers*. 84(3):441-461.
- [23] Simons, H., Kushner, S., Jones, K., James, D. 2003. From evidence-based practice to practice based evidence: the idea of situated generalization. *Research Papers in Education* 18(4):347-364.
- [24] Romm, J.J. 1994. *Lean and Clean Management*.
- [25] Troped, P.J., Wilson, J.S., Matthews, C.E., Cromley, E.K., Melly, S.J. 2010. The build environment and location based physical activity. *American Journal of Preventive Medicine* 38(4):429-438
- Trowbridge M.J., McDonald N. 2008. *Urban Sprawl and Daily Miles Driven by Teenagers in the United States*. *American Journal of Preventive Medicine*. 34(3):202-206.
- U.S. Green Building Council, 2009. Presentations entitled "About USGBC" and "About LEED".

Design of a Universal Solar Power Driven Mobile Phone Charger for Rural Areas of India

Ms Priyanka Roy
 Department of Electrical Engineering
 Department
 Netaji Subhash Engineering College
 Kolkata, India
 E-mail: piu4mail@gmail.com

Ms. Swati Dey
 Department of Applied Electronics
 & Instrumentation Engineering.
 Camellia Institute of Technology,
 Kolkata, India.
 E-mail: swatidey19@gmail.com

Dr. Madhurima Chattopadhyay,
 Department of Applied Electronics
 & Instrumentation Engineering,
 Heritage Institute of Technology,
 Kolkata, India.
 E-mail: madhurima02@gmail.com

Abstract— Rural India is expected to have a rapid growth in mobile phone services in the next four years [8] as cell phones become a vital tool for a civilization for its advancement in every aspect. Government of India statistics reveals that even with the increasing urbanization and migration, 63% of India's population would still be living in rural areas in 2025[6]. Mobile phone subscribers' base has increased from 858.37 to 865.71 million between July to August 2011[9] with an even higher increase in low end subscribers and in rural areas. But there is a huge crisis of electricity distribution in maximum interior parts in India. We want to encourage the usage of solar energy to meet the accelerated demand of mobile phone users. Solar energy is the only form of inexhaustible, non polluted energy source whose costs are continuously declining compared to the other forms of energy. In this present paper a cost effective hardware model of universal mobile charger is presented which consists of a 12-V Saint Gobain Photo Voltaic (PV) solar panel along with a conversion circuitry and a multijack based on the universal standard.

Keywords- *Rural India, Mobile phone subscribers, Solar PV cell, Universal mobile charger.*

I. INTRODUCTION

The world is in the era of a mobile revolution. Unlike the early adopters of mobile telephony, who were elite and comparatively high class professionals, current emerging adopters are from many interior rural parts of our nation. Mobile phones can be used to deliver a better civilized mankind by digitally empowering citizens among every class of people across any nation. The rural segment services includes selling and procurement information and support for farm commodities, educating farming community on best practices, delivery of healthcare and education to remote villages via the mobile network. Here we can see the growth of rural mobile phone users from the graph shown in Figure 1..

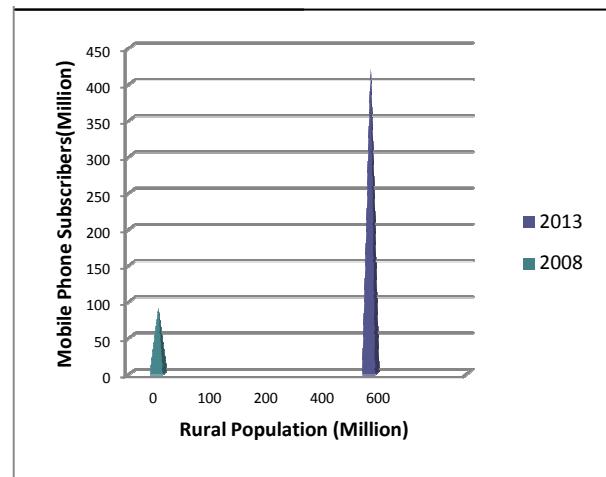


Figure 1: The Growth of Rural Mobile Subscribers

The main problem is access to electricity is scarce or even absent in most of the remote locations and villages. In some interior places, people have to travel for hours to charge their mobile phones where they get access to electricity. That clearly shows the need of cell phones among them. It has been reported that most of the rural portions of our country suffers from extraordinarily power cuts ranging from six to eight hours per day. In some cases, electricity is available for a certain number of hours in a day [7]. So a huge number of Indian populations (as seen in the graph in Figure 2) are under darkness of electricity crisis. A huge area of total population of various developing countries like India is suffering from this crisis. The sky blue line indicates that minimum 72% of electrification rate is essential while the red bordered highlighted country India is one of those countries. Each different colour represents a country, listed in right hand side.

2nd Annual International Conference IEMCON 2012

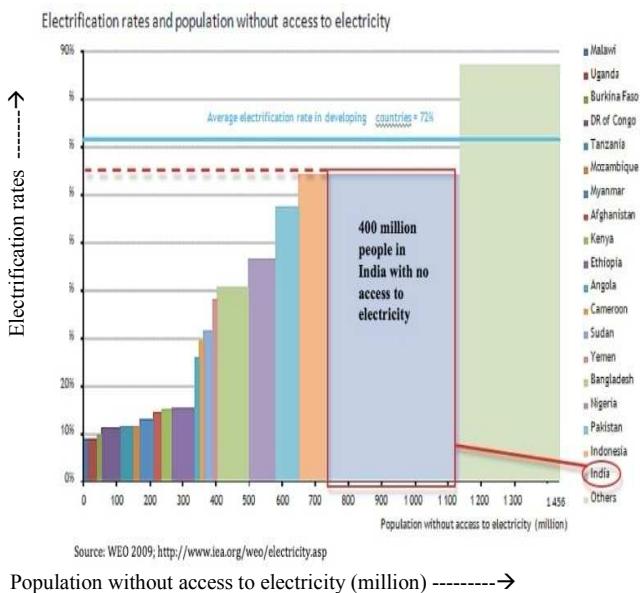


Figure 2: Electrification rates in India in comparison with the rest of the world

Solar power is the most acceptable alternative energy source as the solution for this problem because it is provided with the characteristics of cleanliness, low temperature discharging and pollution free as well as inexhaustible. Indian government is also encouraging the use of these solar energies in various useful fields by granting subsidies. Here from the graph we can get the idea of growing solar energy demand in India.

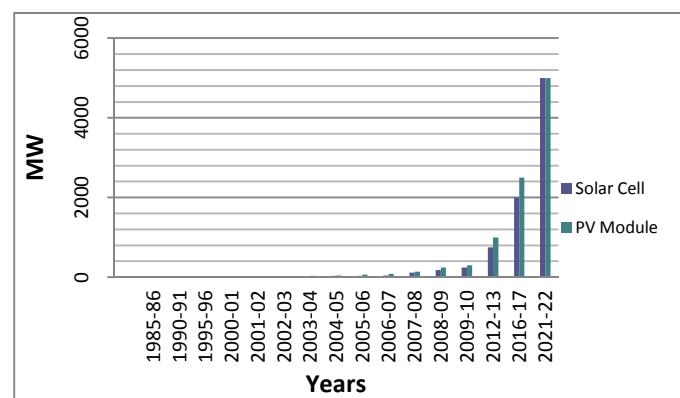


Figure 3: Growth in PV solar cell

Photovoltaic first came into use in 1958 when NASA needed a feasible power source for its space crafts and satellites, and has been used for this purpose ever since. Other current uses of PV solar panels include powering watches and pocket calculators, powering the lamps of some remote lighthouses. India is in a good position for utilizing solar power because of the intense heat. Some regions receive plentiful solar radiation like 6.27 kWh/m^2 per day (Green India Resources). Keeping this in mind we have come with the concept of cost effective

universal mobile phone charger with the use of solar PV cell which will help our rural people to use mobile phones. This charging circuitry is notably diverse, having a wide variety of DC connectors which is compatible with different models of available mobile phones from different manufacturers. This is portable as well as can be used in long travel too.

Previous works on this particular field shows the different projects like T. C. Cheng, J. K. Kuo told about improvement of solar cell based photo voltaic charger system [3]; In 2007, S. S. Rani and company had proposed about mobile communication using solar energy resource, where we can see the conventional electrical charger could be replaced [2], D. Casini & G. Marola [5] explored the way to detect a true fall state of charge avoiding false charge termination due to changing illumination conditions. In the year 2009, Ke Liu and John Makaran, [1] have given the idea of simulation based results of charging a battery through solar panel. Till now research is going on in this particular topic and few companies like nokia, solio, cellboost.etc. have taken initiative to bring it into market.

Keeping that in mind, we have came into the concept in designing such a cost effective circuit which will be able to charge a wide range of cell phones time through solar energy which will be very much useful for the electricity scarce areas of our nation.

II. SYSTEM DESIGN

With the view of previously discussed needs and concepts we have emphasized on few factors while designing the model. The model needs to be cost effective and hazard free; the circuit components should be easy available in the market and simplified enough to be easy repairable; portable and universal, suitable for almost all kinds of market available cell phone models. The overall block diagram of the system is shown below which consists of:

- (i) Solar Panel- from which the solar energy converts into electrical energy and fed it to the (ii) Conversion Circuitry – where the output voltage of the solar panel regulated and maintained in a unidirectional flow and next it is opened as
- (iii) Universal charging outputs - from there finally the (iv) Mobile phones - Battery can be charged.

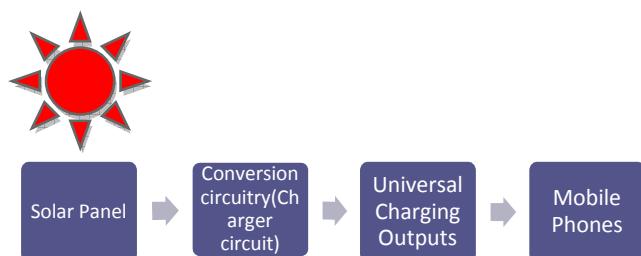


Figure 4: The general overview of the system

The system specifications are shown below:

| S/L Nos: | Components | Specifications |
|----------|-------------------------------|------------------------------|
| 1 | Solar Panel | 12V, 5.6 Watts; Saint Gobain |
| 2 | Voltage Regulator (LM 317) | 12V-37V, 1.5A |
| 3 | Diodes: D1 & D2 | IN 4001 |
| 4 | Transistor | BC 548 |
| 5 | Zener Diode (BZX 83 C series) | 6.8V, 1W |
| 6 | Variac | 1KΩ, 1W |
| 7 | Resistances: R1; R2; R3 | 180Ω, 1W; 1KΩ, 1W; 10Ω, 1W |

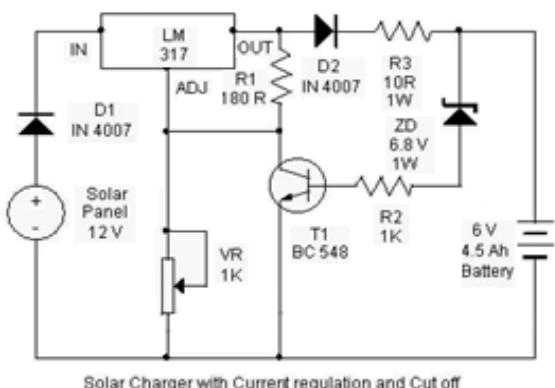


Figure 5: The Circuit Diagram

Circuit Description: The circuit uses a 12V solar panel and a variable voltage regulator IC (LM 317). The solar panel consists of ten solar cells, each rated at 1.2 volts. 12V DC is available from the panel to charge the battery. The charging current passes through D1 to the voltage regulator IC, by using it's adjust pin the output voltage and current can be regulated. Variac is placed between the adjust pin and ground to provide an output voltage of 9 volts to the battery. The resistor R3 restricts the charging current and diode D2 prevents discharge of current from the battery. The transistor T1 along with the zener diode ZD acts as a cut off switch when the battery is full. Normally T1 is off and battery gets charging current. When the terminal voltage of battery rises above 6.8 volts, zener conducts which provides base current to T1. It then turns on grounding the output of LM 317 to stop charging.

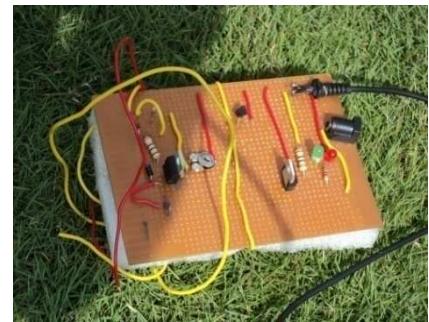


Fig 6: The final hardware circuit (on veroboard)

The Figure 6 is the outcome of the circuit designed on veroboard. It will take the power of 12V solar panel and will pass the regulated output to the universal connectors. The charging sockets are connected to mobile phones finally, and the parameters are measuring through the multimeter as shown in Figure 7.

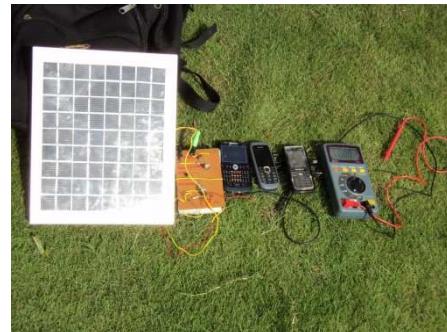


Fig 7 : Circuit is connected with mobile phones

III. RESULT ANALYSIS

One of the most important parts of a project is the result analysis. In order to verify the output voltage with time we have studied the circuitry for a whole day. On 15th April, 2011, we have started to get some measurable output current from morning 6 am onwards. In Figure 8, the current reaches its peak at 1 pm and goes down gradually till the sun sets off

and here the zero along abscissa indicates the 12 O' clock at night.

Here the analysis of results is done by noting down the voltage and current at the output of the solar panel. As the sun intensity varies throughout the day with time, it is always better to measure the voltage at the output terminal of the solar panel and note down the average. The output is shown as below:

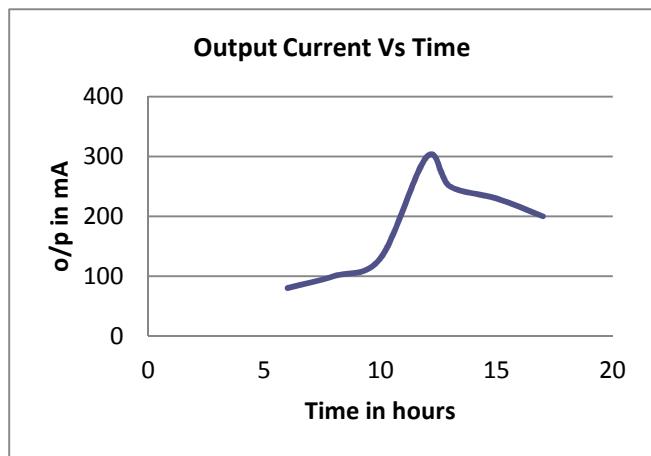


Figure 8: Output current vs. Time Graph

From Figure 8, we can get to see that the solar panel output is steadily increasing from morning to the middle of the day. It reaches at the top position when the Sun shines its best, then it got decreased gradually with sun set.

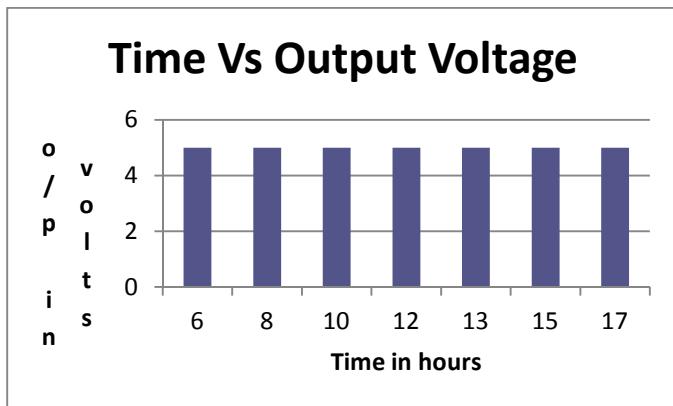


Figure 9: Output Voltage vs. Time Graph

Meanwhile we can see in Figure 9, that the voltage output remains nearly constant throughout the day. The output voltage has been measured in a gap of two hours i.e. we have taken the voltage first in the morning at 6:00 and then at 8:00 and so on.

IV. CONCLUSION

This paper is about our approach to provide our rural Indians a better way to use mobile phones by charging those with solar power in a pocket friendly cost (around Rs.- 350/-). We have used a 12 volt solar panel, a connection circuitry and universal outputs for various available mobile phone models. It can charge multiple mobile sets and spares are easily available too. So that a number of users can be benefited from a single charger set. Therefore, people can use mobile phones even where electricity has not available till date. We can use it whenever alternative power is essential too because of its portability. In future, we have plans for further development in solar powered instruments.

References

- [1] Ke Liu and John Makaran, "Design of a Solar Powered Battery Charger", Oct 2009 IEEE Electrical Power & Energy Conference(EPEC), Pages1-5.
- [2] S. Sheeba Rani, P.M. Beaulah Devamalar, R. Maheswari "Mobile Communication using Solar Energy Resource", IET-UK International Conference on Information & Communication Technology in Electrical Sciences (ICTES 2007), Dr. MGR University, Chennai, Tamil Nadu, India, Dec 20-22,2007,pp.277-281.
- [3] T.C. Cheng, W.C. Chang, K.F. Yarn and J.K. Kuo, "On The Improvement of solar Cell Based Photovoltaic Charger System", 23-26 oct. 2006, pp.986-988.
- [4] F. Clement, M. Neidert, A. Henning, C. Mohr, W. Zhang, B. Thaidigsmann, R. Hoenig, T. Fellmeth, A. Spribile, E. Lohmueller, A. Kreieg, M. Glatthaar, H. Wirth, D. Brio, R. Prieu, M. Menkoe,, K. Meyer, D. Lahmer and H. J. Korokoszinski, "Industrially feasible multi-crystalline metal wrap through (MWT) silicon solar cells exceeding 16% efficiency" 17th International Photovoltaic Science and Engineering Conference, *Solar Energy Materials and Solar Cells*, Volume 93, Issues 6-7, June 2009, Pages 1051-1055.
- [5] D. Casini and G. Marola, "Solar Battery Charger for NiMH Batteries" SPEEDAM 2008, International Symposium on Power Electronics, Electrical Drives, Automation and Motion,pp.1146-1148.
- [6] Ablett, J., Baijal, A., Beinhocker, E., Bose, A., Farrell, D., Gersch U., Greenberg E., Gupta, Shishir, Gupta, Sumit. The 'Bird of Gold': The rise of India's Consumer market. McKinsey & Company, May 2007(www.mckinsey.com).
- [7] Aditya Dev Sood, "The Mobile Development Report", Center for Knowledge Societies(www.cks.in).
- [8] www.asiathisweek.com
- [9] Courtesy: Telecom News Statistics

Energy Consumption Modeling for Maintenance Cost Evaluation in RFID Inventory Tracking

Tanvi Agrawal¹
 Dept. of ITM, NITIE
 Mumbai, India
 agrawal_tanvi@rediffmail.com

Abstract -RFID technology is being widely adopted in various domains e.g. Supply chain Management Inventory tracking, retail environment, public transportation, and security systems to name a few. The inventory management systems involving large number of RFID tags, multiple tags transmit the data simultaneously. One of the key challenges being faced by such systems is loss of energy on account of collisions during simultaneous data transmission. This leads to reduced energy efficiency and in turn a higher operational cost of RFID system. The problem is further aggravated for the systems involving portable hand held/mobile readers. Operational cost being a primary contributor for the Total Cost of Ownership in RFID applications , hence reduction in the former has a direct correlation towards reduction in the efficiency of the system .One of the methods to address this challenge is to use the energy efficient anti-collision protocols. The algorithm reduces the energy consumption in the system resulting in reduced operational cost and in turn enhances the battery life of RFID sensor equipment. This paper discusses energy consumption in RFID applications specifically in the area of inventory tracking with primary focus on operational cost optimization using the energy efficient anti-collision protocols

Keywords— RFID, Transceiver, Energy cost, Radio Model, Anti-collision algorithms, UHF

I. INTRODUCTION

Radio Frequency Identification (RFID) systems are a form of sensor networks that are used to identify physical objects [11], using peer to peer communication mechanism (means that each node is behaving as an information source and sink) [37]. System efficiency is highly dependent on the maintenance costs involving the hardware replacement costs . Hence it is vital to identify reasons leading to need of maintenance related costs

Pradip K Biswas² and A.D. Raoot³
 Dept. of ITM, NITIE
 Mumbai, India
 pradip2468@gmail.com and adraoot@gmail.com

and methods to eliminate the same .In the applications involving portable battery operated radio equipment, collision of the data bits during transmission leads to repetitive transmission and in turn need for a frequent battery replacement and enhanced energy need. As the energy drawn from the battery is inversely related to the battery life [23], the above mentioned collision in the data transmission leads to diminishing battery life. Hence with the application of Anti-collision protocols during design and development of these systems battery life expectancy can be enhanced which in turn can save the battery replacement cost or Recurring cost in RFID applications .

The promising applications of RFID technology is the large scale inventory management needs . These large scale RFID systems require simultaneous answering of tags on shared channel between reader and tags make channel prone to collision [18].Though both active and passive tags can be used in such applications, generally passive RFID tags are being preferred due to low cost [9]. Apart from the commercial difference on the cost the key difference in the passive and active tags is the lack of battery source in the former.Passive tag draws power from the reader's continuous wave and needs to communicate continuously with the reader to exchange commands synchronously. This requires the tag to be on for a longer time adding pressure on battery life of reader [10]. To reduce this load on the energy requirement , anti-collision protocol are being used which operate on the principle of enable the information exchange with the tag with minimal communication and computation overheads needs [16].The performance of these anti-collision algorithms is directly measured by the data transmitted between reader and tags[33]. As the amount of power consumed is influenced by the total

number of replies sent by each of the tags, an efficient protocol will minimize the messages between the tag and tag reader [5]. Networking used in RFID is Wireless LAN. There are three ways to enhance the energy efficiency of wireless Sensor Networks [11].

- i. Low duty Cycle Operation
- ii. Reduce data Volume
- iii. Multi Hop networking to reduce requirement for long range transmission

In the above mentioned 3 options the reduction in data volume can be achieved by minimizing the amount of data bits being transmitted by the tags, which has a negative correlation with the data being lost on account of collision during transmission. These collisions are being minimized with the application of efficient Anti Collision Algorithm.

In this paper we are presenting the Energy consumption modeling for Energy cost evaluation which can be used further for maintenance cost reductions. Maintenance cost and energy consumption in RFID applications. We will be presenting an overview of RFID systems in section II, followed by need of energy efficiency in Section III Cost Model in Section IV followed by conclusion.

II. OVERVIEW OF RFID

Radio-frequency identification (RFID) is a technology that uses radio waves to transfer data from an electronic tag, called RFID tag or label, attached to an object, through a reader for the purpose of identifying and tracking the object. In India allocated band for RFID application is 865-867 MHz, which is UHF band. The technology offers automated identification and data capture, which has significant advantages over the conventional tracking technologies, bar codes [33]. Some of the key benefits which RFID technology offers are given as followed.

- i. RFID Tags can be read rapidly in bulk to provide a nearly simultaneous reading of contents, such as items in a stockroom or in a container.
- ii. RFID Tags can be read in no-line-of-sight conditions (e.g. inside packaging or pallet).
- iii. RFID Tags are more durable than barcodes and can withstand chemical and heat environments that would destroy traditional bar-code labels.

- iv. Tags can potentially contain a greater amount of data compared to barcodes.
- v. Tags do not require any human intervention for data transmission.

As mentioned above the system comprises of readers (Transceivers) and tags (Transponders) which communicate via Ultra High Frequency (UHF) band. The RFID tag includes a small RF transmitter and receiver. An RFID reader transmits an encoded radio signal to interrogate the tag. The tag receives the message and responds with its identification information. Many RFID tags do not use a battery, these are called passive tags . Instead, the tag uses the radio energy transmitted by the reader as its energy source. The communication between a reader and tags is half duplex. The reader initiates an inventory round by sending an unmodulated carrier signal to “energize” the tags. The unmodulated carrier (CW) is sent throughout the inventory round to keep the tags alive. Once the tags are powered up, the reader initiates the communication by modulating its carrier. Generally these passive RFID tags are being preferred for Inventory management purposes due to the lower cost and higher life compared to the other kind of tags having their own energy source called active tags.

RFID reader can be categorized in different types

- i. Handheld reader
- ii. Mobile Reader
- iii. Fixed Reader.

For inventory tracking generally handheld / mobile Readers are being preferred. In this paper we investigate the different anti-Collision algorithms which are based on UHF passive RFID.

III. NEED FOR ENERGY EFFICIENCY IN ALGORITHMS

The two most frequently encountered and measurable metrics of an algorithm includes transmission size and the performance per watt and the total energy, consumed by the chosen hardware implementation, during its required useful lifetime. According to [28] electricity costs impose a substantial strain on the budget of data, so energy-efficient algorithms are the current requirement, as the energy cost to run equipment has grown to be a

major factor. Google engineers, maintaining thousands of servers, warned that if power consumption continues to grow, power costs can easily overtake hardware costs by a large margin. In paper [22] it is stated that RFID creates huge volumes of data that are difficult to manage, according to an estimate an RFID system could generate 10-100 times the data of conventional barcode systems, causing a huge increase in the daily volume of data on the corporate IT system. So there is a tradeoff between Energy consumption and the amount of data collected. Hence to design Energy Efficient Algorithms is the need in RFID as energy consumption as is critical in battery-operated devices.

IV. COST MODEL

The Total Cost of ownership (TCO) of RFID application is depended on following components

- i. RFID hardware middleware and Software (initial deployment costs)
 - ii. RFID Tag Cost
 - iii. Operational (recurring) costs/Maintenance Costs

Operational Cost includes the Maintenance cost involving periodic replacement of batteries/nodes in wireless devices (reader)[2]. Total Maintenance cost C_m of servicing a single node s given by Andre Barroso et al. is given as below.

$$C_m(s) = C_c(s) + C_p(s) + C_a(s) + C_s(s) \dots \dots \dots \quad (I)$$

- **Cross-operation cost (cc(s)):** Cost associated with the infrastructure necessary to service nodes.
 - **Pre-operation cost (cp(s)):** Cost associated with organizing a maintenance operation.
 - **Access cost (ca(s)):** cost associated with one-time resources spent while accessing the sensor to be serviced.
 - **In-situ cost (cs(s)):** Cost associated with one-time resources spent while servicing an individual sensor in its current location in the sensor field. In situ-costs includes the battery and hardware replaced. In the above three costs, Cross Operation cost, Pre operation cost and Access Cost are not related with the hardware replacement costs only the in situ cost deals with the replacement costs.

In [19], through example it has been shown that the maintenance cost is correlated positively with the power consumption of the product and gave the different attributes for maintenance cost, which also includes the lifetime and use time. For battery operated devices whenever the charge remained is, 10% of the initial charge, a maintenance operation is triggered, every maintenance operation incurs a maintenance cost defined by equation (I). During the lifetime of a battery operated device if I maintenance operations are taking place then the total maintenance Cost C_t is given as:

$$C_t(P) = \sum_{i=0}^t C_m$$

Maintenance cost formula given by [20]is given as

$$C_{Maintenance} = [(LC_{Fixed} + (L_T * L_R + C_R))F_R$$

Where

$$C_{\text{Maintenance}} = MC \text{ (corrective)} (\$)$$

$LC_{Fixed} = \text{Fixed labor cost } (\$)$

L_T =Labor time (h)

L_R =Labor rate (\$/h)

C_R = Replacement cost of parts or materials (\$)

F_R =Failure rate

Lifetime formula to calculate the life of battery in terms of hours is given by [2] as number of hours (h) = $(V \cdot mAh) / (\text{Energy Consumed})$. So by calculating the lifetime and remained charge status we can state the requirement for maintenance operation to be triggered or replacement for RFID reader battery. For Wireless devices such as RFID reader we can evaluate the remaining charge through the energy cost evaluation of the algorithm for the battery, which will give us the idea of maintenance cost by summing up all the Operations.

V. ENERGY COST MODELING

RFID systems are classified as radio systems, because they generate and radiate electromagnetic waves [18]. Radiated electromagnetic Energy (Power) creates major concern in case of large RFID systems and becomes critical for RFID readers [33]. Radio energy consumption Model given by [27] is given below.

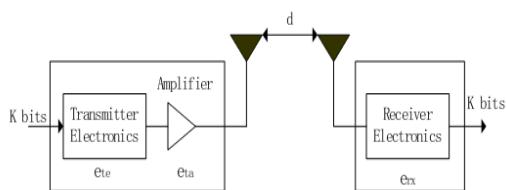


Figure 3. Radio energy consumption model

According to [30] energy consumption model is given as

$$E_{tx} = E_{circuitry} * k + E_{amplify} * k * d^c$$

$$E_{rx} = E_{circuitry} * k$$

- E_{rx} is the energy used while receiving data.
- $E_{circuitry}$ is the energy needed to feed the TX/RX circuits.
- $E_{amplify}$ is the energy required to amplify the TX signal.
- k is the length of the packet to be transmitted.
- d is the distance between source and target.
- c is the path loss exponent.

A. Energy consumption analysis at Circuit level

As RFID is a radio network energy consumption factors in radio environment are given as follows

- Energy consumed by radio transmission
- Energy consumed by radio reception
- Energy consumed by other components (energy required for circuit to be energized)

At circuit level we need to analyze the transceiver and transponder for energy consumption analysis for sending and receiving the command packet format. As we assume that transceiver is the most critical in terms of energy consumption as it has to send the continuous waves to the passive tags .According to [34] the transceiver works in three states: (i) Active state when the signal is transmitted (ii) Sleep state when there is no signal transmission, (iii) Transient state, when the transceiver switches from sleep state to active state, and vice versa. The total energy consumption is given by

$$E_{total} = P_{on} * T_{on} + P_{transient} * T_{transient} + P_{sleep} * T_{sleep}$$

Here as we are talking about the Passive RFID our consideration is more towards active state energy consumption (defined by on state of reader),as reader has to send the Continuous wave due to absence of energy source in Passive tags. So only Reader on period Power consumption defined by ($P_{on} * T_{on}$) is considered rest transient state and sleep state will not be considered because of continuous working of Reader .Energy Cost Calculation in RFID requires investigating fully, the circuit level power consumption in sending corresponding protocol level commands. So we need to focus on the number of components and their radio energy consumption model so that we can find energy consumed at the component level in the circuit, while transmissions are taking place in the form of command packets. So first we need to analyse the number of commands sets between the reader and the tags. To understand the basic energy consumption we need to understand the number of commands sent and energy consumed in sending the command packet by associating it with the power consumed at the circuit level. At circuit level the Transceiver and the Transponder needed to be analyzed for radio energy consumption calculation. Here the Transceiver energy consumption model to analyze circuit level energy consumption as it is assumed that Transceiver plays critical role in energy consumption perspective. Transceiver model presented by [30] is given below based on which we will be analyzing the transceiver energy consumption in RFID scenario.

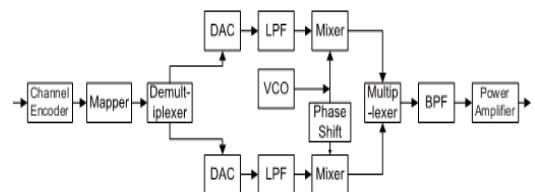


Figure 4. A typical Transmitter Structure

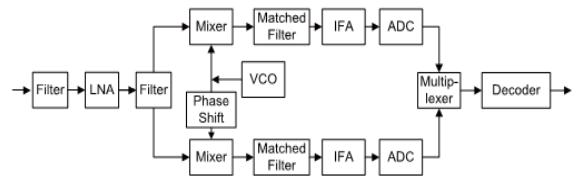


Figure 5. A typical Receiver Structure

According to [29] the power consumption of the circuit components of transmitter and receiver combined in Transceiver is given as

$$P_c = 2 P_{\text{mixer}} + 2 * P_{\text{syn}} + P_{\text{filter}} + P_{\text{DAC}} + P_{\text{LNA}} + P_{\text{ADC}} + P_v$$

Where P_{mixer} is the power consumption of the mixers, P_{syn} is the power consumption of the frequency synthesizer, P_{filter} is the power consumption of the filters, and P_{LNA} is the power consumption of low noise amplifier P_{DAC} and P_{ADC} represent the power consumption of the DAC and the ADC, respectively. P_v is the power consumption of the Viterbi decoder [29]. Below is given the commands sequence according to EPC class 1 Gen 2 specifications, between reader and the tag, while inventory is taking place. First Reader sends the Select command to select particular set of tags and then reader sends the Query command to the tags for identification purpose and then if single tag sends the RN-16 in return then no collision if multiple RN-16s are sent through the tags then collisions takes place and reader sends Queryrep again and again until the collision is detected, and if any tag get identified then the ACK command is sent otherwise NAK if received EPC is invalid. Meanwhile between command packets readers sends the unmodulated carrier waves also known as continuous waves.

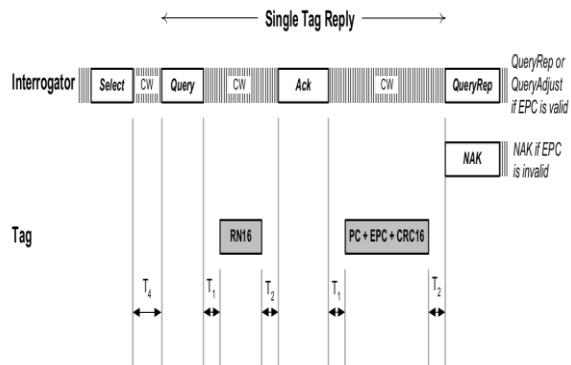


Figure 6. Interrogator and Tag commands sequence (Single Tag Reply or No Collision case)

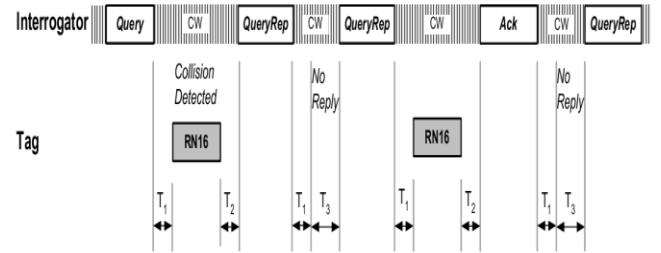


Figure 7. Interrogator and Tag commands sequence (Multiple Tag Reply or Collision case)

B. Energy consumption analysis at protocol level

At protocol level to understand the energy consumed at the protocol level we need to know the command sequence and the command packet format. So that we can find the energy consumed in each packet by associated power consumed at circuit level multiplied by corresponding timing required in sending each command packet. Flow chart overview and Proposed Command packet format is given below for consideration of command Packet transactions between reader and the tags according to EPC class 1 Gen-2 Specifications.

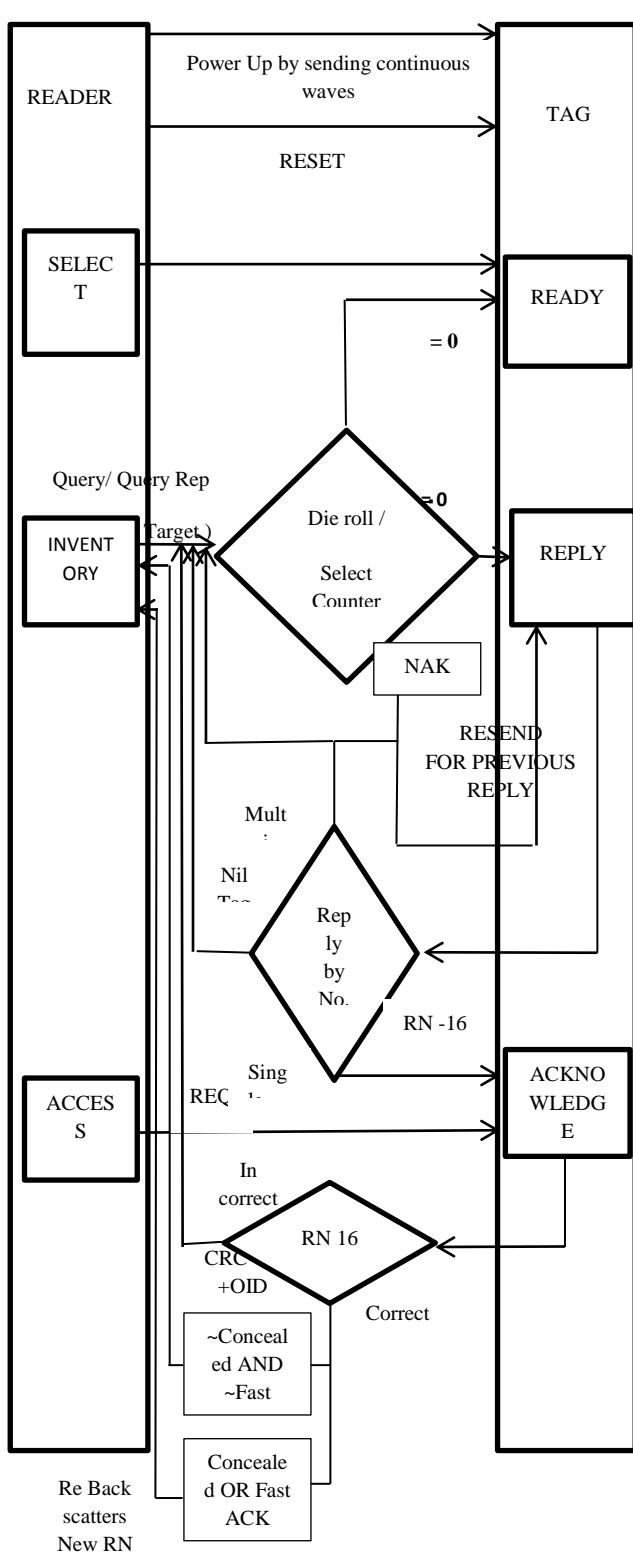


Figure-8a. Command sets Flow Chart

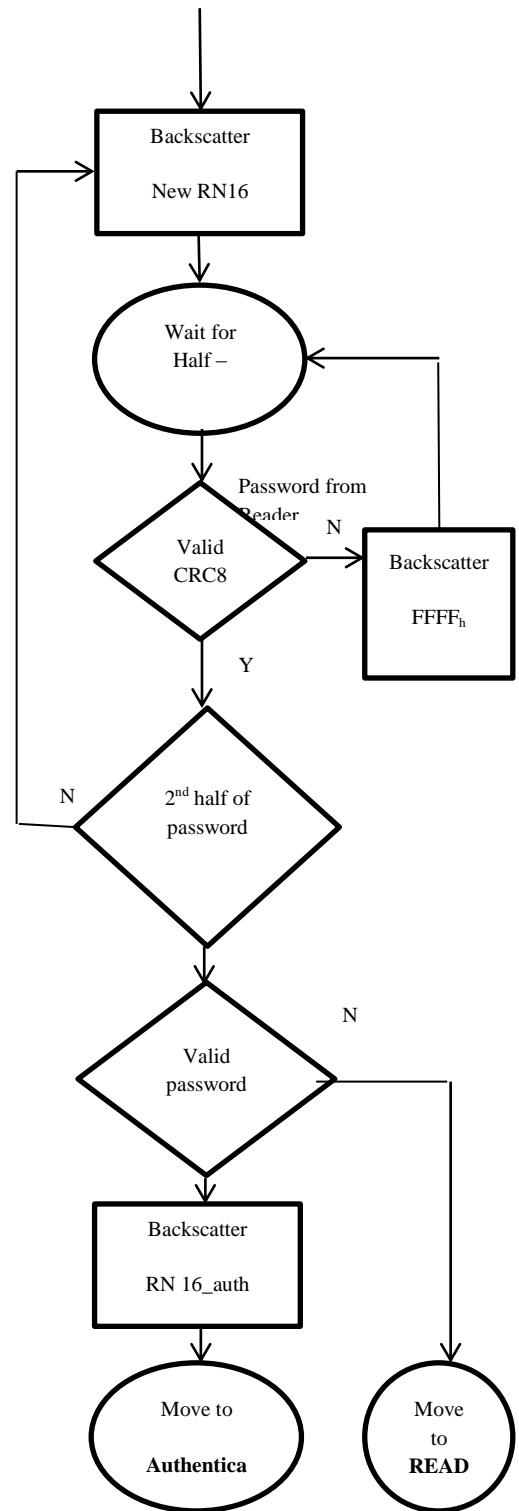


Figure-8b. Command sets Flow Chart (contd ..)

2nd Annual International Conference IEMCON 2012

TABLE I. Timing/Power Required in sending the Command Packet

| Name Of Command | Time Required in sending the Command | Power Consumed | |
|---|--|----------------|--|
| TARI (Reference Timing to send the data zero) Interrogator to tag) | 6.25 μ s | - | |
| T _B , R-T(Symbol length reader to tag signaling) | 25 μ s | - | |
| T _B T-R(Symbol length from tag to reader signaling) | 3.125ms | - | |
| T _r (Rise Time) | 500 μ s | | |
| T _s (Settling Time) | 1500 μ s | | |
| T1(Time Required from Interrogator transmission to tag Response) | 7.0 T _B ,T-R | - | |
| T2(Time required from the last falling edge of last bit of tag response to the first falling edge of the Interrogator Transmission) | 3.5 T _B ,R-T | - | |
| T3(Time from tag Response to reader transmission) | 3.5 T _B ,R-T | - | |
| T4(Minimum time between reader transmissions) | 5.5 T _B ,R-T | - | |
| T _b (Time required for transmitter/ receiver to start) | 5 μ s | | |
| Preamble | Reader to ,tag(5*TARI) Tag to reader (6*TARI) | Pct | |
| | | | |
| Commands sent by Interrogators | | | |
| Interrogator to tag data rate 40 Kbps | | | |
| CW after first Query Timing will be (No collision/ Collision) | T1+T2+Reply Data Packet Timing | Pct | |
| CW after Query (When Collision is detected in Previous Query) | T1+T3 | Pct | |
| CW after(Reset) | T4 | Pct | |
| CW when noreply from tag | T1+T3 | Pct | |
| CW after(ACK command) | T1+T2+CRC-16+OID | Pct | |
| T _{Reset} (Time required in sending Reset command) | 5 TARI(Preamble)+8 TARI(Command Code)+10.5 TARI(CRC 8,01011101) | Pct | |
| T _{Query} (Time required in sending the Query command) | 5 TARI(Preamble)+2 TARI(Command Code)+1/1.5 TARI for s-s parameter+2 TARI(ID/T selected as 0,0)+1 TARI(Mode FM0) +2 TARI(x 1 data rate)+5 TARI (X Chosen as 1010)+1.5 TARI(1 for odd Parity) | Pct | |
| T _{Queryrep} (Time required in sending Queryrep command) | 5 TARI(Preamble)+2.5 TARI(Command Code) | Pct | |
| T _{ACK} (Time required in sending the Acknowledge command) | 5 TARI+4.5 TARI(Command code)+1/1.5 TARI for fast ACK/fast ACK+(TARI reference timing required to send RN-16 bits as per Round) | Pct | |
| T _{Resend} (Time required in sending the Resend command) | 5 TARI(Preamble)+4.5 TARI (Command Code)+4 TARI(Pad) | Pct | |
| T _{NAK} (Time required in sending the NAK command) | 5 TARI (Preamble)+5 TARI(Command Code) | Pct | |
| T _{Select} (Time required in sending the Select Command) | 5 TARI(Preamble)+5 TARI(Command code)+1/1.5 TARI (Action deassert S/ Assert S)+2 TARI(Scope as 0,0 for action as 0 optional)+2.5 TARI (MemBank for OID)+Tari Reference Timing required to send the EBV bits depending on the round | Pct | |
| Commands sent by tags | | | |
| Tag to reader data rate – 80Kbps(symbol period 12.5 ms) | | | |
| Tag Response for Query T _{RN-16} | 5 TARI(Preamble)+ Time required send RN-16 | Pcr | |
| Tag Response for Query rep T _{RN-16} | 5 TARI (Preamble)+Time required to send new RN-16 | Pcr | |
| Tag Response for ACK T _{OID} | 5 TARI + Reference timing required to send the RN-16 | Pcr | |
| Tag Response for an Resend | 5 TARI + Variable Response | Pcr | |

$P_{ct} = P_{filter} + P_{mixer} + P_{amp} (P_t * \beta) + P_{LNA} + P_{syn}$ (Power Consumption of circuit components in transmitting the command packet)

$P_{cr} = P_{filter} + P_{mixer} + P_{LNA} + P_{syn}$ (Power Consumption of components in receiving the command packet)

Here below is the command packet format given for the transactions to take place between reader and the tag.

Assuming there are $m-1$ unsuccessful transmissions in before a single tag identification followed by m^{th} successful transmission of a single tag, the Frame packet format in Transmission/Reception of a single interrogation cycle is given below.

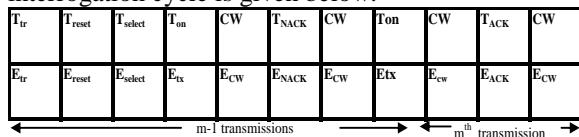


Figure 9: Proposed Frame Packet format for reader Transmissions

here $T_{on} = T_{query/Queryrep}$

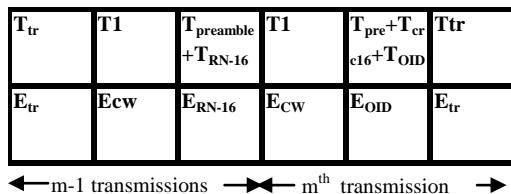


Figure 10: Frame Packet format for Tag Transmissions/ Reader Reception

Now the Transmitted and received energy consumption is given as taking the consideration of hardware components.

$$E_t(m) = (2 * E_{CW} + E_{tx} + E_{LN} + E_{RESET} + E_{Query/Queryrep} + E_{select}) * (m-1) + (2 * E_{tr} + 2 * E_{CW} + E_{tx} + E_{LN})$$

$$E_r(m) = (2 * E_{CW} + E_{OID} + E_{RN-16} + E_{RX})(m-1) + 2 * E_{tr} + E_{tx}^{ack}$$

E_{LN} is the energy consumption during the T_{ACK} period at the transmitter and the , $E_{CW}, E_{LN}, E_{reset}, E_{Query/Queryrep}, E_{select}$ are the respectively the energy consumed ,in continuous wave sending, in transmitting the Reset, Query/Queryrep,Select command packets , from reader to tag. E_{tr} is the initial energy consumption taken by the transceiver circuit

$$E_{tr} = P_{syn} * T_{tr}$$

$$E_{CW} = P_{syn} * \text{Time} \quad (\text{As per Table I})$$

$$E_{tx} = [(1 + \beta) 2BN0G_\gamma/G_c + P_{ct}] * T_{on}$$

$$E_{tx}^{ACK} = [(1 + \beta) 2BN0G_\gamma/G_c + P_{ct}] * T_{ACK}$$

$$E_{rx} = P_{cr} * T_{on}$$

$$E_{RESET} = P_{ct} * T_{RESET}$$

$$E_{QUERY/QUERYREP} = P_{ct} * T_{Query/Queryrep}$$

$$E_{SELECT} = P_{ct} * T_{SELECT}$$

$$E_{LN} = (P_{cr} - P_v) * T_{ACK}$$

$$E_{ACK} = P_{cr} * T_{ACK}$$

Here $\gamma = Pr / (2BN0)$ is known as SNR per symbol, where Pr is the received power, B is the signal bandwidth, and $N0$ is the spectral power density of an AWGN channel. It is easy to find $\gamma = f(Pb)$ and $\beta = (\varepsilon/\rho) - 1$, ε is the peak-to-average ratio, and ρ and the drain efficiency. ε and ρ are both determined by the modulation scheme used. The Total energy consumption will be calculated by summing all the transmissions and receptions for total interrogation cycles. The Total energy consumption will be calculated by summing all the transmissions and receptions for total interrogation cycles [30].

$$E = \sum_{m=1}^{\infty} [(E_t(m) + E_r(m))]$$

VI. CONCLUSION

RFID is a technology which provides potential advantages in supply chain for inventory control management. Here in this paper we have given a framework to evaluate the energy cost in transmitting the commands by considering the circuit level power consumption so based on the energy consumed we can further calculate the maintenance cost by calculating the charge remained in the battery. Maintenance is required at the stage when the 10% charge is remained. As energy consumption is related to the maintenance cost also as if we can reduce the consumed energy we can save the maintenance requirement for the battery or the replacement of the battery by saving its life.

REFERENCES

- [1] Alien Technologies, (2007) Common RFID Implementation Issues: 10 Considerations for Deployment. Available at http://www.alientechnology.com/docs/WP_RFID_Implementation_Issues.pdf.
- [2] Amit Sharma, Kshitij Shinghal, Neelam Srivastava, Raghuvir Singh (2011), 'Energy Management for Wireless Sensor Network Nodes' International Journal of Advances in Engineering & Technology (IJAET) Vol. 1, Issue 1, pp.7-13.
- [3] Andre Barroso , UtzRoedig and Cormac J. Sreenan (2003) 'Maintenance Awareness in Wireless Sensor Networks' CiteSeerX^β available at citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.101&rep=rep1&format=pdf, pp. 2-4.
- [4] AysegulSarac , Nabil Absi, Stephane Dauzere-Peres(2010), Elsevier International Journal of production Economics 'A literature review on the impact of RFID technologies on supply chain management',pp 78-79.
- [5] Building Radio frequency Identification for the Global Environment (BRIDGE) Patrick Schmitt, Florian Michahelles (2008) Economic Impact of RFID Report IST-2005-033546.
- [6] Cherian Abraham, Vinayahuja (2002) 'Inventory management using passive RFID tags: A survey',CiteSeerX^β available at <http://www.cs.rutgers.edu/~badri/553dir/papers/MCPaperFinalcoldetect.pdf>, pp 2-2.
- [7] Ching-Nung Yang, Jyun-Yan He (2011)'An Effective 16-bit Random Number Aided Query Tree Algorithm for RFID Tag Anti-Collision' IEEE Communication Letters, pp1-2.
- [8] Chiu, S. Kipnis, I. Loyer, M. Rapp, J. Westberg, D. Johansson, J. Johansson (2007), P. 'A 900 MHz UHF RFID Reader Transceiver IC, IEEE journal of Solid State Circuits, pp 2822
- [9] Chow, H.K.H., Choy, K.L., Lee, W.B., Lau, K.C.,(2006) 'Design of a RFID case-based resource management system for warehouse operations'. Expert Systems with Applications 30, pp 561–576.
- [10] Dayanasari Abdul Hadi, Norhayati Soin (2010) 'A Proposed Low Power Voltage Multiplier for Passive UHF RFID Transponder' ,IEEE International Conference on Semiconductor Electronics, pp. 397-400
- [11] Deva Seetharam, Richard Fletcher 'Battery Powered RFID' TagSense Inc.432 Columbia St, Cambridge. Available at alumni.media.mit.edu/~deva/papers/senseid.pdf
- [12] Dr. P.C.Jain, K.P. Vijaygopalan (2010) 'RFID and Wireless Sensor Networks' Proceedings of ASCNT, pp 1-11
- [13] EPC™ Radio-Frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860 MHz – 960 MHz Version 1.0.9(2005). Available at www.engr.sjsu.edu/.../AT_wp_EPCGlobal_WEB.pdf.
- [14] Guangsong Yang, Mingbo Xiao, Chaoyang Chen (2007) 'A simple energy balancing method in RFID sensor Network' IEEE workshop on Anti-counterfeiting Security, Identification Xiamen, Fujian.
- [15] Hector Gonzalez Jiawei Han Xiaolei Li Diego Klabjan (2006) 'Warehousing and Analyzing the Massive RFID Data sets' IEEE proceedings of 22nd international conference on data engineering , pp 1
- [16] Institute for prospective Technological Studies (ipts) Technical report Series , Marc van Lieshout, Luigi Grossi, Graziellaspinelli, Sandra Helmus,LindaKool, Leo Kool, Leo Pennings,RoelStap,ThijsVeugen, Bram van der Waaij , Claudio Borean(2007) RFID Technologies: Emerging Issues Challenges and Policy Options ,EUR 2270 EN pp 33-35.
- [17] Jae-Hawn Chang and LeandrosTassiulas(2000) 'Energy Conserving Routing in Wireless Ad- hoc Networks' Proceedings of IEEE Infocom Tel Aviv, Israel, March
- [18] Jihoon Myung, Wonjun Lee, Timothy K. Shih (2006) ' An Adaptive Memory less Protocol for RFID Tag Collision Arbitration'. IEEE Transactions on Multimedia.
- [19] Joshua Rhodes (2007), 'RFID networks and Protocol Standards', CPET 384 –Wide Area Networks.
- [20] Kwang-KyuSeo, Beum Jun Ahn (2006) 'A learning algorithm based estimation method for maintenance cost of product concepts' Elsevier International Journal of Computers and Industrial Engineering Volume 50 Issue 1,pp 70-73
- [21] Lei Pan, Hongyi Wu (2011) 'Smart Trend-Traversal: A Low Delay and Energy Tag Arbitration Protocol for Large RFID Systems', IEEE Transactions on Wireless Communications Vol. 10. No. 11,pp 2571 2573
- [22] Li, S., Visich, J.K., Khumawala, B.M. and Zhang, C. (2006), 'Radio frequency identification technology: applications, technical challenges and strategies', Sensor Review, pp. 193-202.
- [23] Mohsen Attaran (2007) ' RFID: an enabler of supply chain operations' An International Journal of Supply Chain Management, Vol. 12 Issue: 4, pp.249 – 257
- [24] NamboodiriVinod, GaoLixin (2010), 'Energy Aware Tag anti-collision Protocols for RFID systems', IEEE Transactions on Mobile computing.
- [25] Okkyeong Bang, Ji Hwan Choi, Dongwook Lee, Hyuckjae Lee (2009), 'Efficient Novel Anti-Collision Protocols for Passive RFID Tags' Auto-ID Labs White Paper WP-HARDWARE-050 available at pp 3-4
- [26] PrithwishBasu, Jason Redi (2004) 'Effect of Overhearing Transmissions on Energy Efficiency in Dense Sensor Networks' In Proc. International Society for Pediatric Neurosurgery.
- [27] RFID Applied Banks, J., Hanny, D., Pachano, M.A., Thompson, L.G.,(2007) John Wiley& Sons, Inc.
- [28] Ruihua Zhang, Zhiping Jia, Dongfeng Yuan(2008), 'Analysis of Lifetime of Large Wireless Networks Based on Multiple Battery Levels', International Journal of Communications, Network and System Science ,pp 136-143.
- [29] Susanne Albers (2010) 'Energy Efficient Algorithms' ACM Magazine Archive. Available at <http://cacm.acm.org/magazines/2010/5/8721-energy-efficient-algorithms/fulltext>
- [30] Tianqi Wang, Wendi Heinzelman, and AlirezaSeyedi(2008) 'Minimization of Transceiver Energy Consumption in Wireless SensorNetworks with AWGN Channels', Paper Presented at the Forty-Sixth IEEE Annual AllertonConference.USA UIUC ILLINOIS pp. 62-64
- [31] Tomas sanchezlopez, daeyoungkim, gonzaloahuertacanepa and koudjokoumadi (2009) 'Integrating Wireless Sensors and RFID Tags into Energy-Efficient and Dynamic Context Networks', The computer journal, vol. 52 no. 2
- [32] Venture Development Corporation (VDC). <http://www.vdc-corp.com/>.
- [33] Weilian Su Alchazidis, N. Ha, T.T (2010)' Multiple RFID Tags Access Algorithm' IEEE transactions on Mobile computing, pp174
- [34] Xunteng Xu, Lin Gut, Jianping Wang and Guoliang Xing (2010) 'Negotiate Power and Performance in the Reality of

- RFID Systems' IEEE international conference pervasive computing and communication, pp 1-2
- [35] Y. Li, B. Bakkaloglu and C. Chakrabarti (2006), 'A Comprehensive Energy Consumption Model and Energy Quality Evaluation of Wireless Transceiver Front ends' ,pp 262-267
- [36] Yuan-Hsin Chen, Shi-Jinn Horng, Ray-Shine Run, Jui-Lin Lai, Rong-Jian Chen, Wei-Chih Chen, Yi Pan, and Terano Takao(2010) 'A Novel Anti-Collision Algorithm in RFID Systems for Identifying Passive Tags' IEEE Transactions on Industrial Informatics ,pp105.
- [37] Volkan Rodoplu, Teresa H. Meng (1999) 'Minimum Energy Mobile Wireless Networks' IEEE journal on selected areas in Communications, Vol. 17, No. 8, pp 1333-1344.

APPLICATION FOR IT INDUSTRY

Date : 18th January,2012

Chairman : Prof. Tapan Kr. Chatterjee

Co-Chairman : Prof. B.B.Ghosh

Rapporter : Prof. Panchali Bhattacharya

| Paper ID | Paper Title | Authors |
|----------|--|---|
| 16 | Accelerating Electrical Fields in Radio Frequency Standing Wave Linacs | Vinod Babu Pusuluri and Rajat Roy |
| 20 | Mathematical Modeling of a Cylindrical Pipe Structure by Bessel Filter | Paritosh Bhattacharya, Sabyasachi Mukhopadhyay and Bankim Bihari Ghosh |
| 22 | A Study on Optimal Reactive Power Dispatch Using Artificial Bee Colony Algorithm | Tanmoy Malakar and Ashoke Kumar Sinha |
| 27 | A Novel Control Algorithm for Self Supported Dynamic Voltage Restorer | Himadri Ghosh, Pradip Kumar Saha and Goutam Kumar Panda |
| 31 | Study The Nonlinear Behaviour of PFC Boost Converter | Arnab Ghosh, Abhisek Pal, Pradip Saha and Gautam Panda |
| 35 | Control of Chaos in DC-DC Boost Converter with Non-smooth DC Supply | Subhankar Dam, Abhrajit Saha, Dr. Pradip Kumar Saha and Dr. Goutam Kumar Panda |
| 73 | Understanding some Physico-Chemical Components for Estimation of Eutrophication of an Aquatic System | Susmita Bhattacharya, Prabir Kumar Das, Dr. Phani Bhusan Ghosh |
| 96 | Two-level Secure Re-routing (TSR) in Mobile Ad Hoc Networks | Himadri Nath Saha , Dr. Debika Bhattacharyya , Dr A.K. Bandhyopadhyay, Dr. P. K. Banerjee |
| 100 | Degumming Jatropha to Make It Fit for Use As Biofuel | S.K.Das,Prabir Kumar Das, Dr. B.B.Ghosh, Dr. Satyajit Chakrabarti |
| 101 | Impact Of Bio-disel on The Transport Sector of Kolkata Megacity | S.K.Das,Prabir Kumar Das, Dr. B.B.Ghosh, Dr. Satyajit Chakrabarti |

DEGUMMING JATROPHA TO MAKE IT FIT FOR USE AS BIO FUEL

Authors:

1. S. K. Das

Institute of Engineering & Management, Salt Lake, Kolkata

2. Prabir Kumar Das

Institute of Engineering & Management, Salt Lake, Kolkata

3. Dr. Satyajit Chakrabarti

Institute of Engineering & Management, Salt Lake, Kolkata

4. Dr. Bankim Bihari Ghosh.

Institute of Engineering & Management, Salt Lake, Kolkata

Abstract: *Jatropha is a drought resistant perennial growing shrub, tolerant to extremes, suitable to tropical and non tropical climate and considerable climatic changes. It possesses the ability to reclaim the arid or semi-arid land to fertile character. Jatropha oil when blended with diesel to form a renewable fuel called biodiesel, the lubricating properties of the fuel increases. This bio-fuel can be used directly in any existing diesel engine without any modification of its accessories. Jatropha oil has got multi advantages viz.it is energy independent; pollution-free air and so less global warming; environment- friendly and ideal for heavily polluted cities and so on. However, it has got two disadvantages, it has got high viscosity; low Cetane number. These two drawbacks can be overcome through chemical process called degumming which is an appropriate chemical process to reduce the viscosity of crude jatropha oil to make it fit for diesel engine's fuel*

farmers can produce more. Farming can no longer be treated as unskilled work done by the uneducated conventional farmers. A number of institutions are devoted now to conduct research on Agronomics. Farming has entered into a new era with a new height of dimension and growth.

1.2 Jatropha is a drought resistant perennial growing shrub, non demanding, tolerant to extremes, suitable to tropical and non tropical climate and considerable climatic changes, even up to light frost. It grows considerably easily and survives producing seeds for nearly 50 years.

1.3 Impacts of Jatropha using as bio fuel covers the social, ecological and economic impacts. The overall advantage of the system is that all the processing procedure and the added value can be kept within the rural area or even within one village. A Seedling will start yielding seeds after a year of plantation. A jatropha plant requires 2m x 2m area so that 2500 plants can be grown in 1 hectare, taking into account of 20% mortality rate of nursery seeds. Flowering occurs during the wet season and two flowering peaks are generally seen. The seeds mature about 3 months after flowering. Jatropha can survive with minimum inputs and propagates easily. The seeds have an oil content of 37%. Jatropha oil can be combusted as fuel without being refined. Yield generates @ 15 million tons of seeds from cultivation of 10 million hectares

1.0 INTRODUCTION

1.1 Over the generations a farmer yields crops having inherited the occupation percolated from his ancestors. By the application of modern technology and science,

of waste land i.e. 1.5 tons per hectare. Such a yield can compensate for 1/10th requirement of mineral diesel in India.

1.4 In the coming years Jatropha oil will be very popular in India during innovation of renewable energy. According to India's Planning Commission statistics three/four years before, 130 million hectares waste land is available in the country. Out of which 33 million hectares are available for reclamation through tree plantation. The plantation of Jatropha gained importance in India through generation of awareness among the conscious citizen about the necessity of introduction of renewable energy to restrict the import of mineral oil from the foreign countries and also to conserve the natural resources. It is heartening to note that the Government of India is going to promote Jatropha cultivation in more than 4 lakh hectares to commercialize the production of bio-diesel as envisaged in the National Mission on bio-diesel. The Planning Commission has already accorded in principle, approval of Rs.50 crore Jatropha nurseries. (Report: The Statesman DT. 23.08.06.). At present jatropha are being planted in some areas viz. Erode of Tamilnadu and Bankura of West Bengal from which nuts yield good quality of jatropha oil. However, it is necessary to create good seed banks in India for future plantation of Jatropha Curcas because we have to achieve both bio fuel to maintain conducive environment and food for our survival. That is why we should not encroach agricultural land to grow bio fuel plants and bring scarcity of food. Accordingly a judicious planning is necessary to design the need of the area of land for planting jatropha in our country.



SEEDS



PLANTATION

2.1 Jatropha Curcas is a small tree or shrub with smooth grey bark which exudes whitish colored watery latex, when cut. Jatropha grows between three to five meters height. However, it can attain a height of eight to ten meters depending on favorable conditions. Jatropha can be grown in marginal soil with least or no tillage activities. Jatropha plantation can be done abundantly in the barren land because of its less moisture and manure needs. It possesses the ability to reclaim the arid or semi-arid land. Besides, its plantation can protect soil erosion and sand dunes. The plant has large green to pale green leaves. Flowers are grown terminally. Fruits are produced in winter when the shrub is leafless. Jatropha produces seeds for about 50 years with an oil content of 37% or so.

2.2 The seeds become mature when the capsule changes from green to yellow after two to four months. The by-products are seed cakes which are rich in nitrogen, phosphorus and potassium and used as an organic fertilizer. Oil can be extracted from the Jatropha nuts after two to five years of planting. The nuts yield tons of oil annually depending on the tilling activities.

2.3 The oil can be combusted as fuel without being refined. The oil burns with clear smoke free flame and an ideal fuel for use in diesel engines. The Jatropha oil can also be blended economically in the ranges of 5 to 20 percent or even more (to be designed) with the conventional fuel for use as bio-diesel in the diesel engines.

2.0 ABOUT JATROPHA PLANTS



FRUITS

3.0 PROPERTIES OF JATROPHA OIL

- (i) Jatropha oil when blended with diesel to form a renewable fuel called bio-fuel, the lubricating properties of the fuel increases. The superior lubricating properties of the bio-fuel increases the efficiency of the diesel engines compared to that of conventional fossil fuel.
- (ii) Higher flash point of bio-diesel helps the owner to store it safely. Long term storage may pose problem of bio-degradation
- (iii) The bio-diesel molecules are simple hydrocarbon chains containing no sulphur or aromatic substances as associated with fossil fuel.
- (iv) Higher content of oxygen in bio-diesel enables complete combustion of hydrocarbons
Resulting in less emission

4.0 ADVANTAGES OF JATROPHA OIL

Multiple advantages of bio-fuel can be summarized as under:-

- (i) Bio-diesel is the most valuable form of renewable energy that can be used directly in any existing diesel engine without any modification of its accessories.
- (ii) It is energy independent.
- (iii) Cleaner air and so less global warming.
- (iv) Environment-friendly and ideal for heavily polluted cities.
- (v) Bio-diesel produces 80% less CO₂ and 100% less SO₂ emission.
- (vi) It can be used alone or blended with mineral oil in a suitable ratio of 5 to 20 per cent.
- (vii) Proper degumming of oil makes the cost of bio-diesel cheaper than the conventional fuel.
- (viii) Bio-diesel extends the life of the diesel engine.

- (ix) By its use the mineral oil can be conserved for future generation.

However, crude Jatropha oil has got two disadvantages:-

- (i) It has got high viscosity.
- (ii) Low Cetane number for use in diesel engine

These two drawbacks can be overcome through chemical process. Degumming is the appropriate chemical process to reduce the viscosity of crude jatropha oil to make it fit for diesel engine's fuel. Trans-esterification is another process to reduce the viscosity of jatropha, the scope of which has not been discussed here as the write up is prepared based on research work on 'Degumming of Jatropha' followed by experiments to find the performance and emission characteristics of diesel engine run by jatropha based bio-diesel.

5.0 DEGUMMING OF CRUDE JATROPHA OIL IN THE LABORATORY

Degumming is a chemical treatment of crude jatropha oil to remove phosphates, waxes and other impurities. Degumming converts the phosphatides to hydrated gums which are insoluble in oil and readily separated as sludge.

The research work on 'Degumming' was carried out in the engineering chemistry laboratory of 'Institute of Engineering and Management', Salt Lake, Kolkata-700091

1st sample:

Crude jatropha is placed in a measuring jar for a quantity of 95 ml.

Phosphoric acid mixed @ 5% v/v for the quantity of 05 ml.

Laboratory temperature recorded at 1 PM. On date 25th April 2009 = 31 degree centigrade

The liquid mixture of 100 ml. is then placed in a conical flask, stirred and heated over an electric heater to about 50 degree centigrade resulting in slight change in color after 10 minutes. It is dark brown

almost same of the color of crude jatropha. The conical flask containing the heated mixture is allowed to cool and kept for degumming process for about seven days or so.

2nd sample:

Crude jatropha is placed in a measuring jar for a quantity of 55.5 ml.

Phosphoric acid mixed @ 7.5 % v/v for the quantity of 04.5 ml.

Laboratory temperature recorded at 1.30 PM on 30.04.09 =37 degree centigrade.

The liquid mixture of 60 ml. is placed in a conical flask and the experiment is conducted in the same way of the 1st sample and kept for degumming process for seven days or so for settling the sludge. Color seen almost same as crude one.

3rd sample:

Phosphoric acid mixed @ 10 % v/v i.e. 06 ml. to crude jatropha of 54 ml and the conical flask containing 60 ml liquid mixture is subjected to the same experiment on 30.04.09 (lab temp. 37 degree C.) and kept for degumming for seven days or so. Color almost same of crude jatropha.

After degumming for more than a week, the three samples are filtered to separate the sludge. The filtered samples each of quantity 10 ml. is separately passed through a fixed Oswald Viscometer to observe the time of flow for each sample. The viscometer is cleaned and dried after passing the 1st sample and before passing the second sample and so for the third. Distilled water of 10ml. volume is also passed through the same viscometer (after it is cleaned and dried) and the time of flow is recorded. Lab temp noted = 29 degree centigrade on the date of experiment, (14.05.09.). A special pycnometer of 20 ml capacity is chosen to weigh the distilled water, crude and the degummed jatropha at different percentages cleaning and drying each time for weighing the different solvents.

6.0 CALCULATION OF VISCOSITY OF DEGUMMED JATROPHA

Detailed calculation has been given in Appendix-A

FIG.1 SHOWING CURVE FOR VISCOSITY VS PERCENTAGE OF DEGUMMING OF JATROPHA

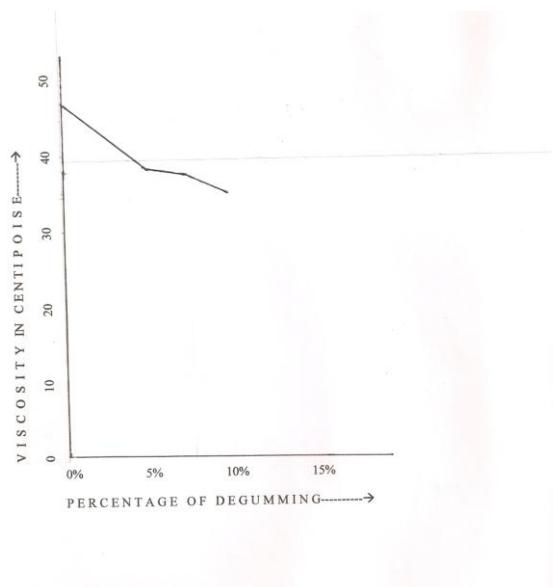


FIG.1

Viscosity of crude jatropha is 46.805 cp. By degumming with 5% phosphoric acid viscosity reduces to 37.703 cp. Degumming with 7.5% and 10%, viscosity reduces to 37.055 cp and 35.566 cp respectively. So the rate of decline of the viscosity curve is predominant in the range from 0 to 5%. The curve is almost flat between 5% to 7.5%. The reduction is not very significant at 10%. Considering all the points it is decided to blend degummed jatropha in different proportion to the mineral diesel and observe the performance the diesel engine and the emission using the blended bio-diesel as a fuel.

7.0 CONCLUSION

7.1 Degumming jatropha is an essential process for manufacturing biodiesel which is free from residual vegetable oil, gums, water, catalyst, methanol, glycerine etc.

7.2. Employment potential can be created for the rural community in the plantation of jatropha followed by picking up seeds, extraction of oil and finally degumming the oil.

7.3 Degummed oil when blended to conventional diesel can be used as a safe fuel for all kinds of diesel engines and generators with a significant reduction of atmospheric pollution. Pollution created by three wheeler vans run by the adulterated diesel or kerosene in the rural areas can be reduced by using the degummed jatropha oil directly as fuels.

8.0 REFERENCES

- [1] P. C. Roy & B. B. Ghosh: "Use of Non- Edible Vegetable Oil (Bio-diesel) in Diesel Engines, Exhaust Emission"; Proceedings of the 3rd BSME- ASME International Conference on Thermal Engineering 20-22 December, 2006, Dhaka.
- [2] S. K. Halder, B. B. Ghosh, A. Nag."Studies on the performance and emission characteristics of a diesel engine using three degummed non-edible vegetable oils" Accepted in 'Biomass and Bioenergy', Elsevier publishers, manuscript no: JBB-D-07-00151
- [3] S. K. Das, "Impact of Bio-Diesel on the Transport Sector of Kolkata" published in the Environmental journal of Institution of Engineers (India) Vol.90, Sept.2009.
- [4] S. K. Das,"Impact of Bio-Diesel on Railways" published in the Rail Transport Journal of The Institute of Rail Transport, New Delhi-110001, Vol.XIX No 1 Jan-March 2010, pp, 5-8.

IMPACT OF BIO-DIESEL ON THE TRANSPORT SECTOR OF KOLKATA MEGACITY

Authors:

1. S. K. Das

Institute of Engineering & Management, Salt Lake, Kolkata

2. Prabir Kumar Das

Institute of Engineering & Management, Salt Lake, Kolkata

3. Dr. Satyajit Chakrabarti

Institute of Engineering & Management, Salt Lake, Kolkata

4. Dr. Bankim Bihari Ghosh.

Institute of Engineering & Management, Salt Lake, Kolkata

ABSTRACT: Is it India or any other country of south-east or west, power is generated predominantly by burning fossil fuels emitting the green house gases? Impact of Climate change and the spiraling increase of cost of petroleum oil are the main reasons to shift to renewable energy fuels for which the government of all the countries has support. As a primary safeguard against excess emission and a secondary safeguard against excess consumption, the replacement of fossil fuels by renewable energy fuels of low emission level cannot be ignored. Bio-diesel which is made by blending Jatropha oil to the mineral diesel fuel in right proportion can be considered as one of the alternative transport fuels in India. The traffic pollution of the Kolkata mega city can be brought down to a tolerable limit if biodiesel is used as fuel for all the diesel vehicles of the mega city.

1.0 INTRODUCTION

1.1 The bygone days of steam engines of the 1960's and 1970's are over. Huge consumption of coal needed to run those engines can be saved now. The existing thermal power stations along with setting up of new ones can neither be stopped nor abandoned. Because the thermal source of energy bypasses all

other sources to compensate the large deficit. Is it India or any other country of south-east or west, lion's share of power is being generated by burning coal emitting the constituents of green house gases. Mining of coal is a must for the thermal power predominating over the other sources of energy. The way the miners are exploiting the coal, the reserve of coal beneath the earth will bring acute shortage to conserve for the future generations. Conservation of the natural mineral sources for the Generation Next is a social obligation for the planners of either the developed or the developing nations.

1.2 Conventional energy resources are the mineral resources i.e. fossil fuels (coal, natural gas and crude petroleum products) which can not be replenished/renewed once they are exhausted. The researchers around the world are working tirelessly to generate renewable energy resources from solar, wind, hydro, sea wave, biomass etc. It is not known exactly what amount of fossil fuel can be conserved by utilizing the non conventional resources. India needs to spend a good amount of foreign exchange for importing huge quantity of crude petroleum from other countries to meet the demand of transport fuels

in the country. So in the coming years the need of exploring edible or non edible vegetable fuels to run automobiles can not be ignored for better performance and emission characteristics of the engines. Impact of Climate change and the spiraling increase of cost of petroleum oil are the main reasons to shift to renewable energy fuels for which the government of all the countries has support.

2.0 WHY INDIA PREFERS NON-EDIBLE VEGETABLE OIL AS RENEWABLE FUEL

2.1 There are various bio-plants which yield either edible or non edible oils. The crude oils are refined to a standard quality to blend with mineral diesel in a right proportion to make biodiesel for use in diesel engines to have better performance of the engine. A right blended biodiesel with good oxygen content burn better than diesel (lesser oxygen content) during combustion in IC engines with comparatively lesser green house emissions.

2.2 There are edible vegetable oils such as sunflower, rapeseed, soybean, palm etc. used for cooking in the domestic and the commercial kitchens. These oils can be effectively blended with diesel for running IC engines. But the cost of these edible vegetable oils will be exorbitant for Indians to use as transport fuels.

2.3 The western countries use to derive renewable fuels from the vegetable edible oil bio-plant such as rapeseed, sunflower, soybean etc which they use to grow in the agricultural field. India can not afford to do so on two counts, viz. (a) the cost of the edible oil will be exorbitant for use as a transport fuel and (b) growing such plants in the agricultural land at the cost of other food crops may bring food scarcity for the 1.2 billions Indians. So India bank on non edible vegetable oil viz. jatropha Curcas, karanja etc plant oil to process for renewable transport fuel in the country.



FRUITS

SEEDS
PLANTATION

2.4 Jatropha Curcas is a multipurpose small tree or shrub with smooth grey bark. Jatropha grows between three to five meters height. However, it can attain a height of eight to ten meters depending on climatic condition. Jatropha is a drought resistant perennial plant which can be grown in marginal soil with least tillage activities. Besides, its plantation can prevent soil erosion and can fertile arid or semi-arid land. Fruits are produced in winter when the shrub is leafless. Jatropha produces seeds for about 50 years with an oil content of 37% or so. The oil can be combusted as fuel without being refined. The oil burns with clear smoke free flame and an ideal fuel for use in diesel engines. The Jatropha oil can also be blended economically in the ranges of 5 to 20 percent or even more (to be designed) with the conventional diesel for use as bio-diesel in the diesel engine.

2.5 Jatropha oil when blended with diesel to form a renewable fuel called biodiesel, the lubricating

properties of the fuel increases. The superior lubricating properties of the biodiesel increases the efficiency of the diesel engines compared to that of conventional diesel. Higher content of oxygen in biodiesel enables complete combustion of hydrocarbons and so less emissions from the transport sector. However, Jatropha oil has got two disadvantages; It has got high viscosity; Low Cetane number. These two drawbacks can be overcome through a chemical process called Degumming.

2.6 Crude jatropha oil of high viscosity either alone or its blend with diesel in different proportion, if used as a combustion fuel fail to deliver the desired performance and emission characteristics of an IC diesel engine without modification of the engine. So the viscosity is required to be reduced by degumming and neutralizing process for making the jatropha based biodiesel, an ideal combustion fuel.

2.7 Degumming is a treatment to eliminate phospholipids by thermal treatment with water and other degumming agents viz. phosphoric acid. Crude jatropha oil contains gums which is phospholipids. Therefore degumming is required to remove phospholipids and for lowering down phosphorus content. Degumming is the appropriate chemical process to remove excess phosphorus and free fatty acids (FFA) to reduce the viscosity of crude jatropha oil to make it fit for combustion fuel used in a diesel engine. Viscosity can also be reduced by transesterification using methanol in the presence of sodium hydroxide in a biodiesel plant.

PICTURES (BELOW) SHOWING DEGUMMED AND REFINED JATROPHA OIL READY FOR BLENDING WITH DIESEL



Degummed J-oil



Refined/ J-oil

3.0 GOVERNMENT POLICY ON BIO-FUEL

Government policy is, in general to expand market of biofuel production. Government all over world including those in China, India, Brazil, the US and the EU have enacted mandatory target for use of biofuel. The reasons for Government to pursue a switch from mineral oil to biofuel are threefold:

3.1 Energy Security: With price of oil at over US\$100 per barrel and future supplies uncertain, the countries are seeking alternative energy sources to increase long term energy security and minimize their currency.

3.2 Rural development: This technology has been provided great opportunity to the rural farmers for the production of biofuel in their areas. It provides better opportunity and long term security for farmers and rural employees in processing bio-fuel for marketing.

3.3 Export Development: Countries with favorable endowment of land, labor and trade condition, biofuel has got a great opportunity to develop export markets.

4.0 MANY FOLD BENEFITS OF JATROPHA BASED BIODIESEL

4.1 Environmental point of view:

- (i) Reduction of emitted particles (carbon monoxide, nitrogen oxide, sulphur dioxide etc.) from burning of fuel.
- (ii) Less deposition of green house gases in the atmosphere

- (iii) Less smoke resulting in less odour and color.

4.2 Engine point of view:

- (i) Reduces engine noise
- (ii) Reduces engine temperature
- (iii) Reduces engine vibration
- (iv) Increases engine power
- (v) Increases engine life.

4.3 Economic point of view:

- (i) Low oil consumption
- (ii) Saves money
- (iii) Direct /Indirect employment generation
- (iv) Less import of crude mineral oil
- (v) Increases foreign exchange reserve of the country.

5.0 BIODIESEL IS THE APPROPRIATE TRANSPORT FUEL FOR THE MEGA CITY OF KOLKATA

5.1 The mega city of Kolkata is facing severe traffic pollution due to its acute shortage of surface traffic access which is 6.2% only. The traffic access can be increased by demolishing the structures and buildings on either side of the traffic roads. But such demolition will invite not only legal complication but also political compulsion. So the option left with the authority is to build the traffic infrastructures either underground or over ground which should, no doubt environment-friendly traffic corridors for rail or road. In this regard the Government of West Bengal has already undertaken the construction of underground/ over ground metro rail corridors and over ground flyovers in and around the city.

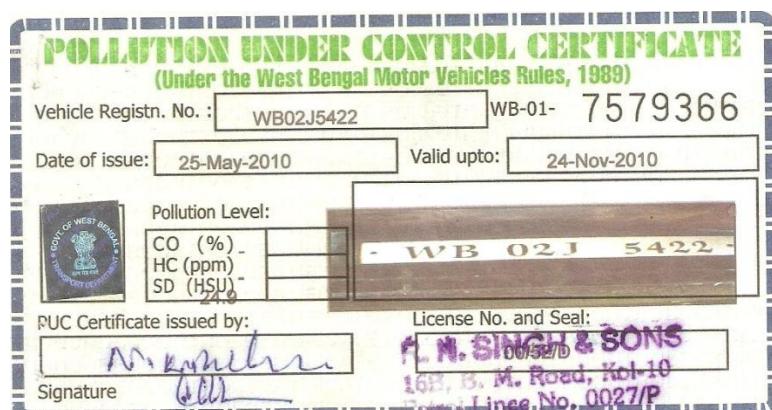
5.2 The next option left to reduce traffic pollution of the Kolkata is to run the diesel driven cars, buses and trucks of the mega city with biodiesel which can be

made by blending refined jatropha oil with the mineral diesel in a right proportion. Jatropha Curcas based biodiesel is a renewable energy fuel for diesel engines with /without minor modification of the engine depending on the proportion of blending.

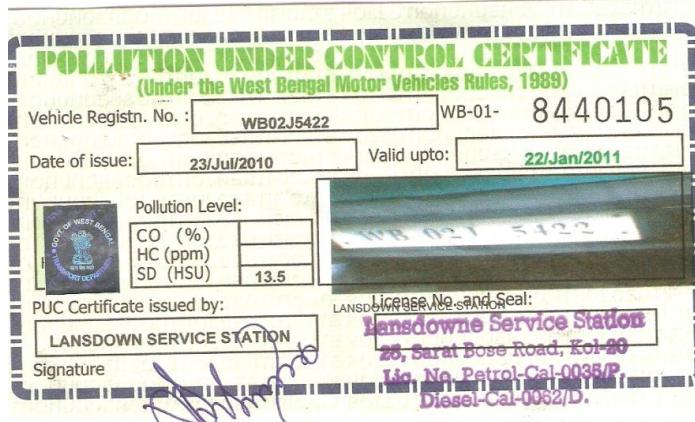
6.0 MOTORING BY A TATA INDICA (DIESEL) CAR IN KOLKATA USING JATROPHA BASED BIODIESEL PROCESSED BY “SUNPLANT AGRO LTD”(AN ISO 9001-2008 CERTIFIED COMPANY)

6.1 The author Prof. S. K. Das has the experience to run his Tata Indica diesel car with the jatropha based biodiesel of B5 formulation (5% jatropha oil + 95% diesel) in Kolkata. The author had checked the emission of his car from the emission testing center authorized by the GOWB run by diesel and the biodiesel. The results of the tests supported by the pollution certificate are as follows:

First test done on 25th May 2010 run with Mineral Diesel. Smoke Density 24.9 HSU



Second test done on 23rd July 2010 after 2 Months trial run with Biodiesel of B5 Formulation. Smoke Density 13.5 HSU.



6.2 It is apparent that pollution level has come down nearly 50% for biodiesel compared to mineral diesel. Odor of mineral diesel which was coming earlier disappeared after running with biodiesel processed by Sun-plant Agro Ltd (An ISO 9001:2008 Certified Company). 4 Dr. Suresh Sarkar Road, Kolkata-70014.

6.3 Besides, the author, using biodiesel as a fuel in his car, got better mileage leading to the saving in cost to the tune of 25% or so.

7.0 CONCLUSION

7.1 It is a social obligation for the planners of all the countries of the world to conserve the natural mineral sources for the Generation Next. Impact of Climate change and the spiraling increase of cost of petroleum oil compelled the government to shift to renewable energy fuels for which the government of all the countries has support.

7.2 The cost of the vegetable edible oil is costly for Indians for use as transport fuel and growing such plants in the agricultural land is not pragmatic (will

bring food scarcity for the 1.2 billions Indians). So India bank on non edible vegetable oil viz. jatropha Curcas, karanja etc plant oil to process for renewable transport fuel in the country.

7.3 If biodiesel is used as the combustion fuel of the entire diesel vehicles of the mega city it will no doubt reduce the traffic pollution of Kolkata to a tolerable limit.

7.4 Experiment has shown that use of biodiesel of B5 formulation (5% bio-oil +95% diesel) as a transport fuel in diesel cars brings not only reduction of atmospheric pollution but also the economy of fuel.

8.0 REFERENCES

- [1] E. M. Kondili, J. K. Kaldellis; "Bio-fuel Implementation in East Europe: Current Status and Future Prospects" Renewable and Sustainable Energy Reviews11 (2007) 2137-2151, Greece, March28-May10, 2006
- [2] Kahraman Bozbasi; "Bio-diesel as an alternative Motor fuel: Production and Policies in the European Union", Renewable and Sustainable Energy Reviews11 (2008) 542-552, Turkey, June 16-17, 2005.
- [3] P. C Roy and B. B. Ghosh: "Use of Non-Edible Vegetable Oil (Bio-diesel) in Diesel engine, Exhaust Emission", Proceedings of the 3rd BSME- ASME International conference on Thermal Engineering 20-22 December,2006, Dhaka, Bangladesh
- [4] Pandey, K. C., Pandey, K. P. ,Ghosh, B. B., and Ramanujam, S., "Critical Review on Vegetable Oils as Substituted Fuel for Diesel Engines" , The Proceedings of the XVI National conference on IC Engines and Combustion, January, 20-22, 2000, Jadabpur University, Calcutta.

[5] S. K. Das, "Impact of Bio-diesel on the Transport Sector of Kolkata" published in the journal of The Institution of Engineers (India) EN. Vol.90, September 2009, ISSN0251-110X, PP.24-26.

[6] S. K. Das,"Impact of Bio-Diesel on Railways" published in the Rail Transport Journal of The Institute of Rail Transport, NewDelhi-110001, Vol.XIX No 1 Jan-March 2010, pp, 5-8.

- [1] E. M. Kondili, J. K. Kaldellis; “Bio-fuel Implementation in East Europe: Current Status and Future Prospects” Renewable and Sustainable Energy Reviews 11 (2007) 2137-2151, Greece, March 28-May 10, 2006
- [2] Kahraman Bozbas; “Bio-diesel as an alternative Motor fuel: Production and Policies in the European Union”, Renewable and Sustainable Energy Reviews 11 (2008) 542-552, Turkey, June 16-17, 2005.
- [3] P. C Roy and B. B. Ghosh: “Use of Non-Edible Vegetable Oil (Bio-diesel) in Diesel engine, Exhaust Emission”, Proceedings of the 3rd BSME- ASME International conference on Thermal Engineering 20-22 December, 2006, Dhaka, Bangladesh
- [4] Pandey, K. C., Pandey, K. P. ,Ghosh, B. B., and Ramanujam, S., “Critical Review on Vegetable Oils as Substituted Fuel for Diesel Engines” , The Proceedings of the XVI National conference on IC Engines and Combustion, January, 20-22, 2000, Jadabpur University, Calcutta.
- [5] S. K. Das, “Impact of Bio-diesel on the Transport Sector of Kolkata” published in the journal of The Institution of Engineers (India) EN. Vol.90, September 2009, ISSN0251-110X, PP.24-26.
- [6] S. K. Das,”Impact of Bio-Diesel on Railways” published in the Rail Transport Journal of The Institute of Rail Transport, New Delhi-110001, Vol.XIX No 1 Jan-March 2010, pp, 5-8.

Accelerating electric fields in radio frequency Standing Wave LINACS.

*Vinod Babu Pusuluri, Dept. of E&ECE

IIT Kharagpur, India-721302

*Corresponding author e-mail: vinodbec@gmail.com

Rajat Roy, Dept. of E&ECE

IIT Kharagpur, India-721302.

Abstract:- Accelerating electric field in standing wave linac has been calculated using HFSS software. The consequences of this which were not available from the older calculations has also been discussed.

Keywords:- Linear accelerators, Microwave, Numerical field calculations ,High frequency structure simulator.

1. Introduction.

S band microwave linear accelerators (Linacs) have been developed by S.A.M.E.E.R [1] for use in cancer therapy. These are standing wave type of structures where the adjacent accelerating cavities are excited in opposite phases. The dimensions of the cavities are so adjusted that it takes an accelerating electron to cross the cavity in exactly the same time as it takes to reverse the phase of the electromagnetic field. Thus an accelerating electron will always find an accelerating field. The structure is shown in fig. 1a and a schematic of the cross section in fig 1b. In fig. 2 the cross section of a single cavity is shown. Numerical methods using the finite element and finite difference techniques [2,3] have been used in the past to calculate the cavity parameters such as resonant frequency and coupling to the side cavities [4]. However the electric field distribution especially in the axial direction that causes the acceleration has been largely unreported in these calculations. In this paper we use the H.F.S.S. [5] software based on the finite element method to calculate these fields. In future a better appreciation of the dynamics of the electron beam will be obtained if we know where

exactly in a cavity the acceleration takes place and subsequently any improvement in the design of the cavity shape.



Fig-1a: Linac tube.

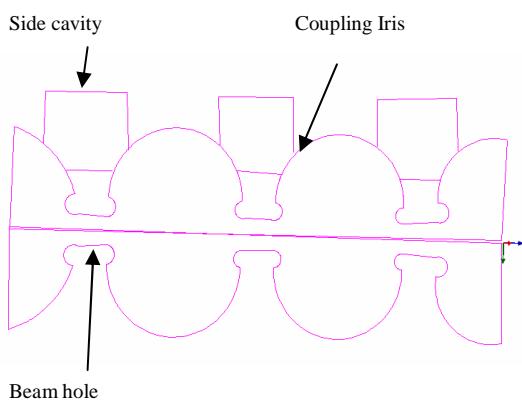


Fig-1b. Schematic of side coupled structure.

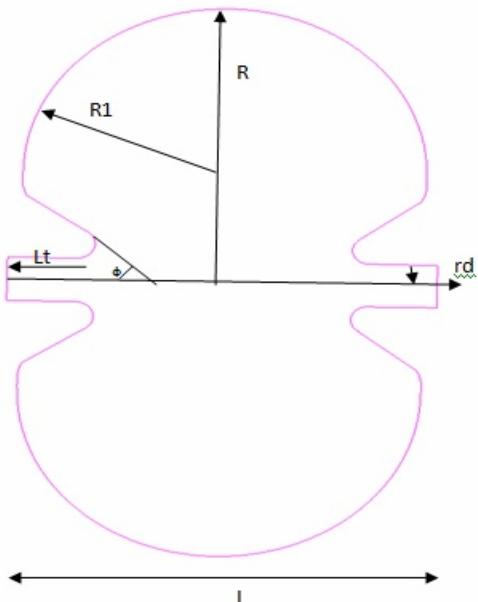


Fig-2 Cross section of single cavity.

Here the above diagram the dimensions are given in [Ref-2] those are $L=50\text{mm}$, $\theta = 30^\circ$, $rd=3\text{mm}$, $R=38.9\text{mm}$, $R1=23.5\text{mm}$ and $Lt=9.1\text{mm}$. And the overall view of the cross section of the system is as shown in fig-1b .

2. Boundary conditions used in the HFSS simulation and results.

The two dimensional structure shown in fig. 2 is rotated about the axis shown to obtain the three dimensional figure used in the HFSS simulation (see fig. 3a and 3b). This is a single cavity which has perfect electric conductor boundaries on all sides except on the axial hole through which the electron moves from one cavity to the next. Since the adjacent accelerating cavities are excited in opposite phases as already stated the assumption of a perfect magnetic conductor along the hole is justified from symmetry. The results of the computation by HFSS eigen-mode solver shows a resonant frequency of 3.01036 GHz with the dimensions shown in fig. 2. The electric field in the cavity along with magnified views of the same are shown in figures 4, 5 and 6.

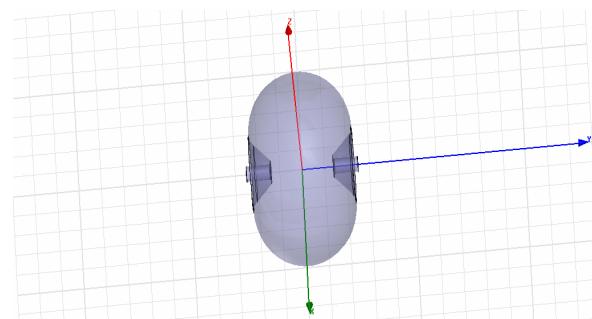


Fig-3a .Front view of the cavity.

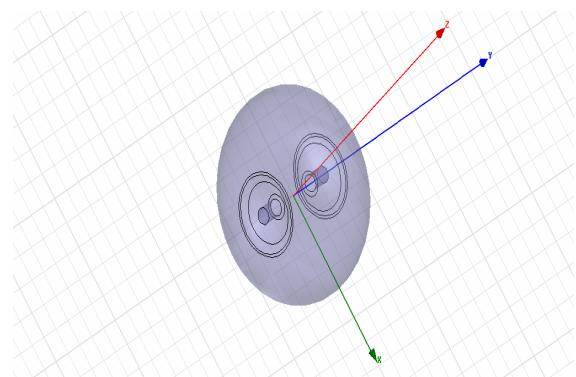


Fig-3b Side view of the cavity.

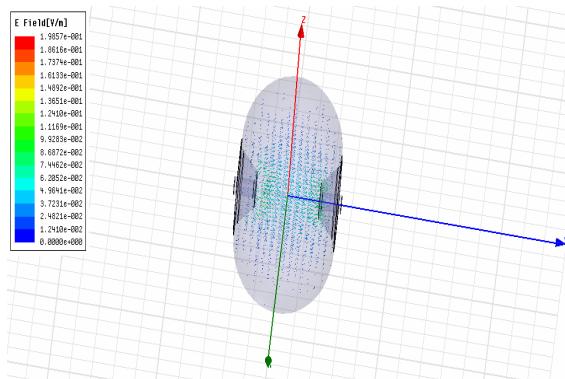


Fig-4 The field distribution inside the cavity.

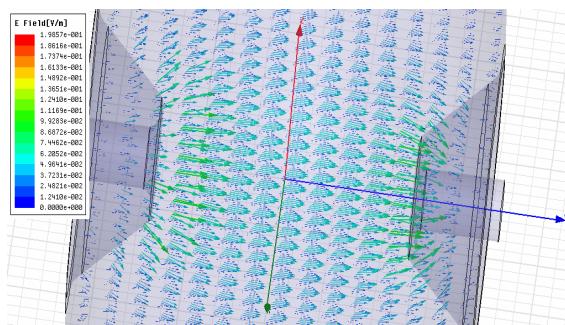


Fig-5. Expanded view of fig-4.

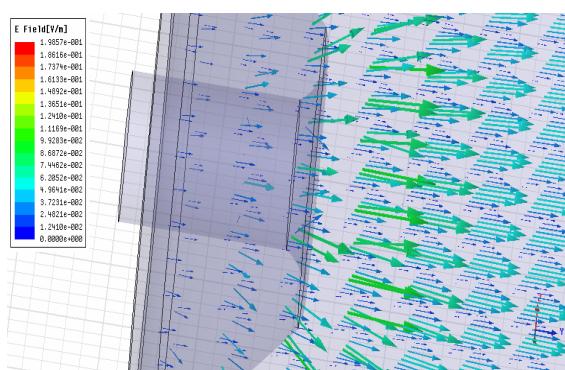


Fig-6.More expanded view of fig-4

3. Conclusions and future research.

From figs. 4,5 and 6 the maximum field is seen in the cone region of the cavity structure. This means the acceleration does not take place uniformly in the cavity and hence this can have some consequences for the electron beam dynamics. Also there is scope for changing the shape of the cavity near the conical structure. These can be the subject of future study.

References:

1. R .Krishnan, A.P Deshpande, T.S Dixit,S.Chavan,C.S Nainwad. S-band 6MeV linac,Proceedings of PAC09,Vancouver B.c Canada.
2. Rajat Roy and O shanker ,”Evaluation of Microwave linac cavity field parameters by the finite element method,” *Indian Journals of pure and applied Physics* vol.30,pp.207-210,1992.
3. S S Bhide and O Shanker , “MULA program 840-646 for LINAC resonant cavities”, *Indian J.Phys.62A(6)*, 1988.
4. Rajat Roy and o.Shanker ‘Caluculation of Inter cavity coupling coefficient for side coupled standing wave linear accelerator’ *IEEE transactions on Microwave theory and technologies* vol 41.No.6/7 ,June/july 1993.
5. Ansoft HFSS (High Frequency Structure Simulator) version-11.

Mathematical modeling of a cylindrical pipe structure by Bessel filter

Sabyasachi Mukhopadhyay
4th year, B.Tech(ECE)
CEM, Kolaghat
sabyasachi.unique@gmail.com

Dr. Paritosh Bhattacharya
CEM, Kolaghat
p_bhattacharya2001@yahoo.co.in

Prof.(Dr.) B.B.Ghosh
Professor, IEM Salt Lake
Ex-Professor IIT Kharagpur, India
bankimghosh@hotmail.com

Abstract- The suitable and developed design of a muffler helps to reduce the noise level. For this purpose it is necessary to have the Bessel filter response over a cylindrical pipe structure, which is nothing but a Infinite Impulse Response (IIR) Filter in nature. In the current work, a sound field propagation performance in a duct with circular cross section is shown to determine the sound pressure level in a muffler but at first the Bessel filter response is shown for that particular circular cross section case.

NOMENCLATURES:

| | |
|----------------------|---------------------------|
| IIR: | Infinite Impulse Response |
| r, Φ, z: | Cylindrical Coordinates |
| p: | Sound Pressure |
| k : | Wave Number |
| k_r, k_Φ, k_z : | Components of Wave Number |

Keyword – Muffler, Reactive Muffler, Bessel Function, Bessel Filter, IIR Filter.

I. INTRODUCTION

The sound wave propagation in a cylindrical tube is a fundamental and classical problem which can be applicable for cylindrical waveguides in case of the equivalent Quantum structure. So, for proper understanding of that problem Bessel function has the key role. The famous names like - Helmholtz, Kirchhoff, Rayleigh - are connected with fundamental area of the wave propagation in a cylindrical pipe structure [1]. The full Kirchhoff solution of a viscous and heat-conducting fluid in rigid circular tubes was later developed in two directions: analytical approximations of a very

complicated transcendental equations for various regions of "wide", "narrow", "wide-narrow", "very wide" tubes and an extension of the theory for non-circular tubes or higher modes in circular tubes [1, 2]. There are five different design criterions in mufflers design [3, 4]. These are Acoustical Criterion, Aero dynamical Criterion, Mechanical Criterion, Geometrical Criterion and Economical Criterion. The acoustical criterion specifies the minimum noise reduction required from the muffler as a function of frequency. Aero dynamical Criterion specifies the maximum acceptable pressure drop through the muffler at given temperature and mass flow. The Mechanical criterion specifies the materials from which the muffler is fabricated or designed [5, 6, and 7]. So that it is durable and requires less maintenance. This is especially important in case of involving high temperature exhaust or corrosive gases or the gaseous flow is carrying solid particles in suspension that might be deposited on the inner surface of the wall of the muffler and reduces the muffler effectiveness. Geometrical Criterion specifies the maximum allowable value and restriction on shape [11, 12]. The Economical Criterion is vital in the market place. A muffler must be inexpensive as possible while designing initial cost as well as operating cost must be considered. For having the successful muffler response at first the necessity of characterization of the Bessel filter is very necessary. For characterization of a Bessel Filter the knowledge about the IIR filter is very necessary [13].

II. THE SOUND FIELD IN A DUCT WITH CIRCULAR CROSS SECTION

In order to determine the sound field in a duct with circular cross section one must express the Helmholtz

equation in a cylindrical coordinate system is shown in fig. 1

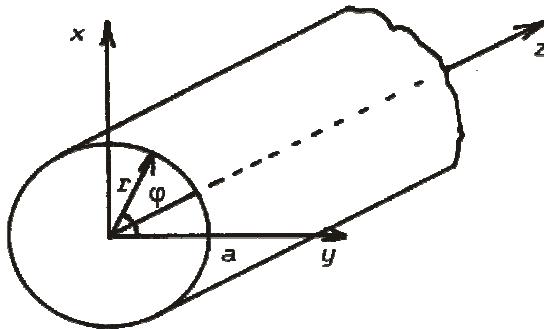


Fig.1

The boundary condition for a tube with rigid walls is

$$\left. \frac{\partial p}{\partial r} \right|_{r=0} = 0$$

It is now assumed that the sound pressure can be expressed as the product of functions that depend only on one coordinate:

$$\hat{p}(r, \varphi, z, t) = p_r(r)p_\varphi(\varphi)p_z(z)e^{j\alpha t}$$

Insertion gives

$$\begin{aligned} & \frac{1}{p_r(r)} \frac{\partial^2 p_r(r)}{\partial r^2} + \frac{1}{r} \frac{1}{p_r(r)} \frac{\partial p_r(r)}{\partial r} + \\ & \frac{1}{r^2} \frac{1}{p_\varphi(\varphi)} \frac{\partial^2 p_\varphi(\varphi)}{\partial \varphi^2} + \frac{1}{p_z(z)} \frac{\partial^2 p_z(z)}{\partial z^2} + k^2 = 0 \end{aligned}$$

It is apparent that the last term but one left-hand side is a function only of z that equals a sum of various terms that are independent of z ; therefore it must be a constant,

$$-k_z^2 :$$

$$\frac{\partial^2 p_z(z)}{\partial z^2} + k_z^2 p_z(z) = 0$$

Inserting and multiplying with r^2 gives

$$\begin{aligned} & \frac{r^2}{p_r(r)} \frac{\partial^2 p_r(r)}{\partial r^2} + \frac{r}{p_r(r)} \frac{\partial p_r(r)}{\partial r} + \frac{1}{p_\varphi(\varphi)} \frac{\partial^2 p_\varphi(\varphi)}{\partial \varphi^2} \\ & + r^2(k^2 - k_z^2) = 0 \end{aligned}$$

A similar argument now shows

$$\text{that } \frac{\partial^2 p_\varphi(\varphi)}{\partial \varphi^2} + k_\varphi^2 p_\varphi(\varphi) = 0 \dots \dots \dots (1)$$

Where $-k_\varphi^2$ is the third term of eq. (1) and therefore

$$\frac{\partial^2 p_r(r)}{\partial r^2} + \frac{1}{r} \frac{\partial p_r(r)}{\partial r} + p_r(r)(k^2 - k_z^2 - \frac{k_\varphi^2}{r^2}) = 0 \dots \dots \dots (2)$$

Equations (1) and (2) can be identified as versions of the one-dimensional Helmholtz equation with the general solutions $p_z(z) = Ae^{-jk_z z} + Be^{jk_z z}$

$$\text{and } p_\varphi(\varphi) = Ce^{-jk_\varphi \varphi} + De^{jk_\varphi \varphi} \dots \dots \dots (3)$$

Inspection of eq.(3) leads to the conclusion that k_φ must be an integer, since the function p_φ should obviously be

$$p_\varphi(\varphi) = p_\varphi(\varphi + 2\pi) \quad \text{periodic with period } 2\pi :$$

Accordingly, we shall use the symbol m instead of k_φ for this integer. Moreover, we can combine the two exponentials in eq.(3)

$$p_\varphi(\varphi) = E \cos(m\varphi + \varphi_m) \quad \text{Finaly, we introduce the transversal wavenumber } k_r :$$

$$k_r^2 = k^2 - k_z^2.$$

Equation now becomes

$$\frac{d^2 p_r(r)}{dr^2} + \frac{1}{r} \frac{dp_r(r)}{dr} + p_r(r) \left(k_r^2 - \frac{m^2}{r^2} \right) = 0$$

This differential equation is known as Bessel's equation, with general solution

$$p_r(r) = F J_m(k_r r) + G N_m(k_r r), \dots \dots \dots (4)$$

Where J_m is the Bessel function of order m and N_m is the Neumann function of order m .

Only the Bessel functions are finite at $r = 0$; therefore G must be zero.

From the Boundary condition, eq.(4) it follows that

$$\frac{dJ_m(k_r r)}{dr} \Big|_{r=a} = J'_m(k_r a) = 0 .$$

This equation can be satisfied only for certain

discrete values of k_r , denoted k_{rmn} with $n=0,1,2,\dots$. This solution can be written as

$$\hat{p}(r, \varphi, z, t) = \sum_{m=-\infty}^{\infty} \sum_{n=0}^{\infty} \Lambda_{mn} \sqrt{\varepsilon_m} J_m(k_{rmn} r) \cosh(m\varphi + \varphi_m) (p_{mn} e^{j(\omega t - k_{rmn} z)} + p_{mn}^* e^{-j(\omega t + k_{rmn} z)}) ,$$

Where $k_{zmn} = (k^2 - k_{rmn}^2)^{1/2}$.
and

$\Lambda_{mn} = (1 - m^2 / (k_{rmn} a)^2)^{-1/2} / J_m(k_{rmn} a)$
is a normalized constant. Each term in the sum
represents a mode, cf. The modes corresponding to
 $m = 0$ are axisymmetric. With

$m = n = 0$ simplifies to

$$\hat{p}(r, \varphi, z, t) = p_+ e^{j(\omega t - kz)} + p_- e^{j(\omega t + kz)},$$

which is recognized as the one-dimensional sound
field. In general case there are m modal planes
where the sound pressure is zero, and n coaxial
nodal cylindrical surfaces.

| $n \setminus m$ | 0 | 1 | 2 | 3 |
|-----------------|---------|---------|---------|---------|
| 0 | 0 | 1.8412 | 3.0542 | 4.2012 |
| 1 | 3.8317 | 5.3314 | 6.7061 | 8.0152 |
| 2 | 7.0156 | 8.5263 | 9.9695 | 11.3459 |
| 3 | 10.1735 | 11.7060 | 13.1704 | 14.5859 |

III. RESULTS AND DISCUSSIONS

Fig.2 represents the graph of X Vs $J_n(x)$. Here we can see that J_0 that has the maximum value 1 and all the others value declines gradually from 0.5.

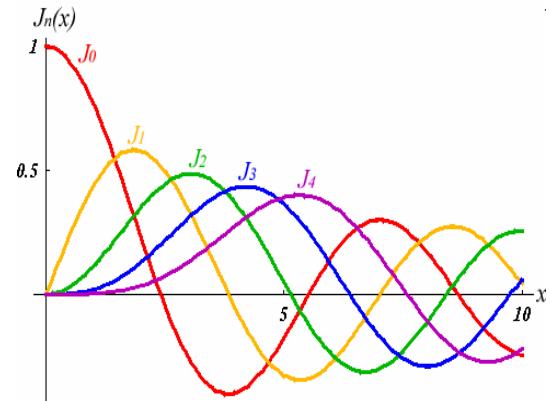


Fig2- A 1st order Bessel function

We show the 1st order Bessel function response in Fig2. Now we are interested to have the plot of a Bessel filter in order to characterize the sound wave in a cylindrical pipe. Actually the Bessel filter is a type of IIR filter i.e, an infinite impulse response filter for designing the highly developed non-linear phase analog filters. The utility of that filter designing is the less memory required for storage of coefficients and the more efficient for the computational purpose. Below Fig.3 shows the Bessel filter response for a cylindrical pipe.

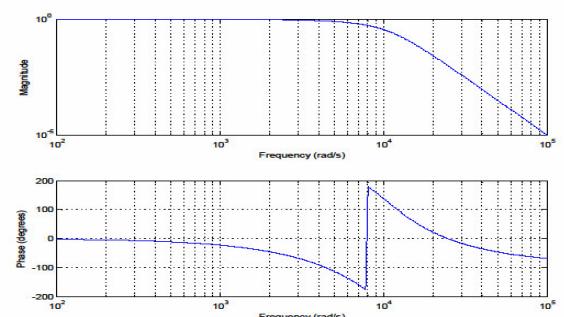
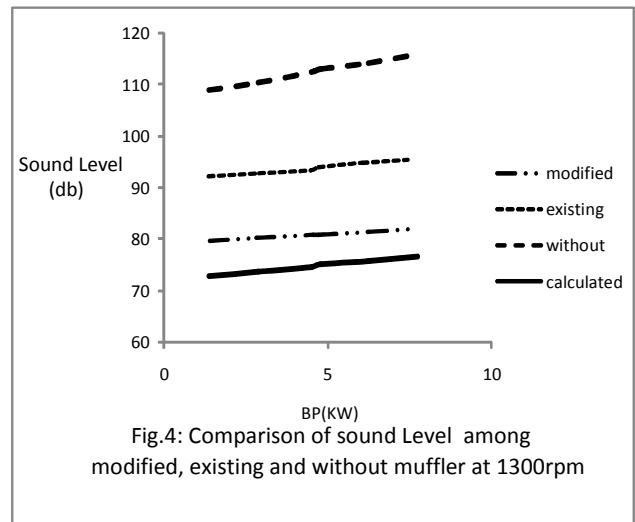


Fig.3- The Bessel filter magnitude and phase response for a cylindrical pipe.

From the Fig.3, we can see that the magnitude response declines gradually after 10^5 rad/sec. It helps us to study the 1st order Bessel characteristics shown in Fig.2 where we see that J_0 that has the maximum value 1 and the others value declines gradually from 0.5. The Bessel filter also shows us the phase characteristic where we observe a breakdown near 10^4 rad/sec and shows an anti-symmetric nature. Though we are basically concentrating on the magnitude characteristics of the Bessel filter for Reactive Muffler here.

It is observed from the Fig.4 that the Sound Pressure Level (SPL) recorded without muffler, existing muffler, modified muffler and the SPL calculated from theoretical modeling are 114.2 db, 94.9 db, 81.3 db and 75.647 db respectively. It is quite interesting to note that Sound Pressure Level by theoretical muffler is 6.95% less than the experimental muffler, indicating that assumptions taken in the theoretical modeling are reasonable.



V. CONCLUSIONS

Actually the Bessel filter used for designing the highly developed non-linear phase analog filters. The utility of that filter designing is the less memory required for storage of coefficients and the more efficient for the computational purpose. Sound pressure propagation performance in a duct with

circular cross section can easily be reduced by using Bessel Filter.

REFERENCES

- [1] Tijdeman H. On the propagation of sound waves in cylindrical tubes J. Sound Vib. 1975. Vol. 39(1). P.1-33.
- [2] Bruneau A. M., Bruneau M., Herzog P., Kergomard J. Boundary layer attenuation of higher order modes in waveguides. J. Sound Vib. 1987. Vol. 119(1). P.15-27.
- [3] P. Bhattacharya, R.Panua, B.B.Ghosh, P.K.Bose "Design of reactive muffler for study on the noise level and performance of a two cylinder four stroke 16 h.p diesel engine" Noise & Vibration Worldwide, Volume 39, Number 8, September 2008, pp. 24-27.
- [4] B.B.Ghosh, P. Bhattacharya, R.Panua, P.K.Bose "Prediction of Noise Level by Mathematical Modeling in the Exhaust Muffler and Validation of these Analytical Results with the Experimental Results for 4-Stroke Diesel Engine, in Advances in Applied Mathematical Analysis (AAMA),Volume 2, Number 1 (2007), pp 41-49.
- [5] B.B.Ghosh, P. Bhattacharya, P.K.Bose, M.Y. Vaijanapurkar "Analytical and Experimental Study on Noise Attenuation of Multi-cylinder Diesel Engine with Existing and Modified Muffler" Noise & Vibration Worldwide, Volume 37, Number 9, October 2006 , pp. 21-28(8).
- [6] Austen, A.E.W and Priede, T.(1986), "Noise of automotive diesel engine, its causes reduction, S.A.E. transaction", vol.74, paper 1000A.
- [7] Alfredson. R. J. and Davies, P.O.A.L. (1971). "The performance of exhaust silencer components, Journal of sound and vibration", 17, pp.175-196.
- [8] Belgaunkar, B. M, Somayajulu, K.D.S.R. and Mukherjee, Subrata (1969), "A Study of engine exhaust noise and silencer performance", N.V.R.L Report, I.I.T. Kharagpur.
- [9] Bender Erich, K and Brammer, A.J (1975), "Internal combustion engine intake and exhaust system noise, Journal of acoustical society of America", vol. 58, (1), pp. 22-30.
- [10] Blair, J.P and Speckho, J.A (1972), "Sound pressure level generated by internal combustion engine exhaust system ", S.A.E Transactions, paper 720155.
- [11] Crocker, M.J and price, A.R. Noise and Vibration Control vol-I, First edition, CRC Press Inc. Pp42-44 and 100-123.
- [12] Davies, P.O.A.L.(1964), "The design of silencers for internal combustion engine", Journal of Sound and Vibration.1.pp.195-201.
- [13] Proakis(2007). "Digital Signal Processing", Pearson Education India, ISBN 8131710009, 9788131710005.

A study on optimal reactive power dispatch using artificial bee colony algorithm

T. Malakar

Electrical Engineering Department
National Institute of Technology Silchar
m_tanmoy1@rediffmail.com

A. K. Sinha

Electrical Engineering Department
National Institute of Technology Silchar
ashokesinha2001@yahoo.co.in

Abstract— This paper presents an Artificial Bee Colony (ABC) based novel search algorithm for solving simple Optimal Reactive Power Dispatch (ORPD) problem for a power system network under its varied operating conditions. The purpose of this study is to investigate the performance of the search algorithm by adjusting the voltage control devices present in the network in minimizing network losses which promotes a flatter voltage profile. Four different cases of the power system have been studied considering the network operational limitations. The ORPD algorithm is tested on IEEE 30 bus system. The static solution of the power system network is obtained for each of the cases and compared with those reported from the ABC based ORPD solution.

Keywords-Optimal Reactive Power Dispatch (ORPD), Optimal Power Flow (OPF), Artificial Bee Colony algorithm (ABC), Voltage Profile.

I. INTRODUCTION

The reactive power management and voltage control is a very important issue in power system operation and control. A significant attention have been made by the researcher in Optimal Reactive Power Dispatch (ORPD) since last many years as ORPD can improve the voltage profile of the power system network by proper dispatch of voltage control devices. An ORPD is not much different then a conventional Optimal Power flow (OPF) problem. An optimal power flow generally deals with combined optimization of real and reactive power. The problem of OPF is to determine the state of the electric power system that optimizes a given objective function and satisfies a set of physical and operating constraints [1-5]. Whereas ORPD problems deals with rescheduling of voltage control devices like tap position of tap changing transformers (OLTC), VAR output of switchable capacitors (SC) in such a manner to promote a flatter voltage profile that leads to appreciable power saving on account of reduced system losses. Basically, ORPD problem is a problem of operation and planning. The ORPD problem has been investigated in a distribution substation for maintaining a specified voltage profile considering physical limitations of OLTC and SC into account in [6]. Classical coordination equation for optimal reactive power dispatch has been used to optimize real and reactive power losses in [7]. A simulation study is presented in [8] for the optimal reactive power dispatch problem considering costs of adjustments of voltage control devices

over a specified period of power system operation. A modified genetic algorithm is used for the optimization purpose.

In this paper, a simple ORPD problem has been studied considering four different operating condition of a power system network to investigate the performance of the voltage control devices in promoting a flatter voltage profile. The optimization algorithm used to solve the ORPD problem is Artificial Bee colony Algorithm (ABC) as is the most recently developed algorithm [9]. The algorithm is implemented in MATLAB and tested on standard IEEE test cases. The results obtained from this ORPD problem has been compared with those obtained without ORPD.

The rest of the paper is arranged like this; the mathematical problem has been formulated in section II, the algorithm (ABC) used for solving the ORPD has been demonstrated in section III, the implementation and result discussion have been performed in section IV and the conclusions are drawn in section V.

II. PROBLEM FORMULATION

For Optimal Reactive Power Dispatch (ORPD), the objective function is chosen as real power loss of a given power system; which ensures optimal utilization of reactive or voltage control sources. The ORPD problem is no different than a conventional Optimal Power Flow (OPF) problem. Therefore, mathematically the problem may be stated as follows:

Minimize,

$$F(x, u) \quad (1)$$

Subject to

$$g(x, u) = 0 \quad (2)$$

$$h(x, u) \leq 0 \quad (3)$$

Where, $F(x, u)$ is the objective function under consideration. x is the vector of dependent variables; for example, slack bus power, bus voltage angles, load bus voltage magnitudes etc. u is the vector of control variables. PV bus voltage magnitude, generated power, tap position of tap changers, shunt compensators output etc are represented as the

control variables. $g(x, u)$ represents the real and reactive power balance equation. $h(x, u)$ represent network's components operational limit.

A. Objective Function

The objective of real power loss minimization is done by selecting the best combination of control variables, which minimizes the total real power loss of the network simultaneously satisfying all the network constraints. Mathematically it can be expressed as:

$$P_{\text{loss}} = \sum_{j=1}^m \text{loss}_j \quad (4)$$

where m is the number of branches present in the network and loss_j is the real power loss of branch j .

B. Constraints

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

The proposed ORPD problem deals with several physical and operational constraints.

1) Power balance constraints

The total power generated by the units must be equal to the sum of total load demand and total real power loss in the transmission lines. Hence the equality constraint equations are:

$$P_{Gi} - P_{Di} - \left| V_i \right| \sum_{k=1}^{NB} \left| V_k \right| \{ G_{ik} \cos(\theta_i - \theta_k) + B_{ik} \sin(\theta_i - \theta_k) \} = 0$$

$$Q_{Gi} - Q_{Di} - \left| V_i \right| \sum_{k=1}^{NB} \left| V_k \right| \{ G_{ik} \sin(\theta_i - \theta_k) - B_{ik} \cos(\theta_i - \theta_k) \} = 0$$

Where, P_{Gi} , Q_{Gi} , P_{Di} and Q_{Di} are real and reactive power generation and demand of i^{th} thermal unit respectively. $|V_i|$, $|V_k|$ are voltage magnitude of bus i and bus k respectively. G_{ik} , B_{ik} are conductance and susceptance of the transmission line between bus i and bus k respectively. θ_i , θ_k are voltage angles of bus i and bus k .

2) Generation capacity constraints

The real power output of generating units must be restricted within their respective limits in the scheduling process and may be written as:

$$P_{Gi}^{\min} \leq P_{Gi} \leq P_{Gi}^{\max}$$

Here, P_{Gi} , P_{Gi}^{\min} and P_{Gi}^{\max} are the real power output of unit i and its minimum and maximum value respectively.

3) Other constraints

The other operational constraints involved in OARD problem with minimum control movements are generator reactive power for all PV buses (nPV), bus voltage magnitudes of all PQ buses (nPQ), line flow limits. Mathematically these can be expressed as:

$$Q_{Gj}^{\min} \leq Q_{Gj} \leq Q_{Gj}^{\max}$$

$$j = 1, 2, \dots, nPV$$

$$\left| V_j^{\min} \right| \leq |V_j| \leq \left| V_j^{\max} \right|$$

$$j = 1, 2, \dots, nPV$$

$$\left| T_i^{\min} \right| \leq |T_i| \leq \left| T_i^{\max} \right|$$

$$i = 1, 2, \dots, nT$$

$$\left| Q_{ci}^{\min} \right| \leq |Q_{ci}| \leq \left| Q_{ci}^{\max} \right|$$

$$i = 1, 2, \dots, nC$$

$$S_{ij}^k \leq S_{ij}^{\max}$$

Where, Q_{Gj} , Q_{Gj}^{\min} and Q_{Gj}^{\max} are reactive power generation of j^{th} generator and its minimum and maximum value respectively. V_j , V_j^{\min} and V_j^{\max} are voltages at j^{th} PV bus and its minimum and maximum values respectively. T_i , T_i^{\min} and T_i^{\max} are tap position of i^{th} tap changer and its minimum and maximum values respectively. Q_{ci} , Q_{ci}^{\min} and Q_{ci}^{\max} are the reactive power output of shunt capacitor i and its extreme values respectively. S_{ij}^k and S_{ij}^{\max} are the MVA flow of the line between bus i and bus j and its maximum value respectively.

III. ARTIFICIAL BEE COLONY ALGORITHM

The Artificial Bees' Colony algorithm is biologically inspired technique of swarm intelligence for searching. It is all about honey bees' work distribution and collective foraging strategy to accumulate extra nectar for their survival in winter season. The bees feed on nectar as a source of energy in their lives and use pollen as a source of protein in the rearing larvae. The colony of artificial bees' algorithm consists of three groups of bees: employed bees, onlookers and scouts. First half of the colony consists of employed artificial bees and the second half constitutes the artificial onlookers. Instead of initiating exploration by all bees, some dedicated explorer bees are appointed (employed bees) to explore the "profitability" of flower patches in the surrounding environment. This profitability accounts various parameters such as amount of nectar in flower patches, sugar content in nectar, distance of

flower patches from the bee hive etc. For every food source there is one employed bee or in other words, the number of employed bees' is equal to the number of food sources around the hive. The employed bee whose food source has been abandoned by the bees' becomes a scout bee. In the ABC, a food source position represents a possible solution to the problem to be optimized and the nectar amount of a food source corresponds to the quality (fitness) of the associated solution. The mathematical formulation of the algorithm regarding the behaviour of bees in distributing their work to optimize the collection of nectar has been presented in [9, 10].

Pseudo code of ABC algorithm

1. Initialize the random population of solutions X_{ij} (Flower patch positions).
 2. Evaluate the population.
 3. Produce new solution Y_{ij} in the neighbourhood of X_{ij} for the employed bees' by the following equation.
- $$Y_{ij} = X_{ij} + rand * (X_{ij} - X_{kj}) \quad (\text{ii})$$
- where, X_{kj} is the randomly selected solution.
4. Apply the greedy selection process between Y_{ij} and X_{ij} .
 5. Calculate the probability values ρ_i for the solution X_i by means of their fitness values.

$$\rho_i = \frac{\text{fitness}_i}{\sum_{j=1}^N \text{fitness}_j} \quad (\text{ii})$$

6. Find new solutions for the onlooker based on the probability ρ_i and evaluate them.
7. Apply the greedy selection process between new and old solutions.
8. Determine the abandoned solution (source), if exists and replace it with a new randomly produced solution for the scout.
9. Memorize the best solution (food source position) achieved so far.
10. Repeat step 3 to 9 until termination criterion is satisfied.

IV. RESULT AND DISCUSSION

The primary objective of reactive power dispatch is to reduce the real power loss involved in power system by dispatching the reactive power sources most optimally under varied operating conditions. Modern power system has to operate under different operating conditions starting from light load to high load. As the system loading increases the real

power loss incurred in the system also increases. A proper reactive power dispatch can significantly reduce the real power loss in the power system network particularly when the system is highly stressed. In this paper, four different operating conditions/cases of the power system have been analyzed. Each case is solved using ORPD and without ORPD and the results have been compared. The cases without ORPD have been solved by running simple Newton's Raphson power flow method and the ORPD problem has been solved using ABC based algorithm. The algorithm is implemented on MATLAB, version 7.4 for solving Optimal Reactive Power Dispatch (ORPD) to determine optimal operating point of voltage control devices and is experimented on IEEE 30 bus test case. Two different types of voltage control devices are present for the test system; one is tap changing transformers and the other is switchable capacitors. Along with the voltage control devices generators terminal voltages are also considered as variables in this formulation. The control variables were considered as both continuous and discrete. The network, load, generation data are taken from ref. [5]. The system is having six generators (at buses 1, 2, 13, 22, 23, 27), four tap changing transformers (between buses 6-9, 6-10, 4-12, 27-28) and two swichable capacitor banks (at buses 5 and 24). The operating range of all transformers is set between 0.9- 1.05 with a discrete step size of 0.01 and the range of capacitor banks are considered between 0- 40 Mvar with a step size of 1. Results presented for the ORPD problem are the minimum value (power loss) after executing 20 trial runs for each test case. The ABC algorithm parameters used in the computation are mentioned below.

| Colony Size | No of Flower Patches | Foraging Cycle |
|-------------|----------------------|----------------|
| 30 | 11 | 100 |

Case-1: Base Load Condition

In this case, the power system is considered to be operating under base loading condition. The real power loss, as reported in Table I is found to be 2.77 MW with corresponding average voltage of 0.9819 p.u. without ORPD i.e when normal power flow is run for the system whereas ORPD results 2.46 MW of real power loss with a corresponding average voltage of 1.001 p.u. The best operating point for the voltage control devices are presented in Table I. The variations of the nodal voltages with the two methods are plotted in figure 1.

TABLE I. BASE LOAD CONDITION

| Variables | Without ORPD (PF) | ABC based ORPD |
|------------|-------------------|----------------|
| V_{g2} | 1.0 | 0.99 |
| V_{g13} | 1.0 | 1.05 |
| V_{g22} | 1.0 | 1.02 |
| V_{g23} | 1.0 | 1.02 |
| V_{g27} | 1.0 | 1.02 |
| T6-9 | 1.01 | 0.96 |
| T6-10 | 0.96 | 0.97 |
| T4-12 | 1.01 | 1.01 |
| T27-28 | 0.96 | 1.04 |
| SC5 | 19 | 6 |
| SC24 | 4 | 8 |
| Ploss (MW) | 2.77 | 2.46 |
| Vavg (p.u) | 0.9819 | 1.001 |

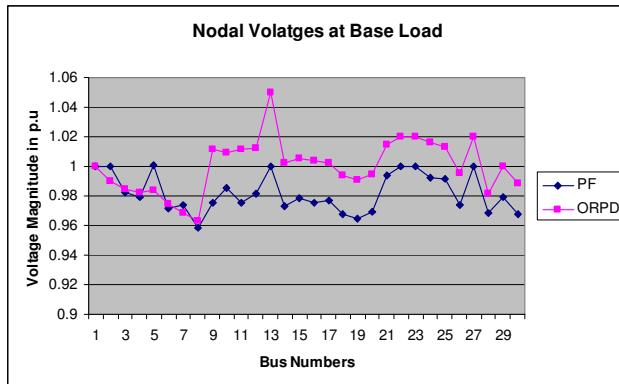


Figure 1.

A. Case 2: 110% of Base Load Condition

The loading of the power system has been increased by 10% for both real and reactive power in this case. However, it is assumed to maintain the power factor constant while increasing the load. The detail results are presented in Table II.

TABLE II. 110% OF BASE LOAD CONDITION.

| Variables | Without ORPD (PF) | ABC based ORPD |
|------------|-------------------|----------------|
| Vg2 | 1.0 | 0.997 |
| Vg13 | 1.0 | 1.05 |
| Vg22 | 1.0 | 1.016 |
| Vg23 | 1.0 | 1.02 |
| Vg27 | 1.0 | 1.03 |
| T6-9 | 1.01 | 0.96 |
| T6-10 | 0.96 | 0.96 |
| T4-12 | 1.01 | 0.99 |
| T27-28 | 0.96 | 1.03 |
| SC5 | 19 | 8 |
| SC24 | 4 | 10 |
| Ploss (MW) | 3.30 | 2.9 |
| Vavg (p.u) | 0.9815 | 1.002 |

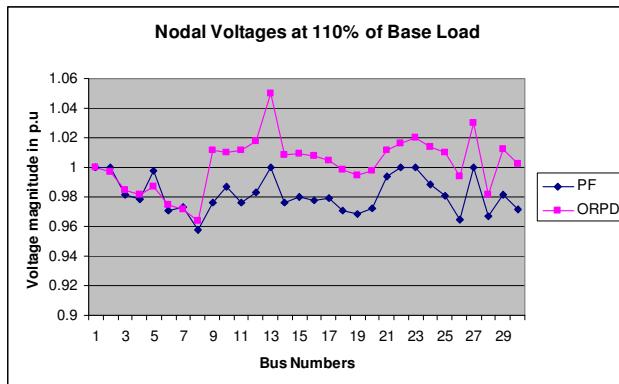


Figure 2.

It is observed that real power loss has decreased from 3.30 MW to 2.90 MW with the proposed ORPD approach with an additional boost in maintaining voltage better among the buses in the system. The optimal operating point of voltage control

devices are marked in the Table. Figure 2 shows how ORPD maintains voltage at this loading condition then compared to the case without ORPD.

B. Case 3: 120% of Base Load Condition

The loading of the power system has further increased to 120% for both real and reactive power in this case for the analysis purpose. Again, the power factor has been maintained constant while increasing the load. From the results reported in Table III, it is observed that real power loss has decreased in this case also with the ORPD algorithm from 4.21 MW to 3.74 MW. The optimal operating point of voltage control devices are also shown in the Table. Figure 3 shows how ORPD maintains voltage at this loading condition differently then the case without ORPD.

TABLE III 120% OF BASE LOAD CONDITION

| Variables | Without ORPD (PF) | ABC based ORPD |
|------------|-------------------|----------------|
| Vg2 | 1.0 | 0.996 |
| Vg13 | 1.0 | 1.05 |
| Vg22 | 1.0 | 1.01 |
| Vg23 | 1.0 | 1.01 |
| Vg27 | 1.0 | 1.02 |
| T6-9 | 1.01 | 0.94 |
| T6-10 | 0.96 | 0.99 |
| T4-12 | 1.01 | 1.0 |
| T27-28 | 0.96 | 1.04 |
| SC5 | 19 | 10 |
| SC24 | 4 | 10 |
| Ploss (MW) | 4.21 | 3.743 |
| Vavg (p.u) | 0.9798 | 0.9963 |

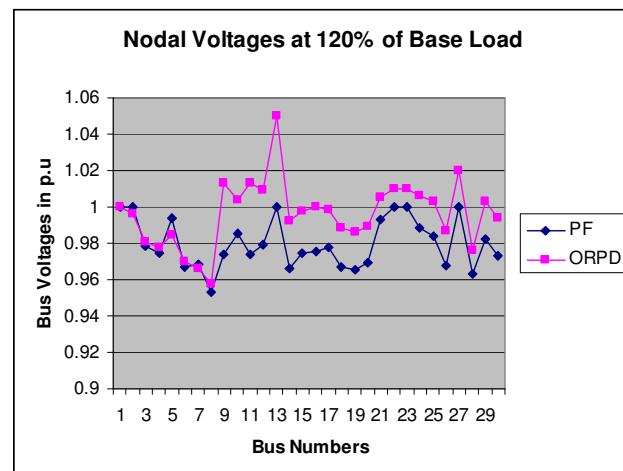


Figure 3.

C. Case 4: 150% of Base Load Condition

The loading of the power system has now been increased to 150% keeping the power factor constant like earlier in order to investigate the impact of voltage control devices. It is observed from Table IV that real power loss has decreased in this case also with the ORPD algorithm from 11.17 MW to 9.90 MW. The optimal operating point of voltage control devices are also

shown in Table IV. It is observed from Table IV that with the proposed ORPD problem, in order to reduce the real power loss incurred by the system, the reactive power output of the switchable capacitors has increased to almost double of its value for the optimization. From the point of view of tap changing transformers, the off-nominal turns ratios have increased to maintain better voltage profile for the system. There is also a boost up observed in PV bus voltage magnitude from Table IV. The overall effect can be viewed from Figure 4 which shows how ORPD maintains voltage at this loading condition better than the case without ORPD.

TABLE IV 150% OF BASE LOAD CONDITION

| Variables | Without ORPD (PF) | ABC based ORPD |
|------------|-------------------|----------------|
| Vg2 | 1.0 | 0.99 |
| Vg13 | 1.0 | 1.05 |
| Vg22 | 1.0 | 1.01 |
| Vg23 | 1.0 | 1.02 |
| Vg27 | 1.0 | 1.03 |
| T6-9 | 1.01 | 0.96 |
| T6-10 | 0.96 | 0.97 |
| T4-12 | 1.01 | 1.02 |
| T27-28 | 0.96 | 1.05 |
| SC5 | 19 | 40 |
| SC24 | 4 | 8 |
| Ploss (MW) | 11.172 | 9.90 |
| Vavg (p.u) | 0.9708 | 0.9928 |

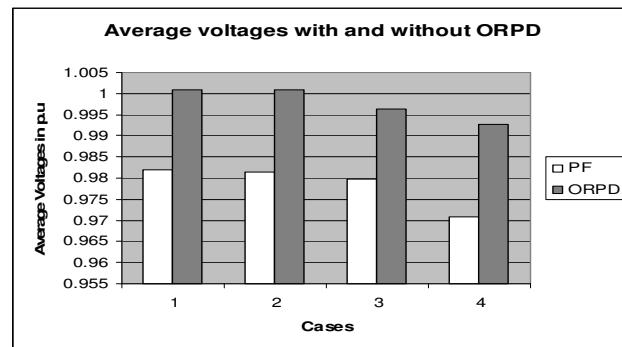


Figure 5.

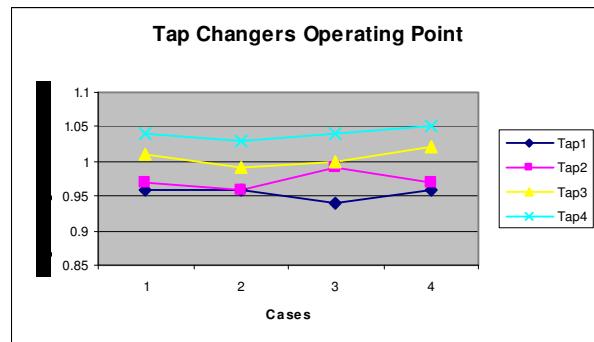


Figure 6.

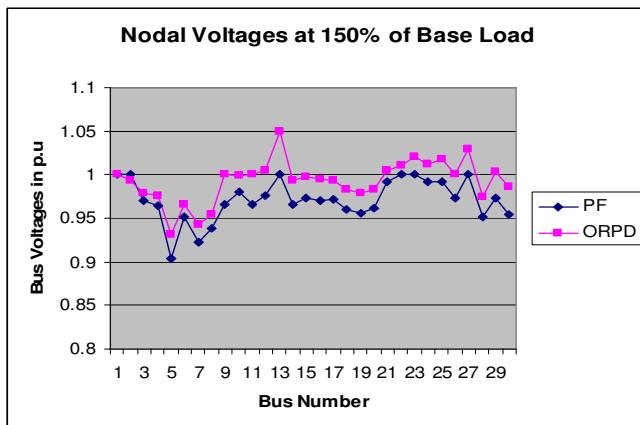


Figure 4.

The p.u average voltages for each of the cases have been plotted in figure 5 to show the effect of ORPD in maintaining the voltage. In order to investigate the change in the operating point of voltage controlling devices on the increase of load of the power system, the optimal operating point of tap changers and switchable capacitors are plotted in figure 6 and figure 7 respectively. It is observed that the output for both types of voltage control devices has increased with the increase of load which is obvious.

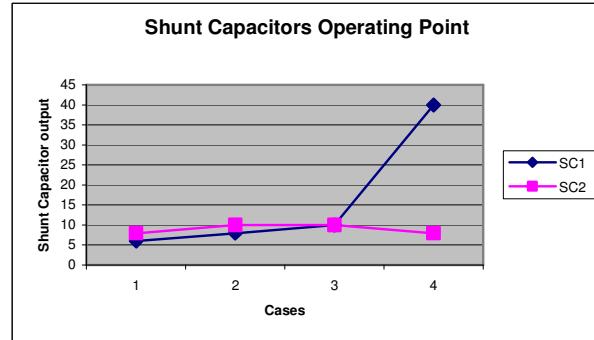


Figure 7.

V. CONCLUSION

In this paper an Optimal Reactive Power Dispatch (ORPD) problem has been solved using artificial bee colony algorithm. The algorithm has been implemented on MATLAB and has been tested on an IEEE test system in order to investigate the performance of the algorithm in adjusting the voltage control devices in minimizing real power loss in a power system under varied network conditions. The results have been compared with the solution obtained without solving ORPD and it is observed that ABC based ORPD can effectively find the

2nd Annual International Conference IEMCON 2012

optimal operating points of the voltage control devices under varied network conditions.

REFERENCES

- [1] Dommel H, W. F. Tinney, "Optimal Power Flow Solutions., IEEE Transactions on Power Apparatus and Systems. Vol: PAS-87, pp 1866-1876, 1968.
- [2] Bjelogrlic M, Calovic M. S, Babic B. S, Ristanovic P., Application of newton's optimal Power Flow in Voltage/Reactive Power Control. 1989 IEEE.
- [3] Pai M. A., Computer Techniques in Power system Analysis. Second Edition, Tata McGraw-Hill.
- [4] Allen J. Wood, Bruce F. Wollenberg., Power Generation Operation and control. John Wiley & Sons, Inc.
- [5] M. R. AlRashidi, M. E. El-Hawary, "Hybrid particle Swarm Optimization Approach for Solving the Discrete OPF Problem considering the Valve Loading Effects". IEEE Transactions on Power systems, vol. 22, No. 4.,2007.
- [6] Lu F.-C, Hsu Y.-Y., Reactive power / voltage control in a distribution substation using dynamic programming., IEE Proc.- Gener. Transm. Distrib. Vol. 142, No. 6, 1995.
- [7] Prof. J. Nanda, L. Hari, M. L. Kothari, " Challenging algoithm for optimal reactive power dispatch through classical co-ordination equations", IEE Procedings-C, Vol. 139, No.2, March 1992.
- [8] Yong-jun Zhang and Zhen Ren, "Optimal reactive power dispatch considering costs of adjusting the control devices", IEEE Trans on Power Systems, Vol. 20, No 3, 2005.
- [9] D. Karaboga, B. Basturk., On the performance of artifitial bee colony (ABC) algorithm., Elsevier Applied Soft Computing 8 (2008).
- [10] Teodorovic D, Lucic P, Markovic G, Orco M. D., Bee colony optimization : Principles and Applications., NEUREL 2006, IEEE.

A novel Control Algorithm for Self Supported Dynamic voltage Restorer(DVR)

Himadri Ghosh

Electrical Engineering Department
Jalpaiguri Government Engineering College
e-mail: joybiet.123@gmail.com

Abstract—Power quality is one of the major concerns in the present era. The problem of voltage sags and swells and its major impact on sensitive loads are well known. To solve this problem, custom power devices are used. One of those devices is the Dynamic Voltage Restorer (DVR), which is one of the most efficient and effective modern custom power devices used in power distribution networks. A new control algorithm for the DVR is proposed in this paper to regulate the load terminal voltage during sag, swell in the voltage at the point of common coupling (PCC).

This new control algorithm is based on synchronous reference frame theory (SRF) along with PI controller is used for the generation of reference voltages for a dynamic voltage restorer (DVR). These voltages, when injected in series with a distribution feeder by a voltage source inverter (VSI) with PWM control, can regulate the voltage at the load terminals against any power quality problem in the source side. It first analyzes the power circuit of the system in order to come up with appropriate control limitations and control targets for the compensation voltage control through the DVR. The control of the DVR is implemented through derived reference load terminal voltages. The proposed control scheme is simple to design. Simulation results carried out by MATLAB with its Simulink and Sim Power System (SPS) toolboxes to verify the performance of the proposed method.

Keywords- Power Quality, DVR, voltage sags/swells, VSI, Synchronous Reference Frame Theory, MATLAB/SIMULINK

I. INTRODUCTION

Power distribution systems, ideally, should provide their customers with an uninterrupted flow of energy at smooth sinusoidal voltage at the contracted magnitude level and frequency [1] however, in practice, power systems, especially the distribution systems, have numerous nonlinear loads, which significantly affect the quality of power supplies. As a result of the nonlinear loads, the purity of the waveform of supplies is lost. This ends up producing many power quality problems. Apart from nonlinear loads, some system events, both usual (e.g. capacitor switching, motor starting) and unusual (e.g. faults) could also inflict power quality problems [2]. Power quality phenomenon or power quality disturbance can be defined as the deviation of the voltage and the current from its ideal waveform. Faults at either the transmission or distribution level may cause voltage sag or swell in the entire system or a large part of it. Also, under heavy load conditions, a significant voltage drop may occur in the system. Voltage

Pradip Kumar Saha & Goutam Kumar Panda

Electrical Engineering Department
Jalpaiguri Government Engineering College
e-mail: p_ksaha@rediffmail.com, g_panda@rediffmail.com

sag and swell can cause sensitive equipment to fail, shutdown and create a large current unbalance. These effects can incur a lot of expensive from the customer and cause equipment damage [1]. The voltage dip magnitude is ranged from 10% to 90% of nominal voltage and with duration from half a cycle to 1 min and swell is defined as an increase in rms voltage or current at the power frequency for durations from 0.5 cycles to 1 min. Typical magnitudes are between 1.1 and 1.8 p.u[2].

There are many different methods to mitigate voltage sags and swells, but the use of a custom power device is considered to be the most efficient method, e.g. FACTS for transmission systems which improve the power transfer capabilities and stability margins. The term custom power pertains to the use of power electronics controller in a distribution system [10], especially, to deal with various power quality problems. Custom power assures customers to get pre-specified quality and reliability of supply. This pre-specified quality may contain a combination of specifications of the following: low phase unbalance, no power interruptions, low flicker at the load voltage, and low harmonic distortion in load voltage, magnitude and duration of over voltages and under voltages within specified limits, acceptance of fluctuations, and poor factor loads without significant effect on the terminal voltage

There are different types of Custom Power devices used in electrical network to improve power quality problems. Each of the devices has its own benefits and limitations. A few of these reasons are as follows. The SVC pre-dates the DVR, but the DVR is still preferred because the SVC has no ability to control active power flow [3]. Another reason include that the DVR has a higher energy capacity compared to the SMES and UPS devices. Furthermore, the DVR is smaller in size and cost is less compared to the DSTATCOM and other custom power devices. Based on these reasons, it is no surprise that the DVR is widely considered as an effective custom power device in mitigating voltage sags. In addition to voltage sags and swells compensation, DVR can also add other features such as harmonics and Power Factor correction. Compared to the other devices, the DVR is clearly considered to be one of the best economic solutions for its size and capabilities [4].

The voltage injection schemes and design of the self-supported DVR and the different control strategies for the controllers of the DVR have been discussed in [14-16]. E.g., the instantaneous reactive power theory (IRPT) [16], adaline based fundamental extraction have been implemented in [14]. Instantaneous symmetrical component theory, space vector

modulation, synchronous reference frame theory (SRFT) based control techniques for a DVR are reported in this literature. The SRFT based algorithm reported in [15] involves conversion of three phase voltages from the stationary frame to rotating frame and vice versa. The IRPT algorithm has been used in [16] based on unit templates and instantaneous symmetrical component theory. In this paper, a new control algorithm is suggested based on SRF theory which includes P-I Controller for the generation of reference V_d and V_q . Reference load signal generation involves the conversion from three-phase to two-phase and vice versa. Moreover low pass filters are essential part of this algorithm which has slow dynamic response of the compensator.

The organization of the paper is as follows. In section II, the constructional part of the DVR is briefly described, the operating principle and the voltage injection capabilities of the DVR is discussed in section III, proposed control algorithm enumerated in section IV and the detailed description of MATLAB Simulation model along with its performance in electrical network discussed in section V and section VI respectively.

II. DYNAMIC VOLTAGE RESTORER(DVR)

A Dynamic Voltage Restorer (DVR) is a recently proposed series connected solid state device that injects voltage into the system in order to regulate the load side voltage. It is normally installed in a distribution system between the supply and critical load feeder [9]. A DVR can restore the load side voltage to the balanced sinusoidal voltage with desired amplitude even when the source voltages are unbalanced or distorted. The basic structure of a DVR is shown in Fig.1.

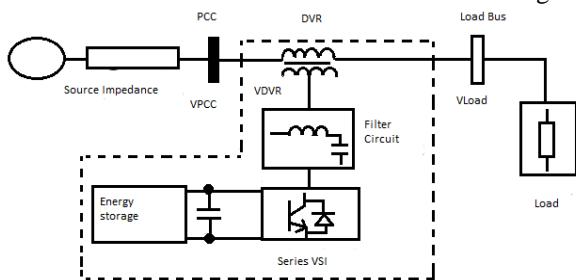


Figure-1: DVR series connected topology

It is divided into six categories[9]: (i) *Injection Transformer*: The Injection / Booster transformer is a specially designed transformer that attempts to limit the coupling of noise and transforms energy from the primary side to the secondary side.(ii) *Harmonic Filters*: Filters are used to convert the inverted PWM waveform into a sinusoidal waveform. This is achieved by eliminating the unwanted harmonic components generated by VSI action. Higher order harmonic components distort the compensated output voltage. (iii) *Inverter*: A VSI is a power electronic system which consists of a storage device and switching device, which can generate a sinusoidal voltage at any required frequency, magnitude, and phase angle from dc storage. IGBT is a three terminal controllable switch that combines the fast switching times of the MOSFET with the high voltage capabilities of the GTO used as a switching device in VSI [6]. (iv) *Energy Storage Unit*: The purpose is to

supply the necessary energy to the VSI via a dc link for the generation of injected voltages. Batteries, SMES and super capacitors can be used as energy storage devices. (v) *Capacitor*: DVR has a large DC capacitor to ensure stiff DC voltage input to inverter. (vi) *By-Pass Switch*: If the over current on the load side exceeds a permissible limit due to short circuit on the load or large inrush current, the DVR will be isolated from the system by using the bypass switches and supplying another path for current.

III. OPERATION OF DVR

The schematic diagram of a self-supported DVR is shown in Figure-2[5]. Three phase source voltages (V_{sa} , V_{sb} , V_{sc}) are connected to the 3-phase critical load through series impedance (Z_{sa} , Z_{sb} , Z_{sc}) and an injection transformer in each phase. The terminal voltages (V_{ta} , V_{tb} , V_{tc}) have power quality problems and the DVR injects compensating voltages (V_{ca} , V_{cb} , V_{cc}) through an injection transformer to get undistorted and balanced load voltages (V_{la} , V_{lb} , V_{lc}). The DVR is implemented using a three leg voltage source inverter with IGBTs along with a dc capacitor (C_{dc}). A ripple filter (L_r , C_r) is used to filter the switching ripple in the injected voltage. The considered load, sensitive to power quality problems is a three-phase balanced lagging power factor load. A self-supported DVR does not need any active power during steady state because the voltage injected is in quadrature with the feeder current.

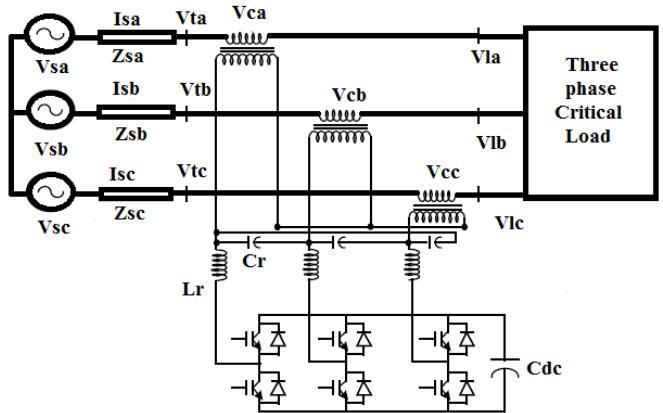


Figure-2: Schematic diagram of self-supported DVR

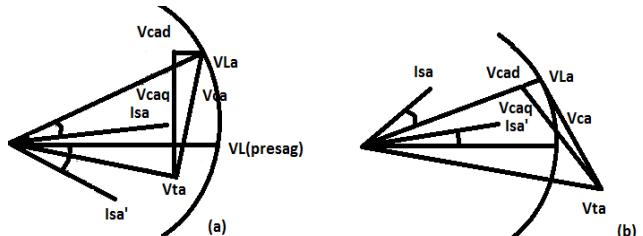


Figure-3: Phasor Diagram for (a) Voltage Sag (b) Voltage Swell

The DVR operation for the compensation of sag, swell in supply voltages is shown in Figure-3. Before sag the load voltages and currents are represented as V_L (presag) and I_{sa}' as shown in Figure-3(a). After the sag event, the terminal voltage (V_{ta}) is gets lower in magnitude and lags the presag voltage by some angle. The DVR injects a compensating voltage (V_{ca}) to

maintain the load voltage (V_L) at the rated magnitude. V_{Ca} has two components, V_{CAd} and V_{Caq} . The voltage in-phase with the current (V_{CAd}) is required to regulate the dc bus voltage and also to meet the power loss in the VSI of DVR and an injection transformer [5]. The voltage in quadrature with the current (V_{Caq}) is required to regulate the load voltage (V_L) at constant magnitude. During swell event, the injected voltage (V_{Ca}) is such that the load voltage lies on the locus of the circle as shown in Figure-3(b).

IV. CONTROL OF DVR

The compensation for voltage sags using a DVR can be performed by injecting/absorbing reactive power or real power. When the injected voltage is in quadrature with the current at the fundamental frequency, compensation is achieved by injecting reactive power and the DVR is self-supported with dc bus. But, if the injected voltage is in phase with the current, DVR injects real power and hence a battery is required at the dc side of VSI. The control technique adopted should consider the limitations such as the voltage injection capability (inverter and transformer rating) and optimization of the size of energy storage [4].

Figure-4 shows the control block of the DVR in which the synchronous reference frame (SRF) theory is used for the control of self-supported DVR. The voltages at PCC (V_t) are converted to the rotating reference frame using the abc-dq0 conversion. The harmonics and the oscillatory components of voltages are eliminated using low pass filters (LPF). The components of voltages in d-axis and q-axis are,

$$\begin{aligned} V_{sd} &= V_{sd} \text{ dc} + V_{sd} \text{ ac} \\ V_{sq} &= V_{sq} \text{ dc} + V_{sq} \text{ ac} \end{aligned}$$

The compensating strategy for compensation of voltage quality problems considers that the load terminal voltage should be of rated magnitude and undistorted.

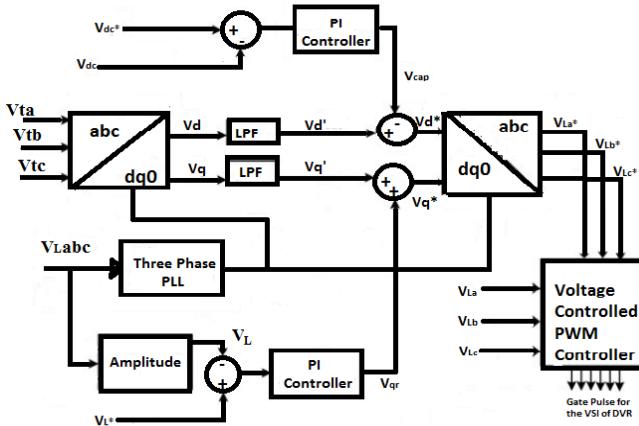


Figure-4: Control Block of DVR using SRF method of Control

The sag and swell in terminal voltages are compensated by controlling the DVR and the proposed algorithm inherently provides a self-supporting dc bus for the DVR. Three-phase reference supply voltages (V_{La*} , V_{Lb*} , V_{Lc*}) are derived using the sensed load voltages (V_{La} , V_{Lb} , V_{Lc}), terminal voltages (V_{ta} , V_{tb} , V_{tc}) and dc bus voltage (V_{dc}) of the DVR as feedback signals. The synchronous reference frame theory

based method is used to obtain the direct axis (V_d) and quadrature axis (V_q) components of the load voltage. The load voltages in the three-phases are converted into the d-q-0 frame using the Park's transformation as,

$$\begin{bmatrix} V_d \\ V_q \\ V_0 \end{bmatrix} = \begin{bmatrix} \cos \theta & \cos(\theta - \frac{2\pi}{3}) & \cos(\theta + \frac{2\pi}{3}) \\ \sin \theta & \sin(\theta - \frac{2\pi}{3}) & \sin(\theta + \frac{2\pi}{3}) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix}$$

A three-phase PLL (phase locked loop) is used to synchronize these signals with the terminal voltages (V_{ta} , V_{tb} , V_{tc}). The d-q components are then passed through low pass filters to extract the dc components of V_d and V_q . In order to maintain the DC bus voltage of the self-supported DVR, the error between the reference dc capacitor voltage (V_{dc*}) and the sensed dc bus voltage (V_{dc}) of DVR is given to a PI controller of which output (V_{cad}) is considered as the loss component of voltage and is added to the dc component of V_d to generate V_d* . The reference d-axis load voltage is therefore as,

$$V_{ld*} = V_{sd} \text{ dc} - V_{loss}$$

Similarly, a second PI controller is used to regulate the amplitude of the load voltage (V_L). The amplitude of load voltage (V_L) at point of common coupling is calculated from the ac voltages (V_{La} , V_{Lb} , V_{Lc}) as,

$$V_L = \sqrt{\left(\frac{2}{3}\right) \sqrt{V_{La}^2 + V_{Lb}^2 + V_{Lc}^2}}$$

The amplitude of the load terminal voltage (V_L) is employed over the reference amplitude (V_L*) and the output of PI controller is considered as the reactive component of voltage (V_{qr}) for voltage regulation of load terminal voltage added with the dc component of V_q to generate V_q* . The reference q-axis load voltage is therefore as,

$$V_{lq*} = V_{sq} \text{ dc} + V_{qr}$$

The resultant voltages (V_d* , V_q* , V_0) are again converted into the reference supply currents using the reverse Park's transformation. Reference supply voltages (V_{La*} , V_{Lb*} , V_{Lc*}) and the sensed load voltages (V_{La} , V_{Lb} , V_{Lc}) are used in PWM current controller to generate gating pulses for the switches. The PWM controller operates at a frequency of 1080Hz and the gating signals are given to the three-leg VSI for the control of supply voltages.

V. MATLAB MODELLING AND SIMULATION

The DVR is modeled and simulated using the MATLAB and its Simulink and Sim Power System toolboxes. The MATLAB model of the DVR connected system [14] is shown in Fig.-5. The three-phase programmable source is connected to the three-phase load through the DVR in order to generate sag, swell and harmonics in supply side. The considered load is a lagging power factor load. The VSI of the DVR is connected to the system using an injection transformer. In addition, a ripple filter for filtering the switching ripple in the terminal

voltage is connected across the terminals of the secondary of the transformer. The dc bus capacitor of DVR is selected based on the transient energy requirement and the dc bus voltage is selected based on the injection voltage level. The dc capacitor decides the ripple content in the dc voltage. The system data are given in Appendix.

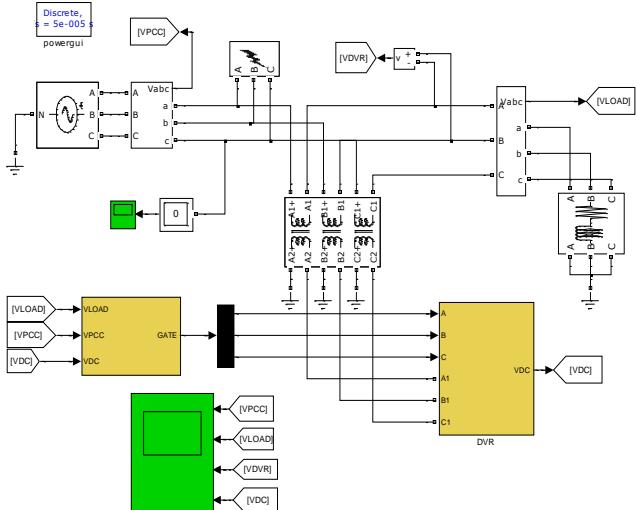


Figure-5: MATLAB model of DVR connected system

The control algorithm for the DVR is simulated in MATLAB. The control algorithm shown in Fig.-4 is modeled for DVR control of Fig.-5. The reference load voltages are derived from the sensed terminal voltages, load supply voltages and the dc bus voltage of the DVR. A pulse width modulation (PWM) controller is used over the reference and sensed load voltages to generate gate signals for the IGBT's of the VSI.

VI. PERFORMANCE OF THE DVR SYSTEM

The performance of the DVR is demonstrated for different supply voltage disturbances such as sag and swells in supply voltage [3]. A case of Three-phase voltage sag is simulated and the results are shown in Figure-6. Figure-6(a) shows 50% voltage sag initiated at 150 ms and it is kept until 300 ms, with total voltage sag duration of 150 ms. Figure-6(b) and (c) show the voltage injected by the DVR and the compensated load voltage, respectively. As a result of DVR, the load voltage is kept at 1 p.u. throughout the simulation, including the voltage sag period. It is observed that during normal operation, the DVR is not operational. It quickly injects necessary voltage components to smoothen the load voltage upon detecting voltage sag.

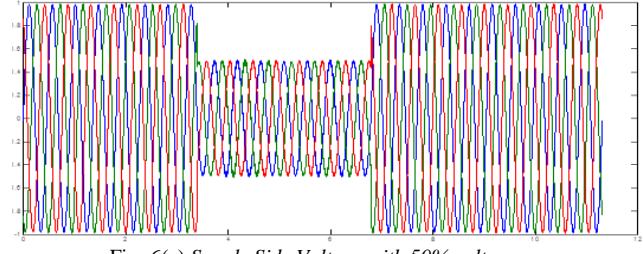


Fig.-6(a) Supply Side Voltage with 50% voltage sag

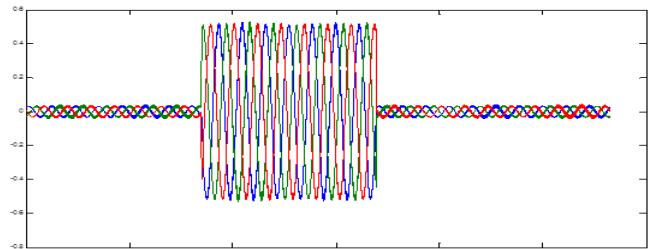


Fig.-6(b) Compensated Voltage injected by DVR

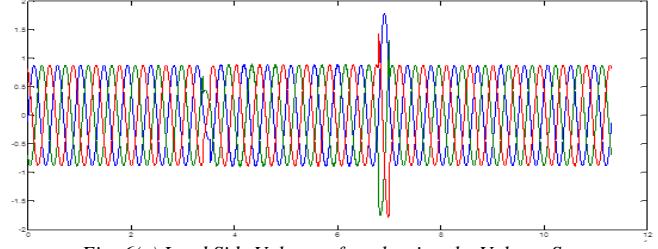


Fig.-6(c) Load Side Voltage after clearing the Voltage Sag

Figure-7(a) shows the transient performance of the system under voltage sag and swell conditions. At 0.12 sec, a sag in supply voltage is created for the duration of 0.13sec and at 0.3 sec a swell in supply is created for the duration of 0.1sec. It is observed that the load voltage is regulated to constant amplitude under both sag and swell conditions. Figure-7(b) shows the in-phase injection of voltage by the DVR. The load voltage is maintained sinusoidal by injecting proper compensation voltage by the DVR shown in Figure-7(c).

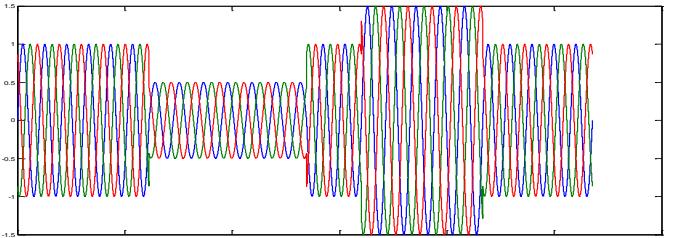


Fig.-7(a) Supply Side Voltage with 50% voltage Sag and 50% voltage Swell

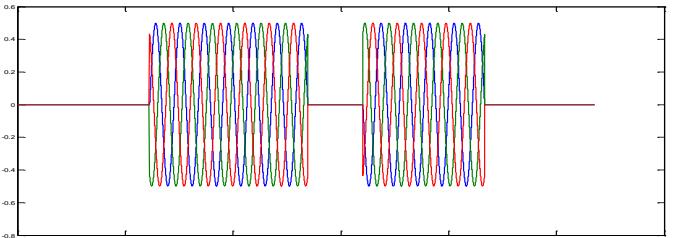


Fig.-7(b) Compensated Voltage injected by DVR for compensating voltage

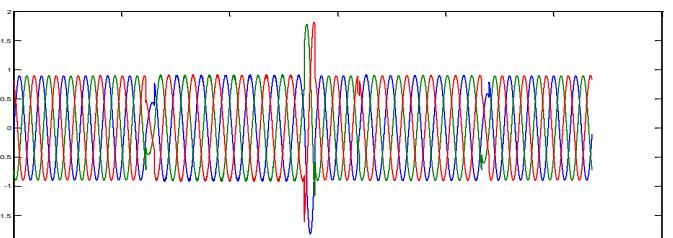


Fig.-7(c) Load Side Voltage after clearing the Voltage Sag and Swell

Figure-8(a) shows the first simulation was done with no DVR and a three phase fault is applied to the system at point

with fault resistance of 0.66 U for a time duration of 200 ms. Figure-8(b) shows The second simulation is carried out at the same scenario as above but a DVR is now introduced at the load side to compensate the voltage sag occurred due to the three phase fault applied.

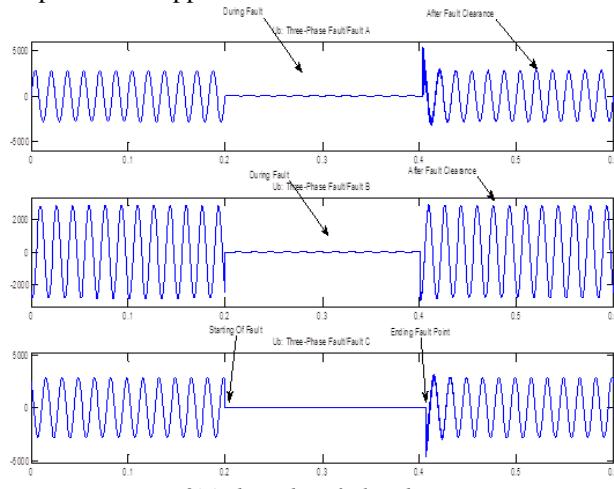


Fig.-8(a) Three phase fault without DVR

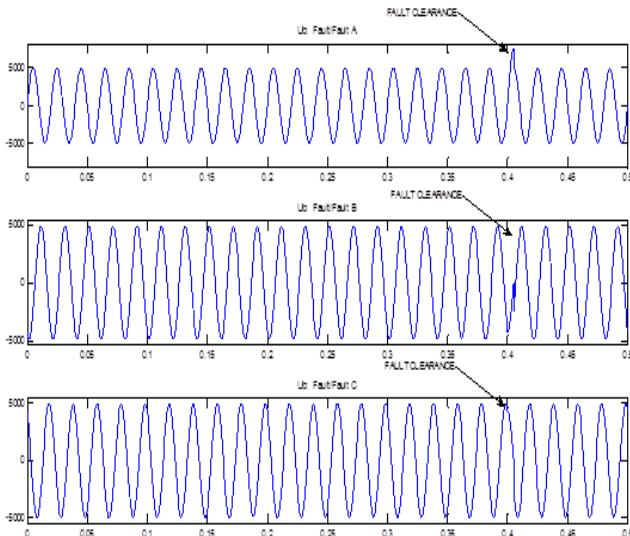


Fig.-8 Three phase fault with DVR voltage compensation

The harmonic compensation in load voltage achieved and depicted in Figure-9(a) and (b). The terminal voltage is distorted by adding 5th harmonic inversely proportional to their harmonic number. The load voltage is sinusoidal and constant in magnitude due to the injection of opposite harmonic voltage by DVR.

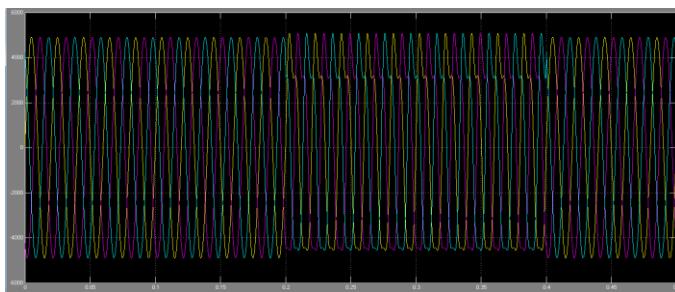


Fig.-9(a) Supply Side Voltage adding with Harmonics

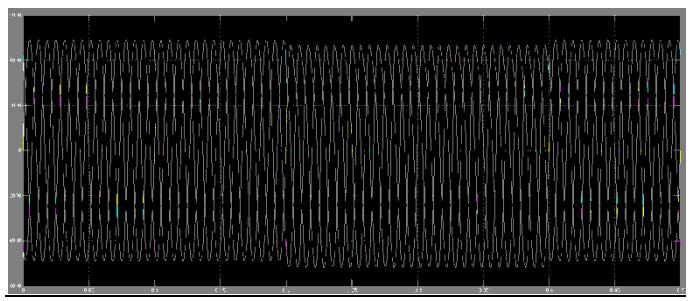


Fig.-9(b) Load Side Voltage after Harmonics Compensation

VII. CONCLUSION

The study of DVR is focused on applying a new sag detection method and a new compensating voltage generation method. The conventional sag detection method is unable to detect the voltage sags lower than a definite level. To overcome the disadvantages of the conventional sag detection method, the proposed method is used in this paper. With the proposed method, the controller is able to detect different types of power quality problems without an error and injects the appropriate voltage component to correct immediately any abnormality in the terminal voltage to keep the load voltage balanced and constant at the nominal value. Simulation and experimental results show that, the proposed DVR successfully protects the most critical load against voltage sags. Moreover it has been found that DVR is capable of providing a self-support to its dc bus by taking active power from the ac line.

APPENDIX

AC line voltage: 415 V, 50 Hz

Load: 10KVA, 0.80 pf lag

PI Controller: $K_p = 5$ $K_i = 120$

DC voltage of DVR: 300V

Harmonic Filter: $L_r = 2.0\text{mH}$, $C_r = 10\mu\text{F}$, $R_r = 4.8\Omega$

PWM Switching Frequency: 1080Hz

Injection Transformer: Turns Ratio=1:1

Inverter: IGBT based 3 arms, 6 Pulse, Frequency =1080 Hz,

Sample Time= 5 μs

DC Bus Capacitance of DVR: 1000 μF

REFERENCES

- [1] Power Quality Problems and New Solutions by A. de Almeida, L. Moreira, J. Delgado
- [2] Math H.J. Bollen, Understanding power quality problems: voltage sags and interruptions, IEEE Press, New York, 2000
- [3] Performance of DVR under different voltage sag and swell conditions by T. Devaraju, V. C. Reddy and M. Vijaya Kumar
- [4] Voltage Quality Improvement Using DVR by Chellali BENACHAIBA, Brahim FERDI
- [5] Adaline-Based Control of Capacitor Supported DVR for Distribution System Bhim Singh*, P. Jayaprakash†, and D. P. Kothari
- [6] Control of three level inverter based DVR by 1S.LEELA, S.DASH

2nd Annual International Conference IEMCON 2012

- [7] Rating and Design Issues of DVR Injection Transformer by Sasitharan S., Mahesh K. Mishra, *Member, IEEE*, B. Kalyan Kumar, and Jayashankar V.,*member, IEEE*
- [8] Particle Swarm Optimization based Dynamic Voltage Restorer
- [9] Power Quality Enhancement Using Custom Power Devices by A. Ghosh and G. Ledwich. 2002. Kluwer Academic Publishers.
- [10] FACTS controllers in power transmission and distribution by K. R. Padiyar
- [11] S. Chen, G. Joos, L. Lopes, and W. Guo. 2000. A nonlinear control method of dynamic voltage restorers in 2002 IEEE 33rd Annual Power Electronics Specialists Conference. pp. 88- 93.
- [12] Voltage quality enhancement with custom power park by Mehmet Emin MERAL
- [13] The Effect of DVR Location for Enhancing Voltage Sag by Rohanim Ibrahim, Ahmed M. A. Haidar, Zahim M., Herbert Iu
- [14] Modeling and simulation for voltage sags/swells mitigation using DVR by Rosli omar, Nasrudin Abd Rahim, Marizan Sulaiman
- [15] "Control of Reduced Rating Dynamic Voltage Restorer with Battery Energy Storage System, " by B. Sing, P. Joyprakash
- [16] H. Akagi, E H Watanabe and M Arede, Instantaneous power theory and applications to power conditioning, John Wiley and Sons, New Jersey, 2007.

Himadri Ghosh receives his B.Tech in Electrical Engineering from Birbhum Institute of Engineering and Technology under West Bengal University of Technology (WBUT) and currently pursuing M.Tech (Final Year) in Power Electronics and Drives at Jalpaiguri Govt. Engineering College. His research interests include Power electronics, Power System.

Dr. Pradip Kr. Saha received his B.E degree in Electrical Engineering from BE College, Shibpur and M.Tech degree in Machine Drives and Power Electronics from IIT Kharagpur and Ph.D. from NBU. He is currently a professor and Head in Electrical Engineering Dept. at Jalpaiguri Govt. Engineering College. His research interests include Power Electronics, Machine Drives, and CHAOS in Power Electronics.

Dr. Gautam Kr. Panda received his B.E degree in Electrical Engineering from Jalpaiguri Govt. Eng. College and M.E.E. degree in Machine Drives and Power Electronics Jadavpur University and Ph.D. from NBU. He is currently a professor in Electrical Engineering Dept. at Jalpaiguri Govt. Engineering College. His research interests include Electrical Machines & Drives, Nonlinear Dynamics and CHAOS.

Study The Nonlinear Behaviour of PFC Boost Converter

Arnab Ghosh

Department of Electrical Engineering
Jalpaiguri Govt. Engineering College
Jalpaiguri, India
e-mail: arnab.4u2010@gmail.com

Abhisek Pal

Department of Electrical Engineering
Jalpaiguri Govt. Engineering College
Jalpaiguri, India
e-mail: abhisekpal.ee@gmail.com

Dr. P. K. Saha

Prof. and Head of the Department of Electrical Engineering
Jalpaiguri Govt. Engineering College
Jalpaiguri, India
e-mail: p_ksaha@rediffmail.com

Dr. G. K. Panda

Professor of the Department of Electrical Engineering
Jalpaiguri Govt. Engineering College
Jalpaiguri, India
e-mail: g_panda@rediffmail.com

Abstract—With rapid development in power semiconductor devices, the usage of power electronic systems has expanded to new and wide application range that include residential, commercial, aerospace and many others. However, their nonlinear behavior puts a question mark on their high efficiency. This paper aims to develop a circuit for PFC boost converter to observe nonlinear behaviors. It is clear that the output storage capacitor is a main contributing parameter on the system stability, therefore, bifurcation maps are developed to determine the accurate minimum output capacitance value that assures the system stability under all operating conditions.

Keywords- *PFC boost converter; Phase Plane Trajectories ; Bifurcation diagrams.*

I. INTRODUCTION

The power electronic engineers observe some strange phenomena noise like oscillation. Actually power electronics system can exhibit a variety of nonlinear behaviours because of periodic switching of the circuits. This kind of nonlinearity is the main cause of harmonics generation i.e. degradation of input power factor. In the last decade, bifurcation and chaotic phenomena have been reported in some type of DC-DC converters [1][3][6]. Here we are discussing about some nonlinear phenomena of Power Factor Corrected (PFC) Boost Converter.

The operation of the boost PFC converter [2][4] has been analyzed in details by many researchers. In practical circuits, it is much more difficult to arrange pure DC source, as well as the setup is much more expensive. So we are considering rectified dc in spite of pure DC. This DC contains several kinds of harmonics. We consider this kind of converter circuit and want to maintain the input power factor high.

Recently, power-factor-correction (PFC) circuits are widely used in power electronics. One of the most common and more attractive circuits than other is the boost power factor correction circuit operating at the

continuous conduction mode (CCM)[3] because of the unity power factor and the reduced current stress. But it has been found that the PFC boost converter may operate in discontinuous conduction mode (DCM) near the crossover point of the line voltage, in particular, the PFC converter may operate in DCM during the whole line cycle under the light load condition. It has been found that the transition from CCM to DCM can cause bifurcation and chaos.

Most prior researches introduced some assumptions that force the time-varying PFC system to be linear. They linearised the system as their assumption. They assumed a very huge output capacitance (not acceptable in industry) and it resulted in the time-invariant feedback signal that neglected the time-varying effect. Also, they replaced the input voltage with its root mean square (r.m.s.) value, neglecting the effect of its amplitude variation. Then, they introduced a small-signal equivalent circuit and the stability was examined by this linear model.

The PFC converter is nonlinear system[4] due to a multiplier using and a large variation of duty cycle. There is also present nonlinear term in its state equations. Here we will observe the nonlinear phenomena and bifurcation of this converter.

II. PFC BOOST CONVERTER AND ITS PROPOSED MODEL

Modelling, simulation and circuit analysis are done by MATLAB respectively. These not only help in developing a deeper understanding of PFC converters but are also extremely important tools for design verification and performance evaluation. These techniques help in the evaluation of a system without risking the huge cost and effort of developing and testing an actual converter.

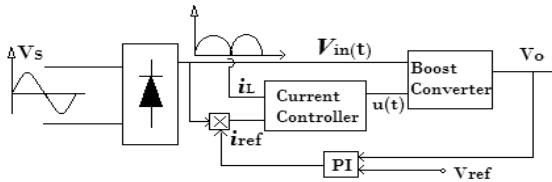


Fig. 1: Block Diagram of a boost PFC circuit

From the above Fig.1 the single phase sinusoidal voltage source v_s is rectified by diode bridge and the rippled DC voltage v_{in} is fed to the boost converter. The output voltage v_o (ripple is presented as the value of the capacitor is taken small) is obtained from load side. The output voltage v_o is compared with a reference voltage (DC) v_{ref} . We use an integral controller to get steady state value of error signal. i_{ref} is obtained after combining the the result of controller, v_{in} and inductor current (i_L). Now i_{ref} or i_L^* is compared with i_L . The duty cycle is maintained by the result of the comparator. The clock period and the value of the inductor are so chosen that the inductor current never falls to zero.

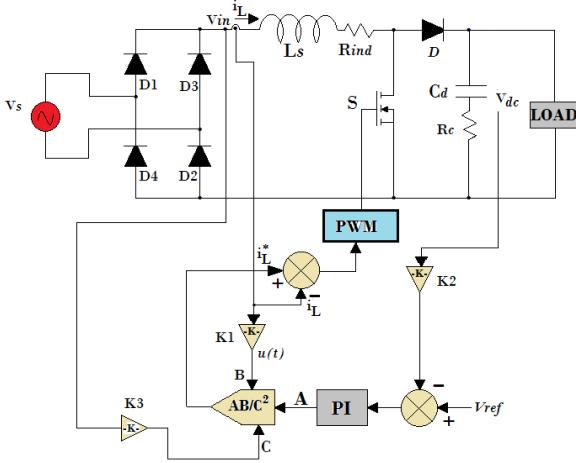


Fig. 2: Boost PFC ac-dc regulator under fixed-frequency Current Mode Control

Depending upon the block diagram we design the above model and derived the several expressions[2][4] which are given below.

Supply system:

Under normal operating conditions the supply system can be modelled as a sinusoidal voltage source of amplitude v_m and frequency f_s . The instantaneous voltage is:

$$v_s(t) = v_m \sin \omega t \quad (1)$$

where $\omega = 2\pi f_s$ electrical radians/second and t is instantaneous time.

In some topologies, the input is rectified line voltage $v_d(t)$ which can be given as:

$$v_d(t) = |v_s(t)| = |v_m \sin \omega t| \quad (2)$$

From the sensed supply voltage, an input-voltage template $u(t)$ is estimated for converter topologies with AC side inductor as:

$$u(t) = v_s(t)/v_m \quad (3)$$

The input-voltage template for converter topologies with a DC side inductor is obtained from:

$$u(t) = |v_s(t)|/v_m \quad (4)$$

Feedback controller:

PFC converters, like most power electronics systems, cannot function without feedback control. Fig.1 shows a block diagram of typical control scheme for PFC converters – the current mode control [3]. This control scheme ensures regulated DC output voltage at high input power factor. The output DC voltage regulator generates a current command, which is the amount of current required to regulate the output voltage to its reference value. The output of the DC voltage regulator is then multiplied with a template of input voltage to generate an input current reference. This current reference has the magnitude required to maintain the output DC voltage close to its reference value and has the shape and phase of the input voltage – an essential condition for high input power factor operation.

(i) Output voltage controller:

A proportional integral (PI) voltage controller is selected for zero steady-state error in DC voltage (rippled in nature) regulation. The o/p capacitor voltage v_{dc} (or v_o) is sensed and compared with the set reference voltage v_{ref} . The resulting voltage error $v_{e(n)}$ at the nth sampling instant is:

$$v_{e(n)} = v_{ref} - v_{dc(n)} \quad (5)$$

The output of the PI voltage regulator $v_o(n)$ at the nth sampling instant of the PI controller will be:

$$v_{o(n)} = v_{o(n-1)} + k_p \{ v_{e(n)} - v_{e(n-1)} \} + k_i v_{e(n)} \quad (6)$$

Here k_p and k_i are the proportional and integral gain constants, respectively. $v_{e(n-1)}$ is the error at the $(n - 1)$ th sampling instant. The output of the controller $v_{o(n)}$ after limiting to a safe permissible value is taken as the amplitude of the input current reference A (Fig. 2).

(ii) Reference current controller:

The input voltage template $u(t)$ obtained from the sensed supply voltage is multiplied by the amplitude of the input current reference A to generate a reference current. The instantaneous value of the reference current is given as:

$$i_L^* = AB / C^2 \quad (7)$$

where B is the input voltage template $u(t)$ and C is the input voltage feed forward component obtained by low-pass filtering the sensed input voltage signal.

Semiconductor switches:

Semiconductor switches, Mosfet S and Diode D are modelled as pure ON-OFF switches. No snubbers or non-idealities in the switches are modelled.

Load:

The converters are modelled as resistive loads having resistance R.

Power circuit:

The power circuit is modelled by first-order differential equations describing the circuit behaviour.

These modelling equations are obtained by application of Kirchoff's and Ohm's laws to the power circuit.

III. STATE EQUATIONS FOR MODELING OF BOOST CONVERTER

There are two states[1][3] of the circuit depending on whether the controlled switch is open or closed. When switch is closed, the current through the inductor rises and any clock pulse arriving during that period is ignored. The switch opens when reaches the reference current. When switch is open, the current falls. The switch closes again upon the arrival of the next clock pulse.

The State Equations during "ON" period

$$\frac{di_L}{dt} = V_{in}/L - (r_i * i)/L \quad (8)$$

$$\frac{dv_c}{dt} = -v_c/C(R + r_c) \quad (9)$$

The State Equations during "OFF" period

$$\frac{di_L}{dt} = V_{in}/L - i*(r_i + R*r_c/(R + r_c))/L - v_c*R/L(R + r_c) \quad (10)$$

$$\frac{dv_c}{dt} = (R*i - v_c)/C(R + r_c) \quad (11)$$

where,

V_{in} =Input Voltage

L = Inductor

C = Capacitor

i_L = Inductor Current

v_c = Capacitor Voltage,

r_i & r_c = Parasitic Elements

IV. SIMULATION OF PFC BOOST CONVERTER

Simulation of PFC Boost Converter is done by MATLAB 7.8R2009a. The model is totally designed by SimPowerSystem and Simulink blocks.

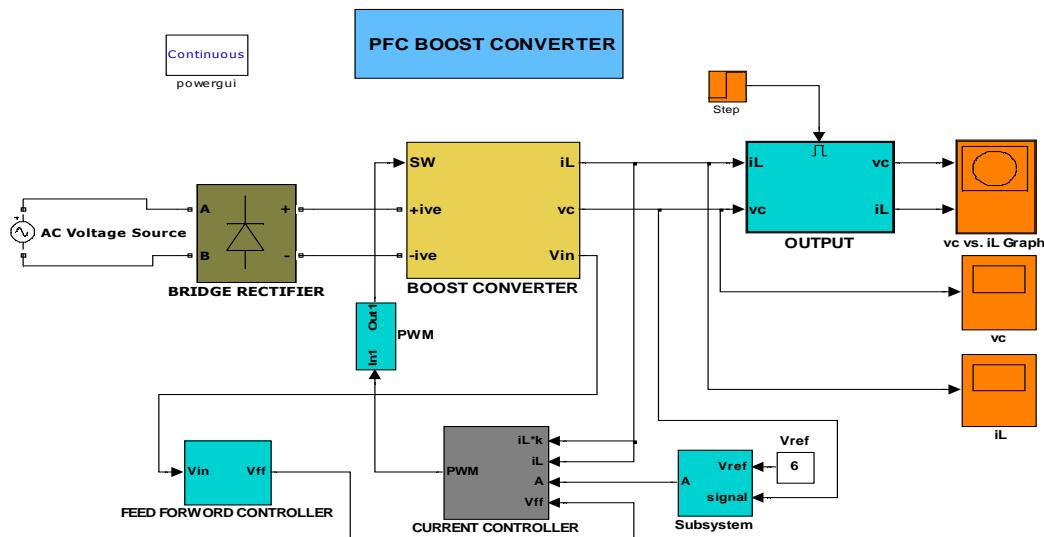


Fig. 3: Simulation of PFC Boost Converter

V. EXPERIMENTAL RESULTS

Here we are varying the value of Current Gain K1 (is placed in Current Controller Block, Fig.2) and we obtain the several periodic behavior of converter.

Case I (Period I Operation)

$V_s=220\sin w t$, $L=40\text{mH}$, $C=100\mu\text{F}$, $R=35\text{ohm}$, $K1=850$

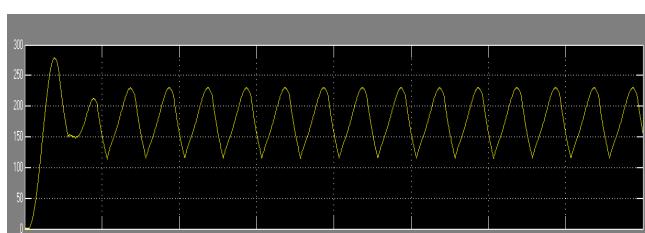


Fig.4(a) : O/P Voltage Waveform at Period I operation ($K = 850$)

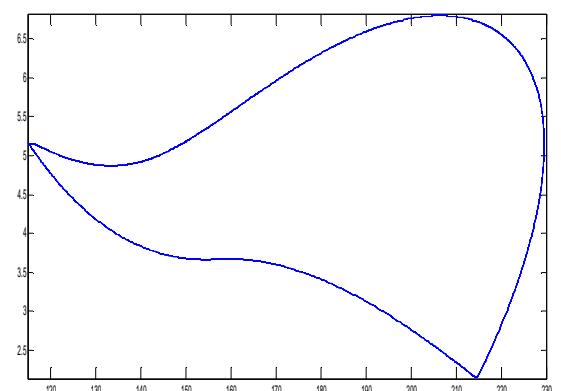


Fig.4(b) : Phase Plane Trajectory(Case I)
Capacitor Voltage vs Inductor Current (Period I)

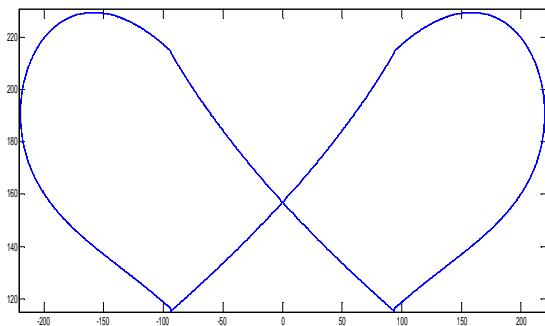


Fig.4(c) : V_{in} vs V_c (Case I)

Case II(Period II Operation)

$V_s = 220 \sin \omega t$, $L = 40 \text{ mH}$, $C = 100 \mu\text{F}$, $R = 35 \text{ ohm}$, $K_1 = 450$

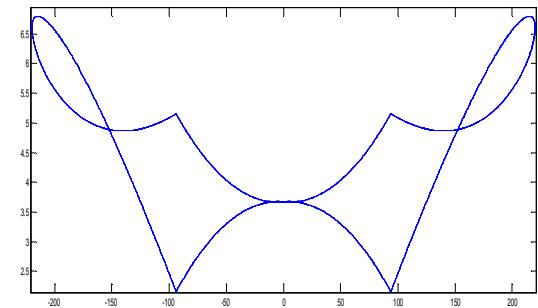


Fig.4(d) : V_{in} vs i_L (Case I)

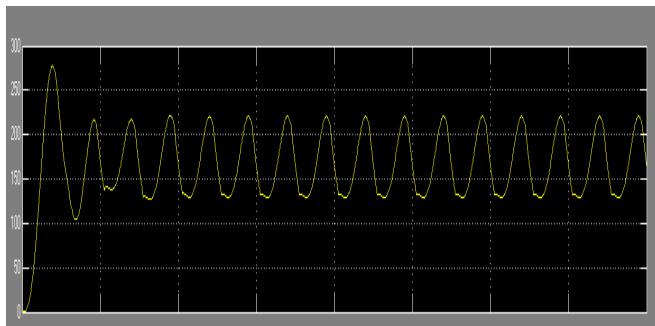


Fig.5(a): O/P Voltage Waveform at Period II operation ($K = 450$)

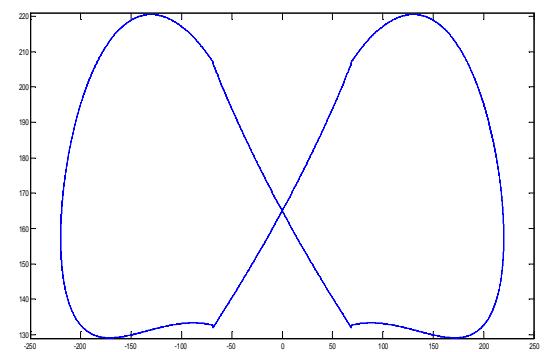


Fig.5(c) : V_{in} vs V_c (Case II)

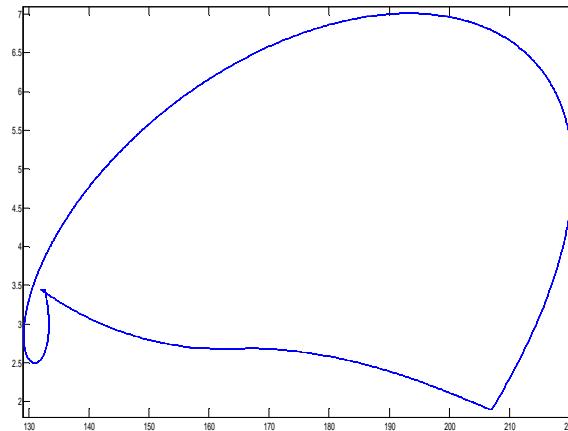


Fig.5(b) : Phase Plane Trajectory(Case II)
Capacitor Voltage vs Inductor Current (Period II)

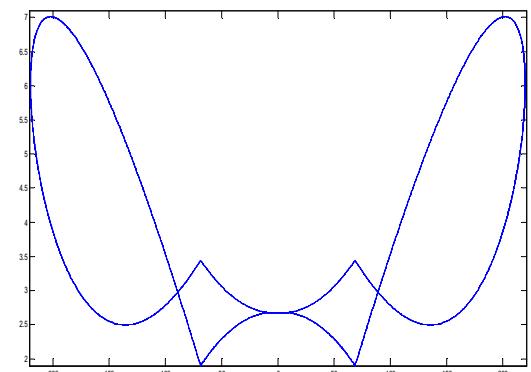


Fig.5(d) : V_{in} vs i_L (Case II)

Analysis of Experimental Results:

Here the state variables are inductor current (i_L) and capacitor voltage (V_c). From above results we see that case I(Fig.4b) is operating at period I condition[5] and case II(Fig.5b) is operating at period II condition[5]. The output voltage waveform in Fig.5(a) is much more ripple free than Fig.4(a). We get better result of output voltage at same value of inductor (L) and capacitor (C), only changing the value of current gain(K_1). The value of Capacitor(C) is chosen small just it operates as a boost converter. If we can decrease more values of current gain K_1 , the system will operate at chaotic region and we can get better ripple free output voltage. This is the main observation that we get

better output voltage profile at least value of capacitor. So, the investment is much more less than other conventional practical instruments.

VI. BIFURCATION DIAGRAMS

Bifurcation diagrams are obtained from FORTRAN and ORIGIN 5.0 software. The data files are obtained after executing the FORTRAN programme of State Equations (8), (9), (10), (11) of PFC Boost Converter and i_{ref} equ (7). This data files are plotted by ORIGIN 5.0. Results are given below.

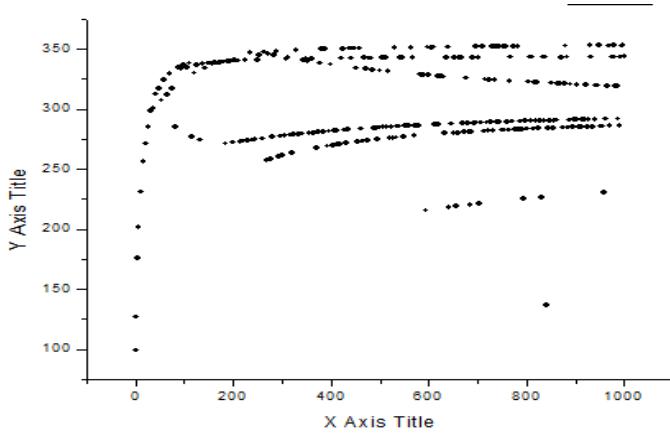


Fig.6(a):R vs Vc (R is varied 1 to 1000ohm with step of 0.5)

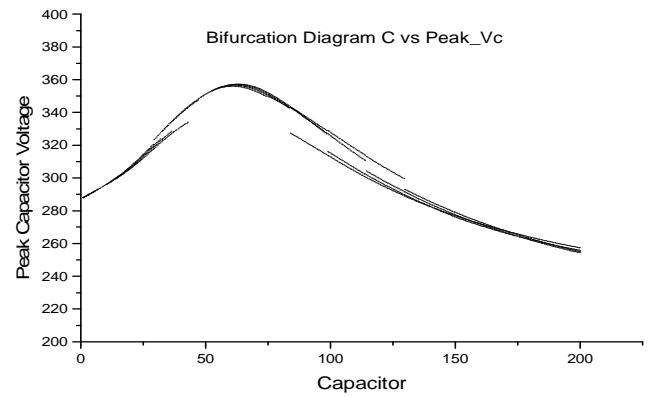


Fig.6(c):C vs Peak_Vc (C is varied 1 to 200ohm with step of 0.5)

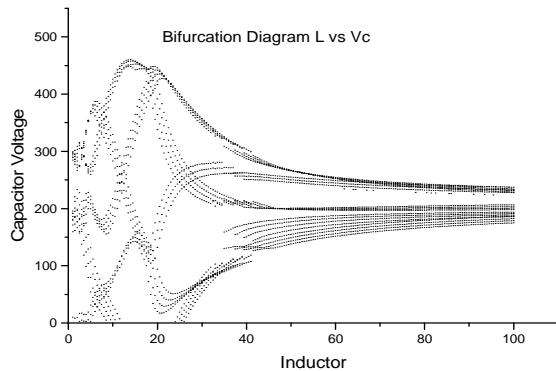


Fig.6(b):L vs Vc (L is varied 1 to 100ohm with step of 0.5)

Analysis of Bifurcation Diagrams:

The above bifurcation diagrams are much more differ from other conventional bifurcation diagrams. In conventional process the border [6] is fixed i.e. I_{ref} (or i_L^*) is constant. In our experiment I_{ref} (or i_L^*) is time varying nature i.e. order is time-varient. It is very difficult to analysis the bifurcation diagram properly. Actually the bifurcation is Period Doubling [5][6] in nature. The analysis is not given here. We are now working on analysis of digrams.

VII. CONCLUSION

The boost PFC converter with current mode control has been examined. Results highlight that the proposed model of practical pfc converter, experimental results and bifurcation diagrams. The value of current gain is decreased; the output capacitor voltage waveform is going to period I to period II i.e. period doubling bifurcation is observed. But the main benefit is the output voltage ripple is going less than the previous. In a DC/DC converter system, the input voltage is constant and therefore the dynamical behavior is periodic with the switching frequency. On the other hand, the input voltage of the boost AC/DC PFC converter system is periodic with the

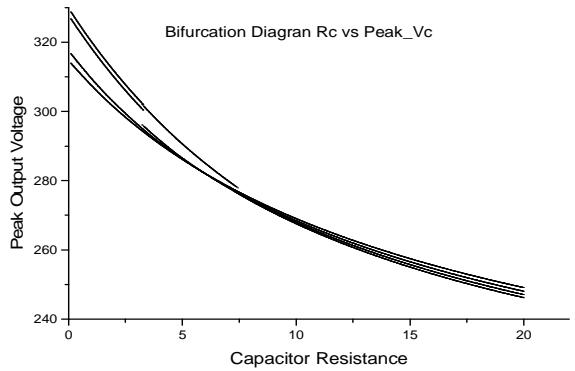


Fig.6(d):Rc vs Peak_Vc (Rc is varied 0.1 to 20ohm with step of 0.01)

line frequency. The results highlight that the dynamical behavior is periodic with the line frequency not with the switching frequency and simulation results are also agree with our statements.

REFERENCES

- [1] Soumitro Banerjee and Krishnendu Chakrabarty, "Nonlinear Modeling and Bifurcations in the Boost Converter" in IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 13, NO. 2, MARCH 1998.
- [2] Ashish Pandey, Dwarka P. Kothari, Ashok K. Mukerjee and Bhim Singh, "Modelling and simulation of power factor corrected AC-DC converters" in International Journal of Electrical Engineering Education.
- [3] Ammar Nimer Natsheh, "ANALYSIS, SIMULATION AND CONTROL OF CHAOTIC BEHAVIOUR IN POWER ELECTRONIC CONVERTERS" in Doctoral Thesis Submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of Loughborough University.
- [4] Yi-Jing Ke, Yu-Fei Zhou and Jun-Ning Chen, "Control Bifurcation in PFC Boost Converter under Peak Current-Mode Control" in IPEMC 2006.
- [5] Soumitro Banerjee, "Dynamics for Engineers" (John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England).
- [6] Soumitro Banerjee and George C. Verghese, "NONLINEAR PHENOMENA IN POWER ELECTRONICS"(Published by John Wiley & Sons, Hoboken, NJ)

Control of Chaos in DC-DC Boost Converter with Non-smooth DC Supply

Subhankar Dam
Birbhum Institute of Engg. & Tech.
subhankardam@gmail.com

Dr. Pradip Kumar Saha
Jalpaiguri Govt. Engg. College
p_ksaha@rediffmail.com

Abhrajit Saha
Birbhum Institute of Engg. & Tech.
saha.abhrajit@gmail.com

Dr. Goutam Kumar Panda
Jalpaiguri Govt. Engg. College
g_panda@rediffmail.com

Abstract-- In this work we have analysed the nonlinearity of dc-dc boost power converter in continuous current mode of operation. The circuit model is used to describe the chaotic phenomenon of this power electronics converter with a nonsmooth dc as input. Finally we have proposed a self-controlling time dependent delayed feedback controller to quench the chaos of the system.

Keywords: boost converter, nonlinear dynamics, chaos, chaos control

I. INTRODUCTION

The on and off time equations of power electronic circuits are inherently linear but nonlinearity is introduced due to nonlinear switching and they exhibit nonlinear phenomena [1]-[4]. It has been proved that these switching circuits are rich source of nonlinearity. A large number of papers has been published after a large investigation. Some researchers were interested to study the methods of control of chaos and that idea to control the nonlinear phenomena in dc-dc converters. In all those studies the source was a pure dc supply. But in practice the dc we get or we derive is not free from ripple as it is obtained from ac-dc rectifier circuit after rectification and filtration. So, a non-smooth dc supply has to be taken for more practical performance analysis. As it is demonstrated that current mode controlled boost converter prone to subharmonics and chaos, a nonlinear model is developed and a detailed study of its bifurcation behaviour of it on variation of its parameter is carried out. [1] & [5] used a large signal analysis for continuous time system while [6] proposed that discrete model should be developed for power electronics circuits as mapping in the form

$$x_{n+1} = f(x_n) \quad (1)$$

Maps are also derived when observations are made at every clock instant instead of every switching instant. All of these above researches are carried out with a smooth dc supply for current controlled boost converter. In this paper we have developed rectified dc supply for input to boost converter to confirm the

very practical situation. Finally we have proposed a time dependent feedback controller [7] to control the nonlinear phenomena and chaos in the converter. The method is based on the stabilization of unstable periodic orbits (UPOs) embedded within a strange attractor. This is achieved by making a small time-dependent perturbation in the form of feedback to an accessible system parameter. The method turned the presence of chaos into an advantage. Due to an infinite number of different UPOs embedded in a strange attractor, a chaotic system can be tuned to a large nos. of distinct periodic regimes by switching the temporal programming of small parameter perturbation to stabilize different periodic orbits. This method has been successfully applied to some experimental experiments.

The rest of the paper is organised as follows: section-II contains the basic boost converter with current limiter controller; section-III delivers the bifurcation phenomenon of boost converter considering different bifurcation parameter; section-IV gives a approach to control the non-linearity using self-controlling delayed feedback controller; finally section-V gives some concluding remarks.

II. BOOST CONVERTER

A boost converter (fig.1) consists of a control switch (s), an uncontrolled switch (D5), an inductor (L), a capacitor (C), a load resistor (R), an internal resistance of the inductor (r_i), an internal resistor of the capacitor (r_c) is taken for our study. The switch is controlled by a feedback path consisting of a comparator, a flip flop and a driver. The supply is given to the converter from a single phase or three phase rectified supply through three phase or single phase rectifier. The comparator compares the inductor current and reference current.

Boost converter has the higher output voltage than the input voltage. when the control switch is turned on, the inductor stores energy from source and during off time of the control switch the inductor transfers the stored energy to the connected capacitor across the load. The circuit is operated in continuous current

mode which means the inductor current never becomes zero.

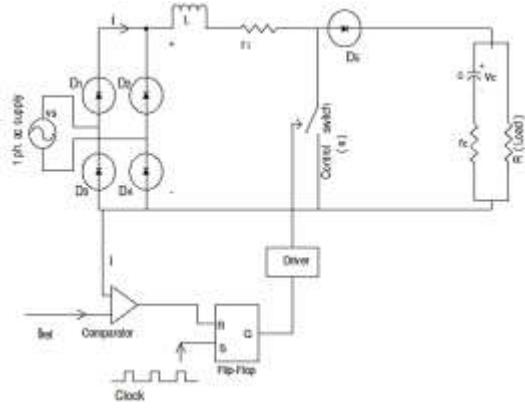


fig1: Basic Boost Converter Circuit (here the input is rectified dc voltage)

The circuit has two states depending on whether the control switch is on or off. When the switch is on the state equations are

$$L \frac{di}{dt} = V_m - ir_i \quad (2)$$

$$C \frac{dv_c}{dt} = -\frac{v_c}{R+r_c} \quad (3)$$

and the state equations during off period (when the switch is off) are

$$L \frac{di}{dt} = V_{in} - ir_i - i \frac{Rr_c}{R+r_c} - v_c \frac{R}{R+r_c} \quad (4)$$

$$C \frac{dv_c}{dt} = -\frac{v_c}{R+r_c} + i \frac{R}{R+r_c} \quad (5)$$

where $V_{in} = |V_m \sin \omega t|$
for single phase full wave $0^\circ < \omega t < 180^\circ$
for three phase full wave $60^\circ < \omega t < 120^\circ$

Switching logic of the circuit is that when switch is closed the inductor current rises and it continues till it continues till it reaches to I_{ref} , ignoring any arriving clock pulse. As soon as the inductor current reaches to I_{ref} , the inductor current falls. The switch closes again the on the arrival of the next clock pulse. It is explained in the following fig2

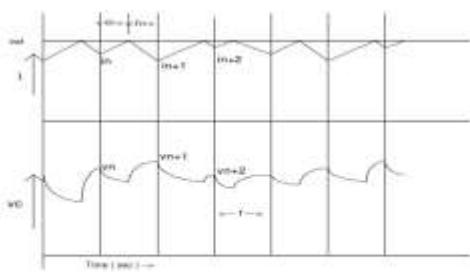


Fig. 2: Time Plot of Capacitor Voltage and Inductor Current of the Boost Converter with Clock Pulses

A. The Model of Boost Converter

An expression is derived for the mapping from one switching instant to the next considering the continuous conduction of the converter. At the starting of on period or off period, the initial condition (i.e. at $t=0$) be $i = i_0$ and $v_c = v_0$. Final condition of one state ate the initial condition of the next other state. The switch turns off when the inductor current reaches the reference current I_{ref} . Using Laplace equations and considering initial conditions, the expressions of time response for both on time and off time are established. During on period, the equations of v_c and i are obtained from (2) and (3) are

$$v_c(t) = v_0 e^{-t/(R+r_c)} \quad (6)$$

$$i(t) = \frac{V_m \omega L}{\omega^2 L^2 + r_i^2} [e^{-r_i t / L} - \cos \omega t + \frac{r_i}{L} \sin \omega t] - i_0 e^{-r_i t / L} \quad (7)$$

During off period the equations of v_c and i are obtained from (4) and (5) as follows

$$\begin{aligned} i(t) = & \frac{V_m \omega}{Rk_3} \{n - C(R + r_c)m\omega\} \sin \omega t + \frac{V_m \omega}{Rk_3} \{m + \\ & nC(R + r_c)\} \cos \omega t + [(0.5K_6 - 1) \left(\frac{V_m \omega m}{k_3} + \right. \\ & \left. 1 \right) - \left\{ \frac{V_m \omega}{k_3} (0.5mk_6 - 1) \right. \\ & \left. - \frac{L}{k_3 i_0} + \frac{1}{Lk_8} (0.5Lk_6 - k_2) \right\}] e^{-\frac{k_0 t}{2}} \cos k_8 t + \\ & [\left\{ \frac{V_m \omega m}{k_3} + 1 \right\} + (1 - 0.5k_6) \left\{ \frac{V_m \omega}{k_3 k_8} (0.5mk_6 - 1) - \right. \\ & \left. - \frac{L}{k_3 k_8} i_0 + \frac{1}{Lk_8} (0.5Lk_6 - k_2) \right\}] e^{-0.5k_6 t} \sin k_8 t \end{aligned} \quad (8)$$

$$\begin{aligned} v_c(t) = & V_m \omega / k_3 [m \cos \omega t + \left(\frac{n}{\omega} \right) \sin \omega t - \\ & me^{-0.5k_6 t} \cos k_8 t + \left\{ \frac{0.5mk_6 - r}{k_8} \right\} e^{-0.5k_6 t} \sin k_8 t] \\ & - \left(\frac{Li_0}{k_3 k_8} \right) e^{-0.5k_6 t} \sin k_8 t - e^{-0.5k_6 t} \cos k_8 t + \\ & \left(\frac{1}{k_8 L} \right) (0.5Lk_6 - k_2) e^{-0.5k_6 t} \sin k_8 t \end{aligned} \quad (9)$$

where,

$$k_1 = \frac{R}{R + r_c}, k_2 = r_i + r_c k_1, k_3 = CL/k_3$$

$$k_4 = \left(\frac{Ck_2}{k_1} \right) + \frac{L}{R}, k_5 = \left(\frac{k_2}{R} \right) + k_1,$$

$$k_6 = \frac{k_4}{k_3}, k_7 = \frac{k_5}{k_3}, k_8 = \sqrt{k_7^2 - k_8^2}/4,$$

$$m = k_6 / (k_6^2 \omega^2 + k_7^2 + \omega^4 - 2\omega^2 k_7),$$

$$r = (2k_7^2 - 3\omega^2 k_7 + \omega^4 + k_6^2 \omega^2) /$$

$$\{\omega^2 (2\omega^2 k_7 - \omega^4 - k_7^2 - k_6^2 \omega^2)\}$$

$$n = (k_7 - \omega^4) / (2\omega^2 k_7 - \omega^4 - k_7^2 - k_6^2 \omega^2)$$

Here we consider $a^2/4 > b$, which is the most important from a practical point of view,

$$\text{where } a = \frac{L+r_i C(R+r_c)+Rr_c C}{LC(R+r_c)}, \quad b = \frac{R+r_i}{LC(R+r_c)}$$

If at any instant (t_n) the value of v_c and i are v_n and i_n then for the next instant (t_{n+1}) are v_{n+1} and i_{n+1} taking $v_0 = v_n$ and $i_0 = i_n$.

The on-time (t_{on}) can be calculated from (7) and the off time (t_{off}) can be calculated from (8).

The map based model can be used to predict the behaviour of the system under different parameter combinations. Here the boost converter is represented by two ways, first by locating the peak of the capacitor voltage and second by sampling it in synchronism with the clock.

III: BIFURCATION PHENOMENA OF THE BOOST CONVERTER

The operation of the boost converter is considered here from clock point of view. As the clock pulse has an externally determined periodicity as in numbers of clock pulses in a period of waveform. So the sampling is done at each clock pulse. Also we detect the peak capacitor voltage for analysis. The model provides a fast and easy way of obtaining bifurcation diagram. Some initial readings are eliminated due to initial transient. There are nine parameters in the system, input voltage ($V=V_m/\sqrt{2}$) , load resistance (R) , inductance (L), Capacitance (C) , reference current (I_{ref}) , clock frequency, input voltage frequency, parasitic r_i , parasitic r_c .

A. Results

Parameters are set for the converter:

$I_{ref} = 4$ amp, $r_i = 0.1\text{ohm}$, $R=20\text{ohm}$, clock frequency = 500Hz. , input voltage frequency = 50 Hz. $L=27\text{ mH}$, $C=120\mu\text{F}$. For these set of values and with single phase full wave rectified input voltage, the phase plot of state variables is plotted and time responses of the two state variables are plotted. By these figures we observe that the system operates at higher period (i.e. output has subharmonics).

i. V as bifurcation parameter: Keeping other parameters at set values, V is varied from 15 to 77 volts in steps of 0.5 volt. The bifurcation diagram shows different chaotic and frequency regions (fig. 3, 4 & 5) and this is also confirmed by phase plot (fig. 6) and time plots (fig. 7 & 8). Chaotic windows are observed for a range of input voltage. Peak voltage of capacitor is stabilized from 57 volts.

ii. I_{ref} as bifurcation parameter: Keeping all the parameters at values I_{ref} is varied from 1 to 7 amps in step of 0.05 amps. The bifurcation diagrams different chaotic regions and different frequency

regions (fig 9, 10 & 12) which is confirmed by phase plot (fig. 14) and time plots (fig. 11 & 13).

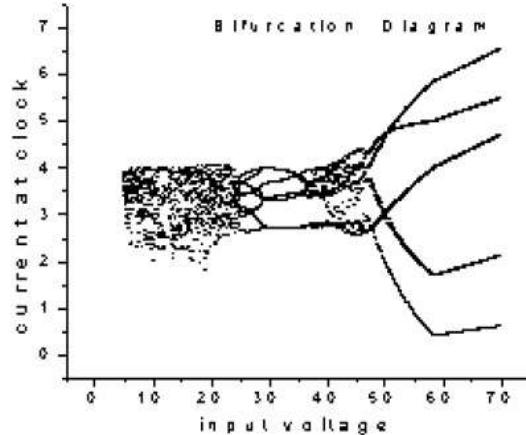


Fig. 3

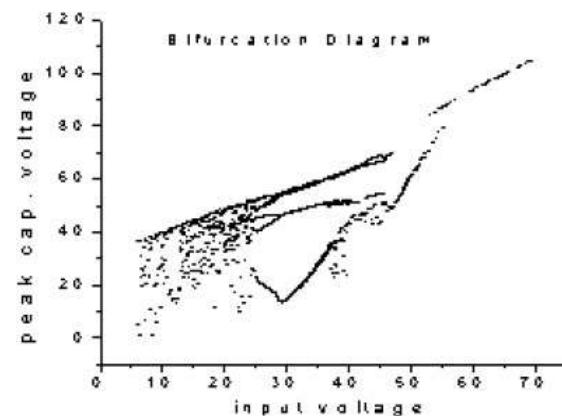
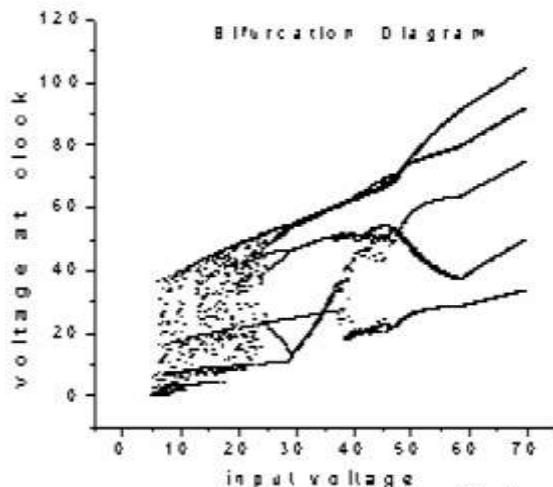


Fig. 5

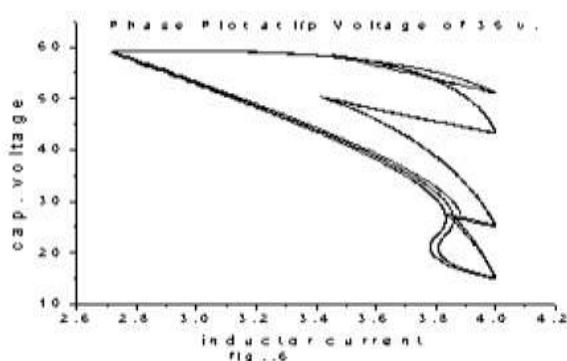


Fig. 6

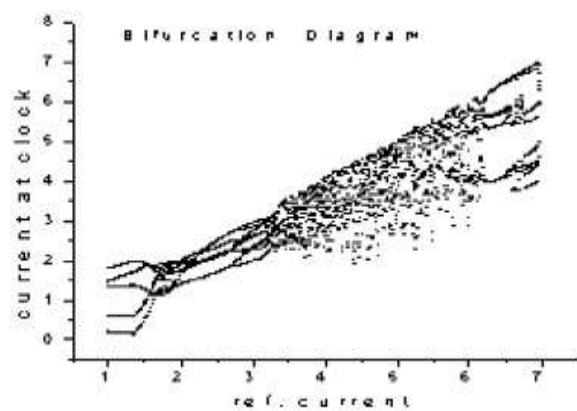


Fig. 10

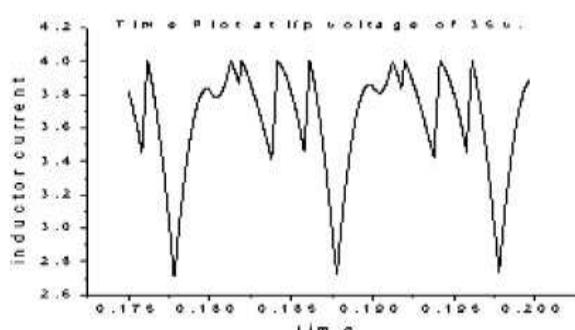


Fig. 7

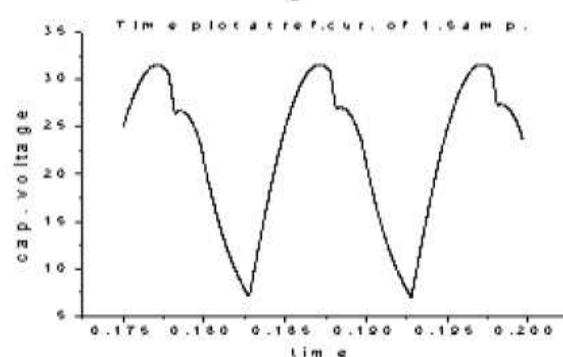


Fig. 11

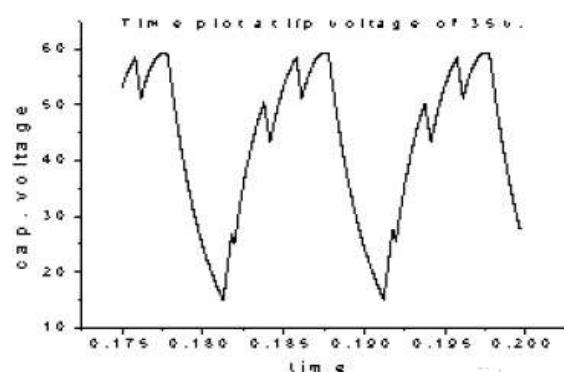


Fig. 8

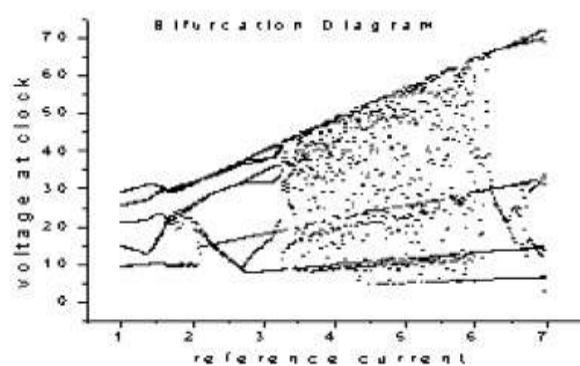


Fig. 12

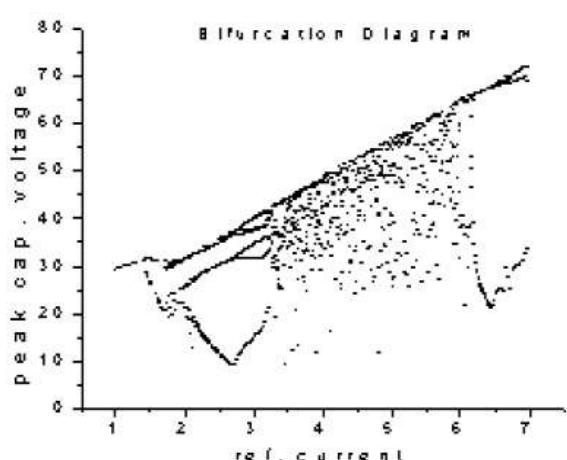


Fig. 9

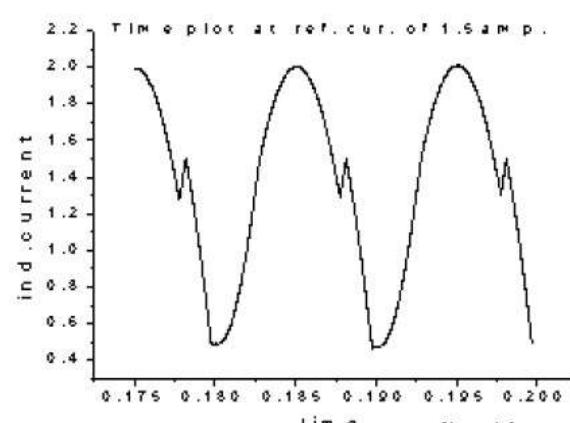


Fig. 13

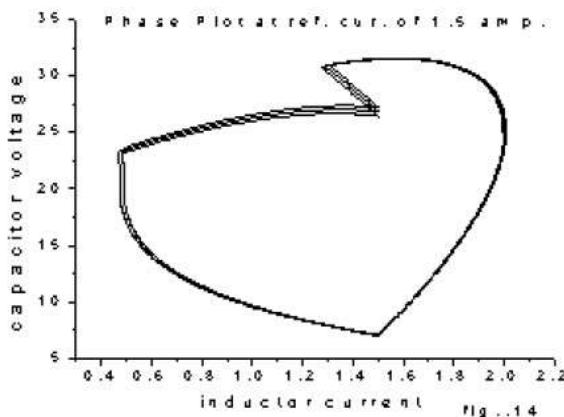


Fig. 14

IV: CONTROL OF CHAOS USING TIME DELAYED FEEDBACK CONTROLLER

A. The Delayed Feedback Control Strategy

The idea of DFC method also known as Pyragas method for the name of it's inventor consists in substituting the external signal (say $y(t)$) for the delayed output signal (say $y(t-\tau)$). In other word we use a perturbation of the form

$$F(t) = K(y(t - \tau) - y(t)) = KD(t)$$

Here τ is the delay time. If the time coincide with the period of the i^{th} UPO (Unstable Periodic Orbit), then the perturbation becomes zero for the solution of the system corresponding to this UPO $y(t)=y_i(t)$. This means that the perturbation in the form above does not change the solution of the system corresponding to this i^{th} UPO. Choosing an appropriate gain K one can achieve the stabilization. No external perturbation or computer is needed for this control. The control is achieved by the use of the output signal, which is fed in a special form into the system input. The difference between the delayed output signal and the output signal itself is used as a control signal. That is how the feedback performs the function of self-control and the method becomes a continuous one for control. Only a simple delay line is required for this feedback. To achieve the stabilization of the desired UPO, two parameters, namely, the time of delay τ and the weight K of the feedback should be adjusted in experiment. The amplitude of the feedback signal can be considered as a criterion of UPO stabilization. When the system moves along it's UPO the amplitude is extremely small. The delayed feedback control is applicable to any dimensional system. Therefore the stabilization

is achieved through additional degrees of freedom change only the Lyapunov exponent of the UPOs, so that they can become stable. The block diagram is shown below

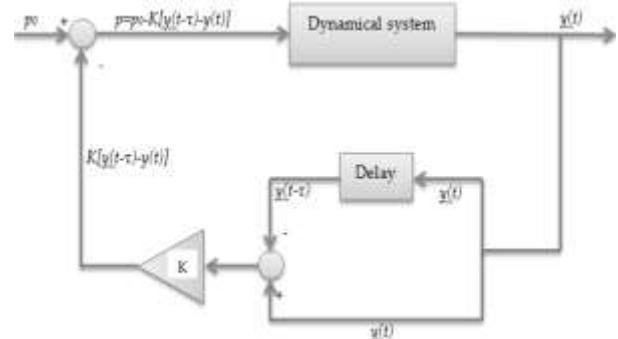


Fig. 15

The DFC method is a reference-free and makes use of a control signal obtained from the difference between the control state of the system and the state of the system delayed by one period of the UPO. Alternatively, the DFC method is referred to as time delayed auto-synchronization.

B. Controlling the Nonlinearity

As shown in the last section the system is operating in higher period which is undesirable from practical point of view because it restricts the operating range of the system. So it is desirable that the system operates in period=1.

In this section, a time-delayed state feedback controller is introduced to control the bifurcation and chaos in the current mode controlled dc-dc boost converter. The capacitor voltage (v_c) is the chosen state and it's time delayed state to constitute the following controller.

$$\frac{dv_c}{dt} = k(v_c(t - \tau) - v_c(t)) \quad (10)$$

The corresponding on time state equations becomes

$$L \frac{di}{dt} = V_m - ir_i \quad (11)$$

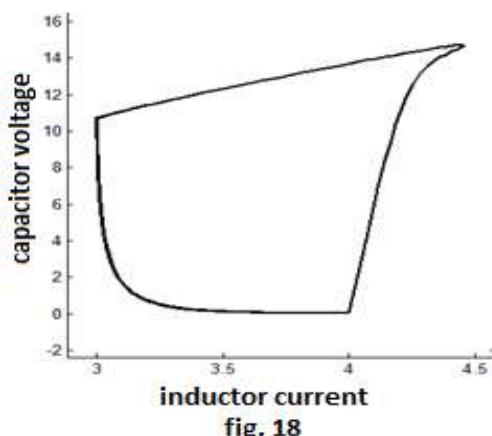
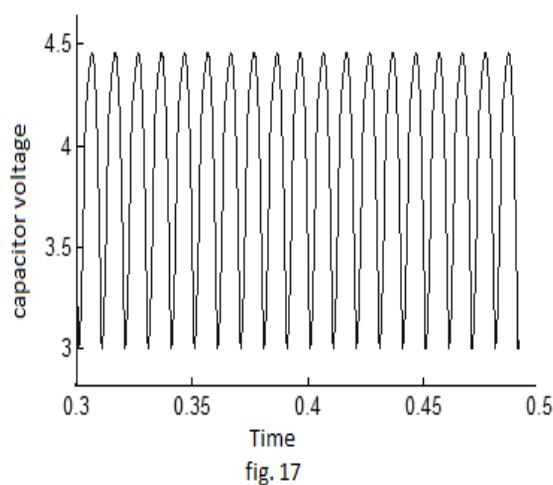
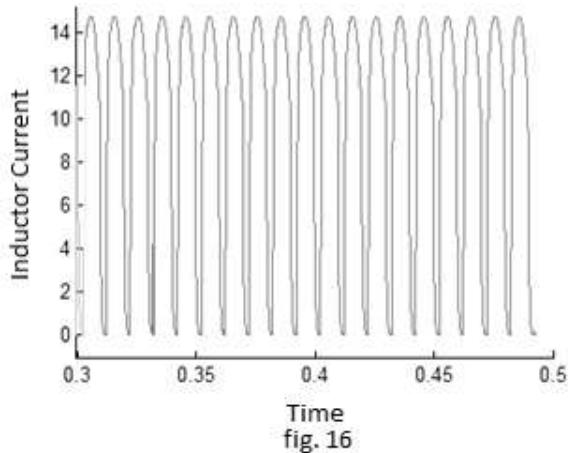
$$C \frac{dv_c}{dt} = -\frac{v_c}{R+r_c} - Ck(v_c(t - \tau) - v_c(t)) \quad (12)$$

and so the off state equations will be

$$L \frac{di}{dt} = V_{in} - ir_i - i \frac{Rr_c}{R+r_c} - v_c \frac{R}{R+r_c} \quad (13)$$

$$C \frac{dv_c}{dt} = -\frac{v_c}{R+r_c} + i \frac{R}{R+r_c} - Ck(v_c(t - \tau) - v_c(t)) \quad (14)$$

Using the same data as for the chaotic boost converter we simulate the system using delay time $\tau = 30$ and feedback gain $K=1000$. The performance of the controller is studied. The time response and phase plot of the capacitor voltage with proposed controller is shown (Fig. 16 , 17 ,& 18) which show that the Fuzzy controller has effectively compensated the bifurcation phenomenon of the system and the system is operated at period 1 as desired by us.



Simulation results shows that the system undergoes a stable limit cycle and the state variables possesses period 1.

V. CONCLUSION

A discrete model of the dc fed boost converter including parasitic effect is presented here single phase rectified dc voltage input, which can be used to plot bifurcation diagrams for analysis and prediction. It is observed that the boost converter exhibits a variety of bifurcation phenomena. There are both period multiple and chaotic regions for different parameter ranges. To control the nonlinear dynamics of the system we build a time delayed state feedback controller which generates a time perturbed version of output signal of any state variable as the control signal. We vary the delay time to control and synchronize the nonlinear dynamics as well as chaos. The performance of the controller is analysed numerically and found satisfactory. It is observed that the controller can retain the signal of any periodic or even aperiodic of any state variable to a periodic one.

REFERENCE

- [1] S.Banerjee & K.Chakraborty, "Nonlinear Modelling and Bifurcations in the Boost Converter," IEEE transactions on power electronics, Vol. 13, No.2 march 1998. pp 252-260.
- [2] J.R.Wood, "Chaos a real phenomenon in power electronics," Appl. Power Electron. Conf. Baltimore.MD.1989.
- [3] J.H.B.Deane and D.C.Hamill, "Instability, Subharmonics and Chaos in Power Electronics Circuits," IEEE Transaction on Power Electronics, Vol.5,1990. pp 260-268.
- [4] D.C.Hamill, "Power Electronics: A field rich in nonlinear dynamics," Nonlinear dynamics of Electronic Systems. Dublin, Irelnad.1995.
- [5] J.H.B.Deane and D.C.Hamill, "Chaotic Behaviour in current mode controlled dc-dc converter," Electron. Lett. Vol. 27, no. 13, 1991, pp 1172-1173.
- [6] J.H.B.Deane, "chaos in a current mode controlled dc-dc converter," IEEE Transaction on Circuits & Systems-I.Vol. 39, 1992, pp-680-683.
- [7] K.Pyragas, "Continuous control of chaos by self – controlling feedback," Phys. Lett. A, vol. 170, no. 6, pp. 421–428, 1992.

Understanding some Physico-Chemical Components for Estimation of Eutrophication of an Aquatic System

Authors Name: Ms. Susmita Bhattacharya

Department of Basic Science & Humanities
Institute of Engineering and management
Kolkata, India
susmitaiem@rediffmail.com

Authors Name: Dr. Phani Bhusan Ghosh
Department of Basic Science & Humanities
Institute of Engineering and management

Kolkata, India
pbghosh2007@rediffmail.com

Authors Name: Mr Prabir Kumar Das
Department of Basic Science & Humanities
Institute of Engineering and management
Kolkata, India
prabir80_iem@rediffmail.com

Abstract

Eutrophication, the natural as well as anthropogenic aging processes of aquatic system is one of the most serious threats throughout the world due to its undesirable consequences. Considerable efforts are on to find out ways and means of its restoration by utilizing ecological engineering processes to bring back the stability of aquatic system. Until and unless, the status of ecosystem in terms of eutrophication is known, no concrete action plan can be either conceived or undertaken. Although several indices are used involving physical, chemical and biological, for its assessment, yet these are time consuming and not cost effective. In the light of the above facts, present study is an attempt to highlight some simple methods comprising of only one or two water quality components for assessing this important phenomenon. The study may serve as a tool to the ecologists/ environmentalists to plan for eco-restoration of the aquatic system.

Keywords- Eutrophication, Assessment, Physico-chemical components, Aquatic system.

Introduction

Eutrophication is a process in which enhanced plant growth or algal bloom is caused by the presence of excess amount of dissolved nutrients of the elements of phosphorous (P) and nitrogen (N) in the system (Hutchison, 1973). It is one of the most serious threats and is considered as the biggest pollution problem, where it can cause anoxia, hypoxia, deterioration of water quality (Smith, 2006) and damages in ecosystem destroying the habitat as well as diversity (Andersen et all, 2006; Bishop et all, 2006)). Research has shown that approximately two thirds of USA's coastal rivers and bays have been moderately to severely affected by eutrophication. The United Nations Environment Programme (UNEP, 2006)

has identified 146 oxygen starved coastal areas in the world. Even the inland and brackish water fish productivity also

reduces after some years of cultivation solely due to this phenomenon. Hence, considerable efforts have been applied in this regard to find out the ways and means of eco-restoration to bring back the stability of the ecosystem. However, no stringent management action plan can ever be conceived until and unless the health of the ecosystem in terms of eutrophication is primarily known. For this, it is necessary to have a clear idea about the physico-chemical changes occurring in the aquatic system at different stages of this phenomenon. Several indices (Karyds, 2009) involving physical, chemical and biological components for its assessment are used, but these are time consuming as well as not cost effective.

In the light of the above facts, present paper is an attempt to highlight some important water quality components for an early as well as easy detection of different trophic stages in terms of eutrophication. The study may serve in one way as a ready-made tool for finding out the exact status of an ecosystem and in other way for taking effective measures of restoration of a deteriorated ecosystem.

Assessment of Eutrophication

In the present study three water quality components are discussed of which the first one dealing with nutritional elements in water and the others are dealing with variation of Dissolved Oxygen (DO) and hydrogen ion concentration (pH) of the medium. These methods are based on the changes of chemical components in aquatic medium during eutrophication. One interesting thing about these parameters is that in cases of pH & DO at least two observations, one at early morning and other at mid afternoon are needed to have the ranges of variation for ecosystem classification. On the

other hand, in case of dissolved nutritional analysis single sampling is enough for this assessment.

From measurement of Nutrients

As it is well known that unusual algal bloom occurs extensively in presence of excess amount of nutrients like Nitrogen (N) and Phosphorus (P) that are the most essential nutritional elements for biota and are incorporated during the production of organic matter. But these nutrients are known to remain in different forms, partly with associated or partly in free inorganic forms. Algae prefer only easily available inorganic forms and thus, these forms of nutrients are called as reactive components. So the implications obtained from the analysis of total nutrients and from analysis of available forms may however, differ in some way. Therefore, analysis of total nutrient content in water provides approximate trophic status of the ecosystem. As for example, total P and N content (Dodds et all, 1998) in microgram per litre in oligotrophic-mesotrophic boundary is 25 and 700, and mesotrophic-eutrophic boundary is 75 and 1500 respectively. This relationship is not found always true, because combined forms of nutrients are not readily used up by algae. On the other hand, the ratio of total free inorganic forms of N ($\text{NO}_2\text{-N}$, $\text{NO}_3\text{-N}$, $\text{NH}_3\text{-N}$) to inorganic P (PO_4^3-P) is known to provide good indicator for eutrophication along with trophic status and could be effectively used. The idea of N/P ratio is originated from the fact of basic elemental composition of organic matter of the aquatic algae as 1P: 16N: 106C in terms of molar ratio (Redfield, 1958). This indicates that the accumulated N/P ratio in cell is about 16 and all types of algae can easily assimilate this proportion of N and P into their body. As a result, the system provides large biodiversity. However, the availability of such requisite type of N/P ratio in water is very rare, because of changing of N and P content due to bio-geochemical processes occurring within the system during passage of time or due to introduction of these nutrients with different ratio from anthropogenic sources. If this ratio becomes very less due to presence of large amount of P, then heterocystous cyanobacteria become active and can increase N value through aerial molecular nitrogen fixation to maintain the ratio of 16 (Wetzel, 2001). This results in proliferation of these algae extensively and releases some compounds of cyanotoxin, which kills the other algae in the system, thereby causing reduction in diversity of the biotic community. If N/P ratio increases due to less supply of available P, then the group of algae which have the capacity to accumulate more P than their requirement can only flourish (Wetzel, 2001) to maintain the ratio with the help of accumulated P and could result in eutrophication of the system. Hence, the estimation of the level of dissolved N and P together with their ratio values is a good measure of assessing the indication of eutrophication. N/P ratio ranging from 12 to 20 represents oligotrophic condition and more or less than this range indicates eutrophication.

From measurement of dissolved oxygen

Extensive algal production is always accomplished with the evolution of O_2 in the medium according to the photosynthetic chemical reaction. But the extent of O_2 production does not remain same during the whole days. It has been known that photosynthesis begins during starting of the day and the rate gradually increases along with the increase in solar light intensity in a day. Accordingly, O_2 production rate also follows the same sequence with comparatively lower values in morning hours, reaches maximum at about 3-4 pm and again decreases after the end of the day due to absence of light. Thus, aquatic system undergoes diel O_2 variation in level and this diel variation is more pronounced during eutrophication (Wetzel, 2001). So from the measurement of O_2 in water at different times of the day and finding out the rate of O_2 evolution, the degree of eutrophication can be estimated easily.

Tropic structure like oligotrophic, mesotrophic and eutrophic of any aquatic system can be more quantitatively assessed by using percent of O_2 saturation instead of O_2 content alone. Because, estimation of O_2 gives idea only of the hygienic condition of water and it neither represent any ecological significance nor inform how much it is either less or more than the actual value that water should contain at the prevailing environmental condition. The percent of O_2 saturation measures the degree of un-saturation or supersaturation in an aquatic system and is calculated as follows.

$$\% \text{ O}_2 \text{ saturation} = C / C_1,$$

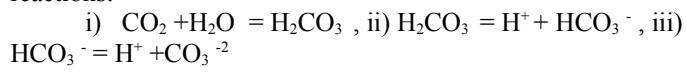
Where, C_1 and C are the concentration of calculated DO at saturated condition (APHA, 2005) and observed DO respectively. During eutrophication, excessive production of organic matter at surface in day time results with higher O_2 concentration and with supersaturation values throughout euphotic zone and at night, bacterial community reduces O_2 content for degradation causing lower percent saturation of O_2 . Hence, large variation from un-saturation to supersaturation indicates eutrophic (higher production) condition of the system and minimum variation of this components is indicative of (un-productive) oligotrophic. So trophic classification (Ghosh et all, 2010) is done as follows,

| % O_2 saturation | tropic status |
|---------------------------|---------------------------------|
| 40—300 | eutropic (highly productive) |
| 70---200 | mesotrophic (medium productive) |
| 80---130 | oligotrophic (un-productive) |

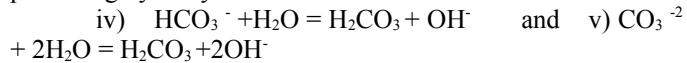
From measurement of pH in water

pH is used as scale of hydrogen ion concentration and is an important environmental factor, the variation of which among other causes, are linked with species composition and life processes of animal and plant communities inhabiting there.

Most natural waters are generally alkaline or near alkaline with pH more than 7. This is due to dissolution of sufficient quantities of carbonates and primary minerals on earth by the influence of dissolved CO₂ in water and usually remains in equilibrium condition as CaCO₃ + CO₂ + H₂O = Ca(HCO₃)₂. This CO₂--HCO₃--CO₃⁻² equilibrium system is the major buffering mechanism of the aquatic system due to following reactions:



Again this bicarbonate and carbonate undergo hydrolysis producing hydroxyl ion in water as follows.



This OH⁻ ion actually serves the buffering purpose and tends to resist change in pH as long as this equilibrium persists. An addition of H⁺ neutralizes OH⁻ ions, but more OH⁻ is formed immediately by reaction of carbonate with water (reaction- v). Similarly when OH⁻ is added, the reaction ii) and iii) will follow and consequently, pH remains essentially unaltered. The solubility of CO₂ increases markedly in water that contains carbonates. A definite amount of free CO₂ will remain in solution after equilibrium is reached and this amount increases rapidly with increasing bicarbonate content derived from carbonates (Wetzel, 2001).

If the observation of pH level at early morning and mid day lie between 7.5 and 8.0, 7.2 and 8.5, and 7.0 and 10.0, then trophic classification could be done as oligotrophic, mesotrophic and eutrophic respectively.

Conclusion

Eutrophication is a major problem of aquatic system and it disturbs the balance of the system creating undesirable consequences. For planning efficient control measure, early detection is very important. For this, we need to know, the exact physico-chemical status of any aquatic system at different trophic levels. From the above discussion it is quite clear that for an aquatic system to reach eutrophic state, nutrient enrichment (especially N:P), level of DO and pH are three major parameters and there is variation in quantity of these parameters at different trophic stages in an aquatic system . These variations can be utilized for estimating the exact trophic status of the system

Acknowledgement

Authors are grateful to the Director and Principal of Institute of Engineering and Management for their constant inspiration and support during the tenure of research.

References:

- APHA. 1992. Standard Method of Examination of Water and Waste water, 18th edition.
- Bishop, M. J, Powers, S.P, Porter,H. J and Peterson, C.H. 2006. Benthic biological effects of seasonal hypoxia in a eutrophic estuary. Predate rapid coastal development, Estuarine Coastal Shelf Science, 70: 415-422.
- Dodds, W.K, Jones, J. R and Welch, E. B. 1998. Suggested classification of stream trophic state: Distributions of temperate stream types by chlorophyll, Tot N and P. Water Research, 32 ; 1455-1462.
- Ghosh P. B, Sarkar, S and Saha, T. 2010. Some useful methods for the assessment of Environmental health of an aquatic system. In : Environmental Concern, D. Dasgupta (editor), Agrobios (India), pp315-323.
- Hutchison, G. E. 1973. Eutrophication. American Scientist, 61: 269-279.
- Karyds, M. 2009. Eutrophication assessment of coastal waters based on indicators: A literature review. Global Nest Journal, 11(40):373-390.
- Redfield, A.c. 1958. The biological control of chemical factors in the environment. American Scientist, 46:205-221.
- Rosemarin, A. 2004. Nitrogen: it pollutes when wasted, Phosphorous: scarce and dwindling. Down to Earth, 13(3): 27-34.
- Smith, V. H.2006. Responses of estuarine and coastal marine phytoplankton to nitrogen and phosphorous enrichment. Limnology and Oceanography, 51: 377-384.
- UNEP, 2006. The state of the marine environment: trends and processes, United Nations Environment Programme, The Hague, pp 28.
- Wetzel, R.G. 2001. Limnology, Lake and rivers ecosystem. Third edition, pp151-158.

Two-level Secure Re-routing (TSR) in Mobile Ad Hoc Networks

Himadri Nath Saha #¹, Dr. Debika Bhattacharyya #², Dr A.K. Bandhyopadhyay*³, Dr. P. K. Banerjee *⁴

Assistant Professor #¹, Professor #², Professor *³, Professor*⁴

Department of Computer Science and Engineering, Institute of Engineering and Management, West Bengal, India #¹, #²

Department of Electronics and Communication Engineering, Jadavpur university, West Bengal, India *⁴

Department of Electronics and Telecommunication, Jadavpur University, West Bengal, India.*³

him_shree_2004@yahoo.com #¹

Abstract— The lack of static infrastructure, open nature and node mobility causes several issues in Mobile Ad Hoc Network (MANET), such as energy utilization, node authentication and secure routing. In this paper we propose a new scheme, Two-level Secure Re-routing (TSR), an attack resilient architecture for Mobile Ad hoc networks. It is significantly different from existing solutions, as it does not focus on any specific attack, but instead, taking a general approach it achieves resilience against a wide range of routing disruption DoS attacks. TSR is a double-layer scheme that detects attacks at the transport layer but responds to them at the network layer. The concept of local supervision using watch-nodes has been incorporated to make it a secure architecture. Also, because routing disruption attacks have a pronounced effect on the size of the TCP congestion window, monitoring the window size is effective in detecting such events. Using both these methods, TSR is able to detect attacks. Once an attack is detected by the local supervision, it is verified by congestion window abnormalities and only then TSR initiates a re-routing process to find a new route. Our simulation results will show how attacks can be detected and data packets are re-routed after attacks have been detected.

Keywords: MANET, grayhole, blackhole, wormhole, jellyfish, DoS attack, rushing attack, watch node, node isolation, alternate route finder.

I. INTRODUCTION

In a Mobile Ad hoc networks (MANET), wireless devices communicate by forwarding packets on behalf of other devices—there is no central base station or fixed infrastructure to handle data routing. MANETs are particularly useful when a fixed infrastructure (e.g., a base station or access point) is impractical due to space or time constraints or when an existing infrastructure is not suitable for the required task. For mission-critical and other information-sensitive applications, the dependability and security aspects of the MANET, including reliability and availability, are of great importance. Denial-of-Service (DoS) attacks are a major threat to MANET security and quite a few of them have been discovered and discussed in the literature. Among them, routing disruption attacks are particularly menacing since they attempt to cause legitimate data packets to be routed in a dysfunctional way [5]. Routing disruption DoS attacks can be divided into three categories based on their different levels of sophistication:

outsider attacks, insider attacks, and protocol-compliant attacks.

In an *outsider attack*, the attackers are assumed to have no knowledge of the keys that are used to encrypt and authenticate the data and routing control packets. Preventing outside attackers from tampering with the data is accomplished by simply employing encryption and authentication schemes [5], [11].

In an *insider attack*, an attacker has compromised or captured a node, thus gaining access to encryption and authentication keys. The primary method of detecting and mitigating insider attacks is to monitor the packet forwarding behavior among the nodes [2], [3], [10], [12]. Also, there are approaches that focus on thwarting specific forms of insider attacks [6], [7].

Protocol-compliant DoS attacks [1] are the most difficult to defend against. In [1], Aad et al. refer to such attacks as “JellyFish” (JF) attacks. While the two types of attacks discussed above disobey protocol rules, JF attacks conform to all routing and forwarding rules. They are also passive, and therefore difficult to detect. A typical target of JF attacks is closed-loop flows that respond to packet delay and loss, such as TCP. Protecting MANETs against JF attacks is a formidable task that has yet to be addressed. It is worth noting that the wormhole attack can be launched even without having access to any cryptographic keys or compromising any legitimate node in the network.

We propose a routing architecture for MANETs, called Two-level Secure Re-routing (TSR), which is resilient against a wide range of attacks, including protocol-compliant attacks. TSR is a double-layer approach that monitors both using watch nodes and through variations in the size of the TCP congestion window to detect abnormalities, and reacts to those abnormalities at the network layer by initiating a re-routing process. The TSR architecture is compatible with on-demand source routing protocols such as Dynamic Source Routing (DSR) [8]. TSR is composed of four modules: nodes required for local supervision, watch nodes (LS) module, the node isolation algorithm (NIA), the congestion window surveillance (CWS) module and the alternate route finder (ARF) module. LS provides secure two-hop neighbor discovery and local monitoring of control traffic to detect malicious nodes. It provides a countermeasure technique that isolates the malicious nodes from the network thereby removing their ability to cause future damage. Only when LS discovers any

malicious node in the route does the CWS come into action. This is because congestion window size may be affected by network congestion, which is not a case of malicious activity. Therefore, CWS is responsible for confirming the presence of any abnormalities that might occur on a route. If any abnormalities are detected, CWS invokes ARF to build a new route.

The remainder of this paper is organized as follows. Section II introduces TSR and its modules. We present the simulation results in Section III. Finally, we conclude the paper in Section IV.

II. THE TWO-LEVEL SECURE RE-ROUTING ARCHITECTURE

A. Assumptions and Overview of TSR

We assume all links in the network to be bidirectional. We only consider network-layer route disruption attacks and disregard attacks to the physical or link layer of a wireless network.

In a communication session, we assume that both the source node and the destination node are trustworthy but intermediate nodes are not. It is assumed that all control packets used in TSR are authenticated via certain security mechanism (e.g., [5], [11]). Under this assumption, TSR is inherently resistant to outsider attacks. We focus our discussions on insider attacks and protocol-compliant attacks. A route with one or more malicious nodes is considered an “infected” route. In this paper, we focus on TCP as it is the most widely-used transport-layer protocol and it is the attack point of JF attacks. TSR uses a CWS module to observe the variations in the size of the TCP congestion window. If the variation indicates an abnormality, an alarm is raised to activate the ARF module. The ARF module in turn finds a new route. TSR strengthens the two vulnerabilities of DSR. First, TSR fortifies the “passive” re-routing approach of DSR by supporting both passive re-routing and a form of “active” re-routing. In DSR, the source node passively waits for a ROUTE ERROR packet to trigger a re-routing process. TSR enables the CWS module to actively initiate a re-routing process when network abnormalities are detected. Therefore, even if a malicious node on an infected route drops ROUTE ERROR packets, the source node is able to initiate re-routing. Second, TSR facilitates the process of identifying a valid route when a new route has to be found. The ARF module of TSR disables duplicate suppression when abnormalities are detected and considers route history in the re-routing process.

B. Local Supervision (LS)

Building neighbor lists. This protocol is used to build the data structure of the first-hop neighbors of each node and the neighbors of each neighbor. The data structure is used in local monitoring to detect malicious nodes and in local response to isolate these nodes. A neighbor of a node, X, is any node that lies within the transmission range of X. As soon as a node, say A, is deployed in the field, it does a one-hop broadcast of a HELLO message. Any node, say B, that hears the message, sends back a reply to A. Node A accepts all the replies that arrive within a timeout. For each reply, A adds the responder to its neighbor list R_A . Then, A does a one-hop broadcast of a message containing the list R_A . When B hears the broadcast, it stores R_A . Hence, at the end of this neighbor discovery process, each node has a list of its direct neighbors and the neighbors of each of its direct neighbors. This process is

performed only once in the lifetime of a node and is assumed to be secure. Henceforth, a node will not accept a packet from a node that is not a neighbor, nor forward to a node that is not a neighbor. Also, second-hop neighbor information is used to determine if a forwarded packet comes from a neighbor of the forwarder. If a node C receives a packet forwarded by B purporting to come from A in the previous hop, C discards the packet if A is not a second-hop neighbor. Finally, A activates local monitoring immediately after building its first and second-hop neighbor lists.

Local supervision: Local supervision starts immediately after the completion of neighbor discovery. It uses a collaborative detection strategy, where a node monitors the traffic going in and out of its neighbors. For a node, say a, to be able to monitor a node say, b, a must be a neighbor of both b and the previous hop from b, say d. If this is satisfied, we call a the watch node of b over the link from d to b. This implies that a is the watch node for its entire outgoing links. For example, in Fig. 1, nodes M, N, and X are the watch nodes of A over the link from X to A. Information for each packet sent from X to A is saved in a watch buffer at each guard. The information includes the packet identification and type, the packet source, the packet destination, the packet’s immediate sender (X), and the packet’s immediate receiver (A). The guards expect that A will forward the packet toward the ultimate destination, unless A is itself the destination. Each entry in the watch buffer is time stamped with a time threshold, s, by which A must forward the packet. Each packet forwarded by A with X as a previous hop is checked for the corresponding information in the watch buffer. A malicious counter ($MalNo(i,j)$) is maintained at each watch node, i, for a node, j, at the receiving end of each link that i is monitoring over a sliding window of length $Twin$ that slides by d units. $MalNo(i,j)$ is incremented for any malicious activity of j that is detected by i. The increment to $MalNo$ depends on the nature of the malicious activity detected, e.g., V_f for fabricating and V_d for dropping a control packet, being higher for more severe infractions. To account for intermittent natural failures that can occur at legitimate nodes, a node is determined to be misbehaving, only if the $MalNo$ goes above a threshold (C_i) over T_{win} time units. Of course, due to the network topology there may not be any watch node for a given link. In that case, malicious behavior cannot be detected. Now we give the isolation and the response algorithm that applies across all the attack modes.

C. Node Isolation Algorithm(NIA)

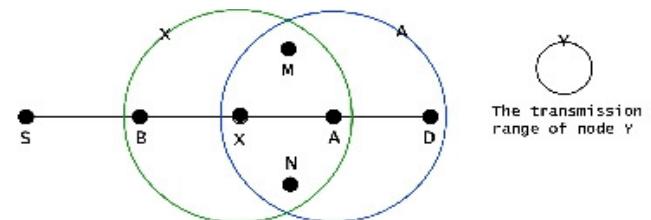


Fig.1. X, M, and N are guards of node A over the link from X to A.

- When $MalNo(a,A)$ crosses C_t , a revokes A from its neighbor list, and sends to each neighbor of A, say D, an authenticated alert message indicating A is a suspected malicious node. This communication is authenticated using the shared key between a and D to prevent false accusations. Alternately, if the clocks of all the nodes in the network are

loosely synchronized, a can do authenticated local two-hop broadcast as in [13] to inform the neighbors of A.

2. When D gets the alert, it verifies the authenticity of the alert message, that a is a first-hop neighbor of node A, and that A is D's neighbor. It then stores the identity of a in an alert buffer associated with A.

3. When D gets enough alert messages,(c) about A it then calls the Congestion Window Surveillance(CWS) module to further verify if A is a malicious node or not. We call c the detection confidence index of D. The detection confidence represents the minimum number of guard nodes that must report that a certain node, j, is malicious for a neighbor, i, of that node to take up further monitoring of it, if i does not directly detect j. Note that the number of guards that report malicious activity is cumulative over time. A single node, due to the authentication mechanism, cannot generate more than one acceptable alert. Framing is the process by which an innocent node is proved to be malicious by a quorum of malicious nodes. A small value of C increases the chance of successful framing of good nodes, while a large value of c increases the rate of harm a malicious node causes in the network before being locally detected. If we set c to be infinity it means that a node only trusts itself in revoking a suspicious node and thus the local framing probability goes to zero. False alarm, distinct from framing, is caused by a (legitimate) node mistaking another (legitimate) node to be malicious because of imperfections in the wireless channel, e.g., node i does not observe node j dutifully forwarding a packet.

4. If there are a large number of malicious nodes (greater than the detection confidence index of D) in the neighborhood of D, then they may cumulatively try to frame an innocent node by sending alert messages one after another. So, after this step we have kept another checking through CWS which will confirm the presence of any malicious activity in the network.

D. Congestion Window Surveillance (CWS)

Before giving the rationale behind the CWS module, we introduce the following theorem.

Theorem I: Suppose that M denotes the mean value of the measured congestion window size in number of TCP segments, assuming that the TCP segment size is fixed. In addition, suppose that the congestion window size M at time t is represented by M_t , and the window size in the next RTT is given by M_{t+1} . Then the probability of M_{t+1} being less than or equal to M_t satisfies the following relation:

$$P[M_{t+1} \leq M_t] < 8/(3M^2 - 2M) \quad .(1)$$

The proof of the theorem is given in [4]. In [14], Fu et al. conducted comprehensive experiments and simulations in order to observe TCP performance in various MANET environments differing in topology and traffic load. It was found that TCP maintains a relatively large average congestion window size. Specifically, the average measured window size M was found to between 12 and 26. Similar but more extensive simulations were conducted and it was observed in [4] that $M \geq 4$ for every simulation run. Applying Theorem I with $M \geq 4$, we get $P[M_{t+1} \leq M_t] < 1/5$. This probability is relatively small. Therefore, if we assume the events in which $M_{t+1} \leq M_t$ are independent, then it would be unlikely that multiple events will take place in a short time duration. Suppose there are consecutive K RTTs being monitored, and

let p denote the number of RTTs with $M_{t+i} \leq M_t$. Given a constant $p_0 \in \{1, ..., K\}$, the probability $P[p \geq p_0]$ is lower-bounded as follows:

$$P[p \geq p_0] < \sum_{i=p_0}^K \binom{K}{i} \left(\frac{1}{5}\right)^i \left(\frac{4}{5}\right)^{K-i} \quad .(2)$$

The summation of the above inequality is a very small value when p_0 is close to K. For example, when $p_0 = 8$ and $K = 10$, $P[p \geq p_0] < 7.8 \times 10^{-5}$. However, the value of $P[p \geq p_0]$ can be relatively large when a network is under a routing disruption attack. For instance, in a Blackhole attack, an adversary node drops every data packet it receives.

This prevents the TCP window size from increasing, and thus results in $P[p \geq p_0] = 1$. Even in low-frequency attacks, such as Grayhole or Jellyfish attacks, the value of $P[p \geq p_0]$ may increase noticeably. Hence, it seems possible to use the value of $P[p \geq p_0]$ to detect the existence of attacks. Therefore, the event $p \leq p_0$, which we refer to as the *stagnant window event* hereafter, serves as an effective indicator of attack existence.

E. Alternate Route Finder (ARF)

ARF enhances the re-routing functions of DSR in three aspects. First, when CWS confirms any abnormalities in a route, ARF tags the current route as an infected route. If there is no non-tagged route available in the source node's cache, ARF initiates an active re-routing process rather than just responding passively to ROUTE ERROR packets. Second, during the active re-routing process, the source node collects more routing information by disabling duplicate suppression. Third, with the route information in its cache, the source node selects a new route using the following ARF algorithm.

The ARF algorithm first marks off the infected node(s), then tries to find an alternative route between the sender and the receiver. If there are multiple routes between the sender and the receiver then the algorithm will choose the one with the least number of nodes i.e. the route with the minimum hop count. But if no other route exists between the sender and the receiver apart from the one containing the malicious node then the algorithm reports so. The pseudo code implementation of the algorithm is given below.

```
//Input: [s1, r1] s1 is sender, r1 is receiver
//Output: is the result of running the ARF algorithm
void pathfinder(s1,r1)
{
    path = new int[n1+1]; //stores the path from
    npath = new int[n1+1]; sender to receiver
    rpath = new int[n1+1];
    prev = new int[n1+1]; //stores the previous hop
    visited = new int[n1+1]; //corresponding node
    value is set to 1
    prev[0] = visited[0] = 0; //if visited once
    enqueue(s1); //add s1 to the queue
    visited[s1] = 1;
    do
    {
        p = dequeue();
        if (p == r1)
            break;
        if (p != 0)
        {
            for(i=1;i<=n1;i++)
            {
                if (nodes[p][i] == 1 && visited[i] != 1)
                {
                    visited[i] = 1;
                    npath[i] = p;
                    if (i == n1-1)
                        rpath[i] = p;
                    else
                        prev[i+1] = p;
                }
            }
        }
    }
}
```

```

prev[i] = p;
enqueue(i);
}
}
} while (p != 0);
if (p != r1)
System.exit (0); //No alternate route found, so
exit the system
int v = prev[r1];
rpath[0] = v;
int i = 0;
while (v != s1)
{
rpath[++i] = prev[v];
v = prev[v];
}
for (i=0;rpath[i]!=0;i++)
p = k = i-1;
for (i=0;i<=p;i++,k--)
path[i] = rpath[k];
path[i] = r1;
for (i=0;path[i]!=0;i++)
print(path[i]); //print the path array which contains
//the nodes which are included in the path
}

```

III. SIMULATION RESULTS

The platform to perform the simulation is designed in Java. In the simulation we have shown the detection and mitigation of four attacks, namely Blackhole Attack, Jellyfish Attack, Framing Attack or Blackmail Attack and Grayhole Attack.

In the simulation we have assumed the initial congestion window size to be 5 and the time required to transmit a packet between two different neighbors of the same node to be identical.

A. Initial Path Establishment

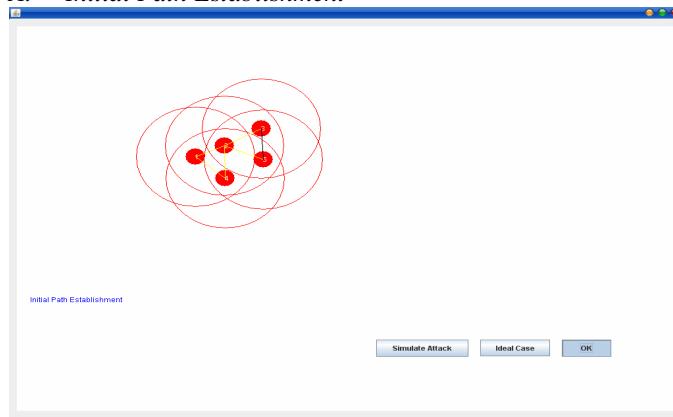


Fig. 2. Initial Paths are set between two neighbor nodes.

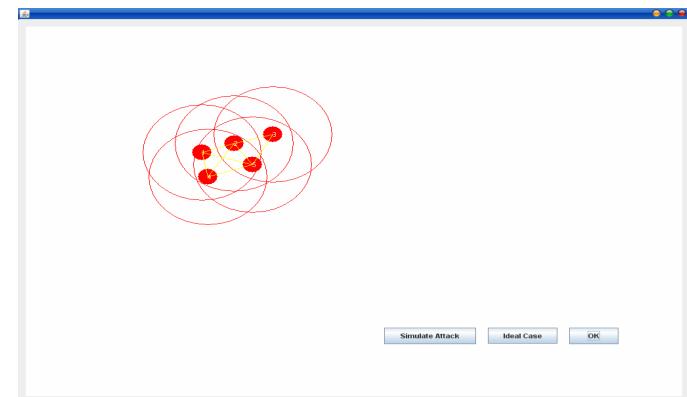


Fig. 3 The yellow line shows the paths between two nodes.

B. Ideal Case

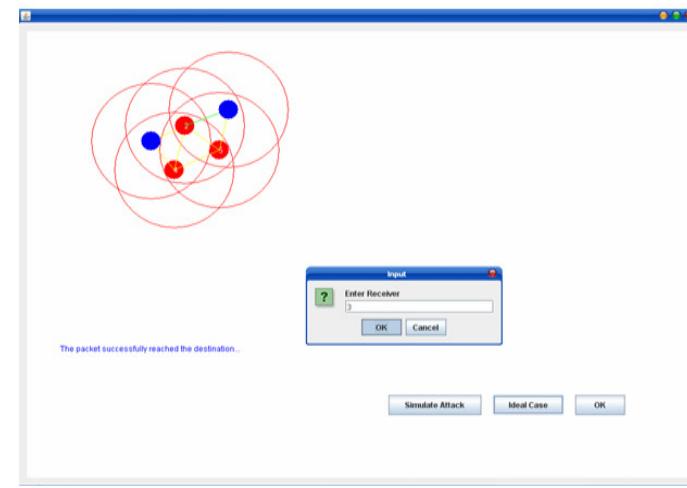


Fig. 4, In the ideal case, there is no malicious node in the path so packet reaches destination successfully. Here, if there are several routes between source and destination, the route with smallest hop count is selected. In this example, the source is node1 and destination is node3 and the smallest route is 1-2-3 with hope count 3.

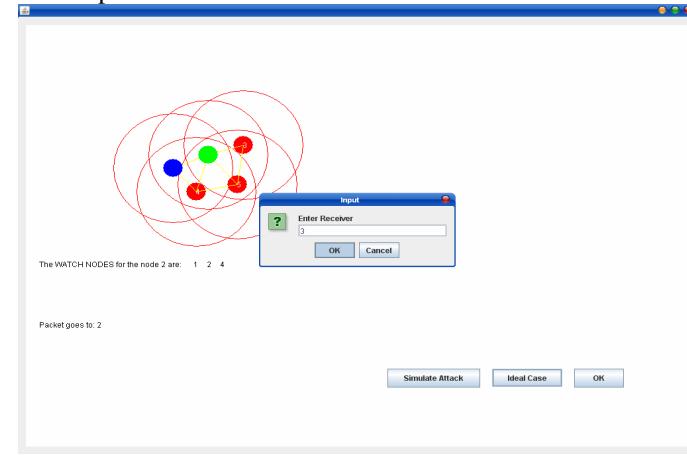


Fig. 5, After packet is sent to node2 from node1, the common neighbor nodes of node1 and node2 (nodes 1, 2, 4) act as watch-nodes for node2 to watch whether it forwards the packet within the timestamp.

C. Black-hole Attack

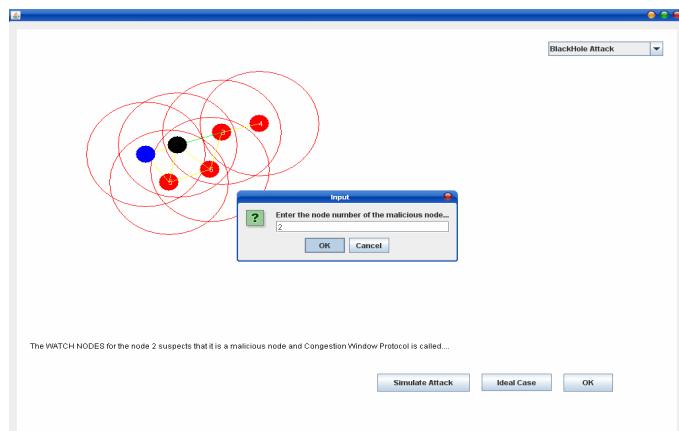


Fig. 6, Here, the source is node1 and destination is node4. The shortest path is 1-2-3-4. Now node2 is a malicious one which incorporates Black-hole Attack. Packet is sent from node1 to node2. The watch nodes for node2 suspect it as a malicious one, so congestion window protocol is called.

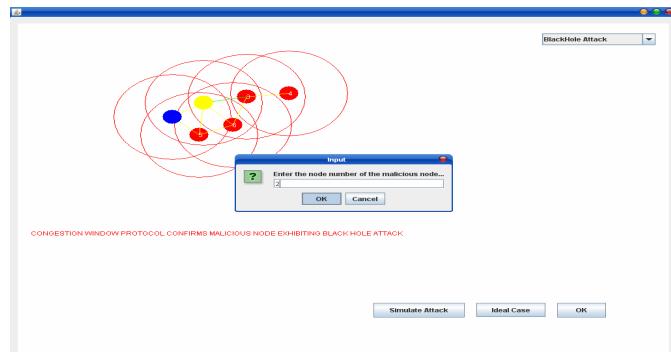


Fig. 7, Congestion window protocol confirms node2 to be malicious.

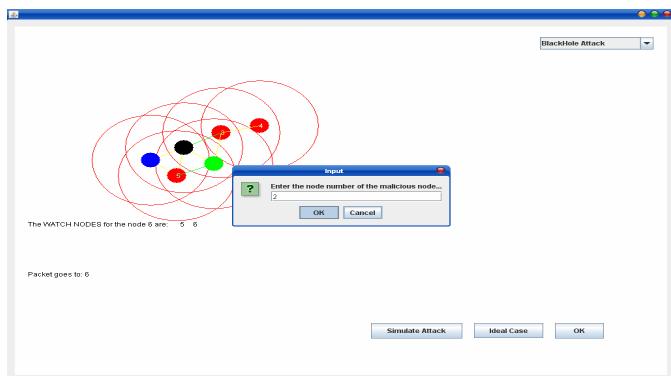


Fig. 8, Alternate Route Finder Algorithm is called. The algorithm returns 1-5-6-3-4 as the alternate shortest route.

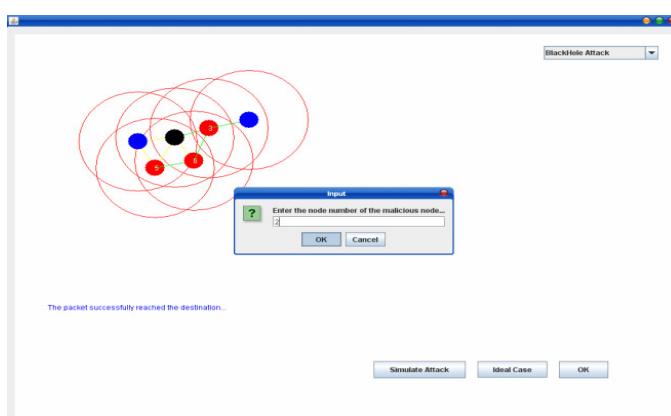


Fig. 9, Packet sent successfully.

D. Jellyfish Attack

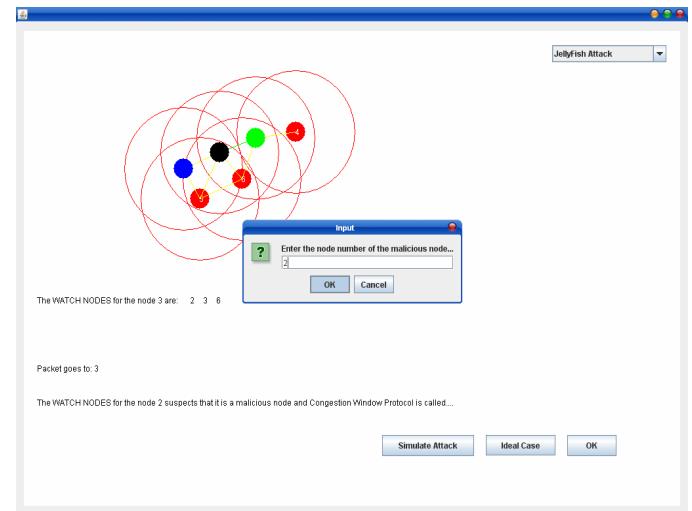


Fig. 10, Here, node2 incorporates Jellyfish Attack so delay is added at node2. Though the packet is sent to node3, watch nodes of node2 suspect it to be a malicious one and Congestion Window Surveillance protocol is called.

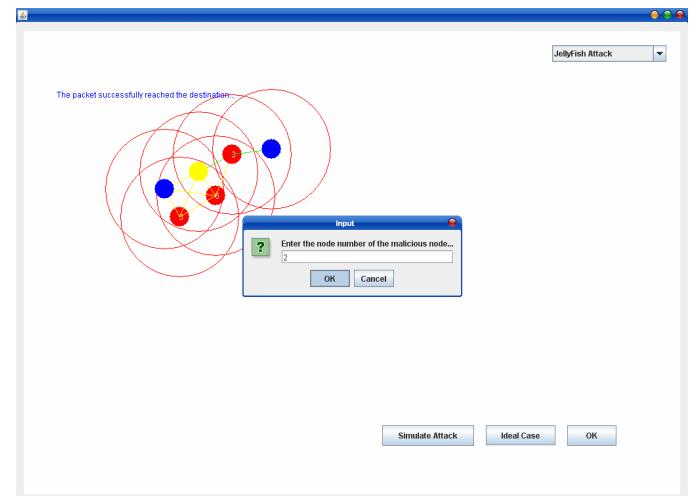


Fig. 11, Congestion window Surveillance protocol confirms node2 as a malicious one.

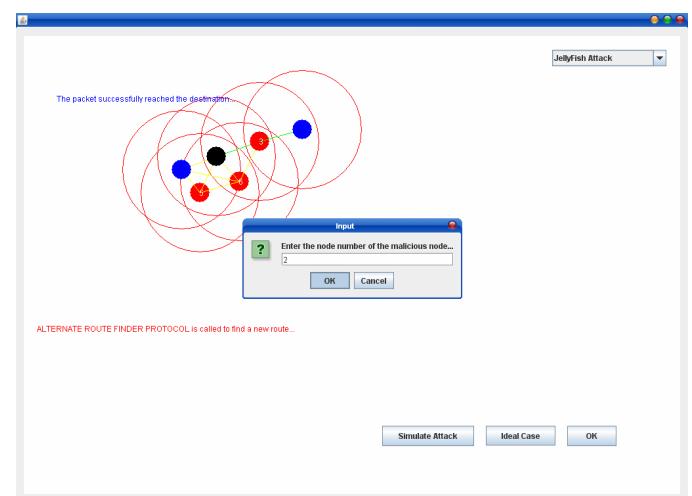


Fig. 12, Alternate Route Finder Algorithm is called.

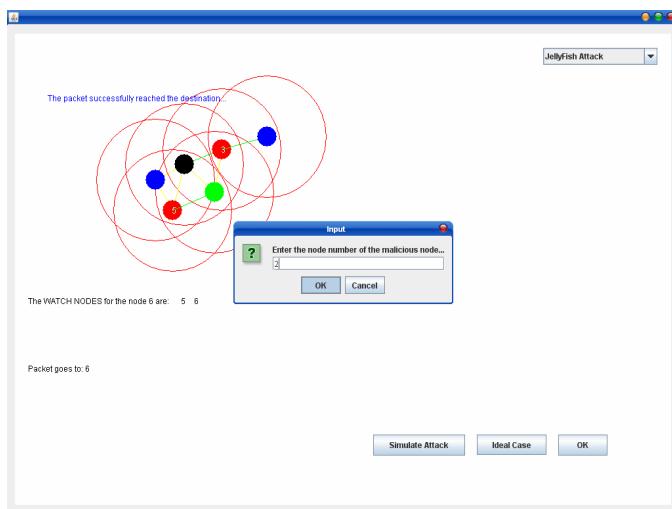


Fig. 13, The alternate shortest route is 1-5-6-3-4.

E. Framing Attack

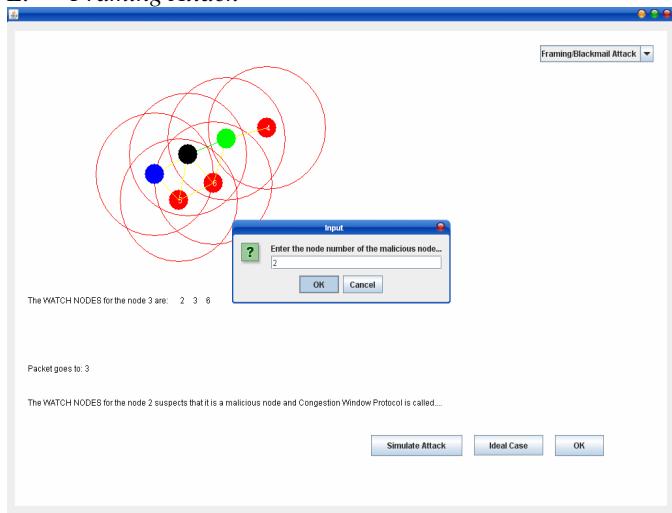


Fig. 14, Here, Framing Attack is simulated. The watch nodes for node2 suspect it as a malicious one and Congestion Window protocol is called.

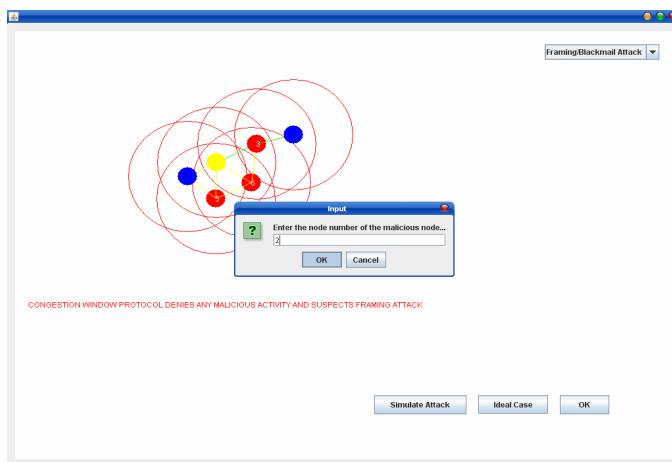


Fig. 15, As the neighbors of node2 framed the attack against node2, so congestion window detects it and announces node2 as an innocent one because packet reaches node4 successfully.

IV. CONCLUSION

This paper presented a novel routing architecture for MANETs called TSR that is attack resilient. TSR employs a Two-level approach in that it uses LS and CWS modules to detect network abnormalities (either attack or dysfunctional events) at the transport layer and responds to them by using the ARF module to execute re-routing at the network layer. Our analysis shows that TSR is resilient against a variety of insider attacks as well as protocol-compliant attacks. As part of our future work, we will explore the possibility of adapting the principles of TSR to routing protocols other than DSR. Alternatively, we can show the application of local monitoring for mitigating the wormhole attack, this approach is general and can be extended to detect other control attacks like the Sybil and the sinkhole attacks by changing the kind of information that is maintained in the watch buffers and the checks run on them.

REFERENCES

- [1] Aad, J. Hubaux, and E. W. Knightly, "Denial of service resilience in ad hoc networks," Proc. MobiCom, Sep. 2004, pp. 202-215.
- [2] B. Awerbuch, D. Holzer, C. Nita-Rotaru, and H. Rubens, "An on-demand secure routing protocol resilient to Byzantine failures," Proc. WiSe, Sep. 2002, pp. 21-30.
- [3] S. Buchegger and J.-Y. Le Boudec, "Nodes bearing grudges: towards routing security, fairness, and robustness in mobile ad hoc networks," Proc. 10th Euromicro Workshop on Parallel, Distributed and Networkbased Processing, Jan. 2002, pp. 403-410.
- [4] Ruiliang Chen, Jung-Min Park, and Michael Snow, "CARE: Enhancing Denial-of-Service Resilience in Mobile Ad Hoc Networks".
- [5] Y.-C. Hu, A. Perrig, and D. B. Johnson, "ANIAdne: A Secure on-demand routing protocol for ad hoc networks", Proc. MobiCom, Sep. 2002, pp. 12-23.
- [6] Y.-C. Hu, A. Perrig, and D.B. Johnson, "Packet leashes: a defense against wormhole attacks in wireless networks," Proc. INFOCOM, Mar. 2003, pp.1976-1986.
- [7] Y.-C. Hu, A. Perrig, and D. B. Johnson, "Rushing attacks and defense in wireless ad hoc network routing protocols," Proc. WiSe, Sep. 2003, pp. 30-40.
- [8] D. B. Johnson, D. A. Maltz, and Y.-C. Hu, The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR) (Internet-Draft), Mobile Ad-hoc Network (MANET) Working Group, IETF, July 2004.
- [9] J. Kong, X. Hong, J. Park, Y. Yi, M. Gerla, L'Hospital: Self-healing Secure Routing for Mobile Ad-hoc Networks, Technical Report TR-040055, Computer Science Department, University of California, Los Angeles, Jan. 2005.
- [10] S. Marti, T. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," Proc. MobiCom, Aug. 2000, pp. 255-265.
- [11] P. Papadimitratos and Z. Haas, "Secure routing for mobile ad hoc networks," Proc. CNDS, Jan. 2002.
- [12] W. Yu, Y. Sun, and K. J. R. Liu, "HADOF: Defense against routing disruptions in mobile ad hoc networks," Proc. INFOCOM, Mar. 2005, pp. 1252-1261.
- [13] D. Liu, P. Ning, Efficient distribution of key chain commitments for broadcast authentication in distributed sensor networks, in: Proceedings of the 10th Annual Network and Distributed System Security Symposium (NDSS), February 2003, pp. 263-276.
- [14] Z. Fu, P. Zerfos, H. Luo, S. Lu, L. Zhang, and M. Gerla, "The impact of multihop wireless channel on TCP throughput and loss," Proc. INFOCOM, Mar. 2003, pp. 1744-1753.

PATTERN RECOGNITION IN ELECTRICAL & OPTIMAL DOMAIN

Date : 17th January,2012

Chairman : Prof. Somnath Sarkar

Co-Chairman : Prof. P.K.Sinha Roy

Rapporter : Prof. Parna Guha Bhattacharya

| Paper ID | Paper Title | Authors |
|----------|---|---|
| 23 | Bidirectional Pattern Matching for DNA Sequences | Tamal Chkrabarti and Devadatta Sinha |
| 28 | PAPR Reduction of an Optically Converted Pre-coded OFDM signal by SOA | Abhirup Das Barman and Ipsita Sengupta |
| 38 | An Efficient Method For Error Correction in Digital Signal Processing Perspective | Dr.Asok Kumar, Rupam Das and Dr. Anup Kumar Bhattacharjee, Mrinmoy Sarkar |
| 48 | Reduction of Power consumption with the Incrementation of LED | Sourav Sarkar and Amitava Sengupta |
| 50 | Effect of Conduction Band Nonparabolicity on Transmission Coefficient of Double Quantum Well Tripple Barrier Structure with Parrabolic Geometry | Souvik Mukherjee, Rajit Karmakar, Arindam Biswas and Arpan Deyasi |
| 51 | An Efficient NLP Based Approach To Personalized Information Retrieval In Web Mining | Subhabrata Sengupta, Sourav Saha, Sukalyan Goswami |
| 55 | Automatic Schema Matching and Mapping | Soumi Dutta, Deepsikha Mullick and Moumita Roy |
| 87 | Key Frame Detection Using Two-Step Corner Selection | Monika Agarwal, Satyabrata Maity, Amlan Chakrabarti |
| 90 | Fiber Optics Laser Systems in Cardio Vascular Diseases | Debadyoti Ghosh, Md. Niyaz Mallick, Farha Serin Naaz, Alok Kumar Das. |
| 99 | Multi-Sensor Data Fusion Using Particle Filter to Estimate Nonlinear Recursive Bayesian Model for Object Tracking | Aninda Mondal,Moloy Kumar Chowdhury |
| 102 | Harmonics and Interharmonics Estimation Of A Passive Magnetic Fault Current Limiter Using Biorthogonal Wavelet Transform | Raju Patwary, A.B.Chowdhury, D.Roy, S.Rudra |

HARMONICS AND INTERHARMONICS ESTIMATION OF A PASSIVE MAGNETIC FAULT CURRENT LIMITER USING BIORTHOGONAL WAVELET TRANSFORM

Raju Patwary¹, A.B. Choudhury², D.Roy³, S.Rudra⁴

Abstract- This paper presents the harmonics and inter-harmonics analysis of a passive magnetic fault current limiter (PMFCL). This device limits the current at post fault without affecting the pre fault state of the system. The harmonics and inter-harmonics estimation of a non-stationary signal generated by PMFCL were investigated using Biorthogonal Wavelet transform and FFT. Continuous wavelet transform (CWT) and discrete wavelet packet transform(DWPT) has been applied for estimation of harmonics and interharmonics in PMFCL under normal and faulted condition.

Key Words: PMFCL, Harmonics, Interharmonics, Continuous Wavelet Transform.

I. INTRODUCTION

With increase in electricity demand the power system fault level increases. Traditional techniques of fixing fault current are reported in [1]. However, all these techniques are suffering with certain disadvantages. If Relay coil, circuit breaker has some disadvantages itself and fault has occurred in that system, then it cannot minimize the catastrophic effect of the fault, also it cannot operate further until the faults are cleared from the system. Magnetic fault current limiter offers a promising solution to this burning problem of traditional protection equipments; it limits fault current without affecting the normal state. This helps us to reduce the fault level of a system without upgrading the existing switchgear. It does not require any sensing and control devices, as well as it can operate without any additional power supply or any auxiliary winding like other fault current limiters. From the earlier research initiatives, it has been cleared that the other fault current limiters require all the auxiliary accessories mentioned above. Based on this we can say that it is more advantageous to use PMFCL rather than any other fault current limiter. In this paper, the harmonics and inter-harmonics estimation of a non-stationary signal in presence of PMFCL are analyzed. This investigation is required as many solid-state fault current limiters [5] inject undesirable harmonics in the system, thus affecting the quality of power. The investigation of the harmonics and inter-harmonics content of a non-stationary signal in presence of PMFCL under normal and faulty condition ware estimated based on Biorthogonal wavelet transform. This wavelets are especially well suited for the analysis of transient time-varying signals than any other wavelet like mayer, morlet, maxicanhat etc. We have studied all the existing wavelet techniques but this Biorthogonal wavelet transform offers results that are more satisfactory.

II. MAGNETIC FAULT CURRENT LIMITER

A PMFCL consists of a permanent magnet (NdFeB) and two low saturation flux density materials such as a ferrite core as shown in Fig-1. [2] The coils are connected such that at any instant of the A.C. cycle, the M.M.F generated by the A.C. source and the permanent magnet assist each other in one core and opposes each other in the another and both cores remain saturated below a specific value of applied current.

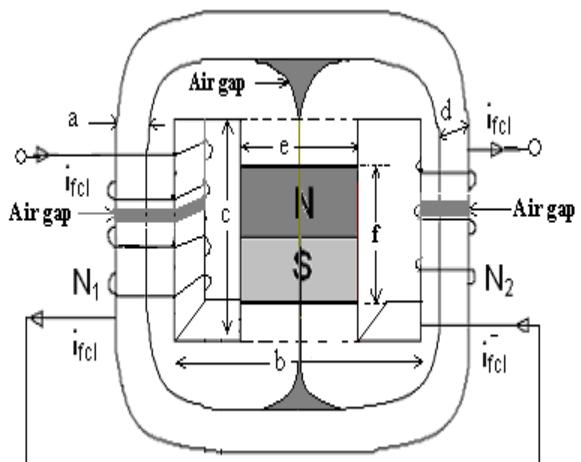


Fig.1 Parallel Biased PMFCL

The permanent magnet will bias the core and forces it to operate in saturation. [3, 4] Magnetic current limiter behaves as a variable inductor. Under normal operating condition below the rated current (maximum nominal current), PMFCL offers low saturated inductance L_s . Under fault condition ($\text{Max nominal current} < I \leq \text{Max fault current}$), the inductance offered by the PMFCL is

the high unsaturated inductance L_u . The design parameters are such that $L_u >> L_s$. Hence under faulty condition the effective inductive reactance of FCL is very high, thus limiting the current.

III. RESPONSE OF PMFCL

A prototype laboratory model of PMFCL was designed with the data shown in Table-1. Software simulations were carried out to obtain the graphs shown in Figs. 3.

Table-1: Design Data

| | |
|------------------|---|
| Core | $a=7.5\text{mm}; b=24\text{ mm}; d = 20\text{mm}; c=48\text{mm}, \mu_{rs}$ (saturated permeability) $= 87.6 * \mu_0$; μ_{ru} (unsaturated permeability) $= 12740 * \mu_0$; B_k (Knee flux density) $= 0.35\text{ T}$; Ferrite core |
| Permanent Magnet | $e = 15\text{ mm}; f = 5\text{mm}, \mu_{ru} = 2.4 \times 10^6\text{ H/m, } H_c$ (coactivity) $= 0.8 * 10^6\text{ A/m}$, NdFeB (Neodymium Iron Boron) |
| Others | No. of turns= 100 per limb; $V_s = 15\text{V r.m.s.}$; $R_{LN} = 11\Omega$ (Normal load resistance) $R_{LF}=1.7\Omega$ (fault resistance) |

Fig.2 shows the simulation circuit of PMFCL. Under normal operating condition when switch 'S' is open the load across

PMFCL is R_{LN} . In fault condition, the switch 'S' is closed and the load across PMFCL is R_{LN} parallel with R_{LF} . Fault is created at time $t=0.02\text{ sec.}$ by closing the switch 'S'. From Fig. 3 it is evident that the fault current is significantly limited by the use of PMFCL[4].

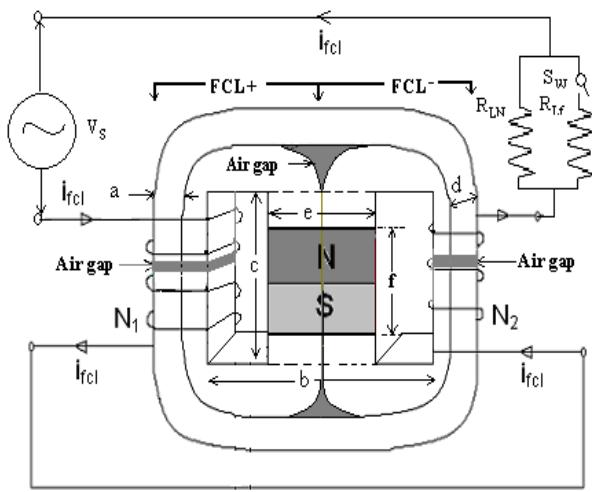


Fig.2 Simulation Circuit of PMFCL

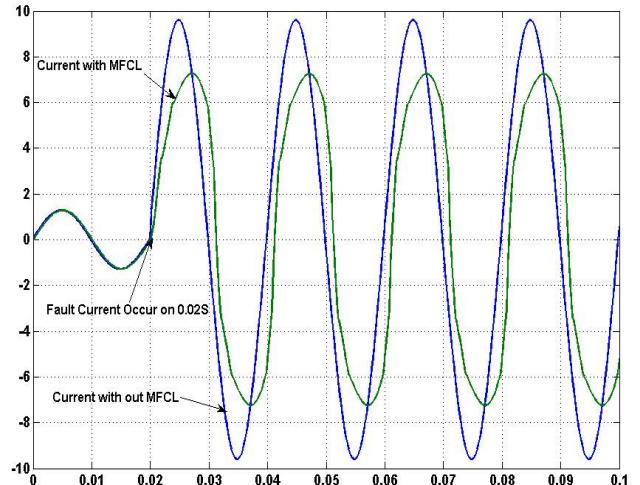


Fig.3 PMFCL Current Vs Time

Thus the PMFCL limits the fault current significantly and voltage regulation at normal state is also good. It is to be investigated whether the presence of PMFCL in a network injects any undesirable harmonics or interharmonics in the supply system to degrade the quality of power.

IV. HARMONICS AND INTERHARMONICS ANALYSIS

Traditionally the Fast Fourier Transform (FFT) were been used to analyze harmonics in power system. The FFT is able to give accurate results only if the analyzed waveform fulfills certain conditions [8]. FFT analysis cannot give any inter-harmonics estimation, because due to spectral leakage FFT produces enormous errors. Wavelet transforms were introduced for harmonic measurements of power systems. The wavelet transform is capable of accurately extracting the information contained in any distorted waveform with high resolution in both time and frequency. For harmonic measurement purposes, the frequency spectrum can be divided into uniform bands. This can only be realized using the Wavelet Packet Transform (WPT).[6] In this paper the harmonic estimation of PMFCL current signal is presented based on discrete wavelet packet transform (DWPT) and continuous wavelet transform (CWT). To achieve good time and frequency resolution of a no stationary signal wavelet was developed. The continuous wavelet transform (CWT) of a signal $f(t)$ is given by

$$W(a,b) = \int_t f(t) \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) dt \quad (1)$$

Where wavelet were denoted by

$$\psi_{a,b}(t) = \int \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) dt \quad \text{----- (2)}$$

Here ‘ a ’ is dilation or scaling parameter and ‘ b ’ is translation or location parameter.

By choosing, $a = a_0^m$, $b = nb_0$ and $t=kT$ in equation (1), where $T=1.0$ and m, n, k . are integer values. The discrete wavelet transform is given by

$$DWT(m, n) = \left(\sum_k x(k) \psi[(k - nb_0)/a_0^m] \right) / a_0^m \quad \text{---- (3)}$$

For computational efficiency a_0 and b_0 are set to 2 and 1 respectively. Multilevel filter bank is implemented on equation (3) to get discrete wavelet packet transform

$$\psi_n(t) = 2 \sum_{k=0}^{2N-1} L(k) \psi_n(2t-k) \quad \text{----- (4)}$$

$$\psi_{2n+1}(t) = 2 \sum_{k=0}^{2N-1} H(k) \psi_n(2t-k) \quad \text{----- (5)}$$

The wavelet packets ψ_n were generated from the linear combination of scaled and translated version of mother wavelet $\psi_1(t)$ and its scaling function $\psi_0(t)$. L and H is low pass and high pass filter of length $2N$ corresponds to mother wavelet.

In this paper, we have used Biorthogonal (bior3.3) wavelet, which is compactly supported orthonormal wavelets, making discrete wavelet analysis practicable. These wavelets have no explicit expression except for bior3.3, which is the Biorthogonal wavelet. However, the square modulus of the transfer function of h is explicit and faulty sample. If the polynomial P(y) can be expressed as

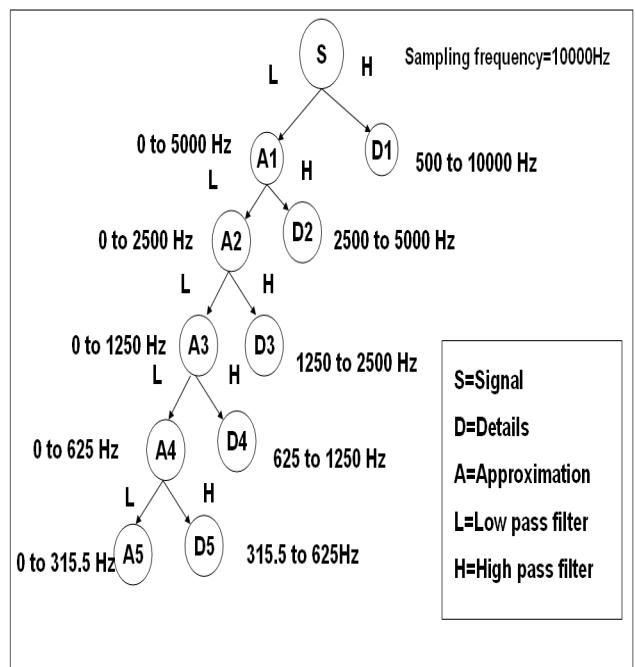
$$P(y) = \sum_{k=0}^{N-1} C_k y^{N-1+k}$$

$$|m_0(w)|^2 = \left(\cos^2\left(\frac{w}{2}\right) \right)^N P\left(\sin^2\left(\frac{w}{2}\right)\right)$$

Then

$$m_0(w) = \frac{1}{\sqrt{2}} \sum_{k=0}^{2N-1} h_k e^{-ikw}$$

Here, N is the number of vanishing moments.



Dig: 4: DWPT Filter bank

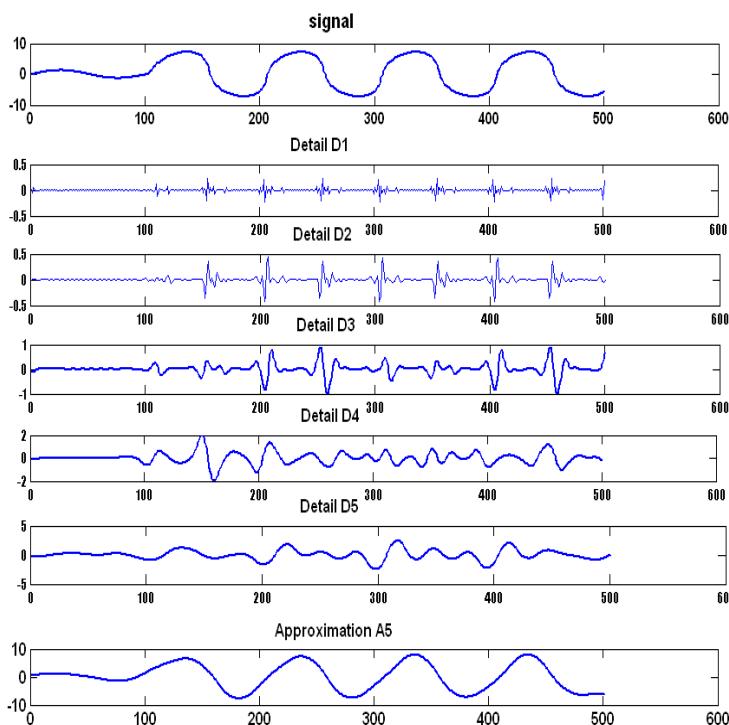


Fig. 5 The first 6 decomposed components of PMFCL current

Discrete wavelet packet in this paper shown in Fig 4. The discrete wavelet packet transform gives a uniform band frequency at each level, which is very suitable for harmonics and inter-harmonics studies. After applying the DWPT at level 5 on PMFCL current signal, we get the desired details and approximation as shown in Fig. 5. It is evident that at normal condition (0-0.02 s or 0-100

Hz) no significant harmonic were been generated. Harmonics generated mainly in fault condition (0.02-0.1s or 0to500hz). It is observed that at fault condition the magnitude of higher harmonics is very less and the harmonics and inter-harmonics of high magnitude occur mainly in frequency band of 0-315 Hz (refer table-2).

Table-2: Sub-band frequency and magnitude

| Sub-band | Frequency range (Hz) | Magnitude Range (A) | |
|----------|----------------------|---------------------|--------------|
| | | Normal State | Faulty state |
| D1 | 5000-10000 | 0.00 | 0.25 |
| D2 | 2500-5000 | 0.00 | 0.49 |
| D3 | 1250-2500 | 0.00 | 0.9 |
| D4 | 625-1250 | 0.02 | 2.0 |
| D5 | 315.5-625 | 0.14 | 3.0 |
| A5 | 0-315.5 | 1.8 | 9 |

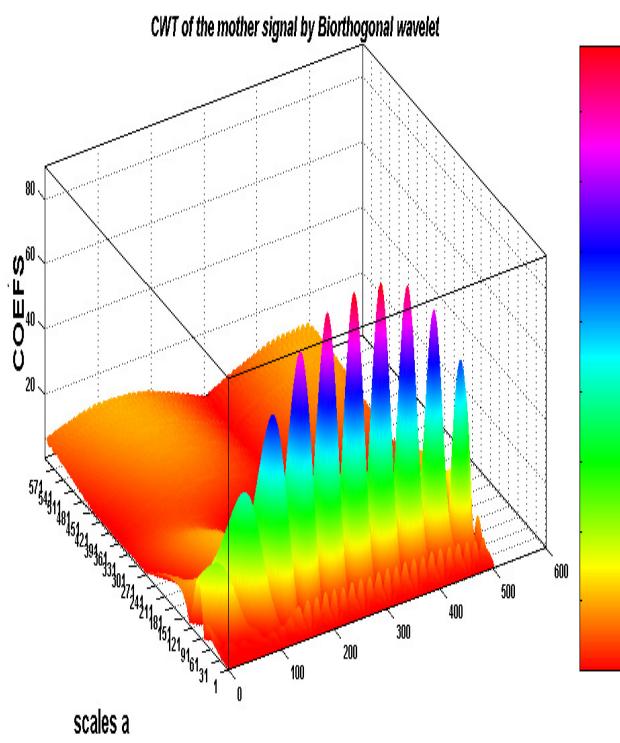


Fig.6 Time-frequency-amplitude plot by CWT

Fig.-6 shows the continuous wavelet transform (CWT) of the current signal of PMFCL as shown in Fig- 3 with Biorthogonal wavelet. Here we observe the time-freq-magnitude plot of the PMFCL current signal. From 0-0.02 is the normal state and 0.02-0.1 is the faulty state.

We have observed that the magnitude of fundamental component (50 Hz) is high and the higher order component has very negligible magnitude. There is some amount of sub harmonics components whose magnitude is also very less. Thus, CWT supports the conclusion drawn from DWPT. Because from the CWT images we can identify the location of the faults, but the approximate magnitude of the fault can't be evaluated from CWT result.

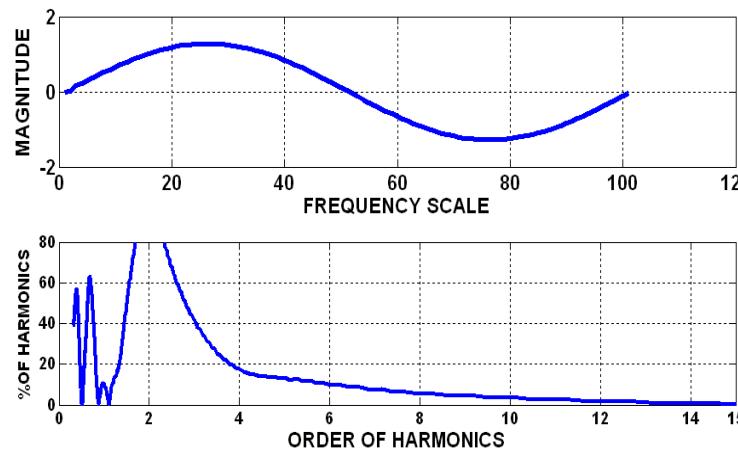


Fig.7 Frequency spectrum of PMFCL current at normal state by Wavelet

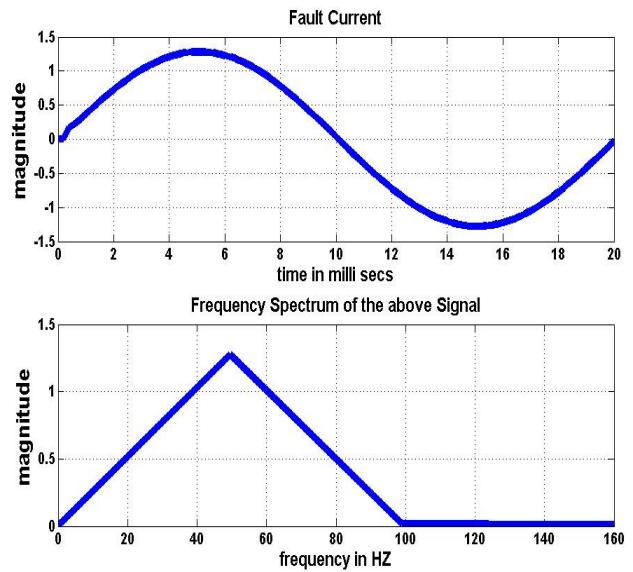


Fig.8 Frequency spectrum of PMFCL current at normal state by FFT

To know the exact harmonic and inter harmonic components and their magnitude in the non-stationary signal generated by PMFCL wavelet transform and Fast Fourier Transform were been carried out. For frequency spectrum, analysis 600 numbers of samples were been

taken for normal condition and faulted condition as shown in Fig-7 & 9.

Fig 7 & 8 shows the frequency spectrum for PMFCL current at normal state by WT and FFT respectively. Here we observe that the 1st harmonic component mainly constitutes the PMFCL current signal at normal condition. There is almost no significant higher order harmonics in the PMFCL current at normal state.

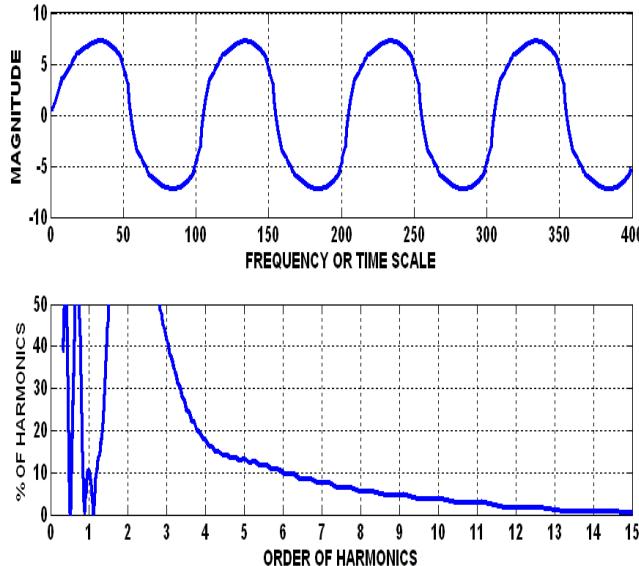


Fig.9 Frequency spectrum of PMFCL current at faulty state by Wavelet

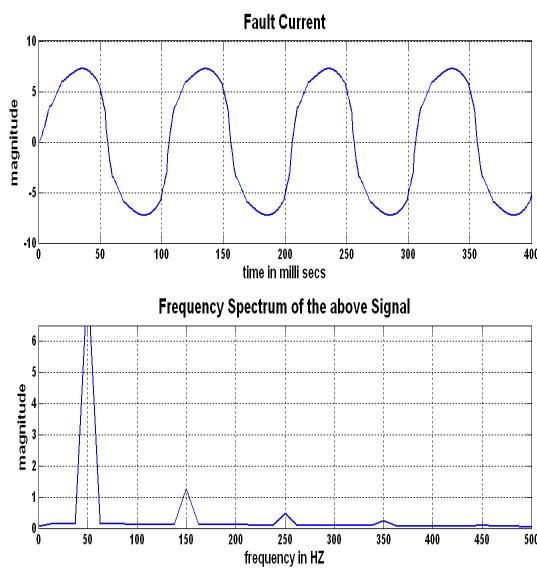


Fig.10 Frequency spectrum of PMFCL current at faulty state by FFT

Fig 9& 10 shows the frequency spectrum of PMFCL current signal at faulty state by WT and FFT respectively.

Table-3: Harmonics and Interharmonics in PMFCL current at fault:-

| Frequency (Hz) | Order | Magnitude (%) |
|----------------|-------|---------------|
| 25 | 0.5 | 10 |
| 150 | 3.0 | 39 |
| 175 | 3.5 | 26 |
| 200 | 4.0 | 18 |
| 250 | 5.0 | 12 |
| 275 | 5.5 | 10 |
| 350 | 7.0 | 8 |

From Table-3 it is evident that at faulty state some amount of higher order harmonics and inter-harmonics (mainly 3rd, 5th and 7th) induced in the PMFCL current signal. But their magnitude is within tolerable limit. The PMFCL current composed with mainly 1st harmonic component.

V. CONCLUSIONS

This paper has investigated the problems associated with the harmonics generation in a fault current limiter. A prototype laboratory model and simulation have been developed to observe the performance of PMFCL. We observe that almost all power electronics and solid-state devices inject harmonics to supply system, which in turn degrades the quality of power. On the contrary, the passive magnetic fault current limiter (PMFCL) did not generate any significant harmonics in normal state (pre-fault). Under fault condition (post-fault), some amount of harmonics was generated but they were not rich in magnitude. In passive magnetic fault current limiter, harmonics were generated due to irregular flux distribution in core in faulty state. Again irregular flux distribution occurs due to transition of PMFCL transfer characteristic from saturated region to unsaturated region. So some amount of harmonics will be there in the network due to presence of PMFCL at fault condition. However, the harmonic contents can be minimized by adopting a good and optimum design procedure of the PMFCL assembly. With the proposed design of permanent magnet fault current limiter, it can be observed that the magnitude of fault current is limited within a safe level.

VI. REFERENCES

- [1] S. C. Mukhopadhyay, F. P. Dawson, M. Iwahara, and S. Yamada, "Analysis, design and experimental results for a passive current

2nd Annual International Conference IEMCON 2012

limiting device, *IEE Proc.—Electr. Power Appl.*, vol. 146, no. 3, pp. 309–316, May 1999.

[2] M. Iwahara, S. C. Mukhopadhyay, S. Yamada, and F. P. Dawson, “A full passive fault current limiter using the permanent magnet for the burden critical equipments,” in *J. Nonlinear Electromagnetic Systems*: IOS Press, 1998, vol. 13, pp. 349–352.

[3] M. Iwahara and E. Miyazawa, “A numerical method for calculation of electromagnetic circuits using the tableau approach,” *IEEE TransMagn.* Vol. 19, pp. 2457–2459, Nov. 1983.

[4] Iwahara M., Yamada S., Dawson F. P., Fillion G., “A Passive Current Limiter for Power Semiconductor Protection,” Industrial Applications Society Conference Proceeding, San Diego, 1996, Vol. 111, pp. 1298-1 301

[5] M. M. R. Ahmed, Ghani A. Putrus, Li Ran and Lejun Xiao, “Harmonic Analysis and Improvement of a New Solid-State Fault Current Limiter” IEEE transaction on industry applications, vol. 40, No.4, July/August 2004

[6] Kit Po Wong, Van Long Pham “ Analysing Power system waveforms using Wavelet-Transform approach” Proceedings of the 5th International Conference on Advances in Power System Control, Operation and Management, APSCOM 2000, Hong Kong, October 2000.

[7] A.B.Choudhury ,D.Roy, and M.Iwahara, ”Field Distribution and Performance Analysis of a Passive Magnetic Fault-current Limiter under Transient Conditions” Electric Power Components and System, 37:pp-1195-1207,2009,

[8] A.B.Choudhury ,D.Roy, M.Iwahara and S. Yamada ”Harmonics and Interharmonics Estimation of a Passive Magnetic Fault Current Limiter Using Wavelet Transform” Jurnal of the Japan Society of Applied Electromagnetics and Mechanics, vol.17.No.4(2009),pp-606-610

-
1. Raju Patwary, Assistant Professor, Department of Electrical Engineering, Calcutta Institute of Engineering and Management, Tollygunge, Kolkata, West Bengal, India. Email: rajpat_2008@yahoo.com
 2. A.B. Choudhury, senior lecturer, Department of Electrical Engineering, Bengal Engineering and Science University, Shibpur, Howrah, West Bengal, India. Email: ab_choudhury@yahoo.com
 3. Dr. D. Roy, assistant professor, Department of Electrical Engineering, Bengal Engineering and Science University, Shibpur, Howrah, West Bengal, India. Email: dbr_roy@yahoo.co.in
 4. S.Rudra, Inspire Fellow in Jadavpur University, Kolkata 70032, West Bengal, India. Email: Shubhobrata.rudra@gmail.com.

Bidirectional Pattern Matching for DNA Sequences

Tamal Chakrabarti

Department of Computer Science & Engineering, Institute of Engineering & Management,

Salt Lake Electronics Complex, Kolkata-700 091, INDIA

tamalc@gmail.com

Devadatta Sinha

Department of Computer Science & Engineering, Calcutta University,

92, Acharya Prafulla Chandra Road, Kolkata – 700009, INDIA

devadatta.sinha@gmail.com

Abstract—Biology in the 21st century is being transformed from a purely laboratory-based science to a computational and information science. Finding all occurrences of a DNA pattern in a DNA Sequence is a problem that arises frequently in Bio-informatics. As the DNA sequences are typically very long (The human genome alone has approximately 3 billion DNA base pairs), matching patterns can be an expensive operation. The need for fast and accurate DNA pattern matching algorithms has increased with the advent of modern computers. This paper introduces a bidirectional pattern matcher algorithm for a fast and efficient solution to the DNA Pattern Matching problem.

Keywords—

DNA, Gene, Bio-informatics, Pattern Matching, Algorithms

I. INTRODUCTION

The task of finding the first, or all of the occurrences of pattern in a text —*string searching*^[8] — arises in many computing applications. A string searching algorithm aligns the pattern with the beginning of the text and keeps shifting the pattern forward until a match or the end of the text is reached^[7]. Complex methods of string matching^[3] today are used in locating DNA sequences (genetics), fingerprint assessment (criminology), soil patterns (geology), retinal blood vessel assessment (medicine), design of assembly lines to improved flow (business), continuous speech profiles and many other fields.

Strings arise very naturally in biology^[10]: an organism's genome – its full set of genetic material - is divided up into giant linear DNA molecules known as chromosomes, each of which serves conceptually as a one-dimensional chemical storage device. Indeed, it does not obscure reality very much to think of it as an enormous linear tape, containing a string over the alphabet {A, C, G, T}. The string of symbols encodes the instructions for building protein molecules; using a chemical mechanism for reading portions of the chromosome, a cell can construct proteins that in turn control its metabolism.

DNA is an acronym for the molecule deoxyribonucleic acid. Deoxyribonucleic acid (DNA) is called a double helix because it is a double-stranded molecule. DNA contains three components: deoxyribose (a five-carbon sugar), a series of phosphate groups, and four nitrogenous bases, (nitrogen compounds that contain bases). The four bases in DNA are adenine (A), thymine (T), guanine (G), and cytosine (C). The deoxyribose combines with phosphates to form the backbone of DNA. Adenine (A) forms a base pair with thymine (T), and guanine (G) with cytosine (C) in DNA.

A DNA sequence is a representation of the genetic code contained within an organism^[6]. Molecular biology researchers have great need to compare portions of DNA sequences. DNA pattern matching is an identification of a pattern in one or more sections of genetic code. Biologists use the pattern matching algorithms to discover evolutionary divergence, the origins of disease, and ways to apply genetic codes from one organism into another. A DNA sequence is our representation of a string of nucleotides contained in a strand of DNA. For example: ATGCGATACAAGTTGTGA represents a string of the nucleotides A, G, C, and T.

Traditional algorithms for matching DNA patterns use well-known approaches, such as the Knuth-Morris-Pratt algorithm^[4] or the Boyer-Moore algorithm^[2]. But since the DNA sequences are typically very long, pattern matching for DNA becomes a costly affair. The process of matching of DNA patterns can be improved by introducing a method of parallelism, based on a bidirectional approach, which eventually decreases the run-time.

II. DNA PATTERN MATCHING

DNA is described as the “blueprint of life, the molecular thread in the nucleus of each of our cells which guides the assembly of molecules and complete living organisms”^[5]. As mentioned earlier, DNA is represented as four letters or base pairs, A, C, G and T. These stand for Adenine, Cytosine, Guanine and Thymine respectively. A DNA sequence is a one-dimensional sequence consisting of these four letters. A

sequence can be made up of several thousand letters. Figure 1 provides examples of two sequences.

| |
|------------|
| GAGATCGCG |
| GAGAGAGAGC |

Figure 1: Examples of two DNA poly-nucleotides

A DNA pattern is also made up of the same letters as the sequences. However they are considerably shorter in length, ranging from four to just over one hundred letters in length. An example of a short DNA pattern is shown in figure below.

| |
|-------|
| GATTA |
|-------|

Figure 2: A DNA Pattern

DNA sequences are searched for specific DNA patterns and this needs to be done efficiently, which would entail string matching [1]. We formalize the string-matching problem (i.e. finding a pattern in a sequence) as follows.

Formal Statement of the problem:

Let us assume that the sequence is an array $T[1 \dots n]$ of length n and that the pattern is an array $P[1 \dots m]$ of length m , where $n > m$. We say that pattern P occurs with shift s in sequence T if $0 \leq s \leq n - m$ and $T[s + 1 \dots s + m] = P[1 \dots m]$.

The string matching algorithms are useful in detecting irregularities in a gene. A gene is normally a stretch of DNA that holds the information to build and maintain an organism's cells and pass genetic traits to offspring. The occurrence of unwanted or mutated sequences in the gene causes disease [9]. This irrelevant DNA takes the form of X, which can be any of the four letters that make up DNA. Figure 3 shows one of the sequences from Figure 1, with irrelevant Xs.

| |
|-----------------|
| GAxGATxxCGxxxCG |
|-----------------|

Figure 3: A DNA sequence containing irrelevant data

A sequence could contain no Xs, a single X or a group of consecutive Xs. It is up to the researcher to determine whether a sequence that contains a pattern that is split by an X is valid or not.

III. RELATED WORK

In 1977, Knuth, Morris and Pratt published an efficient algorithm [4] to solve the string matching problem. The main idea of the KMP algorithm is to pre-process the pattern string P so as to compute a failure function f that indicates the proper shift of P so that, to the largest extent possible, we can reuse previously performed comparisons. The KMP pattern matching algorithm incrementally processes the text string T comparing it to the pattern string P . Each time there is a match, we increment the current indices. On the other hand, if there is a mismatch and we have previously made progress in

P , then we consult the failure function to determine the new index in P where we need to continue checking P against T . Otherwise (there was a mismatch and we are at the beginning of P), we simply increment the index for T (and keep the index variable for P at its beginning). We repeat this process until we find a match of P in T or the index for T reaches n , the length of T (indicating that we did not find the pattern P in T). This algorithm runs in $O(m + n)$ time, where m is the size of the pattern and n is the size of the sequence.

IV. PROPOSED APPROACH

Let A be an alphabet and $x = x_0 \dots x_{k-1}$, be a string of length k over A , and let ϵ be an empty string. We will first look at a few basic definitions.

A prefix of x is a substring u with

$$u = x_0 \dots x_{b-1} \text{ where } b \in \{0, \dots, k\}$$

i.e. x starts with u .

A suffix of x is a substring u with

$$u = x_{k-b} \dots x_{k-1} \text{ where } b \in \{0, \dots, k\}$$

i.e. x ends with u .

A prefix u of x or a suffix u of x is called a proper prefix or suffix, respectively, if $u \neq x$, i.e. if its length b is less than k .

A border of x is a substring r with

$$r = x_0 \dots x_{b-1} \text{ and } r = x_{k-b} \dots x_{k-1}$$

$$\text{Where } b \in \{0, \dots, k-1\}$$

A border of x is a substring that is both proper prefix and proper suffix of x . We call its length b the width of the border.

For example, Let $x = abacab$. The proper prefixes of x are

$$\epsilon, a, ab, aba, abac, abaca$$

The proper suffixes of x are

$$\epsilon, b, ab, cab, acab, bacab$$

The borders of x are

$$\epsilon, ab$$

The border ϵ has width 0, the border ab has width 2.

The empty string ϵ is always a border of x , for all $x \in A^*$, where A^* denotes all strings over A .

The empty string ϵ itself has no border.

The idea behind the algorithm is to find out all positions in the sequence, where a match with the pattern ends. Since we know the length of the pattern, we can find out the starting position of every match. The approach is feasible when we apply the concept of automaton. The automaton used in the algorithm is an array of pointers (which represents the internal rules) and a separate external pointer to some index of that

array (which represents the current state). The position of the external pointer changes according to the current state and the set of rules, when the next character from the sequence is presented to the automaton. Eventually when the final state is reached, a match is found.

Let us consider the pattern AGAGAC, and then all the prefixes are given below:

0. ϵ
1. A
2. AG
3. AGA
4. AGAG
5. AGAGA
6. AGAGAC

Let us now list down the longest proper suffix of each such prefix, which is also a prefix to it, i.e. the longest border of each prefix.

0. ϵ
1. ϵ
2. ϵ
3. A
4. AG
5. AGA
6. ϵ

Let us assume that, at some point we come across a partial match up to AGAGA, which means we also have a partial match of AGA and A. Depending on the next character of the sequence, one of the following three cases may occur:

- i. Say, the next character is C, and then in this case we conclude that a full match is found.
- ii. Say, the next character is G, and then the partial match can't be extended. So we try with the largest border of this match, which is AGA. Now G can be fitted and we move ahead with the new found partial match AGAG.
- iii. Say, the next character is T, and then there is no way to extend the current partial match. So we skip it and try to match the next character of the sequence with the first character of the pattern. But now the current partial match is ϵ .

In order to build the automaton, we build an array F[]. Each entry of F identifies the length of the longest border of the pattern. In other words F[i] is the length of the length of the next best partial match for the string of length i. For the pattern AGAGAC the array F will have the following values:

$$\begin{aligned} F[0] &= 0 \\ F[1] &= 0 \\ F[2] &= 0 \\ F[3] &= 1 \\ F[4] &= 2 \\ F[5] &= 3 \\ F[6] &= 0 \end{aligned}$$

The array F can be pre-computed by comparing the pattern against itself. This information can then be used to construct the DNA pattern matcher.

V. BIDIRECTIONAL PATTERN MATCHING

We can further improve the performance of the pattern matching process, by adapting a bi-directional approach. In this approach, we divide the sequence into two parts by cutting it along the middle. Then we search for the pattern from left to right in the first part and search for the reverse pattern from right to left in the second part.



Figure 4: Bidirectional Pattern Matcher

To identify the existence of the pattern across the cut-point both left-to-right and right-to-left pattern matchers count the number of characters matched at the termination point. Let l be the length of the partial match by the left-to-right process and r is the length of the partial match by the right-to-left process. Then if $(l + r) \geq m$, then we conclude that there is a matching pattern across the cut point.

For example let us consider the pattern AGAGAC. Then the left-to-right pattern matcher searches for AGAGAC in the first half of the sequence, and the right-to-left pattern matcher searches for CAGAGA (which is the reverse of the pattern) in the second half of the sequence. Let us assume that the first process found a partial match of size 4, and the second process found a partial match of size 2 in either side of the cut point, as depicted in the figure below.



Figure 5: Finding a match across the cut point

In this case since $(4 + 2) = 6 =$ the size of the pattern (6), we conclude that there is a match across the cut point.

As another example, let us consider the pattern AGAGAG. Then the left-to-right pattern matcher searches for AGAGAG in the first half of the sequence, and the right-to-left pattern matcher searches for GAGAGA (which is the reverse of the pattern) in the second half of the sequence. Let us assume that the first process found a partial match of size 4, and the second process found a partial match of size 4 in either side of the cut point, as depicted in the figure below.



Figure 6: Multiple matches across the cut point

In this case, we find multiple matches across the cut point, one by considering 4 characters to the left of the cut point and 2 characters to the right of the cut point. And the other by

considering 2 characters to the left of the cut point and 4 characters to the right of the cut point.

From the above discussion, we observe that the following algorithm can be used to compute all the pattern matches across the cut point.

Let us assume:

- l = length of partial match to the left of the cut point
- r = length of partial match to the right of the cut point
- p = index of the cut point
- m = length of the pattern

```
Algorithm MatchAcrossCutPoint(l, r, m, p)
Begin
    if ((l + r) ≥ m)
        Begin
            // A pattern match exists across the cut point
            while (l > 0 AND r ≥ (m - l))
                Begin
                    Report a match with shift  $p - l + 1$ 
                     $l = F[l]$ 
                End (while)
            End (if)
End
```

This bi-directional algorithm can be implemented by running the left-to-right and right-to-left pattern matchers in two separate threads in a dual processor system. Since the two pattern matchers work independent of each-other, we can achieve parallelism, thereby increasing the efficiency of the pattern matching process. Two threads are made to report the size of partial matches on either side of the cut point, which can be then used to compute matches across the cut point.

VI. RESULTS

We implemented the bi-directional algorithm in a dual processor system using two threads for left-to-right and right-to-left pattern matching. We observed the time taken by the algorithm for different lengths of the pattern and the sequence. The results were then compared with the time taken by the Knuth-Morris-Pratt algorithm for equal sequence and pattern lengths. The tests were conducted under the same environment, with all the other factors remaining same. The following diagram shows the result of our tests for sequence sizes in the range 50000 to 130000, and a fixed pattern size of 100 nucleotides.

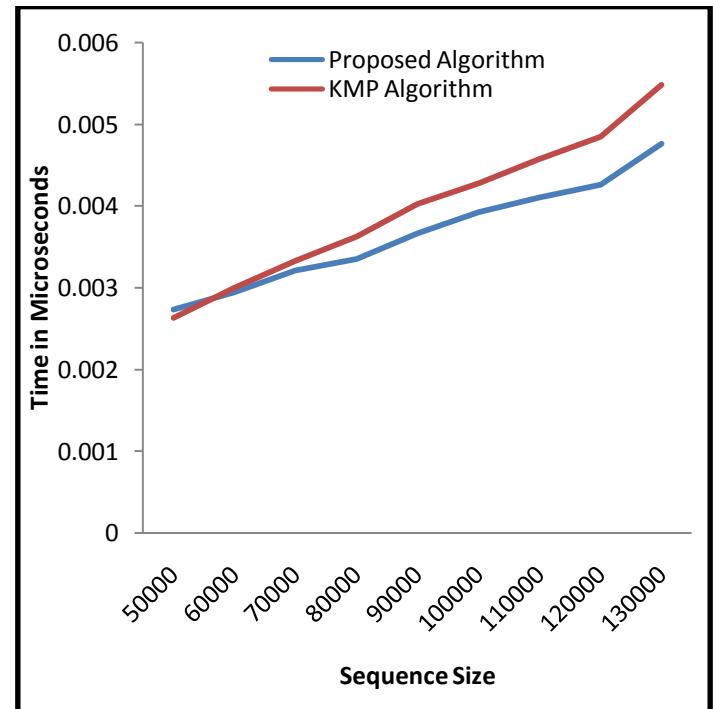


Figure 7: Comparison between proposed algorithm and KMP algorithm, for a pattern size of 100

As can be seen, the proposed approach works much faster than the KMP algorithm when the sequence size is greater than 50000.

VII. CONCLUSIONS

In this paper, we introduced a bidirectional approach, to reduce the time of the DNA pattern matching process. We conducted a comparative analysis of the run time of the traditional KMP Algorithm vs. the proposed approach. The results show that the proposed approach works well, when the sequence sizes are appreciable, i.e. greater than 50000. For smaller sequences, conventional approach, like the KMP algorithm, can give similar time, but larger the size of the sequence, better is time benefit of using the bi-directional technique.

At the end it may be mentioned that the subject opens up a wide scope of investigative study with a view to explore better improvement, if any. The authors suggest the following areas of further research:

- Experimenting with real data
- Comparing our approach, with other approaches, to exploit their complementary strengths
- Exploiting higher levels of parallelism
- Making a more intelligent choice of the cut-points of the DNA sequences
- Use a divide and conquer strategy to further sub-divide the sequence in smaller sub-sequences

REFERENCES

- [1] Baeza-Yates. R. A. String searching algorithms revisited. *Lecture Notes in Computer Science*, 382:75–96, 1989
- [2] Boyer R. and Moore J.S. A fast string searching algorithm. *Comm. of the ACM*, 20:762–772, 1977.
- [3] Colussi. L. Fastest pattern matching in strings. *Journal of Algorithms*, 16:163–189, March 1994.
- [4] Knuth D., Morris J. and Pratt V. Fast Pattern matching in strings, SIAM Journal of computer science, pp323 – 350, 1977.
- [5] Lander E.S., Langridge R., and Saccoccia D.M.. Mapping and Interpreting Biological Information. *Communications of the ACM*, 34(11):33 – 39, November 1991.
- [6] Mount David W., Bioinformatics – Sequence and Genome Analysis, Cold Spring Harbor Laboratory Press, 2001.
- [7] Navarro G. and M. Raffinot. Fast and SimpleCharacter Classes and Bounded Gaps PatternMatching, With Application to Protein Searching.In Annual Conference on Research inComputational Molecular Biology, Montreal,Canada, 2001.
- [8] Navarro. G. A Guided Tour to Approximate String Matching. *ACM Computing Surveys*,33(1):31–88, March 2001.
- [9] Rajesh S., Prathima S., Reddy L.S.S., Unusual Pattern Detection in DNA Database Using KMP Algorithm, *International Journal of Computer Applications* (0975 - 8887) Volume 1 – No. 22, 2010
- [10] Smith-Kearny. P. Molecular Genetics. Macmillan Education Ltd, London, 1991.

PAPR reduction of an optically converted pre-coded OFDM signal by SOA

Abhirup Das Barman

University of Calcutta ,

Institute of Radio Physics and Electronics
92, Acharya Prafulla Chandra Road
Kolkata 700009, W.B. India
abhirup_rp@yahoo.com

Ipsita Sengupta

Govt. College of Engg. and Leather Technology
ECE Department
Block-LB, Sector-III, Salt Lake City
Kolkata, 700098, W.B.
ipsen1980@yahoo.co.in

Abstract— In this paper we have investigated optical conversion of an OFDM modulated signal by using a Semiconductor Optical Amplifier (SOA). Use of a pre-coding algorithm reduces peak to average power ratio (PAPR) of OFDM signal in electrical domain. The algorithm is able to provide a useful diversity technique on a single optical carrier. Detail investigations exposed the existence of an optimum operating condition of a SOA for which minimum distortion is obtained at the receiver. Use of this algorithm offers a reduction in PAPR of OFDM signal, which in turn lowers the requirement of CW optical input power to the SOA.

Index Terms— Semiconductor Optical Amplifier (SOA), PAPR, frequency diversity, signal line rate.

I. INTRODUCTION

Use of OFDM in optical domain has attracted research interest because of its strong incoherent tolerance to chromatic dispersion, low system cost and high signal transmission capacity [1]-[2]. In metropolitan/access network deployment of multimode fiber (MMF) is cost effective. However, MMF suffers from intermodal dispersion which causes pulse spreading at the fiber output. This is due to optical pulse arrival at the fiber output at different time interval. This inter symbol interference (ISI) in a fiber can be mitigated by adopting optical OFDM transmission technique as well as providing diversity on a single optical carrier [3]-[4]. Also OFDM has the ability to mitigate the effect of multipath propagation without complex equalization filters [5].

Use of semiconductor optical amplifier (SOA) in metro/access network is considered actively for its potentially low cost with integration capability into optical network [6]. In this work, optical conversion of OFDM signal by a SOA is investigated. A challenge for OFDM system is the accommodation of large dynamic range of the signal caused by PAPR due to the fact that, OFDM signal has a large variation between the average signal power and the maximum signal power. A large dynamic range is inherent to multicarrier modulation having essentially independent subcarriers [7]. As a result, the subcarriers can add constructively or destructively, which may contribute to large variation in signal power. This problem can not be mitigated by driving the SOA deep into the saturation. Hence, PAPR

reduction along with diversity gain algorithm using DSP technique is implemented in electrical OFDM (E-OFDM) signal domain. The resulting E-OFDM signal is then used to modulate the bias current of a SOA. Performance of the scheme is evaluated in terms of extinction ratio and signal line rate determined from the demodulated signal at the receiver. The behaviour of SOA is simulated by multistage Reservoir modeling of SOA [8]-[9]. Simulation results help in finding an optimum operating condition of the SOA for which minimum distortion of the received signal occurs at the receiver. The diversity gain can be achieved at the cost of decoding delay without sacrificing symbol rate. It has been observed that the introduction of the algorithm reduces PAPR by 2.1 dB. The performance of SOA is examined with the application of this OFDM with frequency diversity and it has been shown that due to the reduced PAPR obtained from this scheme, input CW optical power to SOA can be reduced by 0.7dB.

II.A. THEORY AND MODELING

The block diagram for generation of intensity modulated OFDM signal by SOA is shown in Fig.1. An electrically generated OFDM signal is superimposed over a dc signal. This signal is then used to directly drive the SOA. Optical gain of SOA is altered by the electrical driving current. A cw wave at the input to SOA provides an optical carrier which is then modulated according to the bias current modulation pattern. The dc bias and the amplitude of the driving current are adjusted to operate the SOA at the optimum operating conditions. At the receiver, the optically modulated signal is detected and the data is recovered following an inverse procedure of OFDM modulation. 16-QAM modulation is employed for signal mapping. The input to the IFFT is the complex vector $\mathbf{X} = [X_0 \ X_1 \ X_2 \dots X_{N-1}]^T$, the vector has length N, where N is the number of subcarriers. Each of the element of \mathbf{X} is a complex number representing a particular QAM constellation point. In order to generate real valued OFDM signal, \mathbf{X} is constrained to have Hermitian symmetry [5]:

$$X_0 = X_{N/2} = 0 \text{ and } X_{N-K} = X_K^*$$

The real valued OFDM signal in time domain at the output of the IFFT block is represented by $\mathbf{x} = [x_0 \ x_1 \ x_2 \dots x_{N-1}]^T$.

For analysis of the behaviour of SOA, we have used the powerful Reservoir modeling of SOA [8]-[9]. The single stage

reservoir dynamics with ASE depleting channel can be described by the following ordinary differential equation

$$\frac{dr(t)}{dt} = \frac{I}{q} - \frac{r(t)}{\tau} - S^{in}(G(r) - 1) - S_{ASE}^{in}(G_{ASE}(r) - 1) \quad (1)$$

and the reservoir $r(t)$ is defined as

$$r(t) = A \int_0^L N(z, t) dz, \quad (1a)$$

where A and L are the cross sectional area and length of the active region in the SOA, respectively. $N(z, t)$ is the carrier density along the longitudinal direction z and t denotes the time variable. $r(t)$ in Eq. (1a) denotes the total number of carriers in the active region (volume) of SOA. S^{in} and S_{ASE}^{in} in Eq. (1) are the input signal and the initial ASE photon flux respectively. $G(r)$ is the amplifier gain at the signal wavelength and $G_{ASE}(r)$ is the gain at the wavelength of the ASE depleting channel.

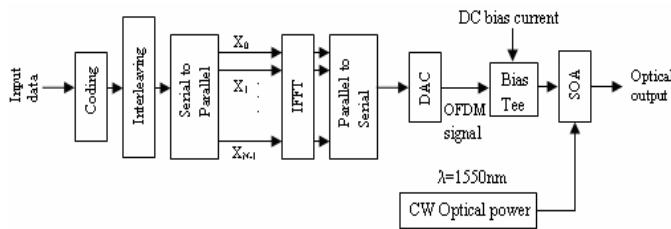


Fig.1: Block diagram for generation of OFDM intensity modulated signal by SOA.

The other terms in Eq. (1) are the bias current I , the electronic charge q , and the carrier lifetime τ .

The amplifier gain can be expressed as [1]:

$$G(r, \lambda_j) = \exp(a_j(\lambda_j)r(t) - r_{0j}(\lambda_j)), \quad (1b)$$

where a_j and r_{0j} are the two dimensionless parameters and explained in detail in [9]. Eq. (1) is solved in order to obtain $r(t)$. Since $r(t)$ is defined as the average carrier density over the length of SOA, the amplifier can be thought of as a point amplifier in the time domain. Hence after obtaining the amplifier gain from (1b) the output power can be found out.

In order to accurately study the carrier dynamics for long SOAs, the length of the SOA is segmented into a number of smaller sections. The reservoir rate equation in (1) is then applied to each section of the amplifier. The rate equation for i^{th} section can now be represented as follows:

$$\frac{dr_i(t)}{dt} = \frac{I}{q} - \frac{r_i(t)}{\tau_i} - S_i^{in}(G(r_i) - 1) + S_{ASE,i}^{in}(G_{ASE}(r_i) - 1) \quad (2)$$

Thus, each section of the SOA can now be treated as an individual traveling type SOA cavity and the rate equations are solved in cascade in order to find the dynamic characteristics $r(t)$ of the propagating signal.

Due to nonlinear relations between SOA gain and its bias current, output of SOA does not show the input OFDM modulation pattern accurately. Severe amount of distortion arises specially at higher peaks of the input OFDM. This phenomenon may be attributed to high PAPR of input OFDM signal. To measure the performance degradation at the output of the SOA, extinction ratio is defined as [10]:

$$R_{ext} = \frac{\sum_{i=1}^{K_1} A^2(i\Delta T) \Big|_{A^2(i\Delta T) > \bar{P}}}{\sum_{i=1}^{K_1} A^2(i\Delta T) \Big|_{A^2(i\Delta T) < \bar{P}}} \frac{K_1}{\sum_{i=1}^{K_2} A^2(i\Delta T) \Big|_{A^2(i\Delta T) < \bar{P}}} K_2 \quad (3)$$

Where,

$$\bar{P} = \frac{\sum_{m=1}^{K_1+K_2} A^2(m\Delta T)}{K_1 + K_2} \quad (4)$$

Here, $A(i\Delta T)$ is the i^{th} signal sample amplitude.

ΔT = sampling interval, $K_1 + K_2$ = Total number of samples.

$A^2 \geq \bar{P}$ for K_1 samples and $A^2 < \bar{P}$ for K_2 samples.

The signal line rate can be calculated from the following expression [10]:

$$R_{sig} = \frac{\sum_{k=2}^{2N} n_k}{T_b} \quad (5)$$

Where, n_k is the total number of binary bits conveyed by the k^{th} sub carrier within one symbol period T_b .

Signal line rate and extinction ratio of the output modulated signal are closely dependent on input parameters, such as input cw optical power and input bias current. Transmission performance is governed by the SOA effective carrier lifetime τ_e which is defined as [11]:

$$\tau_e = \frac{1}{\frac{1}{\tau_c} + \frac{P_{out}}{E_{sat}}} \quad (6)$$

Where, P_{out} is SOA output power and E_{sat} represents SOA saturation energy and τ_c denotes SOA carrier lifetime.

IIB. BRIEF DESCRIPTION OF THE OFDM SYSTEM WITH FREQUENCY DIVERSITY AND LOW PAPR

In this simplified OFDM system, the N -dimensional OFDM symbol at the l -th symbol period is denoted by

$$S_l = \text{IFFT}[X_l X_{l+1}],$$

where, $X_{l+c} = [X_{l+c}(0) \ X_{l+c}(1) \ \dots \ X_{l+c}(N/2-1)]$ ($c=0,1$) of $N/2$ dimensional vector of the OFDM symbol and during the next $(l+1)$ -th symbol period $S_{l+1} = \text{IFFT}[X_{l+1} \ X_l]$ is transmitted.

The transmitted OFDM signal in time domain $s_l(n)$ for $0 \leq n \leq N-1$ is given by

$$s_l(n) = \frac{1}{\sqrt{N}} \left[\sum_{k=0}^{N/2-1} X_l(k) e^{j2\pi kn/N} + \sum_{k=N/2}^{N-1} X_{l+1}(k-N/2) e^{j2\pi kn/N} \right] \quad (7)$$

Similarly, the N -point IFFT output sequence of $s_{l+1}(n)$ can be represented as

$$\begin{aligned} s_{l+1}(n) &= \frac{1}{\sqrt{N}} \left[\sum_{k=0}^{N/2-1} X_{l+1}(k) e^{j2\pi kn/N} + \sum_{k=N/2}^{N-1} X_l(k-N/2) e^{j2\pi kn/N} \right] \\ &= s_l(n) e^{j\pi n} \end{aligned} \quad (8)$$

In ordinary OFDM system, the IFFT is used in each OFDM symbol but in this method, IFFT is used once every two symbols. Thus diversity is provided in the signal. $s_{l+1}(n)$ is obtained from only phase rotation of $s_l(n)$ without IFFT. The phase rotation is required only for odd indexed sample of $s_{l+1}(n)$, whereas phase of even indexed sample is untouched. Peak to average power ratio (PAPR) for a signal $x(t)$ is defined

$$\text{as PAPR} = \frac{\max |s_l(n)|^2}{P_{av}}. \quad (9)$$

We have examined the performance of the SOA in terms of extinction ratio and signal line rate with the PAPR reduction algorithm in OFDM input signal with frequency diversity.

RESULTS AND DISCUSSION

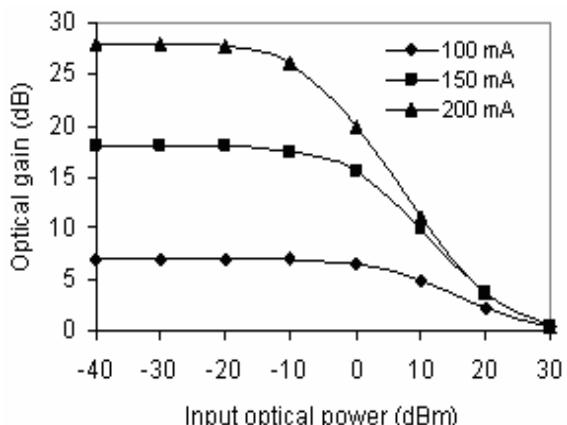


Fig.4: SOA gain vs optical input power plot with bias current as parameters.

A stream of input binary data is first modulated by 16-QAM to produce complex symbols which are fed into IFFT block of 128 sub-carriers in accordance to the algorithm described in

section II B to produce OFDM modulated electrical signal. The resulting OFDM signal modulates the bias current of SOA. DC gain variation of SOA is plotted against input optical power and input bias current and is shown in Fig.4 and Fig.5 respectively. It is seen from Fig.5 that, SOA gain is linearly varying with bias current when input optical power is more or less above 20dBm.

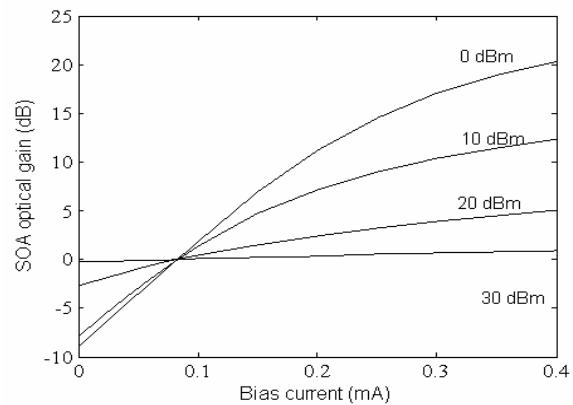


Fig.5: SOA gain vs input bias current plot at Pin: 0dBm, 10 dBm, 20dBm and 30dBm.

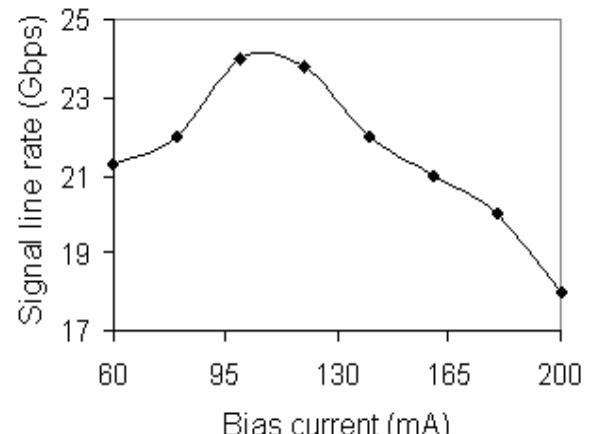


Fig.6: Signal line rate vs. input bias current plot at 20dBm input optical power.

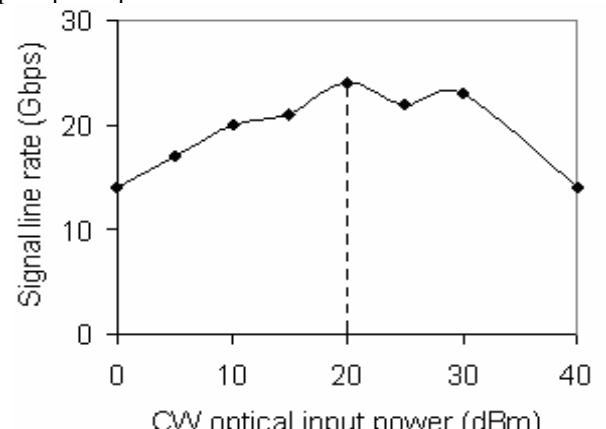


Fig.7: Signal line rate vs. cw optical input power plot at 100mA bias current.

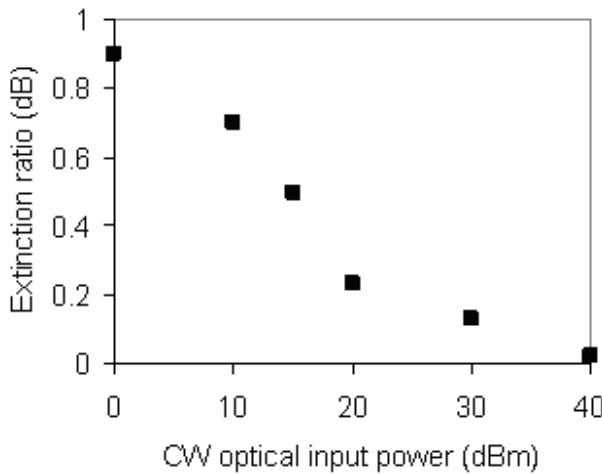


Fig.8: Extinction ratio of SOA output signal as a function of optical input power at bias current of 100mA.

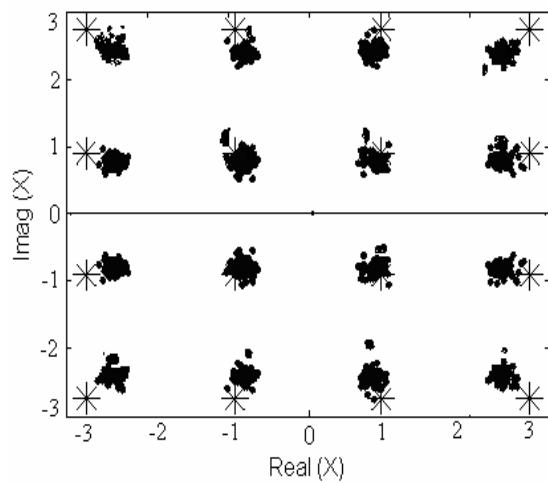


Fig.9: Dynamic constellation shaping for 16-QAM modulation.

Higher the input optical power, more linear is the gain; however this is at the cost of gain reduction. Under such operating conditions, the SOA is strongly saturated. For the purpose of performance evaluation, extinction ratio and signal line rate are calculated from Eq.(3) and (5) respectively. Results are plotted in Fig.6 and Fig.7. From Fig.6 and Fig.7 it is seen that signal line rate is maximum, namely 24.1 Gbps at a bias current of 100mA and an optical power of 20dBm. For bias currents below 100mA, gain is negative and intensity modulated SOA output is clipped. For a bias current of 100mA, when output optical power is more than 20dBm, as seen from Fig.7, signal line rate decreases. This is because, no gain from SOA is then available and consequently signal line rate decreases. Extinction ratio at the SOA output is calculated against cw input optical power. More the input power the less is the SOA gain and lower is the extinction ratio. The result is shown in Fig.8. It is observed from the Fig. 8 that 0.21dB extinction ratio is obtained at 20dBm input optical power and

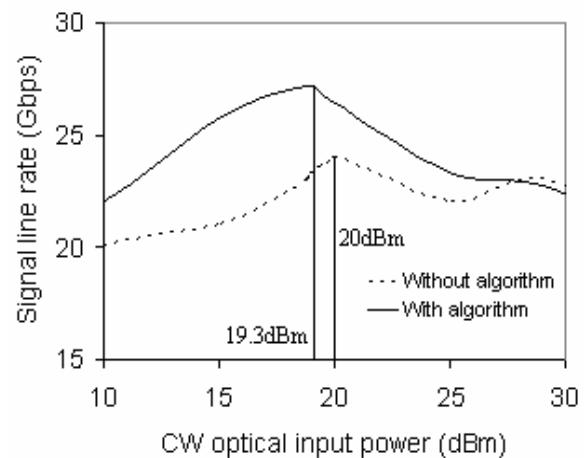


Fig.10: Signal line rate vs. cw optical input power plot at 100mA bias current with application of PAPR reduction algorithm.

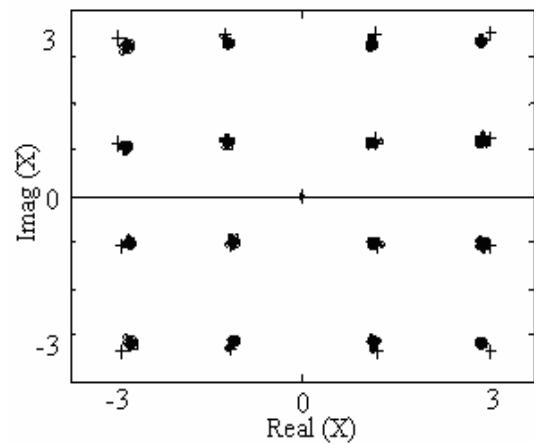


Fig.11: Dynamic constellation shaping for demodulated signal using PAPR reduction algorithm

TABLE-1

| | No. of sub carriers (N) | PAPR of input OFDM (dB) | Optimum input optical power (dBm) |
|-------------------------|-------------------------|-------------------------|-----------------------------------|
| Without any algorithm | 128 | 14.46 | 20 |
| With proposed algorithm | 128 | 12.35 | 19.3 |

at 100mA bias current without any frequency diversity algorithm. Fig.9 shows signal constellation plot for the demodulated symbols at the receiver under the above mentioned operating conditions. The optimum operating condition thus obtained from our simulation is in close

agreement with the experimental results reported by J. L. Wei et.al. (Fig.(3), [10]).

We have investigated the entire performance mentioned above in presence of the algorithm and it has been found that with 0.01% PAPR of new input OFDM symbols can reduce PAPR from 14.46dB to 12.35dB (as shown in Table-1). Fig.10 shows improved performance in terms of higher signal line rate of 27.2 Gbps at input optical power of 19.3dBm and bias current of 100mA. It is observed that application of our PAPR reduction algorithm offers a reduction of input CW optical power to the SOA by 0.7dB (from 20dBm to 19.3dBm). Decoding is carried out at the receiver output and it is observed that improved constellation is achieved than that obtained without the algorithm. This is shown in Fig.11 where a sharp improvement is observed in comparison to Fig.9.

IV. CONCLUSION

In this paper, intensity modulated optical OFDM signal generated by SOA is investigated. Detail analysis showed that an optimum operating condition of SOA can be achieved for which minimum distortion occurs at the receiver. Performance of intensity modulation is evaluated in terms of signal line rate and extinction ratio. Results indicated that 20dBm CW optical input power and 100mA bias current to SOA provides the optimum operating condition. Besides, we have employed frequency diversity and PAPR reduction algorithm into the OFDM system. Introduction of this algorithm reduces the PAPR by 2.11dB. This reduction in PAPR causes lowering of 0.7dB input CW optical power to SOA and improves signal line rate to 27.2 Gbps. OFDM system thus generated will be specially useful for optical transmission through multimode fibers.

REFERENCES

- [1] W. Shieh, "OFDM for flexible high-speed optical networks", *J. Lightw. Technol.*, vol. 29, no. 10, pp.1560–1577, May, 2011.
- [2] C. W. Chow, C. H. Yeh, C. H. Wang, C. L. Wu, S. Chi and C. Lin, "Studies of OFDM signal for broadband optical access networks", *IEEE J. of Sel. Areas in Comm.*, vol. 28, no. 6, Aug., pp.800-807, 2010.
- [3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications", *J. Sel. Area in Comm.*, vol. 16, no. 8, pp. 1451–1458, Oct, 1998.
- [4] Z. Liu, Y. Xin, and G. B. Giannakis, "Space-Time-Frequency coded OFDM over frequency selective fading channels", *IEEE T. on Sig. Proc.*, vol. 50, no. 10, pp. 2465–2476, Oct, 2002.
- [5] Jean Armstrong, "OFDM for Optical Communications", *J. Lightw. Technol.*, vol. 27, no. 3, pp.189–204, Feb., 2009.
- [6] H. Ochiai, "Performance analysis of deliberately clipped OFDM signals", *IEEE Transactions on Comm.*, vol. 50, no. 1, Jan., pp.89-101, 2002.
- [7] Kwang-Cheng Chen and Ramjee Prasad, *Cognitive Radio Network*, Wiley.
- [8] W. Mathlouthi, P. Lemieux, M. Salsi, A. Vannucci, A. Bogoni, and A. Rusch, "Fast and efficient dynamic WDM semiconductor optical amplifier model", *J. Lightwave Technology*, vol.24, no.11, pp. 4353-4365, Nov, 2006.
- [9] A. Das Barman, M. Scalfardi, S. Debnath, L. Poti, and A. Bogoni, "Design tool and its experimental validation for SOA-based photonic signal processing" *Optical Fiber Technology*, Elsevier, vol. 15 no. 1 , pp. 39-49, 2009.
- [10] J. L. Wei, A. Hamie, R. P. Giddings, and J. M. Tang, "Semiconductor optical amplifier-enabled intensity modulation of adaptively modulated optical OFDM signals in SMF-based IMDD systems", *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3678–3688, Aug., 2009.
- [11] M. J. O'Mahony, "Semiconductor laser optical amplifier for use in future fiber systems", *J. Lightwave Technol.* Vol. 6, no. 4, pp. 531–544, 1988.

An Efficient Method For Error Correction in Digital Signal Processing Perspective

Dr.Asok Kumar
 Proff & HOD in ECE Dept
 Asansol Engineering College
 Asansol, India
 asok_km@rediffmail.com

Rupam Das
 Lecturer in ECE Dept
 Asansol Engineering College
 Asansol, India
 das_rupam@rediffmail.com

Mrinmoy Sarkar
 Proff &HOD in ECE Dept
 Bankura Unnayani Institute of Engineering
 Bankura, India
 Mrinmoy.s97@rediffmail.com

Dr. ANUP KUMAR BHATTACHARJEE
 Proff in ECE Dept
 National Institute of Technology
 Durgapur, India
 akbece12@yahoo.com

Abstract— Compressive sampling(CS) is a recent development in digital signal processing that offers the potential of high resolution capture of physical signals from relatively few measurements, typically well below the number expected from the requirements of the Shannon/Nyquist sampling theorem. In this paper we have introduced error-correcting codes that can be defined over the real-number and complex-number fields. The possibility of utilizing real number arithmetic permits the codes to be implemented with operations normally available in standard programmable digital signal processors by digital signal processing methods. Hadamard and discrete Fourier transform codes are presented for block coding, and the latter are presented for the cyclic codes and the class of BCH codes. It is shown that maximum distance separable real-number BCH (N, K) codes exist for all nontrivial values of N and K . A large class of block and convolutional real-number single-error-correcting codes, derived from similar codes over Galois field $GF(p)$, are presented. Both block and convolution codes are seen to be describable by the z-transform in a manner which emphasizes their similarities to conventional digital signal processing structures such as digital filters and digital filter banks. Thus we introduce some methods to encode and decode the class of real and complex numbers by using signal processing methods while using error control coding techniques.[6].

Keywords- DFT code, Galois field $GF(p)$,CS

I. INTRODUCTION

Compressive sampling, also known as compressed sensing or CS theory asserts that one can recover certain signals from far fewer samples or measurements than traditional methods use. To make this possible, CS relies on two principles :sparsity, which pertains to the signals of interest, and incoherence, which pertains to the sensing modality. CS

principles (sparsity, randomness, and convex optimization) can be turned around and applied to design fast error correcting codes over the real's to protect from errors during transmission. Applications of the discrete Fourier transform in the complex field occur throughout the subject of signal processing. Fourier transforms also exist in the Galois field $GF(q)$ and can play an important role in the study and processing of $GF(q)$ -valued signals, that is, of code words. By using the Fourier transform, the ideas of coding theory can be described in a setting that is much closer to the methods of signal processing. Digital transforms like Fourier, cosine, Hadamard, and Karhunen- Loeve transforms being familiar examples of the convolution coding and differential pulse code modulation (DPCM) coders with linear predictors being the most familiar example of the latter. These coding methods are done over the real or complex number fields, and are therefore readily implemented with standard digital signal processors. An error-correcting code, on the other hand, is usually implemented over the binary or other finite field and, as a consequence, requires special hardware for implementation. The use of codes over the real and real numbers for error correction, as opposed to the well-known binary and other finite-field codes, is introduced in this paper. However, many of the well-known algebraic principles of error correction hold over the real-number fields as well as over the finite fields where they are normally employed, and these principles are therefore directly applicable. On the other hand, many of the combinatorial arguments prevalent in the theory of error-correcting codes over finite fields do not carry over to infinite fields. The codes are anticipated to have applications similar to those of Reed-Solomon (RS) codes which perform corrections on bytes of data rather than on a bit-by-bit basis. The reason for this similarity is that in a digital signal processing (DSP) system, each sample value of the real-number or complex-

number sequence is quantized to a byte of data, frequently 16 bits or larger in VLSI signal processors. The codes allow for correction of a certain number of erroneous samples, hence bytes, in a string of otherwise correct samples, just as RS codes do. Typical examples might be the correction of burst errors due to tape dropout, switching, fading, or static. An advantage of a real-number code is that the user of a channel can implement an outer code himself to meet his particular specifications with the facilities he might be expected to have available, namely, digital signal processors at the terminal points of a system.[6],[7].

II. .ENCODING :

A. IMPLEMENTATION ON REAL NUMBER

An error-control code must be judged not only by its rate and minimum distance, but also by whether a decoder can be built for it economically. Usually, there are many ways to decode a given code. A designer can choose among several decoding algorithms and their variations, and must be familiar with all of them so that the best one for a particular application can be chosen. The choice will depend not only on the code parameters, such as block length and minimum distance, but also on how the implementation is to be divided between hardware and software, on the required decoding speed, and even on the economics of available circuit components. In this paper we initially converted the real number sequences into binary format. By using the following methods we transformed the characteristic and mantissa parts of a certain rational numbers like 2.34, 5.6785, etc. The characteristic part can be transformed by 2's complement method while the mantissa part is extracted and is dealt separately. Now for transforming the mantissa part to binary code we followed certain rules .Suppose we have a base ten real number, and want to convert it to its base. On the left side of the decimal point, you perform successive DIVISION in this conversion, and there is a sort of symmetry here because on the right hand side of the decimal point, you perform successive MULTIPLICATION. The goal is to "promote" numbers to the left side of the decimal. Now we multiply over and over by 2 repeatedly. Anytime our answer has content on the left side of the decimal point, we have promoted part of the number. We stop when you reach a zero, and read our answer "down". Now, we go FOREVER without reaching a zero, so usually the question will indicate how many places we need to deal with. Let's look at some examples to get this matter simple.

Eg 1. Determine to eight digits the binary representation of 0.4.

0.4*2=0.8 - Nothing promoted, bring down the .8
 0.8*2=1.6 - Promoted a "1"... drop it, bring the .6
 0.6*2=1.2 - Promoted a "1", drop it, bring down the .2
 0.2*2=0.4 - Nothing promoted
 0.4*2=0.8 - Nothing promoted
 0.8*2=1.6 - ETC... Keep going to support 8 digits
 0.6*2=1.2

$$0.2*2=0.4$$

$$0.4*2=0.8$$

The answer consists of the "Promoted" Column, read top to bottom: 011001100. [1],[2]

USE OF DSP ENCODING:

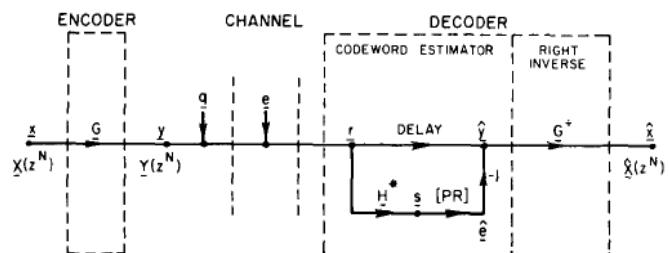


Fig-1:Encoder and Decoder for a real number code.

Now once the real numbers have been converted to binary digits, we use various Error control techniques to encode the digits so that it can be transmitted over the erroneous channel following Shannon's limit of channel capacity with less noise and very limited bandwidth. In our paper we have used Error control coding techniques such as linear block codes, cyclic codes and convolution codes. By these techniques we convert the binary digits into codeword which consists of the binary digits and the parity bits that can be used to check the error in the codeword. From the Fig. 1 we see that in the encoder section Z-transform is applied to the incoming input bit sequence $x(n)$. The resulting transformed sequence $X(z)$ is multiplied with the generator matrix G and we calculate the resulting codeword denoted by $Y(z)$ which is in Z – domain or frequency domain. The transformed sequence is transmitted through the channel which is normally an AWGN (Additive White Gaussian Noise) channel. Thus the noisy sequence is transmitted which provides security amongst the user against hacking, crosstalk and authenticity.[4]

B. DECODING: RECEPTION OF SIGNAL

The erroneous bit pattern of the real number that is been received by the decoder in the channel is finally checked for errors. The entire sequence is received by RAKE receivers that work on the principle of maximum likelihood probability. The signal with the best strength is chosen for decoding and further calculation. This enables complete security amongst the data and thus reduction in noise and providing excellent authenticity. As shown in the Fig.1 the decoder section consists of a delay estimator which is mainly used for calculating the syndrome of the receiver'd signal with the help of the transpose of parity check matrix H . The AWGN noise persisting in the signal is thus removed and a pure codeword including the parity bits is extracted for further computation. The codeword sequence is again de convoluted with the generator matrix G and thus we get the bits in time domain as $x(n)$. We use here syndrome decoding and calculation for knowing the exact position of the error in the bit and thus

correcting the bit in that position. For the convolution coding technique we can use Viterbi decoding method or ML decoding techniques. After we decoded the corrected codeword vectors by the decoding techniques, we convert the received binary digits again to the original real numbers. The characteristic part can be converted to integer by the binary to decimal conversion while the mantissa part can be transformed

| | | | | | | | |
|----|-----|------|-------|--------|---------|----------|---------------|
| .5 | .25 | .125 | .0625 | .03125 | .015625 | .0078125 | .003906 25 |
| 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |
| 0 | .5 | .125 | 0 | 0 | .015625 | .0078125 | 0 |

$$0.25 + 0.125 + 0.015625 + 0.0078125 = \mathbf{0.3984}$$

III. ALGORITHM:

A. ENCODING

```

fprintf('\n\n ENCODING \n\n');
for n=1:1:a
    b = mod (n,a);
    if (b+1)>= a
        b;
    else
        b=0;
    end
end

c = a - b ;
fprintf('The characteristic is %d',b);
fprintf(' and the mantissa is %1.8f',c);

fprintf('\n\nThe binary equivalent of characteristic is ');
p = dec2bin(b);
disp(p);

fprintf('\nThe binary equivalent of mantissa is ');

for m = 1:1:10
    c = c*2;

    if 0 < c & c < 1
        d = 0;
    else
        d = 1;
    end

    while c >= 1
        by 2
        c = c - 1 ;
    end
end

fprintf('%1.0f ',d);
t(m)= d;
end

fprintf('\n\nThe complete encoded binary equivalent of
the real number is \n');
```

```

disp(p);
fprintf('.');
fprintf("%1.0f",t);
```

B. DECODING

```

fprintf('\n\n DECODING \n');
s=0;
fprintf('\n\nThe approximate decoded decimal equivalent
of the binary encoded mantissa is ');
for j = 1:1:10
    s = s + t(j) .* 2^(-j);
    time

    if j==8
        fprintf(' %1.10f',s );
    end
    end

x = bin2dec(p);
y = x+s;
fprintf('\n\nThe complete approximate decoded real
number is %1.10f',y);
```

IV. CONCLUSION:

The new field of error correction over the real numbers appears to offer many opportunities for improving systems performance employing DSP concepts. Coding for error correction over the real and complex numbers is feasible, and the codes can be implemented in standard digital signal processors using primarily linear shift-invariant structures except for the nonlinear error estimators. Many of the algebraic concepts of error correction over finite fields are directly applicable to real-number coding, but it was seen to be advantageous to also introduce the z-transform and other digital signal processing concepts to provide a unified framework for describing both convolution and block codes. Block transforms and their resulting frequency domain view points were seen to be useful both for discovering and describing block codes; in addition, frequency domain concepts and digital filter bank descriptions were seen to be useful for describing convolution codes and are logical extensions of the block transform descriptions. Many of the well known algebraic principles of error correction codes hold over the fields of real number and these principles are therefore directly applicable . The signal processing techniques which are introduced in this paper are appropriate for different noisy environment

V. REFERENCES:

- [1]Richard.E. Blahut, " *Algebraic Codes for Data Transmission*", Cambridge University press, The Edinburgh Building, Cambridge U.K., 2003, Pg. 131-172 .
- [2]T. G. Marshall, Jr., " *Methods for error correction with digital signal processors*," Houghton, MI, Aug. 1982, pp. 1-5.
- [3] W.W Peterson and E.J Weldon Jr, Error Correcting Codes,Cambridge , MA, MIT Press,1972
- [4]R.E.Blahut,Theory and Practice of Error Control Codes,reading MA:Addison – Wesley,1983

2nd Annual International Conference IEMCON 2012

- [5]S.Jafarzadeh,M.Khatami, and F.Marvasti,Senior Member, Decoding Real Numbered Block and Convolutional Codes with Erasure and Impulsive Noise Channels;,(EUSIPCO-2010)
- [6]F.Marvasti,A.Minimi,AHaddadi,M.Soltanolkotabi,B.h.Kahalaj,A.Aldroubi, S.Holm,S.Sanei,J.Chambera.'Aunified Approach to Sparce Signal Processing,",Feb 2009
- [7]Emmanuel J.candes,Michal b.Wkin,"An Introduction To Compressive Sampling, IEEE Signal Processing Magazine, March,2008

Reduction in Power Consumption With the Implementation of LED

Sourav Sarkar*(4th Year), Amitava Sengupta**

#ECE Department,
Institute of Engineering & Management
West Bengal University of Technology
BF 142 Sector I, Salt Lake Kolkata, West Bengal 700064, INDIA
Email id : saurav.sarkar@ymail.com

**Philips Electronics India Ltd.(Kolkata- Eastern Regional Office)
Lighting Division
7, Justice Chandra Madhab Road
Kolkata-700020
Email id : amitava.sengupta@philips.com

Abstract :

This paper mainly deals with the reduction in power consumption & light power density with the application & implementation of Light Emitting Diode by replacing Incandescent lighting sources. For a proper design shaping of light emitting diodes, the exact knowledge of the current distribution in the active area is essential. On the one hand, one has to achieve a uniform current density over the chip area, on the other hand, one has to avoid current crowding in the neighborhood of the electrical contacts. Proper readings with the help of DIALUX software has been taken & it is clearly shown how power consumption & light power density, of the same area with the same no luminaires ,has decreased to a considerable amount.

Keywords :

Basics, Design with Incandescent, Design with LED, Comparison by Tables, Inference

INTRODUCTION :

Despite the fact that light emitting diodes (LEDs) have been present in the market place for many years, there are only a few efforts made to model their characteristics with the help of the finite element method. The reason for

this may be the strong non-linearity of the current-voltage characteristic of the p-n-junction that makes it sometimes difficult to achieve convergence in the models. In this paper an LED model for the current density distribution ,power distribution & uniformity has been presented with the help of DIALUX. A light-emitting diode consists of a n- and a p-doped semiconductor region with metallic n- and p-contacts as electrodes. In the junction of the doped regions the electrons, which are generated in the n-doped region, and the holes, which are generated in the p-doped region, recombine and generate photons. As the photons should not be absorbed by the contacts, the n- and p-contact must have small area and should not be on top of each other. To ensure a proper current flow, the n- and p-doped regions (called spreading layers) must have a certain thickness, so that the current can spread over the whole chip.

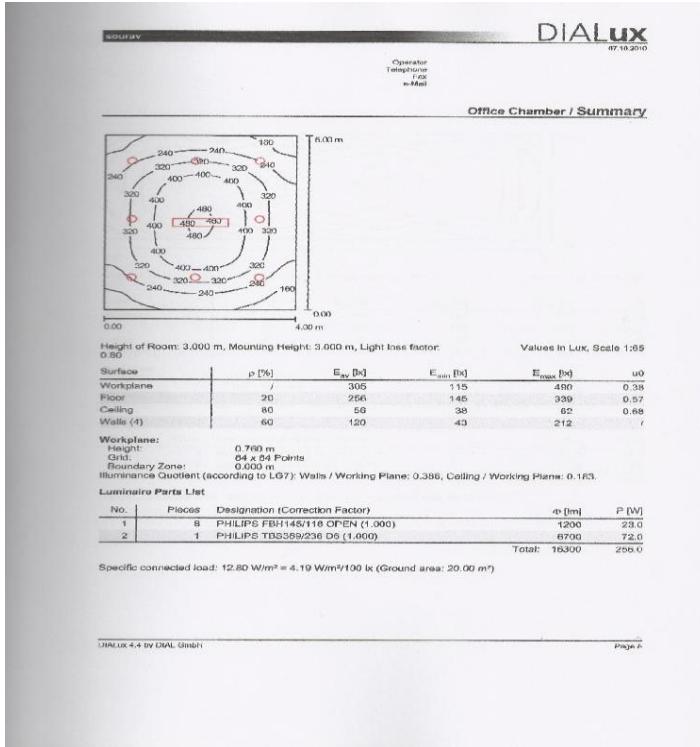
Comparative study is being done using DIALux. DIALux is a free light calculation program which is primarily used by lighting designers , consultants, architects , lighting technicians and electricians. It can be used to calculate indoor and outdoor lighting. Indoor enable it to calculate the lighting levels in a given room, thereby satisfying the requirements of the DS-700 . Indoor workplaces DS EN 12193 , Sports Lighting DS / EN 1838 , Emergency lighting DS-705 , Artificial lighting in the dental surgery .

2nd Annual International Conference IEMCON 2012

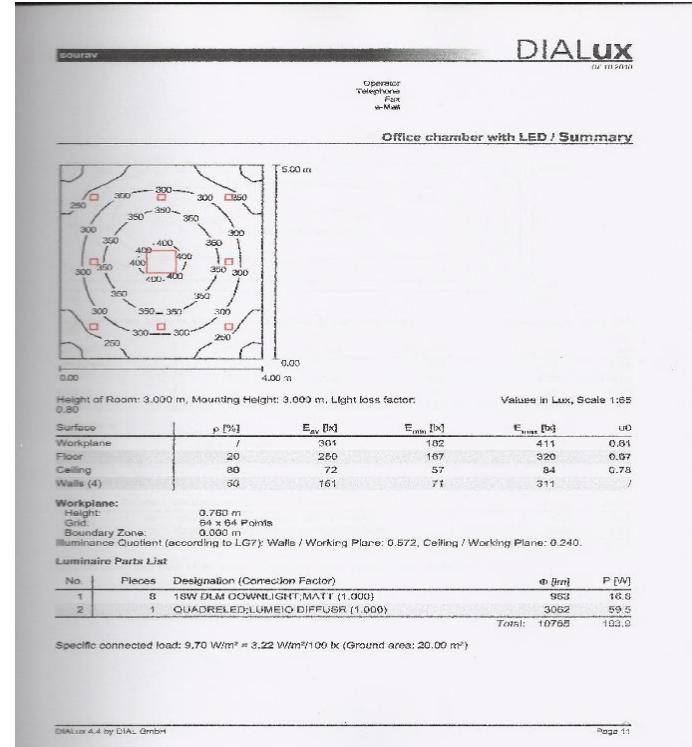
Outdoor enable it to calculate light for outdoor workplaces, stadiums, roads, paths, squares, etc. and

meet DS EN 12464-2-2007 , light at jobs outdoor jobs DS EN 12193 and Light and Lighting, Sports

Office chamber with Incandescent :

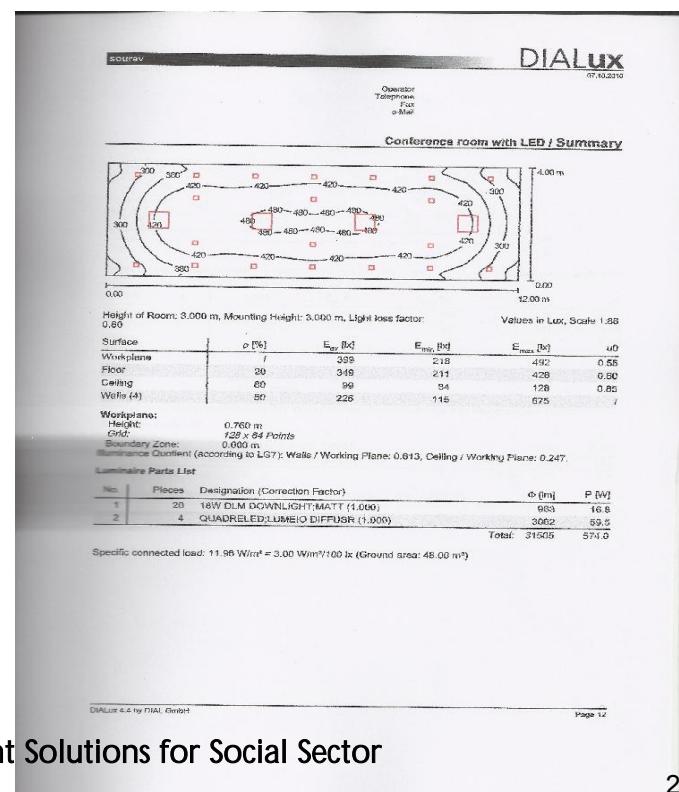
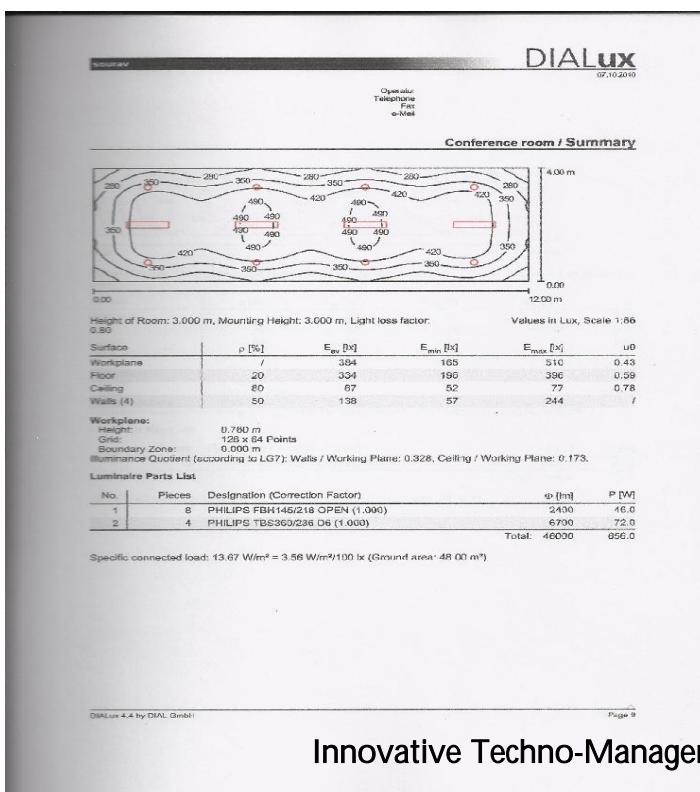


Office chamber with LED



Conference room with Incandescent :

Conference room with LED :



Office Chamber with Conventional Design :

An office chamber with the required dimensions of length 5m,breadth 4m & height 3m was taken in course of the experiment and collection of research data.The conventional luminaires which were present at the time of data collection TBS 369/236 D6 * 1 and FBH 145/218 * 8(both were manufactured by PHILIPS).These luminaires can be replaced by other ones provided they satisfy the required parameters corresponding to the above mentioned ones or they may be as close as possible.The lux level of those 9 luminaires summed upto 305 lx,providing an uniformity of just 0.38.The total power consumption of the lighting sources were 256 watts and the light power density was calculated upto 12.5 watts/sq.m.Here the uniformity obtained was not at all satisfactory due to the uneven distribution of light rays from the incandescent luminaires.The power consumed was pretty much higher than what was expected and thus the consumption cost was huge.So,to provide the necessary means to reduce the power consumption and the usage cost the same arrangements were checked and verified with LED luminaires,which showed drastic reduction in power consumption,a satisfactory increase of uniformity,an accepted level of LPD which was equivalent to international ratings and industry standards.

Office Chamber with LED Design :

The same office chamber was illuminated with the LED luminaires LED QUADRA * 1 and LEDO BBS 145 * 8 which have the or closest lux ratings with the Incandescent luminaires.Using the LED luminaires led to the decrease in the lux level,LPD,total power consumption and satisfactory increase in uniformity which is a necessity for good lighting planning measures.the tables give a detailed analysis of the reports which assure that the implementation of LED luminaires instead of incandescent ones reduces the total power consumption thus decreasing the cost and helping in the trade – off.

Conference Room with Conventional Design :

An office chamber with the required dimensions of length 12m,breadth 4m & height 3m was taken in course of the experiment and collection of research data.The conventional luminaires which were present at the time of data collection TBS 369/236 D6 * 4 and FBH 145/218 * 8(both were manufactured by PHILIPS).These luminaires can be replaced by other ones provided they satisfy the required parameters corresponding to the above mentioned ones or they may be as close as possible.The lux level of those 12 luminaires summed upto 384 lx,providing an uniformity of just 0.43.The total power consumption of the lighting sources were 656 watts and the light power density was calculated upto 13.66 watts/sq.m.Here the uniformity obtained was not at all satisfactory due to the uneven distribution of light rays from the incandescent luminaires.The power consumed was pretty much higher than what was expected and thus the consumption cost was huge.So,to provide the necessary means to reduce the power consumption and the usage cost the same arrangements were checked and verified with LED luminaires,which showed drastic reduction in power consumption,a satisfactory increase of uniformity,an accepted level of LPD which was equivalent to international ratings and industry standards.

Conference Room with LED Design :

The same office chamber was illuminated with the LED luminaires LED QUADRA * 4 and LEDO BBS 145 * 8 which have the or closest lux ratings with the Incandescent luminaires.Using the LED luminaires led to the decrease in the lux level,LPD,total power consumption and satisfactory increase in uniformity which is a necessity for good lighting planning measures.the tables give a detailed analysis of the reports which assure that the implementation of LED luminaires instead of incandescent ones reduces the total power consumption thus decreasing the cost and helping in the trade – off.

Corridor with Conventional Design :

An office chamber with the required dimensions of length 5m,breadth 4m & height 3m was taken in course of the

experiment and collection of research data. The conventional luminaires which were present at the time of data collection FBH 145/218 * 6 (manufactured by PHILIPS). These luminaires can be replaced by other ones provided they satisfy the required parameters corresponding to the above mentioned ones or they may be as close as possible. The lux level of those 6 luminaires summed upto 140 lx, providing an uniformity of just 0.57. The total power consumption of the lighting sources were 276 watts and the light power density was calculated upto 13.35 watts/sq.m. Here the uniformity obtained was not at all satisfactory due to the uneven distribution of light rays from the incandescent luminaires. The power consumed was pretty much higher than what was expected and thus the consumption cost was huge. So, to provide the necessary means to reduce the power consumption and the usage cost the same arrangements were checked and

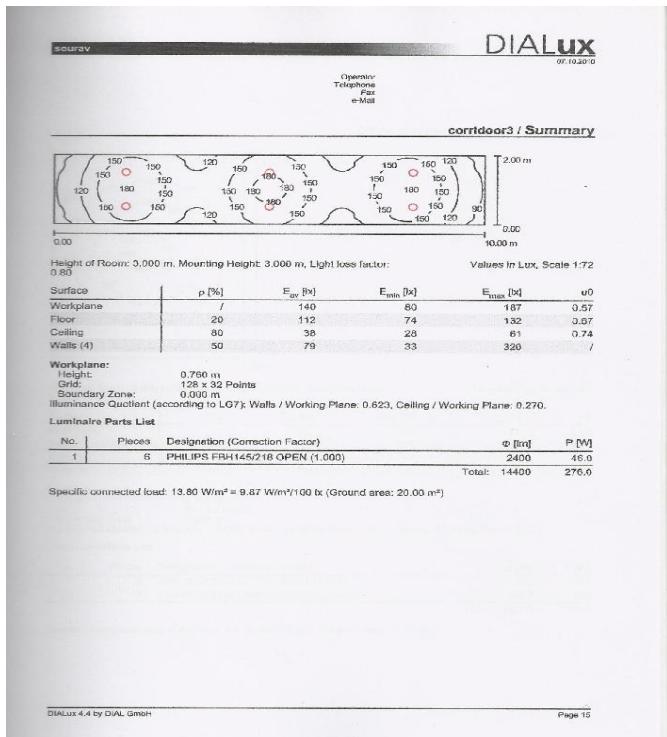
verified with LED luminaires, which showed drastic reduction in power consumption, a satisfactory increase of uniformity, an accepted level of LPD which was equivalent to international ratings and industry standards.

Corridor with LED Design:

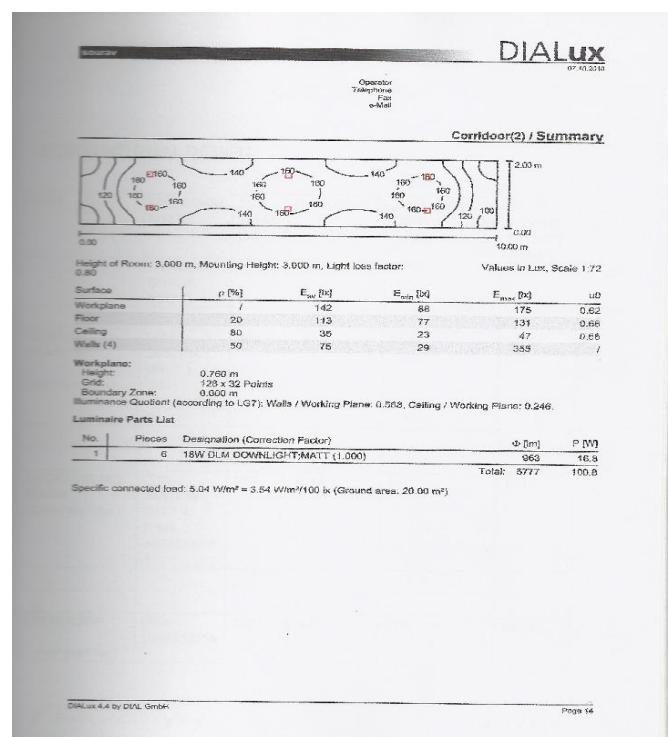
The same office chamber was illuminated with the LED luminaires LED QUADRA * 1 and LEDO BBS 145 * 8 which have the or closest lux ratings with the Incandescent luminaires. Using the LED luminaires led to the decrease in the lux level, LPD, total power consumption and satisfactory increase in uniformity which is a necessity for good lighting planning measures. The tables give a detailed analysis of the reports which assure that the implementation of LED luminaires instead of incandescent ones reduces the total power consumption thus decreasing the cost and helping in the trade – off.

2nd Annual International Conference IEMCON 2012

Corridor with Incandescent :



Corridor with LED :



CONVENTIONAL DESIGN :

| AREA | LUMINAIRE USED | LUX LEVEL (lx) | UNIFORMITY | TOTAL POWER CONSUMPTION (watt) | LIGHT POWER DENSITY (watt/sq.m) |
|-------------------------------|--|----------------|------------|--------------------------------|---------------------------------|
| | ITEM/QTY (PHILIPS MAKE) | | | | |
| 1. Office Chamber (5m*4m*3m) | (TBS 369/236 D6) * 1 (FBH 145/218) * 8 | 305 | 0.38 | 256 | 12.5 |
| 2.CONFERENCE ROOM (12m*4m*3m) | (TBS 369/236 D6) * 4 (FBH 145/218) * 8 | 384 | 0.43 | 656 | 13.66 |
| 3.CORRIDOR (12m*4m*3m) | (FBH 145/218)* 6 | 140 | 0.57 | 276 | 13.35 |

| AREA | LUMINAIRE USED | LUX LEVEL (lx) | UNIFORMITY | TOTAL POWER CONSUMPTION (WATT) | LIGHT POWER DENSITY (WATT/S Q.M) |
|-------------------------------|-------------------------------------|----------------|------------|--------------------------------|----------------------------------|
| | ITEM/QTY (PHILIPS MAKE) | | | | |
| 1.OFFICE CHAMBER (5m*4m*3m) | (LED QUADRA) * 1 (LEDO BBS 145) * 8 | 301 | 0.61 | 193.9 | 9.65 |
| 2.CONFERENCE ROOM (12m*4m*3m) | (LED QUADRA) * 4 (LEDO BBS 145) * 8 | 374 | 0.57 | 558 | 11.625 |
| 3.CORRIDOR (12m*4m*3m) | (BBS 145) * 6 | 142 | 0.62 | 100.8 | 5.4 |

2nd Annual International Conference IEMCON 2012

EXPERIMENTAL OBSERVATIONS :

| AREA | LED DESIGN | | CONVENTIONAL DESIGN | | % DECREASE IN TOTAL POWER CONSUMPTION | %DECREE IN LIGHT POWER DENSITY |
|--------------------------------|--------------------------------|----------------------------------|--------------------------------|----------------------------------|---------------------------------------|--------------------------------|
| | TOTAL POWER CONSUMPTION (WATT) | LIGHT POWER DENSITY (WATT/S Q.M) | TOTAL POWER CONSUMPTION (WATT) | LIGHT POWER DENSITY (WATT/S Q.M) | | |
| 1.OFFICE CHAMBER (5m*4 m*3m) | 193.9 | 9.65 | 256 | 12.5 | 24.25 | 22.8 |
| 2.CONFERENCE ROOM (12m*4 m*3m) | 558 | 11.625 | 656 | 13.66 | 14.93 | 14.89 |
| 3.CORRIDOR (12m*4 m*3m) | 100.8 | 5.4 | 276 | 13.35 | 63.47 | 59.55 |

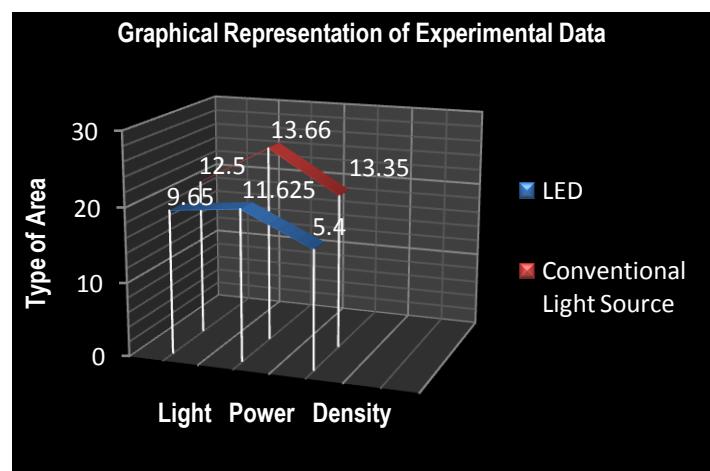
INFERENCE FROM EXPERIMENTAL OBSERVATIONS:

Since the wattage of LED lamps are lower than the incandescent bulbs and since the power consumption is also less so the change in light power density (LPD) is so much deviating in case of LED design from Conventional design.

Light power density for a particular area is mathematically expressed as,

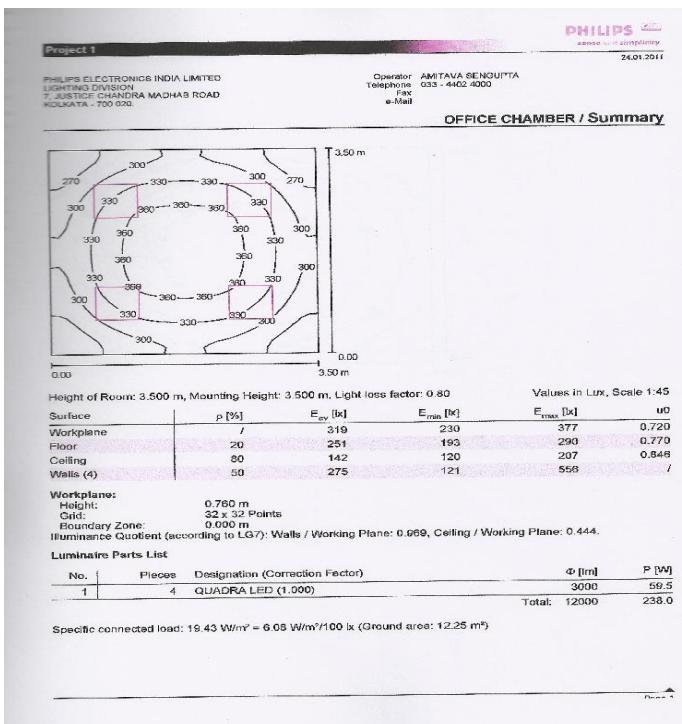
LPD = watt consumed / sq.m

Since the wattage(watt) of LED lamps decreases so the Light Power Density of LED lamps decreases from incandescent bulbs.

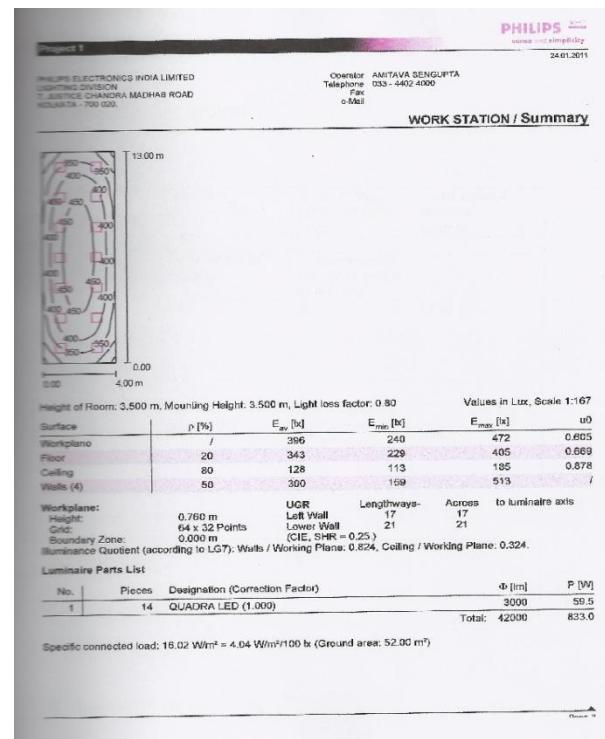


CASE STUDY :

Office chamber with LED :



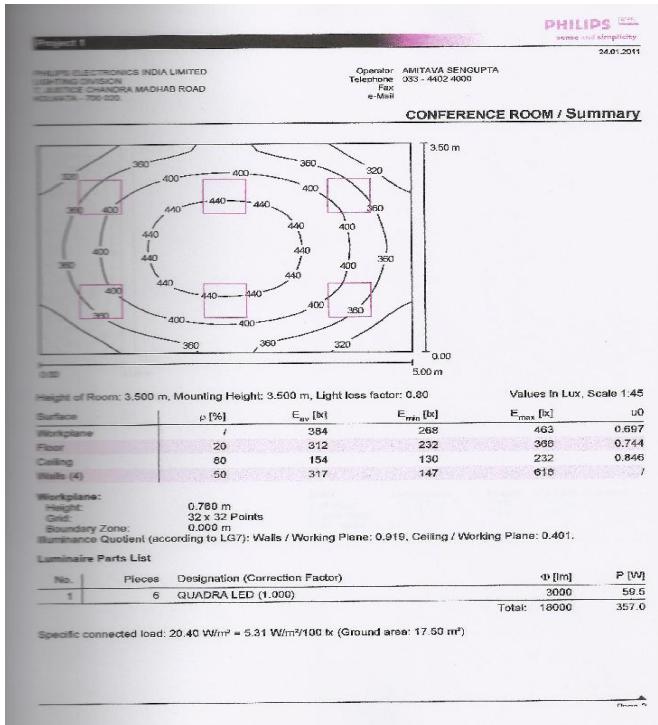
Work Station with LED :



2nd Annual International Conference IEMCON 2012

Conference room with LED :

CONVENTIONAL DESIGN :



| AREA | LUX LEVEL (lx) | TOTAL POWER CONSUMPTION (watt) | LIGHT POWER DENSITY (watt/s) |
|------------------------------------|--|--------------------------------|------------------------------|
| (DATA GIVEN BY PHILIPS) | (ACCORDING TO THE DATA GIVEN BY PHILIPS) | | |
| 1. Office Chamber (3.5m*3.5m*3.5m) | 350 | 264 | 21.12 |
| 2.CONFERENCE ROOM (5m*3.5m*3.5m) | 400 | 462 | 26.40 |
| 3.CORRIDOR (13m*4m*3.5m) | 400 | 1078 | 20.730 |

LED DESIGN :

EXPERIMENTAL OBSERVATIONS :

| AREA | LUMINAIRE USED | LUX LEVEL (LX) | UNIFORMITY | TOTAL POWER CONSUMPTION (WATT) | LIGHT POWER DENSITY (WATT/SQ.M) |
|-----------------------------------|-------------------------|----------------|------------|--------------------------------|---------------------------------|
| | ITEM/QTY (PHILIPS MAKE) | | | | |
| 1.OFFICE CHAMBER (3.5m*3.5m*3.5m) | (QUADRAL ED)* 4 | 319 | 0.720 | 238 | 19.42 |
| 2.CONFERENCE ROOM (5m*3.5m*3.5m) | (QUADRA LED) * 6 | 384 | 0.697 | 357 | 20.4 |
| 3.WORK STATION (13m*4m*3.5m) | (QUADRA LED) * 14 | 396 | 0.605 | 833 | 16.01 |

| AREA | LED DESIGN | | CONVENTIONAL DESIGN | | % DECREASE IN TOTAL POWER CONSUMPTION | %DECR EASE IN LIGHT POWER DENSITY |
|-----------------------------------|--------------------------------|---------------------------------|--------------------------------|---------------------------------|---------------------------------------|-----------------------------------|
| | TOTAL POWER CONSUMPTION (WATT) | LIGHT POWER DENSITY (WATT/SQ.M) | TOTAL POWER CONSUMPTION (WATT) | LIGHT POWER DENSITY (WATT/SQ.M) | | |
| 1.OFFICE CHAMBER (3.5m*3.5m*3.5m) | 238 | 19.42 | 264 | 21.12 | 9.8 | 8.04 |
| 2.CONFERENCE ROOM (5m*3.5m*3.5m) | 357 | 20.40 | 462 | 26.40 | 22.727 | 22.727 |
| 3.WORK STATION (13m*4m*3.5m) | 833 | 16.01 | 1078 | 20.730 | 22.727 | 22.768 |

Case Study :

During the case study all the conventional data were provided by PHILIPS.Only the data were collected based on the designs of the conventional ones.

Office Chamber with Conventional Design :

An office chamber with the required dimensions of length 3.5m,breadth 3.5m & height 3.5m was taken in course of the experiment and collection of research data.The conventional luminaires which were present at the time of data collection were manufactured by PHILIPS.These luminaires can be replaced by other ones provided they satisfy the required parameters corresponding to the above mentioned ones or they may be as close as possible.The lux level of those 4 luminaires summed upto 350 lxThe total power consumption of the lighting sources were 264 watts and the light power density was calculated upto 21.12 watts/sq.m.Here the uniformity obtained was not at all satisfactory due to the uneven distribution of light rays from the incandescent luminaires.The power consumed was pretty much higher than what was expected and thus the consumption cost was huge.So,to provide the necessary means to reduce the power consumption and the usage cost the same arrangements were checked and verified with LED luminaires,which showed drastic reduction in power consumption,a satisfactory increase of uniformity,an accepted level of LPD which was equivalent to international ratings and industry standards.

Conference Room with Conventional Design :

An office chamber with the required dimensions of length 5m,breadth 3.5m & height 3.5m was taken in course of the experiment and collection of research data.The conventional luminaires which were present at the time of data collection were manufactured by PHILIPS.These luminaires can be replaced by other ones provided they satisfy the required parameters corresponding to the above mentioned ones or they may be as close as possible.The lux level of those 6 luminaires summed upto 400 lxThe

total power consumption of the lighting sources were 462 watts and the light power density was calculated upto 26.40 watts/sq.m.Here the uniformity obtained was not at all satisfactory due to the uneven distribution of light rays from the incandescent luminaires.The power consumed was pretty much higher than what was expected and thus the consumption cost was huge.So,to provide the necessary means to reduce the power consumption and the usage cost the same arrangements were checked and verified with LED luminaires,which showed drastic reduction in power consumption,a satisfactory increase of uniformity,an accepted level of LPD which was equivalent to international ratings and industry standards.

Corridor with Conventional Design :

An office chamber with the required dimensions of length 13m,breadth 4m & height 3.5m was taken in course of the experiment and collection of research data.The conventional luminaires which were present at the time of data collection were manufactured by PHILIPS.These luminaires can be replaced by other ones provided they satisfy the required parameters corresponding to the above mentioned ones or they may be as close as possible.The lux level of those 4 luminaires summed upto 400 lxThe total power consumption of the lighting sources were 1078 watts and the light power density was calculated upto 20.730 watts/sq.m.Here the uniformity obtained was not at all satisfactory due to the uneven distribution of light rays from the incandescent luminaires.The power consumed was pretty much higher than what was expected and thus the consumption cost was huge.So,to provide the necessary means to reduce the power consumption and the usage cost the same arrangements were checked and verified with LED luminaires,which showed drastic reduction in power consumption,a satisfactory increase of uniformity,an accepted level of LPD which was equivalent to international ratings and industry standards.

Office Chamber with LED :

An office chamber with the required dimensions of length 3.5m,breadth 3.5m & height 3.5m was taken in course of the experiment and collection of research data.The LED luminaires which were present at the time of data collection QUADRALED * 4(manufactured by PHILIPS).These luminaires can be replaced by other ones provided they satisfy the required parameters corresponding to the above mentioned ones or they may be as close as possible.The lux level of those 4 luminaires summed upto 319 lx,providing an uniformity of just 0.720.The total power consumption of the lighting sources were 238 watts and the light power density was calculated upto 19.42 watts/sq.m.Here the uniformity obtained was satisfactory due to the even distribution of light rays from the LED luminaires.The power consumed was pretty much lower than what was expected and thus the consumption cost was huge.So,to provide the necessary means to reduce the power consumption and the usage cost the same arrangements were checked and verified with LED luminaires,which showed drastic reduction in power consumption,a satisfactory increase of uniformity,an accepted level of LPD which was equivalent to international ratings and industry standards.

Conference Room with LED :

An office chamber with the required dimensions of length 5m,breadth 3.5m & height 3.5m was taken in course of the experiment and collection of research data.The LED luminaires which were present at the time of data collection QUADRALED * 6(manufactured by PHILIPS).These luminaires can be replaced by other ones provided they satisfy the required parameters corresponding to the above mentioned ones or they may be as close as possible.The lux level of those 4 luminaires summed upto 384 lx,providing an uniformity of just 0.697.The total power consumption of the lighting sources were 357 watts and the light power density was calculated upto 20.4 watts/sq.m.Here the uniformity obtained was

satisfactory due to the even distribution of light rays from the LED luminaires.The power consumed was pretty much lower than what was expected and thus the consumption cost was huge.So,to provide the necessary means to reduce the power consumption and the usage cost the same arrangements were checked and verified with LED luminaires,which showed drastic reduction in power consumption,a satisfactory increase of uniformity,an accepted level of LPD which was equivalent to international ratings and industry standards.

Corridor with LED :

An office chamber with the required dimensions of length 13m,breadth 4m & height 3.5m was taken in course of the experiment and collection of research data.The LED luminaires which were present at the time of data collection QUADRALED * 4(manufactured by PHILIPS).These luminaires can be replaced by other ones provided they satisfy the required parameters corresponding to the above mentioned ones or they may be as close as possible.The lux level of those 4 luminaires summed upto 396 lx,providing an uniformity of just 0.605.The total power consumption of the lighting sources were 833 watts and the light power density was calculated upto 16.01 watts/sq.m.Here the uniformity obtained was satisfactory due to the even distribution of light rays from the LED luminaires.The power consumed was pretty much lower than what was expected and thus the consumption cost was huge.So,to provide the necessary means to reduce the power consumption and the usage cost the same arrangements were checked and verified with LED luminaires,which showed drastic reduction in power consumption,a satisfactory increase of uniformity,an accepted level of LPD which was equivalent to international ratings and industry standards.

INFERENCE FROM EXPERIMENTAL OBSERVATIONS:

Since the wattage of LED lamps are lower than the incandescent bulbs and since the power consumption is also less so the change in light power density (LPD) is so much deviating in case of LED design from Conventional design.

Light power density for a particular area is mathematically expressed as,

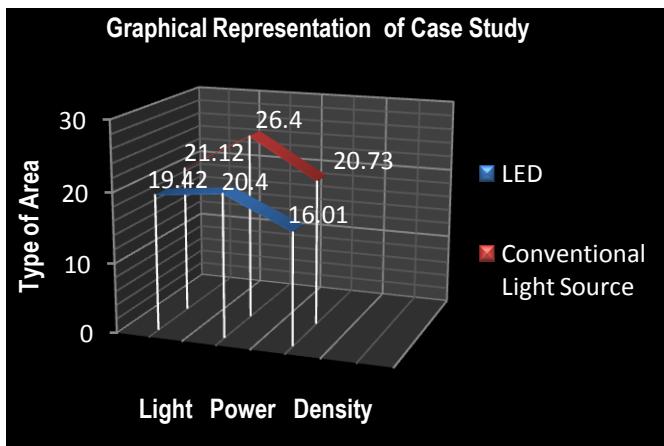
$$\text{LPD} = \text{watt consumed / sq.m}$$

Since the wattage(watt) of LED lamps decreases so the Light Power Density of LED lamps decreases from incandescent bulbs.

characteristic for the pn-junction. The simulated quantities are the area light power density-characteristic, the wattage consumed. The agreement between the simulated and experimental results is very good. It is shown that the uniformity can be optimised significantly with the help of the simulations.

REFERENCES :

1. Philips study material
2. Philips technical report
3. Philips statistics
4. Philips databook
5. Philips datasheet
6. www.philips.co.in
7. COMSOL Users Conference 2007 Grenoble



CONCLUSION :

In this paper a finite element model for the current light intensity distribution across an LED chip is presented. The model has been implemented in DIALUX, using a

two domain approach for the n- and p-layers with a boundary condition containing an effective current-voltage

2nd Annual International Conference IEMCON 2012

----- REVIEW 1 -----

PAPER: 48

TITLE: Reduction of Power Consumption with the Implementation of LED

AUTHORS: Sourav Sarkar and Amitava Sengupta

Question:

1 Relevance to the conference:

2 Originality:

3 Detailed Comments:

The study of reduction for power consumption using LED by replacing incandescent light sources is very important for different applications. The authors can compare their results with other results taking other references. The authors can reduce the size of the paper by reducing the size of the figures.

3 Detailed Comments:

The study of reduction for power consumption using LED by replacing incandescent light sources is very important for different applications. The authors can compare their results with other results taking other references. The authors can reduce the size of the paper by reducing the size of the figures.

4 Overall Rating: Good

5 Comments to Program Chair:

4 Overall Rating: Good

5 Comments to Program Chair:

----- REVIEW 2 -----

PAPER: 48

TITLE: Reduction of Power Consumption with the Implementation of LED

AUTHORS: Sourav Sarkar and Amitava Sengupta

Question:

1 Relevance to the conference:

2 Originality:

Effect of Conduction Band Nonparabolicity on Transmission Coefficient of Double Quantum Well Triple Barrier Structure with Parabolic Geometry

Souvik Mukherjee

U.G Student, Dept. of Electronics & Comm. Engineering
RCC Institute of Information Technology
Kolkata, INDIA
souvik.rcc029@gmail.com

Rajit Karmakar

U.G Student, Dept. of Electronics & Comm. Engineering
RCC Institute of Information Technology
Kolkata, INDIA
rajit.karmakar@gmail.com

Arindam Biswas

U.G Student, Dept. of Electronics & Comm. Engineering
RCC Institute of Information Technology
Kolkata, INDIA
arindamb75@gmail.com

Arpan Deyasi

Assistant Professor, Dept. of Electronics & Comm. Engg
Institute of Information Technology
Kolkata, INDIA
deyasi_arpan@yahoo.co.in

Abstract—Resonant tunneling probability of a parabolic double quantum well triple barrier structure in presence of finite thin contact barriers is computed including conduction band nonparabolicity for GaAs/Al_xGa_{1-x}As material composition along with consideration of effective mass mismatch at junctions following BenDaniel Duke boundary condition and band discontinuity. Accounting nonparabolicity factor of Al_xGa_{1-x}As material is also associated with spin-orbit splitting factor, providing closer image of realistic band structure which is important from accuracy point of view of estimated result. Modified dispersion relationship upto first order is taken into account for simulation purpose, and transmission coefficient is obtained by propagation matrix method. Effect of different barrier thicknesses and well widths are independently studied on transmission coefficient, and also for a specified structure, material composition of barrier layers is varied to observe the tunneling effect.

Keywords- *Band Nonparabolicity, Double Quantum Well Triple Barrier Structure, Parabolic Geometry, Transmission Coefficient*

I. INTRODUCTION

Accurate design of resonant tunneling semiconductor nanodevices is dependent on dimensional configuration and material composition of different layers, as well as realistic band structure of the materials. This greatly affects the performance of the novel quantum electronic and photonic structures [1]-[2], which may bring the beacon of next generation VLSI technology. Carrier confinement along reduced dimensions can be realized by quantum wells, wires and dots, where dimension of confinement is comparable to the de-Broglie wavelength. Since in 1-D confined structures, displacement of energy levels from band-edge should be considered for tailoring transport properties, so conduction-band nonparabolicity plays a very important role to study

quantum phenomena even in absence of bias. Thus precise estimation of transmission coefficient is essential for the device with incorporation of physical parameters [3]-[4] when relative interdependency of these data should be taken into account for mathematical modeling.

Miller [5] investigated the effect of conduction band nonparabolicity on eigenenergies considering the origin of Nonparabolicity effect from Kane's two-band model, Hiroshima [6] considered band nonparabolicity as a function of material parameters of the ultrathin layers. Nelson [7] considered energy-dependent effective mass approach to evaluate the same, and finite element method was introduced by Dave [8] for more precise estimation.

Theoretical investigation on parabolic quantum well structure was carried out by Cruz [9], while its practical importance was studied in designing optical transmitter [10]-[11] by calculating absorption coefficient for interband optical transitions. Intraband transitions in parabolic multiquantum well structures are calculated for designing novel infrared detector [12].

Different numerical methods are already adopted by theoretical researchers to analyze double quantum well triple barrier structure such as Variational method [13], Fourier Series method [14], Propagation Matrix method [15], Finite Element Method [8], Finite Difference method [16] etc. Among all the methods, Propagation Matrix method is considered by the authors as one of the accurate techniques to compute transmission coefficient of Double Quantum Well Triple Barrier structure for parabolic well shape even with the incorporation of band nonparabolicity.

The paper deals with theoretical investigation of transmission probability of DQWTB structure considering parabolic geometry and conduction band nonparabolicity for GaAs/Al_xGa_{1-x}As material composition. Effective mass

mismatch is considered at junctions following BenDaniel Duke boundary condition [17] and conduction band discontinuity is also considered as function of barrier material. Contact and middle barrier widths are always considered less than well width. Both dimensional configuration and material composition is varied to see the change in transmission coefficient, and band nonparabolicity is considered to make a comparative analysis with the case when the effect is absent. First order E- κ relation is taken into account to consider the effect. Logarithmic scale of transmission coefficient is chosen to study the generated profiles.

II. MATHEMATICAL MODELING

Motion of a single electron in one dimension can be computed by using Schrödinger's time-independent equation:

$$-\frac{\hbar^2}{2m^*} \frac{d^2}{dz^2} \psi(z) + V(z)\psi(z) = E(z)\psi(z) \quad (1)$$

Incorporating the concept of effective mass mismatch, i.e. spatial variation of effective mass in Schrödinger's equation, we obtain

$$-\frac{\hbar^2}{2} \frac{\partial}{\partial z} \left[\frac{1}{m^*(z)} \frac{\partial}{\partial z} \psi(z) \right] + V(z)\psi(z) = E(z)\psi(z) \quad (2)$$

Solution of Schrödinger's equation (2) requires envelope function approximation that is both $\psi(z)$ and $(1/m^*)(\partial\psi(z)/\partial z)$ are continuous by considering electron transport across the heterojunction.

In the barrier and well regions, modified Schrödinger's equation's are-

$$-\frac{\hbar^2}{2} \frac{\partial}{\partial z} \left[\frac{1}{m_b^*(z)} \frac{\partial}{\partial z} \psi(z) \right] + V_b(z)\psi(z) = E(z)\psi(z) \quad (3)$$

and

$$-\frac{\hbar^2}{2} \frac{\partial}{\partial z} \left[\frac{1}{m_w^*(z)} \frac{\partial}{\partial z} \psi(z) \right] + V_w(z)\psi(z) = E(z)\psi(z) \quad (4)$$

where m_b^* & m_w^* are the effective masses of barrier and well regions, and V_b & V_w are potentials respectively. For the double quantum well triple barrier structure under consideration as shown in fig, wavevector for the problem may be defined as:

$$k_j = \left[\frac{2m^*(E - qV_j)}{\hbar^2} \right]^{0.5} \quad (5)$$

The wave functions in regions j and $j+1$ are

$$\psi_j = A_j \exp[ik_j x] + B_j \exp[-ik_j x] \quad (6.1)$$

$$\psi_{j+1} = C_{j+1} \exp[ik_{j+1} x] + D_j \exp[-ik_{j+1} x] \quad (6.2)$$

where A and C are coefficients for the wave function traveling left to right in regions j and $j+1$ respectively; B and D are the corresponding right-to-left traveling-wave coefficients.

We assume that BenDaniel Duke boundary conditions are satisfied in all the junctions.

Propagation between potential steps separated by distance L_j carries phase information only so that

$$\psi A_j \exp[ik_j L_j] = \psi C_j \quad (7.1)$$

and

$$\psi B_j \exp[-ik_j L_j] = \psi D_j \quad (7.2)$$

This can be formulated as-

$$\begin{bmatrix} A_j \\ B_j \end{bmatrix} = M \begin{bmatrix} C_{j+1} \\ D_{j+1} \end{bmatrix} \quad (8)$$

where

$$M = \begin{bmatrix} \exp[-ik_j L_j] & 0 \\ 0 & \exp[ik_j L_j] \end{bmatrix} \quad (9)$$

Thus, propagation matrix for j^{th} region can be written as:

$$P_j = \frac{1}{2} \begin{bmatrix} \left(1 + \frac{m_j}{m_{j+1}} \frac{k_{j+1}}{k_j}\right) \exp[-ik_j L_j] & \left(1 - \frac{m_j}{m_{j+1}} \frac{k_{j+1}}{k_j}\right) \exp[-ik_j L_j] \\ \left(1 - \frac{m_j}{m_{j+1}} \frac{k_{j+1}}{k_j}\right) \exp[ik_j L_j] & \left(1 + \frac{m_j}{m_{j+1}} \frac{k_{j+1}}{k_j}\right) \exp[ik_j L_j] \end{bmatrix} \quad (10)$$

E- κ dispersion relationship for barrier material consists the nonparabolicity factor-

$$E(\kappa) = \frac{\hbar^2}{2m^*} \kappa^2 (1 - \gamma \kappa^2) \quad (11)$$

where γ is the coefficient of conduction band nonparabolicity, and is a function of material composition and spin-orbit splitting factor, may be written as-

$$\gamma = \frac{\hbar^2}{2m^*} \left(1 - \frac{m^*}{m_0}\right)^2 \left(\frac{3+4y+2y^2}{3+5y+2y^2} \right) \left(\frac{1}{E_g} \right) \quad (12)$$

Here we consider only first higher-order term of the series. The wave functions are subjected to the variable effective mass condition at different materials where they are constant throughout the defined regions. For asymmetric potential structure, κ should have different values depending on the material composition.

A flow of propagation matrix in different region of the concerned structure provides the transmission matrix which can be defined as-

$$T(E) = \frac{1}{(P_{11})^2} \quad (13)$$

III. RESULTS & DISCUSSIONS

Numerical investigation is carried on the double quantum well triple barrier structure when well s have parabolic

geometry and is made of lower bandgap GaAs material, and barriers are made of higher bandgap $\text{Al}_x\text{Ga}_{1-x}\text{As}$ material. At each junction, effective mass mismatch is considered following BenDaniel Duke boundary condition and conduction band discontinuity is also taken into account. Band nonparabolicity plays a vital role following first order E-k relationship.

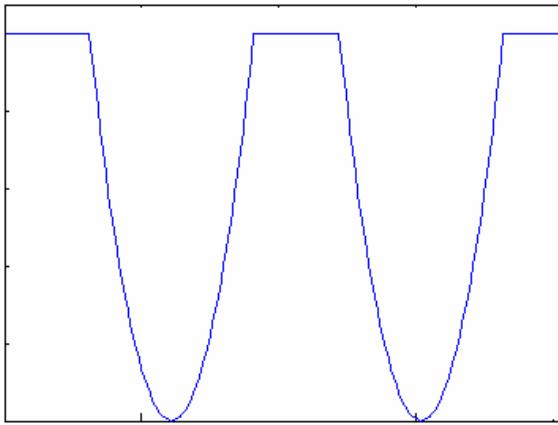


Fig 1: Schematic picture of double quantum well triple barrier structure having parabolic geometry with GaAs/Al_xGa_{1-x}As composition

Figure 2 shows the effect of well width variation on transmission coefficient profile in presence and absence of band nonparabolicity for specified material composition. It is observed that with increasing well width, transmission probability is enhanced. Also for specified well and barrier dimensions, band nonparabolicity increases the resonant tunneling probability which is verified by the origin of peak.

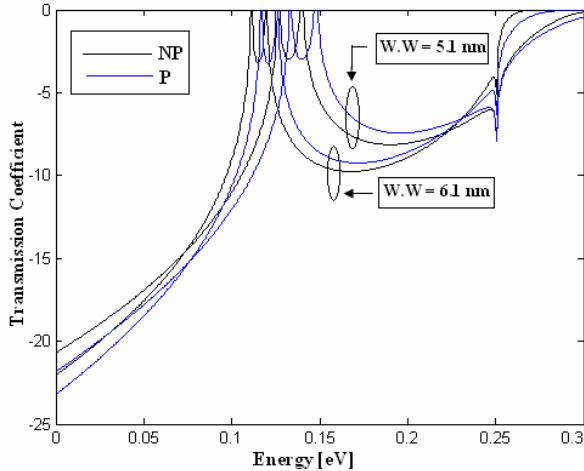


Figure 2: Comparative study of transmission coefficient profile with energy for different well width but specified barrier width and material composition with and without band nonparabolicity

Similarly, keeping well width constant and varying barrier width, it is seen that transmission coefficient decreases with increasing barrier width. Here also inclusion of band

nonparabolicity increases the transmission probability, as can be observed from fig 3.

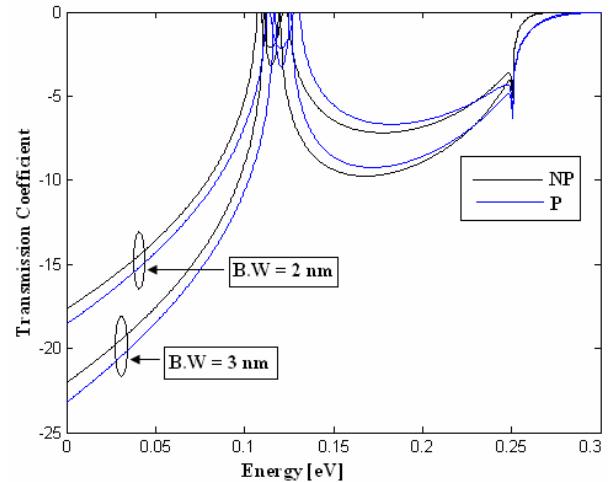


Figure 3: Comparative study of transmission coefficient profile with energy for different barrier width but specified well width and material composition with and without band nonparabolicity

If dimensional configuration remains unchanged and material composition of barrier layer is varied, then it may be observed that nonparabolicity effect becomes dominant for higher barrier width, i.e., for higher Al mole fraction. This is shown in fig 4.

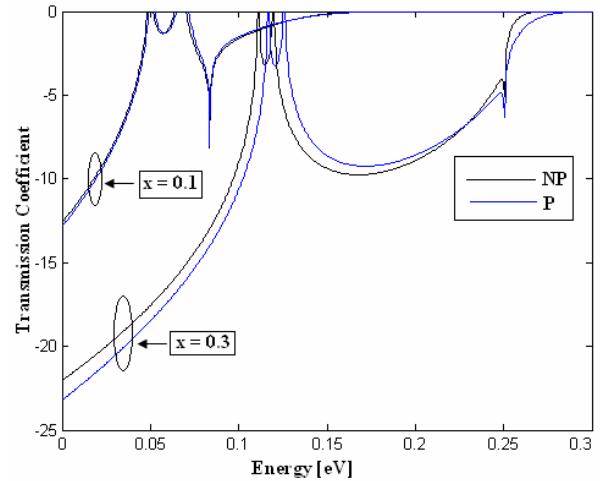


Figure 4: Comparative study of transmission coefficient profile with energy for different Al mole fraction but specified barrier and well widths with and without band nonparabolicity

Fig 5 shows that profile of tunneling probability for different barrier composition when variation is limited within type-I heterostructure. By keeping barrier thickness and well width constant, it can be observed that transmission coefficient reduces and existence of quasi-states are revealed from origin of secondary peaks. This is due to the fact that with increase of Al mole fraction, both potential barrier (i.e. conduction band

discontinuity) and effective mass mismatch at the junctions increase, and hence tunneling probability reduces.

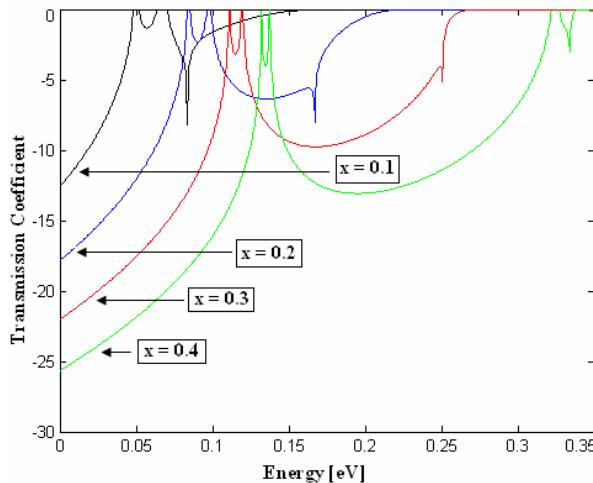


Figure 5: Transmission coefficient profile with energy for different Al mole fraction but specified barrier and well widths including band nonparabolicity

IV. CONCLUSION

Electronic and photonic properties of parabolic double well triple barrier structure can be predicted from knowledge of transmission coefficient which is dependent on dimensional configuration and material composition. This analysis is more critical and has better practical importance when parabolic geometry is considered compared to rectangular configuration as step potential is ideal and fabrication limitations in existing growth technology makes the well structure distorted. Band nonparabolicity is a critical parameter considering the precision of theoretical result, and hence should be considered when simulation is performed as it deals with more accurate dispersion relation including spin-orbit splitting factor. Comparative result shows that inclusion of Nonparabolicity-factor makes a shift of origin of eigenstates towards lower energy value. Al composition is varied within type-I heterostructure limitation so that structure can be made potentially different. This analysis can be further extended to analyze complex one-dimensional confined structures having arbitrary potential profiles.

REFERENCES

- [1] G. B. Morrison and D. T. Cassidy, "A Probability-Amplitude Transfer-Matrix Method for Calculating the Distribution of Light in Semiconductor Lasers", *IEEE Journal of Quantum Electronics*, Vol. 39, pp. 431-437, 2003.
- [2] D. Joel and M. R. Singh, "Resonant Tunneling in Photonic Double Quantum Well Heterostructures", *Nanoscale Research Letters*, Vol. 5, pp. 484-488, 2010.
- [3] S. Adachi, "GaAs, AlAs and $\text{Al}_x\text{Ga}_{1-x}\text{As}$: Material Parameters for use in Research and Device Applications", *Journal of Applied Physics*, vol. 58, R1, 1985.
- [4] W. P. Yuen, "Exact Analytic Analysis of Finite Parabolic Quantum Wells With and Without a Static Electric Field", *Physical Review B*, vol. 48, pp. 17316-17320, 1993.
- [5] R. C. Miller, D. A. Kleinman and A. C. Gossard, "Energy-Gap Discontinuities and Effective Masses for GaAs-Al_xGa_{1-x}As Quantum Wells", *Physical Review B*, Vol. 29, pp. 7085-7087, 1984.
- [6] T. Hiroshima and R. Lang, "Effect of Conduction Band Nonparabolicity on Quantized Energy Levels of a Quantum Well", *Applied Physics Letters*, Vol. 49, pp. 456-457, 1986.
- [7] D. P. Nelson, R. C. Miller and D. A. Kleinman, "Band Nonparabolicity Effects in Semiconductor Quantum Wells", *Physical Review B*, Vol. 35, pp. 7770-7773, 1987.
- [8] D. P. Dave, "Numerical Technique to Calculate Eigenenergies and Eigenstates of Quantum Wells with Arbitrary Potential Profiles", *IET Electronics Letters*, Vol. 27, pp. 1735-1737, 1991.
- [9] H. Cruz, "Resonant Tunneling through Parabolic Quantum Wells Achieved by Means of Short Period Superlattices", *Solid State Communications*, Vol. 85, pp. 65-68, 1993.
- [10] S. H. Pan, "Interband Optical Transitions of a Parabolic Quantum Well in the Presence of an Applied Midinfrared Field", *Physical Review B*, Vol. 48, pp. 2292-2297, 1993.
- [11] S. L. Chuang and D. Ahn, "Optical Transitions in a Parabolic Quantum Well With an Applied Electric Field -Analytical Solutions", *Journal of Applied Physics*, Vol. 65, pp. 2822-2826, 1989.
- [12] R. P. G. Karunasiri and K. L. Wang, "Infrared Absorption in Parabolic Multiquantum Well Structures", *Superlattices and Microstructures*, Vol. 4, pp. 661-664, 1988.
- [13] T. Yamaguchi, M. Kato and K. Tada, "A New Variational Method for Calculating Eigenenergy in Various Quantum Wells on the Basis of Effective Well Width", *IEEE LEOS Conference Proceedings*, Vol. 2, pp. 293-294, 1994.
- [14] D. L. Mathine, S. K. Myjak and G. N. Maracas, "A Computational Fourier Series Solution of the BenDaniel-Duke Hamiltonian for Arbitrary Shaped Quantum Wells", *IEEE Journal of Quantum Electronics*, Vol. 31, pp. 1216-1222, 1995.
- [15] J. G. S. Demers and R. Maciejko, "Propagation Matrix Formalism and Efficient Linear Potential Solution to Schrödinger's Equation", *Journal of Applied Physics*, Vol. 90, pp. 6120-6129, 2001.
- [16] T. P. Horikis, "Eigenstate Calculation of Arbitrary Quantum Structures", *Physics Letters A*, Vol. 359, pp. 345-348, 2006.
- [17] D J BenDaniel and C B Duke, "Space-Charge Effects on Electron Tunneling", *Physical Review*, Vol. 152, pp. 683-692, 1996.

AN EFFICIENT NLP BASED APPROACH TO PERSONALIZED INFORMATION RETRIEVAL IN WEB MINING

Sourav Saha

**Compure Science & Engineering,
Institute of
Engineering & Management
Kolkata, India
sourav.saha1977@gmail.com**

Subhabrata Sengupta

**Compure Science & Engineering,
Institute of
Engineering & Management
Kolkata, India
ssg365@gmail.com**

Sukalyan Goswami

**Compure Science & Engineering,
Institute of
Engineering & Management
Kolkata, India
sukalyan.goswami@gmail.com**

Abstract

In the field of Information retrieval, the main objective is to find meaningful as well as most relevant information with respect to some queries. But, the main problem regarding retrieval has always been, that the search area is so vast that it has become very difficult to retrieve pertinent information efficiently. It has been observed that, the generic ontological organizational information causes unnecessary extra CPU cost, while the user queries mostly target a specific domain. In addition another challenging problem in this regard is the key-phrase extraction from the query which has also an important role for pin pointing searching to a specific retrieval-domain. By focusing, these limitations and challenges, we have targeted our information retrieval system particularly towards Academic Institutional Domain in order to cater to the demand of various institution related queries. The institutional information has been organized as per ontological relationship among various categories and a natural language parser has been used during key phrase extraction for efficient retrieval of web links most preferably ordered with respect to the degree of relevance to the queries.

Keywords

Semantic web, Search Engine, Personalization, NLP, Web-Link Categorization, Parse Tree

1. INTRODUCTION

The World Wide Web serves as a huge, widely distributed, global information service center for news, advertisements, consumer information, financial management, education, government, e-commerce and many other information services. With the explosive growth of information available on the World Wide Web, various intelligent web services have been developed to help user's access relevant information from the Web. Web search has its root in **information retrieval** (or IR for short) [1]. Traditional IR assumes that the basic information

unit is a document, and a large collection of documents is available to form the text database. Retrieving information simply means finding a set of documents that is relevant to the user query. A ranking of the set of documents is usually also performed according to their relevance scores to the query [2]. The most commonly used query format is a list of keywords, which are also called "**Terms**". IR is different from data retrieval in databases using SQL queries because the data in databases are highly structured and stored in relational tables, while information in text is unstructured [4]. Web personalization is a strategy, a marketing tool, and an art. Personalization requires implicitly or explicitly collecting visitor information and leveraging that knowledge in your content delivery framework to manipulate what information you present to your users and how you present it [6]. Web mining aims to discover useful information or knowledge from the Web hyperlink structure, page content, and usage data. Although Web mining uses many data mining techniques, as mentioned above it is not purely an application of traditional data mining due to the heterogeneity and semi-structured or unstructured nature of the Web data.

2. PROBLEM DESCRIPTION

Information retrieval (IR) is the study of helping users to find information that matches their information needs. Technically, IR studies the acquisition, organization, storage, retrieval, and distribution of information. Historically, IR is about document retrieval, emphasizing document as the basic unit. In the **figure 1**, the user with information need issues a query (user query) to the retrieval system through the query operations module. The retrieval module uses the document index to retrieve those documents that contain some query terms (such documents are likely to be relevant to the query), compute relevance scores for them, and then rank the retrieved documents according to the scores. The ranked documents are then presented to the user. The document collection is also

called the text database, which is indexed by the indexer for efficient retrieval.

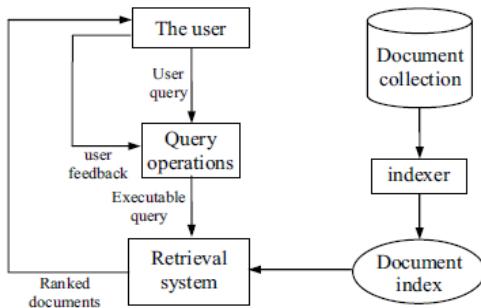


Figure 1: A general IR system architecture

The current World Wide Web presents a large amount of information in a format tailored for viewing and understanding by humans. People can surf from link to link, query information by search engines, or attempt to access websites by domain name. As web pages retrieved are only meaningful to humans, to the software that processes the contents, they are no more than a string of random characters. Software programs cannot just load a random document, web page, or file and understand the contents of that document. While the software could make assumptions based on **HTML (Hyper Text Markup Language)** or **XML (Extensible Markup Language)** tags, a programmer would have to get involved and determine the meaning, or semantics, of each tag. From a computer's perspective, the World Wide Web is a confusable, unstructured and less machine-readable mess. As such, a method for representing knowledge, such that software programs can understand, share and exchange knowledge, is needed.

3. PROPOSED SYSTEM

The following section describes the design of the proposed architecture. The schematic diagram of the architecture is presented in **figure 2**. In general, the web links in a website are associated with contextual labels. During web site registration process, the web sites will be parsed by the HTML link parser to retrieve the web links along with their contextual labels. Meaningful key phrases will be extracted from the contextual labels and subsequently the key-phrases will be categorized as per the relevance [8]. The key phrases, categories are organized in an xml based data repository maintaining the ontological structure of the information pieces. During information retrieval process, the user feeds a specific query with respect to the intended domain. A parse tree will be generated from the given query by the NLP parser [5]. The parse tree yields an expression generated via DFS (Depth First Search) traversal of semantic tree corresponding to the query. The DFS expression represents POS (Parts of

Speech) tagged phrases organized in a tree as per their dependencies with each other. The key phrase extraction module extracts key phrases by parsing the POS tagged expression in an order depicting the degree of relevance with the query. Subsequently, the extracted key phrases are provided to the key phrase categorization module to categorize each phrase to a priority category. Each pair of key phrase, category is used by the proposed retrieval system to retrieve the web links ordered with respect to their relevance to the query [7]. The retrieval module consults the xml based web link repository to match the key phrase-category pair with the stored key phrases. As mentioned earlier, each stored key phrase is also associated with relevant web links, which aids the retrieval module to list the relevant web links.

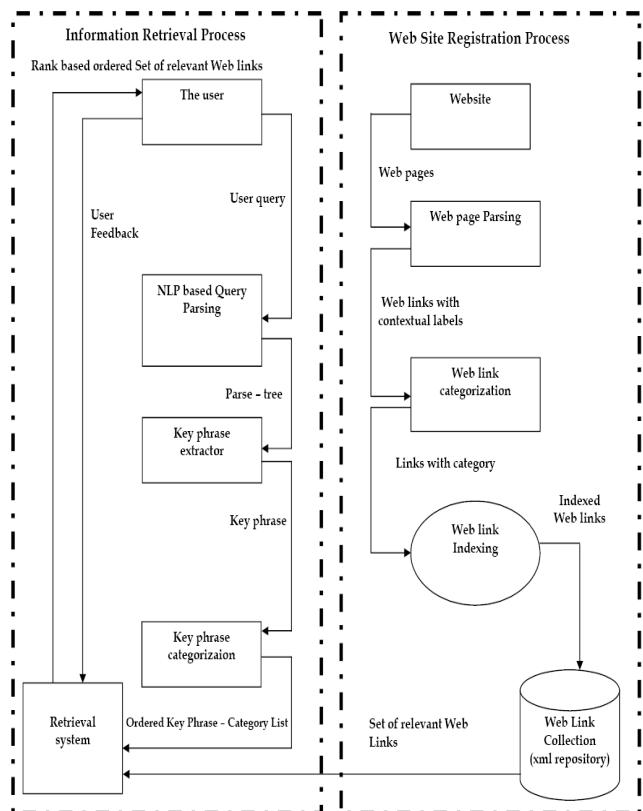


Figure 2: Architecture of the Proposed System

4. ALGORITHM FOR IMPLEMENTATION OF PROPOSED SYSTEM

URL Registration Process

Step 1: Input the URL.

Step 2: For every single web page belonging to the specified website, do the following:

Step 2.1: Tokenize every elements of the webpage.

Step 2.2: For every token check, whether it refers to link or not.

Step 2.3: If the token refers to a link, extract the Link and the label, associated with the link. Otherwise, the token will be ignored. Store the valid links and its associated.

Step 3: For each stored link at **Step 2**, do the following:

Step 3.1: Classify the associated label to one of the domain category with help of a classifier, which has been previously trained with selective categories of the information retrieval domain. The training of the classifier has been carried out as follows:

Step 3.1.1: Identify probable categories belonging to the search domain.

Step 3.1.2: Collect as many as possible key phrases for every category.

Step 3.1.3: Train the classifier with the categorized collection of key phrases.

Step 3.2: Populate the database, with link, and its associated category.

Information Retrieval Process

Step 1: Input the user query.

Step 2: The query string will be parsed by the NLP (Natural Language Processing) parser.

Step 3: Traverse the parse tree generated at Step 2. For each noun phrase, do the following steps:

Step 3.1: For each node check, whether it is noun Phrase, and has not already been explored, then, do the following steps:

Step 3.1.1: Find the relative position of the noun Phrase in the query.

Step 3.1.2: Assign a degree of importance to the Noun phrase based on its depth of its node in the parse tree.

Step 3.1.3: Store each noun phrase along with its degree of importance.

Step 4: Order the list of noun phrases with respect to the degree of importance.

Step 5: For every key noun phrase, stored in the ordered list, do the following:

Step 5.1: Preprocess the noun phrase to trim it to the standard key phrase. Classify the noun phrase, to one of the domain – category with help of a classifier, which has been previously trained with elective categories of the information retrieval domain. Store the key-phrase along with its category.

Step 6: For every successive pair of key-phrase and the corresponding category do the following

Step 6.1: Look up the xml having web-resources organized in such a way that depicts the relationship among categories.

Step 6.2: Find the closest match of the noun-phrases with the key- phrases stored in the xml.

Step 6.3: Retrieve the links associated with the matched key-phrases. Store the link in hash-map and for each collision in the hash map against the link increment the degree of relevance associated with the link.

Step 7: Present the links ordered in accordance with their associated degree of relevance.

5. ILLUSTRATION OF THE PROPOSED ALGORITHM

Search String: “Where can I get the syllabus for mtech in CSE?”

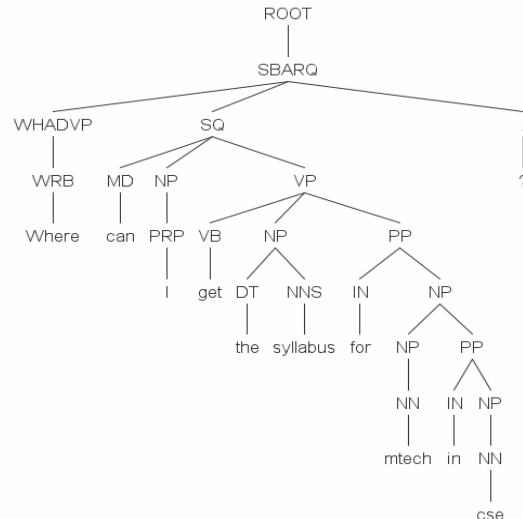


Figure 3: NLP based Parse Tree

DFS(Depth First Search) traversal expression of parse tree:

Parse Tree: (ROOT (SBARQ (WHADVP (WRB Where)) (SQ (MD can) (NP (PRP I)) (VP (VB get) (NP (DT the) (NNS syllabus)) (PP (IN for) (NP (NP (NN mtech)) (PP (IN in) (NP (NN cse))))))) (. ?)))

5.1 Key Phrase Extraction:

KeyPhrase: syllabus
KeyPhrase: mtech
KeyPhrase: cse

5.2 Key Phrase Categorization:

| | |
|---------------------|--------------------|
| KeyPhrase: syllabus | Category: syllabus |
| KeyPhrase: mtech | Category: course |
| KeyPhrase: cse | Category: subject |

Web Link Retrieval:

| Url | Link Description |
|------|---|
| url1 | Link to btech cse syllabus |
| url2 | Link to btech information technology syllabus |
| url3 | Link to mtech cse syllabus |
| url4 | Link to mtech information technology syllabus |

Table 1: A Link Description Table

KeyPhrase: syllabus Category: syllabus

Retrieved Links: url1, url2, url3, url4

KeyPhrase: mtech Category: course

Retrieved Links: url3, url4

KeyPhrase: cse Category: subject

Retrieved Links: url3

5.3 Final list of web links in order of relevance:

1. url3: Link to mtech cse syllabus

Degree of relevance = frequency of occurrence = 3

2. url4: Link to mtech information technology syllabus

Degree of relevance = frequency of occurrence = 2

3. url1: Link to btech cse syllabus

Degree of relevance = frequency of occurrence = 1

4. url2: Link to btech information technology syllabus

Degree of relevance = frequency of occurrence = 1

6. EXPERIMENTATION AND RESULTS

This section describes the experimental details as well as the result of the proposed information retrieval scheme. The experimentation has been carried out on a particular domain to retrieve relevant web links based on user given query. As detailed in the previous section, the proposed scheme attempts to extract the meaningful key-phrases with the help of natural language based parsing. As per our observations, the extracted key phrases closely match with human cognition system leading to efficient retrieval of relevant web links from the xml based database. The proposed scheme also orders the web links in accordance with their degree of relevance to the query. This implies the higher the relevance of a link, the higher will be its possibility being at the top of the list.

6.1 Experimental Setup

The experimentation has been carried out on mining various academic institutional websites. The primary focus of the scheme is to collect various web links on academic courses and categorize them as per subjects. The xml as shown in the figure 4 presents the ontological organization of key phrases(fractional view) along with relevant web links. Every entity in the xml belongs to a category among “syllabus” “course” “subject”. The relevant web links with respect to an entity plays key role in retrieving the web links.

The following table shows the query set presented to our retrieval system. All the queries are related to the retrieval of syllabus oriented web links.

```

<db>
  <entity>
    <id>1</id>
    <key-phrase>syllabus</key-phrase>
    <category>syllabus</category>
    <link>1, 2, 3, 4</link>
    <entity>
      <id>1.1</id>
      <key-phrase>btech,
undergraduate</key-phrase>
      <category>course</category>
      <link>1, 2</link>
      <entity>
        <id>1.1.1</id>
        <key-phrase>cse computer
science</key-phrase>
      <category>subject</category>
      <link>1</link>
    </entity>
    <entity>
      <id>1.1.2</id>
      <key-phrase>it information
technology</key-phrase>
    <category>subject</category>
    <link>2</link>
  </entity>
</entity>

```

Figure 4: xml database

6.2 Performance Analysis

The performance metric of the retrieval system, has been based on the divergence of the retrieved web links as compared to the order of ground truth(Human Cognition Interpretation) interpreted web links list. The more the order of the list gets diverted from the ground truth order, the less will be the performance score of the retrieval system. This implies the performance score of the system is directly proportional to the degree of relevance of the web link order. Equation 2 has been used as the merit score of the proposed system. The deviation of a particular web link order with respect to the ground truth is multiplied by a weight factor to reflect the degree of relevance of the link with its query.

$$\text{Performance Score} = 1 - (\text{Divergence} / \text{Maximum Divergence})$$

$$\text{Divergence} = \sum_{i=1}^{N(\text{No. of links})} W_i * \text{abs}(\text{Link Order}_i \text{ in human interpretation list} - \text{Link Order}_i \text{ in Proposed Algorithm List})$$

Where, $W_i = (\text{No. of links} - i + 1)$

$$\text{Maximum Divergence} = \sum W_i * (\text{No. of links} - i)$$

| Query No. | Query |
|-----------|--|
| 1 | Where can I get the syllabus for mtech in cse? |
| 2 | Where can I get mtech cse syllabus? |
| 3 | Show the cse syllabus. |
| 4 | How can I find the undergraduate syllabus in information-technology? |
| 5. | I want to download the syllabus for cse in mtech? |

Table 2: Query lists

The web links and its urls with reference to the course related syllabus are listed in **Table 3**. The performance score of the proposed scheme with regards to the queries listed in **Table 4**.

| URL | Link Description |
|------|---|
| url1 | Link to btech cse syllabus |
| url2 | Link to btech information technology syllabus |
| url3 | Link to mtech cse syllabus |
| url4 | Link to mtech information technology syllabus |

Table 3: URL and corresponding web links

| Query No. | Ordered Web Links by our proposed algorithm | Ordered Web Links by human interpretation | Performance Score |
|------------------|--|--|--------------------------|
| 1. | Link: url3 Link: url1 Link: url4 Link: url2 | Link: url3 Link: url4 Link: url1 Link: url2 | 75% |
| 2. | Link: url3 Link: url4 Link: url1 Link: url2 | Link: url3 Link: url4 Link: url1 Link: url2 | 100% |

Table 4: URL and corresponding web links

7. CONCLUSION

Proposed system is designed and developed with the intentions to build an intelligent information retrieval system. In this proposal, we have extracted all the links from a designated domain (Academic Institutions). All the links are mapped with their suitable categories and indexed in an xml based repository. The query string gets parsed by a natural language based parser to yield a semantic parse tree. The semantic interdependence among the noun phrases is exploited by our key-phrase extraction module to produce most relevant key-phrases. We have implemented key-phrase categorization module with the help of vector space classifier. The extracted key-phrases are categorized via the vector space classifier to limit the searching to the domains specified by the categories. The effective searching by our retrieval system attempts to yield the relevant web links ordered in accordance with their degree of relevance to the query.

8. FUTURE SCOPES

It has been observed, during the experimentation that the link levels are not self explanatory to the satisfactory extent. Therefore, the web site registration module needs to be improved by considering the contextual information around the web links in addition to the link levels. So far, we have utilized a Natural Language Parser developed by Stanford Group, which has been found to be constrained to some extent with respect to the search query in our retrieval system.

9. REFERENCES

- [1] A. M. Riad. et.al. PSSE: An Architecture for a Personalized Semantic Search Engine. International Journal on Advances in Information Sciences and Service Sciences Volume 2, Number 1, March 2010, pp.102 – 112
- [2] Myungjin Lee. et.al. Semantic Association-Based Search and Visualization Method on the Semantic Web Portal. International Journal of Computer Networks & Communications (IJCNC), Vol.2, No.1, January 2010, pp.140 – 152
- [3] Ulf Brefeld. et.al. Document Assignment in Multi-site Search Engines. WSDM'11, ACM 978-1-4503-0493-1/11/02, February 2011, pp.575 – 584
- [4] Fabrizio Lamberti. A Relation-Based Page Rank Algorithm for Semantic Web Search Engines. IEEE Transactions on Knowledge and Data Engineering, Volume 21, Number 1, January 2009, pp. 123 – 136
- [5] Nguyen Tuan Dang, and Do Thi Thanh Tuyen. Natural Language Question Answering Model Applied To Document Retrieval System. World Academy of Science, Engineering and Technology 51 2009, pp.36 – 39
- [6] A.Jebaraj Ratnakumar. An Implementation of Web Personalization Using Web Mining Techniques. Journal of Theoretical and Applied Information Technology, 2005, pp. 68 – 73
- [7] Gauri Rao, Chanchal Agarwal, Snehal Chaudhry, Nikita Kulkarni, Dr. S.H. Patil. Natural Language Query Processing Using Semantic Grammar (IJCSE) International Journal on Computer Science and Engineering Volume 02, Number 02, 2010, pp.219-223.

AUTOMATIC SCHEMA MATCHING AND MAPPING APPROACH

Wordnet & Classification based schema matching Approach

Soumi Dutta

Department of Computer Science &
Engineering
Institute of Engineering and Management
Kolkata, India
soumi.it@gmail.com

Deepshikha Mullick

Department of Information Technology
Institute of Engineering and Management
Kolkata, India
deepshi.m@gmail.com

Moumita Roy

Department of Computer Application and
M.Sc.
Institute of Engineering and Management
Kolkata, India
moumita.majumder@gmail.com

Abstract — A common task in many database applications is the migration of data from multiple sources (.xml, .sql, .csv file formats) into a new one (relational dbms format). This requires identifying semantically related elements of the source and target formats and the creation of mapping expressions to transform instances/data of those elements from the source format to the target format. Currently, Data Migration is typically done manually, a tedious and time-consuming process, which is difficult to scale to a large number of data sources. In this paper, a new semi-automatic approach to determining semantic correspondences between schema elements for data migration applications is described. Semantic schema matching is based on the two ideas:

(i) Structured matching (ii) Unstructured matching (conceptual form). The project advances the state of the art with a set of new techniques exploiting sample instances, ontology, classification and reuse of existing mappings to detect not only element correspondences but also their mapping expressions.

Schema matching is one of the key challenges in data migration as well as information integration. To ease the problem, many automated solutions have been proposed. Most of the existing solutions mainly rely upon textual similarity of the data to be matched. However, there exist instances of the schema-matching problem for which they do not even apply.

One of the common challenges to automating matching and mapping has been previously classified especially for relational DB schemas recognizing schematic vs. semantic differences/heterogeneity.

Keywords: Schema matching, Schema mapping, Data Migration, Data Integration, Data warehouse, Ontology, Semantic Web.

I. INTRODUCTION

In most schema integration or data migration systems schema matching is a fundamental problem in many applications, such as integration of semantic web-oriented data, e-commerce, schema evolution and migration, application evolution, data warehousing, database design, web site creation and management, and component-based development. Here a new approach is introduced to replace existing schema matching and mapping methodologies. During the migration process, data needs to be extracted from the source format, transformed and loaded into the target format. This process requires solving two difficult tasks:

a) Schema matching to identify schematic or semantically related elements between the source and target systems in a structured (similar) or unstructured (conceptual – synonym, abbreviation, ontological, classification) way.

b) Schema mapping discovery to determine mapping elements useful for transforming data from the source format to the target format.

Obviously, manually specifying schema matches is a time consuming, error-prone, and therefore expensive process. None of the schema matching methods reviewed in this paper have yet reached the stage of being completely automatic, and human intervention is still required. In fact, some researchers (for example, Yan, 2001, Ram 2004, Aumueller 2005, and Halevy 2005) do not foresee fully automatic schema matching as a possibility, and orient their research towards assisting human-performed schema matching. Schema matching has been identified as a problem that is “AI complete”, meaning that it is as difficult to replicate as human intelligence (Bernstein 2004).

Moreover, there is a linear relation between the level of effort and the number of matches to be performed, a growing problem due to the rapidly increasing number of web data sources and warehouse data to integrate. A faster and less effort-intensive integration approach is needed. This requires automated support for schema matching and this is the aim of this work.

II. LITERATURE REVIEW

A. Schema:

Schema is the structure of a database system, described in a formal language supported by the database management system (DBMS). A schema is defined in text database language; the term is often used to refer to a graphical depiction of the database structure. In a relational database, the schema defines the tables, the fields in each table, and the relationships between fields and tables.

A schema is a collection of schema objects. Examples of schema objects include tables, views, sequences, synonyms, indexes, clusters, database links, snapshots, procedures, functions, and packages. Schema can represent in two ways:

The logical schema was the way data were represented to conform to the constraints of a particular approach to database management. At that time the choices were hierarchical and network.

Describing the logical schema, however, still did not describe how physically data would be stored on disk drives. That is the domain of the *physical schema*.

B. Schema Matching & Mapping:

Schema matching is the process of identifying that two objects are semantically related. *Match* is a schema manipulation operation that takes two schemas as input and returns a mapping that identifies corresponding elements in the two schemas. Schema matching is a critical step in many applications: in E-business, to help map messages between different XML formats; in data warehouses, to map data sources into warehouse schemas; and in mediators, to identify points of integration between heterogeneous databases. Schema matching is primarily studied as a piece of these other applications. For example, schema integration uses matching to find similar structures in heterogeneous schemas, which are then used as integration points. Data translation uses matching to find simple data transformations. Given the importance of XML message mapping, we expect to see match solutions to appear next in this context.

Schema matching is challenging for many reasons. Most importantly, even schemas for identical concepts may have structural and naming differences. Schemas may model similar but non-identical content. They may be expressed in different data models. They may use similar words to have different meanings.

C. Example :

These examples provide some illustrations of problems encountered when attempting to match schemas. In Figure 1 below, for example, most methods would have little difficulty determining Zip from schema A and ZipCode from schema B as matches. This is an example of a match that could have been determined by examining just the element names. It is also an example of a direct match, sometimes also termed a match of

1:1 cardinality, meaning that a single element from one schema is matched to a single element in another schema.

Schema A

| First | Last | Address | Zip | Phone |
|--------|-----------|------------------|------|------------|
| Carol | Frenditte | Washington , MA | 2051 | 6174432974 |
| Allen | LeBlanc | Union St., MA | 2053 | 6174432643 |
| Thomas | Gutierrez | Chestnut St., RI | 2054 | 6174453925 |

Schema B

| Name | Address | City | State | Zip |
|------------|------------|------------|-------|------|
| William J | Union St. | Easton | MA | 2053 |
| Robert Ian | Elm St. | Washington | RI | 2793 |
| Nancy Lan | Dodley St. | Providence | MA | 2045 |

Figure 1. Simple schema matching based on element name

Much research has been directed at developing direct matches. In the example above, the address may require an indirect match, sometimes termed a match of cardinality 1: n, to match Address in schema A with Address, City, and State in schema B. Also, the first and last name elements from schema A would need to be concatenated to match the name element in schema B. Even so, one schema might have additional information not in the other, such as schema B which includes prefix and suffix as part of the name element. In this case, only a partial match would be possible between the two schemas. This can also be seen with the element Phone in schema A, which has no corresponding match in schema B.

Moreover, some matches may complex matches, also termed matches of m:n cardinality, where multiple elements in one schema must be matched to multiple elements in the other. In the example above, if schema B represented name as prefix, first, middle initial, last, and suffix, then such a match would be required.

Considering matches of m:n cardinality increases the complexity of schema matching exponentially. Further, because data may need to be transformed in addition to only being matched through schema elements, the complexity of schema matching in general can be considered to be unbounded (Doan 2005). Consider the example in Figure 2 in which each schema has:

Schema C

| Item | Qty | Price | Taxes |
|------|-----|----------|--------|
| 1405 | 5 | \$110.00 | \$6.00 |
| 1982 | 3 | \$45.00 | \$2.25 |
| 2023 | 1 | \$18.00 | \$9.00 |

Schema D

| Item | Quantity | Price | Total |
|-------|----------|---------|---------|
| A110C | 2 | \$11.00 | \$22.00 |
| AV99x | 4 | \$9.00 | \$36.00 |
| AL129 | 5 | \$18.00 | \$90.00 |

Figure 2. An example of complex schema matching

2nd Annual International Conference IEMCON 2012

An element named Price, yet they are not semantically equivalent. In order to match these two schemas, a matcher would have to not only discover this semantic difference, but evaluate Price and also Price + Taxes in schema C as matches for Total in schema D.

This example also shows how instance data adds information to the schema matching process, and how approaches based only on examining the schema itself may not be adequate.

This example also illustrates how data types might be used to improve the accuracy of matches. In this case, both Qty and Quantity have the same data type (integer), and this piece of schema information might be used to augment other information in ascertaining matches.

The examples above are considered to be element-based schema matching, in that they determine matches without knowledge of the database structure. However, in many cases a consideration of database structure should improve schema matching.

Schema E:

| Badge | Name | DeptID | HomePhone |
|-------|-----------------|--------|------------|
| 41723 | Katherine Baker | 172 | 5082307682 |
| 56784 | Mark Bharati | 189 | 6179641242 |
| 66010 | Edward Waters | 189 | 9788917692 |

| DeptID | DeptName |
|--------|------------|
| 172 | Purchasing |
| 189 | Marketing |

Schema F:

| EmployeeID | Name | Department |
|------------|-----------------|------------|
| 13572 | Kevin Li | Accounting |
| 20473 | Julie McCormack | HR |
| 33717 | Fran Liebowitz | HR |

| EmployeeID | PhoneType | PhoneNumber |
|------------|-----------|-------------|
| 13572 | Home | 4015629982 |
| 13572 | Cell | 4018849010 |
| 13572 | Emergency | 4012399497 |
| 20473 | Home | 5083431884 |
| 20473 | Cell | 5087898386 |
| 33717 | Home | 6172847702 |

Figure 3. Element Matching

The example in Figure 3 shows how the appropriate elements to match may need to be determined through an examination of each schema structure. The department names to be matched would need to come from the department entity in schema E and the employee entity in schema F. Consideration of the database structure would help enable the match of DeptName rather than DeptID from schema E since DeptName is presumably the same data type as Department in schema F, whereas DeptID would not be.

III. PROBLEM STATEMENT

Many solutions are proposed for the Relational Schema matching problem but there is still no complete automatic schema matching solution. In almost all cases, schema matching requires manual effort explicitly. One of the main problems in schema matching is that the existing solutions mostly depend on web vocabulary, mainly WordNet dictionary. In a case where WordNet can't help, the matching fails. Also, WordNet database is so huge that we get large number of synonyms/sense/phrase for a single element. From that subset, identifying the actual match is time consuming and increases complexity. For examples, the synonym/sense/phrase of an element could be indefinite and WordNet could not recognize it. Also, abbreviations are not available in WordNet, which needs to be given manually. In these cases data type matching and structural matching are not effective, because the syntactic matching is the leading part of matching the semantic meaning of two elements. So, syntactic or structural matching is useless here, since a manual intervention is required.

IV. SOLUTION

A. Schema Compatibility Architecture:

Let us assume that the source and target schema are following the same data model or architecture. The most preferred architecture is RDBMS as it is structured and helps faster searching methods using indexing. Thus, other probable formats need to be converted in RDBMS format. For CSV files, equivalent database tables need to be created. For XML formats, a parse tree needs to be created along with the database table. Example:

```
<BookStore>
    <Book>
        <Name> Let us C </Name>
        <Author> Kanetkar </Author>
        <Subject> Computer Science </Subject>
        <Price> 300 </Price>
    </Book>
    <Book>
        <Name> Pointers in C </Name>
        <Author> Kanetkar </Author>
        <Subject> Computer Science </Subject>
        <Price> 250 </Price>
    </Book>
</BookStore>
```

Figure 4. XML File

| Name | Author | Subject | Price |
|---------------|----------|------------------|-------|
| Let us C | Kanetkar | Computer Science | 300 |
| Pointers in C | Kanetkar | Computer Science | 250 |

Figure 5. RDBMS Table

B. Explicit Inputs:

1. Abbreviations are not available in WordNet dictionary. If any abbreviation is being used repetitively in place of an element, user can explicitly mention the abbreviations and the corresponding elements to get perfect match output, which will be helpful for further data integration or migration.
2. WordNet dictionary contains set of synonyms along with sense number for an element. Synonyms of elements can be used from WordNet dictionary for matching of source with target system.
3. User can also manually specify synonyms where it is known to avoid the use of WordNet dictionary and thus pacify the process.
4. Classification of abbreviation and synonym can be a probable alternative of manual data entry.
 - a) To create an abbreviation class, at least one schema matching is required. After manual entry of abbreviation and matching of elements of source and target system, matched abbreviations will automatically migrate to an abbreviation class. In future, this abbreviation class can be used for element matching prior to user input.
 - b) Similar to abbreviation class, synonym class also requires at least one schema matching. Synonym class will be created automatically migrating manual and/or WordNet synonym after matching of elements of source and target system.
5. After a match is found, it is easier to find the mapping expression. Using the found mapping expression further data can be transformed into a single schema from source and target schema. This is called data integration.

There can be cases where multiple elements together can be mapped to a single element. Referring to Figure 1, “Address” of schema A can be mapped to “Address, City, State” of schema B. In this case, along with the mapping expression, specific integration protocol is needed to get the proper integrated data as output. There can be other integration protocols used for different data integration cases along with the mapping expression, like splitting of Name into first name and last name (Ref. Figure 1). While implementation, these integration protocols need to be predefined in the system.

V. EXPERIMENTAL RESULT

For Schema Matching, initially string matching is required to match source and target schema. As example, we have considered two different schemas, Emp as source schema and

Employee as target schema. After string matching, we will get the result as shown below.

Figure.6 this is the result using exact and partial String Matching.

| Emp | Employee |
|---------------|------------------|
| 1. Emp_ID | 4. ID |
| 2. First_Name | 2. Name |
| 3. Last_Name | 3. Post |
| 4. Position | 4. Dept |
| 5. Department | 5. Gender |
| 6. Salary | 6. Phone |
| 7. Gender | 7. Country |
| 8. Address | 8. State |
| 9. Blood_grp | 9. City |
| 10. Ph | 10. Street |
| | 11. Pin |
| | 12. Remuneration |

Match Found: 2

So, we can see two string matching results are found in the example. To get better accuracy, mostly dictionaries are used. Here we are considering WordNet as it is one which is mostly used. Let us take an example - we take an attribute ‘Salary’ from source schema **Emp** and find it’s synonym using WordNet dictionary. The result set will have wage, pay, earnings, payroll, paysheet, regular payment, remuneration. Result set contains 7 synonyms out of which the exact match is ‘remuneration’. To find this exact match each synonym from the result set will be matched against all the attributes in the target schema ‘Employee’. So, the number of iteration, along with the time complexity, may be high for a single match.

Figure.7 This is the result using WordNet for Synonym Based Matching.

| Emp | Employee |
|---------------|------------------|
| 1. Emp_ID | 1. ID |
| 2. First_Name | 2. Name |
| 3. Last_Name | 3. Post |
| 4. Position | 4. Dept |
| 5. Department | 5. Gender |
| 6. Salary | 6. Phone |
| 7. Gender | 7. Country |
| 8. Address | 8. State |
| 9. Blood_grp | 9. City |
| 10. Ph | 10. Street |
| | 11. Pin |
| | 12. Remuneration |

Match Found: 2

2nd Annual International Conference IEMCON 2012

Figure 8 Final Mapping Expression: This is the final mapping expression after finding string matching and synonym matching from WordNet dictionary.

| Emp | Employee |
|---------------|------------------|
| 1. Emp_ID | 1. ID |
| 2. First_Name | 2. Name |
| 3. Last_Name | 3. Post |
| 4. Position | 4. Dept |
| 5. Department | 5. Gender |
| 6. Salary | 6. Phone |
| 7. Gender | 7. Country |
| 8. Address | 8. State |
| 9. Blood_grp | 9. City |
| 10. Ph | 10. Street |
| | 11. Pin |
| | 12. Remuneration |

Total Match Found: 4

Expected Results for proposed approach:

As previously discussed, using classification concept we have to create two separate classes for synonym and abbreviation. Abbreviation class is necessary as they are not predefined in any dictionary and for an organization, a set of abbreviations are mostly used for specific words, as ph for phone, mob for mobile, DOB for date of birth etc. The synonym class is created from the frequently used synonyms for each word in the schema.

Figure 9 & 10 Proposed Synonym and Abbreviation class of the previous example.

Synonym Class:

| ID | Word | Synonym |
|-----------|-------------|----------------|
| 1 | Post | Place |
| 2 | Post | Position |
| 3 | Salary | Remuneration |
| 4 | Designation | Post |
| 5 | Salary | Earning |
| 6 | Pay | Salary |

Abbreviation Class:

| ID | Word | Abbreviation |
|-----------|-------------|---------------------|
| 1 | First_Name | Name |
| 2 | Last_Name | Name |
| 3 | Department | Dept |
| 4 | Phone | Ph |
| 5 | Country | Address |
| 6 | State | Address |
| 7 | City | Address |
| 8 | Street | Address |
| 9 | Pin | Address |

In Figure.9 & 10 the result of exact and partial String Matching is shown which we have to consider first before considering the classes. After string matching, we will use synonym class for finding matches which will reduce the number of comparisons to find the match with respect to web dictionaries like WordNet. The result of synonym matching will be same as Fig. with reduced time complexity.

Next, we will use abbreviation class to find matching for the abbreviated words in source and target schema. Fig. is the result of the abbreviation based matching.

| Emp | Employee |
|---------------|------------------|
| 1. Emp_ID | 1. ID |
| 2. First_Name | 2. Name |
| 3. Last_Name | 3. Post |
| 4. Position | 4. Dept |
| 5. Department | 5. Gender |
| 6. Salary | 6. Phone |
| 7. Gender | 7. Country |
| 8. Address | 8. State |
| 9. Blood_grp | 9. City |
| 10. Ph | 10. Street |
| | 11. Pin |
| | 12. Remuneration |

Match Found: 9

Figure: 10 this is the final result considering string matching, abbreviation class matching and synonym class matching.

| Emp | Employee |
|---------------|------------------|
| 1. Emp_ID | 1. ID |
| 2. First_Name | 2. Name |
| 3. Last_Name | 3. Post |
| 4. Position | 4. Dept |
| 5. Department | 5. Gender |
| 6. Salary | 6. Phone |
| 7. Gender | 7. Country |
| 8. Address | 8. State |
| 9. Blood_grp | 9. City |
| 10. Ph | 10. Street |
| | 11. Pin |
| | 12. Remuneration |

Total Match Found: 13

We can conclude that using the proposed approach Schema Matching & Mapping Process can be enhanced and also reduce time complexity.

VI. CONCLUSION

Today schema matching is a basic task in almost every data intensive distributed application, namely enterprise information integration, collaborating web services, ontology based agents communication. Schema matching is a basic problem in much database application, such as heterogeneous database integration, E-commerce, data warehousing, data mining and semantic query processing. In this paper an automatic schema matching and mapping technique is proposed. For distributed database and data mining integration is necessary. So, Schema matching and mapping is required for integration purpose. So the challenge is to find the accurate and maximum semantic matches and to define the mapping expression.

In this paper we proposed a broad overview of the steps of schema matching and mapping, in the large scale schema integration and data mining. Beyond manual effort we can include the concept of WordNet dictionary, classification to find exact matches and mapping expression. Along with synonym matches abbreviation is also considered. In future we can also include the concept of cluster analysis for abbreviation as it is not predefined in any dictionary.

In the future, we would like to introduce quantitative work on the relative performance and accuracy of different approaches based on above mentioned algorithm. Such results could tell us which of the existing approaches dominate the others and could help identify weaknesses in the existing approaches that suggest opportunities for future research.

VII. REFERENCES

- [1] Alexe, B. M. Gubanov, M. A. Hernandez, H. Ho, J.-W. Huang, Y. Katsis, L. Popa, B. Saha, and I. Stanoi. Simplifying Information Integration: Object-Based Flowof-Mappings Framework for Integration. Proc. BIRTE, 108–121. Springer, 2009.
- [2] Algergawy, A., E. Schallehn, and G. Saake: Improving XML schema matching performance using Prüfer sequences. Data Knowl. Eng. 68(8), 728-747, 2009.
- [3] Aumueller, D., H.H. Do, S. Massmann, and E. Rahm: Schema and ontology matching with COMA++. Proc. SIGMOD, demo paper, 906-908, 2005.
- [4] Bellahsene, Z., A. Bonifati, F. Duchateau, and Y. Velegrakis: On Evaluating Schema Matching and Mapping. In: Z. Bellahsene, A. Bonifati, E.Rahm (eds), Schema Matching and Mapping, Springer, 2011.
- [5] Bellahsene, Z., A. Bonifati, and E. Rahm (editors), *Schema Matching and Mapping*, Springer, 2011.
- [6] Bergamaschi, S., S. Castano, and M. Vincini: Semantic Integration of Semistructured and Structured Data Sources. SIGMOD Record 28(1), 54-59, 1999.
- [7] Bernstein, P.A., T. J. Green, S. Melnik, and A. Nash: Implementing mapping composition. VLDB J. 17(2), 333-353, 2008.
- [8] Bernstein, P.A., L.M. Haas, M. Jarke, E. Rahm, and G. Wiederhold: Panel: Is Generic Metadata Management Feasible? Proc. VLDB, 660-662, 2000.
- [9] Bernstein, P.A., A.Y. Halevy, and R. Pottinger: A Vision of Management of Complex Models. SIGMOD Record 29(4), 55-63, 2000.
- [10] Bernstein, P.A. and S. Melnik: Model Management 2.0: Manipulating Richer Mappings. Proc. SIGMOD, 1-12 2007.
- [11] Bernstein, P.A., S. Melnik, and J.E. Churchill: Incremental schema matching. Proc. VLDB, demo paper, 1167-1170, 2006.
- [12] Bernstein, P.A., S. Melnik, M. Petropoulos, and C. Quix: Industrial-Strength Schema Matching. ACM SIGMOD Record 33(4), 38-43, 2004.
- [13] Bizer, C., T. Heath, and T. Berners-Lee: Linked Data – The Story So Far. Int. J. Semantic Web Inf. Syst. 5(3), 1-22, 2009.
- [14] Bohannon, P. E. Elnahrawy, W. Fan, and M. Flaster: Putting context into schema matching. Proc. VLDB, 307- 318, 2006.
- [15] Chiticariu, L., P. G. Kolaitis, and L. Popa: Interactive generation of integrated schemas. Proc. SIGMOD, 833- 846, 2008.
- [16] Cruz, I.F., F.P. Antonelli, and C. Stroe: AgreementMaker: Efficient Matching for Large Real-World Schemas and Ontologies. PVLDB 2(2), demo paper, 1586-1589, 2009.
- [17] Dhamankar, R., Y. Lee, A-H. Doan, A.Y. Halevy, and P. Domingos: iMAP: Discovering Complex Mappings between Database Schemas. Proc. SIGMOD, 383-394, 2004.
- [18] Do, H.H., S. Melnik, and E. Rahm: Comparison of Schema Matching Evaluations. In: Web, Web-Services, and Database Systems, Springer LNCS 2593, 221-237, 2003.
- [19] Do, H.H. and E. Rahm: COMA – A System for Flexible Combination of Schema Matching Approaches. Proc. VLDB, 610-621, 2002.
- [20] Do, H.H. and E. Rahm: Matching large schemas: Approaches and evaluation. Inf. Syst. 32(6), 857-885, 2007.

- [21] Doan, A-H., P. Domingos, and A.Y. Halevy: Reconciling the Schemas of Disparate Data Sources: A Machine- Learning Approach. Proc. SIGMOD, 509-520, 2001.
- [22] Doan, A-H., J. Madhavan, P. Domingos, and A. Y. Halevy: Learning to Map between Ontologies on the Semantic Web. Proc. WWW, 662-673, 2002.
- [23] Ehrig M., and S. Staab: Quick ontology matching. Proc. Int. Conf. Semantic Web (ICSW), Springer LNCS 3298, 683-697, 2004.
- [24] Ehrig M., S. Staab, and Y. Sure: Bootstrapping Ontology Alignment Methods with APPFEL. Proc. Int. Conf. Semantic Web (ICSW), Springer LNCS 3729, 186-200, 2005.
- [25] Elmeleegy, H., M. Ouzzani, and A.K. Elmagarmid: Usage- Based Schema Matching. Proc. ICDE, 20-29, 2008
- [26] Euzenat, J. and P. Shvaiko, *Ontology Matching*, Springer, 2007.
- [27] Euzenat, J., A. Ferrara, C. Meilicke, J. Pane, F. Scharffe, P. Shvaiko, H. Stuckenschmidt, O. Šváb-Zamazal, V. Svátek and C. Trojahn dos Santos: Final Results of the Ontology Alignment Evaluation Initiative 2010. Proc. 5th ISWC workshop on Ontology Matching, 2010.
- [28] Fagin, R.: Inverting schema mappings. ACM Trans. Database Syst. 32(4), Article 25, 2007.
- [29] Fagin, R., P.G. Kolaitis, L. Popa, and W.-C. Tan: Quasiinverses of schema mappings. ACM Trans. Database Syst. 33(2), Article 11, 2008.
- [30] Fagin, R., P.G. Kolaitis, L. Popa, and W.-C. Tan: Schema Mapping Evolution Through Composition and Inversion. In: Z. Bellahsene, A. Bonifati, E. Rahm (eds), Schema Matching and Mapping, Springer, 191-222, 2011.
- [31] Falconer, S.M. and N.F. Noy: Interactive Techniques to Support Ontology Matching. In: Z. Bellahsene, A. Bonifati, E. Rahm (eds), Schema Matching and Mapping, Springer, 29-52, 2011.
- [32] Gal, A.: Managing uncertainty in schema matching with top-k schema mappings. J. Data Semantics 6, 90-114, 2006.
- [33] Gal, A.: Enhancing the capabilities of attribute correspondences. In: Z. Bellahsene, A. Bonifati, E. Rahm (eds), Schema Matching and Mapping, Springer, 53-74, 2011.
- [34] Gross, A., M. Hartung, T. Kirsten, and E. Rahm: On Matching Large Life Science Ontologies in Parallel. Proc. 7th Int. Conf. Data Integration in the Life Sciences (DILS), Springer LNCS 6254, 35-49, 2010.
- [35] Giunchiglia F., P. Shvaiko, and M. Yatskevich: Semantic schema matching. Proc. OTM Conf. (CoopIS), 347-365, 2005.
- [36] Hartung, M., J. Terwilliger, and E. Rahm: Recent advances in schema and ontology evolution. In: Z. Bellahsene, A. Bonifati, E. Rahm (eds), Schema Matching and Mapping, Springer, 149-190, 2011.
- [37] He, B. and K. C.-C. Chang: Statistical Schema Matching across Web Query Interfaces. Proc. SIGMOD, 217-228, 2003.
- [38] He, H., W. Meng, C.T. Yu, and Z. Wu: Automatic integration of Web search interfaces with WISE-Integrator. VLDB J. 13(3), 256-273, 2004.
- [39] Hu, W., Y. Qu, and G. Cheng: Matching large ontologies: A divide-and-conquer-approach. Data Knowl. Eng. 67(1), 140-160, 2008.
- [40] Jean-Mary, Y.R., E.P. Shironoshita, and M.R. Kabuka: Ontology matching with semantic verification, J. Web Semantics 7(3):235–251, 2009.
- [41] Koudas, N., A. Marathe, and D. Srivastava: Flexible String Matching Against Large Databases in Practice. Proc. VLDB, 1078-1086, 2004.
- [42] Lambrix, P., H. Tan, and W. Xu: Literature-based Alignment of Ontologies. Proc. ICSW Ontology Matching (OM) workshop, 2008.
- [43] Lee, Y., M. Sayyadian, A. Doan, and A. Rosenthal: eTuner: tuning schema matching software using synthetic scenarios. VLDB Journal 16(1), 97-122, 2007.
- [44] Li, J., J.Tang, Y. Li, and Q. Luo. RiMOM: A Dynamic Multistrategy Ontology Alignment Framework. IEEE Trans. Knowl. Data Eng. 21(8), 1218-1232, 2009.
- [45] Madhavan, J., P. A. Bernstein, and E. Rahm: Generic Schema Matching with Cupid. Proc. VLDB, 49-58, 2001.
- [46] Madhavan, J., P.A. Bernstein, A. Doan, and A.Y. Halevy: Corpus-based Schema Matching. Proc. ICDE, 57-68, 2005.
- [47] Madhavan, J., D. Ko, L. Kot, V. Ganapathy, A. Rasmussen, and A. Y. Halevy: Google's Deep Web crawl. PVLDB 1(2), 1241-1252, 2008.
- [48] Magnani, M., N. Rizopoulos, P. McBrien, and D. Montesi: Schema integration based on uncertain semantic mappings. Proc. ER, Springer LNCS 3716, 31-45, 2005.

Key Frame Detection Using Two-Step Corner Selection

Monika Agarwal
 Electronics and Communication
 Department
 National Institute of Technology
 Durgapur, West Bengal, India
 monika.agarwal1990@gmail.com

Satyabrata Maity
 A.K.Choudhury School of Information
 Technology, University of Calcutta
 Kolkata, India
 satyabrata.maity@gmail.com

Amlan Chakrabarti
 A.K.Choudhury School of Information
 Technology, University of Calcutta
 Kolkata, India
 acakes@caluniv.ac.in

Abstract— A summarized video helps the viewer to get an overall idea of its content prior to watching the entire video. This not only saves time but also bandwidth required for data transfer, resulting in easy and faster video browsing. Apart from browsing a video in a smaller amount of time, summarized video information plays an important role in identifying the key event frames of a video which have lot of need in security and surveillance applications. Any video contains a large number of similar image frames which attributes to its continuity. Selecting the key-frames refers to selection of those image frames which reveal the vital information of the video without any repetition. In this paper, a method of key-frame selection for summarization of a video has been proposed, based upon the idea that large difference in the number of corner points between two images will indicate considerable change in object or event captured by the two images. The corner point detection is carried out using two-step corner selection method. The grouping of frames and key frame detection was done using Disturbance Ratio Method. The work is novel in the sense, that a statistical measure related to the variation in the corner point information has been used to identify the key frames of a video, which is a new approach in video summarization. This method was applied on some example videos, showing the effectiveness of this technique.

Keywords- Disturbance ratio, corner detection, edge detection, summarization

I. INTRODUCTION

Video Summarization is the process of generating an easily interpreted synopsis of a video by abstracting the main occurrences, scenes or objects in the video. The three main features, which a summarized video must have, are: i. *It must contain high priority entities and contents from a video*, ii. *The summary itself should exhibit a reasonable degree of continuity* and iii. *It should be free from repetition* [10]. This all sums up to the need that summarization of any video requires the detection of key frames. A video is basically composed of a large number of image frames; with frames appearing at a rate say 30-35 frames per second on the screen to generate a video experience. To maintain the continuity in the video, the frames closer to each

other in time interval, do not show appreciable information change between them. A group of frames selected in such a way that each frame carries unique information, not conveyed by any other frame in that group and the group in all captures the overall message conveyed by the video, are said to be the key frames of the video. A group of similar set of image frames is called scene and a sudden change from one scene to another determines a scene transition. So the collection of frames, one from each scene occurring in a video very efficiently serves as the key frames of the video. In create a summarization we have to add some intermediate frames to maintain the continuity.

Because of its vast application, several works have been done in order to find an efficient method for video summarization. In [1], we find color based selection of key frames and [2] was based on motion matrix which was function of optical flow within the frames. Object based selection is proposed in [3], where the key frames were categorized as key frame by event and key frame by action, based on change of region and feature extraction. Vector space based key frame detection [4] represented a frame as a point in multidimensional space and adopted curve splitting algorithm, detecting key frames based on curve properties. While in [5] singular value decomposition is used for detection of key frames, they used color histogram to represent video frames.

The key contributions of this paper can be summarized as follows:

- We have introduced the idea of using corner point detection as a feature that can be utilized to measure the change of information between the image frames in a video.
- A statistical measure of the corner information variation, between the image frames, has been utilized in grouping the frames and as well as detecting the key frame.
- Our methodology resulted to a much reduced size of the summarized video, which shows the effectiveness of the technique.

2nd Annual International Conference IEMCON 2012

The organization of the paper is as follows. Section II discusses our proposed methodology, results and discussions appear in Section III and concluding remarks are briefed in Section IV.

II. THE PROPOSED METHODOLOGY

The key to a good video-summarization is selection of the parameter to compare the information variation between the image frames. The grouping of image frames into scenes largely depends on this parameter. It determines the extent to which the algorithm will be able to distinguish between similar and dissimilar scenes. Corners are an important feature of an image and can be robustly detected. Corner points in an image deliver a great deal of information about the image, so does the number of corner points. In a video, we have a number of similar images close together, so the probability that a large change in the number of corners between two successive image-frames will mean a change in scene, and the probability that the similar frames will have almost same number of corner points, is quite high. Thus the difference between numbers of corner points for two consecutive image frames is the description of their similarity quotient. Larger the difference in the number of corner points, less similar the two frames are.

So, we first count the number of corners in each frame of the video. Then, we group the image frames into groups of similar frames. This is done using the Disturbance Ratio method proposed later. Key frames are detected by extracting one frame from each group.

For the detection of corner points we propose a two step corner selection method. The conventional corner detection algorithm detects a large number of false corners. The two step method reduces the number of false corners to a large extent making it more reliable. In this method, instead of directly applying the corner detection algorithm we first pre-process the image using an edge detection algorithm and then apply the corner detection algorithm.

A. Two-Step Corner Selection

A corner is but an intersection of two edges. This means that true corner points will lie on the edges only (in case of isolated point, it can itself be considered as an edge).

We adopt novel method of detecting a corner point by first processing the image using an edge-detector and then detecting the corners in this edge-detected image. A good edge-detection algorithm gives us an image describing only the objects in the image. Such an image is largely independent of brightness changes and additive noise. Detecting corners in the edge-detected image is found to give much better result than applying the corner detection algorithm directly as shown in "Fig. 1(a)" and "Fig. 1(b)".

The two mostly commonly used edge detection algorithms are Canny edge-detection algorithm [7] and Sobel edge-detection algorithm [9].

Although the Canny edge-detector gives more accurate result, it is computationally very complex and hence have more computation cost, making it unsuitable for application on a video, which contains a large number of image-frames. Whereas Sobel edge detector is less complex, fast and gives satisfactory results [9]. Since our consideration is large difference in the number of corners and not the exact corner points, the inaccuracy to this

extent is ignorable. Thus we adopt the Sobel Edge Detection algorithm.

Sobel Edge Detection

The Sobel operator performs a 2-D spatial gradient measurement on an image. Typically it is used to find the approximate absolute gradient magnitude at each point in an input gray scale image.

Steps:

- The Sobel edge detector uses a pair of 3x3 convolution masks, one estimating the gradient in the x-direction (columns) and the other estimating the gradient in the y-direction (rows).
- A convolution mask is usually much smaller than the actual image. As a result, the mask is slid over the image, manipulating a square of pixels at a time.
- If we define A as the source image, and G_x and G_y are two images which at each point contain the horizontal and vertical derivative approximations.

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A \text{ and } G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * A \quad (1)$$

- The magnitude of the gradient is then calculated using the formula:

$$|G| = G_x + G_y \quad (2)$$

- An approximate magnitude can be calculated using:
 $G = |G_x| + |G_y| \quad (3)$

After the edge-detected image is obtained, corner-detection is performed on the image. The several corner detection algorithms are the Moravec's corner detection algorithm [8] and Harris corner detection algorithm [8]. While Moravec's is the earliest of all corner detection algorithms, Harris method is an improved method.

Harris Corner Detection

This corner detection is an improved version of the Moravec's Corner Detection. In this method, instead of using shifted patches, differential of the corner score is directly considered.

For a shift of the window (u, v) under consideration, by say (x, y), the weighted sum of squared differences in intensity I is calculated as:

$$S(x, y) = \sum_u \sum_v w(u, v)(I(u+x, v+y) - I(u, v))^2 \quad (4)$$

Using Taylor's Expansion of I(u+x, v+y) this can be approximated to :

$$S(x, y) \approx \sum_u \sum_v w(u, v)(I(u, v)x + I(u, v)y)^2 \quad (5)$$

Where I_x and I_y are partial derivatives of I(x,y).

The above expression in matrix form is given by:

$$S(x, y) \approx \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} \quad (6)$$

Where,

$$A = \sum_u \sum_v w(u, v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix} \quad (7)$$

2nd Annual International Conference IEMCON 2012

A is called the Harris matrix. $\langle Ix^2 \rangle$ and $\langle I_y^2 \rangle$ denotes averaging over the window (u,v) . If a Gaussian window $w(u,v)$ is used the result is isotropic, thus overcoming the defect of Moravec's algorithm which was anisotropic in response.

The eigen values of A , define the cornerness but its computation is highly complex. Hence, Harris and Stephen introduced a parameter depending upon the eigen values of A , as a measure of corner response:

$$M_c = \lambda_1 \lambda_2 - \kappa(\lambda_1 + \lambda_2)^2 = \det(A) - \kappa \text{trace}^2(A) \quad (8)$$

M_c is largely positive for a corner, small in magnitude for a flat region and largely negative in case on an edge.

Harris corner detection is widely used corner detection algorithm for its largely improved results.

It is based on the idea that a corner is a point with strong brightness changes in orthogonal or near orthogonal directions. Since it is based on the intensity of pixels it is very sensitive to additive noise and brightness changes, and hence results in a large number of false corner points.

"Fig. 1", depicts the improvement in the corners detected in the same image when on one image only corner detection algorithm was applied "Fig. 1a" and the other was processed by using the proposed two step corner selection algorithm "Fig. 1b".

Two-step corner selection method has the following advantages:

1. The detection of corner points is not affected by the brightness change within the image but only depends on the objects and their orientation, so the change in the image frames depends on the change of objects and their orientation. Thus this parameter becomes suitable for reliable shot detection mechanism.
2. The large reduction in the number of false corners makes it a true descriptor of the image frame and hence suitable for key frame detection.

B. Detection of Key Frames

We use the two step corner selection algorithm to detect corners in all the image frames comprising the video and keep a record of the number of corner points in each frame naming it, the corner value of that particular frame. This corner value is now used to compare consecutive image frames whose difference in magnitude is inversely proportional to the similarity between the two images. "Fig. 2" shows a graph of corner values against frame number for a video sequence. Next step in the detection of key frames is grouping of frames into scenes. According to the proposed idea, scenes are groups of consecutive frames with almost same number of corner values. We group the frames by the Disturbance Ratio (DR) method [10]. This method measures the disturbance in a video as the ratio of the range of corner values to the standard deviation "(11)". Range measures the difference between the highest and the lowest value, hence a measure of largest variation. Large value of DR is obtained from large value of range and low standard deviation. Low standard deviation implies negligible varieties but large range implies a major gap between the lowest and the largest value. This indicates the presence of spikes. A threshold value directly proportional to DR measure is fixed "(12)", which will determine the allowable amount of disturbance in consecutive frames to group them into a scene. In a video, say there are n numbers of image frames.

Let $x(i)$ be the corner value of image frame i where $1 \leq i \leq n$;

$$\text{Global standard deviation} = \frac{1}{n} \sqrt{\sum_{i=1}^n (x(i) - \bar{x})^2} \quad (9)$$

$$\text{Global range} = \text{global maximum} - \text{global minimum} \quad (10)$$

$$\text{DR} = (\text{Global range}) / (\text{Global standard deviation}) \quad (11)$$

$$\text{Threshold (T)} = k \times \text{DR} \quad (12)$$

K = multiplication factor (to be adjusted depending upon the video by trial and error basis)

The steps for grouping are as follows:

- (1) Starting with the first element as a group in itself, the average and the standard deviation for the group is calculated. The average is the element itself and standard deviation in this case is thus zero.
- (2) The next element is converged into the group, the average and the standard deviation of the group thus formed is computed. The standard deviation of the group is now compared with the threshold value.

Figure 1: COMPARISON BETWEEN HARRIS CORNER DETECTION TWO STEP CORNER SELECTION ALGORITHMS.



Figure 1(a): Corner points detected by Harris corner detection algorithm.



Figure 1(b): Corner points as detected by two step corner detection algorithm.

2nd Annual International Conference IEMCON 2012

- Greater standard deviation means greater difference in values and hence implies disturbance within the group.
- (3) The next frame is converged into the group and the process repeated. As soon as the group standard deviation becomes larger than the threshold value, the group breaks and the last frame in that group is identified as a key frame.
 - (4) Fresh grouping and calculation starts from the frame next to the last key frame identified. The process is repeated over all the corner values.
 - (5) Last frame of the video has no succeeding frame for comparison and is always detected as a key frame.

III. RESULTS AND DISCUSSIONS

The proposed method of key frame detection was applied on several videos using MATLAB R2010a software. The keyframes detected for two videos [11] is shown in “Fig. 3”. “Fig. 2” shows

Figure 2: VARIATION OF NUMBER OF CORNER POINTS WITH IMAGE FRAME.

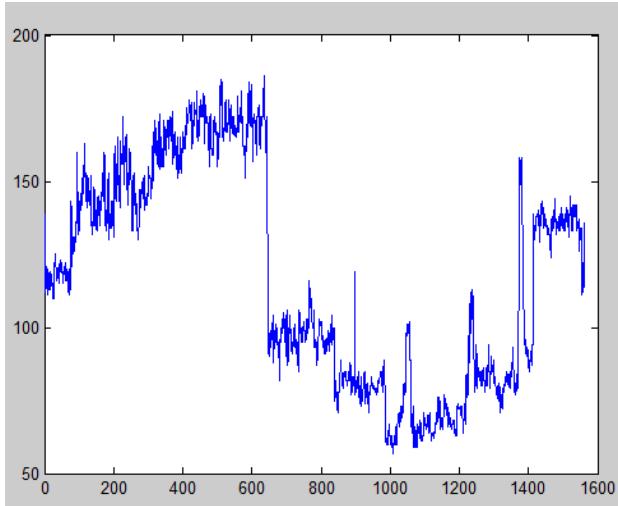


Figure 2a: Variation of number of corner points with the image frame for video1

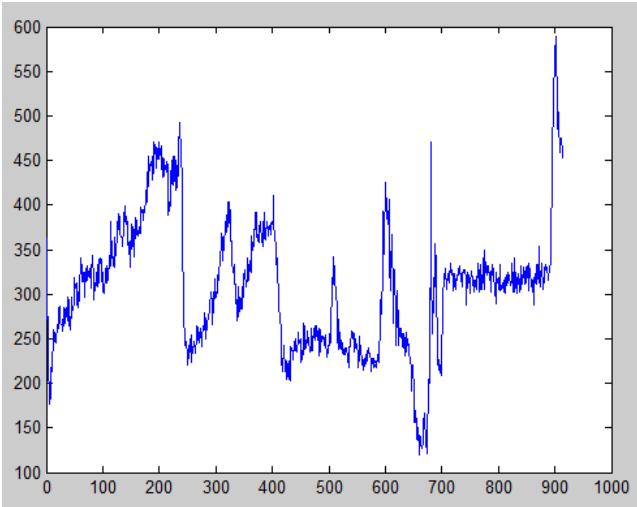


Figure 2b: Variation of number of corner points with the image frame for video2

the variation of number of corner points with the image frames for the two videos. Large variance in consecutive values indicates a scene transition. The multiplicative factor K was found out by trial and error method. “Fig. 3a” shows the key frames detected for video1 for which the multiplicative factor K was found to be 5. “Fig. 3b” shows the detected key frames for video2 for $K = 3$. Table 1 shows results of our proposed summarization technique applied on the two example videos. The results prove that our technique is efficient.

Figure 3: KEY FRAMES DETECTED FOR SAMPLE VIDEOS.



Figure 3a: Key Frames detected for video1



Figure 3b: Key Frames detected for video2

TABLE I: RESULTS OF THE SUMMARIZATION METHOD APPLIED OVER THE TWO EXAMPLE VIDEOS

| Index | Video | File Type | Total # Frames | # Key Frames Detected | % Reduction In Video Size |
|-------|--------|-----------|----------------|-----------------------|---------------------------|
| 1 | Video1 | WMV | 507 | 6 | 98.82 |
| 2 | Video2 | MPG | 913 | 10 | 98.90 |

2nd Annual International Conference IEMCON 2012

IV. CONCLUSION

In this paper a novel method of detecting the key frames in a video has been proposed, based on the corner feature. The corner detection based approach is a new work in this domain, and the results prove the efficiency of the approach. This work can be further improved by devising an algorithm where the multiplication factor 'K' can be defined as a function of the video or image frame features. The multiplication factor in that case can be dynamically allocated, resulting in self-adaptive key frame detection algorithm requiring no external input.

REFERENCES

- [1] Latecki.L.J, de Wildt. D, Jianying Hu, "Extraction of key-frames from video using optimal color composition matching and polygon simplification" in Multimedia Signal Processing,2001 IEEE Fourth Workshop on 2001
- [2] W. Wolf, "Key frame selection using motion analysis", Proceedings of ICASSP'00, vol.2, pp. 1228-1231, 1996
- [3] Changick Kim, Jenq-Neng Hwang, "Object-based Video Abstraction for Video Surveillance Systems", IEEE Transactions on circuits and Systems for video technology, vol. 12, No. 12, 2002
- [4] Li Zhao; Wei Qi; Li, S.Z.; Yang, S.Q.; Zhang H.J.; "Content-based retrieval of video shot using the improved nearest feature line method". Proceedings of ICASSP'01, vol. 3, pp. 1625-1628, 2001
- [5] Yihong Gong and Xin Liu,"Video summarization using singular value decomposition" in Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on 2000
- [6] H.J Zhang, "An integrated system for content-based video retrieval and browsing". Pattern Recognition, vol.30, no.4, pp.643-658, 1997.
- [7] Canny, John, "A computational approach to edge detection" in Pattern Analysis and Machine Intelligence, IEEE Transactions on November 1986.
- [8] C. Harris and M. Stephens (1988). "A combined corner and edge detector". Proceedings of the 4th Alvey Vision Conference. pp. 147–151
- [9] Sharifi.M, Fathy.M, Mahmoudi.M.T, "A classified and a comparative study of edge detection algorithms" in Information Technology: Coding and Computing, 2002. Proceedings International Conference on April 2002.
- [10] S. Maity, A.Chakrabarti and D.Bhattacharjee; "An Innovative Technique for Adaptive Video Summarization", SPRINGER, ICIP, Bangalore, August,2011
- [11] www.youtube.com/user/2000turtle

Fiber Optics Laser Systems in Cardio Vascular Diseases

Debadyoti Ghosh, Md. Niyaz Mallick, Farha Serin Naaz, Alok Kumar Das.

Electronics & Communication Engineering Department,
Institute of Engineering & Management, Salt Lake, Kolkata, India
debadyoti@gmail.com

ABSTRACT – The purpose of the paper is to design and develop ordered-fiber optic bundles as diagnostic tools called as optical fiber sensors. The study involves improving the sensitivity, the accuracy and cost reductions. With the development of modern medical lasers and laser sensors, therapy has gained an increasing role in the wide spectrum of treatment modalities. Also in cardio vascular diseases laser techniques have become very interesting. The Nd:YAG laser, Dye laser, Excimer laser, CO₂ laser are the most important surgical lasers. Possibility of transmitting its light through flexible fibers allows wide variation of applications. Here we have also shown the construction of Q-switched pulsed mode Nd:YAG laser which is used in surgical sensors. This laser is mostly known for its ability of volume coagulation, which is due to its large optical penetration depth. It shows a reaction depth which can be varied in the widest range of all medical lasers by using the appropriate parameters. The applications of different lasers in medical science and the fabrications of sensor fiber and power fiber are also discussed.

Keywords – Ordered bundle optical fibers, Q-switched pulsed mode Nd:YAG laser, optical sensor fibers, optical power fibers and cardio vascular diseases.

I. INTRODUCTION

During the past few decades, cardiology has become a prominent specialty in medicine. Cardio vascular disease is one of the principle causes of death worldwide. Major surgery and other invasive modalities are often required to diagnose and treat the disease. The heart, coronary arteries, and peripheral arteries all constitute a rather complex system of tubes, pumps, and valves which lends itself to the use of fiber optic investigation and treatment. Laser fiber optic system can be useful to diagnose and treat cardiovascular disease. In this paper, the design of ordered-fiber optic bundles as diagnostic tools called as optical fiber sensors has been shown. The study involves

improving the sensitivity, the accuracy and cost reductions. The laser beam through the transmitted fiber illuminates diagnostic elements of biological fluid like blood, urine or tissue. The reflected (or the luminescent) light from the biological fluid or the tissue is collected and transmitted through the second fiber to a detector. The signal obtained from the detector is proportional to the reflection by blood of particular color. If the process is repeated for a number of colors, the reflected spectrum of blood is obtained. Such a spectrum can reveal the oxygen spectrum of blood. In this case, the ends of the fibers are base and diagnosis is made by direct interaction between light and blood named as direct fiber optic sensors.

II. OPTICAL FIBER

An optical fiber is a flexible, transparent and cylindrical dielectric waveguide (nonconducting waveguide) made of a pure glass (silica) not much wider than a human hair that transmits light along its axis (between the two ends of the fiber), by the process of total internal reflection. The fiber consists of a core surrounded by a cladding layer, both of which are made of dielectric materials. To confine the optical signal in the core, the refractive index of the core must be greater than that of the cladding. The boundary between the core and cladding may either be abrupt, in step-index fiber, or gradual, in graded-index fiber. Signals can travel along the fibers with less loss and less electromagnetic interference. Fibers that support many propagation paths or transverse modes are called multi-mode fibers (MMF), while those that only support a single mode are called single-mode fibers (SMF). Multi-mode fibers generally have larger core diameters, and are normally used as therapy for medical purpose. While a singal mode fiber is used in communication. But the bundle of fibers made of a number of fibers having small core diameters can be used for medical diagnosis sensing.

There are two conditions necessary for total internal reflection to occur -

- The refractive index of the first medium is greater than the refractive index of the second medium ($n_1 > n_2$).
- The angle of incidence must be greater than the critical angle ($i > c$).

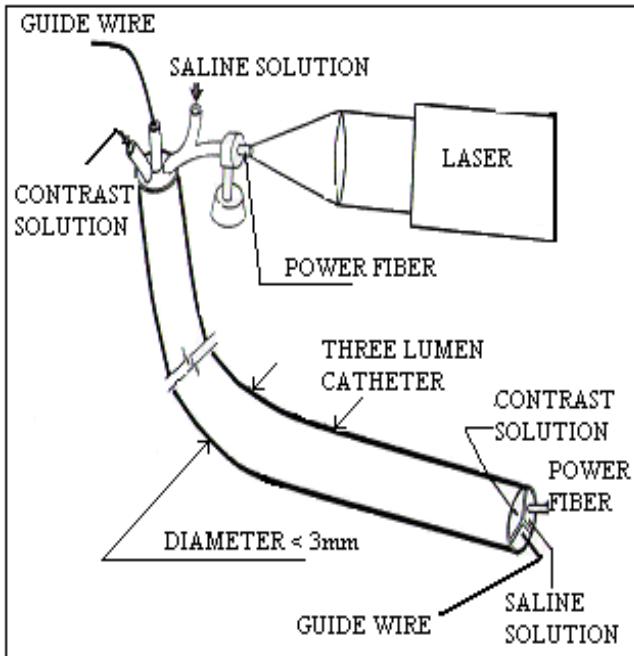


Figure 1. Laser catheter for cardiology

III. FABRICATION OF ORDERED BUNDLES FOR SENSING

An ordered bundle may be fabricated by winding a clad fiber on a precision drum as shown in Fig.2. The windings are carefully aligned – unlike the case of nonordered bundle. After windings one layer onto a drum, other layers are continuously wound, one on top of the other, in an orderly fashion. The whole bundle is cut; the fibers on each end of the cut are glued together, and the ends are polished. Tens of thousands of fibers with diameters of $10\mu\text{m}$ - $20\mu\text{m}$ can be bunched into one bundle. Alternatively a small bundle of fibers called multifiber can be fabricated. The multi-fiber thus contains many smaller fibers inside. A typical multifiber may contain 100 fibers, each with diameter of 5 um. The outer diameter is relatively small, and multifiber is thus quite flexible. Several multifibers can be aligned in a bundle. The end result is a full ordered bundle, which can have a high number of fibers and high resolution. The number of fibers in multifiber is sometimes large (e.g., 1000).

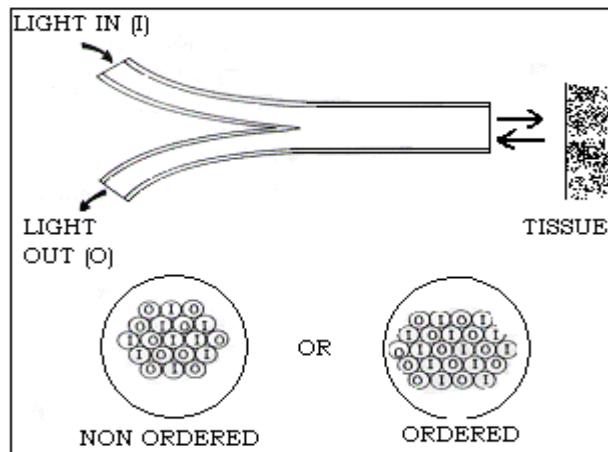


Figure 2. Construction of fiber

IV. POWER FIBER FOR LASER SURGERY, TISSUE WELDING AND ABLATION

The simple system consists of a laser, an optical fiber for transmitting the laser power and a lens for coupling the laser into the fiber as shown in Fig.3. Each fiber has an acceptance angle and all rays of light that impinge on the input end of the fiber at a smaller angle can be propagated through the fiber. A laser beam whose diameter is often several millimeters has to be coupled into the core of a fiber whose diameter is often a fraction of a millimeter. For this purpose short, medium or long focal length lenses can be used.

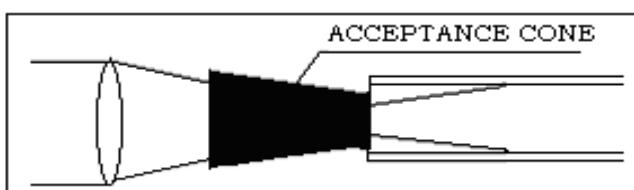


Figure 3. Coupling of laser light into a power fiber

The table shows some of the continuous lasers that can be used in medical science:

| Laser | Wavelength (μm) | Power (W) | Fiber Transmission |
|-----------------|------------------------------|-----------|--------------------|
| CO ₂ | 10.6 | 0.5 - 50 | No |
| Nd-YAG | 1.064 | 0.5 – 100 | Yes |
| Argon | 488 or 514 | 01 – 10 | Yes |
| Dye | Tunable | 0.05 – 5 | Yes |
| Excimer | 0.353 | 01 – 10 | Yes |

The table shows some of the pulsed lasers that can be used in medical science:

| Laser | Wavelength (μm) | Pulse duration | Energy/Pulse(J) | Fiber Transmission |
|--------------|------------------------|-----------------------|------------------------|---------------------------|
| Nd-YAG (QS) | 1.064 | ns | 0.1 - 1 | No |
| Nd-YAG | 1.064 | μ s | 0.1 - 1 | Yes |
| Dye | 0.39-0.44 | μ s | 0.01-0.1 | Yes |
| Excimer | 0.353 | ns | 0.01-0.1 | Yes |

V. ENDOSCOPIC LASER SYSTEM IN CARDIOLOGY

The fiber optical laser systems used in cardiology are as follows:

- Guidance
- Imaging
- Diagnosis
- Therapy

Guidance: Optical fibers used for therapeutic or diagnostic purposes are placed near a blockage in an artery with endoscope. Its position is monitored by X-ray fluoroscopy. When the guide wire is in place, the physician slides a thin catheter over it and into the coronary artery. The guide wire is then pulled out. A thin optical fiber can be inserted into the same catheter and pushed all the way into the artery until its distal tip is brought into contact with the atherosclerotic blockage.

Imaging: The development of thin and ultra thin endoscopes paved the way for fiber optic imaging inside blood vessels. These endoscopes however are not as rigid as guide wire. The fiber optic imaging is performed with regular white light in order to see a detailed image through a bundle. There are two basic requirements: one is high level of illumination, and another is high resolution.

a.High level of illumination: The individual fibers must have a large numerical aperture for a high quality image. Therefore the core diameter d must be large and cladding layer thickness (D-d) is preferably small, where D is the overall diameter of each fiber.

b.High resolution: Resolution is the optical property that makes it possible to see small details of the image. The Resolution of an optical system can be measured by

imaging a line pattern similar to the bar code lines which are found in most items in the supermarket.

Diagnosis: Optical fiber sensors can be inserted into blood vessels or the heart via thin catheters or thin endoscopes (as shown in figure.1).

Therapy: Laser angioplasty has been performed to recognize blockages in the coronary / peripheral arteries in 1000s of patients. Laser can also be used for endarterectomy or for tissue welding in the cardiovascular system.

VI. DIFFERENT TYPES OF LASERS AND THEIR CLINICAL APPLICATIONS

Lasers are useful for surgery because

- They can shine high intensity light that can be focused for cutting or destroying tissue.
- They produce heat that causes the tissue around the cut to seal and prevents bleeding.
- The beam is narrow and can therefore make very precise and accurate cuts.

There are various types of lasers used in medicine and surgery. Some of them and their applications have been enlisted below –

| Specialty | Laser | Applications |
|------------------|--------------------------|---|
| Cardiology | Excimer, Ar, Dye, Nd:YAG | Laser angioplasty, endarterectomy. |
| Dentistry | CO2 Nd:YAG | Soft tissue surgery. Caries removal. |
| Dermatology | Ar, Dye | Port wine stains & strawberry marks. Varicose vein & tattoo excision, spider nevi. |
| General surgery | CO2 Nd:YAG | General cutting tool, welding. Cholecystectomy. |
| Ophthalmology | Excimer Ar ion Nd:YAG | Corneal surgery. Retinal detachment, iridectomy. Posterior capsulotomy; iridectomy; vitreous bands. |
| Orthopedics | CO2 Ho:YAG | Arthroscopic surgery; bone tumor excision. |
| Oncology | Dye Nd:YAG | Photodynamic therapy of tumors. Hyperthermia, thermotherapy of tumors. |

| | | |
|---------|--------|--|
| Urology | Ar ion | Urethral stricture, bladder hemorrhage. |
| | Dye | Laser lithotripsy (kidney stones). |
| | Nd:YAG | Bladder bleeding, bladder tumor therapy. |
| | CO2 | Renal resection, penile carcinoma. |

VII. Q-SWITCHED PULSED MODE ND: YAG LASER

Lasers based on YAG crystals are usually used for solid-state lasers made of neodymium-doped YAG (Nd: YAG, more precisely Nd³⁺: YAG). YAG is the acronym for yttrium aluminum garnet ($\text{Y}_3\text{Al}_5\text{O}_{12}$), a synthetic crystal material which became popular in the form of laser crystals in the 1960s. Pulsed Nd: YAG lasers are typically operated in Q-switching mode: An optical switch is inserted in the laser cavity waiting for a maximum population inversion in the neodymium ions before it opens. Then the light wave can move through the cavity, depopulating the excited laser medium at maximum population inversion. In this Q-switched mode, output powers of 250 megawatts and pulse durations of 10 to 25 nanoseconds have been achieved. The high-intensity pulses may be efficiently frequency doubled to generate laser light at 532 nm, or higher harmonics at 355 and 266 nm. YAG is a host medium with favorable properties, particularly for high-power lasers and Q-switched lasers emitting at 1064 nm. Nd: YAG absorbs mostly in the bands between 730–760 nm and 790–820 nm.

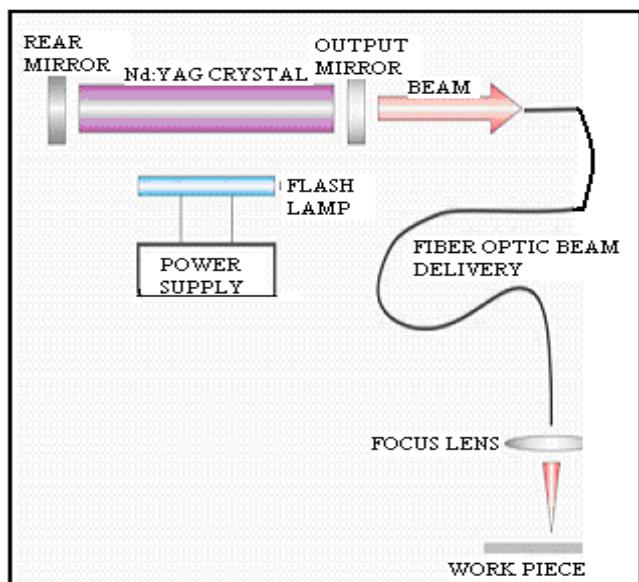


Figure 4. Construction of laser

VIII. DESIGN OF OPTICAL SENSOR FIBER

(a) Sensor Fiber

When light is focused on the end face of the whole bundle, each of the cores transmits light (as shown in Fig.2). The area occupied by the cores is thus used for light transmission, whereas the area occupied by the cladding transmits little light. If the overall diameter of each fiber is D and the core diameter is d , the ratio of the core area to that of the whole fiber is approximately $0.9(d^2/D^2)$.

Considering the standard 1mm bundle diameter, and considering core diameter of 40 μm for each fiber,

$$\text{Numbers of optical fiber} = \pi (1000)^2 / \pi (40)^2 = 625$$

Therefore the actual area of the sensor fiber

$$= 0.9 \times 625 \times \pi \times (20)^2 = 0.707 \text{ mm}^2.$$

(b) Power fiber

The consideration in designing the light guiding fiber optic bundle is the numerical aperture (NA) of the individual fiber. In order to collect as much light as possible from a light source the individual fibers in the bundle must have a large NA; a wide cone of light must be collected by each fiber. The collection efficiency is actually proportional to the NA. We know for a fiber the critical angle, numerical aperture and angle of acceptance are expressed as

$$\text{Angle of acceptance} = \sin^{-1} (\text{NA}) \dots \dots \dots \text{(iii)}$$

Where n_1 and n_2 are core refractive index and cladding refractive index respectively.

Thus for $n_1 = 1.5$ and $n_2 = 1.47$,

$$\text{Critical angle } \Phi_c = \sin^{-1}(n_2/n_1) = \sin^{-1}(1.47/1.50) = 78.5^\circ$$

$$\text{Numerical aperture} = \text{NA} = (n_1^2 - n_2^2)^{\frac{1}{2}}$$

$$= (1.50^2 - 1.47^2)^{1/2} = 0.30$$

$$\text{Angle of acceptance} = \sin^{-1} \text{NA}$$

$$= \sin^{-1} 0.30$$

$$= 17.4^{\circ}$$

The ability to collect light from a source or collection efficiency per unit area = $k (NA)^2 = k (0.3)^2$, where k is constant depending upon the types of fiber used.

IX. CONCLUSION

Here we have shown the design and fabrication principles of sensor and power fibers to be used in different medical applications. We have selected the silicon based optical fiber taking core diameter as 40 μm . and overall diameter 1mm having the numbers of fiber around 1000 for sensor fibers to form a bundled fiber of a length up to two meter. For the power fiber we have taken the same length having numerical aperture = 0.30, critical angle = 78.5° & Angle of acceptance = 17.4° , with a core diameter is 300 μm . However further research work on this topic may add a new feather to the field of medical surgery.

X. REFERENCES

- [1] J. E. Geusic et al., "Laser oscillations in Nd-doped yttrium aluminum, yttrium gallium and gadolinium garnets", *Appl. Phys. Lett.* 4 (10), 182 (1964)
- [2] D. Y. Shen et al., "Highly efficient in-band pumped Er:YAG laser with 60 W of output at 1645 nm", *Opt. Lett.* 31 (6), 754 (2006).
- [3] J. W. Kim et al., "Fiber-laser-pumped Er:YAG lasers", *IEEE Sel. Top. Quantum Electron.* 15 (2), 361 (2009).
- [4] Li Chaoyang et al., "106.5 W high beam quality diode-side-pumped Nd:YAG laser at 1123 nm", *Opt. Express* 18 (8), 7923 (2010).
- [5] X. Délen et al., "34 W continuous wave Nd:YAG single crystal fiber laser emitting at 946 nm", *Appl. Phys. B* 104 (1), 1 (2011).
- [6] Allan W.B.(1973). *Fiber Optics: Theory and practice*. New York: Plenum Publishing.
- [7] Allard, E.C. (1990) *Fiber Optics Handbook*. New York: McGraw-Hill.
- [8] Cherin, A.H. (1983). *An Introduction to Optical Fibers*. New York: McGraw-Hill.

[9] Kao,K.C, and Hockham, G.A (1966). Dielectric fiber surface waveguides for optical frequencies. *Proc.IEEE* 113.1151.

[10] All-Solid-State Nd:YAG Laser Operating at 1064 nm and 1319 nm under 885 nm Thermally Boosted Pumping.

Ding Xin^{1,2}, Chen Na^{1,2}, Sheng Quan^{1,2}, Yu Xuan-Yi³, Xu Xiao-Yan^{1,2}, Wen Wu-Qi^{1,2}, Zhou Rui^{1,2}, Wang Peng^{1,2} and Yao Jian-Quan^{1,2}

[11] "Q switching of a diode-pumped Nd:YAG laser with GaAs", T. T. Kajava* and Alexander L. Gaeta, School of Applied and Engineering Physics, Cornell University,Ithaca, New York 14853,Received February 20, 1996.

[12] Berci G, Forde KA. History of endoscopy: what lessons have we learned from the past? *Surg Endosc.* 2000;14(1):5–15. [PubMed].

[13] Sivak MV. Gastrointestinal endoscopy: past and future. *Gut.* 2006;55(8):1061–4. [PMC free article] [PubMed].

[14] Kodama T, Tatsumi Y, Sato H, Imamura Y, Koshitani T, Abe M, Kato K, Uehira H, Horii Y, Yamane Y, Yamagishi H. Initial experience with a new peroral electronic panreatoscope with an accessory channel. *Gastrointest Endosc.* 2004;59(7):895–900. [PubMed].

Multi-Sensor Data Fusion using Particle Filter to Estimate Non-linear Recursive Bayesian Model for Object Tracking

Anindya Mondal

Department of Electronics and Communication Engineering
Heritage Institute of Technology
Kolkata, India
mondalanindya@ymail.com

Moloy Kumar Chowdhury

Department of Electronics and Communication Engineering
Heritage Institute of Technology
Kolkata, India
moleykumar.chowdhury@gmail.com

Abstract— Multi-Sensor Data Fusion (MSDF) is very rapidly emerging as an independent discipline to be reckoned with and finds ever-increasing applications in many biomedical, automation industry, aerospace, robotics, and environmental engineering processes and systems, in addition to typical defense applications. MSDF offers one or more of the following benefits: more spatial coverage of the object under observation, redundancy of measurements, robustness of the system's performance and higher accuracy (basically reduced uncertainty of prediction) of inferences, and an overall assured performance of the sensor-integrated systems. In this paper, we used Multi-Sensor Data Fusion (MSDF) using Particle Filter to Estimate Non-linear Recursive Bayesian Model for Object Tracking and studied its performance from extensive Monte-Carlo simulations.

Keywords- *Object tracking; Particle Filter; Multi-sensor Data Fusion(MSDF); Bayesian Model.*

I. INTRODUCTION

In statistics, particle filters, also known as Sequential Monte Carlo methods (SMC), are sophisticated model estimation techniques based on simulation [1]. Particle filters have important applications in econometrics [2], and in other fields.

Particle filters are usually used to estimate Bayesian models in which the latent variables are connected in a Markov chain — similar to a hidden Markov model (HMM), but typically where the state space of the latent variables is continuous rather than discrete, and not sufficiently restricted to make exact inference tractable (as, for example, in a linear dynamical system, where the state space of the latent variables is restricted to Gaussian distributions and hence exact inference can be done efficiently using a Kalman filter). In the context of HMMs and related models, "filtering" refers to determining the distribution of a latent variable at a specific time, given all observations up to that time; particle filters are so named because they allow for approximate "filtering" (in the sense just given) using a set of "particles" (differently-weighted samples of the distribution).

Particle filters are the sequential ("on-line") analogue of Markov chain Monte Carlo (MCMC) batch methods and are often similar to importance sampling methods. Well-designed particle filters [13] can often be much faster than MCMC. They are often an alternative to the Extended Kalman filter (EKF) or Unscented Kalman filter (UKF) with the advantage that, with sufficient samples, they approach the Bayesian optimal estimate, so they can be made more accurate than either the EKF or UKF. However, when the simulated sample is not sufficiently large, they might suffer from sample impoverishment. The approaches can also be combined by using a version of the Kalman filter as a proposal distribution for the particle filter.

Sensor networks and intelligent arrays of micro-sensors have broad range of applications including information gathering and data fusion for modeling an environment, surveillance, active monitoring of forests & agricultural lands, health-care applications, collaborative information processing, and control of smart materials with embedded sensors. Multi-sensor data fusion is an emerging technology applied to Department of Defense (DoD) areas such as automated target recognition, battlefield surveillance, and guidance and control of autonomous vehicles, and to non-DoD applications such as monitoring of complex machinery, medical diagnosis, and smart buildings. Techniques for multi-sensor data fusion are drawn from a wide range of areas including artificial intelligence, pattern recognition, statistical estimation and other areas [3, 4, 5, 6, 7].

The problem of tracking moving objects—including targets, mobile robots, and other vehicles—using measurements from sensors is of considerable interest in many military and civil applications that use radar, sonar systems, and electro-optical tracking systems (EOTs) for tracking flight testing of aircrafts, such as missiles, unmanned aerial vehicles, micro- or mini-air vehicles, and rotorcrafts [10]. It is also useful in nonmilitary applications such as robotics, air traffic control and management, air surveillance, and ground vehicle tracking. In practice, scenarios for target tracking could include

maneuvering, crossing, and splitting (meeting and separating) targets [8, 9]. Various algorithms are available to achieve target tracking for such scenarios. The selection of the algorithms is generally application dependent and is also based on the merits of the algorithm, complexity of the problem, and computational burden.

II. OVERVIEW OF MULTI-SENSOR DATA FUSION

Sensor fusion is the combining of sensory data or data derived from sensory data from disparate sources such that the resulting information is in some sense better than would be possible when these sources were used individually.

To ensure that systems are operating within defined conditions, measurements are taken which, when analyzed, enable decisions to be made based on condition. These measurements can produce data that are either very similar, often from the same sensor, or completely different from different techniques. Experienced engineers and analysts have traditionally undertaken the analysis of this data. However, with the increased computer power and development of new and novel detection systems, the data produced needs to be handled in a robust and logical manner. As such computer systems have been developed that are capable of extracting meaningful information from the recorded data. The integration of data, recorded from a multiple sensor system, together with knowledge, is known as data fusion. Data fusion first appeared in the literature in the 1960s, as mathematical models for data manipulation. It was implemented in the US in the 1970s in the fields of robotics and defense [13]. In 1986 the US Department of Defense established the Data Fusion Sub-Panel of the Joint Directors of Laboratories (JDL) to address some of the main issues in data fusion and chart the new field in an effort to unify the terminology and procedures. The present applications of data fusion span a wide range of areas: maintenance engineering, robotics, pattern recognition and radar tracking, mine detection and other military applications, remote sensing, traffic control, aerospace systems, law enforcement, medicine, finance, metrology, and geo-science.

Multi-sensor data fusion (MSDF) is defined as the process of integrating information from multiple sources to produce the most specific and comprehensive unified data about an entity, activity or event. MSDF is expected to result in robust operational performance, extended spatial coverage, increased confidence, reduced ambiguity, improved detection performance, enhanced spatial resolution, improved system operational reliability and increased dimensionality (Hall 1992)[11, 12]. At a basic level the processing operations are dominated by numeric procedures involving linear and nonlinear estimation techniques, pattern recognition processes and various statistical operations. For a flight test range the tracking of the flight vehicle and sensor fusion are of great importance.

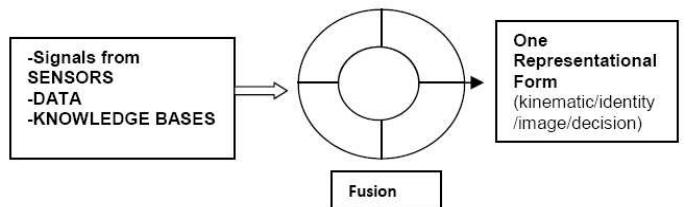


Figure 1: Chain of data fusion From Data Fusion to result

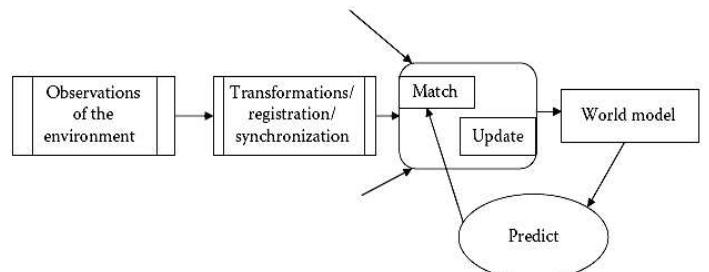


Figure2: Dynamic world-model framework

III. STRATEGIES FOR TARGET TRACKING AND DATA FUSION

Target tracking comprises estimation of the current state of a target, usually based on noisy measurements. The problem is complex even for single target tracking because of uncertainties in the target's mathematical model, especially for maneuvering targets, and process/state and measurement noises. The complexity of the tracking problem increases for multiple-targets using measurements from multiple sensors. The importance of data fusion for tracking stems from the following considerations. For an aircraft observed by pulsed radar and an infrared (IR) imaging sensor, the radar provides the ability to accurately determine the aircraft's range; however, it has only a limited ability to determine the angular direction, whereas the IR sensor can accurately determine the aircraft's angular direction, but not the range. If these two types of observations are correctly associated and fused, this will provide an improved determination of range and direction than could be obtained by either of the two sensors alone. Based on an observation of the attributes of an object, the identity of the object can be determined and the observations of angular direction, range, and range rate may be converted to estimate the target's position and velocity.

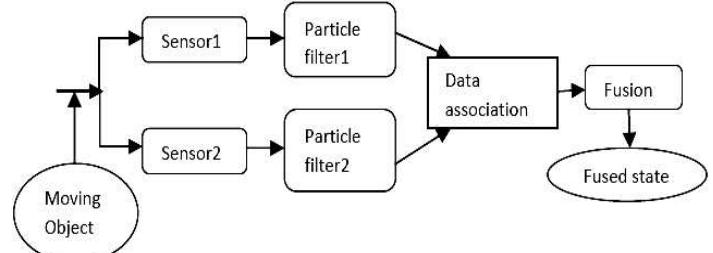


Figure 3: Concept of sensor data fusion for target tracking

IV. BAYESIAN FUSION METHODS

The Bayesian approach involves the definition of priors (a priori probabilities), their specifications, and computations of the posteriors [13]. Primarily, the probability theory is based on crisp logic, comprising “zero” or “one” (yes or no, on or off, -1 or +1). It does not consider any third possibility, because the probability definition is based on set theory, which is in turn based on crisp logic; it considers only the probability of occurrence or non occurrence of an event.

A. Bayesian Method

Let us define $p(A)$ as the probability of occurrence of an event A , and $p(A|B)$ as the probability of occurrence of two events A and B . Then, the conditional probability of occurrence of A , given that the event B has already occurred, can be related as follows:

$$p(A, B) = p(A|B)p(B) \quad (1)$$

We will also notice that because $p(A, B) = p(B, A)$,

$$p(B, A) = p(B|A)p(A) \quad (2)$$

$$p(A|B) = p(B|A)p(A) / p(B) \quad (3)$$

The above relation can also be written as follows:

$$p(A|B) = \frac{p(B|A)p(A)}{\sum_i p(B|A_i)p(A_i)} \quad (4)$$

The denominator acts as a normalization factor if there are several events of A that can be distinguished from B in a few ways. The above equation is known as Bayes' rule. Replacing A with x and B with z , we obtain the following relation:

$$p(x|z) = \frac{p(z|x)p(x)}{\sum_i p(z|x_i)p(x_i)} \quad (5)$$

The items x and z are regarded as random variables, and x is a state or parameter of the system and z is the sensor measurements. Thus, the Bayes' theorem is interpreted as the computation of the posterior probability, given the prior probability of the state or parameter ($p(x)$), and the observation probability ($p(z|x)$): the value of x that maximizes the term ($x|data$). The maximum likelihood is related to this term if $p(x)$ is considered a uniform distribution: the value of x that maximizes the term ($data|x$). The term $p(z|x)$ assumes the role of a sensor model in the following way:

- (1) First build a sensor model: fix x , and then find the probability density function (pdf) on z .
- (2) Use the sensor model: observe z , and then find the pdf on x .
- (3) For each fixed value of x , a distribution in z is defined.
- (4) As x varies, a family of distributions in z is formulated.

For the observation z of a target-tracking problem with state x , the Gaussian-observation model is given as a function of both z and x as shown below:

$$p(z|x) = \frac{1}{\sqrt{2\pi}\sigma_z^2} \exp\left[-\frac{(z-x)^2}{2\sigma_z^2}\right] \quad (6)$$

When the model is built, the state is fixed, and the distribution is then a function of z . When the observations are made, the distribution is a function of x . The prior $p(x)$ is given as

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma_x^2} \exp\left[-\frac{(x-x_p)^2}{2\sigma_x^2}\right] \quad (7)$$

Then, using the Bayes' rule, the posterior is given as follows after noting the observation:

$$\begin{aligned} p(x|z) &= \text{Const.} \frac{1}{\sqrt{2\pi}\sigma_z^2} \exp\left[-\frac{(z-x)^2}{2\sigma_z^2}\right] \frac{1}{\sqrt{2\pi}\sigma_x^2} \exp\left[-\frac{(x-x_p)^2}{2\sigma_x^2}\right] \\ &= \frac{1}{\sqrt{2\pi}\sigma_z^2} \exp\left[-\frac{(z-\bar{x})^2}{2\sigma^2}\right] \end{aligned} \quad (8)$$

Where

$$\bar{x} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2} z + \frac{\sigma_z^2}{\sigma_x^2 + \sigma_z^2} x_p \quad (9)$$

and

$$\sigma^2 = \frac{\sigma_x^2 \sigma_z^2}{\sigma_x^2 + \sigma_z^2} z = \left(\frac{1}{\sigma_z^2} + \frac{1}{\sigma_x^2} \right)^{-1} \quad (10)$$

From Bayes' rule, we obtain the following general application rule for the independent likelihood pool:

$$p(x|Z^n) = \{p(Z^n)\}^{-1} p(x) \prod_{i=1}^n p(z_i|x) \quad (11)$$

The conditional probabilities of $p(z/x)$ are stored “a priori” as functions of z and x . For fusion, it is assumed that the information obtained from different sources and sensors is independent. Only the underlying state is common between the sources. For a set of observations, we thus obtain:

$$p(x|Z^n) = \{p(Z^n)\}^{-1} p(Z^n|x)p(x) \quad (12)$$

$$p(x|Z^n) = \frac{p(z_1, \dots, z_n|x)p(x)}{p(z_1, \dots, z_n)} \quad (13)$$

The joint distribution of Z should be known completely. From the foregoing discussion, it is clear that Bayes' formula can be used for sensor DF.

B. Bayesian Method for Fusion of Data from Two Sensors

Here we assume that for sensor 1, the new set Z_1^l is obtained from the current measurement z_1^l and the old dataset Z_0^l , and for sensor 2, the new set Z_2^l is obtained from the current measurement z_2^l and the old dataset Z_0^l . Essentially, at the fusion node, the probability of x (a particular aircraft) being one of the three aircraft types is to be computed based on the latest set of data: $p(x|Z_1^l Z_2^l)$. Using Bayes' rule, we obtain the following relationship:

$$\begin{aligned} p(x|Z_1^1 Z_1^2) &= p(x|z_1^1 z_1^2 Z_0^1 Z_0^2) \\ &= \frac{p(z_1^1 z_1^2|x, Z_0^1 Z_0^2) p(x|Z_0^1 Z_0^2)}{p(z_1^1 z_1^2|Z_0^1 Z_0^2)} \end{aligned} \quad (14)$$

Because the sensor measurements are assumed independent, we obtain:

$$p(x|Z_1^1 Z_1^2) = \frac{p(z_1^1|x, Z_0^1) p(z_1^2|x, Z_0^2) p(x|Z_0^1 Z_0^2)}{p(z_1^1 z_1^2|Z_0^1 Z_0^2)} \quad (15)$$

On the basis of the Equation 15, we derive the equation

$$p(x|Z_1^1 Z_1^2) = \frac{p(x|Z_1^1) p(z_1^1|Z_0^1) p(x|Z_1^2) p(z_1^2|Z_0^2) p(x|Z_0^1 Z_0^2)}{p(x|Z_0^1) p(x|Z_0^2) p(z_1^1 z_1^2|Z_0^1 Z_0^2)} \quad (16)$$

Finally, at the fusion node, the posterior probability for the aircraft x is given as follows:

$$p(x|Z_1^1 Z_1^2) = \frac{p(x|Z_1^1) p(x|Z_1^2) p(x|Z_0^1 Z_0^2)}{p(x|Z_0^1) p(x|Z_0^2)} \quad (17)$$

Thus, Equation 17 yields the required fusion solution using the Bayesian approach.

V. RESULTS OF COMPUTER SIMULATIONS

We present here a non-linear model which will illustrate the operation of particle filter in recursive Bayesian networks. The non-linear model is represented by the following equations, assuming w_k and v_k are zero mean Gaussian white noise with variances 10 and 1 respectively and the initial state is considered at 0.1:

$$x_k = .5x_{k-1} + 25\left(\frac{x_{k-1}}{(1+x_{k-1}^2)}\right) + 8\cos(1.2k - 1.2) + w_k \quad (18)$$

$$y_k = \left(\frac{x_k^2}{20}\right) + v_k$$

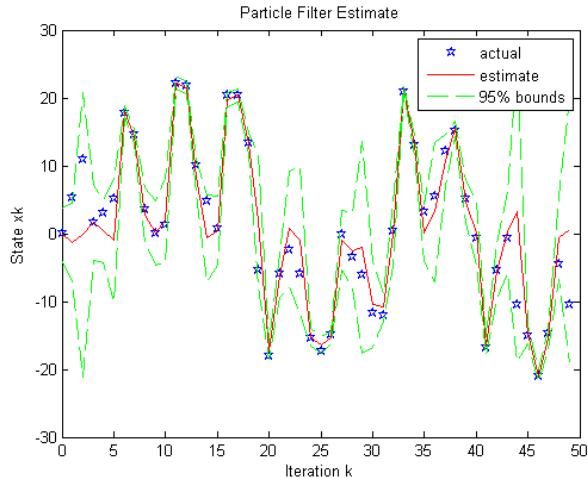


Figure 4: Particle Filter estimate of Posterior Mean

The Fig.4 gives the particle filter estimate about the posterior mean, where the state x_k is plotted against iteration number k. The blue points gives the actual value while the red solid line gives the plot of the estimate as decided by the

particle filter. The green dashed line gives value of the 95% bound region.

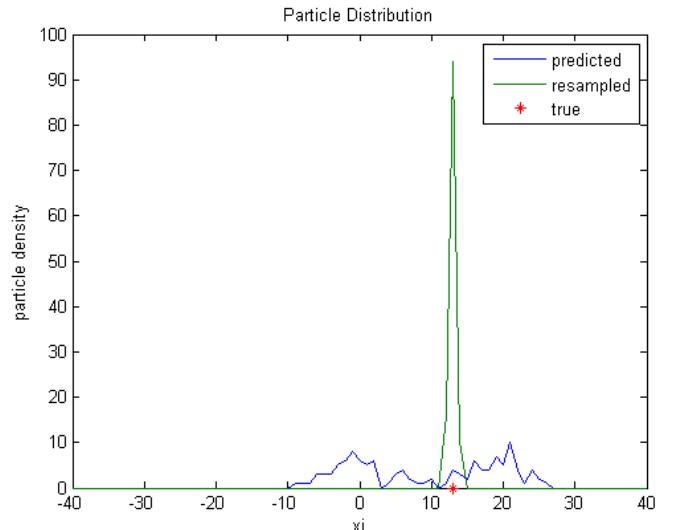


Figure 5: Particle Distribution

The Fig.5 above shows the probability distribution function (PDF) of the particle filter. Comparing Fig.1 and Fig.2, we see that although the green dashed lines in Fig.1 gives estimate of the 95% probability region, these may not represent the true 95% region. At a couple of time instants, the actual state is just outside these percentile estimates, and quite often, it is close to one of the limits.

The Fig. 6 displays the likelihood distribution and shows that it is bimodal, since there are two peaks.

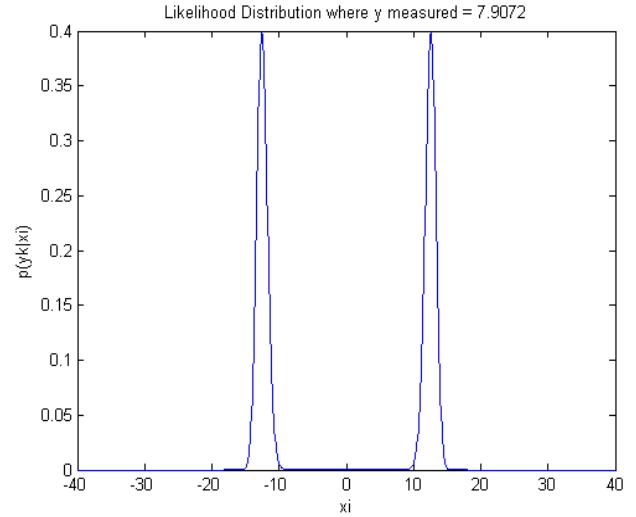


Figure 6: Likelihood Distribution

Next, we show the simulation in terms of actual object tracking in Figs. 7 and 8. The target moves in the x-y plane and two sensors are used in this simulation. The sensors are depicted by green circles while the true positions are shown in red dots. The particle filter displays the estimate of the track in blue solid lines.

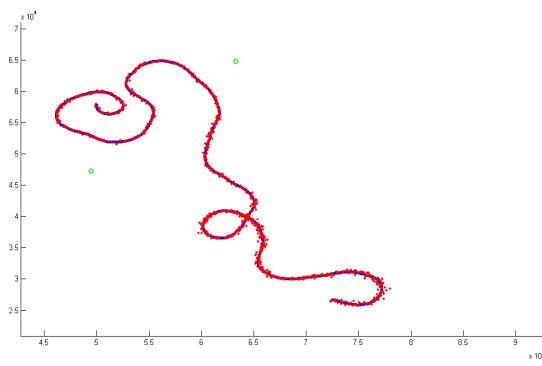


Figure 7: Movement of target at track1

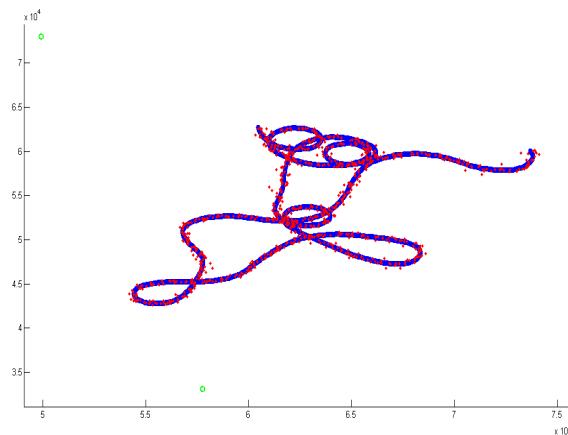


Figure 8: Movement of target at track2

VI. CONCLUSIONS:

In this paper, we presented the Bayesian method for data fusion from multiple sensors. The particle filter has been used for implementing the non-linear recursive Bayesian models. One of the primary advantages of this technique is that the method is not restricted by assumptions of linearity or Gaussian noise and this approach to non-linear Bayesian state estimation may be applied to any state transition or measurement model. The required posterior distribution is produced as a set of samples, and even considering a finite value of sample points, the approach is far superior to other standard filters, like the extended and unscented Kalman filter.

The estimate of the posterior mean shows a greater degree of convergence for the particle filter. Again, from the trajectory of the tracks, it can be commented that after an initial period of uncertainty, the particle filter quickly homes onto the target, unlike the case of the Extended Kalman filter which studies have shown to diverge.

However, further work with extensive Monte Carlo simulations is necessary to derive practical methods and to provide a quantitative assessment of filter performance for important non-linear problems.

REFERENCES

- [1] A. Doucet, N. D. Freitas, N. J. Gordon, (2001). "Sequential Monte Carlo Methods in Practice", Springer.
- [2] T. Flury and N. Shephard, "Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models," 2008, OFRC Working Papers Series 2008fe32, Oxford Financial Research Centre
- [3] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar. Next century challenges: scalable coordination in sensor networks. Proc. of Mobile Computing and Networking, pages 263–270, 1999.
- [4] W. R. Heinzelman, J. Kulik, and H. Balakrishnan. Adaptive protocols for information dissemination in wireless sensor networks. Proc. of Mobile Computing and Networking, pages 174–185, 1999.
- [5] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. B. Srivastava. Coverage problems in wireless ad-hoc sensor networks. Proceedings of IEEE Infocom, 3:1380–1387, April 2001.
- [6] M. Chu, H. Haussecker, and F. Zhao. Scalable information-derived sensor querying and routing for ad hoc heterogeneous sensor networks. International Journal of High-Performance Computing Applications, 16(3):293–313, 2002.
- [7] I. Akyildiz, W. Su, Y. Sankarasubramniam, and E. Cayirci. A survey on sensor networks. IEEE Communications Magazine, pages 102–114, August 2002.
- [8] J. Cortes, S. Martinez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. The IEEE Trans. on Robotics and Automation, 20(2):243–255, April 2004.
- [9] S. Martinez and F. Bullo. Optimal sensor placement and motion coordination for target tracking. Automatica (submitted), 2004.
- [10] J. R. Raol, "Multi-Sensor Data Fusion with MATLAB®" © 2010 by Taylor and Francis Group, LLC.
- [11] G. Girija, J R Raol, R. A. Raj and S. Kashyap, "Tracking filter and multi-sensor data fusion" Sādhānā, Vol. 25, Part 2, April 2000, pp. 159–167. ©Printed in India.
- [12] D. L. Hall, Senior Member, IEEE, And J. Llinas, "An Introduction to Multisensor Data Fusion" © 1997 IEEE, Proceedings Of The IEEE, Vol. 85, No. 1, January 1997.
- [13] H. F. Durrant-Whyte, Australian Centre for Field Robotics The University of Sydney NSW 2006 Australia, "Multi Sensor Data Fusion" © Hugh Durrant-Whyte 2001.



Anindya Mondal (M'87) was born in Howrah on 12th May 1987. He had completed B.Sc in PHYSICS from Narasinha Dutt College, Howrah, West Bengal under University of Calcutta. He had completed M.Sc in ELECTRONIC SCIENCE from Dinabandhu Andrews College, Kolkata, West Bengal under University of Calcutta. He is at present final year M.Tech student in Electronics & Communication Engineering Dept. at Heritage Institute of Technology Kolkata, West Bengal under West Bengal University of Technology



Moloy Kumar Chowdhury (M'78) was born in Kolkata on 17th December, 1978. He has done B.E. in Electronics & Communication Engg. Dept. from Sikkim Manipal Institute of Technology in 2001 and pursued M.E. in Electronics and Telecommunication Engg. Dept. from Jadavpur University in 2010. He was a faculty at M.C.K.V. Institute of Engg. for almost five years and now currently serving as Assistant Professor in the Electronics & Comm. Engg. Dept. at Heritage Institute of Technology, Kolkata, West Bengal, India. He had secured 96 percentile in GATE-2008 in Electronics & Communication Engg. He has published a few papers in several International

2nd Annual International Conference IEMCON 2012

and National Conferences. He has currently submitted a chapter proposal on Ultra Wideband for book publication with Intech Publications. His research interests include Wireless mobile communications systems, Ultra Wideband, MIMO, Multi-Sensor Data Fusion, Software Defined Radio, Fuzzy applications.

Mr. Chowdhury had organized IESPC-2011 (IEEE EDS Student Paper Conference) at Heritage Institute of Technology, Kolkata, West Bengal, India in his capacity as Co-convenor in 2011.

FINANCE & MANAGEMENT

Date : 18th January,2012

Chairman : Prof. D.K.Banwet

Co-Chairman : Prof. Rahul Gupta

Rapporter : Prof. Paramita Mukherjee

| Paper ID | Paper Title | Authors |
|----------|---|--|
| 3 | A Study on Common Finance Engineering in India | N.Singha Roy,S.S. Roy, S.B.Chowdhury and Dr. Subhamoy singha Roy |
| 6 | Analysing Existing Online Business Models and Suggesting an Ideally Secured and Integrated Website Layout | Reetwika Banerjee and Tamal Kanti Mukherjee |
| 10 | Fuzzy Mixed Integer Linear Programming for Air Vehicles Operations Optimization | Dr Arindam Chaudhuri, Dr Dipak Chatterjee and Ritesh Rajput |
| 82 | Cloud Computing: emphasizing its possible roles and importance in Information Systems and Centers | Prantosh Kumar Paul, Dr. Dipok Chatterjee and Bhaskar Karn |
| 83 | Information Science Education and Research | Prantosh Kumar Paul, Dr. Dipok Chatterjee and Bhaskar Karn |
| 85 | A Simple Approach of Building a Learning Managemnt System(LMS) to Expeiment with Intelligent Tutoring | Lopa Mandal, Avik Dutta, Rahul Roy and Mohit Daga |
| 89 | FII's and Indian Stock Markets- Emeging Issues From A Literatur Review Post-2000 | Ranjan Dasgupta and Sujit Dutta |

Fuzzy Mixed Integer Linear Programming for Air Vehicles Operations Optimization

Arindam Chaudhuri
Assistant Professor,
Computer Science Engineering,
NIIT University,
Neemrana, Rajasthan
arindam_chau@yahoo.co.in

Dipak Chatterjee
Principal and Professor,
Mathematics,
Institute of Engineering and
Management,
Kolkata, West Bengal

Ritesh Rajput
BTech Student,
3rd Year,
Computer Science Engineering,
NIIT University,
Neemrana, Rajasthan

Abstract—Multiple Air Vehicles (AVs) to prosecute geographically dispersed targets is an important optimization problem. Associated multiple tasks viz., target classification, attack and verification are successively performed on each target. The optimal minimum time performance of these tasks requires cooperation among vehicles such that critical time constraints are satisfied i.e. target must be classified before it can be attacked and AV is sent to target area to verify its destruction after target has been attacked. Here, optimal task scheduling problem from Indian Air Force is formulated as Fuzzy Mixed Integer Linear Programming (FMILP) problem. The solution assigns all tasks to vehicles and performs scheduling in an optimal manner including scheduled staged departure times. Coupled tasks involving time and task order constraints are addressed. When AVs have sufficient endurance, existence of optimal solution is guaranteed. The solution developed can serve as an effective heuristic for different categories of AV optimization problems.

Keywords—Fuzzy Mixed Integer Linear Programming; Air Vehicles; Planning; Scheduling

I. INTRODUCTION

Unmanned Air Vehicle (UAV) is an aircraft with no onboard pilot which can be remote controlled or fly autonomously based on pre-programmed flight plans with complex dynamic automation systems. Optimizing air to ground operations of such air vehicles is an important decision making problem [7]. A more challenging scenario is multiple UAVs required to service geographically dispersed targets. Multiple tasks are required to be successively performed on each target, viz. targets to be classified, attacked and damage inflicted on targets must be assessed. The floating time constraints are critical. A target cannot be attacked before it is classified and UAV sent to target area only after target attack has been executed. Multi role UAVs are considered such that each UAV can perform all tasks.

Small UAVs such as autonomous Wide Area Search Munitions (WASM) are deployed in groups from aircraft flying at higher altitudes. They are individually capable of autonomous searching, recognizing and attacking targets. The ability to network involving communication of target information to one another and consequent co-operation greatly improves effectiveness of future UAV teams. The problem comprises of planning performance of UAVs' tasks such that critical timing constraints are satisfied. This calls for optimal assignment and scheduling [1].

In a time phased network optimization model was used to perform task allocation for team of UAVs [9]. The model is run simultaneously on all air vehicles at discrete points in time and assigns each vehicle one or more tasks each time it is executed. The network optimization model is iteratively implemented such that all known targets are prosecuted by resulting allocation. The model is solved each time new information is brought in the system, because a new target has been discovered or known target's status is changed, thus achieving feedback action. Classification, attack and battle damage assessment tasks can all be assigned to different vehicles when a target is found resulting in target being more quickly serviced. A single vehicle can be given multiple task assignments to be performed in succession if that is more efficient than having multiple vehicles perform tasks individually. In [9], variable path lengths are included to guarantee that feasible trajectories calculated for all tasks. This method is computationally efficient and scales well. However, the iterative procedure is heuristic and suboptimal. Tabu Search can be used to solve difficult combinatorial optimization problems such as vehicle routing problem with fixed time windows [2], [3], [4], [5], [11].

In this work, a solution method for AV routing combinatorial optimization problem with floating time windows from Indian Air Force is developed. The work addresses optimal formulation for solving coupled

multiple-assignment and scheduling problem. Time variables used are continuous in nature. They are modeled using fuzzy membership functions to tackle the inherent uncertainty involved. This leads to the formulation of AV optimization problem as Fuzzy Mixed Integer Linear Programming (FMILP) problem [8], [12] and allows optimal solution to be obtained while satisfying all time constraints. The entire FMILP problem formulation is developed using MATLAB. The mathematical formulation presented can be used to solve optimally for some realistic problem sizes. The method presented here allows staged departure times before vehicles begin their tour of targets and tasks. FMILP formulation is flexible enough to allow consideration of different cost functions such as, mission completion in minimum time, shortest total path lengths travelled by vehicles or maximization of number of AV that survive the mission. The method also accommodates fixed time windows as in vehicle routing problem [2], [5], [11]. FMILP formulation also incorporates dynamic and logical constraints on task performance as in scheduling problems [7]. Without using fixed time windows, feasibility is guaranteed as long as number of air vehicles exceeds number of targets, even with three or more tasks per target provided air vehicles have sufficient endurance.

This paper is organized as follows. In section II, UAV optimization problem is given. This is followed by mathematical formulation of AV optimization problem in section III. In the next section, experimental results are given. In section V, a discussion on problem size is illustrated. Finally, in section VI conclusions are given.

II. UNMANNED AIR VEHICLES OPTIMIZATION PROBLEM

In this section AV optimization problem scenario is presented which is adapted from Indian Air Force. Consider n geographically dispersed targets with known position w ($w \geq n+1$) air vehicles (AV). We then have $n+w+1$ nodes: n target nodes, w source nodes (points of origins of AVs) and one sink node. Nodes $1, \dots, n$ are located at n target positions. Nodes $n+1, \dots, n+w$ are located at vehicle initial positions. Node $n+w+1$ is sink. An AV with no future target assignments is relegated to sink i.e. will continue to search. A vehicle located at sink cannot be reassigned.

The flight time of AV v from node i to node j to perform task k at node j is:

$$\tilde{t}_{ij}^{(v,k)} \geq 0; i = 1, \dots, n+w; j = 1, \dots, n; v = 1, \dots, w \quad (1)$$

which is modeled as the following fuzzy membership function [10]:

$$\tilde{t}_{ij}^{(v,k)} = \begin{cases} 1 + (z_j - c_{ij}) / b_{ij}^-, & -b_{ij}^- \leq (z_j - c_{ij}) \leq 0 \\ 1 - (z_j - c_{ij}) / b_{ij}^+, & 0 \leq (z_j - c_{ij}) \leq b_{ij}^+ \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Here, z_j is j^{th} input of i^{th} fuzzy membership function, c_{ij} is modal point, b_{ij}^- its lower half width and b_{ij}^+ upper half width. The membership function attains value 1 when input is c_{ij} . As input decreases from c_{ij} membership function value decreases to 0 at $c_{ij} - b_{ij}^-$ and remains at 0 for all inputs less than $c_{ij} - b_{ij}^-$. As input increases from c_{ij} , membership function value decreases to 0 at $c_{ij} + b_{ij}^+$ and remains at 0 for all inputs greater than $c_{ij} + b_{ij}^+$. The obvious choice of this membership function arises because there is only one output here viz. time.

The time to travel from node i to node j depends on particular AV v 's airspeed and assigned task k . Three tasks must be performed on each target viz. (a) classification (b) attack (c) target damage assessment (verification). Furthermore, once an AV attacks a target, it is destroyed and can no longer perform additional tasks. This is certainly the case for powered munitions and WASM mission, but if AV was a reusable aircraft, the problem formulation would have to be modified and account for AV's depletion of its store of ammunition following each attack. Three tasks must be performed on each target in order listed. This results in critical timing constraints. The problem inherits some aspects of job shop scheduling [7]. In optimization problem considered, the number of pre-specified problem parameters is $3wn + 3n(n-1)w = 3n^2w$. When Euclidean

distances are used the dimension of parameter space is reduced appreciably. The endurance of AV v is $T_v, v = 1, \dots, w$.

III. MATHEMATICAL FORMULATION OF AIR VEHICLES OPTIMIZATION PROBLEM

In this section, the mathematical formulation of AV optimization problem given in section 2 is illustrated. The problem is modeled using FMILP technique. The FMILP model uses a discrete approximation of real world based on nodes that represent start and end positions for segments of UAVs path. Nodes representing target positions range from $1, \dots, n$ and nodes for initial UAV positions range from $n+1, \dots, n+w$. There is also an additional node for sink $n+w+1$. The sink node is used when no further assignment for UAV is present. It goes to sink when it has completed all its tasks or when not assigned to another task. Generally, when a UAV enters sink it is used for performing search of battle space. FMILP model requires information on costs or times for UAV to fly from one node to another node. The flight times are represented by $\tilde{t}_{ij}^{(v,k)} \geq 0$ as given in Equation (2), the time it takes UAV v to fly from node i to node j to perform task k .

A. Decision Variables

The binary decision variable $x_{ij}^{(v,k)} = 1$ if AV v is assigned to fly from node i to node j to perform task k at node j and 0 otherwise; $i = 1, \dots, n+w; j = 1, \dots, n; v = 1, \dots, w; k = 1, 2, 3$. For $i = n+1, \dots, n+w$, only $x_{n+v,j}^{(v,k)}$ exist. These variables correspond with first task assignment each vehicle receives starting from its unique source node. Only vehicle v can do a task starting from node $n+v$. For task assignments $k = 1, 3, i \neq j$ and for task assignment $k = 2, i = j$ is allowed. The later allows for an AV to perform target classification task and immediately thereafter attack the target. Thus, there are $wn(3n+1)$ binary decision variables. Another

decision variable $x_{i,n+w+1}^{(v,k)} = 1$ if AV v is assigned to fly node i to sink node $n+w+1$ and 0 otherwise $v = 1, \dots, w; i = 1, \dots, n+w$. This adds $w(n+1)$ binary decision variables. Entering sink can also be thought of as being reassigned to search task. Time of performance of task k on target j is $\tilde{t}_j^{(k)} > 0; k = 1, 2, 3; j = 1, \dots, n$ which is modeled as the following fuzzy membership function [6]:

$$\tilde{t}_j^{(k)} = \begin{cases} 0 & , x = a \\ x/(x_i - a) + a/(a - x_i) & , a \leq x \leq x_i \\ 0 & , x > b \end{cases} \quad (3)$$

These are continuous decision variables which are $3n$ in number. There are also w additional continuous decision variables: time AV v leaves node $j = n+v$ is $\tilde{t}_v; v = 1, \dots, w$ which is also modeled as fuzzy membership function given by Equation (3). There are $w[n(3n+2)+1]$ and $3n+w$ binary and continuous non-negative decision variables.

B. Cost Functions

The cost functions involved for AV optimization problem include the following:

(a) Minimize the total flight time of the AVs:

$$J = \sum_{k=1}^3 \sum_{v=1}^w \sum_{i=1}^{n+w} \sum_{j=1}^n \tilde{t}_{i,j}^{(v,k)} x_{i,j}^{(v,k)} \quad (4)$$

(b) Alternatively, minimize total engagement time. Target j is visited for last time at time $\tilde{t}_j^{(3)}$. Let \tilde{t}_f be the time at which all targets have been through verification. Considering an additional continuous decision variable $\tilde{t}_f \in \mathbb{R}_+$. The cost function is then $J = tf$ and we minimize J subject to constraints

$$\tilde{t}_j^{(3)} \leq \tilde{t}_f; j = 1, \dots, n \quad (5)$$

A small weightage to time of performance of each individual task is added to encourage each task to be completed as quickly as possible. Then,

$$J = \tilde{t}_f + \tilde{c}_j^{(k)} \tilde{t}_j^{(k)}; j = 1, \dots, n; k = 1, 2, 3 \quad (6)$$

where, $\tilde{c}_j^{(k)} > 0$ is small weight on completion time of each individual task and modeled as fuzzy membership function given by Equation (3).

- (c) The number n_s of surviving UAVs that end up in sink and the corresponding cost function is given by,

$$J = n_s = \sum_{v=1}^w \sum_{i=1}^{n+w} x_{i,n+w+1}^{(v)} \quad (7)$$

and optimization problem considered is $\max J = n_s$.

C. Constraints

The corresponding constraints for AV optimization problem require inclusion of all issues which are critical to automatically enforcing desired vehicle behavior.

- (a) Mission completion constraints: It requires that all three tasks are performed on each target exactly one time. The following must hold as evident for linear assignment problems [1], [7] yielding $3n$ constraints.

$$\sum_{v=1}^w \sum_{i=1, i \neq j}^{n+w} x_{ij}^{(v,k)} = 1 \quad k = 1, 3; \quad j = 1, \dots, n \quad (8)$$

$$\sum_{v=1}^w \sum_{i=1, i \neq j}^{n+w} x_{ij}^{(v,k)} = 1 \quad j = 1, \dots, n \quad (9)$$

- (b) Not more than one AV is assigned to perform a specific task k on a specified target j . This also yields $3n$ constraints.

$$(c) \quad \sum_{i=1, i \neq j}^{n+w} x_{ij}^{(v,k)} \leq 1 \quad v = 1, \dots, w; \quad j = 1, \dots, n; \quad k = 1, 3 \quad (10)$$

$$\sum_{i=1}^{n+w} x_{ij}^{(v,2)} \leq 1$$

$$v = 1, \dots, w; \quad j = 1, \dots, n \quad (11)$$

- (c) AV v coming from outside can visit target j at most once.

$$\sum_{k=1}^3 \sum_{i=1, i \neq j}^{n+w} x_{ij}^{(v,k)} \leq 1$$

$$v = 1, \dots, w; \quad j = 1, \dots, n \quad (12)$$

In addition each AV v can only enter the sink once. This yields $w(n+1)$ constraints.

$$\sum_{i=1}^{n+w} x_{i,n+w+1}^{(v)} \leq 1 \quad v = 1, \dots, w \quad (13)$$

- (d) An AV v leaves node j at most once yielding nw constraints.

$$\sum_{k=1}^3 \sum_{i=1, i \neq j}^n x_{ij}^{(v,k)} + x_{j,n+w+1}^{(v)} \leq 1$$

$$v = 1, \dots, w; \quad j = 1, \dots, n \quad (14)$$

The constraints (c) and (d) eliminate the possibility of loops.

- (e) Perishable Munition: An AV v can be assigned to attack at most one target yielding w constraints.

$$\sum_{j=1}^n \sum_{i=1}^{n+w} x_{ij}^{(v,2)} \leq 1 \quad \forall v = 1, \dots, w \quad (15)$$

- (f) If AV v is assigned to fly to target j for verification. It cannot possibly be assigned to attack target j .

$$\sum_{i=1, i \neq j}^{n+w} x_{ij}^{(v,2)} \leq 1 - \sum_{i=1, i \neq j}^{n+w} x_{ij}^{(v,3)}$$

$$v = 1, \dots, w; \quad j = 1, \dots, n \quad (16)$$

- (g) Continuity Constraints: These constraints ensure proper flow balance is maintained at each node.

- (i) If AV v enters target node j for purpose of performing task 3 it must also exit target node j .

$$\sum_{i=1, i \neq j}^{n+w} x_{ij}^{(v,3)} \leq \sum_{k=1}^3 \sum_{i=1, i \neq j}^n x_{ji}^{(v,k)} + x_{j,n+w+1}^{(v)}$$

$$v = 1, \dots, w; \quad j = 1, \dots, n \quad (17)$$

- (ii) If AV v enters target node j for purpose of performing task 1 it must also exit target node j .

$$\sum_{i=1, i \neq j}^{n+w} x_{ij}^{(v,1)} \leq \sum_{k=1}^3 \sum_{i=1, i \neq j}^n x_{ji}^{(v,k)} + x_{j,j}^{(v,2)} + x_{j,n+w+1}^{(v)}$$

$$v = 1, \dots, w; \quad j = 1, \dots, n \quad (18)$$

(iii) As a munition is perishable if AV v is assigned to fly to target node j to perform task $k = 2$ then at any other point in time, AV v cannot be assigned to fly from target j to target $i, i \neq j$ to perform any other task k at target i . AV v can enter target j not more than once.

$$\sum_{k=1}^3 \sum_{i=1, i \neq j}^n x_{j,i}^{(v,k)} + x_{j,n+w+1}^{(v)} \leq 1 - \sum_{i=1}^{n+w} x_{i,j}^{(v,2)}$$

$$v = 1, \dots, w; j = 1, \dots, n \quad (19)$$

(iv) If AV v is not assigned to visit node j , then it cannot be assigned to fly out of node j .

$$\sum_{k=1}^3 \sum_{i=1, i \neq j}^n x_{i,j}^{(v,k)} + x_{j,n+w+1}^{(v)} \leq \sum_{k=1}^3 \sum_{i=1, i \neq j}^{n+w} x_{i,j}^{(v,k)}$$

$$v = 1, \dots, w; j = 1, \dots, n \quad (20)$$

(v) All AVs leave source nodes. An AV leaves source node even if this entails direct assignment to sink.

$$\sum_{k=1}^3 \sum_{j=1}^n x_{n+v,j}^{(v,k)} + x_{n+v,n+w+1}^{(v)} = 1$$

$$\forall v = 1, \dots, w \quad (21)$$

(vi) An AV cannot attack target node i coming from target node i unless it entered target node i to perform classification.

$$x_{i,i}^{(v,2)} \leq \sum_{j=1}^{n+w} x_{j,i}^{(v,1)}$$

$$\forall i = 1, \dots, n; v = 1, \dots, w \quad (22)$$

(h) Timing constraints: Nonlinear equations enforcing timing constraints can be formulated represented as FMILP. Thus, assuming

$$T \equiv \max_v \{\tilde{T}_v\}_{v=1}^w \quad (23)$$

Then the linear timing constraints take the following form:

$$\tilde{t}_j^{(k)} \leq \tilde{t}_i^{(1)} + \tilde{t}_{i,j}^{(v,k)} + \left(2 - x_{i,j}^{(v,k)} - \sum_{l=1, l \neq i}^{n+w} x_{l,i}^{(v,1)} \right) wT$$

$$(24)$$

$$\tilde{t}_j^{(k)} \geq \tilde{t}_i^{(1)} + \tilde{t}_{i,j}^{(v,k)} - \left(2 - x_{i,j}^{(v,k)} - \sum_{l=1, l \neq i}^{n+w} x_{l,i}^{(v,1)} \right) wT$$

$$(25)$$

$$\tilde{t}_j^{(k)} \leq \tilde{t}_i^{(3)} + \tilde{t}_{i,j}^{(v,k)} + \left(2 - x_{i,j}^{(v,k)} - \sum_{l=1, l \neq i}^{n+w} x_{l,i}^{(v,3)} \right) wT \quad (26)$$

$$\tilde{t}_j^{(k)} \geq \tilde{t}_i^{(3)} + \tilde{t}_{i,j}^{(v,k)} - \left(2 - x_{i,j}^{(v,k)} - \sum_{l=1, l \neq i}^{n+w} x_{l,i}^{(v,3)} \right) wT$$

$$(27)$$

$$\forall i = 1, \dots, n; j = 1, \dots, n; i \neq j;$$

$$v = 1, \dots, w; k = 1, 3$$

In addition to this we also have the following timing constraints:

$$\tilde{t}_j^{(2)} \leq \tilde{t}_i^{(1)} + \tilde{t}_{i,j}^{(v,2)} + \left(2 - x_{i,j}^{(v,2)} - \sum_{l=1, l \neq i}^{n+w} x_{l,i}^{(v,1)} \right) wT$$

$$(28)$$

$$\tilde{t}_j^{(2)} \geq \tilde{t}_i^{(1)} + \tilde{t}_{i,j}^{(v,2)} - \left(2 - x_{i,j}^{(v,2)} - \sum_{l=1, l \neq i}^{n+w} x_{l,i}^{(v,1)} \right) wT$$

$$(29)$$

$$\tilde{t}_j^{(2)} \leq \tilde{t}_i^{(3)} + \tilde{t}_{i,j}^{(v,2)} + \left(2 - x_{i,j}^{(v,2)} - \sum_{l=1, l \neq i}^{n+w} x_{l,i}^{(v,3)} \right) wT$$

$$(30)$$

$$\tilde{t}_j^{(2)} \geq \tilde{t}_i^{(3)} + \tilde{t}_{i,j}^{(v,2)} - \left(2 - x_{i,j}^{(v,2)} - \sum_{l=1, l \neq i}^{n+w} x_{l,i}^{(v,3)} \right) wT$$

$$(31)$$

$$\forall j = 1, \dots, n; k = 1, 2, 3; v = 1, \dots, w$$

Also

$$\tilde{t}_j^{(k)} \leq \tilde{t}_v + \tilde{t}_{n+v,j}^{(v,k)} + (1 - x_{n+v,j}^{(v,k)}) wT$$

$$(32)$$

$$\tilde{t}_j^{(k)} \leq \tilde{t}_v + \tilde{t}_{n+v,j}^{(v,k)} - (1 - x_{n+v,j}^{(v,k)}) wT$$

$$(33)$$

$$\forall j = 1, \dots, n; k = 1, 2, 3; v = 1, \dots, w$$

These timing constraints operate in pairs. The inequalities (32) – (33) are not validated for assignments $x_{i,j}^{(v,k)}$ that do not occur but become soft equality constraint for assignments that occur. Thus, the time that a task k is performed on target j by AV v will be equal to the time that preceding task was performed by AV v at node i plus the time it will take AV v to fly from node i to node j . A similar constraint applies if AV v left its source node $n+v$ to fly to node j . Also

$$t_j^{(1)} \leq t_j^{(2)} < t_j^{(3)}; j = 1, \dots, n$$

$$(i) \quad (34)$$

The timing constraints add $2n [w(6n - 1) + 1]$ linear inequality constraints. The constraints (24) – (34) are critical for FMILP formulation of AV optimization problem.

IV. EXPERIMENTAL RESULTS

In this section one real life example from Indian Air Force comprising of three UAVs and one Target to illustrate the mathematical formulation of AV optimization problem presented in section III.

A. Three UAVs and One Target

We consider the case of three AVs and one target i.e. $w = 3$ and $n = 1$. Though the number of variables in the problem is small enough, this can easily be extended to large number of variables at the cost of additional computational complexity. There are 18 binary and 6 continuous decision variables which are given in equations below. Minimizing time the final task occurs will add an additional continuous decision variable for a total of 25 decision variables.

$$\begin{aligned} (x_1, \dots, x_5) &= (x_{1,1}^{(1,2)}, x_{1,1}^{(2,2)}, x_{1,1}^{(3,2)}, x_{2,1}^{(1,1)}, x_{2,1}^{(1,2)}) \\ (x_6, \dots, x_{10}) &= (x_{2,1}^{(1,3)}, x_{3,1}^{(2,1)}, x_{3,1}^{(2,2)}, x_{3,1}^{(2,3)}, x_{4,1}^{(3,1)}) \\ (x_{11}, \dots, x_{15}) &= (x_{4,1}^{(3,2)}, x_{4,1}^{(3,3)}, x_{1,5}^{(1)}, x_{1,5}^{(2)}, x_{1,5}^{(3)}) \\ (x_{16}, x_{17}, x_{18}) &= (x_{2,5}^{(1)}, x_{3,5}^{(2)}, x_{4,6}^{(3)}) \end{aligned} \quad (35)$$

$$(x_{19}, \dots, x_{25}) = (\tilde{t}_1, \tilde{t}_2, \tilde{t}_3, \tilde{t}_1^{(1)}, \tilde{t}_1^{(2)}, \tilde{t}_1^{(3)}, \tilde{t}) \quad (36)$$

We wish to minimize

$$J = x_{25} + 0.1(x_{22} + x_{23} + x_{24}) \quad (37)$$

subject to following constraints:

Constraint 1:

$$\begin{aligned} x_4 + x_7 + x_{10} &= 1 & x_6 + x_9 + x_{12} &= 1 \\ x_1 + x_2 + x_3 + x_5 + x_8 + x_{11} &= 1 \end{aligned} \quad (38)$$

Constraint 2:

$$\begin{aligned} x_4 + x_7 + x_{10} &\leq 1 & x_6 + x_9 + x_{12} &\leq 1 \\ x_1 + x_2 + x_3 + x_5 + x_8 + x_{11} &\leq 1 \end{aligned} \quad (39)$$

Constraint 3:

$$\begin{aligned} x_4 + x_5 + x_6 &\leq 1 & x_7 + x_8 + x_9 &\leq 1 \\ x_{10} + x_{11} + x_{12} &\leq 1 \end{aligned} \quad (40)$$

Constraint 7(a):

$$\begin{aligned} x_6 &\leq x_{13} & x_9 &\leq x_{14} \\ x_{12} &\leq x_{15} \end{aligned} \quad (41)$$

Constraint 7(b):

$$\begin{aligned} x_4 &\leq x_1 + x_{13} & x_7 &\leq x_2 + x_{14} \\ x_{10} &\leq x_3 + x_{15} \end{aligned} \quad (42)$$

Constraint 7(c):

$$\begin{aligned} x_1 + x_5 + x_{13} &\leq 1 & x_2 + x_8 + x_{14} &\leq 1 \\ x_3 + x_{11} + x_{15} &\leq 1 \end{aligned} \quad (43)$$

Constraint 7(d):

$$\begin{aligned} x_{13} &\leq x_4 + x_6 & x_{14} &\leq x_7 + x_9 \\ x_{15} &\leq x_{10} + x_{12} \end{aligned} \quad (44)$$

Constraint 7(e):

$$\begin{aligned} x_4 + x_5 + x_6 + x_{16} &= 1 \\ x_7 + x_8 + x_9 + x_{17} &= 1 \\ x_{10} + x_{11} + x_{12} + x_{18} &= 1 \end{aligned} \quad (45)$$

Constraint 7(f):

$$\begin{aligned} x_1 &\leq x_4 & x_2 &\leq x_7 \\ x_3 &\leq x_{10} \end{aligned} \quad (46)$$

It is to be noted that constraints (d)–(f) are eliminated in one-target case. With only one target, constraints associated with Equations (24)–(27), (30) and (31) are ineffective, which results in following timing constraints. From Equations (28) and (29),

$$\begin{aligned} x_{23} &\leq x_{22} + \tilde{t}_{1,1}^{(1,2)} + (2 - x_1 - x_4)wT \\ x_{23} &\geq x_{22} + \tilde{t}_{1,1}^{(1,2)} - (2 - x_1 - x_4)wT \\ x_{23} &\leq x_{22} + \tilde{t}_{1,1}^{(2,2)} + (2 - x_2 - x_7)wT \\ x_{23} &\geq x_{22} + \tilde{t}_{1,1}^{(2,2)} - (2 - x_2 - x_7)wT \\ x_{23} &\leq x_{22} + \tilde{t}_{1,1}^{(3,2)} + (2 - x_3 - x_{10})wT \\ x_{23} &\geq x_{22} + \tilde{t}_{1,1}^{(3,2)} - (2 - x_3 - x_{10})wT \end{aligned} \quad (47)$$

From Equations (32) and (33),

$$\begin{aligned} x_{22} &\leq x_{19} + \tilde{t}_{2,1}^{(1,1)} + (1 - x_4)wT \\ x_{22} &\geq x_{19} + \tilde{t}_{2,1}^{(1,1)} - (1 - x_4)wT \end{aligned}$$

$$\begin{aligned} x_{22} &\leq x_{20} + \tilde{t}_{3,1}^{(2,1)} + (1-x_7)wT \\ x_{22} &\geq x_{20} + \tilde{t}_{3,1}^{(2,1)} - (1-x_7)wT \end{aligned} \quad (48)$$

$$\begin{aligned} x_{22} &\leq x_{21} + \tilde{t}_{4,1}^{(3,1)} + (1-x_{10})wT \\ x_{22} &\geq x_{21} + \tilde{t}_{4,1}^{(3,1)} - (1-x_{10})wT \end{aligned}$$

and

$$\begin{aligned} x_{23} &\leq x_{19} + \tilde{t}_{2,1}^{(1,2)} + (1-x_5)wT \\ x_{23} &\geq x_{19} + \tilde{t}_{2,1}^{(1,2)} - (1-x_5)wT \\ x_{23} &\leq x_{20} + \tilde{t}_{3,1}^{(2,2)} + (1-x_8)wT \end{aligned}$$

$$x_{23} \geq x_{20} + \tilde{t}_{3,1}^{(2,1)} - (1-x_8)wT \quad (49)$$

$$\begin{aligned} x_{23} &\leq x_{21} + \tilde{t}_{4,1}^{(3,2)} + (1-x_{11})wT \\ x_{22} &\geq x_{21} + \tilde{t}_{4,1}^{(3,2)} - (1-x_{11})wT \end{aligned}$$

and finally

$$\begin{aligned} x_{24} &\leq x_{19} + \tilde{t}_{2,1}^{(1,3)} + (1-x_6)wT \\ x_{24} &\geq x_{19} + \tilde{t}_{2,1}^{(1,3)} - (1-x_6)wT \\ x_{24} &\leq x_{20} + \tilde{t}_{3,1}^{(2,3)} + (1-x_9)wT \end{aligned}$$

$$x_{24} \geq x_{20} + \tilde{t}_{3,1}^{(2,3)} - (1-x_9)wT \quad (50)$$

$$\begin{aligned} x_{24} &\leq x_{21} + \tilde{t}_{4,1}^{(3,3)} + (1-x_{12})wT \\ x_{24} &\geq x_{21} + \tilde{t}_{4,1}^{(3,3)} - (1-x_{12})wT \end{aligned}$$

Also from Equation (34), we have

$$x_{22} \leq x_{23} - \varepsilon \quad (51)$$

$$x_{23} \leq x_{24} - \varepsilon \quad (52)$$

where, $\varepsilon > 0$ is small positive constant which is modeled using fuzzy membership function. Considering $\varepsilon = 0.1$ we enforce a small delay between each task being performed on a target. Finally, from Equation (5) we have,

$$x_{24} \leq x_{25} \quad (53)$$

Thus, full set of constraints contain 6 constraints, 51 inequality constraints for 57 total constraints. Assuming a simplifying situation that time to travel from node i to j to perform task k is independent of which task is required and vehicle performing the task. Then, $\tilde{t}_{i,j}^{(v,k)} = \tilde{t}_{i,j}$. For example,

$$t_{1,1} = 0.11 \quad t_{2,1} = 3.8 \quad t_{3,1} = 4.24 \quad t_{4,1} = 5.38$$

We set $T = 100$ as endurance of all AVs, so that endurance is not a constraint and feasibility is guaranteed. Then optimal assignment is as follows:

$$\begin{aligned} x_i &= 1; & i &= 1,4,10,14,18 \\ x_i &= 0; & i &= 2,3,5,\dots,8,9,\dots,13,15,16,17 \\ x_i &= 0; & i &= 19,20,21 \\ x_{22} &= 3.7 & x_{23} &= 3.81 \\ x_{24} &= 4.24 & x_{25} &= 4.24 \end{aligned}$$

This corresponds with all three vehicles immediately leaving their source nodes ($x_{19} - x_{21} = 0$) and vehicle 1 performing classification and attack on target at $t = 3.7$ and $t = 3.81$ respectively with vehicle 2 performing verification at $T = 4.24$. Vehicle 3 flies in direction to sink. Suppose it takes longer for a vehicle that has just classified a target to complete an attack on that target. Then we have the initial conditions as:

$$t_{1,1} = 1.1 \quad t_{2,1} = 3.8 \quad t_{3,1} = 4.24 \quad t_{4,1} = 5.38$$

In this case, assignment is identical, except that attack occurs at $t = 4.64$ and verification at $t = 4.69$. This example illustrates the situation where verification had to be delayed such that it took place after attack. Finally, suppose that vehicle 3 is closer to target initially, the initial conditions are:

$$t_{1,1} = 1.1 \quad t_{2,1} = 3.8 \quad t_{3,1} = 4.24 \quad t_{4,1} = 4.50$$

Then optimal assignment is as follows:

$$\begin{aligned} x_i &= 1; & i &= 4,7,12,13,15 \\ x_i &= 0; & i &= 1,2,3,5,6,8,9,10,11,14,16,17,18,19 \\ x_i &= 0; & i &= 19,20,21 \\ x_{22} &= 3.7 & x_{23} &= 4.24 \\ x_{24} &= 4.50 & x_{25} &= 4.50 \end{aligned}$$

This assignment requires all three vehicles to immediately leave their source nodes and proceed to target. Vehicle 1 performed classification, vehicle 2 attack and vehicle 3 verification tasks. WASMs 1 and 3 then proceed to sink i.e., continue to search for other targets.

V. DISCUSSION ON PROBLEM SIZE

Considering the AV optimization problem in section 3 and corresponding illustrative example in section 4, a discussion on problem size is presented here. For n targets, w vehicles and $m = 3$ tasks per target, problem size comprises of $n(n - 1)wm + nwm + 2nw + mn + 2w + 1$ decision variables. Of these $3 + nm + 1$ are continuous timing variables and rest is binary decision variables. The number of constraints also grows rapidly. There are $12(n - 1)nw + 9nw + 2nwm + 2nm + 3w$ constraints, of which $mn + w$ are equality constraints and $7nw + 2w$ inequality non-timing constraints and $12(n - 1)nw + 2nw + 2nwm + mn$ inequality timing constraints. The size of FMILP increases rapidly as problem size increases. Some practical sized problems are amenable to optimal solution with this formulation. For $n = 2, w = 3$, there are 51 binary decision variables, 10 continuous decision variables, 9 linear equality constraints and 174 linear inequality constraints. For $n = 2, w = 4$, there are 68 binary decision variables, 11 continuous decision variables, 10 linear equality constraints and 230 linear inequality constraints. For $n = 2, w = 5$, there are 85 binary decision variables, 12 continuous decision variables, 11 linear equality constraints and 286 linear inequality constraints. For $n = 3, w = 4$, there are 136 binary decision variables, 14 continuous decision variables, 13 linear equality constraints and 485 linear inequality constraints. The growth of constraints and variables is linear in number of vehicles but quadratic in number of targets. For many operational scenarios involving large problem sizes optimal solutions cannot be found within available computation time and the problem becomes intractable in nature leading to NP Complete problem.

VI. CONCLUSION

In this work, FMILP model is presented to find an optimal solution to NP Complete complex multiple task assignment and scheduling problem where tasks are coupled by timing and task precedence constraints. With due consideration to linearity and inherent uncertainty in real life data involved, AV optimization problem is formulated as FMILP by treating the time variable as continuous one. The

mathematical formulation allows staged AV departure times to guarantee that time constraints are satisfied and incorporate varying task completion times into optimization model. Here feasibility is guaranteed provided that number of UAVs employed exceeds number of targets and UAV endurance is sufficiently high. The formulation is flexible enough to allow for various alternative performance functionals. The rigorous formulation allows true optimal solution for challenging assignment and scheduling problem. Experimental results are presented for practical real life problem from Indian Air Force which demonstrates the efficiency of FMILP problem formulation. The solution developed can serve as an effective heuristic for different categories of AV optimization problems.

REFERENCES

- [1] Conway, R. W., Maxwell, W. L., Miller, L. W.: *Theory of Scheduling*. Courier Dover Publications (2003).
- [2] Gendreau, M., Hertz, A., Laporte, G.: A Tabu Search Heuristic for Vehicle Routing Problem. *Management Science*, 40, 1276 – 1289 (1994).
- [3] Glover, F., Laguna, M.: *Tabu Search*. Kluwer Academic Publishers: Boston, MA, USA (1997).
- [4] Hooker, T. N.: Testing Heuristics – We have it all wrong. *Journal of Heuristics*, 1, 33 – 44 (1995).
- [5] O'Rourke, K. P., Bailey, T. G., Hill, R., Carlton, W. B.: Dynamic Routing of Unmanned Ariel Vehicles using reactive Tabu Search. *Military Operations Research Journal*, 6, 33 – 42 (2000).
- [6] Pho, N. V.: Formulation of Membership Function and Determination of Input of Fuzzy Loads in Structural Fuzzy Analyzing Problem. *Vietnam Journal of Mechanics*, 29(2), 117 – 126 (2007).
- [7] Pinedo, M.: *Scheduling*. Prentice Hall: Upper Saddle, NJ, USA (2002).
- [8] Ross, T. J.: *Fuzzy Logic with Engineering Applications*. John Wiley and Sons, Third Edition, New Delhi, India (2004).
- [9] Schumacher, C., Chandler, P., Pachter, M., Pachter, L. S.: Unmanned Air Vehicles Task Assignment with Timing Constraints via Mixed Integer Linear Programming. *Proceedings of Third Unmanned Unlimited Technical Conference*, American Institute of Aeronautics and Astronautics, Reston, VA, USA, AIAA Paper Number 6410 (2004).
- [10] Simon, D.: Sum Normal Optimization of Fuzzy Membership Functions. *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems*, 1(1), 100 – 122 (1993).
- [11] Toth, P., Vigo, D.: *The Vehicle Routing Problem*. SIAM: Philadelphia, PA, USA (2002).
- [12] Yen, J., Langari, R.: *Fuzzy Logic: Intelligence, Control and Information*. Pearson Education, New Delhi, India (2005).

This document was created with Win2PDF available at <http://www.win2pdf.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.
This page will not be added after purchasing Win2PDF.

A Study on Common Finance Engineering in India

Nilanjan Singha Roy

Techno India College of Technology, New Town, Kolkata-700158,India

Sourabh Singha Roy

EIILM, 6 Waterloo street, Kolkata -700069,India

S.Basu chowdhury

Centre for Management Studies, JIS College of Engineering, Kalyani, Nadia-741235, India

Subhamoy Singha Roy[‡]

Department of Physics , JIS College of Engineering, West Bengal University of Technology, Kalyani, Nadia-741235, India

Abstract— The recent circumstances of the Indian capital market reflect that, industry as a whole has been reckoned as a pace setter in the capital market with significant impact in the capital market stability. The capital market has recently recorded a historic height and according to the market analysis's, this is largely the handy work of the CF manufacturing. On the other hand, to what extent this robust activity of the CF industry in the capital market is to the benefit of corporate sector or to the benefit to the small investors is a basic question of examination. In this context we first note that the overall growth of the CF industry in different phases and analyzed the resource mobilization by CF's in India. We also study the dealt with the classification of the MF industry and discuss the private sector CFs of different varieties. It also covers foreign MFs and their collaborations. Finally we have explained the growing impact of mutual funds in Indian capital market and examine some rules and regulations of mutual funds's in India.

in the Indian MF industry. By the turn of the century MF industry comprising government, private and foreign agencies have begun to emerge in the capital market as dominant players. The recent situation of the Indian capital market reflect that, MF industry as a whole has been reckoned as a pace setter in the capital market with significant impact in the capital market stability. The capital market has recently recorded a historic height and according to the market analysts, this is largely the handy work of the MF industry. However, to what extent this robust activity of the MF industry in the capital market is to the benefit of corporate sector or to the benefit to the small investors is a basic question of examination. In this paper we note the overall growth of the common fund industry in different cases are discusses in section A ,B,C,D,E,F,G and section H.

II. THEORETICAL BACKGROUND

A. Over all growth of common fund industry in India

Common Fund (CF) industry in India has actually started its journey from 1964. The growth of mutual fund industry in India took place in several phases. The initial phase was started in 1964 and continued till 1987. The year 1964 is the red lettered year in the history of mutual fund industry in India which makes the establishment of Unit Trust of India (UTI), under the Act of Parliament. UTI was set up by RBI and functioned under its regulatory and administrative control. In the year, 1964, UTI launched its first mutual fund scheme named Unit Scheme- 1964 (US-64) which is the 'Flagship' Scheme of UTI and by the end of March, 1987, it held assets worth of Rs 4,564 crores under its management.

The second phase of mutual fund (Common fund) industry in India is the period from 1987 to 1993. This period is characterized by the entrance of Public sector mutual

[‡]Corresponding author.

E-mail address: ssroy.science@gmail.com

2nd Annual International Conference IEMCON 2012

funds in the market. State Bank of India Mutual Fund fired the first salvo in June, 1987 and afterwards followed by many other nationalized banks like, Canara Bank Mutual Fund, Punjab National Bank Mutual Fund, Indian Bank Mutual Fund, Bank of India Mutual Fund, Bank of Baroda Mutual Fund and financial institutions like LIC of India and GIC of India. By the end of March, 1993 the Indian CF industry held assets worth of Rs.47,000 crores under its management.

The third phase is from 1993 to 2003 which is marked by entry of private sector in the CF industry with a large number of choices of fund families to the investors. In the year 1993, the first common fund regulation was laid down which necessarily required all CFs except UTI, to be registered and governed under SEBI (Mutual Fund) Regulations. In 1996, this regulation was replaced with a more comprehensive and revised SEBI (Mutual Fund) Regulations. This phase witnessed numerous mergers and acquisitions and a large inflow of funds by several foreign CF in the industry. The number of mutual fund houses increased to 33 with total assets of Rs. 1, 21, 805 crores by the end of January, 2003.

The fourth phase started in February, 2003, marked with the bifurcation of UTI into two separate entities; the specified undertaking of the UTI (UTI-I) with assets worth Rs. 29,835 crores under its own management by the end of January 2003, and the UTI Mutual Fund Ltd. (UTI-II), funded by SBI, PNB, BOB and LIC (This bifurcation of has been treated separately in separate chapter) With all these developments and amendments the MF industry in India has entered into its present consolidation and expansion phase. By the end of September, 2004, there were 29 MFs with assets under management of Rs. 1, 53,108 crores under 421 schemes as the following chart indicates.

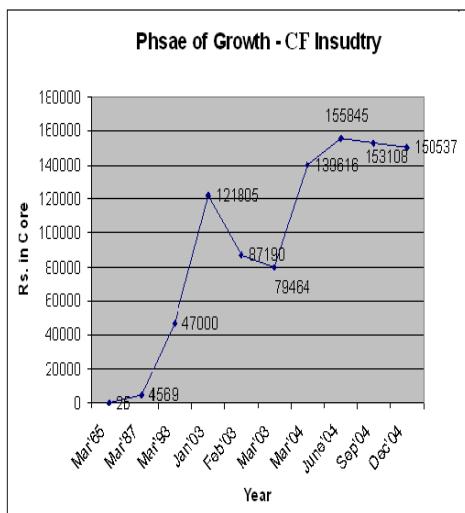


CHART-1

B. Resource mobilization

Recent capital market index indicates that the CF industry has become a magnetic institution for resource mobilization. The Table below indicates that the CF industry since 2000, mobilized huge resources and the private sector CFs has out paced the public sector CFs.

TABL— 1
Resource Mobilisation by Common Funds (Rs. Core)

| Year | Public Sector Mutual Funds | | | | Private Sector Common Funds | Grand Total (S+G) |
|-----------|----------------------------|----------------|---------------------|----------|-----------------------------|-------------------|
| | Bank Sponsored | F.I. Sponsored | Unit Trust of India | Total | | |
| | (2) | (3) | (4) | (3+4) | (5) | (7) |
| 1995-96 | 113(4) | 235(3) | -6314(1) | -5966(8) | 133(11) | -5833(19) |
| 1996-97 | 6(3) | 137(2) | -3043(1)@ | -2900(6) | 864(17) | -2034(23) |
| 1997-98 | 237(2) | 203(3) | 2875(1) | -3315(6) | 749(15) | 4064(21) |
| 1998-99 | -88(2) | 547(3) | 170(1) | 629(6) | 2067(16) | 2695(22) |
| 1999-2000 | 336(6) | 295(3) | 4548(1) | 5179(10) | 16937(27) | 22117(37) |
| 2000-01 | 248(6) | 1273(3) | 322(1) | 1843(10) | 9292(27) | 11135 |
| 2001-02 | 863(6) | 407(3) | -7284(1) | 6014(10) | 16134(27) | 10120(37) |
| 2002-03P | 1033(4) | 862(3) | -9734(1) | -7539(8) | 12122(26) | 4532(34) |
| 2003-04P | 2635(4) | - | * | 4812(7) | 42873(24) | 47684(31) |

Phase

- Data pertain to period February 1, 2003 to March 2004, being first year of operation after the bifurcation of erstwhile UTI into UTI Mutual Funds and specialized undertaking of Unit Trust of India.

Notes:

1. For UTI, the figures are net sales (with premium) under all domestic schemes and for other mutual funds. Figures represent net sales under all ongoing schemes.
2. Date exclude amount mobilized by off-shore funds and through roll over schemes.
3. Data within parentheses relate to the number of mutual funds which mobilized resource during the year.

Source: UTI and representative Common Funds.

Some notable feature of growth and factors contributing that growth may be mentioned:

One important aspect is that there is a lot of difference in MF industry between pre-reform and post-reform phases. As we have seen the early phase had a slow growth because the numbers of players were limited and the resources mobilised by them were also insignificant.

Secondly, the entry of private mutual funds meant that not only the market share but also the mind set-up of the investors have been captured by them. These new players went on a high scale marketing to create awareness among the investors about the advantages of investing in CFs. They

2nd Annual International Conference IEMCON 2012

came out with innovative tailor made products to suit the varied needs of the investors. In post1993, the market saw profuse innovative products catering to the different needs of the investors, such as, tax savings schemes, balanced schemes, sector specific funds, age specific schemes, gender specific schemes etc. They started offering different investment horizons to suit investors' short term and long term plans and financial needs. The tax concessions of various schemes have largely influenced the urban investors and contributed to the growth of this industry.

Thirdly, the entry of new foreign investors, particularly after 2000, as we have seen on the table has brought about a phenomenal change in volume and diversity of the CF industry activity.

Fourthly, one important recent change is the trends towards consolidations by mergers and acquisitions of the CF players. This is particularly a trend set after the emergence of the foreign CFs.

Fifthly, this trend has set a new tone of competition to the CF industry. This is the outcome of liberalization of the capital market. As the trend shows, the public sector CFs are generally losing grounds to the private sector CFs. This trend indicates that the CF industry in India is rapidly becoming a notable part of global capital market.

The growth of CF industry has an important impact on gross domestic savings including corporate, government, and household savings. As the following table indicates, Gross Domestic Savings (GDS) as a percentage of GDP in 1995 was 24.8; of this total savings, bank deposits captured 38.2 percent and AUM of the MFs share 7.4 percent. In March, 2001, share of bank deposits rose to 45.7 percent and AUM by CFs was 4.3 percent of the total GDS which was 23.4 percent of the GDP in the same year. The important point to note in this connection is that domestic savings slightly decreased in March, 2003, (24.2 percent of GDP). Volume of GDS increased and total bank deposits also increased significantly. By March, 2003, the relative position of bank deposits and CFs as competitive pool of GDS indicates that strong pull of bank deposits sharing 54.9 percent of GDS. That the CFs have secured a notable position of attraction of Household Savings is clear from this, but the pull of banks is far stronger than that.

C. Classification of mutual fund industry

The dimension of the CF industry as it stands currently can be easily accessed from the accompanying chart. The Bank sponsored Mutual Funds are four; viz. Bank of Baroda Asset Management Co. Ltd., Can bank Investment Management Services Ltd., State Bank of India Funds Management Ltd., UTI Asset Management Company Pvt. Ltd. The combined assets under management of these four companies are Rs 26079 crores or 16.73 per cent of the total asset under management by CF industry.

The Financial Institution sponsored MFs are three; viz. GIC Asset Management Co. Ltd., IL& FS Asset

Management Co. Ltd. and Jeevan Bima Sahayog Asset Management Co. Ltd. The total assets under management of these companies are paltry amount of Rs 7179 crores or 4.60 percent of the asset under management by CF industry.

TALBT— 2

| Year | Gross Domestic Savings | Bank Deposit (in cr.) | | | | Common Funds | | |
|----------|------------------------------|-----------------------|---------|--------|---------|--------------|--------|--------|
| | | Term | Savings | Demand | Total | No | Scheme | AUM |
| Mar-95 | 251463 | 218632 | 91324 | 76903 | 386859 | 22 | 167 | 74513 |
| % of GDP | 24.8 | 21.6 | 9.0 | 7.6 | 38.2 | | | 7.4 |
| Mar-98 | 352178 | 356008 | 139964 | 102513 | 598485 | 38 | 259 | 97228 |
| % of GDP | 23.1 | 23.4 | 9.2 | 6.7 | 39.3 | | | 6.4 |
| Mar-2000 | 466640 | 494078 | 191900 | 127366 | 813344 | 38 | 330 | 107946 |
| % of GDP | 24.1 | 25.5 | 9.9 | 6.6 | 42.0 | | | 5.5 |
| Mar-01 | 491761 | 597084 | 222982 | 142552 | 962618 | 39 | 394 | 90587 |
| % of GDP | 23.4 | 28.4 | 10.6 | 6.8 | 45.7 | | | 4.3 |
| Mar-02 | 535185 | 797403 | 255598 | 152929 | 1205930 | 38 | 417 | 100594 |
| % of GDP | 23.5 | 34.9 | 11.2 | 6.7 | 52.8 | | | 4.4 |
| Mar-03 | 597697 | 888987 | 302303 | 164590 | 1355880 | 38 | 406 | 109299 |
| % of GDP | 24.2 | 36.0 | 12.2 | 6.7 | 54.9 | | | 4.4 |

These two groups of companies together have control over control over around 21 percent of the total asset under management by CF industry by June, 30, 2004. An inter group management of these companies indicate that Bank sponsored FI sponsored companies have only one- fifth of the market share of the CF industry. It is significant to note that the UTI AMC which was one time worth of the CF industry has now a market share of Rs 18875 crores or 12.11 per cent of the total asset under management of CF industry and 72.37 per cent of bank sponsored Asset Management Company. We must remember that UTI AMC is virtually is a bank sponsored AMC as it has been promoted by a consortium of three banks as mentioned above.

Another group of CF industry is the private sector companies, separately discussed below. This group together control over 78 percent of the overall market share of the MF industry. We discussed these segments of the CF industry. The UTI, being the pioneer entity in the market has been treated in two chapters separately. The bank sponsored MFs are treated in another chapter. The LIC sponsored AMC has also been treated separately. These two groups are our major focus of our study. However, as shown above the private sector CF industry has virtually outstripped all proportions in its different ramification. This has been treated in next section.

2nd Annual International Conference IEMCON 2012

D. Private sector common fund

As mentioned above the emergence of the private sector CF has completely changed the character of the MF industry in our country. The private sector CF are characterized in three categories.

- a. Indian private sector CF
- b. Joint ventures— predominantly Indian
- c. Joint venture— predominantly Foreign.

As it stands at present, the whole lot of private sector MFs comprises 22 out of 29 CFs in India. The list given below present the current position of the Asset Under Management of these companies and their relative position with other Bank and FIs sponsored mutual funds.

As stated above, the entire CF industry veering around the private sector unit groups under different descriptions. We also seen that in the entire industry the predominantly foreign but joint ventures around 39 percent compared to around 21 percent share of bank FI sponsored together. We thus find that in the market status liberalization has strongly molded the CF industry. The industry initially sponsored by the government has now been handed over to private sector and that too with predominant with foreign companies' control. This has a serious implication for resource mobilization and investment pattern of the Indian capital market. This is further starkly appears when we look to the leadership status of the company.

Leadership Status

In the status analysis of the different CF companies it appears that two leading predominantly foreign CFs (viz. Franklin Templeton Asset Management (India) Pvt. Ltd. and Prudential ICICI Asset Management Co. Ltd.) cover

around 21 percent of the total asset managed by the entire CF industry and 56 percent of the group share. Side by side, the leading players of India dominated private CF, viz. HDFC Asset Management Co. Ltd has itself hold the control over 10 percent of the total asset under MF industry in India. A further compression of the groups show that the MF industry in India virtually controlled by 5 leading players with total market share of 50 percent. Within this group only Reliance has 7 percent share but others have shares 10 percent or more.

The game in the CF market therefore, is confined to very limited number of players. Their collusion or conflict is very likely to generate unforeseen turmoil in the capital market. As the Indian capital market under liberalization has become subjected to the forces with both domestic and foreign origin leading CF companies will be bent on taking the advantages of the market forces and hence to attain maximum gain under conditions of imperfect market information.

TABLE – 3
ASSETS UNDER MANAGEMENT AS MEMBER
WISE
(Rs. In Crores)

| Sr. No. | Name of the Asset Management Company | Assets under Mnagement |
|---------|---|------------------------|
| A | Bank Sponsored | |
| 1 | B OF Asset Management Co Ltd. | 419 |
| 2 | Canbank Investment Management Services Ltd. | 1489 |
| 3 | SB IFunds Management Company Pvt. Ltd. | 5296 |
| 4 | UTI Asset Management Company Pvt. Ltd. | 18875 |
| | Total A | 26079 |
| B | Institutions | |
| 1 | GIC Assets Management Co. Ltd. | 204 |
| 2 | IL & FS Asset Management Co. Ltd. | 2015 |
| 3 | Jeevan Bima Sahayog Asset Management | 4960 |
| | Total – B | 7179 |
| C | Private Sector | |
| (i) | Indian | |
| 1 | Benchmark Asset Management Co. Pvt. Ltd. | 78 |
| 2 | Cholamandalam Asset Management Co. Ltd. | 1243 |
| 3 | Escorts Asset Management Ltd. | 131 |
| 4 | JM Capital Management Pvt. Ltd. | 4111 |
| 5 | Kotak Mahindra Asset Management Co. Ltd. | 5651 |
| 6 | Reliance Capital Asset Management Ltd. | 11204 |
| 7 | Sahara Asset Management Co. Pvt. Ltd. | 345 |
| 8 | Sundaram Asset Management Co. Ltd. | 1841 |
| 9 | Tata Asset Management Pvt.Ltd. | 5537 |
| | Total C(i) | 30141 |
| (ii) | Joint Ventures – Predominantly Indian | |
| 1 | Birla Sun Life Asset Management Co. Ltd. | 9397 |
| 2 | Credit Capital Asset Management Co. Ltd. | 131 |
| 3 | DSP Merrill Lynch Fund Managers Ltd. | 6389 |
| 4 | HDFC Asset Management Co. Ltd. | 16105 |
| | Total C (ii) | 32022 |
| (iii) | Joint Ventures – Predominantly Foreign | |
| 1 | Alliance Capital Asset Management (India) Pvt. Ltd. | 2446 |
| 2 | Deutsche Asset Management (India) Pvt. Ltd. | 2266 |
| 3 | Franklin Templeton Asset Management (India) Pvt. Ltd. | 17342 |
| 4 | HSBC Asset Management (India) Private Ltd. | 5463 |
| 5 | ING Investment Management (India) Pvt. Ltd. | 1472 |
| 6 | Morgan Stanley Investment Management Pvt. Ltd. | 1095 |
| 7 | Prudential ICICI Asset Management Co. Ltd. | 16071 |
| 8 | Principal Asset Management Co. Pvt. Ltd. | 4825 |
| 9 | Standard Chartered Asset Management Co. Pvt. Ltd. | 9444 |
| | Total C(iii) | 60424 |
| 22 | Total (i+ii+iii) | 122587 |
| 29 | Total (A+B+C) | 155845 |

2nd Annual International Conference IEMCON 2012

TABLE -4

Leadership Status:

| Name of the Sponsor | Name of the AMC | In the group(%) | Overall (%) |
|-------------------------------------|--|-----------------|-------------|
| Bank sponsored | i) UTI Assets Management Company ii) SBI Funds Management Ltd. | 72% | 12% |
| | | 20% | 3% |
| FI Sponsored | i) Jeevan Bima Shayog Assets Management Company Ltd. ii) IL & FS Assets Management Co. Ltd. | 69% | 3% |
| | | 28% | 1% |
| Private sector- Indian | i) Reliance Capital Asset Management Ltd. ii) Kotak Mahindra Asset Management Co. Ltd. | 37% | 7% |
| | | 19% | 4% |
| Joint venture- Predominantly Indian | i) HDFC Asset Management Co. Ltd. ii) Birla Sun Life Asset Management Co. Ltd. | 50% | 10% |
| Joint Ventures – Predominantly | i) Franklin Templeton Asset Management (INDIA Pvt. Ltd.) ii) Prudential ICICI Asset Management Co. Ltd. | 29% 29% | 6% 11% |
| Total - | | 27% | 10% |
| | | | 67% |

E. Recent trend of mutual fund industry

The common fund industry in India has undergone a most successful phase in the recent times; the industry has never been better. Today common funds have entered all the segments of capital markets be it equity, debt, money market, gilt or hybrid. Retail investors have put billions of money into funds and have reaped handsome rewards in the form of double-digit or even triple-digit dividends. With the record breaking performance of the Mutual Fund industry, the economists expected GDP rising to 8% by the end March, 2007, the confidence of investors in the market has increased and a large number of Foreign Institutional Investors (FIIs) including mutual, hedge and pension funds are diverting their money to the Indian capital market. The overseas investment in the country has accumulated to \$22bn since 2003, leading to more than double the increase in sensex over the past two and half years. US-based mutual funds targeting emerging markets like India have accumulated over \$5bn from investors till June, 2005, accounting to nearby ten times the raised in 2002. the total number of FIIs registered with SEBI has increased to 685 out of which almost 150 new FIIs have registered in 2005. IN addition to American and British FIIs, the FIIs from Canada, Denmark, Sweden, Belgium Ireland, Italy and Japan have made their presence felt and are playing a major role in the present up-surging market. Several leading international pension funds like Cal PERS (California Public Employees' Retirement System),the UN Pension Funds, the US State Government Pension Funds, the General Motors Employees' Pension Fund etc. have also registered as FIIs with SEBI. All this growth in the mutual fund industry is the outcome of India's robust capital market which has taken this sector to the peak of its performance. According to the estimates of the Association of Mutual Funds in India (AMFI), with the spurt in stock

market the Mutual Fund Industry witnessed a growth of 7% in July 2005. Many funds have been coming up with New Fund Offerings (NFOs). Providing the investors with more lucrative returns the equity mutual funds have topped this list of funds in terms of returns offering, i.e.; between 8 and 10% a month.

After achieving such great success in the retail sector, the Mutual Funds are now looking forward to the overseas market, aiming at the affluent Indian Community, estimated to be around 25 million, living across the world, remitting around \$20bn to India annually. In addition to this the Mutual Funds in India are also planning to increase their exposure in the infrastructure sector. There are plans to invest around \$10bn for the construction of new roads, highways, seaboards, airports etc, over the next three years in the country. Funds such as Tata Mutual Fund, DSP Merrill Lynch and Prudential ICICI have already come up with their infrastructure funds and many more are expected to come. The performance of these infrastructural funds has been reasonably good with 'The Infrastructure Growth and Economic Reforms' (or TIGER) fund by DSP Merrill Lynch giving a return of 15.86% in the first of 2005 and of 52.64% for the 12 month period; Tata Infrastructure Fund gave a return of 11.75% in the same quarter.

F. Money market mutual fund

MMMFs operation is a unique addition to institution ventures introduced by the RBI in 1991 of the recommendation of the Tusk Force set up for that purpose.

The RBI has appeared on September 3, 1991 a Tusk force on MMMMF under the chairmanship of D.Basu, deputy Managing director of SBI. The terms of reference of the T.F were

- A. To set up detail guidelines from the view point of implementation of the schemes of MMMFs;
- B. To devise the precise forms of the negotiable instruments;
- C. To formulate the instrument guidelines of MMMF in with the broad outline in the credit policy circular.

The MMMF are particularly designed to enable the individuals to park their funds temporarily in Money Market instruments until some long term investment avenues open up. They will bring within the reach of some individual investors the MM Instruments which hitherto they could not invest in because of the large minimum amount required for that. MMMF assured the individuals high liquidity and relatively better yields coupled with safety of investment. It is expected that MMMFs will stimulate the financial savings and enhance the share of such savings of the household sector. MMMFs will therefore, improve the cash management of the individuals who invest in them.

New money market instruments include Commercial Papers (CPs), Commercial Bills (CBs), Treasury Bills (T-Bills), Government Securities (G-Sec),

2nd Annual International Conference IEMCON 2012

Call and Notice money, Certificate of Deposits (CDs), Usance Bills and any other similar instruments as specified by the RBI from time to time.

The minimum investment in CDs and CPs is Rs 25 lacs each. Similarly, for participation in the auction of 182-Days T- Bills minimum investment amount is prescribed as Rs 1 lac. For the participation of an entity in Bills Rediscounting Schemes and call money market, an investment of Rs 20 crores per transaction for any lender is required. The MMMF instruments like CDs and CPs etc. are thus meant for raising bulk funds as such, the access to these institutions has been confined to banks, FIs and some large PSUs and business houses, as such the individuals and small bodies could not mobilize enough resources for investment in money market instruments. Besides, they are not able to obtain optimum yields on their investment due to lack of required expertise and diversified portfolio of investment. The transaction costs for such operations are also very high.

Therefore, in order to provide access to various money market instruments for small investments, particularly, the individuals the MMMF schemes have been formulated. It is designed to collect small lots from individuals and others for investing in various specified short term money market instruments of high quality. MMMFs, therefore, offer an economic advantage of block purchase of money market instruments and diversification of investment which would ensure optimum yields and instruments.

Net income earned by the MMMFs on their investment portfolios would be distributed among the fund participants. The return on investment in MMMFs would depend upon the prevailing interest rate on various money market instruments. The primary objective of the MMMFs is to provide high rate of return on short term investment consistent with liquidity and preservation of capital through maintenance of high quality investment portfolio.

An important feature the MMMFs is the association of MMMFs with commercial banks. The RBI considered that it would be desirable to allow commercial banks and their subsidiaries to set up MMMFs. The Task Force was of the view that MMMFs could be set up as a part of a bank. (They are considered as in-house MMMFs.) In the USA many MMMFs have been set up as a part of bank because banks have suitable infrastructure for MF operation at a reduced cost of administration. The Task Force was of the opinion that if the MMMFs are set up as in-house body they would be subject to Cash Reserve Ratio (CRR) and Statutory Liquidity Ratio (SLR).

MMMFs scheme is usually in two forms; such as, Money Market Deposit Account (MMDA) and Money Market Mutual Funds(MMMFs). Where in the case of MMDAs the investors are issued deposit receipts, in the case of MMMFs they are issued instruments in the forms of

share units or Certificate of Deposits (CDs). MMDAs would not be eligible for tax exemption but would be under insurance cover. The Task Force recommended that the units of shares of MMMFs be exempted from stamp duty both at the time of issue of instruments and at the time of issue transfer instruments at the secondary market.

The schemes launched by MMMFs would be either as 'close'-ended or 'open'-ended funds. The close ended are set up for a fixed period. The open ended funds have no ceiling on the size of the fund and the investors have facility to invest or disinvest their investment whenever they like. Thus the open ended funds are always available for investment of short term funds and also have better liquidity. The regular flow fund would enable to fund managers to diversify investment which may ensure better return.

MMMFs are very popular in the US and UK as a vehicle for retail investors to arbitrage between the low bank deposit rate and higher money market interest rate, price decline in the equity markets and rising interest rate made the MMMF a popular investment avenue. By 1980, MMMFs were managed about 57% of the total assets of the MF industry in the USA. In India the RBI has set up MMMFs in April, 1992, as stated above for there were no regulatory restrictions on the minimum size or tenure of investment in MMMFs. Here is a conflict of interest between the banks and Money Markets (MMs). To avoid the competition the RBI has aligned the lock-in period for investment in MMMFs with the minimum maturity period of bank fixed deposits. The minimum tenure and lock-in period for MM instruments like CPs and CDS were also aligned with the minimum maturity period of bank deposits. MFs set up a bank for FIs have access to call market. MF set up a private sector was allowed to lend call money market and bill rediscounting by the RBI in june, 1995. To develop MM the RBI allowed non- bank participants to lend the call money markets, through primary dealers and DFHI. The minimum lending size for call money market reduced gradually. The restricted access for non-bank entities to lend in the call notice money markets is being phased out by RBI. CFs have been allowed to participate in Repo market from March, 2003, to provide an alternative to lending in the call market. These measures taken by RBI have resulted in a vibrant money market. As it appears from the average amount outstanding of selected instruments given in the table below:-

TABLE — 5
Average Amount Outstanding of Few Market Instruments
(Rs. Core)

| Year | CP | 14 T Bill | 91 T Bill | 182 T Bill | 364 T Bill | Call Money | CD |
|-----------|-------|-----------|-----------|------------|------------|------------|-------|
| 1997-98 | 2,793 | 2,425 | 3,588 | - | 15,567 | 25,043 | 9,405 |
| 1998-99 | 4,585 | 625 | 4,132 | - | 8,242 | 26,228 | 6,506 |
| 1999-2000 | 7,014 | 631 | 2,025 | 1,167 | 13,483 | 23,205 | 1,861 |
| 2000-01 | 6,751 | 396 | 1,800 | 1,300 | 13,417 | 30,320 | 155 |
| 2001-02 | 7,927 | 50 | 4,842 | 433 | 18,244 | 35,166 | 965 |
| 2002-03 | 8,268 | - | 6,617 | - | 23,282 | 29,421 | 1,224 |

2nd Annual International Conference IEMCON 2012

Impact of MMMFs in India

The impact of MMMF is very significant in the financial sector in India. There is a phenomenal rise in Asset Under Management (AUM) of MFs. As on 31 March, 2004, the AUM of MFs was Rs 41,704 crores. The reason for this growth of the MF industry is rapid enhancement of MMMFs and at the same time continues fall of interest rate in India, and a consequent result in capital appreciation in debt instrument. Buoyancy in the equity market is also an important factor for more inflow of money in the MF industry.

However, the real test of the MF industry performance would come when the interest rate rises. We are keenly watching to see if 2003-04 was the year of turned around for MMMFs or if it was just a sporadic event.

Impact on household savings:

It was envisaged that MMMFs would stimulate gross Household Financial savings (HFs) which is around 45% of Gross Domestic savings (GDs) in India. The share of HFs in MFs has been declining in recent years. HFs in UTI was 4.3% in 1994 and 1.2% in rest of the MFs. It moved to -0.6% and 1.7% respectively in 2002. It is clear that other MFs failed to contribute significantly at least at the macro level. Secondly, with the decline of HFs in UTI, the share of MF industry in HFs has shown a massive decline in the said period. This is caused mainly due to the debacle of UTI when investors were frightened about to invest in CF industry. However, the picture is changing, the share of HFs in other MFs has been increasing in recent period. The impact of MMMFs on GDs has been significant. Since the MMMFs commenced operation from 1996-1997 GDs has been consistently rising. But compared to the growth of bank deposit the contribution of MMMFs is low.

In fact the size of the Indian CF industry is small when compared to its global peers. However, the relative performance of MMMFs within the CF industry is better and constitutes around 25% of the size of the MF industry as in September, 2003.

G. Growing impact of common funds in indian capital market

Common Funds are specialized institutions that collectively manage the funds obtained from investors through portfolio of securities and capital gain. These specialized institutions are managed by professionals and have safety, liquidity and returns to the investors.

Common Funds play an important role in developing the capital markets by providing a lot of capital information. There is also need of large number of domestic mutual funds. The domestic mutual funds industry can enrich the liquidity and stability of India's stock and bond markets. CFs in general should further broaden the appeal and participation

of all Indian investors in their own securities market and ultimately the personal stake of all investors in the growth of the Indian economy. In this regard Indian regulatory authorities also have a crucial role to play.

For any growing economy, especially a rapidly emerging one with more demands for infrastructure and technology will require enormous sums of capital. The days of World Bank loans to build showcase factories in remote villages is not the paradigm for the coming generations of economic expansions. Therefore, the emerging economics of the world, such as India, will increasingly rely on private capital resources, either through direct investment capital project finance or domestic capital markets. All these three sources of capital are essential to the long-term success of providing funds for capital market activity.

The ideal role of private sector capital market is three fold. First, an emerging capital market requires sustainable domestic demand for its expanding supply of debt and equity issues. Secondly, the best source of domestic demand in India is the role of wealth held by individual savers who have begun to share in India's growing economy. Lastly, the best practical way to direct these into the capital market is through MFs, which can serve as a conduit for dis-intermediating bank deposits into stocks and bonds. The process of institutional dis-intermediation of time deposits into CFs benefits the capital market because it provides a sustainable domestic source of demand. There are several latent benefits for its markets from this process. An indigenous CF industry provides valuable liquidity latent benefits for the markets from this process. An indigenous CF industry provides valuable liquidity to domestic capital market when foreign investors and individual speculators exit and re-enter the marketing. Further, a domestic institutional class of investors will in all investor discipline requiring increased disclosure of Corporate Financial information flow and increased investor reliance on economic fundamentals. An institutional investor class within India supplied with capital from individual savers purchasing CFs will provide for a more efficient and liquid capital market overall. Further, the process of investing through CFs will benefit the individual savers in India. It is because of the individual saver who purchase shares in CFs acquire access to capital market investments. Therefore they enjoy a better opportunity to match their long-term liabilities as retirement and education with long-term investment assets such as stocks and bonds. Thus India will require its nascent CF industry to mature and develop to provide the necessary link between the private savings and its capital markets. This demand for India's stock and bonds as well as encouraging long-term savings presently held in short-ten-n deposits.

H. Regulations of the common funds

The broad principals of the regulation CFs relates to the money market regulations of RBI. Money market are

2nd Annual International Conference IEMCON 2012

expected to perform the following broad functions.

- a) To serve as an equilibrating mechanism between the supply of and demand for short term funds.
- b) To serve as a focal point the central bank intervention for influencing the liquidity and general level of interest rate and
- c) Also to provide access to the fund users for borrowing and investing at an efficient market clearing price.

Private sector CFs and CFs set up by public sector banks and FIs come under the regulatory aspect of SEBI. The guidelines designated by RBI as regulator of MMMFs set up by banks. But SEBI is the regulator of MMMFs schemes of other non- bank CFs. It is specified that SEBI regulations pertaining to MMMFs shall be in conformity with the guidelines for MMMFs to be announced by RBI.

We have seen above the broad characteristics of Indian CF industry which consists of both public and private sector unites. Also we have seen the changing characteristics particularly the relative strength and weakness of the different groups of CF unites. This has important implication for money market operations in particular, and the financial activities in general which calls for principal regulation by a regulatory authority.

III. RESULT & DISCUSSIONS

Although relatively recent, CF industry becomes important players in Indian capital market after liberalization. Starting with the limping growth with UTI (1964) the industry has a registered phenomenal progress by the end of 90's, when UTI has itself broken asunder and was bifurcated to two companies. The recent growth of the industry is largely attributed to emergence of private CF players of which foreign dominated joint venture companies begun to steal the march. The rapid compression of the industry has virtually led it to the monopoly control of 4 to 5 players of which the first player is UTI is thoroughly witted out. This monopoly trend of the MF industry signals the source of the market sensitivity of the capital market presently, appearing in the stock market.

ACKNOWLEDGEMENT

The authors thank to the Associate Editor and a referee for the insightful comments that helped improvement of the paper.

REFERENCES

- [1] Chander, Ramesh (2002), Performance Appraisal of Mutual Funds in India, Excel Books, New Delhi.
- [2] Gupta, Amitabh (2002) Performance Evaluation of Mutual Funds, Anmol Publications, New Delhi.
- [3] Reddy, Venugopal. Y (2000), Monetary & Financial Sector Reforms In India, UBS Publishers Distributors Ltd., New Delhi.
- [4] Reserve Bank of India (1992), A Report of the Task Force of the MMMFs, Mumbai.
- [5] Reserve Bank of India (2001-02), Report on Currency & Finance, Mumbai.
- [6] UTI Mutual Fund (2005) "Abridged Annual Report (2004.05), Mumbai,
- [7] Government of India (2004), "The Unit Trust of India (Transfer of Undertaking and Repeal) Act, 2002," New Delhi.
- [8] UTI Mutual Fund (2004), "Abridged Annual Report (2003-04)". Mumbai.
- [9] RBI (2002), "Report on Currency & Finance (2001-2002)".
- [10] Reddy, Y Veringopal (2000) "Monetary and Financial Sector Reforms in India." UBS Publishers' Distributors Ltd., Delhi.
- [11] Tripathy Nalini Prava (2007): Financial Investments and Services p:174-184
- [12] S.S.Roy, P.hD thesis " Changing role of mutual funds in the context of financial perform with particular reference to the public sector mutual funds (UTI,LICI &BANKS sponsored mutual funds), Rabindra Bharati University, 2008.India.
- [13] Datt Ruddar and Sundaram K.P.M (2009): 59th revised edition, Indian Economy: p: 344-347, 758-762.
- [14] Todaro Michael P (1989): 3rd edition, Economic Development in the third World, p: 436-438
- [15] Viz Madhu (2006): 2nd edition, International Financial Management, p: 339-350
- [16] Niti Bhasin (2008):1st edition, Foreign Investment in India, p: 190-196.
- [17] Kamlesh Gakhar (2006): 1st edition, Foreign Direct Investment in India, Policies, Trend and Outlook, p: 185-187.
- [18] Paul Justin(2007): 3rd edition, International Business, p: 203-218
- [19] Sodersten Bo (1989): 2nd edition, International Economics, p: 286-310
- [20] Ghosh Alok (1992): Industrial Policy- A Critical Appraisal, The Economic Studies, vol-31, No-1, p: 81-83
- [21] Economic survey (2006-07), p:153
- [22] Economic survey (2007-08), Report of Central Statistical Organization.
- [23] Datt Ruddar and Sundaram K.P.M (2006): 52nd revised edition, Indian Economy: p: 268-270, 338-346, 789-791.

Analysing Existing Online Business Models and Suggesting an Ideally Secured and Integrated Website Layout

Reetwika Banerjee

*Information Security Manager
Tata Consultancy Services Ltd
Mob – +918697577452
reetwika.banerjee@tcsin.com*

Tamal Kanti Mukherjee

*Senior Consultant
Capgemini India Pvt Ltd
Mob – +919051654459
tamal.mukherjee@capgemini.com*

Abstract

In the modern world everything is getting changed. The world of business is also no exception. Previously, business meant only the ‘Brick and Mortar’ model where the owners of business used to have only physical presence through their stores. Slowly, time changed and with the advent of technology, ‘Brick and Click’ model of business became popular with both offline and online presences. Later, once the concept of E-commerce gained universal recognition, several online business models came into picture. Keeping the current trends in mind, all of them face some issues or the other. The aim of our research paper is to come up with an ideally secured and integrated website layout, suitable to any online business model.

1. Introduction

1.1. Objectives

- i. To study all the existing online business models
- ii. To study in detail some practical examples of each of the models
- iii. To analyse them based on the 7Cs of Internet Marketing
- iv. To do SWOT Analysis for each of the 7C scrutiny
- v. To list down the issues in the form of a common SWOT matrix
- vi. To suggest an ideal 7C model
- vii. To design an ideally secured and integrated website layout suitable to any online business model

1.2. Methodology

The methodology we followed in this particular work started by picking up real life examples of websites which follow these online business models. We selected one website each for the different online business models. After the selection, we analysed each of the websites on the

basis of seven ‘C’s of internet marketing. The step-by-step procedure followed is explained as below:-

- i. Doing a detailed literature review from various online and offline sources on the existing online business models.
- ii. Carrying out an even detailed literature review to come out with some practical examples of each of the models.
- iii. Analysing all these existing models on a common platform using the 7Cs of internet marketing.
- iv. Making use of the marketing tool SWOT Analysis for each of the models.
- v. Listing down all the findings in a consolidated SWOT matrix.
- vi. Suggesting an ideal 7C model, removing the weaknesses and threats from the holistic SWOT matrix.

2. Literature Review

2.1. Definition of Online Business Model

Dave Chaffey (2006) defines an online business model as “*the summary of how a web-company will generate a profit identifying its core product or service offering, target customers in different web markets, position in the competitive online marketplace or value chain and its projections for revenue and costs*”. Timmers (1999) has a fuller definition of online business model as “*the architecture for product, service and information flows, including a description of the various online business actors and their roles, and a description of the potential benefits for the various online business actors and a description of the sources of revenue*”.

2.2. E-Commerce Considerations

- **Picking a business product** ^[1]: Among the most common are selling physical, shippable products; professional services; business or consumer information; and/or attracting

- advertising revenues by delivering free information.
- **Choosing a domain name:** Keeping it simple and easy to remember and register it with a reputable domain registrar.
- **Setting up a merchant account:** Choosing a credit card clearing service so that it can handle online transactions.
- **Take care in designing the web pages:** Templates that include these processes have been developed by many software companies and services providers and can be outsourced saving the time and effort required for designing these features.

2.3. Existing Online Business Models

Every online business is based on a specific business model – be it an auction website, affiliate business or a social network. The models are implemented in a variety of ways, as described below with examples. Moreover, a firm may combine several different models as part of its overall internet business strategy.

Brokerage Model: Brokers are market-makers. They bring buyers and sellers together and facilitate transactions. Usually a broker charges a commission for each transaction it enables, which may vary from business to business. It is categorized into three different types – Business to Business (B2B), Business to Customers (B2C) and Customers to Customers (C2C). Example: Orbitz

Advertising Model: The web advertising model is an extension of the traditional media broadcast model. The broadcaster, in this case, a web site, provides content and/or services blend with e-advertisements. These e-banner ads are the sources of their revenues. It works best when the traffic volume is large and highly specialized. The different Advertising Model categories include – Portal, Classifieds, User Registration and Ultramercials. Example: Monster

Merchant Model: This model consists of wholesalers and retailers of goods and services. Sales may be made based on list prices or through auction. The different Merchant Model categories include - Virtual Merchant, Catalogue Merchant, Click and Mortar and Bit Vendor. Example: Barnes & Noble

Manufacturer (Direct) Model: The manufacturer reaches the buyers directly and thereby compresses the distribution channel. The manufacturer model can be based on efficiency, improved customer service, and a better understanding of customer preferences. Example: Dell

Affiliate Model: It provides purchase opportunities wherever people may be surfing. It does this by offering financial incentives (in the form of a percentage of revenue) to affiliated partner sites. Example: Amazon

Community Model: It is based on user loyalty. The user invests both time and emotion and the revenue can be earned accordingly. The different Community Model categories include - Open Source, Open Content and Social Networking. Example: Flickr

Subscription Model: Users are charged periodic – daily, monthly or annual – fees to subscribe to a service. It can be a combination of free and premium contents. Example: Netflix

2.4. Internet Marketing Security Concerns

Information security concerns are of great importance and the online businesses have been working hard to develop solutions.^[1] Encryption, Multifactor Authentication, Firewalls, Secured payment Gateways etc are some of the primal security controls implemented for dealing with privacy and security concerns on the internet.

2.5. Internet Marketing Analysis Tools

In this paper, we have made use of two internet marketing analysis tools – 7Cs of internet marketing and SWOT Analysis.

2.5.1. 7Cs of Internet Marketing

Using this tool, the websites of each of seven online business models are analysed from both the customer's and provider's perspectives. The 7Cs of Internet Marketing based on which the analysis of each of the above business models are done include^[2]:

C#1 – Context: A website's layout and overall visual design needs to be uncluttered, easy to read and navigate, the colour scheme needs to be appropriate for the marketing design. Having some white space will also aid in the overall design and readability.

C#2 – Communicate: It governs how the company talks to its customers. It can be done through signing up for special offers, email newsletters, contests, surveys, live chat with company representatives, and company contact information.

C#3 – Commerce: If the website is intended for commercial transactions, then it has to be 3D secured. And it must be communicated to the customers. Most websites use a 'lock' symbol at

the corner to indicate that it has been encrypted. SSL standard is recently being adopted by Google.

C#4 – Connection: Any links that lead the customer away from the website falls under the connection category.

C#5 – Content: The text, graphics, sound, music, and/or videos that are presented are classified under the content header.

C#6 – Community: The website may allow interaction between customers through message boards and live chat rooms. This helps in providing a round the clock help desk support to the customers.

C#7 – Customization: Companies can allow customers to personalize aspects of the website or it may tailor itself to different users.

2.5.2. SWOT Analysis

A SWOT analysis is a tool that helps to evaluate the Strengths, Weaknesses, Opportunities, and Threats (SWOT) involved in any business enterprise. The details are described below:-

Strengths (S) Strengths are considered mostly internal. They include all the core powers and potencies the business may have.

Weaknesses (W) They are the internal resistive factors which the business will need to address to run successfully. For a start-up business, an example might be a lack of experience in the selected industry.

Opportunities (O) They are the external advantageous factors including Political, Economical, Social and Technological which may facilitate smooth running of the business.

Threats (T) They are the external hiccups including Political, Economical, Social and Technological which may hamper smooth running of the business.

3. Analysis of Work Done

We followed the following steps to reach our final deliverable:-

- i. List down one practical example each for all the seven existing online business models.
- ii. Analyse all seven of the examples selected based on the 7Cs of Internet Marketing.
- iii. On the basis of the 7C analysis, come up with a SWOT for each of them.
- iv. Combining all the 7 SWOT analysis of all the seven different business models and come up with a consolidated SWOT.
- v. Removing the weaknesses and threats from this consolidated SWOT and suggest an ideal 7C online business model.

3.1. Selected List of Practical Examples

Table 1: Selected list of Practical Examples

| Online Business Model | Website Chosen |
|------------------------------|-----------------------|
| Brokerage Model | Orbitz |
| Advertising Model | Monster |
| Merchant Model | Barnes & Noble |
| Manufacturing Model | Dell |
| Affiliate Model | Amazon |
| Community Model | Flickr |
| Subscription Model | Netflix |

3.2. 7C Analysis of Orbitz.com

Orbitz Worldwide Inc is an internet travel company headquartered in Chicago, Illinois. Through its

primary website Orbitz.com, it enables travellers to research, plan and book a broad range of travel products, facilitating 1.5 million flight searches and one million hotel searches every day.

| C# | Seven Cs | Description |
|-----------|-----------------|---|
| C1 | Context | It has a very simple look and feel. Nothing so eye-catching could be found at the first go, but all the basic information is readily available. However, certain points need to be noted as below:- |

| | | |
|----|---------------|--|
| | | <ul style="list-style-type: none"> • Search and Navigation are convenient but the search results could be made clearer. • The latest deals are displayed too prominently at the top of the home page. This may deflect the customers away from the Orbitz services and end up the customer in losing interest. • A unique way of displaying hot deals. It is mentioned as “<i>Deal Detector</i>”. Wherever this is visible, it is understood that the customer can calculate an approximate price for his deals. But the way of display is not so prominent to catch customer’s initial attention. • Most noticeable is the fact that in certain web pages, the alignments are not proper and some of the data got overlapped. |
| C2 | Communicate | <ul style="list-style-type: none"> • Communication facilities are available. Any registered user can ‘Write a Review’ or provide ‘Feedback’ to the company at any time, but only after logging in to the website. Also, he/she can call up the helpdesk dialling the numbers available under the ‘Customer Support’ tab. • Also any interested customer can fix his deals under ‘Deal Detector’. Once suitable deals come to Orbitz, they will inform the customer with auto-generated alerts. Though the facility is available, but the efficiency is quite on the lower side. |
| C3 | Commerce | This site is used entirely for commercial purposes. In fact it acts as the broker between the various flights, cars, hotels and other service providers forming the common platform between them and their potential customers. So, the payment gateway must be secured. They follow a 3D secured digital payment gateway for all the financial transactions done online through Orbitz. |
| C4 | Connection | Apart from links to Facebook and Twitter fan page of Orbitz, there are a few more connection links provided – like Hotel Price Assurance, Top Destinations to Google ads etc. |
| C5 | Content | The site does not have any attractive video or audio files. All data are based on plain text, well aligned except in a few pages (as mentioned above). The in-house ads are presented through simple coloured graphics. |
| C6 | Community | There is no live chat facility available, nor can any customer communicate with other customers through Orbitz’s web portal. Socializing is not allowed through Orbitz. |
| C7 | Customization | Some basic customization of user profiles are allowed wherein the customer can check, cancel or view his trips under ‘My Trips’ tab, can update his deals through ‘Traveller Updates’ and can go for changing his account details through ‘My Account’. No further profile customizations are enabled. |

3.3. SWOT Analysis of Orbitz.com

| <u>Strengths (S)</u> | <u>Opportunities (O)</u> |
|---|--|
| 1. Communication is easy through the website 2. 3D secured digital payment gateway is used 3. Basic customization is allowed. | |
| <u>Weaknesses (W)</u> | <u>Threats (T)</u> |
| 1. Not eye catching 2. Lacks Audio-Video content 3. Socializing is not allowed through Orbitz | 1. A lot of outside links are present in the website |

3.4. 7C Analysis of Monster.com

It is one of the employment websites in the world, owned by Monster Worldwide Inc. It was created in 1999 by the merger of The Monster Board (TMB) and Online Career Centre (OCC).

Monster.com is primarily used to help those seeking work to find job openings that match their skills and location.

| C# | Seven Cs | Description |
|-----------|-----------------|--|
| C1 | Context | The homepage is too cluttered and poorly designed. The initial attention of a visitor can be rarely achieved through such an unattractive web layout. |
| C2 | Communicate | Communication is quite poor from Monster's side which should be one of the prime features available in any Advertising model. <ul style="list-style-type: none"> • Emails, Newsletters and SMS alerts are the only ways through which Monster communicates to the job seekers. • For corporate customers or employers, a separate gateway is provided which drives them into their pool of registered job seekers. • There is no way a job seeker can communicate directly with Monster people other than uploading his/her resume. There is a 'Help' tab, but it is rarely responded. Also, no live assistance is available. |
| C3 | Commerce | No payments are made through this web portal except paid resumes. The payment gateway is 3d secure. |
| C4 | Connection | Some out bound links are posted at the homepage, which keeps popping after every 60 seconds. |
| C5 | Content | The site contents are too poorly displayed in plain text. No audio or video files are there. |
| C6 | Community | Community is provided under the 'Communities' tab, but the facility is rarely used by the users. |
| C7 | Customization | Resume and profile customizations are widely available. |

3.5. SWOT Analysis of Monster.com

| Strengths (S) | Opportunities (O) |
|--|--|
| 1. The payment gateway is 3d secure 2. Community is provided under the 'Communities' tab 3. Resume and profile customizations are widely available | |
| Weaknesses (W) | Threats (T) |
| 1. The homepage is too cluttered 2. Communication is quite poor from Monster's side 3. Contents are too poorly displayed in plain text | 1. Some out bound links are posted at the homepage |

3.6. 7C Analysis of Barnes&Noble.com

It is the largest book retailer in the United States, operating mainly through chain of bookstores, headquartered at Lower Manhattan, New York

City. Most of the stores also sell magazines, newspapers, DVDs, gifts, games and music.

2nd Annual International Conference IEMCON 2012

| <u>C#</u> | <u>Seven Cs</u> | <u>Description</u> |
|-----------|-----------------|---|
| C1 | Context | The homepage layout is quite attractive. It helps in attracting the initial attention of the visitor. <ul style="list-style-type: none"> • All the newly available books are displayed attractively in the homepage. • Free membership offer is prominently highlighted which is one of the key features to attract new customers. • Free shipment facility is available for online orderers. But the link is not quite prominent. • B&N is new into the publishing sector of ebooks. The link to which is highlighted under ‘Pub-It’ link. |
| C2 | Communicate | B&N communicates to its registered customers through Newsletters and Emails. |
| C3 | Commerce | A very secure 3D payment gateway is available. A special B&N Master Card is offered to its members which is an unique feature of B&N. |
| C4 | Connection | It has no out bound links in the homepage. All the links refer to the in house offers and deals only. |
| C5 | Content | A number of uniquely attractive facilities and features are available in B&N including – ‘Nook’, ‘Pub-It’, ‘B&N Studio’, ‘Teens Corner’ etc apart from traditional Books, Textbooks, Children Books etc. |
| C6 | Community | Users cannot form communities themselves. |
| C7 | Customization | High customization is available. Users can create their profiles, add or remove items from their carts and view them anytime they want. |

3.7. SWOT Analysis of Barnes&Noble.com

| <u>Strengths (S)</u> | <u>Opportunities (O)</u> |
|--|---|
| 1. The homepage layout is quite attractive 2. B&N communicates to its registered customers through Newsletters and Emails. 3. A very secure 3D payment gateway is available. 4. uniquely attractive facilities and features are available | 1. It has no out bound links in the homepage. |
| <u>Weaknesses (W)</u> | <u>Threats (T)</u> |
| 1. Users cannot form communities themselves | |

3.8. 7C Analysis of Dell.com

laptops and desktops for household, SME business, Public sector and Large enterprises.

Dell specializes in manufacturing of customized

| <u>C#</u> | <u>Seven Cs</u> | <u>Description</u> |
|-----------|-----------------|--|
| C1 | Context | The web layout is too simple which may be improved for better customer turn over. The tabs are well organized, but the major portion of the homepage is made occupied by the new deal news. |
| C2 | Communicate | Dell’s communication portal is very fast and efficient which earns it innumerable repeat and new customers throughout the world. Emails, Newsletter, SMS alerts along with tele-calling services are available. But nowhere in the website is it clearly |

| | | |
|----|---------------|---|
| | | mentioned that it offers its book retailing services only in US. |
| C3 | Commerce | The payment gateway is 3D secured and absolutely safe. Different product rates are offered for Home, SME business, Public sector and Large enterprises. The price factor is maintained quite at a competitive scale owing to its direct model. |
| C4 | Connection | No out bound links are present. |
| C5 | Content | Basic and add on information are well represented. <ul style="list-style-type: none"> • Dell offers different product range and categories to household and corporate customers. The differentiation is made clearly under different tabs. • For knowing the company information, its different offices and contact information, one needs to scroll down. Generally, for the visitors this information is need driven. So, scrolling down can be a good option and optimum space utilization is done. • The new deals are displayed in simple yet attractive language like – ‘Run your business leaner; the tag line itself gives an overall idea of the new product. |
| C6 | Community | Separate accounts are made available for Premium and Standard users. The members can interact among themselves, not live but offline. |
| C7 | Customization | Poor customization features are enabled. |

3.9. SWOT Analysis of Dell.com

| <u>Strengths (S)</u> | <u>Opportunities (O)</u> |
|--|------------------------------------|
| 1. Dell's communication portal is very fast 2. The payment gateway is 3D secured 3. Basic and add on information are well represented. 4. The members can interact among themselves | 1. No out bound links are present. |
| <u>Weaknesses (W)</u> | <u>Threats (T)</u> |
| 1.Poor customization features | |

3.10. 7C Analysis of Amazon.com

Amazon is arguably the most widespread online bookseller. Along with online books, it popularized

the concept of Kindle as well. Kindle is that device which helps one read digital books.

| <u>C#</u> | <u>Seven Cs</u> | <u>Description</u> |
|-----------|-----------------|--|
| C1 | Context | The web layout is fantastic in all respects – be it from basic to featured information, new deals information, payment information, special offers, current deals etc. The list of tabs at the left top corner of the homepage gives a clear view of all the products available to the visitor. New visitor can go for a sign up, the link to which is provided at the top itself. |
| C2 | Communicate | Personalized recommendations are made for registered users, which are clearly mentioned at the top of the homepage. Rest emails and newsletters are always available. |
| C3 | Commerce | Amazon has special VISA cards for its premium users and also provides gift cards which are available with free one day shipping guarantee. |

| | | |
|----|---------------|---|
| C4 | Connection | Some out bound links are available but not so highlighted, like there is a link to Levis offers at the right side of the webpage. |
| C5 | Content | The entire content is beautifully organized and well represented in simple yet perfect way. |
| C6 | Community | No live chats are enabled. |
| C7 | Customization | Highly customized user profiles can be created, right from creating personalized carts, deals and gifts. |

3.11. SWOT Analysis of Amazon.com

| <u>Strengths (S)</u> | <u>Opportunities (O)</u> |
|---|----------------------------------|
| 1. The list of tabs at the left top corner of the homepage gives a clear view 2. Personalized recommendations are made for registered users 3. The entire content is beautifully organized 4. Highly customized user profiles can be created | |
| <u>Weaknesses (W)</u> | <u>Threats (T)</u> |
| 1. No live chats are enabled. | 1. Out bound links are available |

3.12. 7C Analysis of Flickr.com

It is an image and video hosting site, web services suite and online community created by Ludicrop in 2004 and later acquired by Yahoo! The site is

widely used by bloggers to host images that they embed in blogs and social media.

| <u>C#</u> | <u>Seven Cs</u> | <u>Description</u> |
|------------------|------------------------|---|
| C1 | Context | The web layout is too simple. No extra information or links are available. Only three tabs are prominently displayed – Sign In, Search and Photo share. |
| C2 | Communicate | Apart from Newsletters, other modes of communication are sparsely enabled. |
| C3 | Commerce | No commercial transactions are carried out. |
| C4 | Connection | No outbound links are posted. |
| C5 | Content | The content is too condensed and almost no web designs and graphics are there. The homepage has an attractive photo, which is tagged as the photo of the month by Flickr, based on maximum votes. |
| C6 | Community | Users and members can form communities and can interact with each other in live chat rooms. |
| C7 | Customization | Customization features are remotely available, where one can only upload, organize and share his/her photos. |

3.13. SWOT Analysis of Flickr.com

| | |
|--|---------------------------------|
| <u>Strengths (S)</u> | <u>Opportunities (O)</u> |
| 1. Users and members can interact with each other in live chat rooms | 1. No outbound links are posted |
| <u>Weaknesses (W)</u> | <u>Threats (T)</u> |
| 1. The web layout is too simple 2. Apart from Newsletters, other modes of communication are sparsely enabled 3. The content contains no web designs and graphics 4. Customization features are remotely available | |

3.14. 7C Analysis of Netflix.com

It offers online flat rate DVD and Blue-Ray disc rental by mail and video streaming only in the USA and Canada. It was established in 1997 and is

headquartered in California. It started its subscription services from 1999.

| <u>C#</u> | <u>Seven Cs</u> | <u>Description</u> |
|------------------|------------------------|--|
| C1 | Context | Very attractive and bright web layout which is easy for any visitor to pick up what he wants. The website design is eye-catching enough to get initial attention from the visitors. The very ad line – ‘Watch as many movies as you want....’ Is clear enough to convey the best deals to its customers at the first go. |
| C2 | Communicate | All communications are only possible if the customer has a valid US mailing address. |
| C3 | Commerce | In any Subscription model, this is the most important aspect. <ul style="list-style-type: none"> • A free trial for one month is provided by Netflix to its potential customers. The payment is made entirely through an online secure gateway. • The users need to subscribe to Netflix at a monthly basis for enjoying continued services. |
| C4 | Connection | There are no out bound connections or links in the homepage. |
| C5 | Content | The tabs – ‘Browse your selection’ and ‘How it Works’ give a clear instruction to the new users about the working and facilities of Netflix customers. Also, anytime the customers can cancel the subscription, the link to which is clearly available at the top. |
| C6 | Community | Users cannot socialize using Netflix web-portal. |
| C7 | Customization | Seldom customizations are enabled. |

3.15. SWOT Analysis of Netflix.com

| | |
|---|--|
| <u>Strengths (S)</u> | <u>Opportunities (O)</u> |
| 1. Very attractive and bright web layout 2. The tabs give clear instructions | 1. There are no out bound connections or links in the homepage |
| | |

| <u>Weaknesses (W)</u> | <u>Threats (T)</u> |
|--|---------------------------|
| <ul style="list-style-type: none"> 1. All communications are only possible if the customer has a valid US mailing address 2. Users cannot socialize using Netflix web-portal 3. Customizations are seldom enabled | |

3.16. SWOT Validation

We conducted a random survey among the Facebook fans of the above websites and requested them to give scores to these websites on the basis of the 7Cs in a scale of 1 to 10. This helped us to understand the relative importance of the strengths and weaknesses of these selected websites and this understanding ultimately led us to the design of the ideal website layout, irrespective of the online business model it follows. The result of the survey is depicted below, with their average overall scores as received from the Facebook respondents:-

Table 2: Average Overall Score

| <u>Selected Website</u> | <u>Average Overall Score</u> |
|--------------------------------|-------------------------------------|
| Orbitz | 4.5 |
| Monster | 3.5 |
| Barnes & Noble | 6.75 |
| Dell | 4.5 |
| Amazon | 7 |
| Flickr | 4.5 |
| Netflix | 7 |

As it can be seen from the above table, the websites with higher strengths received higher average overall scores.

3.17. Graphical Analysis of the Above Results

3.17.1. Plot of Average Overall Score

The schematic below represents the Average Overall Score of all the seven websites, as received from the Facebook fans of the respective sites.

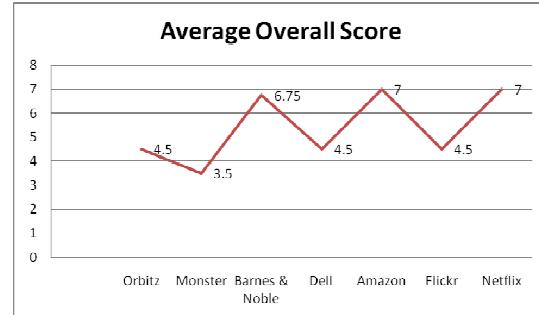


Figure 1: Plot of Average Overall Score

3.17.2. Plot of Facebook 'Likes'

The 7C based above SWOT analyses can also be validated from the Facebook pages of the respective websites. The companies with higher number of Strengths have higher number of 'Likes' in Facebook. The Facebook "Likes" are depicted through the following graphical analysis:-

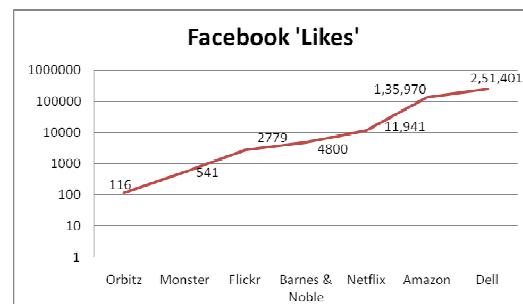


Figure 2: Plot of Facebook 'Likes'

4. Our Findings

4.1. Consolidated SWOT

From the 7C based SWOT Analysis of all the online business models, we found that there are a number of positive and negative factors which must be taken care of while designing any online business, irrespective of the model it follows. We have consolidated all the SWOTs into a master one with all their Strengths, Weaknesses, Opportunities and Threats combined into a single SWOT Matrix as below:-

| <u>Strengths (S)</u> | <u>Opportunities (O)</u> |
|--|--|
| <ol style="list-style-type: none"> 1. Communication is easy through the website 2. 3D secured digital payment gateway is used 3. Customization is allowed 4. Community is provided 5. Attractive homepage layout 6. Personalized recommendations for users 7. Clear instructions to users | <ol style="list-style-type: none"> 1. There are no out bound connections or links in the homepage |
| <u>Weaknesses (W)</u> | <u>Threats (T)</u> |
| <ol style="list-style-type: none"> 1. Not eye catching 2. Lacks Audio-Video content 3. Socializing is not allowed 4. The homepage is cluttered 5. Poor Customization features 6. Lack of communication features | <ol style="list-style-type: none"> 1. Lot of outside links are present in the website |

4.2. Comparison with Facebook ‘Likes’ and Average Overall Scores Received

Table 3: Facebook ‘Likes’ and Overall Score Comparison

| <u>Selected Website</u> | <u>Average Overall Score</u> | <u>Facebook ‘Likes’</u> |
|--------------------------------|-------------------------------------|--------------------------------|
| Orbitz | 4.5 | 116 |
| Monster | 3.5 | 541 |
| Barnes & Noble | 6.75 | 4800 |
| Dell | 4.5 | 251,401 |
| Amazon | 7 | 135,970 |
| Flickr | 4.5 | 2779 |
| Netflix | 7 | 11,941 |

4.3. Graphical Comparison

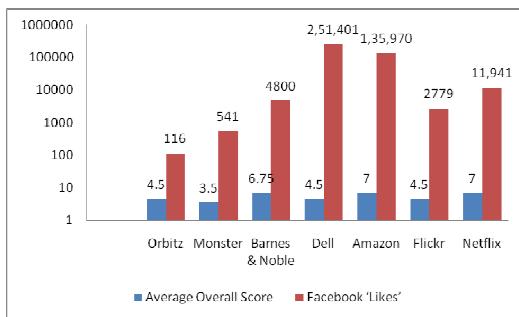


Figure 3: Combined Plot of Facebook ‘Likes’ and Average Overall Score

4.4. Recommendation

As we can see the number of Facebook ‘Likes’ is in synchronization with the Average Overall Score of the websites. Therefore, if the weaknesses and the

threats as arrived from the consolidated SWOT can be removed, an ideal layout of online business can be arrived at, irrespective of the model it follows.

5. Conclusion

By looking at all the existing works, we can conclude that so far no work has been done attempting the integration of all existing online business models. In this research paper, the same has been attempted and the consolidated SWOT can work as the basis of designing an integrated secured web-layout for any organization intending to enter into the online economy.

6. References

- [1] Dave Chaffey, “*Best Practices: E-business and Online Marketing*”, Computer Associates International, June 2005.
- [2] Aden George, “*Seven Cs of Internet Marketing*”, <http://adengeorge.com/>

Cloud Computing: emphasizing its possible roles and importance in Information Systems & Centers

Prantosh Kumar Paul

Research Scholar

EB Swastik Academy of Education & Training (CIIP-121314)

EIILM University, Jorethang, Sikkim, India.

apexpgi@rediffmail.com

Bhaskar Karn

Associate professor

Dept. of Information Technology

Birla Institute of Technology

Mesra, Ranchi

bhaskar@bitmesra.ac.in

Abstract:-Cloud Computing is actually a new kind of supplement, consumption as well as delivery mechanism or model for information technology and system service based on the internet or network technology .Here in this article we describe the concept of cloud, its evolution, benefits and impact on information centre, information analysis centre, information systems. This article describes how the cloud computing can customize and create a new kind of users experience of information centre & systems and also employers of the concerned institutions.

Keywords:-Computing, Cloud IT infrastructure, Collaborative Technologies, Information Centre, Information Systems, IT Trends.

OBJECTIVES:-

- To find out the concept and implication of cloud computing.
- To know the evolution and latest trends of cloud computing.
- To draw a possible picture of cloud computing uses in information centre and systems.
- To find out the strength, weakness, opportunities and threats of cloud computing in terms of its uses in information systems and Networks.

I. INTRODUCTION

Cloud Computing has emerged in the field of Science and Technology. It works wonder in infrastructure, platform, and software and so on. Many benefits can be derived from the cloud computing as it is one of the vital branches of technology. It has already made its mark in Information Technology, Infrastructure, in Business and commerce,

Dipak Chatterjee

Principal

Institute of Engineering & Management (IEM)

Saltlake, Kolkata, India.

d.chatt@rediffmail.com

creating opportunities to save money, to provide and maximize flexibility, efficiency [11]. It also devised a new way of rapid solution in Information Technology. It is obvious that it can create a better way to engaged it self for the benefit of information users in information centre and systems. Cloud Computing is also trying its best to unite individuals and customers in a new way by connecting them with it.

Before going to discuss about aspects and importance of Cloud Computing. Let's cheek out the following general terms and concepts.

Information-In simpler terms, the processed data is information. Information consists of data that have been retrieved, processed or otherwise used for informative or inference purpose, arguments or as a basis for forecasting or decision making.

Information Centre- Information Centre is an information institution which is dedicated to collect, select, organize, manage and disseminate information, depending upon users needs.

Information Networks & Systems-Information Network and System is a kind of consortium which is mainly dedicated to information collection, organization, delivery within a group of similar organization by Sharing Software, Hardware, Content and ideas. It is also an important component and aspect of Information Networks.

Computing- Computing includes designing and building hardware and software systems for a wide range of purposes;

2nd Annual International Conference IEMCON 2012

processing, structuring, and managing various kinds of information; doing scientific studies using computers; making computer systems behave intelligently; creating and using communications and entertainment media; finding and gathering information relevant to any particular purpose, and so on (ACM).abbreviations and Acronyms

II. CLOUD COMPUTING: AN OVERVIEW

Cloud Computing is one type of electricity grid or mechanism in which internet and networks play an important role. In cloud environment it is possible to share resource, software and information to other computers and IT devices, if required. The term cloud is used as a metaphor for the internet in computer networks diagrams as an abstraction of the underlying infrastructure it represents .Lots of benefits are possible today and for almost all kinds of corporation,industries;be it private,public,small or large. In most of the cases cloud computing users do not own the physical infrastructure instead of avoiding capital expenditure by renting usage from third party provider .so here is a chance to save the resources due to its utility computing architecture. As cloud computing infrastructure consists of service delivered through common centre and built on servers, its saves money and time also [11].

environment to client server architecture or model in the early in 1980s and beyond. However it is important to note most cloud computing infrastructure consist of common centre and built on server, the computing history tell us that, the actual term cloud borrows from telephony in that telecom companies, who until the 1990s primary offered dedicated point to point data circuits, began offering virtual private network (VPN) services with comparable quality of services but at a much lower cost. So it is clear that, the main concept of cloud computing is old but with new connectivities, efficiencies and technologies. one of the new aspect of cloud computing is that it create a new set of relationships between individuals and institutions, people, data and technologies taken personally sensisitive information out of a desk drawer or similar and started to store in on their hard drivers.

IV. INFORMATION CENTRE & SYSTEMS: EMPHASIZING CLOUD ENVIRONMENT

Information centre is an information hub, which is mainly dedicated to collect, select organize, manage and dissemination of information, depending upon user needs. Information centre may be on a specific subject or discipline oriented or project oriented including institutional information within a corporate or organization. However if we consider discipline or subject oriented information centre dedicated to the community we can draw a picture, in which we can use cloud computing model or architecture.

V. CLOUD COMPUTING IN INFORMATION CENTRE AND SYSTEMS: AREAS: -

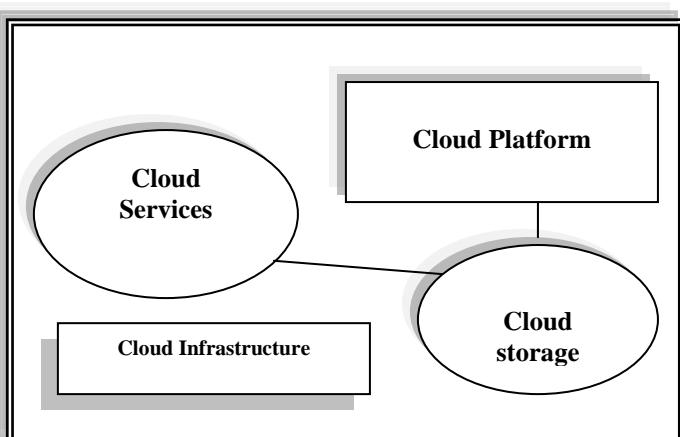
We can see the following areas where the application of cloud computing is possible.

In collecting and selecting information it is possible to use cloud computing models. As the information centre collects information from the various media, medium (like books or documents). So it is no doubt possible to use cloud computing through the peers of information suppliers or information delivery agencies (not raw information delivery agencies). This enables easy and fast information collection. Similarly it is possible to select information which is really needed in peer reviewing with the help of cloud network. And we can reduce the time which we usually take to perform by using cloud system. So ultimately it can save money by avoiding the use of Information Technology Infrastructure within a short span at time. With the help of cloud computing it is also possible to attempt a new form of flexibility that comes from newer technologies.

In terms with information organization and management cloud computing plays an important role by quick information classification and organization.

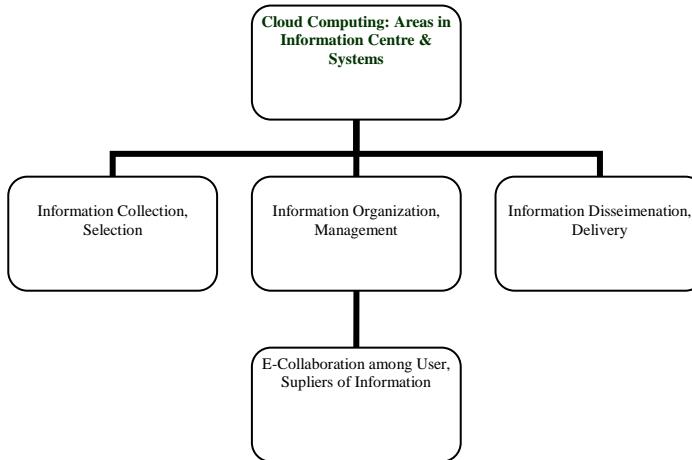
III. CLOUD COMPUTING: EVOLUTION AND PRESENT SCENARIO

Cloud computing the concept is not fully new; this is actually a paradigm shift from mainframe computing



This kind of environment does not give important to capital expenditure on hardware, software and services and is applicable only to pay the used one. The cloud computing is reliable as it is well designed and scalable as it dynamic in nature. Cloud computing also saves future expenses regarding hardware and its maintenance. Cloud computing also keeps robust securities by dint of centralization, but accessible nature [09].

2nd Annual International Conference IEMCON 2012



As a service providing institution the main aim of an information centre is to keep connection with the user. As the centre provides information and documentation flexibly basing on technology.

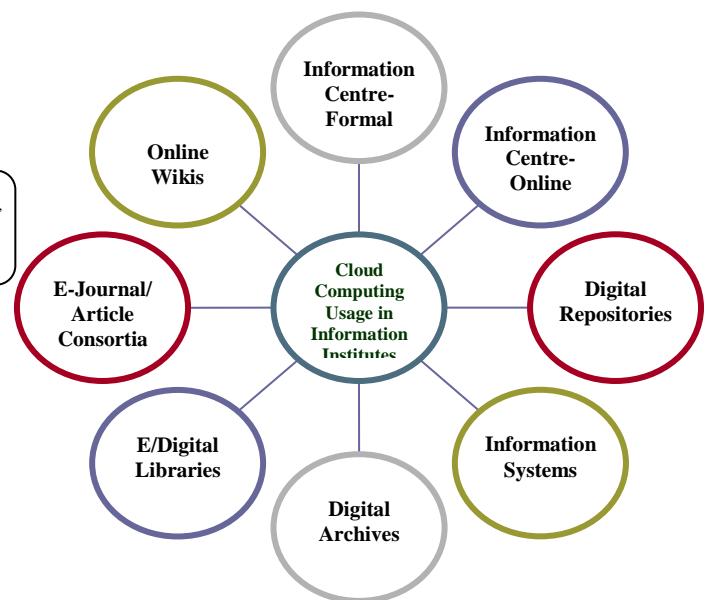
Further cloud computing model can integrate the search techniques, queries and information services of an information centre. So it is possible to increase and simply communicate with users as well as staffs. Ultimately it increases the number of users of the information centre.

Digital repositories or libraries which are fundamentally based on general library system, but digital system can get benefit cloud computing. The registered users can keep connection with the main digital hub. The digital repositories can adopt others system from today's technologies. Cloud environment can reduce the computing cost of digital repositories, e libraries, e- article consortium and others.

The number of users of digital repositories, online information centers will obviously be increased by the cloud computing system.

It is possible to apply cloud computing model in the various sectors of a small/large discipline oriented information systems or consortiums or networks. As information Systems are connected hub of some information centre's, so with cloud computing system/model the information centers can connect strongly with other information centers with the main information systems. This will increase flexibilities of data sharing among the information centers with information centre and information systems or vies versa.

So as a result we ultimately can draw the following benefits from cloud computing to the overall information centres and systems periphery.



Importance and Needs in Information Systems & Centers

- Greater efficiencies in information and documentation system regarding their information collection and dissemination and for good collaboration among the users who are registered with digital services.
- Reduction of cost in information technology infrastructure, hardware, software as there is no need to invest further.
- The deployment of database and software and maintenance of users are hassle free.
- Better connections of the information centre with main information systems.
- It is possible for better collaboration through at any place and any time access to IT for users around the world, especially for digital repository user.
- Cloud Computing provides easy availability of information technology application through subscription or on demand model which is useful for digital repositories, information systems, and online information centre.

VI. FINDINGS:-

- Cloud computing has appeared as one of the important technologies.

2nd Annual International Conference IEMCON 2012

- Cloud computing is facing some challenges from research sector and less government aids or assistance.
- Cloud computing gives opportunities to harness the power technologies in the new, dynamic and creative way without over spending on IT budget.
- Wonderful opportunities are available to use cloud computing systems in the subject based/mission oriented information systems, apart from organizational or corporate information systems.

• The main problems of cloud computing is its utilization in information centre and unwillingness to changes, funding and finance, full and actual delivery of services to the information centre and systems by the IT corporations/companies, who provides cloud computing services.

- However it is possible to use many cloud services, that are already exists as per need of the information centre and systems.

VII. CONCLUSIONS:-

Cloud computing, which is consider as the next era in computing, has the ability to transform conventional Computing to Collaborative, environmental friendly computing. There are some words in fact, that describe the Cloud perfectly: cheaper, faster, flexible, efficient, greener. Frankly speaking enough awareness is still absent in India and abroad in the field of cloud computing in the face of many difficulties and threats. The usage and implementation in cloud computing are really tough and need many aspects like funds, initiatives, more technological collaboration, package software in information centres and systems, documentation centres and related institutes. However with modern research and trends to cloud computing one thing is evident that the future is here.

- [1] <http://cloudcomputing.blogspot.com>
- [2] <http://cloudcomputing.sys-con.com/node/1528536>
- [3] <ftp://public.dhe.ibm.com/common/ssi/sa/wh/n/ciw03067usen/CIW03067USEN.PDF>
- [4] http://en.wikipedia.org/wiki/Cloud_client
- [5] Danielson, Krissi (2008-03-26). "Distinguishing Cloud Computing from Utility Computing". Ebizq.net.
http://www.ebizq.net/blogs/saasweek/2008/03/distinguishing_cloud_computing/. Retrieved 2010-08-22.
- [6] "Cloud Computing: Clash of the clouds". The Economist. 2009-10-15.
http://www.economist.com/displaystory.cfm?story_id=14637206. Retrieved 2009-11-03.
- [7] "National Science Foundation press release. September 2008. "National Science Foundation Awards Millions to Fourteen Universities for Cloud Computing Research." Retrieved 2010-03-01". Nsf.gov.

http://www.nsf.gov/news/news_summ.jsp?cntn_id=114686. Retrieved 2010-08-22.

- [8] Myslewski, Rik (2009-12-02). "Intel puts cloud on single megachip". Theregister.co.uk. http://www.theregister.co.uk/2009/12/02/intel_scc/. Retrieved 2010-08-22.
- [9] http://./AppData/Local/Temp/Temp1_attachments_2011_12_20.zip%23%23%23pkp/11.09.10/cloud/Cloud_computing.htm - cite_ref-20"Nicholas Carr on 'The Big Switch' to cloud computing". Computerworlduk.com.
http://www.computerworlduk.com/technology/internet/applications/insta_nt-expert/index.cfm?articleid=1610. Retrieved 2010-08-22.
- [10] "IEEE Technical Committee on Services Computing". Tab.computer.org. <http://tab.computer.org/tcsc>. Retrieved 2010-08-22.
- [11] "Cloud Computing : the future of computing is here" Microsoft Interface | April - June 2010

Information Science Education and Research: emphasizing contemporary Indian scenario-an overview

Prantosh Kr Paul

Research Scholar

EB Swastik Academy of Education & Training
EIILM University, Jorethang, Sikkim, India
apexpgi@rediffmail.com

Dipak Chaterjee

Principal

Institute of Engineering & Management (IEM)
Saltlake, Kolkata, India.
d.chatt@rediffmail.com

Bhaskar Karn

Associate Professor

Birla Institute of Technology
Mesra, Ranchi, India.
Bhaskar302001@yahoo.co.in

Abstract— Information science is one of the important techno management based subject responsible for social changes. This article gives a state-of-the art overview of Information Science (not LIS) education in India providing review of the scenario of awareness, career prospects and research in respect of Information Science. It also highlights the initiatives and efforts of some Indian institutes by which Information Science has come to the present shape and status, the article also mentions its strengths, weaknesses, opportunities and threats in Indian perspective. For critically analyze we also present the syllabi of Information Science programmes run by some institutes, who believe in modernization and technological implementation on traditional Information Science to create a truly informed society

Keywords- *Information Science, Informatics, Information Studies, Knowledge Management, Library and Information Science, Computer and Information Science.*

I. INTRODUCTION

Information Science is actually a discipline of disciplines. It has come from the subject which deals with information, information services, information activities and digital information. So mainly from library science, archive science, documentation science, computer science it has originated; however this is not truly computer science nor library science. This is an interdisciplinary science; which is a combination of Library Science, Computer Science, Information Technology, Management Science, Operation Research and the subjects that deal with information.

In India, as like most of the Western countries Information Science fall under the Library and Information Science field, but considered as a study of

Information and its various activities involved with Information Technology and Management. The various LIS committees of India treat ‘Information Science’ as a specialized course. So as per Knowledge and scope Information Science course comes under the purview of LIS in contemporary scenario in India.

II. INFORMATION SCIENCE EDUCATION

Especially the subject Information Science with ‘Information Science nomenclature’ was first launched at the University level in the year 1993 by Birla Institute of Technology, Mesra, one of the top ranked university in India. However, two other popular and important central government funded institutes namely DRTC (Documentation Research and Training Center, Bangalore, 1962) under Indian Statistical Institute(ISI) and NISCAIR (National Institute of Scientific Communication and Information Resources, New Delhi), formerly INSDOC, also were engaged in imparting Information Science education and research, but emphasizing social, technical and behavioral dimensions of information, these institutes did not offer ‘degree on Information Science They offer only a course named Associateship in Information Science (equivalent to Masters Degree).

III. DIMENSIONS OF INFORMATION SCIENCE

- In Indian scenario, we can see some of the dimensions of information science education and research. First of all, the universities, colleges and engaged in Library Science (LS) education and research. The next thought is most of the institutes offering Library and Information Science education and Research. The uncommon but hottest thought is offering Information Science, Information and knowledge Management, Information System-education and Research, However in

2nd Annual International Conference IEMCON 2012

respect of Technological involvement; Information Science courses much more IT oriented compare to Library and Information Science courses in India.

IV. UNIVERSITY STRATEGY

After realizing the importance and nature of information science out of two hundred fifty nine (283) state government funded universities only two universities started offering information science degree programmes. The University of Madras (third oldest university in India) is the first to start Library science Diploma courses in India. The first of the state government universities who changed Department name from Department of Library and information Science to 'Department of Information Science' in the year 2002 and offer MSc in Information Science in 2003. There is no doubt it was a great achievement by the University of Madras. After Madras University another government funded university; Viz., West Bengal University of Technology started offering a course in 'Information Science' at the Institute of Engineering and Management, Kolkata in the year 2007.

However out of 44 central universities only one central university viz., Dr. Bhimrao Ambedkar University at Lucknow started 'Information Science' as a post graduate course. It was really a good event for Information Science and its development that, apart from Birla Institute of technology (BIT, Mesra) the government institutes also considered that it was 'Time to change' and "Time to move to Information Science programmes/Movement". In the recent time another renowned deemed university like Birla Institute Technology, Mesra, Manipal University, Karnataka started one 'Information Science' post graduate programme at Manipal Centre for Information Science (MCIS), Manipal. It is also important to note that, apart from the change made in naming department and course; the content and syllabus of the course also were totally changed and involved various Information Technology gradients.

But surprisingly out of these three government universities BabaSahib Bhim Rao Ambedkar University at Lucknow discontinued Information Science nomenclature and again went back to Library and Information Science. In the year 2009 The University of Madras discontinued their MSc in Information Science due to some confidential matters. This fact was not taken sumptuously by the Information scientists and researchers of the country.

V. INFORMATION SCIENCE COURSES

Information Science courses are offered in India mainly by universities directly and not by affiliated institutes or colleges except the Institute of Engineering and Management, Kolkata

which offers Masters Degree in Information Science. One of the famous Information Science and Documentation Research and Training Centre; DRTC, Bangalore offered a course in Information Science earlier but not the Masters Degree Programme. However recently they have introduced MSc degree, but this comes with the nomenclature 'Library and Information Science (LIS)' with much more Information Technology orientation than most other LIS programmes.

Another central government institute NISCAIR (formerly INSDOC) under the Council for Scientific and Industrial Research (CSIR, New Delhi) is also engaged in information science education and research actively, but the fact is that NISCAIR offers a Masters Degree equivalent programme; not at par with MSc/Masters Degree. However unlike DRTC, Bangalore NISCAIR offers Associate Programme in Information Science.

| University/College | Programme Name | Duration | Status |
|-----------------------------------|---------------------------|----------|--------------------------------------|
| BIT Deemed University | MSc – Information Science | 2 years | Regular, from 1993. |
| IEM,WBUT,kolkata | MSc – Information Science | 2 years | Regular, from 2007. |
| Madras University | MSc – Information Science | 2 years | Regular, from 2003, now Discontinue. |
| Dr.BR Ambedkar University,Lacknow | MSc – Information Science | 2 years | Regular, now Discontinue. |
| Manipal University | MSc – Information Science | 2 years | Regular, continue. |

It is noteworthy that Masters Degree courses are offered only by 3 or 4 universities. MSc Information Science course is offered at Birla Institute Technology, Mesra, and at Institute of Engineering and Management, Kolkata from the year 2007 under West Bengal University of Technology, Kolkata and as per data available from the website out of nearly 530 universities; Manipal Centre for Information Science (MCIS), Manipal under Manipal University, Karnataka has recently started offering M.Sc Information Science. But it is surprising to note that not a single university, college, higher educational institute out of 20000+ academic institutions in India offers Bachelors Degree in 'Information Science' or any equivalent programme in information science. Therefore persons who are interested in Information Science education join either Bachelor in Computer Science programme or in

2nd Annual International Conference IEMCON 2012

Information Technology programme or simply in Library and Information Science (LIS) programme. So far as the universities are concerned the following universities of India their offer M.Sc. /Ph.D. in Information Science

So it is a surprising fact that out of 530+ universities as well as Centrally funded institutes only 3 universities offer Information Science Degrees and if we consider about 20000 + higher education institutes, then the fact becomes unbelievable; where as the universities located in the countries like USA, UK, China, Japan and other developed countries or even the universities in the developing countries are coming forward to start courses on Information Science by setting up Departments of Information Science to understand the need of globalization and the significance and importance of information in contemporary management and administration with focus on Information Technology.

VI. ON SYLLABUS

Information Science syllabus presently has two foci; first focus being on IT depended Information Science with grounding of traditional knowledge organization and management (like LIS with higher Informatics Studies); another important focus is on IT depended Information Science with grounding of Information Security and Networking. Frankly speaking, courses with the first focus will be extremely useful in Indian scenario because right now all the Information centers, systems and libraries are not at all computerized. So such institutions need educate IT along with traditional Knowledge of Organization and Information Psychology. We can now come to the conclusion about job prospects of Information Science Graduates after verifying the current courses through the syllabi. The last syllabus of MSc Information Science of Madras University is also given here (now changed to the nomenclature MSc-LIS)

BIT Deemed University: - 1.Foundation of Information Science 2.Theory & Practices of Knowledge Organization and Information Processing 3. Fuzzy Logic & Neural Networks Applications 4 .Internet And Web Technology 5. Data Structures 6.Data Base Management System 7.Operating Systems 8.Computer Communication Networks 9.Management Of Information System & Centers10.Information Psychology &Information Architecture 11.Research Methodology &Quantitative Techniques 12.Digital Library And Multimedia And Other 2 Electives

Manipal University

1.Programming Techniques2 .Computer Architecture 3.Linux & Scripting Languages 4. Operating Systems 5.Computational Mathematics 6.Data Structures & Algorithms7.Database Management Systems 8.Dot Net Technologies 9. Adv. Programming Techniques 10.Computer Graphics11.Computer Networks And other 4 electives, Seminars and Project works

Madras University:-1.Evolution of information science 2.information sources 3.Knowledge organization 4.Introduction to information technology 5.Soft skill 6.Management of information centre Information service 7.Information processing-classification 8.Information and Communication technology 9.Information storage and retrieval 10.Information processing-cataloguing 11.Research methods 12.Preservation of information material 13.Knowledge management 14.Marketing of information 15.Digital libraries And electives, internship.

WB University Of Technology:-1. Mathematics for information science 2.Data structure with c/c++ 3.Computer organization 4.Business communication 5.Library and information science 6.Object technology 7.Operating system and system programming 8.Software engineering 9.Information theory 10.Soft computing 11.Digital image processing 12.Values & ethics in profession 13e-commerce

VII. FACETS AND CHARACTERISTICS

- The institutions offering Information Science, are giving importance to the following course prospect and curriculum designing.
- Develop a syllabus that deals with both traditional librarianship along with digital orientation.
- Give importance to Digital Information Management and System Development.
- Prepare educates who able to join traditional Information service and sector; along with advanced information activities.
- Modernizing whole information sector, activities and retrieval systems.
- Creating interdisciplinary curricula useful for research.
- To prepare graduates who know Computer Science along with Information; its nature, values, characteristics, role and so on.

VIII. JOB OPPORTUNITIES

Job opportunities for Information Science graduates are tremendous due to the nature of this interdisciplinary knowledge cluster. Indeed the curricula are so blended with traditional Information processing and knowledge management with hi-end technological studies that the products turn out to be optimally utilizable. However it is important to note that generally the information science graduates are eligible for the following posts according to Information Science organization, foundation and societies as far as India is concerned. We can categorize the job opportunities as follows-

2nd Annual International Conference IEMCON 2012

Information Officer (CIO) and so on. So here Information Science degree holder has tremendous opportunities to join both government and private organization.

| General | Technological Job or Post | Other Job |
|---------------------|---------------------------|--|
| Librarian | Information Scientist | Teacher/Lecturer of Computing |
| Reference in charge | MIS Professionals | Teacher/Lecturer of LIS |
| Cataloguer | Database Managers | R & D Activities |
| Classificationist | Network Security Expert | Public Relation Officer |
| Indexer | Multimedia | Media Person |
| Abstractor | System Analyst | Archivist |
| Information Officer | Knowledge Engineer | Documentation Officer |
| Information Manager | Knowledge Broker | Educationalist |
| Knowledge Manager | Web Administrator | Consultant |
| | Programmer | Information System Developer & Designer. |

In case of universities and institute of National Importance, Information Science degree holder are eligible for 'Information Scientist' which is consider as special post who understand both information behavior and computing. Generally the institutions needed this post needs a person who have both The LIS and Computing Degree but as a combination of both fields Information Science graduates can join this post as far as India as concerned. As per recommendation of Knowledge Commission which formed under the leadership of Prime Minister of India, India needs near about 1500 new universities. So it is possible for the Information Science Degree holders who know both old and new about information activities to join as various posts which are mentioned above. They can join thousands of Public and Private Academic and Special Libraries and Information Centre located various parts in the country.

NASSCOM, the apex society for Information Technology in India, also mentioned India needs thousands of sophisticated experts who are able to join IT task force/posts apart from Programmer or Software Engineer fields. These posts are Infrastructure Management (Networking Expert), Web Experts, Content Developer and Designer, System Analyst, Chief

IX. RESEARCH POTENTIALS

Most of the Indian universities offering Information Science research programmes offer Doctorate level courses also. No institute offers MPhil in Information Science. .

The main characteristics of the research degree offered by Indian universities are

- (I) Course work integrated degree.
- (II) IT Oriented.
- (III) External thesis evaluation.
- (IV) Focuses on Information Technological involvement in social issues regarding Information systems and services.
- (V) Right now there is no opportunity available

For DSc or Post doctoral research degrees. The institutes, offering Information science programmes, put priorities of research in the following areas-

| Information System Fields | Information Technology Fields | Interdisciplinary Fields |
|---|---|---|
| Information System Development. Information Systems Management. Information Systems Designing. Information Systems Information Systems Securities. Information Systems-Mission Oriented. Health Information Systems | Information Networking. Soft Computing. Data Warehousing. Data Mining. Information Intelligence. Expert Systems. Communication Technologies. MIS Human Computer Interaction. Web Development. Internet and Search Engineering IT Infrastructure Management | Fuzzy Logic. Information Management. Knowledge Management. Information Architecture. Information Psychology. Data Management Information Society. Digital Divide. Digital Repositories. Informatics Metrics Science. |

The above mentioned fields are offered at MCIS, Manipal University, BIT (Deemed University) and Institute of Engineering and Management, West Bengal University of Technology at their research wings. However two other national institutes namely DRTC and NISCAIR are also engaged in active researches and their areas of interest include

2nd Annual International Conference IEMCON 2012

Information Systems, Artificial Intelligence, and Web metrics, Depth Classification, Informatics and others.

X. CONCLUSION

The Information science education and research needs an awareness regarding its benefits, scope and periphery. India needs societies, organizations that will take initiative for standardization of Information Science. its time to move towards launching of information science in the departments of Library Science, Computer Science or merge them together for the betterment of societies. The Society for Information Science, New Delhi and Mysore -based ISIM (International School of Information Management) are giving their best towards motivating and popularizing Information Science, and the I -school concept apart from the institutes offering Information Science course mentioned in this literature.

- [1] Handbook on library and information science. New Delhi: Association of Indian Universities. 1997
- [2] Sharma, Dev Raj. 1997. Doctoral research in library and information science in Indian universities. Palampur: The Author.
- [3] Shera, J.H. 1976. Introduction to library science. Littleton, CO: Libraries Unlimited.
- [4] Steig, Margaret F.. Change and challenge in library and information science education. Chicago, ALA: 96. 1992
- [5] Indian Association of Teachers of Library and Information Science. National Seminar (15: 1997: Mysore). [Resolutions] *Iastlic Newsletter* (March–April): 2. 1998
- [6] Kaula, P.N. "Hundred years of library and information science education" In Kumar, P. S. G. and Vashisth, C.P. Eds., Library and information science in India. New Delhi: Sterling Publishers. 1992
- [7] Keren, Carl. "On information science". *Journal of the American Society for Information Science* 35(2): 137. 1984
- [8] Kumar, P.S.G.. Research in library and information science in India. Indexed by A. Tejomurty and H.R. Chopra. New Delhi. 1987
- [9] Lahiri, Ramansu. "Research in library science in India (1950–95): an account of Ph.D. Programme". *Annals of library science and documentation* 43(2): 59–68. 1996.
- [10] Mitra, C.R. "University-Industry interaction with reference to RECs". *University News* 35(25) June 23:1–4. 1997.
- [11] Prytherch, Ray. "Problems of research". *Information Management Report* (Sept): 18. 1997
- [12] Satija, M.P. "Sources of Indian library and information science dissertations." *Libri* 39(1) March: 71–78. 1989.
- [13] Satija M.P. "Forty years of doctoral research in classification and indexing in India, 1957–1997." *Library Herald* 36(2): 80–87. 1998.
- [14] Paul, P. K "Information Science: Trends and Analysis 2003-07" Dissertation, University of North Bengal, India.2008.
- [15] Sharma, Dev Raj. 1997. Doctoral research in library and information science in Indian universities. Palampur: The Author.
- [16] Shera, J.H. Introduction to library science. Littleton CO: Libraries Unlimited. Steig, Margaret F. 1992. Change and challenge in library and information science education. Chicago, ALA: 96. 1976.
- [17] University Grants Commission (UGC) Curriculum Development Committee on Library and Information Science Report. New Delhi: UGC: 107. 1992
- [18] University Grants Commission (UGC) Curriculum Development Committee on Library and Information Science Report. New Delhi: UGC. 2001
- [19] Varalakshmi, R.S.R.. "Library and information science research in India: subject perspective". *Library science with a slant to documentation* 31(2) June: 91–110. 1994
- [20] Cohen, E. From ugly duckling to swan: Re conceptualizing information systems as a field of the discipline informing science. *Journal of Computing and Information Technology*, 7(3), 213–219. 1999

A simple approach of building a Learning Management System (LMS) to experiment with intelligent tutoring

Lopa Mandal, Avik Dutta, Rahul Roy, Mohit Daga

Department of Information Technology
Institute of Engineering & Management
Kolkata, India

mandal.lopa@gmail.com, d.avik@hotmail.com, rahul_roy10@hotmail.com, mohit.daga@hotmail.com

Abstract— In comparison to traditional classroom teaching, one to one human tutoring has been proved to be much more effective in increasing learning performance. Inspired by the fact, educational researchers have identified Intelligent Tutoring Systems (ITS) as means of providing cost effective yet personalized tuition in many knowledge domains. But, building such a technology enabled, personalized educational environment is a complicated and time consuming task. The current article presents a simple approach of building a Learning Management System (LMS) to give much importance to personalized content delivery with special consideration to the flexibility and domain independence of the LMS to facilitate teachers of different knowledge domains to experiment with intelligent tutoring.

Keywords-intelligent tutoring;LMS; ITS; personalized learning; e- learning

I. INTRODUCTION

It has been observed that students differ from one another mainly in learning styles and capability level. For example, some students may be good in understanding teaching in a particular lecture based mode, but others may require special attention, maybe, an easier diction, or a different media or a different pace. Depending on the learning style, text may not always suffice to satiate the course needs that a student may want. She may learn better if she is shown a video or an animation to understand an extremely complex topic. But in a traditional classroom teaching these different requirements of different students are difficult to cater. Technology enabled education has the scope to incorporate intelligence in education system to provide personalized tutoring in a cost effective way. It can also be a solution for lack of trained teachers in remote places. For implementing technology enabled education, use of Learning Management Systems (LMS) are common. But implementing an LMS is a complex and time consuming task. The present work attempts to build an LMS which is simple and at the same time incorporates the features of personalized learning. The system is independent of the domain to be taught and can be easily authored by teachers of different knowledge domain to build their courses by uploading appropriate course material and also by organizing the course structure to ensure that the students get

the course materials in the means, they are most comfortable with. The teachers can also put different sets of questions of different level of difficulty to assess the progress of students of different capability level. The present work focuses on facilitating teachers of different knowledge domain to experiment with intelligent tutoring in their respective domain.

II. BACKGROUND WORK

The incorporation of intelligence into technology enabled education in order to bring effective changes in education system, started in 1970s. Before that, in the 1960s, researchers created a number of Computer Assisted Instructional (CAI) systems [1]. These programs generated sets of problems designed to enhance student performance in skill-based domains, primarily arithmetic and vocabulary recall. Essentially, these were designed to present the student with a problem, receive and record the student's response, and tabulate the student's overall performance on the task, with little consideration on learning sciences. These usually refer to a frame-based system with hard-coded links, i.e. hypertext with an instructional purpose.

These systems evolved into intelligent computer-aided instruction (ICAI) systems and, then, intelligent tutoring systems (ITS) during 1970s and 1980s. In 1982, Sleeman and Brown first coined the term 'Intelligent Tutoring Systems' (ITS) to describe these evolving systems and distinguish them from the previous CAI systems [2]. Both ICAI and ITS contain explicit knowledge of the subject taught. ITS also attempts to simulate the behavior of an intelligent human tutor in addition to acting as a domain expert. The characteristics of an ITS include the ability to judge a student's knowledge status and learning style, teach a given subject accordingly, to detect student errors, to try to figure out where and how the student made an error, to correct flaws in the student's logic and to clear up any misgivings or misunderstandings the student may have about the material. Sleeman and Brown first use the term 'student model' to describe an abstract representation of the learner within the computer program.

Since the late eighties, the building of ITS got acceleration by the development of cognitive theory of learning viz.

Adaptive Control of Thought (ACT-R) theory of cognition and the wider availability of computers, with adequate speed and memory to run an ITS [3]. While applying a cognitive domain model, Anderson outlined the ACT-R theory of skill knowledge which assumes a fundamental distinction between declarative and procedural knowledge. Declarative knowledge is factual or experiential whereas procedural knowledge is goal oriented and mediates problem solving behaviour. Declarative knowledge can be represented as semantic networks of facts and procedural knowledge can be represented as a set of production rules that associate problem states and problem solving goals with associated action sequences and state changes [4]. Hence, ITS implementations that use a cognitive domain model or support the ACT-R theory generally maintain a knowledge structure of declarative or conceptual knowledge for a particular domain and an associated rule-based production system indicating problem solving procedures normally followed by a relevant domain expert. So, it is observed that, ITS domain models primarily used static question/answer pairings and later semantic networks of related concepts, and production rules to indicate expert problem solving sequences, were incorporated.

III. SYSTEM ARCHITECTURE

The classical ITS architecture has four components, domain model, student mode, teaching model and user interface. The domain model stores the course structure in an organized manner. A course is comprised of different topics. The domain model stores all learning materials that are required to teach the topics along with their pre-requisition information. The student model represents all the information regarding a particular student in the system including his/her profile, learning style/preferences [5], capability level etc. This set of information is used by the system to adapt the system according to the requirement of the student. The teaching module specifies all the decision-making process regarding course planning and adaptation. Finally, the user interfaces have two parts - one for the teacher and the other for the student. The teachers' interface acts as an authoring tool using which a teachers can add and modify the system and its various components [6]. The present work aims to provide the basic requisites of an ITS in a simple and flexible way.

A. Proposed Domain Model

The domain model deals with the course, its structure and components. A course is comprised of different topics. One part of the domain model handles the actual learning materials for each topic, stored as learning objects (LO) in a repository. The other part is concerned with the course structure. It is to be understood that to understand any random topic of a course, certain pre-requisite knowledge is required. For example if a student is required to be taught the topic on 'functions in C', she should be well acquainted with the topic on 'statements and expressions in C' and also the topic on 'data types and variables in C'. So, it is always important that the knowledge be disseminated to a student, follows a certain order. A pre-requisite graph is useful for the purpose. A pre-requisite graph described in Figure 1 is a directed graph where each node is a

topic of a certain course and the directed edges between the nodes depict a prerequisite-of relation between them. An edge from node A to node B implies that, topic A needs to be covered before topic B. After drawing the pre-requisite graph, a path is to be drawn starting from the initial topic to the final topic of a subject. According to Figure 1, four such path is possible which has been shown in Figure 2. It gives the valid sequences in which the topics can be delivered to the students.

The teachers/course authors of the proposed system will be guided to follow any one of the valid paths to specify the sequence in which the learning objects are to be delivered to the students. The authors may also try some other paths from the set of valid paths to deliver the content in a different sequence to experiment with the effectiveness of learning by changing the sequence number of the learning objects.

For each topic, the authors can give more than one learning material which differs from each other in level of difficulty and media to ensure that the students can get material according to their capability level and preferences. For that purpose, while uploading the learning objects, the authors have to provide some information regarding the aforesaid factors, which are stored in the system as metadata. Figure 3 shows the metadata given by the authors. The metadata is used by the teaching module for planning the appropriate content to be delivered to the students [7,8].

B. Proposed Student Model

The student model describes a student by her basic information and level of progress in the system. When a student logs in for the first time, a file is created to monitor her progress throughout her learning process. The student is at first authenticated and the respective courses, which are accessible to her, are shown. Initially the student may select any course. For each course, the student may be provided with learning material of some initial predefined topic according to the preferences given by the student. The student is then evaluated. On the basis of the performance of the test the student may be provided with a learning object on the next topic or maybe presented with an easier learning object on the same topic.

The student is hence assessed on the basis of her test performance. The current state of the student is monitored by

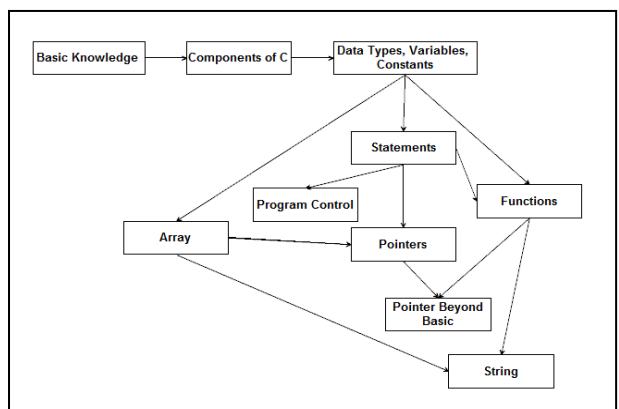


Figure 1. Topic pre-requisite graph

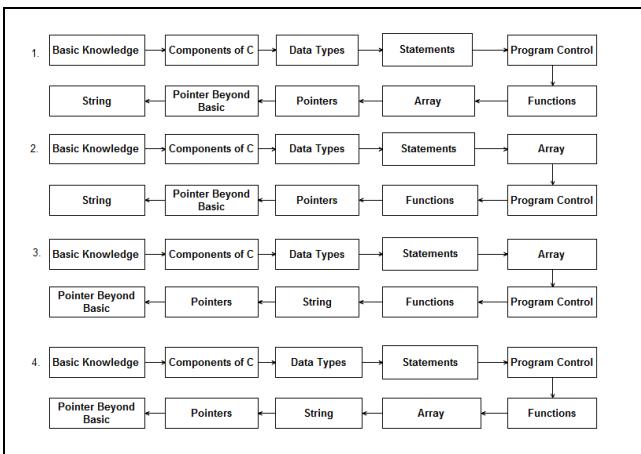


Figure 2. Possible sequences drawn from the pre-requisite graph of Figure 1

the system in terms of topic covered by the students, her capability level and her media preference for the purpose of learning.

C. Proposed Teaching Model

The main goal of the teaching model is to work as a content delivery and evaluation planner for a particular student depending on her profile, preferences and capability level. The content delivery planner starts from the initial topic and terminates when all the topics are delivered according to a

```

...
<course>
    <course_name> Programming in C </course_name>
    <topic>
        <lo_id>1</lo_id>
        <topic_name>Basic Knowledge</topic_name>
        <difficulty_level>easy</difficulty_level>
        <media>text</media>
        <url>Basic1.pdf</url>
        <sequence_no>1</sequence_no>
    </topic>
    <topic>
        <lo_id>2</lo_id>
        <topic_name>Basic Knowledge</topic_name>
        <difficulty_level>hard</difficulty_level>
        <media>text</media>
        <url>Basic2.pdf</url>
        <sequence_no>1</sequence_no>
    </topic>
    <topic>
        <lo_id>3</lo_id>
        <topic_name>Pointer</topic_name>
        <difficulty_level>medium</difficulty_level>
        <media>animation</media>
        <url>Pointer2.swf</url>
        <sequence_no>8</sequence_no>
    </topic>
...

```

Figure 3. A fragment of the course data

sequence specified by the course author. For each topic, more than one learning objects of different level of difficulty and different media are available in the repository. At first, a predefined topic of a particular difficulty level is delivered to a student. For each topic of a particular difficulty level, there is a test set of the same difficulty level, available in the repository. After a topic is taught to a student, the corresponding test set is given to her to evaluate her level of understanding. Depending on her performance, the next learning object is decided to be delivered using the following logic: Certain threshold score is decided. For example 50% is the threshold for the student to graduate to the next topic. If a student scores below the threshold it is assumed that the student does not have proper understanding of the current learning object and hence is not allowed to move on to the next topic. At such a point of time, the student is provided with a learning object based on the same topic but of an easier standard. However an easier standard does not imply that the content coverage of learning object is less. If in the second attempt also the student fails in the test, it is assumed that the given media for the student should be changed. It is to be noted that the present system has got provisions for different media (i.e. audio, video and animation) in which the learning modules are to be presented. If a student scores above the threshold, it is assumed that she has grasped the topic and is eligible to be presented with the next topic. For a score in the range of 50% to 90% the student is provided with the learning object of a medium level for the next topic. However if the student does exceedingly well, i.e. scores above 90%, she is provided with the learning object of the next topic but of a greater level of difficulty. The concept explained above provides an overview of the adaptive/personalizes learning that the system delivers. The process logic of the teaching model is elaborated in structured English in Figure 4.

D. Proposed User Interface

The user interface has two parts – one for the teachers/course authors and the other for the students. The present system provides a user friendly interface for the teachers of different knowledge domain to enter the learning materials. While entering the learning materials the interface

```

Provide a learning object on a topic of a particular difficulty_level
Give a test to the student on the same topic and of same difficulty_level
IF Student has obtained marks less than 50
    IF difficulty_level of the learning object is Easy
        Provide Learning Object for same topic of same difficulty_level but of
        a different media
    ELSE
        Provide Learning Object for same topic and of a lower difficulty_level
    ENDIF
ELSE IF Student has obtained marks between 50 and 90
    Provide Learning Object for next topic and of the same difficulty_level
ELSE IF Student has obtained marks more than 90
    IF Level difficulty_level is Hard
        Provide Learning Object for next topic and of the same difficulty_level
    ELSE
        Provide Learning Object for next topic and of a higher difficulty_level
    ENDIF
ENDIF
Repeat the above steps until all the topics of the course are delivered.

```

Figure 4. Process logic of the teaching model

IV. CONCLUSIONS

The present work elaborates a learning environment which attempts to incorporate some intelligence while delivering learning contents to students and evaluating them depending on their profiles and capability level. The authoring tool of the system allows teachers of different knowledge domain to upload the contents of different level of difficulty and of different media in a very user friendly way. The current work allows the teachers to have some control in changing the sequence of topics which is used by the content delivery planner. This helps teacher to experiment with intelligent tutoring in their respective knowledge domain.

The present work will be extended by conducting survey to measure learning effectiveness of the students by using the current system and improving it accordingly.

REFERENCES

- [1] L.Uhr, "Teaching machine programs that generate problems as a function of interaction with students," Proceedings of the 24th National Conference, 1969, pp. 125-134.
- [2] D. Sleeman and J. S. Brown, "Introduction: Intelligent Tutoring Systems," in Intelligent Tutoring Systems, Ed. Academic Press, 1982, pp. 1-11.
- [3] J.R.Anderson, "The Architecture of Cognition," Cambridge, Massachusetts: Harvard University Press,1983.
- [4] A.T. Corbett and A.Bhatnagar, "Student modeling in the ACT Programming Tutor: Adjusting a procedural learning model with declarative knowledge," Proceedings of the Sixth International Conference on User Modeling, New York: Springer-Verlag Wein,1997.
- [5] R.M. Felder, L.K. Silverman, "Learning and Teaching Styles in Engineering Education," Engr. Education, Vol 78(7), 1988, pp. 674-681.
- [6] T. Murray, "Authoring Intelligent Tutoring Systems: An analysis of the state of the art," Int. Journal of Artificial Intelligence in Education, Vol. 10, 1999, pp. 98-129.
- [7] J. Brase, M. Painter, "Inferring Meta-data for a Semantic Web Peer-to-Peer Environment," Educational Technology & Society, 7(2), 2004, pp.61-67.
- [8] A. Garrido, E. Onaindia, O. Sapena, "Automated planning for personalised course composition," Ninth IEEE International Conference on Advanced Learning Technologies, 2009, pp. 178-182.

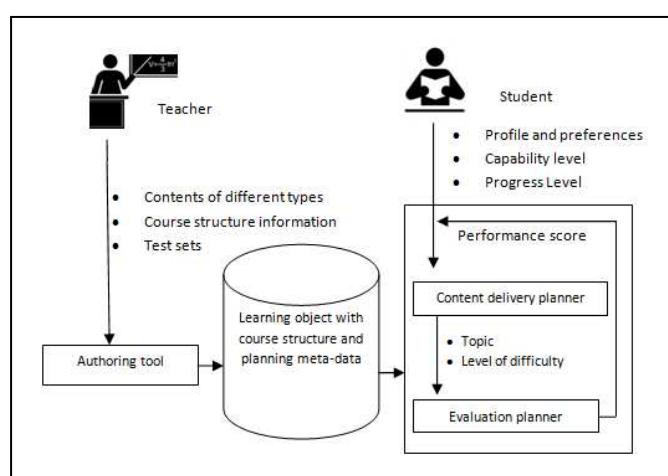


Figure 6. The overall system architecture

FII_s and Indian Stock Markets – Emerging Issues From A Literature Review Post-2000

**Dasgupta Ranjan
& Dutta Sujit***

Like many other emerging countries, India opened up its stock markets in 1992 to FII_s after adopting a liberalized economic policy to counter different crises. From then onwards, the FII_s have taken active role in developing this country. However, there are contradictory viewpoints in this regard. This present study has attempted to review all such viewpoints by providing most of the conceptual and empirical studies on FII_s causes & relationships, nature, events, impact, and extent in regard to Indian stock markets. Such evaluation has raised ten new issues needed to be addressed by the future researchers. In conclusion, this study has pointed out the enormous scope of Indian stock markets in future to attract FII_s investment in larger quantities provided the regulators develop transparent and more authentic stock market situations.

Key Words: FII_s, Indian stock markets, Conceptual, Empirical, Studies, Review, Issues

1) Introduction:

The remarkable economic growth during the past two decades in most of the emerging countries including India had been stimulated by foreign capital inflows from developed countries. Most of the developing countries opened their capital markets to Foreign Investors either because of inflationary pressures, widening current account deficits, exchange depreciation, increase in foreign debt or as a result of liberal economic policy. India faced debt crisis in late 1980s and BOPs crisis in 1990-91. These two problems led to the third one, namely, shortage of investible resources. Thus, the Indian government was forced to initiate financial liberalization programmes. As a result, the foreign investments regime was also liberalized to a great extent [Gordon and Gupta (2003)].

An important milestone in the history of Indian economic reforms happened on September 14, 1992, when the Foreign Institutional Investors (FII_s) were allowed to invest in all the securities traded on the primary and secondary markets, including shares, debentures and warrants issued by companies which were listed or were to be

listed on the stock exchanges in India and in the schemes floated by domestic Mutual Funds (MFs). Ananthanarayanan, Krishnamurthi and Sen [2004], Lalitha [1992] and Prasanna [2008] observed that attracting foreign capital was appeared to be the prime reason for opening up of the stock markets for FII_s.

The impact of FII_s on Indian stock markets all these years has been enormous as the figure has shown that FII_s net investment has risen from Rs.2,595.10 crores in 1993 to Rs.4,44,292.09 crores till November, 2011. The number of FII_s registered with SEBI has also increased from 18 in 1993 to 1743 by November, 2011.

All these facts have suggested that FII_s have played a very important and dominant role in building up Indian economy and also given the country a respectable place in the global community. But, there is a contra view which observed that FII_s brought ‘hot money’ and at the slightest smell of financial crisis around they withdraw their investments from Indian stock markets (see Table 2).

The role of FII_s on the host country stock market has been a subject of interest to academicians and policy makers from decades. In the Indian context, there has been more noise on this issue from different ideological perceptions with very little empirical analysis. In this study, we are

* Asst. Professor
Institute of Engineering & Management, Kolkata,
Saltlake
College of Management

attempting to summarise the viewpoints of the studies post-2000 explaining the causes & relationships, nature, events, impact, and extent of FIIs in Indian stock markets.

2) FIIs – A Brief Review:

An FII is an investor generally of the form of an Institution or Entity situated outside India, which invests money in the financial markets of India. Also, a domestic asset management company or domestic portfolio manager who manages funds raised from outside India for investment in India on behalf of a sub-account shall be deemed to be an FII. These investors must register with the Securities & Exchange Board of India (SEBI) to take part in Indian markets.

Thus, ‘FIIs’ include “Overseas pension funds, mutual funds, investment trust, asset management company, nominee company, bank, institutional portfolio manager, university funds, endowments, foundations, charitable trusts, charitable societies, a trustee or power of attorney holder incorporated or established outside India proposing to make proprietary investments or investments on behalf of a broad-based fund” [Schedule 2 of the Regulation 5(2), of Notification No. 20/RB-2000 dated May 3rd, 2003].

The ceiling for overall investment for FIIs is 24 per cent of the paid up capital (PUC) of any Indian company. The ceiling of 24 per cent for FIIs’ investment could be raised up to sectoral cap/statutory ceiling as prescribed by Government of India (GOI)/Reserve Bank of India (RBI), subject to the approval of the Board and the general body of the company passing a special resolution to that effect.

NSE (2001) observed that in the Indian stock markets FIIs have a disproportionately high level of influence on the market sentiments and price trends. This is so because other market participants perceive the FIIs to be infallible in their assessment of the market and tend to follow the decisions taken by FIIs. This ‘herd instinct’ displayed by other market participants amplifies the importance of FIIs in the domestic stock markets in India.

Table 1 has pointed out that BSE SENSEX and S&P CNX NIFTY Indices value mostly (except 2006-07 and 2010-11 [marginally]) has a positive correlation with FIIs net investment/

different conceptual and empirical Indian studies inflows in Indian stock markets. % change in FII net inflows under Table 1 has also shown that the years (i.e., 2003-04 and 2009-10) in which such change were the highest, it resulted in highest improvement of both the Indices of Indian stock markets in % change terms. However, this was not true in the year 2005-06. So, the impact of the FIIs in the development of the Indices and overall Indian stock markets is quite evident.

Table 2 has also shown the critical role of the FIIs in maintaining the stability and volatility of Indian stock markets. It is shown in Table 2 that in all the heavy down days (i.e., Dates) in Indian stock markets, the FIIs were the net sellers (except 18.01.2007 and 04.02.2011). However, the same is not true in case of MFs.

3) Objectives of the Study:

This present study has the objectives to review the viewpoints of most available conceptual and empirical studies on FIIs in relation to Indian stock markets in the areas of causes & relationships, nature, events, impact, and extent. This study has also aimed at raising the relevant and important issues based on above findings to be addressed by future researchers in their studies.

4) Methodology:

This study has been done on the basis of most available conceptual and empirical studies on FIIs in relation to Indian stock markets during the period of 2000-2010. All such literature has been collected from the published sources and internet. For preparing the Tables, this study has taken help of SEBI Annual Reports, NSE Fact Books, India Index Services and Products Ltd. (IISL), BSE Statistics, RBI Bulletin and different websites, such as, moneycontrol.com.

5) Review of the Literature:

5.1) Causes & Relationships Studies Review:

The question is why the FIIs would invest in an emerging market like India and in which types of shares/companies?

And

How the FIIs are related to market-related factors of Indian stock markets?

2nd Annual International Conference IEMCON 2012

Table 1: FIIs and Indian Stock Markets Information

| Year | FII Net Inflows (in Rs. Crores) | % Change in FII Net Inflows | SENSEX Value | % Change in SENSEX Value | S&P CNX NIFTY Value | % Change in S&P CNX NIFTY Value |
|-----------|------------------------------------|--------------------------------|--------------|-----------------------------|------------------------|------------------------------------|
| 2000-2001 | 9934.50 | | 3604.38 | | 1148.20 | |
| 2001-2002 | 8755.20 | (11.87) | 3469.35 | (3.75) | 1129.55 | (1.62) |
| 2002-2003 | 2527.00 | (71.14) | 3048.72 | (12.12) | 978.02 | (13.42) |
| 2003-2004 | 39960.00 | 1481.32 | 5590.60 | 83.38 | 1771.90 | 81.17 |
| 2004-2005 | 44122.90 | 10.42 | 6492.82 | 16.14 | 2036.00 | 14.90 |
| 2005-2006 | 48801.00 | 10.60 | 11279.96 | 73.73 | 3403.00 | 67.14 |
| 2006-2007 | 25236.00 | (48.29) | 13072.10 | 15.89 | 3822.00 | 12.31 |
| 2007-2008 | 53403.50 | 111.62 | 15644.44 | 19.68 | 4735.00 | 23.89 |
| 2008-2009 | (47706.20) | (189.33) | 9708.50 | (37.94) | 3021.00 | (36.20) |
| 2009-2010 | 110220.60 | 331.04 | 17527.77 | 80.54 | 5249.00 | 73.75 |
| 2010-2011 | 110121.00 | (0.09) | 19445.22 | 10.94 | 5833.80 | 11.14 |

Table 2: FIIs and MFs Role in Indian Stock Markets Crashes (2008-2011)

| Market Crashes Date | FIIs Flows | MFs Flows |
|---------------------|------------|-----------|
| 21/01/2008 | (2425.60) | 2001.70 |
| 24/10/2008 | (1178.00) | (318.70) |
| 17/03/2008 | (632.30) | (102.40) |
| 06/07/2009 | (351.30) | (586.70) |
| 22/01/2008 | (2256.30) | 1195.10 |
| 11/02/2008 | (1845.50) | (570.90) |
| 18/05/2006 | (810.60) | 762.67 |
| 10/10/2008 | (2323.20) | 315.20 |
| 13/03/2008 | (172.50) | (380.60) |
| 17/12/2007 | (1098.70) | (199.00) |
| 07/01/2009 | (1058.40) | (1011.20) |
| 31/03/2007 | - | - |
| 06/10/2008 | (1121.40) | (231.10) |
| 17/10/2007 | (1776.60) | 16.20 |
| 15/09/2008 | (629.30) | 130.90 |
| 22/09/2011 | (1234.60) | (41.00) |
| 18/01/2007 | 58.95 | 58.95 |
| 21/11/2007 | (2222.40) | (151.10) |
| 16/08/2007 | (2849.90) | 239.20 |
| 17/08/2009 | (973.80) | (681.80) |
| 27/06/2008 | (746.10) | 4.10 |
| 24/02/2011 | (2249.80) | 154.00 |
| 12/11/2010 | (683.20) | (183.20) |
| 16/11/2010 | (16.80) | (108.00) |
| 04/02/2011 | 224.40 | (21.20) |

There is statistical evidence that FIIs purchase or sell by taking leads through the movement of stock markets, i.e., FIIs are feedback traders. The statistical evidence also pointed out that when FIIs purchase/sell, there is an influence on the stock market. Consequently, either the stock market rise or fall on account of FIIs activities [Gupta (2011)] [as evident from Tables 1 and 2].

Mukherjee, Bose and Coondoo [2002] also suggested that FIIs inflows and outflows are caused by developments in Indian markets and not by those in markets abroad.

Ahmed, Ashraf and Ahmed [2005] revealed that the FIIs investments are influenced by the previous trading day returns, confirming the positive feedback strategy by them, but they are also influenced by the next trading day returns.

Prasanna [2008] observed that higher BSE SENSEX Index value and high price earnings (PE) ratio are the country level factors attracting more foreign investment in India.

Nidhi [2007] observed that the regulatory environment of the host country (i.e., India) has an important impact on FIIs inflows. As the pace of foreign investment began to accelerate, regulatory policies have changed to keep up with changed domestic scenarios. The above paper also provides a review of these changes.

The factors that influence individual share prices could either be internal factors, such as earnings, dividend, book value, etc. or external factors such as interest rate, government regulations, foreign exchange rate, etc. Several such factors have been identified by previous empirical research [Nirmala, Sanju and Ramachandran (2011)].

Among the financial performance variables the specific share/company returns and Earning per Share (EPS) are more influencing variables on FIIs investment decisions.

Prasanna [2008] also empirically observed that the foreign investment is more in companies with higher volume of publicly held shares. It has also been observed in the corporate governance literature (e.g., Aggarwal, Klapper and Wysocki [2005], etc.) that the FIIs choose the companies where family shareholding of promoters is not substantial. Thus, the

promoters' holdings and the FIIs investments are inversely related.

Many of the existing literature on FIIs in Indian stock markets found that the equity return has a significant and positive impact on the FIIs (e.g., Agarwal [1997]; Chakrabarti [2001]; Trivedi and Nair [2003]; etc.). Mukherjee, Bose and Coondoo [2002] found that the FIIs activities had a strong demonstration effect and was driving the domestic market suggesting that the FIIs flows tend to be caused by return in the domestic market.

Kumar [2001] investigated the effects of FIIs inflows on the Indian stock market represented by the SENSEX using monthly data from January 1993 to December 1997. He inferred that FIIs investments are more driven by Fundamentals and they do not respond to short-term changes or technical position of the market. The study concluded that SENSEX causes Net FII Investment (NFI). This finding is in contradiction with the findings of Rai and Bhanumurthy [2006] who did not find any causation from FIIs to return in BSE using similar data between 1994 and 2002. However, Rai and Bhanumurthy [2006] have also found significant impact of return in BSE on NFI. A study by Panda [2005] also showed that no clear causality is found between FIIs and S&P CNX Nifty.

Ahmed, Ashraf and Ahmed [2005] made a firm level analysis of FII's role in the Indian equity market. At the aggregate level FIIs investments and NSE Nifty seem to have a strong bi-directional causality. At the firm level FIIs are influencing equity returns especially in the government owned companies.

The study of Khan, Rohit, Goyal, Ranjan and Agrawal [2010] has attempted to analyze the relationship between equity returns and FIIs inflows on a new set of data. The period of the study has been taken from 1999 to 2009. Time period up to 2009 is taken to investigate the global crisis and its effect on FII investment in Indian stock Market. Also the analysis is based on the more broad-based National Stock Exchange Index, Nifty, composed of 50 stocks. The NSE has outstripped the BSE in terms of turnover, efficiency and transaction costs providing more liquidity and depth to trading. With strong preference of FIIs for holding shares of large firms and more liquid stocks, the NSE

Nifty appears a more reasonable index to work on than the BSE SENSEX. This paper empirically investigates the causal relationship as well as the degree of interdependency between Nifty and FIIs investment in Indian economy. They found that both Nifty and FII are not normally distributed. Nifty was found to be non-stationary and FII to be stationary at level itself. They also applied correlation test that indicates that Nifty is positively correlated to FII. Correlation between time series is higher in bear phase as compared to bull phase as in bull phase other market participants raise their involvement reducing the influence of FIIs. The correlation is further verified for the direction of influence by the Granger causality test between Nifty and FII. They found that Nifty Granger causes FII in all the four phases whereas reverse causality doesn't hold true. In order to find the short term causal relationship between the two time series, they next applied the Variance Decomposition and Impulse Response tests. They found out that the same result as in the Granger-Causality Test that Nifty causes FII in all the four phases.

Most of the studies generally point to the positive relationship between FIIs investments and movement of the NSE Index, and some also agree on bidirectional causality stating that foreign investors have the ability of playing like market makers given their volume of investments [Babu and Prabheesh (2008)].

But, given the huge volume of investments, foreign investors could play a role of market makers and book their profits, i.e., they can buy financial assets when the prices are declining, thereby jacking-up the asset prices and sell when the asset prices are increasing [Gordon and Gupta (2003)].

Rai and Bhanumurthy [2003] pointed out that the FII, given its short-term nature, might have bi-directional causation with the returns of other domestic financial markets like money market, stock market, foreign exchange market, etc. Hence, understanding the determinants of FII is very important for any emerging economy as it would have larger impact on the domestic financial markets in the short run and real impact in the long run.

Using monthly data from April 1993 to March 2004, Bandhani [2005] observed bi-directional long term causality between FIIs investment

flows and stock prices but no short run causality could be traced between the variables.

The result of the study of Bhattacharya and Mukherjee [2005] also suggested a bi-directional causality between stock price and the net foreign institutional investment, thus implying that the market informational efficiency hypothesis can be rejected for BSE Sensitive Index with respect to the FII.

The possibility of bi-directional relationship between FIIs and the equity returns was also explored by Rai and Bhanumurthy [2006]. They studied the determinants of foreign institutional investment in India during the period 1994-2002. They found, using monthly data that the equity returns is the main driving force for FIIs investment and is significant at all levels. They further studied the impact of news on FII flows and found that the FIIs react more (sell heavily) to bad news than to good news (as evident in Table 2). The study also revealed the positive association of FIIs investment with return on BSE SENSEX, inflation in US (home country) and negative association with inflation in India (host country), return on S&P 500 index, ex-ante risk on BSE and ex-ante risk on S&P 500 index. However, the ex-post risk neither in US nor in India affected FIIs inflows to India. The study also did not find any causation running from FIIs inflow to stock market returns. The study concluded that stabilizing the stock market volatility and minimizing the ex-ante risk would help in attracting more FIIs inflows. Otherwise there would be adverse impact of non-fundamental factors of FIIs behaviour which in turn would affect the real economy in the long-run.

Takeshi [2008] reported unidirectional causality from stock returns to FIIs flows irrelevant of the sample period in India whereas the reverse causality works only post 2003. However, impulse response function shows that the FII investments in India are more stock returns driven. Perhaps the high rates of growth in recent times coupled with an increasing trend in corporate profitability has imparted buoyancy to the stock markets, triggering off return chasing behavior by the FIIs.

According to the study of Kaur and Dhillon [2010], host country stock market returns (i.e., returns on SENSEX) have positive and significant impact whereas home country returns

(i.e., returns on S&P 500 Index) have negative but insignificant influence on FIIs investment inflows in long-run as well as in short-run. In terms of risk attached to returns on securities variability of SENSEX over variability of S&P 500 Index has negative and significant influence on FIIs inflows to India. Market capitalization and stock market turnover of India have positive and insignificant influence on FIIs investment in long-run but positive and significant influence on FIIs investment in short-run.

Rai and Bhanumurthy [2003] by using monthly data, we found that FII inflow depends on stock market returns, inflation rate (both domestic and foreign) and ex-ante risk. In terms of magnitude, the impact of stock market returns and the ex-ante risk turned out to be major determinants of FIIs inflow.

Trivedi and Nair [2003] also investigated the determinants of FIIs flows in India, and the causal relationship between FIIs investment inflows and the risk-returns in the Indian stock markets.

Prasuna [2000] also studied the determinants of FIIs investments in India using monthly data from January 1993 to March 1998. He found that lagged FII investment is significant at 1% level. Also, percentage change in BSE Sensex is also significant at 1%.

However, less attention was given to determine the relation of FIIs investments with other microeconomic variables, such as, market capitalization, total trading value and P/E Ratio of stock market, etc.

Kumar [2006] found a positive relation between FIIs investments with both market capitalization and total trading value. Tripathi [2007] found no significant relation between FIIs with market capitalization. However, the scope of these studies were limited as time period of all these studies was too short (generally 3 to 5 years' monthly data were taken) and the number of variables used for the analysis purpose was less.

Griffin, Nardari and Stulz [2004], found that foreign flows are significant predictors of returns for Korea, Taiwan, Thailand and India, indicating that foreign investors are buying before market index increases. He also found that contemporaneous flows are positive and

highly significant in India. FIIs and Stock Index show positive correlation, but fail to predict the future value.

Gordon and Gupta [2003] studied the relevance of micro and macro economic variables which are expected to affect foreign portfolio investment. They used monthly data on FII equity investments from 1993 to 2001. Prasuna [2000] found Exchange rate, interest rates, forward premium and foreign exchange reserves have been insignificant.

However, Bandhani [2005] found that exchange rate long term granger causes FIIs investment flow and not vice versa.

Nidhi [2007] examined the role of various factors relating to individual firm-level characteristics and macroeconomic-level conditions influencing FIIs investment.

Mukherjee, Bose and Coondoo [2002] tried to identify the major covariates of FIIs flows to India and Mukherjee and Coondoo [2004] examined the nature of volatility of such flows. Factors like the US and world stock price movements, volatility of return in different markets and domestic real and financial variables etc. have minor effects on FIIs flows to India [Mukherjee, Bose and Coondoo (2002)].

Kumar [2001] inferred that FIIs flows do not respond to short-term changes or technical position of the market and they are more driven by fundamentals.

Kaur and Dhillon [2010] observed that among macroeconomic determinants, economic growth of India (IIP) has significant and positive impact on FIIs investment inflows to India both in long-run and short-run. However, all other macroeconomic factors have significant influence only in long-run such as inflation in home country represented by US Producer Price Index (PPI) has significant and positive influence while inflation in India represented by Wholesale Price Index (WPI) has negative and significant influence on FIIs investment in India. This implies that foreign inflation leads to increase in FIIs investment inflows to Indian capital market. However, host country inflation i.e., inflation in India has an adverse impact on FIIs investment.

Using monthly data between May 1993 and Dec. 1999, Chakrabarti [2001] found that FIIs do not

have any informational disadvantage in comparison with domestic investors in India, since the US and world return are not significant in explaining FIIs flows. Besides, changes in country risk ratings for India do not appear to affect the FIIs flows.

A gap analysis of FIIs investment in Indian stock market, on closing data of S & P CNX NIFTY 500 of September 30, 2004, showed that the FIIs investments are highly concentrated in terms of their market value in a very small number of companies and there is a wide gap between the actual investments by FIIs and the investments allowed as per the cap [Sharma (2004)].

Pal [2005] showed that FIIs are the most dominant non-promoter shareholder in most of the SENSEX companies and they also control more tradable shares of SENSEX companies than any other investor groups.

FIIs by adopting a bottom-up approach seem to invest in top-quality, high growth, large cap stocks [Gordan and Gupta (2003)]. Douma, Rejje, and Kabir [2006] and Sytse, George and Rezaul [2003] provided empirical evidence that FIIs in India, invest in large, liquid companies which enable them to exit their positions quickly at relatively lower cost and also that the foreign institutional owners have a larger impact than foreign corporate owners when performance is measured using stock market valuation criterion.

Banaji [2000] emphasizes that the capital market reforms like improved market transparency, automation, dematerialization and regulations on reporting and disclosure standards were initiated because of the presence of the FIIs. But FIIs flows can be considered both as the cause and the effect of such capital market reforms. The market reforms were initiated because of the presence of FIIs and this in turn has lead to increased flows.

Bose and Coondoo [2005] has examined the impact of reforms of the FIIs' investment policy, on FIIs portfolio flows to the Indian stock markets, an aspect, studies on determinants of FIIs flows to India so far have not taken into consideration.

5.2) Nature Studies Review:

The question is whether the FIIs flows cause volatility, destabilization, noise trading, etc. and shows any types of herding behaviour?

Gordan and Gupta [2003] observed that although India receives hardly 1 percent of the FIIs investments in emerging markets, the portfolio flows to India have been less volatile when compared with that of many other emerging markets.

The study of Bansal and Pasricha [2009] had also pointed out that while return declined reasonably after the entry of FIIs, the volatility has been reduced significantly after their entry. Gordan and Gupta [2003] also found that FIIs investment had not resulted in increase in volatility in the Indian stock markets. Besides, FIIs investment flows, there may be other reasons as well that may have some degree of influence on market volatility and return. While the FIIs investment flows and contemporaneous SENSEX, NIFTY, market capitalization and market turnover have been strongly correlated in India, the correlation between FIIs investments and market volatility and market return has been comparatively low. It means volatility in Indian market is not the function of FIIs investment flows. There may be some other reasons which induced the volatility in Indian market over the time.

The direct relation between purchases and returns and the inverse relation between sales and returns, signal that FIIs in the Indian market may be herding. This is consistent with evidence provided by Batra [2003], where using both daily and monthly data, she finds that FIIs tend to herd in the Indian stock markets and the herding measure being high for the monthly horizon.

However, Lakshman, Basu and Vaidyanathan [2009] pointed out that FIIs Flows or normalized FIIs Flows does not significantly impact the herding behaviour; i.e., overall market-level herding is not impacted whether the FIIs Flows increase or decrease.

Batra [2003] also found that there is strong evidence that FIIs have been positive feedback investors and trend chasers at the aggregate level on a daily basis. However, there is no evidence of positive feedback trading on a monthly basis.

Noise trading on the part of FIIs would be less in Indian context since all their trades are delivery based only [Kumar (2005)].

Sivakumar [2003] had analysed the net flows of FIIs over the years, briefly analyses the nature of FIIs flows based on research, explores some determinants of FIIs flows and examines if the overall experience has been stabilising or destabilising for the Indian capital market. Ahmed, Ashraf and Ahmed [2005] confirmed that there has been little destabilizing effect of FIIs flows on individual equity return of the firm during their period of study.

A point also to be noted here is the heavy investment and selling attitude of FII's causing a major hurdle in stabilization of market sentiments [Bohra and Dutt (2011)].

Nidhi [2007] held that the increase in the volume of FIIs inflows in recent years had led to concerns regarding the volatility of these flows, threat of capital flight, its impact on the stock markets and influence of changes in regulatory regimes. The determinants and destinations of these flows and how are they influencing economic development in the country had also been debated.

In Indian context, Apte [2001] investigated the relationship between the volatility of the stock market and the nominal exchange rate of India by using the EGARCH specifications on the daily closing USD/INR exchange rate, BSE 30 (SENSEX) and NIFTY-50 over the period 1991 to 2000. The study suggested that there appears to be a spillover from the foreign exchange market to the stock market but not the reverse.

5.3) Events Studies Review:

The question is whether the FIIs flows are impacted by political, economic, natural, social or any other calamities or events?

Chakrabarti [2001] pointed out that the beta of the Indian market with respect to S&P 500 index seems to affect the FIIs flows inversely, but the effect disappeared in the post-Asian crisis period. There appear to be significant differences in the nature of FIIs flows before and after the Asian crisis. In the post-Asian crisis period i.e. from 1998 onwards, returns on the BSE National Index became the sole driving force behind the FIIs flows.

Batra [2003] observed that in times of pressure in the stock markets on account of a financial crisis in the region there is excessive sell side herding by the FIIs even though the extent of herding on the average and on either side of the market during a crisis may be lower than that in the immediately preceding period.

Gordan and Gupta [2003] observed that the nuclear tests and East Asian crisis did slow down the FIIs flows but as stated by, their effects were short lived.

A practical example in this regard is that Foreign Funds (including the FIIs) withdrew more than Rs.3,200 crores from the Indian stock markets in November 2011 only in mist of concerns over the poor debt crisis in the Eurozone. In the past few days, the stock markets have seen sharp volatility due to the rollback of funds by FIIs and also the depreciation of the rupee.

However, in many headlines it was also pointed out that whether it was the stock market crash of 2001, or the market collapse after India's nuclear test in 1998 or the meltdown after the Asian financial crises of 1997, FIIs were largely net buyers in India. For example, on Black Monday, May 17, 2004, when the stock markets recorded one of their biggest drop in a single day, FIIs were not the culprits. Net sales by FIIs were only about Rs.500 crores per day, a small proportion of the total trading volume of the NSE and BSE, spot and derivatives. It was also noted that while there was a net FIIs outflow of Rs.3,151.3 crores in May 2004, the FIIs were back recording a net investment of Rs.511 crores in June, when the markets were actually lower at about 4820 points.

5.4) Impact Studies Review:

The question is how much impact the FIIs have on Indices returns, individual shares, and overall Indian economic development till date?

Bose and Coondoo's [2005] results strongly suggested that liberalization policies have had the desired expansionary effect and have either increased the mean level of FIIs inflows and for the sensitivity of these flows to a change in BSE return or the inertia of these flows.

Lin and Chen [2006] concluded that the investment performance of FIIs high holding stocks is significantly better than that of FIIs low holding stocks. They presented the evidence that FIIs trading behaviour has generated better returns and portfolio performance since the stock market's full liberalization. This is also evident in stocks listed in SENSEX and S&P CNX NIFTY, e.g., HDFC Bank, ICICI Bank, Infosys, Hero Motocorp Ltd., M&M Ltd., etc. in which FIIs holding is significant.

Morgan Stanley [2002] examined that FIIs have played a very important role in building up India's Forex Reserves, which have enabled a host of economic reforms. FIIs are now important investors in the country's economic growth despite sluggish domestic sentiment. The Morgan Stanley [2002] report also noted that FIIs strongly influence short-term market movements during bear markets. However, the correlation between returns and flows reduces during bull markets as other market participants raise their involvement reducing the influence of FIIs. The results show that the correlation between foreign inflows and market returns is high during bear and weakens with strengthening equity prices due to increased participation by other players.

The study of Bhattacharya and Mukherjee [2005] and Sandhya, Krishnamurti and Sen [2003] found strong evidence consistent with the base-broadening hypothesis consistent with prior work. Sandhya, Krishnamurti and Sen [2003] also did not find compelling confirmation regarding momentum or contrarian strategies being employed by FIIs. Their findings support the price pressure hypothesis. They did not find any substantiation to the claim that foreigners' destabilize the market.

On investigating the impact of trading imbalance across days, Batra [2003] did not find any significant evidence that would make it possible to attribute equity market instability to FIIs. On the issue of market stability, Mazumdar [2004] found that FIIs flows have enhanced liquidity in the Indian stock market.

Panda [2005] found that the returns on Indian stock market indices such as BSE SENSEX and S&P CNX NIFTY are relatively more affected by MFs investments than FIIs investments. In fact, FIIs investments do not affect BSE SENSEX rather it is affected by the latter.

The study of Rajput and Thaker [2008] find that the FIIs significantly affect the S&P CNX NIFTY performance and is one of the important driving forces of the Indian stock market.

5.5) Extent Studies Review:

The question is how much and how far the FIIs investment has its arms-length?

Banaji [2000] pointed out that it is not the market capitalization of Indian stock markets that matters but what is important is the level of the free float, that is, the shares that are actually publicly available for trading.

The regression results of the study of Kumar [2005] showed that the combined might of the FIIs and MFs are a potent force, and their direction in fact can forecast market direction. The Granger causality test under this study however showed that the MFs in fact lead the market rise or fall and FIIs follow suit. This may actually raise questions on the market efficiency but on the contrary, markets become more efficient with the growing presence of FIIs who predominantly go by fundamentals.

The study of Dasgupta and Dutta [2011] pointed out that total trading value of NSE Cash Segment which is an evidence of market liquidity and size has been significantly influenced by MF total trading value, FII total trading value and S&P CNX Nifty Index returns in that order.

6) Issues for Future Researchers:

Some of the findings and issues from the review of available conceptual and empirical studies on FIIs in relation to Indian stock markets are as follows:

- 1) The first issue for further studies and debate is whether FIIs cause BSE SENSEX and S&P CNX NIFTY returns or whether the other way round. A broader stock market perspective (e.g., BSE 200 or BSE 500 returns in terms of FIIs investment) has also to be considered to find the real impact of FIIs on the overall market activities and development.
- 2) Share/company specific studies are very limited in Indian context in relation to FIIs investment. The internal and external factors that prompt the FIIs to invest in a specific share/company at some specific times have to be analysed in details.

- 3)** Although many such studies have pointed out that the FIIs are mostly long-term in their attitude and time-horizon for investment, but, a strong contradictory viewpoint of herding and withdrawal of funds at the slightest possible hint of a turnaround of stock markets is also evident.
- 4)** Another very important issue to analyse for the FIIs investing in Indian stock markets has been the coupling effect of world stock markets with their Indian counterparts. Many macroeconomic determinants, such as, inflation, GDP growth rate, exchange rate, interest rate regimes, etc. of India and some other influential countries of the world in relation to FIIs investment in India need to be intervened and studied empirically.
- 5)** The role of FIIs in driving the volume in Cash and Derivatives Markets of India needs special attention from researchers. Many other microeconomic variables, such as, market capitalization, total trading value and P/E Ratio of Indian stock markets have also been given less importance in FIIs studies.
- 6)** In the modern competitive investment environment through out the world, ‘downgrade/upgrade a country’ impact on FIIs have to be intensely studied. The impact of free-float methodology on the FIIs in respect to individual shares/companies need to be analysed also.
- 7)** Many studies have highlighted that FIIs do not cause any types of market volatility or destabilisation. But, facts and figures (see Table 1 and 2 also) in many cases point out to the opposite viewpoint. Thus, this needs to be proved with advanced statistical techniques.
- 8)** The impact of financial and economic crises on FIIs investment in Indian stock markets has been studied in a few cases. However, the impact of political, social and natural events and calamities on FIIs and their immediate reactions to those have hardly been studied.
- 9)** It is important and relevant to assess the impact of FIIs investment in a specific share/company in terms of its financial performance, dividend policy, stock split decisions, issue of bonus shares, etc. In this regard, the nature of investment strategies adopted by the FIIs overall and share/company specific need to be addressed.
- 10)** FIIs role in the overall economic development of India in building up foreign exchange reserves, eliminating current account deficits and BOPs problems, curbing inflation and favourable exchange rate situations have also yet to be studied.

Future studies on FIIs and Indian stock markets should look into all the above issues. In turn, the regulatory authority, stock exchanges and the Indian Government above all could develop a more authentic yet lucrative policy framework for the FIIs, so that their impact on Indian stock markets would be long-term and solid.

7) Conclusion:

Given that India is one of the fastest growing economies in South Asia, promising a growth of over 8 percent, second only to China, it would not be a surprise to see increased FII flows to India in the future. FIIs are now looking at the economy as a whole, with the macro-economic factors also playing their role in attracting foreign investors. Factors like a strong currency, key reforms in the banking, power and telecommunications and infrastructure sectors, increased consumer spending and stable monetary and fiscal policies are expected to play a major role in attracting FIIs to India in future.

The step of the Securities Exchange Board of India (SEBI) along with the Institute of Chartered Accountants of India (ICAI) to jointly monitor the markets and announces the regulatory measures has been making the Indian companies more transparent and more disciplined.

Last, but not the least, we hope this study will provide a good ground work to carry out more vigorous analysis in this field with more effective statistical tool and with latest data of present period.

References:

- [1] Gordon, J. and Gupta, P., 2003, “Portfolio Flows into India: Do Domestic Fundamental Matter?”, IMF Working Paper, No. 03/20.
- [2] Ananthanarayanan, S., Krishnamurthi, C. and Sen, N., 2004, “Foreign Institutional Investors and Security Returns: Evidence from Indian Stock Exchanges”, Working Paper, ISB (Indian School of Business), Hyderabad.
- [3] Lalitha, S., 1992, “Failure to woo Funds forced move on FIIs”, Observer of Business and Politics, September 16.
- [4] Prasanna, P.K., 2008, “Foreign Institutional Investors: Investment Preferences in India”, JOAAG, Vol. 3., No. 2, P 40-51.

- [5] NSE, 2001, "Indian Securities Market: A Review", Volume IV, National Stock Exchange, Mumbai.
- [6] Gupta, A., 2011, "Does the Stock Market Rise or Fall Due to FIIs in India?", Journal of Arts, Science & Commerce, Vol. 2, No. 2, P 99-107.
- [7] Mukherjee, P., Bose, S. and Coondoo, D., April-September, 2002, "Foreign Institutional Investment in the Indian Equity Market: An Analysis of Daily Flows during January 1999 to May 2002", Money and Finance, ICRA Bulletin, P 21-51.
- [8] Ahmed, K.M, Ashraf S. and A. Ahmed, 2005, "An Empirical Investigation of FII's role in the Indian equity Market", ICFAI Journal of Applied Finance, Vol. 11, No. 8, P 21-33.
- [9] Nidhi, D., 2007, "Foreign Institutional Investment in India", Journal: Research on Indian Stock Volatility, Vol. 14., Emerald Group Publishing Limited.
- [10] Nirmala, P.S., Sanju, P.S. and Ramachandran, M., 2011, "Determinants of Share Prices in India", Journal of Emerging Trends in Economics and Management Sciences, Vol. 2, No. 2, P 124-130.
- [11] Aggarwal, R., Klapper, L. and Wysocki, P. D., 2005, "Portfolio preferences of foreign institutional investors", Journal of Banking and Finance, Vol. 29, No. 12, P 2919-2946.
- [12] Agarwal, R.N., 1997, "Foreign Portfolio Investment in Some Developing Countries: A Study of Determinants and Macroeconomic Impact," Indian Economic Review, Vol. XXXII, No. 2, P 217-229.
- [13] Chakrabarti, R., 2001, "FII Flows to India: Nature and Causes", Money and Finance, Vol. 2, No. 7, P 61-81.
- [14] Trivedi, P. and Nair, A., January 30-February 1, 2003, "Determinants of FII Investment Inflow to India", Presented in 5th Annual conference on Money & Finance in The Indian Economy, Indira Gandhi Institute of Development Research.
- [15] Kumar, S., 2001, "Does the Indian Stock Market Play to the tune of FII Investments? An Empirical Investigation", ICFAI Journal of Applied Finance, Vol. 7, No. 3, P 36-44.
- [16] Rai, K. and N. R. Bhanumurthy, 2006, "Determinants of Foreign Institutional Investment in India: The Role of Risk Return and Inflation", http://iegindia.org/dis_rai_71.pdf. JEL Classification: E44, G15, G 11.
- [17] Panda, C., January, 2005, "An Empirical Analysis of the Impact of FII' Investment on Indian Stock Market", Applied Finance, P 53-61.
- [18] Khan, M.A., Rohit, Goyal, S., Ranjan, V. and Agrawal, G., 2010, "Investigation of Causality between FIIs' Investment and Stock Market Returns", International Research Journal of Finance and Economics, Issue 40, P 100-112.
- [19] Babu M. Suresh and Prabhesh K.P., 2008, "Causal Relationships between Foreign Institutional Investments and Stock Returns in India", International Journal Trade and Global Markets, Vol. 1, No. 3.
- [20] Rai, K. and N.R. Bhanumurthy, 2003, "Determinants of Foreign Institutional Investments in India – The Role of Returns Risk and Inflation", JEL Classification: E44, G15, G11.
- [21] Badhani, K. N., December 17, 2005, "Dynamic Relationship among Stock-Prices, Exchange Rate and Net F.I.I. Investment Flow in India", Paper presented at The Conference on Research in Finance and Accounting Indian Institute of Management, Lucknow.
- [22] Bhattacharya, B. and Mukherjee, J., 2005, "An Analysis of Stock Market Efficiency in the Light of Capital Inflows and Exchange Rate Movements: The Indian Context", P 1-26.
- [23] Takeshi, I., 2008, "The causal relationships in mean and variance between stock returns and Foreign institutional investment in India", IDE discussion paper No.180.
- [24] Kaur, M. and Dhillon, S.S., 2010, "Determinants of Foreign Institutional Investors' Investment in India", Eurasian Journal of Business and Economics, Vol. 3, No. 6, P 57-70.
- [25] Prasuna, C.A., 2000, "Determinants of Foreign Institutional Investment in India", Finance India, Vol. XIV, No. 2, P 411-421.
- [26] Kumar, S., February, 2006, "FII v/s Sensex: An Emerging Paradigm", Treasury Management, P 43-49.
- [27] Tripathi, N.P., August, 2007, "Dynamic Relationship between Stock Market, Market Capitalization and Net FII Investment in India", ICFAI Journal of Applied Finance, P 60-68.
- [28] Griffin, J.M., Nardari, F. and Stulz, R.M., 2004, "Are daily cross-border equity flows pushed or pulled", The Review of Economics and Statistics, Vol. 86, P 641–657.
- [29] Mukherjee, P. and Coondoo, D., 2004, "Volatility of FII in India", Money & Finance, Vol. 2, Nos. 15–16.
- [30] Sharma, L., 2004, "A Gap Analysis of FIIs Investments – An estimation of FIIs Investments Avenues in Indian equity market".
- [31] Pal, P., 2005, "Recent Volatility in Stock Market in India and Foreign Institutional Investors", Economic and Political Weekly.

- [32] Douma, S., Rejie, G. and Kabir, R., 2006, "Foreign and domestic ownership, business groups and firm performance-Evidence from large emerging market", Strategic Management Journal, Vol. 27, No. 7, P 637-657.
- [33] Sytse D., George R. and Rezaul K., 2003, "Foreign and Domestic Ownership, Business Groups and Firm Performance: Evidence from a Large Emerging Market", Tilburg University, Netherlands.
- [34] Banaji, J., 2000, "Investor capitalism and the reshaping of business in India", Queen Elizabeth House (QEH), Oxford University, Working paper No. 54.
- [35] Bose, S. and Coondoo, D., 2005, "The Impact of FII Regulations in India", International Journal of financial market trends, Vol. 30, MCB UP Ltd.
- [36] Bansal, A. and Pasricha, J.S., 2009, "Foreign Institutional Investor's Impact on Stock Prices in India", Journal of Academic Research in Economics, Vol. 1, No. 2, P 181-189.
- [37] Batra, A., 2003, "The Dynamic of Foreign Portfolio Inflows and Equity Returns in India", Working Paper, ICRIER, New Delhi.
- [38] Lakshman, M.V., Basu, S. and Vaidyanathan, R., 2009, "Market-Wide Herding and the Impact of Institutional Investors in the Indian Capital Market", IIM (B) Report, P 1-41.
- [39] Kumar, S.S.S., 2005, "Role of Institutional Investors in Indian Stock Market", International Journal of Management Practices & Contemporary Thoughts, P 76-80.
- [40] Sivakumar, S., 2003, "FIIs: Bane or boon?", Journal of stock market volatility, Vol. 34, MCB UP Ltd.
- [41] Bohra, N.S. and Dutt, A., 2011, "Foreign Institutional Investment in Indian Capital Market: A Study of Last One Decade", International Research Journal of Finance and Economics, Issue 68, P 103-116.
- [42] Apte, P.G., 2001, "Currency Exposure and Stock Prices", Journal of Foreign Exchange and International Finance, Vol. XII, No. 2, P 135-143.
- [43] Lin, A. and Chen, C.Y., 2006, "The impact of qualified FII on Taiwan's Stock Market", Web Journal of Chinese Management Review, Vol.9, No.2, P 1-27.
- [44] Morgan Stanley, 2002, "FII's influence on Stock Market", Journal of impact of Institutional Investors on ism. Vol. 17, Emerald Group Publishing Limited.
- [45] Sandhya, A., Krishnamurti, C. and Sen, N., 2003, "Foreign Institutional Investors and Security Returns: Evidence from Indian Stock Exchanges", P 2-31.
- [46] Mazumdar, T., 2004, "FII Inflows to India; Their effect on stock market liquidity", ICFAI Journal of Applied Finance, Vol. 10, No. 7, P 5-20.
- [47] Rajput, A. and Thaker, K., 2008, "Exchange Rate, FII and Stock Index Relationship in India", Vilakshan, XIMB Journal of Management, P 43-56.
- [48] Dasgupta, R. and Dutta, S., 2011, "The Sustainable Dynamism of National Stock Exchange- The Influencing Factors- An Empirical Study", Paper presented at the 1st International Conference (SPSITM 2011) organized by IEM in collaboration with IEEE, P 1-10.

2nd Annual International Conference IEMCON 2012

----- REVIEW 1 -----

PAPER: 89

TITLE: FIIs and Indian Stock Markets –
Emerging Issues From A Literature Review
Post-2000

AUTHORS: Ranjan Dasgupta and Sujit Dutta

Question:

1 Relevance to the conference:

Very Relevant

2 Originality:

Somewhat original

3 Detailed Comments:

- Good research work

4 Overall Rating:

5/5

5 Comments to Program Chair:

SIGNAL PROCESSING & MOBILE COMMUNICATIONS

Date : 18th January,2012

Chairman : Prof. P.K.Banerjee

Co-Chairman : Prof. Debika Bhattacharyya

Rapporter : Prof. Nilanjana Dutta Roy

| Paper ID | Paper Title | Authors |
|----------|--|---|
| 65 | Study of Different Attacks in Manet with its Detection and Mitigation Schemes | Aniruddha Bhattacharyya, Arnab Banerjee and Dipayan Bose |
| 66 | An Inverse Cell Breathing Based Power Management for Congested Mobile Network | Anwesha Mukherjee,Debashis De |
| 93 | An Efficient Approach for Designing High Performance Squaring Circuit using Vedic Algorithms | Jubin Hazra, Saptarshi Mallick, Joydeep Saha and Biswajit Das |
| 94 | Voronoi Coverage in Swarm Robot Position Control: A New Approach | Paramita Mandal, Rangit Kumar Barai and Madhubanti Maitra |
| 97 | Modeling of Multirate 3-Stage Cascaded-Integrator-Comb Filter for GSM Networks and its FPGA Implementation | Saikat Dutta, Moloy Kumar Chowdhury |

STUDY OF DIFFERENT ATTACKS IN MANET WITH ITS DETECTION AND MITIGATION SCHEMES

Aniruddha Bhattacharyya

Department Of Computer Science & Engineering
Institute Of Engineering & Management, Salt Lake
Kolkata, India
aniruddha.aot@gmail.com

Arnab Banerjee

Department Of Computer Science & Engineering
Institute Of Engineering & Management, Salt Lake
Kolkata, India
arnab.saheb@gmail.com

Dipayan Bose

Department Of Computer Science & Engineering
Institute Of Engineering & Management, Salt Lake
Kolkata, India
connect2dipayan@gmail.com

Abstract— In mobile ad-hoc network (MANET) security is a challenging issue due to its open nature, infrastructure less property and mobility of nodes. In designing a new security mechanism for mobile ad hoc networks, one must consider the attacks variations as well as the characteristics of the attacks that could be launched against the ad hoc networks and existing detection and mitigation schemes.. The discussions of these four aspects are summarized in this paper. This paper also classifies several common attacks against the ad hoc networks routing protocols based upon the techniques that could be used by attackers to exploit routing messages.

Keywords- *adhoc, black hole attack, gray hole attack, fabrication, manet, malicious node, routing attacks*

I. INTRODUCTION

A mobile ad hoc network (MANET) consists of a set of mobile hosts that carry out basic networking functions like packet forwarding, routing, and service discovery without the help of an established infrastructure. Security in MANET is an essential component for basic network functions like packet forwarding and routing. Network operation can be easily jeopardized if countermeasures are not embedded into basic network functions at the early stages of their design. Unlike networks using dedicated nodes to support basic network functions, ad hoc networks carry out those functions by all the

available nodes. This very difference is at the core of the security problems that are specific to ad hoc networks. As opposed to dedicated nodes of a classical network, the nodes of an ad hoc network cannot be trusted for the correct execution of critical network functions. However, similar to other networks, MANET also vulnerable to many security attacks. MANET not only inherits all the security threats faced in both wired and wireless networks, but also introduces security attacks unique to itself [1]. As people will be encouraged to use a secured network, it is important to provide MANET with reliable security mechanisms if we want to see this exciting technology become widely used in a next few years. Before the development of any security measure to secure mobile ad hoc networks, it is important to study the variety of attacks that might be related to such networks. With the knowledge of some common attack issues, researchers might have a better understanding of how mobile ad hoc networks could be threatened by the attackers, and thus might lead to the development of more reliable security measures in protecting them. The purpose of this study is to investigate some of the important issues that might be related to security attacks in mobile ad hoc networks and some of the existing detection and mitigation schemes.

We will then see how attacks against the ad hoc networks may vary depending upon the environment the attacks are

launched in, the communication layer the attacks are targeting, and the level of ad hoc network mechanisms targeted. After considering these three variations, it is also important to investigate the characteristics of attacks against the ad hoc networks. In this paper, we give a special attention to attacks that could be launched against the routing protocols[2]. We identified most of the attacks against ad hoc networks' routing protocols and further classified them based upon the techniques that could be used to exploit routing messages.

Layer Security issues:

| LAYERS | ATTACKS |
|--------------------------|---|
| Application Layer | Repudiation, Data corruption |
| Transport Layer | Session Hijacking, SYN Flooding |
| Network Layer | Gray hole, Black hole, Worm hole, Byzantine, Sybil, Jellyfish, Rushing. |
| Link Layer | Interception, Fabrication, Modification. |
| Physical Layer | Jamming, Sniffing. |

II. ATTACK VARIATIONS

A. Mobile vs Wired Attackers :

Mobile attackers have the same capabilities as that of the other nodes of any particular ad hoc network. Having the same resource limitations, their capabilities to harm the networks operations gets also limited. For instance, with the limited transmitting capabilities and battery power, mobile attackers could only jam the wireless links within its vicinity. They are not capable to launch the network jamming attacks to disrupt the complete network's operations. On the other hand, wired attackers are attackers that are capable of gaining access to the external resources such as electricity. As they have more resources, they can launch more severe attacks in the network, such as jamming the whole network or breaking expensive cryptography algorithms. Existence of the wired attackers in ad hoc networks is always possible as long as the wired attackers are able to locate themselves in the communication range and have access to the wired infrastructures.

B. Passive vs Active Attacks :

Classes of attack might include passive monitoring of communications, active network attacks, close-in attacks, exploitation by insiders, and attacks through the service provider. Information systems and networks offer attractive targets and should be resistant to attack from the full range of threat agents, from hackers to nation-states. A system must be able to limit damage and recover rapidly when attacks occur.

Attacks in ad hoc networks can be broadly classified into two groups:

- Passive attacks: typically involve only eavesdropping of data. Passive attacks result in the disclosure of information or data files to an attacker without the consent or knowledge of the user. Passive interception of network operations enables adversaries to see upcoming actions.
- Active attacks: involve actions performed by attacker, such as replication, modification and deletion of exchanged data. Active attacks include attempts to circumvent or break protection features, to introduce malicious code, and to steal or modify information [3]

III. ACTIVE ATTACK AGAINST ROUTING SCHEMES

Routing is one of the most vital mechanisms in the ad hoc networks. Improper and insecure routing mechanisms will not only degrade the performance of the ad hoc networks, but also will render such networks vulnerable to many security attacks. One of the basic elements in the routing mechanism is the routing message, which is used to establish and maintain relationships between nodes in the networks. The importance of the routing message has made it a main target of the attackers to launch attacks against the ad hoc networks [4][5]. Attacks against the routing messages could be launched in many forms and may include all the characteristics described below. In this work, attacks against routing messages are classified based on the classification suggested by Stallings in [6]. In such classification, information or messages could be deviated from the normal operation flow using modification, interception, interruption or fabrication attacks. In a more severe case, attackers also might use any combination of these attacks to disrupt the normal information flow.

A. Modification

In a message modification attack, adversaries make some changes to the routing messages thus endangering the integrity of the packets in the networks. Since nodes in the ad hoc networks are free to move and self-organize, relationships among nodes might sometimes include the malicious nodes which might exploit the sporadic relationships in the network to participate in the packet forwarding process, and later launch the message modification attacks. Examples of such attacks are packet misrouting and impersonation attacks.

- Packet misrouting attacks: In a packet misrouting attack, malicious nodes reroute traffic from their original path to make them reach the wrong destinations [7]. Attackers might misroute a packet to make it stay in the network longer than its lifetime thus rendering it to be dropped from the network resulting in the source node needing to retransmit the lost packets resulting in consumption of more

bandwidth, as well as increase of the overhead in the network.

- Impersonation attacks: Also called spoofing attacks. In this type of attack malicious node assumes the identity of another node in the network [8]. By impersonating another node, attackers are able to receive routing messages that are directed to the nodes they faked.

B. Interception

Attackers might launch the interception attacks to get an unauthorized access to the routing messages that are not intentionally sent to them. The intercepted packets might be modified jeopardizing the integrity of the packets or might be analyzed violating the confidentiality before being forwarded to the next hop.

- Wormhole attacks: [9] In wormhole attacks, a compromised node in the ad hoc network colludes with external attacker to create a shortcut in the networks to trick the source node to win in the route discovery process and later launch the interception attacks. Packets from these two colluding attackers are usually transmitted using wired connection to create the fastest route from source to the destination node [10][11]. In addition, if the wormhole nodes consistently maintain the bogus routes, they could permanently deny other routes from being established.
- Black hole attacks: In this attack, malicious nodes trick all their neighboring nodes to attract all the routing packets to them. As in the wormhole attacks, malicious nodes could launch the black hole attacks by advertising themselves to the neighboring nodes as having the most optimal route to the requested destinations[12][13][14]. However in the black hole attacks, unlike wormhole attacks, only one attacker is involved and it threatens all its neighboring nodes.
- Gray hole attack: is a variation of the black hole attack, where the malicious node is not initially malicious but turns into malicious sometimes later. This anomalous behavior of malicious nodes prevents a trust based security solution from detecting them before they turn into malicious node. A gray hole may forward all the packets to certain nodes but may drop those packets coming from or destined to some specific nodes. In another variation of this attack, a node may behave maliciously for some time but later on it behaves normally. Due to this uncertainty in behavior of gray hole, this type of attacks are more difficult to detect/prevent compared to black hole attack. Like black holes, cooperative gray hole attacks may be possible against AODV [15][16].
- Routing packet analysis attacks: Since no disruptive action occurs, routing packet analysis could be classified as one of the passive attacks against the ad hoc networks. One way to launch this attack is by exploiting the promiscuous mode employed in the ad

hoc network. In a promiscuous mode, if node A is the neighbor of both nodes B and C at a particular time, node A can always hear the transmissions between node B and node C. By exploiting this nature, node A is able to analyze the overheard packets transmitted between node B and node C. Besides, malicious nodes could also launch this attack by exploiting the nature in a multi hop routing.

C. Fabrication

Malicious nodes fabricate their own packets to cause chaos in the network operations. They could launch the message fabrication attacks by injecting huge packets into the networks such as in the sleep deprivation attacks. Such attacks also might come from the internal misbehaving nodes such as in route salvaging attacks.

- Sleep deprivation attacks [17]: This kind of attack is actually more specific to the MANET. The aim is to drain off limited resources (e.g. battery powers) in the mobile ad hoc nodes, by constantly making them busy to process unnecessary packets. In a routing protocol, sleep deprivation attacks might be launched by flooding the targeted node with unnecessary routing packets.
- Route salvaging attacks: Route salvaging attacks are launched by the greedy internal nodes in the network. In MANET, there is no guarantee that each transmitted packet will successfully reach the desired destination node [18]. It may happen because of the natural network failures or might be due to attacks by the adversaries. Therefore, to salvage their packets from such failures, misbehaving internal nodes might duplicate and retransmit their packets though not sending error messages received. The effects of the route salvaging attacks might be more severe if there are many greedy nodes in the networks. This attack might also cause the consumption of unnecessary bandwidth.

D. Interruption

Interruption attacks are launched to deny routing messages from reaching the destination nodes. Adversaries could do this by either attacking the routing messages or the mobile nodes in the network. For instance, adversaries aiming to interrupt the availability service in the networks might destroy all paths to a particular victim node by using the message modification attacks. In a message fabrication attack, adversaries could overload the networks by injecting huge unnecessary packets. Examples of such attacks are packet dropping attacks, flooding attacks, and lack of cooperation attacks.

- Packet dropping attacks: Direct interruption to the routing messages could be done by using the packet dropping attacks. In a standard packet dropping attack, an adversary collaborates in the route discovery

process and launches constant packet dropping attacks if it is included as one of the intermediate nodes. In addition adversaries might vary their techniques using random, selective, or periodic packet dropping attacks to help their interrupting behavior remain concealed [19][20].

- Flooding attacks: Adversaries also might interrupt the normal operations in the packet forwarding process by flooding the targeted destination nodes with huge unnecessary packets. Nodes under the flooding attacks are unable to receive or forward any packet thus all the packets directed to them will be discarded from network [21][22].
- Lack of cooperation attacks: Lack of cooperation from the internal nodes to participate in the network operations can also be seen as an attempt to launch a refusal of service attack. In such attacks, internal nodes are discouraged to cooperate in the network operations that do not benefit them as it drains off their resources. Misbehaving internal nodes, to save their limited resources, might refuse to forward the other node's packets, or not sending back the route error report to the sender when failing to forward packets, or turning off their devices when not sending any packet in the network.

IV. DETECTION AND MITIGATION SCHEMES

A. Intrusion Detection

The mobile nodes are independent and their movement is not controlled by the system. So the nodes can easily be captured & compromised. The architecture of wireless network also has no physical obstacles, so attacks can come from any directions. Adversaries can exploit the decentralized management, by breaking the cooperative algorithm. To tackle this situation different IDS are available, such as like:

- Standalone IDS: Every node has its own IDS agent, which monitor the node only, if any threat has been detected then it can take protection locally & since there is no cooperation in between the nodes then all decisions are based on information collected by nodes. This approach is not so efficient.
- Cooperative IDS: Wireless ad hoc network is distributed in nature so the detection & response mechanism must be in two phase. Local & Global detection. Every node has IDS agent that detect attacks locally and cooperates with other nodes inside network and inform globally. Significantly, this distributed-cooperative IDS technology must be more stable than standalone IDS and more stable form flat, cluster based network configuration.[23,24]
- Hierarchical IDS: Here also every node has its own IDS agent. Collection of nodes form cluster. Every cluster has a special node known as cluster head.IDS

agent for cluster head in responsible for both local & global Intrusion detection. Layered architecture in wireless ad hoc network can be protected by this approach. Another alternative distributed solution called spontaneous watchdog. Fast sensor ad hoc network without divide them into clusters, some powerful independent, spontaneous nodes are created, known as watchdog, which monitors the communication with their neighbors[25][26].

- Zone based IDS: The local IDS agent used in Zone based Intrusion detection system (ZBIDS).Zone can be formed by geographic partitions. Every node has its two identities, INTRA ZONE and INTERZONE & can be determined by Zone ID. By forwarding HELLO Message, inter and intra zone nodes are determined. A node may change its role by the nature of mobility.[27][28][29]

B. Mitigation Technique

Mitigation technique in ad hoc network guarantees to protect from the attacks, security threats and vulnerabilities, like The Multipath Routing can be effective way to mitigate selective forwarding. Different mitigation techniques for the different types of attacks are:

- Black Hole Attack [26]: (i) Collecting multiple RREP messages (from more than two nodes) and thus hoping multiple redundant paths to the destination node and then buffering the packets until a safe route is found. (ii) Maintaining a table in each node with previous sequence number in increasing order. Each node before forwarding packets increases the sequence number. The sender node broadcasts RREQ to its neighbors and once this RREQ reaches the destination, it replies with a RREP with last packet sequence number. If the intermediate node finds that RREP contains a wrong sequence number, it understands that somewhere something went wrong.
- Gray Hole Attack: Mitigated by priority protocols schemes [31]. Whenever a node enters in a MANET, IP allocation is the first step in which the node will get its IP along with initial priority and the technique of Prime DHCP is adopted [23]. Neighbor Discovery is the second step of the proposed scheme. New node will send the HELLO packets to its neighbors and discover the identity of the neighbors along with their priority. Authentication is the next step of the scheme in which it will broadcast information about its existence and exchange keys with the neighbors according to the scheme of hop-by-hop authentication protocol, HEAP [24], by using a modified HMAC-based algorithm along with two keys and drops any packets that originate from outsides.

- Jellyfish Attack: (i) 2ACK [34]: The basic idea of the 2ACK scheme is that, when a node forwards a data packet successfully over the next hop, the destination node of the next-hop link will send back a special two-hop acknowledgment called 2ACK to indicate that the data packet has been received successfully. Such a 2ACK transmission takes place only for a fraction of the data packets. (ii) Credit based systems [30]: This approach provides incentives for successful transmission of some kind of token or credit which the node might use when it starts sending its own packet. (iii) Reputation based scheme: Here individual nodes collectively detect misbehaving nodes (such as CONFIDANT).[43][44]
- Worm Hole Attack [45] [46]: (i) Geographical leashes & temporal leashes: A leash is added to each packet in order to restrict the distance the packets are allowed to travel. A leash is associated with each hop. Thus, each transmission of a packet requires a new leash. A geographical leash is intended to limit the distance between the transmitter and the receiver of a packet. A temporal leash provides an upper bound on the lifetime of a packet [47]. (ii) Using directional antenna: Using directional antenna restricts the direction of signal propagation through air. This is one of the crude ways of limiting packet dispersion.
- Rushing Attack: (i) SEDYMO [36]: Secured Dynamic MANET On-Demand is similar to DYMO but it dictates that intermediate node must add routing information while broadcasting the routing messages and no intermediate node should delete any routing information from previous sender while broadcasting. It also incorporates hash chains and digital signature to protect the identity. (ii) SRDP [35]: Secure Route Discovery Protocol is security enhanced Dynamic Source routing (DSR) protocol. (iii) SND [29]: Secure Neighbor Detection is another method of verifying each neighbor's identity within a maximum transmission range.
- Cache Poisoning Attack: (i) SAODV [38]: Secure AODV is an extension to AODV protocol that adds each node to exchange signed routing messages. Each node has its own public key which it uses to sign routing messages. Also SAODV uses hop count as a metric for shortest-route as AODV and uses hash chains to secure hop count information in route messages. (ii) ARAN [39]: Authenticated Routing protocol for Ad-hoc Networks uses similar techniques as SAODV. ARAN uses certificates issued by a third party certification authority. (iii) SNRP [40]: Secure Neighbor Routing protocol uses security enhanced Neighbor Lookup Protocol (NLP) to secure MANET routing. Newly added node uses public key to participate in MANET
- Sybil Attack: One way of mitigating this attack is maintaining a chain of trust, so single identity is generated by a hierarchical structure which may be hard to fake. Another approach would be based on signal strength [41][42].

V. CONCLUSION

We have tried to categorize the different types of ad hoc security attacks solely based on their characteristics to considerably reduce the mitigation period. By bringing the attacks under these two broad categories the complicity of naming also reduces. We have also kept a close look on the existing algorithms needed to mitigate the attacks and have tried to bind the attacks into categories according to that. Some attacks have characteristics which makes them unsuitable to be categorized into these categories, so they have been kept away from this topic of discussion for the time being.

Further study is in progress to find out more common characteristics of the attacks to more strongly bind them into these categories and to ably design more powerful algorithm in mitigating DATA and CONTROL traffic attacks.

REFERENCES

- [1] Mohammad Al-Shurman ,Seong-Moo Yoo , Black Hole Attack in Mobile Ad Hoc Networks , ACM-SE 42 Proceedings of the 42nd annual Southeast regional conference, pp. 96-97, 2004
- [2] Songbai Lu,Lingyan Jia,Kwok-Yan Lam,Longxuan Li, SAODV: A MANET Routing Protocol that can Withstand Black Hole Attack, International Conference on Computational Intelligence and Security, 2009, pp. 421-425, 11-14 Dec. 2009
- [3] Piyush Agrawal and R. K. Ghosh, "Cooperative Black and Gray Hole Attacks in Mobile Ad Hoc Networks", ICUIMC '08 Proceedings of the 2nd international conference on Ubiquitous information management and communication, pp 310-314, 2008
- [4] K. Sanzgiri, B. Dahill, B. N. Levine, C. Shields, and E. Belding-Royer, "A Secure Routing Protocol for Ad Hoc Networks," in Proc. of 2002 IEEE International Conference on Network Protocols (ICNP), pp. 778-89, Nov. 12-15, 2002.
- [5] H. Li, Z. Chen and X. Qin, "Secure Routing in Wired Networks and Wireless Ad Hoc Networks," Univ. of Kentucky, Department of Computer Science, Term-paper, 2003.
- [6] W. Stallings, Cryptography and network security, Principles and practice, 2nd ed., Prentice Hall, Inc, 1999, pp. 6-9.
- [7] S. Rajavaram, H. Shah, V. Shanbhag, J. Undercoffer, and A. Joshi, "Neighborhood Watch: An Intrusion Detection and Response Protocol for Mobile Ad Hoc Networks," Student Research Conference, University of Maryland at Baltimore County (UMBC), May 3, 2002.
- [8] A. Burg, "Ad hoc networks specific attacks", Technische Universität München, Institut für Informatik, Seminar Paper, Seminar Ad Hoc Networking: concept, applications, and security, Nov., 2003. pp 12, http://wwwspies.informatik.tu-muenchen.de/lehre/seminare/WS0304/UB-hs/burg-ad_hoc_specific_attacks-paper.pdf
- [9] Johnny Wong, Xia Wang ,An End-to-end Detection of Wormhole Attack in Wireless Ad-hoc Networks, Computer Software and Applications Conference, Annual International , pp. 39-48, July 2007

2nd Annual International Conference IEMCON 2012

- [10] Rashid Hafeez Khokhar, Md Asri Ngadi & Satria Mandala : A Review of Current Routing Attacks in Mobile Ad Hoc Networks, Volume 2 Issue 3, pp 18-29, 2008
- [11] Rutvij H. Jhaveri, Ashish D. Patel, Jatin D. Parmar, Bhavin I. Shah, "MANET Routing Protocols and Wormhole Attack against AODV", Vol 10 Issue 4, pp. 12-18, 2010
- [12] Songbai Lu,Lingyan Jia,Kwok-Yan Lam,Longxuan Li, SAODV: A MANET Routing Protocol that can Withstand Black Hole Attack, International Conference on Computational Intelligence and Security, pp. 421-425, 2009
- [13] Sheenu Sharma, Roopam Gupta, "Simulation Study of BLACK HOLE Attack In MOBILE AD HOC NETWORKS", JESTEC 2009
- [14] Juan-Carlos Ruiz, Jesús Frigina, David de-Andrés, Pedro Gil : Black Hole Attack Injection in Ad hoc Networks.
- [15] M. Girish Chandra, Harihara S.G., Harish Reddy, P. Balamuralidhar,Jaydip Sen: A mechanism for detection of Gray Hole Attack in Mobile Ad Hoc Networks , Proceedings of the 6th International Conference on Information, Communications and Signal Processing (ICICS '07), Singapore, December 2010.
- [16] Vishnu K Amos J Paul , Detection and Removal of Cooperative Black/Gray hole attack in Mobile AdHoc Networks ,International Journal of Computer Applications, Number 22 - Article 8,2010 ,
- [17] Rainer Falk, Hans-Joachim Hof , "Fighting Insomnia: A Secure Wake-Up Scheme for Wireless Sensor Networks" , The International Conference on Emerging Security Information, Systems, and Technologies, June 2009.
- [18] S. Y. Ni, Y. C. Tseng, Y. S. Chen, and J. P. Sheu, "The broadcast storm problem in a mobile ad hoc network," in Proc. of the 5th annual ACM/IEEE international conference on Mobile computing and networking, pp. 151-162, Aug. 15-19, 1999.
- [19] Salmin Sultana,Elisa Bertino,Mohamed Shehab , "A Provenance Based Mechanism to Identify Malicious Packet Dropping Adversaries in Sensor Networks", International Conference on Distributed Computing Systems Workshops, June 2011.
- [20] Shoab A. Khan,Muhammad Zeshan,Attique Ahmed,Ahmad Raza Cheema, "Adding Security against Packet Dropping Attack in Mobile Ad Hoc Networks",International Seminar on Future Information Technology and Management Engineering, November 2008.
- [21] Chris Karlof, David Wagner, Secure routing in wireless sensor networks: attacks and countermeasures , Ad Hoc Networks 1 (2003).
- [22] Joaquin Garcia-Alfaro, Gimer Cervera, Michel Barbeau, Evangelos Kranakis , "Mitigation of Flooding Disruption Attacks in Hierarchical OLSR Networks" , Annual Conference on Communication Networks and Services Research, May 2011.
- [23] International Journal of Computer Applications (0975 – 8887)Volume 8– No.9, October 2010, Comparative study of Distributed Intrusion Detection in Ad-hoc Networks,
- [24] SECURITY FOR WIRELESS AD HOC NETWORKS, Farooq Anjum and Petros Mouchtaris, Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
- [25] Zhihong Qian,Yinan Li , "Mobile Agents-Based Intrusion Detection System for Mobile Ad Hoc Networks", International Conference on Innovative Computing and Communication and Asia-Pacific Conference on Information Technology and Ocean Engineering, January 2010.
- [26] Ze Li,Haiying Shen , "ARM: An Account-Based Hierarchical Reputation Management System for Wireless Ad Hoc Networks", International Conference on Distributed Computing Systems Workshops, June 2008.
- [27] Alert Aggregation in Mobile Ad Hoc Networks, by Bo Sun, Kui Wu, Udo W. PoochWiSE'03, September 19, 2003, San Diego, California, USA. Copyright 2003 ACM 1-58113-769-9/03/0009.
- [28] Guorui Li,Yingfang Fu,Jingsha He , "A Distributed Intrusion Detection Scheme for Mobile Ad Hoc Networks", Annual International Computer Software and Applications Conference, July 2007.
- [29] Xiaoxin Wu,Elisa Bertino , "An Analysis Study on Zone-Based Anonymous Communication in Mobile Ad Hoc Networks", IEEE Transactions on Dependable and Secure Computing , October 2007.
- [30] Jaydip Sen,Sripad Koilakonda,Arijit Ukil, "A Mechanism for Detection of Cooperative Black Hole Attack in Mobile Ad Hoc Networks",International Conference on Intelligent Systems, Modelling and Simulation , January 2011.
- [31] A Priority Based Protocol for Mitigating Different Attacks in Mobile Ad hoc Networks by Himadri Nath Saha, Debika Bhattacharyya & P. K. Banerjee, International Journal of Computer Science & CommunicationVol. 1, No. 2, July-December 2010, pp. 299-302.
- [32] Sanjay Ramaswamy, Huirong Fu, Manohar Sreekantadhy, John Dixon and Kendall Nygard : Prevention of Cooperative Black Hole Attack in Wireless Ad Hoc Networks.pdf.
- [33] Kejun Liu, Jing Deng, Member, IEEE, Pramod K. Varshney, Fellow, IEEE, and Kashyap Balakrishnan, Member, IEEE,"An Acknowledgment-Based Approach for the Detection of Routing Misbehavior in MANETs", IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 6, NO. 5, MAY 2007.
- [34] Zhaoyu Liu, AnthonyW. Joy, Robert A. Thompson,Department of Software and Information Systems,University of North Carolina at Charlotte,{zliu, awjoy, rthompson}@unc.edu, "A Dynamic Trust Model for Mobile Ad Hoc Networks" , 10th IEEE International Workshop.
- [35] Jihye Kim and Gene Tsudik, Computer Science Department, University of California, Irvine, {jihye, gts}@ics.uci.edu, "SRDP: Securing Route Discovery in DSR" , Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'05).
- [36] Helena Rif a-Pous and Jordi Herrera-Joancomart, Universitat Oberta de Catalunya, Rambla del Poblenou, Spain, {hrifa,jherreraj}@uoc.edu, "Secure Dynamic MANET On-demand (SEDYMO) Routing Protocol", Fifth Annual Conference on Communication Networks and Services Research(CNSR'07).
- [37] Yih-Chun Hu,yihchun@cs.cmu.edu, Adrian Perrig, perrig@cmu.edu, David B. Johnson, dbj@cs.rice.edu, "Rushing Attacks and Defense in Wireless Ad Hoc NetworkRouting Protocols", WiSe '03 Proceedings of the 2nd ACM workshop on Wireless security.
- [38] Manel Guerrero Zapata: "Secure Ad hoc On-Demand Distance Vector (SAODV) Routing" September 2006.
- [39] International Journal of Computer Theory and Engineering, Vol. 1, No. 4, October2009 1793-8201, Secure Routing Techniques for MANETs, Dr. Harsh Sadawarti and Anuj K. Gupta, Member, IAENG.
- [40] Ajay Jadhav and Eric E. Johnson, Senior Member, IEEE, "Secure Neighborhood Routing Protocol" , MILCOM 2006 , October 2006.
- [41] D. Llewellyn-Jones,M. Merabti,S. Abbas , "Signal Strength Based Sybil Attack Detection in Wireless Ad Hoc Networks", International Conference on Developments in eSystems Engineering, December 2009.
- [42] John Brooke,Saorsh Hashmi , "Towards Sybil Resistant Authentication in Mobile Ad Hoc Networks", The International Conference on Emerging Security Information, Systems, and Technologies, July 2010.
- [43] N. Bhalaji, A. R. Sivaramkrishnan, Sinchan Banerjee, V. Sundar, and A. Shanmugam, "Trust Enhanced Dynamic Source Routing Protocol for Adhoc Networks", World Academy of Science, Engineering and Technology 49 2009.
- [44] Jean-Yves Le Boudec,Sonja Buchegger , "Performance analysis of the CONFIDANT protocol" , Proceedings of the third ACM international symposium on Mobile ad hoc networking & computing (MobiHoc '02).
- [45] Sun Choi, Doo-young Kim, Do-hyeon Lee, Jae-il Jung , "WAP: Wormhole Attack Prevention Algorithm in Mobile Ad Hoc Networks", International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, June 2008.
- [46] Stephen Mark Glass, Vallipuram Muthukumarasamy, Marius Portmann , "Detecting Man-in-the-Middle and Wormhole Attacks in Wireless Mesh Networks", International Conference on Advanced Information Networking and Applications, May 2009.
- [47] C.E. Perkins and P. Bhagwat, "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers," Proc. ACM SIGCOMM, 1994.

AN INVERSE CELL BREATHING BASED POWER MANAGEMENT FOR CONGESTED MOBILE NETWORK

Anwesha Mukherjee

West Bengal University of Technology, BF-142, Sector-I,
Salt lake, Kolkata-700064, India,
Email-anweshamukherjee2011@gmail.com

Debashis De

West Bengal University of Technology, BF-142, Sector-I,
Salt lake, Kolkata-700064, India
Email-dr.debashis.de@gmail.com

Abstract— This paper presents an innovative power management method based on inverse cell breathing to develop green mobile network. When the traditional cell breathing method fails to deal with high traffic, the proposed inverse cell breathing scheme can handle this situation by dividing the overloaded cell into three sectors. This method gives 14.44-72.22% reduction in total transmitted power by all the base stations contained in a location area than that of traditional cell breathing considering only the lightly loaded cell adjacent to the overloaded sector.

Keywords- Power management; Inverse Cell breathing; Transmitted power; Traffic intensity

I. INTRODUCTION

Every mobile radio communication system is designed with an objective of providing continuous communication to mobile subscribers occupying arbitrary locations in the service area. To provide green mobile network, optimization of the transmitted power of each base station is required without compromising with the quality of service even during peak period. Different type of channel allocation and channel borrowing methods have been proposed in [1], [2], [3] to deal with high traffic load. For a given transmitted power, the achievable signal coverage determines the range of operation of a cell-site transmitter and the size of a cell in a cellular telephony. For the capacity bounds in the existence of interference due to non-ideal Orthogonality of codes in the used CDMA system and background noise methods are proposed in [4]. To address the problem of how a mobile node in a wireless network should vary its transmitter power, thus the energy consumption is minimized, subject to fixed quality of service constraints methods are proposed in [5], [6]. For the purpose of energy saving another power management strategy is proposed that evaluates the usage of partial shutdown strategy for multi-antenna base station systems to allow better adaptation of power to the network traffic required by the users [7].

This paper proposes a green power management as well as channel allocation method during peak period. It provides a new mechanism to serve all the subscribers an interruption less high quality service even in congested network. Organization of the paper is as follows: Section II contains the proposed power management strategy, the corresponding mathematical

expressions; Performance analysis of the proposed method is given in Section III; Section IV contains the Conclusion.

II. PROPOSED POWER MANAGEMENT STRATEGY

The network is partitioned into a number of location areas and each location area contains a number of cells. In this method we have considered a location area containing seven cells.

A. Parameters used

TABLE1 presents the parameters used in power calculation in the proposed method.

TABLE 1: PARAMETERS USED

| Parameter | Definition |
|-----------|--|
| P_t | Transmitted power of a BTS (Base Transceiver Station) |
| P_r | Received power by a MT (Mobile Terminal) |
| λ | Wavelength given by $\frac{c}{f_c}$ where c is the speed of light (3×10^8 km/sec) and f_c is the frequency of the carrier wave (150-1500MHz) |
| R | Radius of a cell |
| L | System loss factor not related to propagation |
| G_t | Transmitter antenna gain |
| G_r | Receiver antenna gain |
| d | T-R separation distance |
| h_{te} | Effective transmitter antenna height (30-200m) |
| h_{re} | Effective receiver antenna height(1-10m) |

B. Inverse Cell Breathing Algorithm

In this algorithm, the current traffic load of each cell is compared against its maximum traffic load that it can handle. If its current traffic load is less than or equals to its maximum allowable traffic, then no cell breathing will be performed. Else if, the cell is overloaded, then it seeks into the adjacent cells whether these cells are lightly loaded. If all of them are lightly

loaded, then the traditional cell breathing can be used [8]. In this method, the overloaded cell reduces its coverage area and the adjacent cells increase their coverage area to capture the customers situated at the border of the overloaded cell [8]. If some of these adjacent cells are fully loaded then traditional cell breathing cannot be applied to deal with such situation. Then the proposed cell breathing method can be used to deal with this high traffic. The proposed cell breathing algorithm is as follows:

1. **If** the current traffic load of a cell > maximum allowable traffic load in that cell,
2. Divide the overloaded cell into three sectors;
3. Identify the adjacent cells of a sector;
4. Divide the location area containing the overloaded cell into three regions where in each region contains a sector and its adjacent cells;
5. In a sector identify the customers located comparatively far from the base station;
6. **If** any one of the cells or both of the cells adjacent to the sector are lightly loaded,
 - a. The fully loaded cell reduces its coverage area and pushed these customers to the lightly loaded cells or cell of that region;
 - b. The lightly loaded cells or cell increase their coverage area to capture the customers;
7. **Else**
 - a. In that location area identify the lightly loaded cells those can capture and provide service to these customers;
 - b. Identify the sector which is adjacent to these cells.
 - c. The customers of the sector who are located comparatively far from the base station adjacent to the lightly loaded cells are identified;
 - d. Channels from the adjacent lightly loaded cell are allocated to these customers and the channels of the overloaded cell already they occupied are released from them;
 - e. The overloaded cell reduces its coverage area and pushed these customers to the lightly loaded cells;
 - f. The adjacent lightly loaded cells capture these customers by increasing their coverage;
 - g. The overloaded cell allocates the released channels to the customers of the previous sector;
8. **End**
9. Check the current traffic load of the cell;
10. **If** current traffic load \leq maximum allowable traffic load,
 - a. Go to Step 15;
11. **Else**
 - a. Steps (5)-(10) will be applied for the next region;
12. **End**
13. **Else**
14. No Cell Breathing;
15. **End**

In Fig. 1, we have divided the location area containing seven cells into three regions. In this figure cell7 is overloaded i.e. its current traffic load > its maximum allowable traffic load. If some of its adjacent cells are also fully loaded then to deal with this situation traditional cell breathing method will not work. To overcome this problem, our proposed cell breathing method can be used. If we divide cell7 into three sectors, then the customers (located comparatively far from the base station) of sector1 can be pushed to the cells adjacent to sector1 i.e. cell1 and cell6 as the arrow indicates in Fig. 1. (a) in case of region 1.

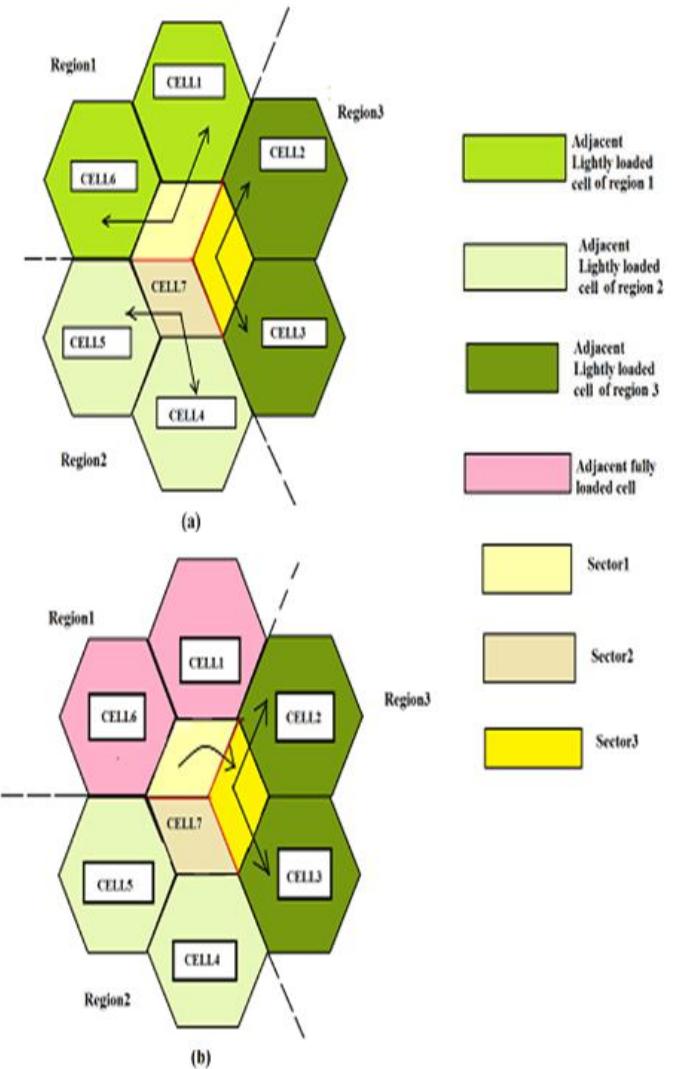


Fig. 1. Proposed Inverse Cell Breathing Method

If one of these two cells is fully loaded, then the customers are pushed only to the lightly loaded cell. Now if both of them are fully loaded, and cell2, cell3 of region3 are lightly loaded, then according to the proposed method, the customers (located comparatively far from the base station) of sector3 are pushed to cell2 and cell3 as the arrow indicates in Fig. 1. (b). Channels from cell2 or cell3 are allocated to these customers and the channels of cell7 already occupied by them are released from them. Released channels are allocated to the customers of sector1 of cell7. In traditional cell breathing all

the adjacent cells of the overloaded cell capture the customers located at the border region of overloaded cell by increasing their coverage area and the overloaded cell reduces its coverage area. In this algorithm the customers of the overloaded cell are pushed to the lightly loaded adjacent cell and the adjacent cell captures these customers. Hence we can conclude that in this proposed method we have applied the opposite cell breathing method and thus it is called inverse cell breathing. In this method all the adjacent cells are not involved, only the lightly loaded adjacent cells are involved.

C. Power Consumption by Base Station in Proposed Method

$$\text{The area of a hexagonal cell is given by, } A = \frac{3\sqrt{3}}{2} R^2 \quad (1)$$

The received power by a MT at a distance d from the BTS of a cell is given by [9],

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \quad (2)$$

The minimum received power by a MT is the received power by it when it is situated at the border region of the cell. For the MT situated at the border region of a cell, the received power is

$$\text{given by, } P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 R^2 L} \quad (3)$$

Hence the transmitted power of each BTS to serve all subscribers is given by, $P_t = \frac{P_r(4\pi)^2 R^2 L}{G_t G_r \lambda^2}$ (4)

Consequently the total transmitted power of that area containing seven cells is given by,

$$P_{tot} = 7 \times P_t = 7 \times \frac{P_r(4\pi)^2 R^2 L}{G_t G_r \lambda^2} \quad (5)$$

If cell₇ is overloaded then if normal cell breathing method is applied the new transmitted power of cell₇ (after reduction in its

$$\text{coverage area}) \text{ will be given by, } P_{t7} = \frac{P_r(4\pi)^2 (R-r)^2 L}{G_t G_r \lambda^2} \quad (6)$$

where r is the reduction in radius and $(R-r)$ is the new radius of cell₇ after reducing its coverage area. Similarly the transmitted power of each adjacent base station (after increase in the coverage area due to cell breathing) will be given by,

$$P_{tadj} = \frac{P_r(4\pi)^2 (R+r)^2 L}{G_t G_r \lambda^2} \quad (7)$$

Hence the total transmitted power by the adjacent cells is,

$$P_{totadj} = 6 \times \frac{P_r(4\pi)^2 (R+r)^2 L}{G_t G_r \lambda^2} \quad (8)$$

Hence the total transmitted power in that area is,

$$P_{totCB} = P_{t7} + P_{totadj} \\ = \frac{P_r(4\pi)^2 (R-r)^2 L}{G_t G_r \lambda^2} + 6 \times \frac{P_r(4\pi)^2 (R+r)^2 L}{G_t G_r \lambda^2} \quad (9)$$

In proposed method, if the number of adjacent lightly loaded cells=m ($1 \leq m \leq 5$) , total transmitted power by the lightly loaded adjacent cells will be given by,

$$P_{totladj} = m \times \frac{P_r(4\pi)^2 (R+r)^2 L}{G_t G_r \lambda^2} \quad (10)$$

Thus total transmitted power in that area will be given by,

$$P_{totpro} = P_{t7} + P_{totladj} \\ = \frac{P_r(4\pi)^2 (R-r)^2 L}{G_t G_r \lambda^2} + m \times \frac{P_r(4\pi)^2 (R+r)^2 L}{G_t G_r \lambda^2} \quad (11)$$

$$\text{As } m < 6, m \times \frac{P_r(4\pi)^2 (R+r)^2 L}{G_t G_r \lambda^2} < 6 \times \frac{P_r(4\pi)^2 (R+r)^2 L}{G_t G_r \lambda^2}. \text{ Hence,}$$

$$\frac{P_r(4\pi)^2 (R-r)^2 L}{G_t G_r \lambda^2} + m \times \frac{P_r(4\pi)^2 (R+r)^2 L}{G_t G_r \lambda^2} < \frac{P_r(4\pi)^2 (R-r)^2 L}{G_t G_r \lambda^2} + 6 \times \frac{P_r(4\pi)^2 (R+r)^2 L}{G_t G_r \lambda^2} \quad (12)$$

$$\text{Thus, } P_{totpro} < P_{totCB} \quad (13)$$

Hence we can conclude that the proposed inverse cell breathing method requires less transmitted power than that of traditional cell breathing method. To consider the real scenario we have considered path loss. According to Hata model the median path loss (in dB) in urban area is given by [9],

$$PL = 69.55 + 26.16 \log f_c - 13.82 \log h_{te} - a(h_{re}) \\ + (44.9 - 6.55 \log h_{te}) \log d \quad (14)$$

$$\text{Where } a(h_{re}) = (1.1 \log f_c - 0.7) h_{re} - (1.56 \log f_c - 0.8) \quad (15)$$

Where $a(h_{re})$ is the correction factor for effective mobile antenna height which is a function of the size of the coverage area. Thus the maximum path loss will be,

$$PL_{max} = 69.55 + 26.16 \log f_c - 13.82 \log h_{te} - a(h_{re}) \\ + (44.9 - 6.55 \log h_{te}) \log R \quad (16)$$

Where R is the radius of the cell.

III. PERFORMANCE ANALYSIS

In this section we have compared our strategy with traditional cell breathing to demonstrate that the proposed method gives better performance than that of the traditional one depending on different data set. We have considered location area containing seven cells. The values of the parameters considered are presented in TABLE2.

TABLE 2: PARAMETER VALUE

| Parameter | Value |
|-----------|-----------------|
| P_r | 1-5.5 microwatt |
| R | 0.5-5km |
| L | 1 |
| G_t | 1dB |
| G_r | 1dB |

| | |
|----------|---------|
| d | 0.5-5km |
| h_{te} | 0.1km |
| h_{re} | 0.005km |

Fig.2 presents the path loss (in dB) calculated using equation (14) with respect to the T-R separation distance (0.5-5km).

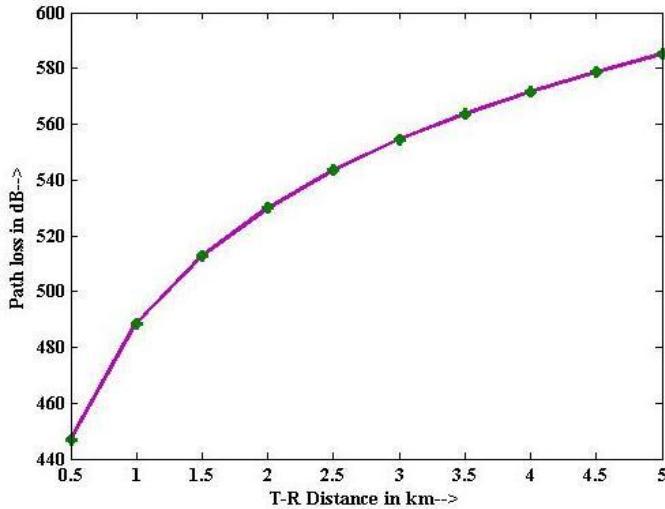


Fig. 2. T-R separation distance vs. Path loss

From equation (14) it is observed that path loss is directly proportional to the T-R separation distance. Fig.2 also establishes this. In Fig.3 we have presented total transmitted power by all the seven BTSS contained in an area using traditional cell breathing calculated using equation (9) and that of in proposed inverse cell breathing calculated using equation (11). In the figure, m ranges from 1 to 5 and it represents the number of lightly loaded cells adjacent to the overloaded cell.

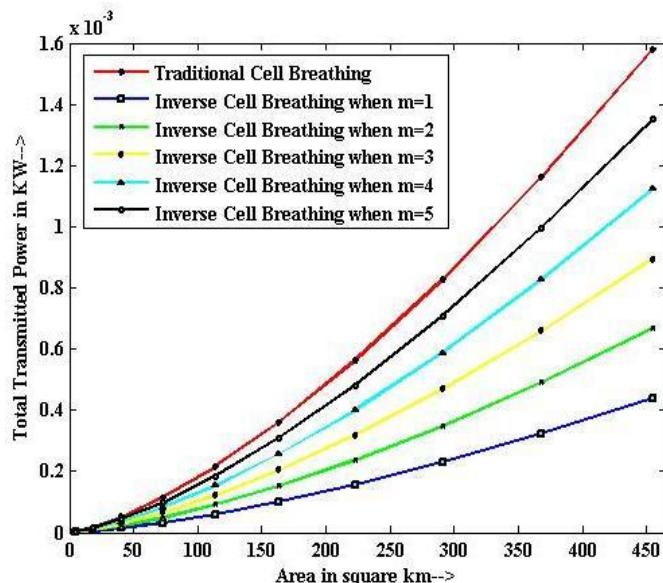


Fig. 3. Total coverage area by seven cells vs. Total transmitted power by all seven BTSS in that area

In section II we have already proved theoretically that our proposed cell breathing method reduces the total transmitted

power by all BTSSs in a location area compared to traditional cell breathing method. Fig.3 also shows that our proposed inverse cell breathing reduces 14.44-72.22% of the total transmitted power in a location area containing seven cells than traditional cell breathing method depending on m as represented in section II. Hence we can conclude that inverse cell breathing offers low power consumption than the traditional one.

IV. CONCLUSION

An inverse cell breathing based power management method is proposed for a congested mobile network. This method gives reduction in transmitted power than traditional cell breathing to develop green mobile network. The power of each base station is calculated based on the coverage area of the cell. To present real scenario we have considered path loss. If the traffic load of a cell exceeds its maximum allowable traffic and most of its adjacent cells are also fully loaded then to overcome this problem the proposed inverse cell breathing method can be used. This method provides 14.44-72.22% reduction in total transmitted power in a location area than that of traditional cell breathing.

ACKNOWLEDGMENT

Authors are grateful to Department of Science and Technology (DST) for sanctioning a research Project under Fast Track Young Scientist Scheme reference no.: SERC/ET-0213/2011 under which this paper has been completed.

REFERENCES

- [1] M. Ghaderi and R. Boutaba, "Call admission control in mobile cellular networks: A comprehensive survey," *Wireless Communications and Mobile Computing*, vol. 6, 2006, pp. 69–93.
- [2] W. Sitao, T. W. S. Chow, K. T. Ng, and K. F. Tsang, "Improvement of borrowing channel assignment for patterned traffic load by online cellular probabilistic self-organizing map," *Neural Computing & Applications*, vol. 15, 2006, pp. 298-309.
- [3] Y. Zhang and E. Salari, "A Hybrid Channel Allocation Algorithm with Priority to Handoff Calls in Mobile Cellular Networks," *Computer Communications*, vol. 32, 2009, pp.880-887.
- [4] A. Zreikat and K. Al-Begain, "Simulation of 3G Networks in Realistic Propagation Environments," *I.J. of SIMULATION*, 2003, vol .4, pp. 21-30.
- [5] J. M. Rulnick and N. Bambos, "Mobile power management for maximum battery life in wireless communication networks," *Fifteenth Annual Joint Conference of the IEEE Computer Societies*, 1996, pp.443-450.
- [6] J. M. Rulnick and N. Bambos, "Power control and time division: the CDMA versus TDMA question," *Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, 1997, pp. 634 – 641.
- [7] I. Humar, Z. Jing, W. Zeshi, X. Lin, "Energy Savings Modeling and Performance Analysis in Multi-Power-State Base Station Systems," *Green Computing and Communications, IEEE/ACM International Conference on & International Conference on Cyber, Physical and Social Computing, 2010* , pp. 474 – 478.
- [8] D. De and S. Chakrabarty, "Methods to Sustain Cellular Channel Availability in High Traffic Location," *URSIGA*, 2011, pp.1-4.
- [9] T. S. Rappaport, *Wireless Communications* (Prentice-Hall), Englewood Cliffs, 1996, NJ.

Key Frame Detection Using Two-Step Corner Selection

Monika Agarwal
 Electronics and Communication
 Department
 National Institute of Technology
 Durgapur, West Bengal, India
 monika.agarwal1990@gmail.com

Satyabrata Maity
 A.K.Choudhury School of Information
 Technology, University of Calcutta
 Kolkata, India
 satyabrata.maity@gmail.com

Amlan Chakrabarti
 A.K.Choudhury School of Information
 Technology, University of Calcutta
 Kolkata, India
 acakes@caluniv.ac.in

Abstract— A summarized video helps the viewer to get an overall idea of its content prior to watching the entire video. This not only saves time but also bandwidth required for data transfer, resulting in easy and faster video browsing. Apart from browsing a video in a smaller amount of time, summarized video information plays an important role in identifying the key event frames of a video which have lot of need in security and surveillance applications. Any video contains a large number of similar image frames which attributes to its continuity. Selecting the key-frames refers to selection of those image frames which reveal the vital information of the video without any repetition. In this paper, a method of key-frame selection for summarization of a video has been proposed, based upon the idea that large difference in the number of corner points between two images will indicate considerable change in object or event captured by the two images. The corner point detection is carried out using two-step corner selection method. The grouping of frames and key frame detection was done using Disturbance Ratio Method. The work is novel in the sense, that a statistical measure related to the variation in the corner point information has been used to identify the key frames of a video, which is a new approach in video summarization. This method was applied on some example videos, showing the effectiveness of this technique.

Keywords- Disturbance ratio, corner detection, edge detection, summarization

I. INTRODUCTION

Video Summarization is the process of generating an easily interpreted synopsis of a video by abstracting the main occurrences, scenes or objects in the video. The three main features, which a summarized video must have, are: i. *It must contain high priority entities and contents from a video*, ii. *The summary itself should exhibit a reasonable degree of continuity* and iii. *It should be free from repetition* [10]. This all sums up to the need that summarization of any video requires the detection of key frames. A video is basically composed of a large number of image frames; with frames appearing at a rate say 30-35 frames per second on the screen to generate a video experience. To maintain the continuity in the video, the frames closer to each

other in time interval, do not show appreciable information change between them. A group of frames selected in such a way that each frame carries unique information, not conveyed by any other frame in that group and the group in all captures the overall message conveyed by the video, are said to be the key frames of the video. A group of similar set of image frames is called scene and a sudden change from one scene to another determines a scene transition. So the collection of frames, one from each scene occurring in a video very efficiently serves as the key frames of the video. In create a summarization we have to add some intermediate frames to maintain the continuity.

Because of its vast application, several works have been done in order to find an efficient method for video summarization. In [1], we find color based selection of key frames and [2] was based on motion matrix which was function of optical flow within the frames. Object based selection is proposed in [3], where the key frames were categorized as key frame by event and key frame by action, based on change of region and feature extraction. Vector space based key frame detection [4] represented a frame as a point in multidimensional space and adopted curve splitting algorithm, detecting key frames based on curve properties. While in [5] singular value decomposition is used for detection of key frames, they used color histogram to represent video frames.

The key contributions of this paper can be summarized as follows:

- We have introduced the idea of using corner point detection as a feature that can be utilized to measure the change of information between the image frames in a video.
- A statistical measure of the corner information variation, between the image frames, has been utilized in grouping the frames and as well as detecting the key frame.
- Our methodology resulted to a much reduced size of the summarized video, which shows the effectiveness of the technique.

The organization of the paper is as follows. Section II discusses our proposed methodology, results and discussions appear in Section III and concluding remarks are briefed in Section IV.

II. THE PROPOSED METHODOLOGY

The key to a good video-summarization is selection of the parameter to compare the information variation between the image frames. The grouping of image frames into scenes largely depends on this parameter. It determines the extent to which the algorithm will be able to distinguish between similar and dissimilar scenes. Corners are an important feature of an image and can be robustly detected. Corner points in an image deliver a great deal of information about the image, so does the number of corner points. In a video, we have a number of similar images close together, so the probability that a large change in the number of corners between two successive image-frames will mean a change in scene, and the probability that the similar frames will have almost same number of corner points, is quite high. Thus the difference between numbers of corner points for two consecutive image frames is the description of their similarity quotient. Larger the difference in the number of corner points, less similar the two frames are.

So, we first count the number of corners in each frame of the video. Then, we group the image frames into groups of similar frames. This is done using the Disturbance Ratio method proposed later. Key frames are detected by extracting one frame from each group.

For the detection of corner points we propose a two step corner selection method. The conventional corner detection algorithm detects a large number of false corners. The two step method reduces the number of false corners to a large extent making it more reliable. In this method, instead of directly applying the corner detection algorithm we first pre-process the image using an edge detection algorithm and then apply the corner detection algorithm.

A. Two-Step Corner Selection

A corner is but an intersection of two edges. This means that true corner points will lie on the edges only (in case of isolated point, it can itself be considered as an edge).

We adopt novel method of detecting a corner point by first processing the image using an edge-detector and then detecting the corners in this edge-detected image. A good edge-detection algorithm gives us an image describing only the objects in the image. Such an image is largely independent of brightness changes and additive noise. Detecting corners in the edge-detected image is found to give much better result than applying the corner detection algorithm directly as shown in "Fig. 1(a)" and "Fig. 1(b)".

The two mostly commonly used edge detection algorithms are Canny edge-detection algorithm [7] and Sobel edge-detection algorithm [9].

Although the Canny edge-detector gives more accurate result, it is computationally very complex and hence have more computation cost, making it unsuitable for application on a video, which contains a large number of image-frames. Whereas Sobel edge detector is less complex, fast and gives satisfactory results [9]. Since our consideration is large difference in the number of corners and not the exact corner points, the inaccuracy to this

extent is ignorable. Thus we adopt the Sobel Edge Detection algorithm.

Sobel Edge Detection

The Sobel operator performs a 2-D spatial gradient measurement on an image. Typically it is used to find the approximate absolute gradient magnitude at each point in an input gray scale image.

Steps:

- The Sobel edge detector uses a pair of 3x3 convolution masks, one estimating the gradient in the x-direction (columns) and the other estimating the gradient in the y-direction (rows).
- A convolution mask is usually much smaller than the actual image. As a result, the mask is slid over the image, manipulating a square of pixels at a time.
- If we define A as the source image, and G_x and G_y are two images which at each point contain the horizontal and vertical derivative approximations.

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A \text{ and } G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * A \quad (1)$$

- The magnitude of the gradient is then calculated using the formula:

$$|G| = G_x + G_y \quad (2)$$

- An approximate magnitude can be calculated using:
 $G = |G_x| + |G_y| \quad (3)$

After the edge-detected image is obtained, corner-detection is performed on the image. The several corner detection algorithms are the Moravec's corner detection algorithm [8] and Harris corner detection algorithm [8]. While Moravec's is the earliest of all corner detection algorithms, Harris method is an improved method.

Harris Corner Detection

This corner detection is an improved version of the Moravec's Corner Detection. In this method, instead of using shifted patches, differential of the corner score is directly considered.

For a shift of the window (u, v) under consideration, by say (x, y), the weighted sum of squared differences in intensity I is calculated as:

$$S(x, y) = \sum_u \sum_v w(u, v)(I(u+x, v+y) - I(u, v))^2 \quad (4)$$

Using Taylor's Expansion of I(u+x, v+y) this can be approximated to :

$$S(x, y) \approx \sum_u \sum_v w(u, v)(I(u, v)x + I(u, v)y)^2 \quad (5)$$

Where I_x and I_y are partial derivatives of I(x,y).

The above expression in matrix form is given by:

$$S(x, y) \approx \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} \quad (6)$$

Where,

$$A = \sum_u \sum_v w(u, v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix} \quad (7)$$

A is called the Harris matrix. $\langle Ix^2 \rangle$ and $\langle I_y^2 \rangle$ denotes averaging over the window (u,v) . If a Gaussian window $w(u,v)$ is used the result is isotropic, thus overcoming the defect of Moravec's algorithm which was anisotropic in response.

The eigen values of A , define the cornerness but its computation is highly complex. Hence, Harris and Stephen introduced a parameter depending upon the eigen values of A , as a measure of corner response:

$$M_c = \lambda_1 \lambda_2 - \kappa(\lambda_1 + \lambda_2)^2 = \det(A) - \kappa \text{trace}^2(A) \quad (8)$$

M_c is largely positive for a corner, small in magnitude for a flat region and largely negative in case on an edge.

Harris corner detection is widely used corner detection algorithm for its largely improved results.

It is based on the idea that a corner is a point with strong brightness changes in orthogonal or near orthogonal directions. Since it is based on the intensity of pixels it is very sensitive to additive noise and brightness changes, and hence results in a large number of false corner points.

"Fig. 1", depicts the improvement in the corners detected in the same image when on one image only corner detection algorithm was applied "Fig. 1a" and the other was processed by using the proposed two step corner selection algorithm "Fig. 1b".

Two-step corner selection method has the following advantages:

1. The detection of corner points is not affected by the brightness change within the image but only depends on the objects and their orientation, so the change in the image frames depends on the change of objects and their orientation. Thus this parameter becomes suitable for reliable shot detection mechanism.
2. The large reduction in the number of false corners makes it a true descriptor of the image frame and hence suitable for key frame detection.

B. Detection of Key Frames

We use the two step corner selection algorithm to detect corners in all the image frames comprising the video and keep a record of the number of corner points in each frame naming it, the corner value of that particular frame. This corner value is now used to compare consecutive image frames whose difference in magnitude is inversely proportional to the similarity between the two images. "Fig. 2" shows a graph of corner values against frame number for a video sequence. Next step in the detection of key frames is grouping of frames into scenes. According to the proposed idea, scenes are groups of consecutive frames with almost same number of corner values. We group the frames by the Disturbance Ratio (DR) method [10]. This method measures the disturbance in a video as the ratio of the range of corner values to the standard deviation "(11)". Range measures the difference between the highest and the lowest value, hence a measure of largest variation. Large value of DR is obtained from large value of range and low standard deviation. Low standard deviation implies negligible varieties but large range implies a major gap between the lowest and the largest value. This indicates the presence of spikes. A threshold value directly proportional to DR measure is fixed "(12)", which will determine the allowable amount of disturbance in consecutive frames to group them into a scene. In a video, say there are n numbers of image frames.

Let $x(i)$ be the corner value of image frame i where $1 \leq i \leq n$;

$$\text{Global standard deviation} = \frac{1}{n} \sqrt{\sum_{i=1}^n (x(i) - \bar{x})^2} \quad (9)$$

$$\text{Global range} = \text{global maximum} - \text{global minimum} \quad (10)$$

$$\text{DR} = (\text{Global range}) / (\text{Global standard deviation}) \quad (11)$$

$$\text{Threshold (T)} = k \times \text{DR} \quad (12)$$

K = multiplication factor (to be adjusted depending upon the video by trial and error basis)

The steps for grouping are as follows:

- (1) Starting with the first element as a group in itself, the average and the standard deviation for the group is calculated. The average is the element itself and standard deviation in this case is thus zero.
- (2) The next element is converged into the group, the average and the standard deviation of the group thus formed is computed. The standard deviation of the group is now compared with the threshold value.

Figure 1: COMPARISON BETWEEN HARRIS CORNER DETECTION TWO STEP CORNER SELECTION ALGORITHMS.



Figure 1(a): Corner points detected by Harris corner detection algorithm.



Figure 1(b): Corner points as detected by two step corner detection algorithm.

- Greater standard deviation means greater difference in values and hence implies disturbance within the group.
- (3) The next frame is converged into the group and the process repeated. As soon as the group standard deviation becomes larger than the threshold value, the group breaks and the last frame in that group is identified as a key frame.
 - (4) Fresh grouping and calculation starts from the frame next to the last key frame identified. The process is repeated over all the corner values.
 - (5) Last frame of the video has no succeeding frame for comparison and is always detected as a key frame.

III. RESULTS AND DISCUSSIONS

The proposed method of key frame detection was applied on several videos using MATLAB R2010a software. The keyframes detected for two videos [11] is shown in "Fig. 3". "Fig. 2" shows

Figure 2: VARIATION OF NUMBER OF CORNER POINTS WITH IMAGE FRAME.

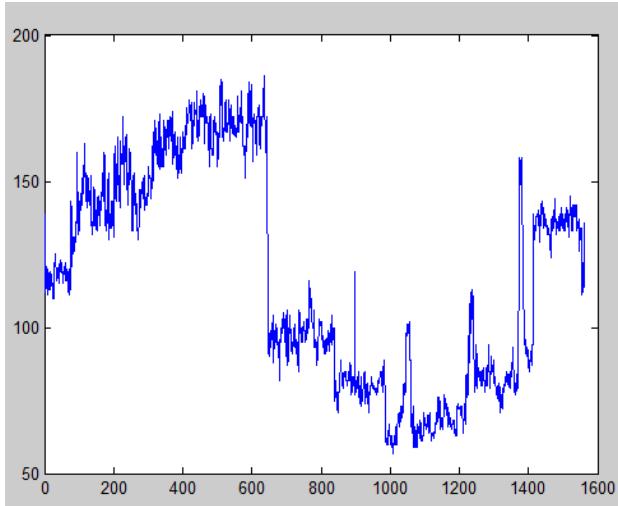


Figure 2a: Variation of number of corner points with the image frame for video1

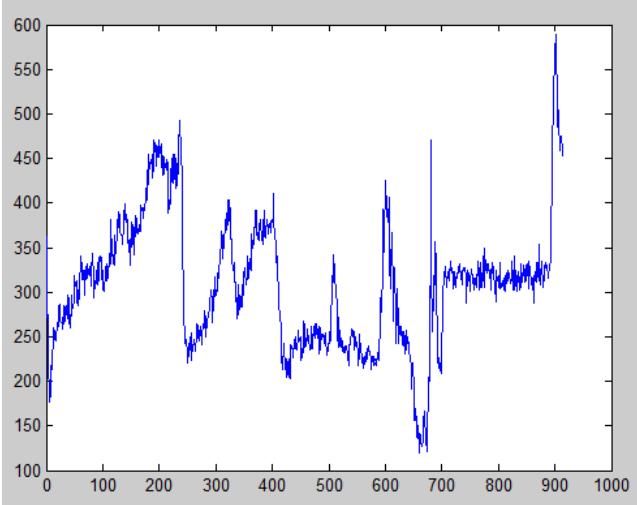


Figure 2b: Variation of number of corner points with the image frame for video2

the variation of number of corner points with the image frames for the two videos. Large variance in consecutive values indicates a scene transition. The multiplicative factor K was found out by trial and error method. "Fig. 3a" shows the key frames detected for video1 for which the multiplicative factor K was found to be 5. "Fig. 3b" shows the detected key frames for video2 for $K = 3$. Table 1 shows results of our proposed summarization technique applied on the two example videos. The results prove that our technique is efficient.

Figure 3: KEY FRAMES DETECTED FOR SAMPLE VIDEOS.



Figure 3a: Key Frames detected for video1



Figure 3b: Key Frames detected for video2

TABLE I: RESULTS OF THE SUMMARIZATION METHOD APPLIED OVER THE TWO EXAMPLE VIDEOS

| Index | Video | File Type | Total # Frames | # Key Frames Detected | % Reduction In Video Size |
|-------|--------|-----------|----------------|-----------------------|---------------------------|
| 1 | Video1 | WMV | 507 | 6 | 98.82 |
| 2 | Video2 | MPG | 913 | 10 | 98.90 |

2nd Annual International Conference IEMCON 2012

IV. CONCLUSION

In this paper a novel method of detecting the key frames in a video has been proposed, based on the corner feature. The corner detection based approach is a new work in this domain, and the results prove the efficiency of the approach. This work can be further improved by devising an algorithm where the multiplication factor 'K' can be defined as a function of the video or image frame features. The multiplication factor in that case can be dynamically allocated, resulting in self-adaptive key frame detection algorithm requiring no external input.

REFERENCES

- [1] Latecki.L.J, de Wildt. D, Jianying Hu, "Extraction of key-frames from video using optimal color composition matching and polygon simplification" in Multimedia Signal Processing,2001 IEEE Fourth Workshop on 2001
- [2] W. Wolf, "Key frame selection using motion analysis", Proceedings of ICASSP'00, vol.2, pp. 1228-1231, 1996
- [3] Changick Kim, Jenq-Neng Hwang, "Object-based Video Abstraction for Video Surveillance Systems", IEEE Transactions on circuits and Systems for video technology, vol. 12, No. 12, 2002
- [4] Li Zhao; Wei Qi; Li, S.Z.; Yang, S.Q.; Zhang H.J.; "Content-based retrieval of video shot using the improved nearest feature line method". Proceedings of ICASSP'01, vol. 3, pp. 1625-1628, 2001
- [5] Yihong Gong and Xin Liu,"Video summarization using singular value decomposition" in Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on 2000
- [6] H.J Zhang, "An integrated system for content-based video retrieval and browsing". Pattern Recognition, vol.30, no.4, pp.643-658, 1997.
- [7] Canny, John, "A computational approach to edge detection" in Pattern Analysis and Machine Intelligence, IEEE Transactions on November 1986.
- [8] C. Harris and M. Stephens (1988). "A combined corner and edge detector". Proceedings of the 4th Alvey Vision Conference. pp. 147–151
- [9] Sharifi.M, Fathy.M, Mahmoudi.M.T, "A classified and a comparative study of edge detection algorithms" in Information Technology: Coding and Computing, 2002. Proceedings International Conference on April 2002.
- [10] S. Maity, A.Chakrabarti and D.Bhattacharjee; "An Innovative Technique for Adaptive Video Summarization", SPRINGER, ICIP, Bangalore, August,2011
- [11] www.youtube.com/user/2000turtle

An Efficient Approach for Designing High Performance Squaring Circuit using Vedic Algorithms

J. Hazra¹, S. Mallick¹, J.Saha¹, B.Das¹

¹Department of Electronics and Communication Engineering, Institute of Engineering and Management, Kolkata-700091, India

Email: jubinhazra.iem@gmail.com, send4saptarshi@gmail.com, joydeepsaha007@gmail.com, 1991biswajitdas@gmail.com,

Abstract— This paper proposes a high speed low power squaring circuit design (ASIC) using ‘Vedic Mathematics’, the ancient methodology of mathematics which has a unique technique of calculations based on 16 ‘Sutras’ (Formulae) by means of which it offers a very efficient approach for practical mathematical applications covering a wide range of areas like algebra, arithmetic, geometry or trigonometry. Propagation delay and dynamic power consumption of a squarer were minimized significantly by removing unnecessary redundancy in partial products. The functionality of these circuits is checked and performance parameters like propagation delay and dynamic power consumption are calculated by Cadence Spice Spectre using 90nm CMOS technology. The propagation delay of the proposed 32bit squarer was only 2.97ns and consumed 23.46μW power. By combining Boolean logic with ancient Vedic mathematics, substantial amount of partial products were eliminated that resulted in considerable amount of reduction in delay and power compared with the mostly used (array, Wallace Tree & Tree based) architectures.

Keywords: Duplex (DY) Methodology, Squarer Circuit, Vedic Mathematics, Yavadunam (YVDN) Sutra.

I. INTRODUCTION

Squaring, a special kind of multiplication, is frequently used in many digital signal processing applications like adaptive filtering, vector quantization, image compression, pattern recognition etc.[1]. Many techniques have so far been proposed by various researchers to implement high speed squarer [1-7]. While, some of the researchers used symmetry [2] or rearrangement [3] of the partial products to improve squarer performance, others [5, 6] took the help of combined technique using Booths [5] or folding number [6] methods for achieving the same. Though most of the investigators [7] used Wallace Tree Multiplier (WTM) to implement squarer owing to advantages like symmetric performance and least propagation delay, few reports [8] are also there on promising performance of squarer using array based multiplier along with compressor.

‘Vedic Mathematics’ is the ancient system of Indian mathematics [9, 12] which has a unique technique of calculations based on 16 ‘Sutras’ (Formulae). Multiplier implementation at the gate level (FPGA) using Vedic Mathematics has already been reported [9], but to the best of our knowledge till date there is no report on transistor level (ASIC) implementation of such squarer. Mehta et al. [9] proposed a multiplier design using ‘Vertically and crosswise’ formula (“UT sutras”). Likewise, a multiplier design using ‘all from 9 and last from 10’ formula (“NND sutras”) has been reported by Tiwari et al [10] in 2009, but without any hardware module implementation in the circuit level. The two main formulae of ‘Vedic mathematic’ called DY (means “duplex”) and YVDN (means “Whatever the extent of its deficiency”) [11] were used to implement the squarer algorithm with two specific objectives: i) Simplicity and modularity in multiplications by removing unnecessary redundancy in partial products ii) The reduction of carry propagation delay for rapid additions.

In this approach, by employing the Vedic mathematics, an ($N \times N$) bit squarer implementation was transformed into just one squarer and one adder implementation. “Duplex” method was used to remove unnecessary redundancy by generating less number of partial products, whereas, YVDN methodology was used for the transformation of multiplication into shifting operation. Compared to existing methods (Wallace tree, Booth’s Array or linear compensation squarer), our approach resulted not only in simplified arithmetic operations, but also in a regular array-like structure leading to substantial improvement in performance. Performance parameters like propagation delay and dynamic switching power consumption of the proposed circuit were calculated by Cadence spice spectre with 90 nm CMOS technology and compared with the other design like Wallace Tree [7], Booth Array [5], array [2] and tree based [8] implementation. The calculated results revealed (32×32) bit squarer have propagation delay of only 2.97 ns with 23.46

μW dynamic switching power which have impressively outperformed some of the existing architectures.

II. MATHEMATICAL FORMULAE OF VEDIC SUTRAS

The contributions of the Indian ancient mathematicians in the world history of mathematical science are not well recognized. However, it was explored [12] the mathematical potentials in the Vedic primers and mathematical operations can be carried out mentally to produce fast answers using the Sutras. In this paper we concentrated on YVDN and DY formulae.

A. YVDN (*Whatever the extent of its deficiency*) Method

The meaning of this sutra is “Whatever the extent of its deficiency”, applicable for generating all the square numbers.

1) Illustration of YVDN Sutra

Fig. 1 represents an example of square generation procedure using YVDN methodology, where example is considered for decimal number system.

Steps for table implementation methods:

- Assume that nearest base is equal to 1000. Subtract 994 from 1000, result is equals to 6.
- Implement the square of 6, i.e. equals to 036 (Here 994 is a 3 digit number, to make a 3 digit number a 0 is padded in front of 36 to make a 3 digit number).
- Subtract 6 from 994 i.e. equals to 988.
- Concatenate the subtracted result with elementary generated squaring result, i.e. result is equals to 998036.

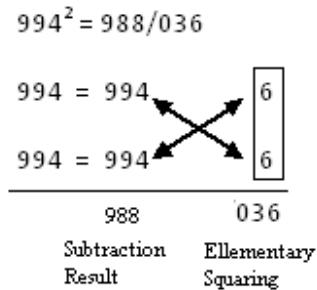


Figure 1. Squarer Implementation Using YVDN Sutra

| |
|-------------------------------|
| Illustration: |
| $5231^2 = 25/20/34/22/13/6/1$ |

Fig2. (a)

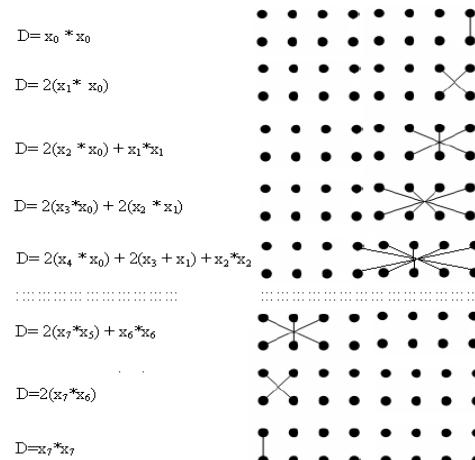


Figure 2. (a) Illustration of DY (Duplex) Method for squaring; (b) Duplex generation procedure

2) Mathematical Description of YVDN Sutra

Let X is an n digit unsigned number, and x_i are the digits, then X can represented as:

$$X = \sum_{i=0}^{n-1} x_i 10^i \quad \text{Where } x_i \in \{0,1,2,\dots,9\} \quad (1)$$

$$X^2 = \left(\sum_{i=0}^{n-1} x_i 10^i \right)^2 \quad (2)$$

The eq. (2) can be reformed for two different cases, i.e. case 1: Number is nearer to base of operation (Radix) and case 2: Number is far away from base of operation (Radix).

Case 1: Equation (2) can be reformulated as (assuming base of operation is equals to 10^n):

$$X^2 = 10^{2n} - 2 \cdot 10^n \cdot (\bar{X}) + (\bar{X})^2 \quad (3)$$

$$X^2 = 10^n \left(10^n - 2 \cdot (\bar{X}) \right) + (\bar{X})^2 \quad (4)$$

$$X^2 = 10^n \left(X - (\bar{X}) \right) + (\bar{X})^2 \quad (5)$$

Where \bar{X} is equal to complement of X . That can be implemented from subtraction result of base of operation and the actual number. For simplification considering Fig. 1, here base of operation is equal to 1000 and actual number to be squared is equals to 994, thus $\bar{X} = 6$.

Case 2: Equation (1) can be reformulated as (assuming base of operation is equals to 10^{n-1}):

$$X = x_{n-1} 10^{n-1} + \sum_{i=0}^{n-2} x_i 10^i \quad (6)$$

$$X^2 = \left(x_{n-1} 10^{n-1} + \sum_{i=0}^{n-2} x_i 10^i \right)^2 \quad (7)$$

$$X^2 = x_{n-1}^2 10^{2(n-1)} + 2x_{n-1} 10^{n-1} \sum_{i=0}^{n-2} x_i 10^i + \left(\sum_{i=0}^{n-2} x_i 10^i \right)^2 \quad (8)$$

$$X^2 = x_{n-1} 10^{n-1} \left(x_{n-1} 10^{n-1} + 2 \sum_{i=0}^{n-2} x_i 10^i \right) + \left(\sum_{i=0}^{n-2} x_i 10^i \right)^2 \quad (9)$$

$$X^2 = x_{n-1} 10^{n-1} \left(X + \sum_{i=0}^{n-2} x_i 10^i \right) + \left(\sum_{i=0}^{n-2} x_i 10^i \right)^2 \quad (10)$$

B. DY (Duplex) Method for Squaring

The meaning of this sutra is “Duplex” and it is applicable to all the squaring operations.

3) Illustration of DY Sutra

Fig. 2 (a) represents the illustration of the duplex method, where example is considered for decimal number system. Fig. 2(b) represents the graphically duplex generation procedure. For example, to implement a square of a number 5231, following steps are executed;

Steps for table implementation methods:

- i) Implement the duplex of 1, i.e., 1.
- ii) Implement the duplex of last two digits. In this example it is taken as 31. And their duplex is $(2*3*1)=6$. Put the value of the duplex before 1.
- iii) Implement the duplex of last 3 digits. In this example it is taken as 231. And their duplex is $(2*(2*1)+32)=13$ Put the value of the duplex before previous duplex, i.e. 6. Here consider only one digit for replacing, i.e., 3, and remaining 2 is the carry term.
- iv) Implement the duplex of last 4 digits. In this example it is taken 5231, and duplex is $(2*(5*1)+2(2*3))=22$. Put the value of the duplex before previous duplex, i.e. 13. Here consider only one digit for replacing, i.e., 2, and remaining 2 is the carry term.
- v) Implement the duplex of the last three terms. In this example it is taken as 523. And their duplex is $(2*(5*3)+22)=34$. Put the value of the duplex before previous duplex, i.e. 22. Here consider only one digit for replacing, i.e., 4, and remaining 3 is the carry term.
- vi) Implement the last two digit, i.e., 52, and duplex is equals to $2(5*2)=20$. Put the value of the duplex before previous duplex, i.e. 34. Here consider only one digit for replacing, i.e., 0, and remaining 2 is the carry term.
- vii) Calculate the duplex of the last digit, i.e. 5, and duplex is 25. Put the value before the previous duplex.
- viii) Arrange the terms, and add them with previous carry.

4) Mathematical Description of DY Sutra

Mathematical formulation of this algorithm is given in eqn. (11-12). Where number is equals to X [eqn (1)] and its square

is equals to X2 [eqn (2)], Then eqn. (2) can be reformulated as:

$$X^2 = \left(\sum_{i=0}^{n-1} x_i 10^i \right)^2 = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} x_i x_j 10^{i+j} \quad (11)$$

Where i and j are dummy variables.

Eqn. (11) is the weighted sum of the partial products. Expanding the eqn. (11)

$$X^2 = x_{n-1}^2 10^{2(n-1)} + 2x_{n-1} x_{n-2} 10^{n(n-1)} + (x_{n-1}^2 + 2x_{n-1} x_{n-2}) 10^{n(n-1)-1} + \dots + 2x_1 x_0 10^4 + x_0^2 \quad (12)$$

III. PROPOSED SQUARER ARCHITECTURE DESIGN

In squaring operation, conventional multipliers are used which eventually increases power and area owing to its unnecessary redundancy. In this paper we focused on the design of application specific high speed squarer with less area and power by reducing the unnecessary redundancy. The proposed architecture is based on the following mathematics, which have been adopted from Vedic mathematics [12] and reformulated in Binary number system for reported squarer implementation.

A. Mathematical Implementation of Vedic Formulae for Binary Number System

1) YVDN (Whatever the extent of its deficiency) Sutra for binary number system

If X is an n bit unsigned binary number, and xi are the bits, then X can represented as:

$$X = \sum_{i=0}^{n-1} x_i 2^i \quad (13)$$

The eqn. (13) can be rewritten as:

$$X = \begin{cases} 2^n - (\bar{X}) & \text{Radix} = 2^n \\ x_{n-1} 2^{n-1} + \sum_{i=0}^{n-2} x_i 2^i & \text{Radix} = 2^{n-1} \end{cases} \quad (14)$$

Case 1: When Radix=2n, then,

$$X^2 = 2^{2n} - 2 \cdot 2^n \cdot (\bar{X}) + (\bar{X})^2 \quad (15)$$

$$X^2 = 2^n \left(2^n - 2 \cdot (\bar{X}) \right) + (\bar{X})^2 \quad (16)$$

$$X^2 = 2^n \left(X - (\bar{X}) \right) + (\bar{X})^2 \quad (17)$$

Case 2: Radix=2n-1

$$X^2 = \left(x_{n-1} 2^{n-1} + \sum_{i=0}^{n-2} x_i 2^i \right)^2 \quad (18)$$

$$X^2 = x_{n-1}^2 2^{2(n-1)} + 2x_{n-1} 10^{n-1} \sum_{i=0}^{n-2} x_i 2^i + \left(\sum_{i=0}^{n-2} x_i 2^i \right)^2 \quad (19)$$

2) DY (Duplex) Methodology for Binary Number System

Decimal representation of this sutra is given in eqn. (12). For binary number system eqn. (12) can be reformulated as:

$$X^2 = \left(\sum_{i=0}^{n-1} x_i 2^i \right)^2 = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} x_i x_j 2^{i+j} \quad (20)$$

Eqn. (2) is the weighted sum of the partial products, assuming I and j are dummy variable. Expanding the eqn. (20)

$$X^2 = x_{n-1}^2 2^{2(n-1)} + 2x_{n-1} x_{n-2} 2^{(n-1)-1} + (x_{n-2}^2 + 2x_{n-2} x_{n-3}) 2^{(n-1)-2} + \dots + 2x_1 x_0 2^1 + x_0^2 \quad (21)$$

This procedure is simply known as array implementation of squarer exploiting addition redundancy [13]. It is an efficient squaring technique when the length of the bit is small. Fig. 3 represents the data structure implementation of the 8×8 squarer from the eqn. (21).

| | X_7 | X_6 | X_5 | X_4 | X_3 | X_2 | X_1 | X_0 | Inputs |
|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | X_7 | X_5 | X_5 | X_4 | X_3 | X_2 | X_1 | X_0 | |
| $X_7 * X_7$ | $X_7 * X_5$ | $X_6 * X_6$ | $X_7 * X_3$ | $X_5 * X_5$ | $X_7 * X_1$ | $X_4 * X_4$ | $X_6 * X_0$ | $X_5 * X_3$ | $X_7 * X_0$ |
| $X_7 * X_4$ | $X_5 * X_4$ | $X_7 * X_2$ | $X_6 * X_2$ | $X_7 * X_0$ | $X_5 * X_1$ | $X_5 * X_3$ | $X_3 * X_1$ | $X_3 * X_0$ | $X_7 * X_0$ |
| $X_6 * X_3$ | $X_5 * X_3$ | $X_5 * X_1$ | $X_4 * X_4$ | $X_4 * X_1$ | $X_4 * X_3$ | $X_2 * X_1$ | $X_2 * X_0$ | $X_1 * X_1$ | $X_5 * X_0$ |
| $X_5 * X_4$ | | $X_5 * X_2$ | | $X_3 * X_2$ | | | | | $X_4 * X_0$ |
| $X_4 * X_3$ | | | | | | | | | |

Fig 3. Data Structure of 8×8 Squarer Generated from Eq. 21

B. Flowchart Diagram for Reported Squarer

YVDN formula for squarer implementation is the combination of elementary squarer (which have been implemented using DY methodology) and subtraction or addition of two numbers. Finally the result is computed through concatenation of the addition or subtraction results with elementary squaring results. Addition or Subtraction results have been positioned to the proper place with the help of combinational shift register circuitry. The architecture that was designed for reported architecture is shown in Fig. 4, can be subdivided into several modules (dedicated for a pre-assigned task). Larger bit length it is subdivided into two parts; based on base of operation (radix). The value of the base of operation (radix) was obtained from the mean value. If the base of operation is higher than the number then it is subtracted from base of operation, and the subtraction has been implemented using 2's complement operation, otherwise the base of operation is subtracted from the given number. Multiplication with the base of operation is computed through word delay circuitry. The output from base of operation and the input number are fed to the subtractor which is generating the residue part. The residue is fed to the squaring block

which has been designed through duplex generation methodology. The addition or the subtraction of the residue with the input is controlled by the carry output of the subtractor. If the carry is low then subtraction operation is executed otherwise addition operation is executed. The result generated by the add/sub unit shifted to the left by combinational circuit shifter and the delay is controlled by the exponent generated from base selection unit (BSU). At last the result generated by elementary squarer and the output of the delayed circuitry are added to get the required squared output.

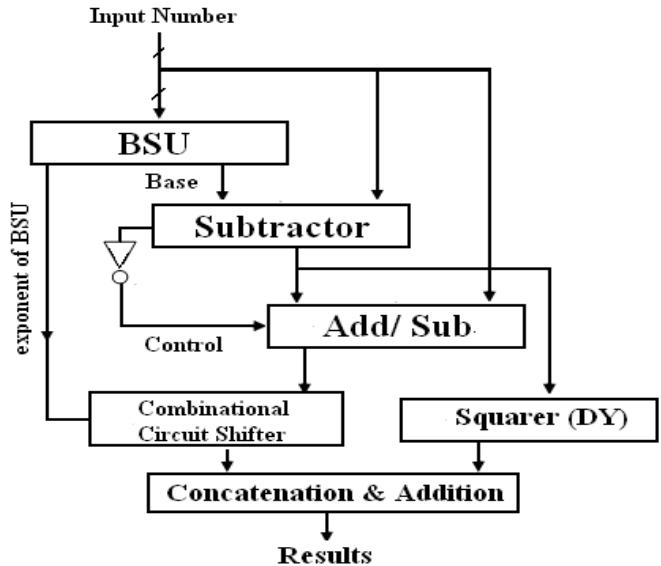


Fig 4. Hardware implementation of Proposed Squarer using YVDN formula

IV. SQUARER IMPLEMENTATION USING DY (DUPLEX) METHOD

In 8-bit squaring operation 34 partial products are generated as shown in Fig. 3. It is also to be noted that the maximum column length is only five. So this reduction process reduces the number of gates, half adders and full adders that eventually reduce the power, area and the computation time. We have minimized the number of adders by introducing special kind of adders that are capable to add more than bits per decade. These adders are called compressors [14]. Binary counter property has been merged with the compressor property to develop high order compressors. Compressors and adders are used carefully so that a minimum number of outputs would be generated. As an example, consider the column number nine of Fig. 3 where five bits are added at the first stage. These five bits could be added by using a full adder and a half adder, but that will generate four (two of each adder) outputs, instead of this we have used one 5-3 compressor generate three outputs that eventually decrease the number of bits for the next stage.

V. SIMULATION RESULTS

All the simulations including squarer of different bit length (up-to 32×32) were performed in Cadence spice spectre simulator using 90nm CMOS technology. Delay and power (switching power) of squarers are shown in Table I. The propagation delay and power are the worst case delay &

power of all possible bit combinations. A comparative study between conventional architecture and proposed architecture for 32×32 bit binary squarer is shown in Table II. It is revealed from table II, that BM [6] architecture consumed more power ($72.13 \mu\text{W}$) and delay (10.92 ns) compared to other conventional architectures. Proposed architecture offered almost 46%, 66%, 73% and 23% faster operation than that of ABS, WTM, BM and TBS architecture respectively. It was further observed that proposed architecture consumed 51%, 65%, 67% & 43% less power compared to ABS, WTM, BM and TBS respectively. E.D.P of the proposed structure is ~89% less compared to TBS which is considered as the best architecture among the conventional structures.

TABLE I. PERFORMANCE PARAMETERS SUCH AS PROPAGATION DELAY AND AVERAGE DYNAMIC POWER CONSUMPTION ANALYSIS OF VEDIC SQUARER

| Squarer Type | Using Vedic Squarer | | |
|--------------|---------------------|------------------------------|--------------------------------|
| | Delay (ns) | Avg. Power (μW) | EDP (10^{-24} J-s) |
| 4×4 | 0.29 | 4.523 | 1.31167 |
| 8×8 | 0.482 | 10.44 | 5.032 |
| 16×16 | 1.88 | 19.32 | 36.3216 |
| 32×32 | 2.97 | 23.46 | 69.6762 |

TABLE II. COMPARISON RESULT ANALYSIS OF DIFFERENT SQUARER IMPLEMENTATION

| Squarer Type | Architecture Used | Delay (ns) | Power (μW) | EDP (10^{-24} J-s) |
|--------------|-------------------|------------|-------------------------|--------------------------------|
| 32×32 | ABS [2] | 4.35 | 48 | 908.28 |
| 32×32 | WTM [7] | 8.86 | 67.39 | 5290.08 |
| 32×32 | BM [6] | 10.92 | 72.13 | 8601.24 |
| 32×32 | TBS[8] | 3.86 | 41.23 | 614.31 |
| 32×32 | Proposed | 2.97 | 23.46 | 69.6762 |

VI. CONCLUSION

In this paper a high speed low power squarer design based on the formulae of the ancient ‘Vedic Mathematics’ is reported. In this work, the squaring operation was divided into two multiplier operation in such a way so that we can achieve the goal without direct generation of partial products and that eventually reduces the operational delay & power consumption. The transistor level implementation was done in Cadence and the results were compared with the mostly used architecture like array based, Wallace tree, and Tree based design. Owing to the substantial reduction in partial product generation stages, the squarer reported here offered significant improvement in speed (~2.97 ns) with appreciably low power consumption (~ $23.5\mu\text{W}$) with reference to the ABS, WTM, and TBS based design. This squarer offered 46%, 66%, 23% improvement in speed and about 51%, 65% & 43% improvement in power consumption compared to ABS, WTM and TBS respectively.

REFERENCES

- [1] V. Garofalo, M. Coppola, D.D. Caro, E. Napoli, N. Petra, and A.G.M. Strollo, “A Novel Truncated Squarer with Linear Compensation Function,” in *Proc. IEEE Int. Symp. on Circuits and Systems*, Paris, May 30-June 02, 2010, pp. 4157-4160.
- [2] J. Pihl, and E. Aas, “A multiplier and square generator for high performance DSP applications,” in *Proc. IEEE 39th Midwest Symp. on Circuits and Systems*, (Ames, IA), Aug. 18-21, 1996, pp. 109-112.
- [3] R.K. Kologota, W.R. Griesbach, and H.R. Scrinivas, “VLSI implementation of a 350MHz $0.35\mu\text{m}$ 8 bit merged squarer”, *Electronics Letters*, vol.34, pp. 47-48, Jan. 1998.
- [4] K.E. Wires, M.J. Schulte, L.P. Marquette, and P.I. Balzola “Combined unsigned and two’s complement squarer’s,” in *Proceedings IEEE 33rd Asilomar Conf. on Signals, Systems and Computers* (Pacific Grove, CA) Oct. 24-27, 1999, pp.1215-1219.
- [5] A.G.M. Strollo, E. Napoli, and D. D. Caro, “New design of squarer circuits using Booth encoding and folding techniques,” in *Proceedings IEEE The 8th Int. Conf. on Electronics, Circuits and Systems*, Sept. 2-5, 2001, pp. 193-196.
- [6] A.G.M. Strollo, and D.D. Caro, “Booth Folding Encoding for High Performance Squarer Circuits,” *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, vol.50, pp. 250-254, May 2003.
- [7] C. S Wallace, “A Suggestion for a Fast Multiplier,” *IEEE Trans. on Computers*, vol. EC13, pp. 14-17, Dec. 1964.
- [8] G. Paul, S. Satpathy, “High performance & area efficient N-bit tree based binary squarer,” in *Proc. of VLSI Design and Test Symposium*, 2006, pp. 247-252.
- [9] P. Mehta, and D. Gawali, “Conventional versus Vedic mathematical method for Hardware implementation of a multiplier,” in *Proc. IEEE Int. Conf. on Advances in Computing, Control, and Telecommunication Technologies*, Trivandrum, Kerala, Dec. 28-29, 2009, pp. 640-642.
- [10]H. D. Tiwari, G. Gankhuyag, C. M. Kim, and Y. B. Cho, “Multiplier design based on ancient Indian Vedic Mathematics,” in *Proc. IEEE Int. SoC Design Conf.* (Busan) Nov. 24-25, 2008, pp. 65-68.
- [11]P. Saha, A. Banerjee, P. Bhattacharyya, and A. Dandapat, “High Speed ASIC Design of Complex Multiplier Using Vedic Mathematics,” in *Proc. (Abstract) IEEE Student Technology Symp. 2011*, (Kharagpur),Jan 14-16, pp. 237-241.
- [12]J.S.S.B. K.T. Maharaja, *Vedic mathematics*, Delhi: Motilal Banarsiidas Publishers Pvt Ltd, 2001.
- [13]A.J. Al-Khalili, and A. Hu, “Design of a 32-Bit squarer, - Exploiting Additive Redundancy,” in *Proc. Int. Symp. on Circuits and Systems*, May 25-28, 2003, pp. V-325-328.
- [14] P.K. Saha, A. Banerjee, and A. Dandapat, “High Speed Low Power Complex Multiplier Design Using Parallel Adders and Subtractors,” *Int. Journal on Electronic and Electrical Engineering*, (IJEEE), vol. 07, no. 11, pp 38-46, Dec. 2009.

VORONOI COVERAGE IN SWARM ROBOT POSITION CONTROL: A NEW APPROACH

Paramita Mandal

*Electrical Engineering Department
Jadavpur University,
Kolkata -700 032, India
Email: eie.paramita@gmail.com*

Ranjit Kumar Barai

*Electrical Engineering Department
Jadavpur University,
Kolkata -700 032, India
Email: ranjit.k.barai@gmail.com,*

Madhubanti Maitra

*Electrical Engineering Department
Jadavpur University,
Kolkata -700 032, India
Email: madhubanti.maitra66@hotmail.com*

Abstract - This paper presents a novel coverage control strategy that allows a group of mobile robots to position themselves such that they can optimize the measurement of sensory information in an environment. The robots use the sensed information to estimate a function indicating relative importance of the different areas in the environment. Their estimate is then used to drive the network to a desirable placement configuration using a computationally simple control law. A feasible control solution has been proposed and numerical simulations results for a swarm of mobile robots are also presented.

Keywords- voronoi; coverage control; swarm robot; position control.

I. INTRODUCTION

In nature, coordinated and cooperative or swarming behavior can be seen in a large number of creatures starting from simple bacteria to mammals. Examples of swarms include the *flocks of birds, schools of fish, herds of animals, and colonies of bacteria*. These biological societies, which consist of relatively simple members (creatures), can collectively perform complex, meaningful, and intelligent tasks. Moreover, such systems could be more *flexible* and *robust*. From the biological side they construct ordered patterns; which is a characteristic of intelligence. Another is the recognition /organization of patterns which swarm do when they optimize a function. There are basically two problems in swarm: (i) *Coverage Control*: Area coverage of multiple agents for spreading out the fronts, and (ii) *Formation Control*: Control and coordination of multiple robots moving in formation. Hazon et al. in [1] have analytically showed that the coverage will be completed as long as a single robot is able to move. The results presented therein show empirical coverage-time by running the algorithm in two different environments and several group sizes. Rekleitis et al. [2] presented an algorithmic solution for distributed complete coverage, and path planning problem. There the coverage is based on a single robot using *Boustrophedon decomposition*. The robots are initially distributed through space and each robot allocated a virtually bounded area to cover. Communication between robots is available without any restrictions. Mannadiar et al.[3] presents a new algorithm based on the *Boustrophedon cellular decomposition* of particular interest of providing a solution that guarantees the complete coverage of the free space by traversing an optimal path in terms of the distance travelled.

They presented an algorithm that encodes the areas (cells) to be covered as edges of the Reeb graph. The optimal solution to the *Chinese Postman Problem (CPP)* is used to calculate an Euler tour, which guarantees complete coverage of the available free space while minimizing the path of the robot. Proof of correctness is provided together with experimental results in different environments. A coverage-based path planning algorithm for multiple robotic agents with the application on the automated inspection of an unknown 2D environment is studied by Wang et al [4]. The proposed path planning algorithm determines a motion path that a robotic agent will follow to sweep and survey all areas of the unknown environment, which is enclosed by the known boundary. The proposed path planning algorithm has been tested and evaluated on the problem of planning path for two types of robotic agents- flying agents and crawling agents in a two-tier hierarchical mission planner to cover various unknown 2D environments. Winward et al. [5] present a team of automated vehicles safely and effectively they must be coordinated to avoid collisions and deadlock situations. Unexpected events may occur during the operation which may affect vehicles' velocities, so the coordination method must be robust with respect to these events. The method's computation speed and solution quality are evaluated through simulation, and compared with two other methods based on common path coordination techniques. Distributed coverage of environments with unknown extension using a team of networked miniature robots is verified analytically and experimentally has been proposed by Rutishauser et al [6]. The proposed algorithm is robust to positional noise and communication loss, and its performance gracefully degrades for communication and localization failures to a lower bound, which is given by the performance of a non-coordinated, randomized solution. Amstutz et al. [7] proposed the system performance is systematically analyzed at two different microscopic modeling levels, using *agent-based, discrete-event and module based, realistic* simulators. Finally, results obtained in simulation are validated using a team of Alice miniature robots involved in a distributed inspection case study. An inspired control strategy by the hunting tactics of ladybugs to simultaneously achieve sensor coverage and exploration of an area with a group of networked robots is presented by Schwager et al [8]. The controller is distributed in that it requires only information local to each robot and adaptive in that it modifies its behavior based on information in the environment. Stability is proven for both cases *adaptive*

and non adaptive with a Lyapunov-type proof. Slotine et al. [9] presented a decentralized controller is presented that causes a network of robots to converge to an optimal sensing configuration, while simultaneously learning the distribution of sensory information in the environment. Convergence and consensus is proven using a *Lyapunov-type proof*. The controller with parameter consensus is shown to perform better than the basic controller in numerical simulations. The control law is *adaptive* in that it uses sensor measurements to learn the distribution of sensory information in the environment. It is *decentralized* in that it requires only information local to each robot. These techniques are suggestive of broader applications of adaptive control methodologies to decentralized control problems in unknown dynamic environments proposed by Rus et al. [10].

In this investigation a novel coverage control strategy has been proposed where the controller can be used by groups of robot swarm to carry out tasks such as environmental monitoring and clean-up, automatic surveillance of rooms, buildings or towns, or search and rescue. This methodology may be useful for scientific studies on geological and ecological scales, and provide tools for a host of security and surveillance applications. Here, we introduce a novel multi-robot on-line coverage algorithm, based on *approximate cell decomposition* [1]. We introduce the algorithm in a form that is feasible for practical implementation on multiple robot platforms with potentially limited computational resources. The algorithm is shown to operate in realistic situations in the presence of noise on sensor measurements and actuator outputs and with asynchronous operation among the robots in the group. Simulation results are analyzed with performance metrics to quantify the performance.

The rest of the paper is organized as follows. The technical approach for the problem under consideration has been elaborated in section II, In Section III, we describe our algorithm and discuss challenges related to practical implementation. In Section IV we give results of two simulation studies and show experimental snapshots. Conclusions and discussion on the overall results are given in Section V.

II. TECHNICAL APPROACH

We model the robots as point mass and moving in a plane. Let there be n robots in a bounded, convex environment $Q \subset \mathbb{R}^N$. An arbitrary point in Q is denoted by q , the position of the i th robot is denoted by $p_i \in Q$. Let $\{V_1, \dots, V_n\}$ be the Voronoi partition of Q , for which the robot positions are the generator points. Specifically,

$$V_i = \{q \in Q \mid \|q - p_i\| \leq \|q - p_j\|, \forall j \neq i\} \quad (1)$$

(Henceforth, $\|\cdot\|$ is used to denote the l^2 -norm). The robots are assumed to have some physical extent; therefore no two robots can be at the same position at the same time, $p_i(t) \neq p_j(t)$ for all $i \neq j$.

Define the sensory function to be a continuous function $\phi : Q \mapsto R_{>0}$ (where $R_{>0}$ is the set of strictly positive real numbers). The sensory function should be thought of as a

weighting of importance over Q . We want to have many robots where $\phi(q)$ is large, and few where it is small. The function $\phi(q)$ is *not known* by the robots in the network, but the robots have sensors that can measure ϕ_{p_i} at the robot's position p_i . The precise definition of the sensory function depends on the desired application. In an application in which a team of robots are used to clean up an oil spill, an appropriate choice for the sensory function would be the concentration of the oil as a function of position in the environment. For a human surveillance application in which robots use audio sensors, $\phi(q)$ may be chosen to be the intensity of the frequency range corresponding to the human voice.

III. VORONOI DIAGRAM

The Voronoi diagram is a versatile geometric structure and has many applications in social geography, physics, astronomy, robotics, and many more fields [1, 2]. Let S denote a set of n sites (e.g., sensor nodes) in the two-dimensional plane. For two distinct sites $p, q \in S$, the *dominance* of p over q is defined as the subset of the plane being at least as close to p as to q , that is,

$$\text{dom}(p, q) = \{x \in R^2 \mid d(x, p) \leq d(x, q)\} \quad (2)$$

where $d(x, p)$ is the Euclidean distance between two points x and p .

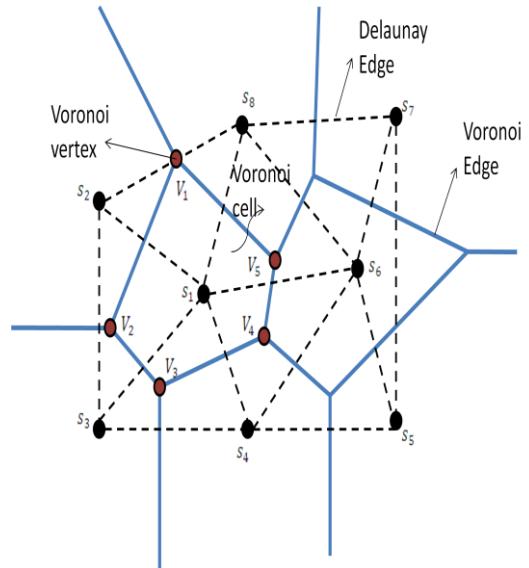


Fig:1 Voronoi Diagram

IV. PROPOSED TECHNIQUE

In our proposed technique, first of all we want to divide the whole area into Voronoi distribution. Then, we specify some robot position in the Voronoi distribution. Next, we compare the Delaunay triangulation for 10 randomly generated points and Voronoi diagram of the same points. Now, by considering the Voronoi cell, examine which robot position is under coverage. Actually, we are determining how much of the area is under coverage. The proposed algorithm has been explained by the following flowchart:

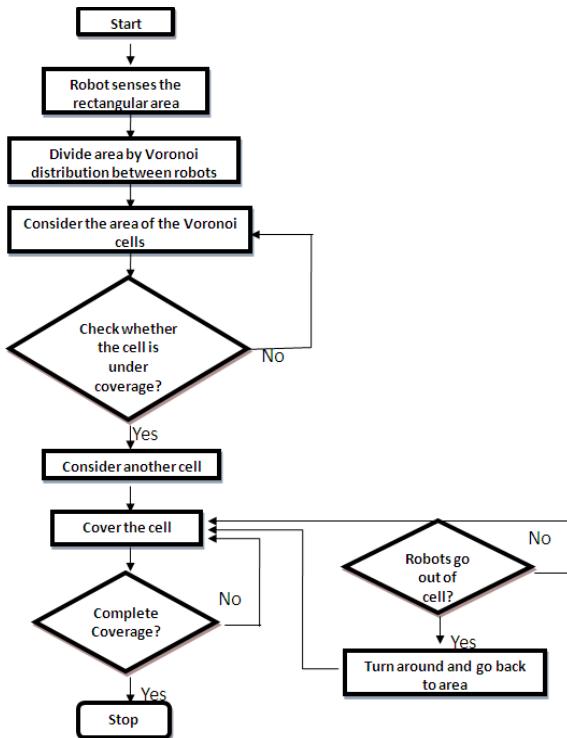


Fig.2 Flow chart of robot coverage in Voronoi cell

V. SIMULATION RESULTS

Simulations were carried out in a Matlab environment. Voronoi computation was carried out in a centralized fashion using standard Matlab Voronoi functions. The main results are depicted in Fig. 3, Fig. 4, and Fig. 5(a) and 5(b). Here, we are considering a rectangular area which has been distributed by Voronoi distribution. Here, the 10 random positions of robots in this specific area. Now, for comparison Delaunay triangulation for 10 randomly generated points and Voronoi diagram of the same points.

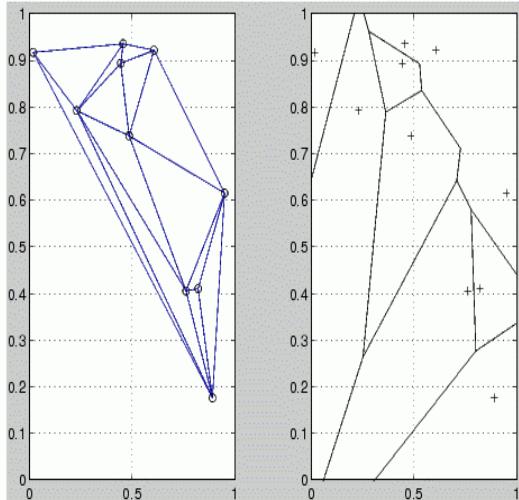


Fig.3 Delaunay triangulation for 10 randomly generated points. Voronoi diagram of the same points.

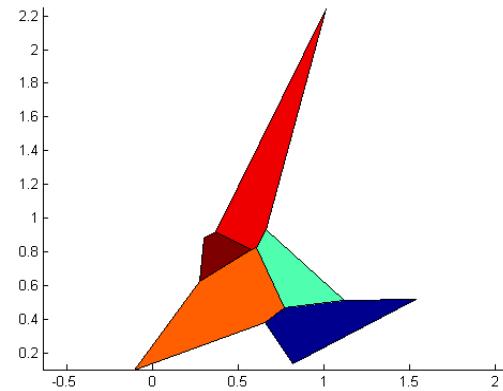


Fig.4 Voronoi and patch to fill the bounded cells of the same Voronoi diagram with color.

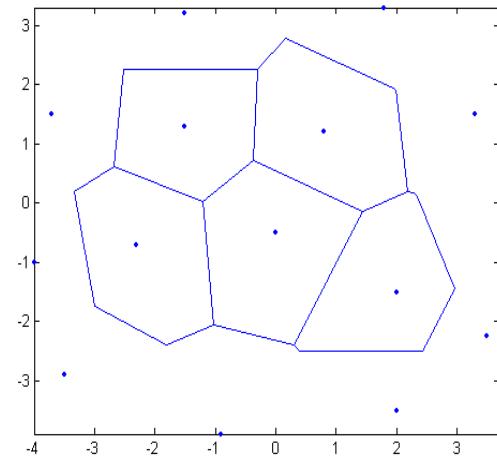


Fig.5(a)

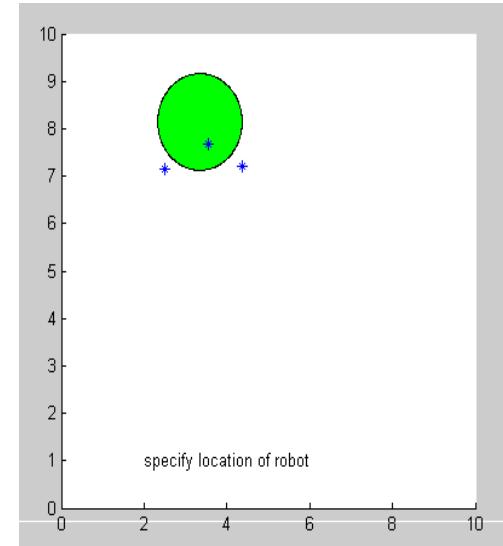


Fig.5(b)

Fig.5(a) and (b) Voronoi cell is under coverage and obstacles are there. By considering the area of the Voronoi cells, we can examine whether the cell is under coverage or not.

V. CONCLUSIONS & FUTURE WORK

This paper shows the coverage percentage of an area and how the area being covered by Voronoi distribution of the robot respectively. The foundation of this algorithm is that in order to achieve good coverage as a team, robots must 'spread out' over the environment, i.e. if robots are too close to each other, their coverage areas overlap resulting in poor overall coverage. We expect that the proposed technique will find broader application beyond the problem chosen here. It appears that consensus algorithms could be a fundamental and practical tool for enabling distributed learning, and has compelling parallels with distributed learning mechanisms in biological systems.

REFERENCES

- [1]. Noam Hazon, Fabrizio Mieli and Gal A. Kaminka, "Towards Robust On-line Multi-Robot Coverage", in *Proc. of the IEEE Int Conf. on Robotics and Automation*, pp. 1710 – 1715, 2006.
- [2]. Ioannis Rekleitis, Ai Peng New and Howie Choset, "Distributed Coverage of Unknown/Unstructured Environment by Mobile Sensor Network" in *Multi-Robot Systems, From Swarms to Intelligent Automata*, vol.III, IV, pp.145-155, 2005.
- [3].Raphael Mannadiar and Ioannis Rekleitis, "Optimal Coverage of a Known Arbitrary Environment", in *2010 IEEE Int. Conf. on Robotics and es*, pp. 5525 – 5530, 2010.
- [4]. Xudong Wang and Vassilis L. Syrmos, "Coverage Path Planning for Multiple Robotic Agent-Based Inspection of an Unknown 2D Environment", in *17th Mediterranean Conf. on Control & Automation*, pp.1295 – 1300, 2009.
- [5]. Garrett Winward and Nicholas S. Flann, "Coordination of Multiple Vehicles for Area Coverage Tasks", in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 1351 – 1356, 2007.
- [6]. Samuel Rutishauser, Nikolaus Correll and Alcherio Martinoli, "Collaborative coverage using a swarm of networked miniature robots", in *Robotics and Autonomous Systems*, vol. 57, pp. 517-525, 2009.
- [7]. Patrick Amstutz, Nikolaus Correll and Alcherio Martinoli, "Distributed Boundary Coverage with a Team of Networked miniature Robots using a Robust market-based Algorithm", in *Springer Science + Business Media B.V. 2009, Ann Math Artif Intell*, 2008, vol. 52, pp.307–333.
- [8]. Mac Schwager, Francesco Bullo, David Skelly and Daniela Rus, "A Ladybug Exploration Strategy for Distributed Adaptive Coverage Control", in *2008 IEEE Int. Conf. on Robotics and Automation*, pp. 2346 – 2353, 2008.
- [9]. Mac Schwager, Jean-Jacques Slotine, and Daniela Rus, "Consensus Learning for Distributed Coverage Control", in *2008 IEEE Int. Conf. on Robotics and Automation*, pp. 1042 – 1048, 2008.
- [10]. Mac Schwager, Daniela Rus and Jean-Jacques Slotine, "Decentralized, Adaptive Coverage Control for Networked Robots", in *2007 IEEE Int. Conf. on Robotics and Automation*, pp. 3289 – 3294, 2007.
- [11]. Bang Wang, "Coverage Control in Sensor Networks", in *Springer-Verlag London Limited*, ISBN 978-1-84996-058-8, 2010.
- [12]. Xingzhen Bai, Shu Li, Changjun Jiang, Zhengzhong Gao, "Coverage Optimization in Wireless Mobile Sensor Networks", in the 5th Int. Conf. on Wireless Communications, Networking and Mobile Computing, Wicom'09,pp. 1-4,2009.
- [13]. Noam Hazon and Gal A. Kaminka, "Redundancy,Efficiencyand Robustness in Multi-Robot Coverage", in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pp. 735-741, 2005.
- [14]. Nikolaus Correll and Alcherio Martinoli, "Robust Distributed Coverage using a Swarm of Miniature Robots", in *2007 IEEE Int. Conf. on Robotics and Automation*, pp. 379-384, 2007.
- [15]. Gyula Simon, Miklos Molnar, Laszlo Gonczy, and Bernard Cousin, "Robust k -Coverage Algorithms for Sensor Networks", *IEEE Trans. on Instrumentation and Measurement*, vol. 57, pp.1741-1748, 2008.

Modeling of Multirate 3-stage Cascaded-Integrator-Comb Filter for GSM Networks and its FPGA Implementation

Saikat Dutta

Department of Electronics and Communication Engineering
Heritage Institute of Technology
Kolkata, India
saikat_dutta007@rediffmail.com

Moloy Kumar Chowdhury

Department of Electronics and Communication Engineering
Heritage Institute of Technology
Kolkata, India
moleykumar.chowdhury@gmail.com

Abstract—The Cascaded-Integrator-Comb filter is an efficient way of implementing low-pass filters and is particularly useful for sampling rate conversion using decimation or interpolation and plays a vital role in many practical down converter systems. The Multirate Filter Design method is used for FIR filters that have very narrow transition bands, or narrow passbands or wide passbands. These FIR filters are in general not practical to design or implement as ordinary time invariant FIR filters due to the extremely long filter lengths. In this paper we study the filter chain of Digital down converter used in GSM networks and study the response of the key component – the CIC filter – response in Matlab and its hardware implementation using Simulink HDL coder

Keywords-Cascaded-Integrator-Comb (CIC), Digital Down-Converter (DDC), Global System for Mobile (GSM), Lowpass filter (LPF).

I. INTRODUCTION

In digital signal processing, a cascaded-integrator-comb (CIC) is an optimized class of finite impulse response filter combined with an interpolator or decimator.

A CIC filter consists of one or more integrator and comb filter pairs. In the case of a decimating CIC, the input signal is fed through one or more cascaded integrators, then a down-sampler, followed by one or more comb sections (equal in number to the number of integrators). An interpolating CIC is simply the reverse of this architecture, with the down-sampler replaced with a zero-stuffer (upsampler).

Multirate Signal Processing consists of using different sample rates within a system to achieve computational efficiencies that are impossible to obtain with a system that operates on a single fixed sample rate. Reducing the sampling rate requires an anti-aliasing filter prior to the decimation to a lower sampling rate. Increasing the sampling rate requires an anti-imaging filter after the interpolation. The two filters are specified using the original lowpass filter specification. To achieve any gain in computational efficiency, the two filters must run at the reduced sampling rates.

II. OVERVIEW OF THE CIC FILTER

It is easy to understand the CIC filter in the frequency domain. The CIC filter is a cascade of digital integrators followed by a cascade of combs (digital differentiators) in equal number. Between the integrators and the combs there is a digital switch or decimator, used to lower the sampling frequency of the combs signal with respect to the sampling frequency of the integrators.

Each integrator contributes to the CIC transfer function with a pole. Each comb section contributes with a zero of order D, where D is the frequency decimation ratio. The CIC transfer function in the Z-plane becomes:

$$H(z) = H_I^N(z) H_C^N(z) = \frac{(1-z^{-D})^N}{(1-z^{-1})^N} = \left[\sum_{k=0}^{D-1} z^{-k} \right]^N$$

We must be careful here because we have two sampling frequencies in the system, related by D. If we evaluate the z-transference at the output sampling frequency $z=\exp(j2\pi fs/D)$, the transference becomes

$$A(f) = \left[\frac{\sin(\pi f)}{\sin(\pi f/D)} \right]^N \approx \left[D \cdot \frac{\sin(\pi f)}{\pi f} \right]^N \quad \text{for } D \gg 1$$

It is important to make few remarks:

- The filter gain is approximately DN.
- The CIC transfer function has nulls at each multiple of the output sampling frequency $f=fs/D$.
- There are only two control parameters the number of integrator/comb stages N and the decimation ratio D. In the Graychip N is fixed at 5.
- The CIC has a wide transition band. Strictly speaking, the passband is small (the amplitude “droops” quickly) and there is substantial aliasing specially around the first mirror image. This signify that the CIC filter must be accompanied of an anti-aliasing filter or be used on narrow-band spectrums.

A. Time domain analysis

Let $u(k)$ be the input to the CIC. The integrator outputs will be called $x_1(k), \dots, x_N(k)$. The decimator output will be called $x_D(k)$, and the comb outputs will be $y_1(k), \dots, y_N(k)$

It is interesting to consider the cases of $N=1$ (one integrator/comb stage) and $N=2$ first. For $N=1$ the CIC is a simple integrator that can only remember the last D inputs.

$$x_1(n) = x_1(n-1) - u(n-1) = \sum_{k=0}^{n-1} u(k) \quad (1)$$

The decimator output only picks one every D of $x_1(k)$ outputs

$$x_D(m) = x_1(Dm) = \sum_{k=0}^{Dm-1} u(k) \quad (2)$$

In consequence, the output of the comb is

$$\begin{aligned} y_1(m) &= x_D(m-1) - x_D(m-2) = \sum_{k=0}^{D(m-1)-1} u(k) - \sum_{k=0}^{D(m-2)-1} u(k) \\ y_1(m) &= \sum_{k=D(m-2)}^{D(m-1)-1} u(k) \end{aligned} \quad (3)$$

If we ignore the filter's delay, each output is D times the average of the last D inputs to the filter.

For two stages the signal is integrated twice before decimating.

$$\begin{aligned} x_2(n) &= x_2(n-1) - x_1(n-1) = \sum_{k=0}^{n-1} x_1(k) \\ x_2(n) &= \sum_{k=0}^{n-1} x_1(k) = \sum_{k=0}^{n-1} \sum_{l=0}^k u(l) \end{aligned} \quad (4)$$

$x_2(k)$ is the running sum of $x_1(k)$ terms which grow in length as k increases.

$$x_D(m) = x_1(Dm) = \sum_{k=0}^{Dm-1} \sum_{l=0}^k u(l) \quad (5)$$

The first comb output is

$$\begin{aligned} y_1(m) &= x_D(m-1) - x_D(m-2) = \sum_{k=0}^{D(m-1)-1} \sum_{l=0}^k u(l) - \sum_{k=0}^{D(m-2)-1} \sum_{l=0}^k u(l) \\ y_1(m) &= \sum_{k=D(m-2)}^{D(m-1)-1} \sum_{l=0}^k u(l) \end{aligned} \quad (6)$$

The second comb output is

$$\begin{aligned} y_2(m) &= y_1(m-1) - y_1(m-2) = \sum_{k=D(m-3)}^{D(m-2)-1} \sum_{l=0}^k u(l) - \sum_{k=D(m-4)}^{D(m-3)-1} \sum_{l=0}^k u(l) \\ y_2(m) &= \sum_{k=D(m-3)}^{D(m-2)-1} \sum_{l=k-D+1}^k u(l) \end{aligned} \quad (7)$$

We observe that $y_2(m)$ is a double sum of input values $u(k)$. The indices of each sum (k and l) have a range equal to D. In consequence each $y_2(m)$ output is the sum of D2 samples.

The decimator output $x_D(\cdot)$ copies one out of every D $x_2(\cdot)$ samples. So each $x_D(\cdot)$ also sums all the $u(\cdot)$'s in a triangle, but the base of the triangle moves by D as $x_D(\cdot)$ increments by 1.

The combs subtract from the "current input" the "previous to the current input".

The second comb proceeds in the same way as the first comb. In this case the second comb subtracts the smaller trapezoid from the bigger trapezoid. It is interesting to note that in this case the polygons are not overlapping. However the $u(\cdot)$ terms contained in the small trapezoid are all included in the terms in the bigger trapezoid. As a result of the second comb operation we get the parallelograms shown in dashed red. We finally arrived to an interesting conclusion!

- As said, the output of the 2-stage CIC filter is the sum of D2 number input samples (i.e. $u(\cdot)$'s).
- Each output is formed by D consecutive rows of Figure 2, having D consecutive $u(\cdot)$'s in each row.
- Each row $u(\cdot)$'s are displaced in time by 1 sample as we move in ascending k.
- The CIC output contains the "last" $2*D-1$ input samples. The "last" means the last minus the CIC internal delay which is equal to the number of integrator/decimator stages.

The last bullet implies that as we increment the decimation factor D we are doing more averaging in the CIC filter. This will impact the signal-to-noise performance of the filter as it will be shown later. But first we need to generalize this result to a CIC filter with N integrator/comb stages.

The output of the Nth integrator is

$$\begin{aligned} x_N(n) &= x_N(n-1) - x_{N-1}(n-1) = \sum_{k_1=0}^{n-1} x_{N-1}(k_1) = \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{k_1} x_{N-1}(k_2) = \dots \\ x_N(n) &= \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{k_1} \dots \sum_{k_N=0}^{k_{N-1}} u(k_{N-1}) \end{aligned} \quad (8)$$

Equation (8) shows that the output of the Nth integrator contains N running sums. The indices of the sums, starting from the right most sum, go from 0 to the value of the index of the previous running sum.

$$x_D(m) = x_N(Dm) = \sum_{k_1=0}^{Dm-1} \sum_{k_2=0}^{k_1} \dots \sum_{k_N=0}^{k_{N-1}} u(k_{N-1})$$

And the output of the last comb is

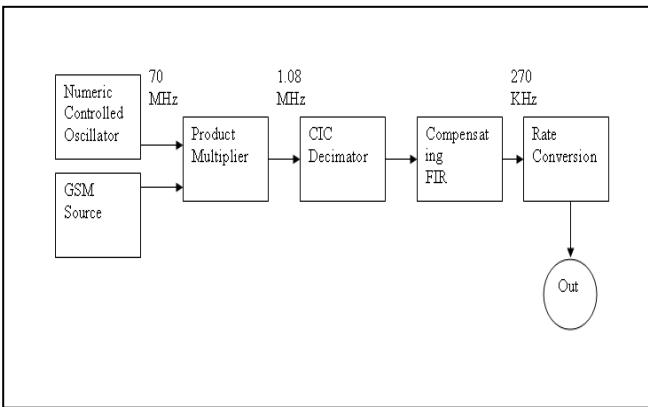
$$y_N(m) = \sum_{k_1=D(m-N)}^{D(m-N+1)-1} \sum_{k_2=k_1-D+1}^{k_1} \dots \sum_{k_{N-1}=k_{N-2}-D+1}^{k_{N-1}} u(k_{N-1}) \quad (9)$$

The output of a N-stage CIC is the sum of DN of the "last" N(D-1) input samples. The delay in the output is N if D>N.

III. DESIGN OF DDC FOR GSM NETWORKS

This section shows the design and analysis of a multistage decimator used in a digital down-converter (DDC) for a GSM signal. The design has been developed so that it can be implemented in a Graychip 4016 multi-standard quad DDC chip.

The DDC is implemented with a three-stage decimator in addition to a numerically-controlled oscillator (NCO) and a mixer. Here, we concentrate on the decimator only. The multistage decimator consists of a five-section Cascaded Integrator-Comb (CIC) decimator followed by two FIR decimators. The first decimator, CFIR, has 21 taps and provides decimation by two. The second decimator, PFIR, has 63 taps and also provides decimation by two.



The first stage of the decimator is a five-section CIC filter that provides a decimation factor of 64. This filter is attractive for high-speed implementations (the input signal in this case is 69.333248 MHz) because it is implemented without the use of multipliers. However, its magnitude response exhibits a significant droop in the passband which need to be corrected. The second stage, CFIR, is used primarily to compensate for the droop in the CIC filter. It also provides some out of band attenuation. The third stage consists of another FIR filter.

IV. RESULTS OF COMPUTER SIMULATIONS

Herein we present the results of computer simulations for the three-stage, mutirate CIC filter designed for GSM networks. CIC-filters are linear phase FIR-filters which realizes a moving average calculation of the order N. The frequency response of the filter is shown below:

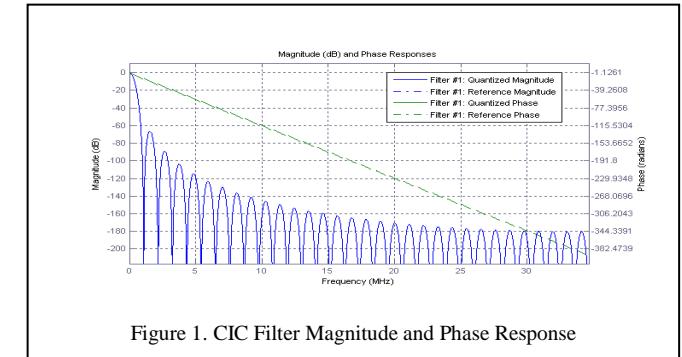


Figure 1. CIC Filter Magnitude and Phase Response

The above filter operates at an input rate of 70 MHz approximately and as shown in Fig. 1, the CIC filter has a huge pass-band gain. This is due to the structural feedback and addition operations.

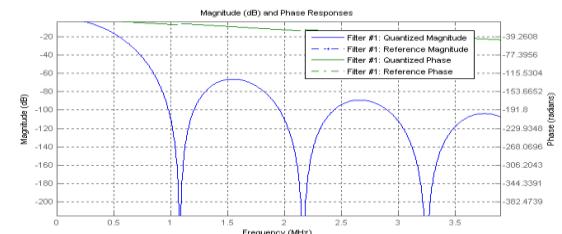


Figure 2. Magnified Response of CIC Filter

The Fig. 2 shows the magnified response in the passband region of the CIC filter. we see that the CIC has about -60 dB of attenuation at 0.8 MHz, which is within the bandwidth of interest. This is because a CIC filter is essentially a cascade of pulse-shaped type filters and therefore has a sinc-like response which causes the droop. The FIR filter in the next stage compensates for this attenuation.

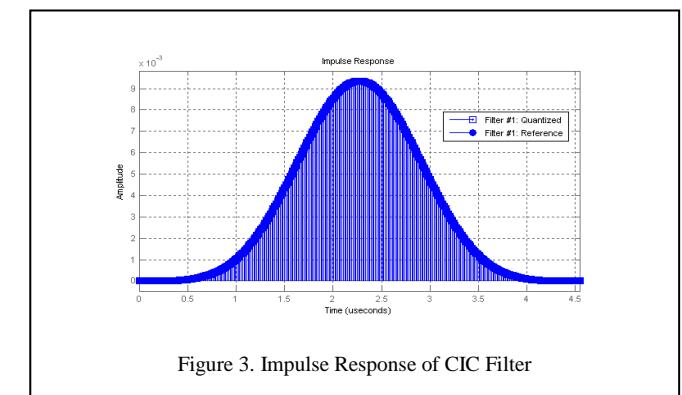


Figure 3. Impulse Response of CIC Filter

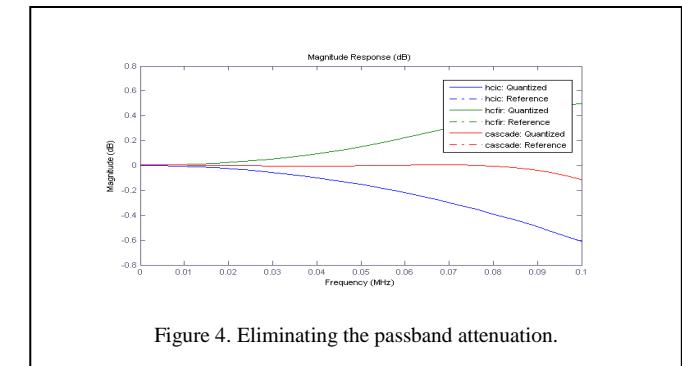


Figure 4. Eliminating the passband attenuation.

The Fig. 3 reveals the impulse response of the CIC filter. Since the CIC has a sinc-function type response, a low-pass filter that has an inverse-sinc response in the passband adjusts for the attenuation. This filter will operate at 1/64th the input sample rate which is 70 MHz, therefore its rate is 1.1 MHz. This is shown in Fig. 4. As we can see in the filter response of the cascade of the two filters, which is between the CIC response and the compensating FIR response, the pass-band attenuation has been eliminated.

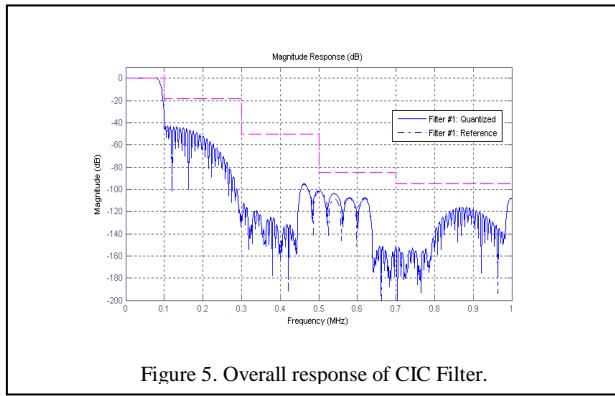


Figure 5. Overall response of CIC Filter.

The Fig. 5 displays the response for the multirate, 3-stage digital down-converter filter chain. We note that the overall filter response is obtained by cascading the normalized CIC and the two FIR filters. The CIC filter is normalized to ensure that the cascaded filter response is normalized to 0 dB. To see if the overall filter response meets the GSM specifications, we can overlay the GSM spectral mask in red dotted line on the filter response. We can see that our overall filter response is within the constraints of the GSM spectral mask. The Fig also ensures that the passband ripple meets the requirement that it is less than 0.1 dB peak-to-peak.

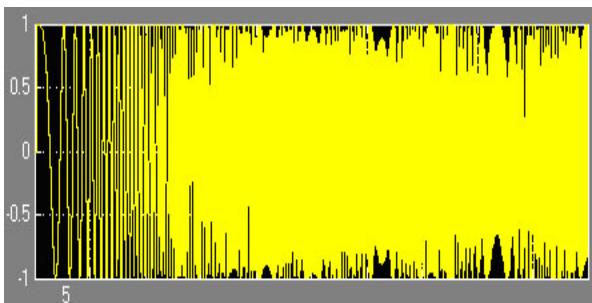


Figure 6(a). Excitation Signal.

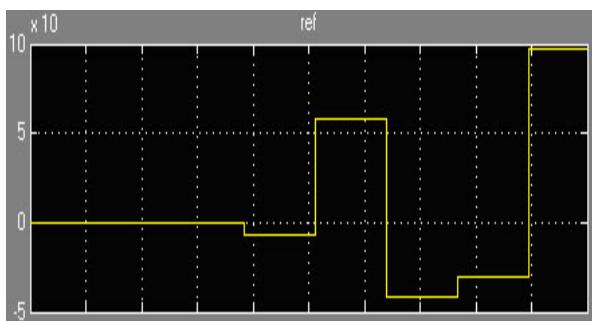


Figure 6(b). Reference Signal.

The Fig. 6(a) shows excitation chirp signal. The next signal labeled "ref" is the reference signal produced by the Simulink behavioral model of the three-stage multirate filter. This is shown in Fig. 6(b).

CONCLUSIONS

In this paper, we presented the results of three-stage, multirate filter chain of a digital down-converter (DDC) required for a GSM application. The spectral plots have been obtained using Matlab 2007b. Then we generated HDL code to implement the filter using Simulink's behavioral model. As the frequency response reveals, it fits well into the GSM mask specified. Further work can be done by verifying the generated HDL code of FDA Simulink coder with Modelsim.

The filter chain herein designed can be further extended to other application areas. More avenues like low-power approach implementation, reconfigurable type and high-speed DDC can also be studied in further extension of this paper.

REFERENCES

- [1] Eugene B. Hogenauer. An Economical Class of Digital Filters for Decimation and Interpolation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP- 29(2):155–162, April 1981.
- [2] Tim Hentschel, Gerhard Fettweis, and Marcus Bronzel. Channelization and Sample Rate Adaptation in Software Radio Terminals. In 3rd ACTS Mobile Communication Summit, pages 121–126, Rhodes, Greece, June 1998.
- [3] James C. Candy. Decimation for Sigma Delta Modulation. *IEEE Transactions on Communications*, COM-34(1):72–76, January 1986. Babic D and Renfors M (2005) Power efficient
- [4] Babic D and Renfors M (2005) Power efficient structures for conversion between arbitrary sampling rates. *IEEE Signal Processing Lett.* 12(1), 1–4.
- [5] Babic D, Saramaki T and Renfors M (2002) Sampling rate conversion between arbitrary sampling rates using polynomial-based interpolation filter. The 2nd Intl. Workshop on Spectral Methods & Multirate Signal Processing. Toulouse, France. pp: 57–64.
- [6] Crochiere RE and Rabiner LR (1983) Multirate digital signal processing. Englewood Cliffs, NJ: Prentice-Hall.
- [7] Farrow C (1998) A continuously variable digital delay element. Proc. IEEE Intl. Symp. on Circuits and Systems (ISCAS88). pp: 2642–2645.
- [8] Hentschel T and Fettweis G (2000) Continuous-time digital filter for sample rate conversion in reconfigurable radio terminals. Proc. of European Wireless Conf. pp: 55–59.
- [9] Hogenauer EB (1981) An economical class of digital filters for decimation and interpolation. *IEEE Transact. on Acoustic, Speech and Signal Proc.* ASSP29(2). pp:155-162.
- [10] Ljiljana M (2009) Multirate filtering for digital signal processing: MATLAB applications. Hershey, PA: Information Science Reference.



Saikat Dutta (M'23) was born in Vishakapatnam on 6th Jan 1988. I have completed B.Tech in ECE from Dumka Institute of Engineering & Technology, Murshidabad, West Bengal under WBUT. I am currently pursuing my final year M.Tech

2nd Annual International Conference IEMCON 2012

degree in ECE from Heritage Institute of Technology Kolkata, West Bengal under WBUT.

He has done a Winter Internship for 15 days from DURGAPUR STEEL PLANT (SAIL) in the year 2008 and Sumeer Internship of 1 month from MONIBA COMPUTER CENTER (IBM) on JAVA in the year 2008.



Moloy Kumar Chowdhury (M'78) was born in Kolkata on 17th December, 1978. He has done B.E. in Electronics & Communication Engg. Dept. from Sikkim Manipal Institute of Technology in 2001 and pursued M.E. in Electronics and Telecommunication Engg. Dept. from Jadavpur University in 2010.

He was a faculty at M.C.K.V. Institute of Engg. for almost five years and now currently serving as Assistant Professor in the Electronics & Comm. Engg. Dept. at Heritage Institute of Technology, Kolkata, West Bengal, India. He had secured 96 percentile in GATE-2008 in Electronics & Communication Engg. He has published a few papers in several International and National Conferences. He has currently submitted a chapter proposal on Ultra Wideband for book publication with Intech Publications. His research interests include Wireless mobile communications systems, Ultra Wideband, MIMO, Multi-Sensor Data Fusion, Software Defined Radio, Fuzzy applications.

Mr. Chowdhury had organized IESPC-2011 (IEEE EDS Student Paper Conference) at Heritage Institute of Technology, Kolkata, west Bengal, India in his capacity as Co-convenor in 2011.

ANTENNA & PROPAGATION

Date : 18th January,2012

Chairman : Prof. Bhaskar Gupta

Co-Chairman : Prof. Ajoy Chakraborty

Rapporter : Prof. Asima Sarkar

| Paper ID | Paper Title | Authors |
|----------|---|---|
| 26 | Elliptical Zeroth Order Resonant Antenna | Sheeja K.L , P.K Sahu and S.K Behera |
| 54 | Design and Development of E-Shaped Array Antenna using Defected Ground Structure at 5.25GHZ Band | Chandan Kumar Ghosh, Biswarup Rana and Susanata Kumar Parui |
| 59 | Hybrid Dielectric Resonator antenna with hexagonal patch for UWB applications | Kumari Padmawati, D.K.Bisoy, S.K Behera |
| 67 | Miniaturization of a Rectangular Micro strip Patch Antenna with Slots and Defected Ground Structure | Mrinmoy Chakraborty and Biswarup Rana |
| 68 | Design of a CPW-Fed Circular Shaped Slot Antenna for Wireless Application | Mrinmoy Chakraborty and Biswarup Rana |
| 69 | Non-Linear Analysis Using MATLAB and Assembly Language Programming to Investigate Possible Cause of Breakdown in power System Network | Swarnankur Ghosh,Pradip Kumar Saha,Gautam Kumar Panda |
| 71 | A 3-D Box Parity Model for Error Detection and Correction in Digital Communication Systems | Avisek Lahiri,Manashi Chakraborty |

Elliptical Zeroth Order Resonant Antenna

Sheeja K.L.¹, P.K.Sahu¹

¹ Department of Electrical Engineering,
N.I.T. Rourkela, Odisha, India.
sheejakl@gmail.com, pksahu@nitrkl.ac.in

S.K. Behera²

² Department of Electronics and Communication
Engineering, N.I.T. Rourkela, Odisha.
prof.s.k.behera@gmail.com

Abstract— In this paper a compact microstrip elliptical antenna is proposed to exhibit zeroth order resonance (ZOR) based on the Composite Right/Left Handed (CRLH) structure. The performance of the proposed antenna is designed and simulated by Ansoft HFSS. The resonant frequency, the gain and the antenna efficiency are observed to be at 3.82GHz, 7.7035dB and 85.505% respectively. The antenna exhibits single mode, omni-directional radiation pattern at its fundamental mode of resonant frequency 3.82GHz. The proposed antenna addresses the important property of being a low profile and also exhibiting a omni-directional field pattern typical of monopole antennas.

Keywords: Zeroth order resonant antenna; backward wave; electrically small antennas.

I. INTRODUCTION

Recently metamaterials and, in particular, left-handed (LH) materials [1], [2] have drawn considerable attention in the physics and engineering communities by offering new concepts such as Left-Handed(LH) propagation, zero propagation constant, backward wave, phase advance and of being electrically small. An interesting approach of these LH metamaterials is the transmission line theory, from which deep physical insight and efficient design tools have been developed [3], [4], [5]. In this approach, it has been shown that a real physical LH material is in fact composite right/left-handed (CRLH) in nature, exhibiting left-handedness at low frequencies and right-handedness at high frequencies due to parasitic right-handed (RH) effects. One practical distributed implementation of a 2D CRLH structure is the so-called mushroom structure as proposed by Sievenpiper et al. [6] in the context of high-impedance surfaces (Bragg frequencies), where it was used for its stop-band characteristics, for instance for the suppression of spurious surface waves in planar antennas. But these ZOR antennas suffer from a major drawback of having a narrow bandwidth and low gain.

In this paper a new elliptical ZOR antenna is proposed. Here a straight forward design approach is adopted. The proposed structure consists of an elliptical unit cell with microstrip line feeding. Here 3 unit cells are used. An effort is being made to circumvent the negative effect of low gain as exhibited by popular CRLH structures. So the proposed geometry exhibits a low profile high gain antenna suitable for wireless applications. Such antennas

are required in various wireless systems such as GSM/DCS cellular communication systems and point to point communication systems.

II. Zeroth Order Resonance Theory

A practical realization of a LH TL,which includes unavoidable RH effects, known as a CRLH TL is able to support an infinite wavelength ($\beta=0$ when $\omega \neq 0$) and therefore can be used to realize the proposed antenna. The dispersion diagram of the CRLH TL unit-cell is shown in Fig. 1(b). The CRLH TL supports a fundamental LH wave (phase advance) at lower frequencies and a RH wave (phase delay) at higher frequencies.

The equivalent circuit model of the CRLH TL unit-cell is shown in Fig. 1(a). By applying periodic boundary conditions (PBCs) related to the Bloch-Floquet theorem, the CRLH TL unit-cell's dispersion relation is determined to be

$$\beta(\omega) = \frac{1}{p} \cos^{-1} \left(1 - \frac{1}{2} \left(\frac{\omega_L^2}{\omega^2} + \frac{\omega^2}{\omega_R^2} - \frac{\omega_L^2}{\omega_{se}^2} - \frac{\omega_L^2}{\omega_{sh}^2} \right) \right) \quad (1)$$

$$\text{where, } \omega_L = \frac{1}{\sqrt{C_L L_L}}, \omega_R = \frac{1}{\sqrt{C_R L_R}} \quad (2)$$

$$\omega_{se} = \frac{1}{\sqrt{C_L L_R}}, \omega_{sh} = \frac{1}{\sqrt{C_R L_L}} \quad (2)$$

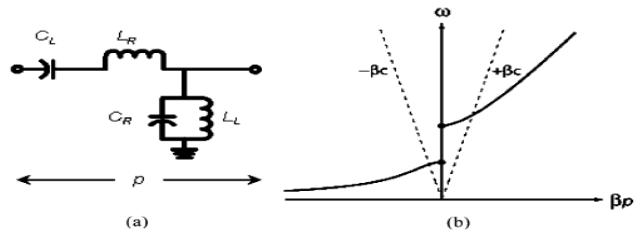


Fig. 1(a)LCunit-cell of length p. Fig. 1(b) Dispersion diagram of CRLH TL

In general, the series resonance and the shunt resonance are not equal and two non-zero frequency points with are present. These two points are referred to as infinite wavelength points and are determined by the series resonance and shunt resonance of the unit-cell as given in (2). By cascading a CRLH TL unit-cell of length, times, a CRLH TL of length can be realized. The CRLH TL can be used as a resonator under the resonance condition

$$\beta_n = \frac{n\pi}{L}$$

where, n is the resonance mode number and can be a positive or negative integer and even zero [6]. In the case where, an infinite wavelength is supported and the resonance condition is independent of the CRLH TL's length (i.e., number of unit-cells, can be arbitrary).

III. Design and Simulation Results

The entire structure was implemented on Rogers RT Duroid 5880(tm) substrate with $\epsilon_R=2.2$ and $h=1.57$ mm. A single feed line is used to excite the antenna. Three elliptical unit cells with major radius 12mm is used. A return loss of -16.6876 dB is experimentally obtained at $f_0= 3.82$ GHz. The antenna size is $\lambda_o/5 \times \lambda_o/5 \times \lambda_o/50$ at f_0 . At f_0 , an infinite wavelength is supported by the CRLH TL and the electric field along the perimeter of the patch formed by the three CRLH unit-cells is a constant. Therefore, the equivalent magnetic current densities along the edges of the patch form a magnetic loop. Therefore, an omnidirectional radiation pattern is expected and is confirmed by the simulated E-radiation pattern at x-y plane and H- radiation pattern at x-z plane of Fig. 4. A maximum gain of 7.7035dB was achieved at f_0 . All simulations were done by using Ansoft HFSS.

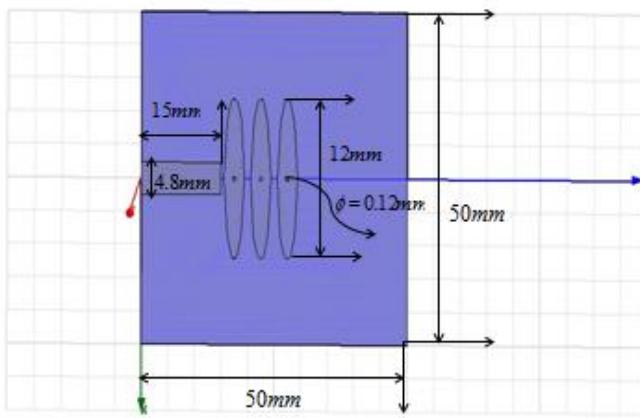


Fig 2. Geometrical model of the antenna with a substrate of size 50mm x 50mm having a permittivity of 2.2, $h=1.57$ mm, gap between successive

unit cell=0.2 mm, gap between the feed and the 1st unit cell=0.1mm, feedline width=4.8mm and via radius=0.12mm.

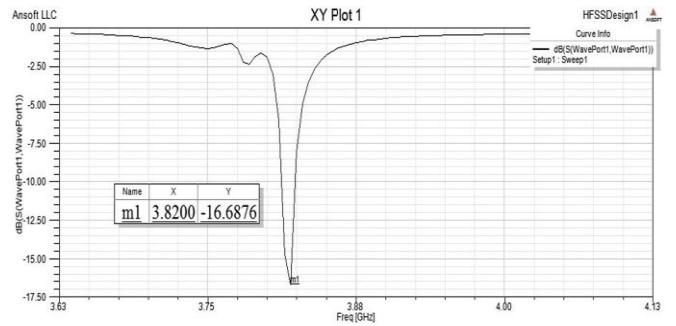


Fig.3 Reflection Coefficient of the proposed antenna at n=0 mode (3.82GHz)

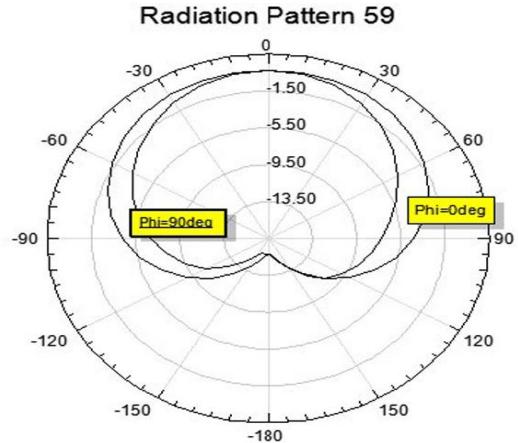
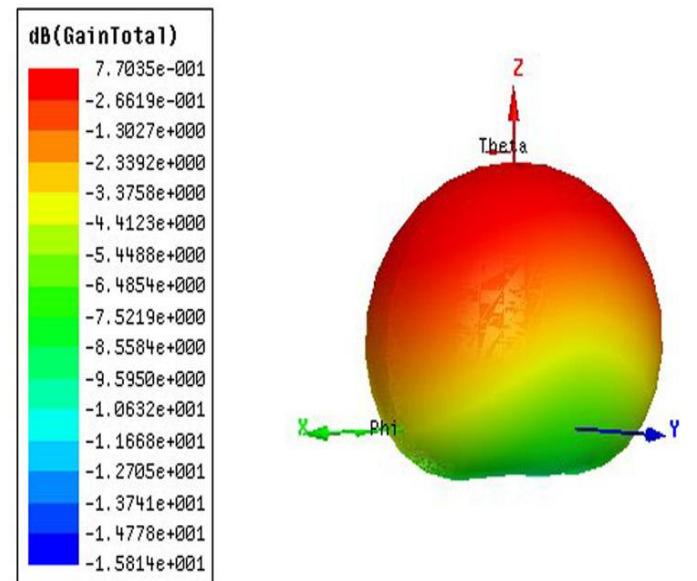


Fig.4 Radiation Pattern of the proposed antenna at n=0 mode (3.82GHz) for Phi=0deg (x-y plane) and 90deg (x-z plane).



IV. Fig.5 Gain of the proposed antenna at n=0 mode (3.82GHz)

IV. CONCLUSION

A new periodic elliptical ZOR antenna is proposed. The infinite wavelength frequency i.e; the resonant frequency is determined by the shunt resonance of the unit-cell. An omnidirectional radiation pattern with a high gain of 7.7035dB and antenna efficiency of 85.505% at the zeroth order mode is obtained. The proposed antenna is especially suitable for point to point wireless communications.

ACKNOWLEDGMENT

The authors would like to thank Prof. D.R.Poddar, Jadavpur University, Kolkata, India for providing simulation facilities at his laboratory to carry out this work.

REFERENCES

- [1] Caloz and T. Itoh, Electromagnetic Metamaterials: Transmission Line Theory and Microwave Applications, John Wiley & Sons, Press, 2006.
- [2] A.Lai, K. M. K. H. Leong, and T. Itoh, "Infinite wavelength resonant antennas with monopolar radiation pattern based on periodic structures," *IEEE Trans. On Antennas and Propagation*, vol.55, no.3, March 2007.
- [3] Lai, C. Caloz, and T. Itoh, "Composite right/left-handed transmission line metamaterials," *IEEE Microwave Mag.*, vol. 5, no. 3, pp. 34–50, Sep. 2004.
- [4] C. Lee, K.M.K.H. Leong, and T. Itoh, "Design of resonant small antenna using composite right/left handed transmission line," *IEEE AP-S/URSI Int. Symp.*, Washington D.C., Jul. 2005.
- [5] C. Caloz and T. Itoh, "Application of the transmission line theory of left-handed (LH) materials to the realization of a microstrip LH transmission line," in *IEEE AP-S Int. Symp.*, 2002, pp. 412–415.
- [6] D. Sievenpiper, L.Zhang, F.J. Broas, N.G.Alexopoulos, and E. Yablonovitch, "High-impedance electromagnetic surface with a forbidden frequency band," *IEEE Trans. Microwave Theory Tech.*, vol. 47,no. 11, pp. 2059-2074, Nov. 1999.

Design and development of E-shaped array antenna using Defected Ground Structure at 5.25 GHz band

¹Chandan Kumar Ghosh ²Biswarup Rana ³Susanta Kumar Parui

^{1,2}Dr.B.C. Roy Engineering College, Durgapur, India

³Bengal Engineering and Science University, Shibpur, India.

¹E-mail: mcet_ckg@yahoo.com³arkap@yahoo.com

Abstract- A DGS integrated E-shaped (2x2) array antenna is presented. It is designed to function in the 5.25 GHz which corresponds to IEEE 802.11a wireless LAN application. The proposed antenna array is a high directive gain, low-cost, low weight base station antenna. The characteristic analysis such as return loss (RL), bandwidth (B.W), and VSWR and radiation pattern of the antenna array have been investigated numerically. In this investigation, VSWR less than 1.2, bandwidth of 240 MHz (For RL>-9.5dB) and antenna gain of 13.45 dBi have been achieved with DGS. The performance of the antenna has been optimized by varying the slot length of the antenna element. Further the improved performances of the array have been achieved by introducing defected ground structure (DGS) structure in between two antenna elements of the array. Numerical study has been carried out by using Zeland make IE3D electromagnetic simulator.

Key words: Antenna array, Antenna gain, Slotted microstrip patch, Wireless LAN, Array optimization, DGS.

I. INTRODUCTION

Over the last two decades the wireless communication system has experienced a significant growth from first generation (1G) analog voice signal to forthcoming forth generation (4G) mobile technology. The motto of 4G communication system is to provide Wi-Fi (Wireless Fidelity) communication network and high quality audio and video services. Today's technology requires high data rate and longer range to provide quality services to the users. For current mobile communication, the diversity scheme has already been implemented to mitigate the fading effects of multipath scenario [1], [2]. In this study, an antenna array has been design to improve the gain, isolation and directivity. This is achieved as a result of the use of multiple antenna elements, exited through single feed point via the transmission line networks.

This article describes the design, analysis and simulation of single element and multi elements patch antenna array for recent wireless communication system which operates at 5.25 GHz band. The antenna is designed, optimized and analyzed with commercially available IE3D electromagnetic simulator. In this investigation, a simple E-shaped patch has been taken for the study and its performance has been improved by using DGS structure [3]-[6]. We focused on

the antenna aspect of array system where four element microstrip patch configuration is proposed. Some good review articles are [7] - [10] where they did not optimize the distance between the antenna elements for reducing the effect of mutual coupling between the elements. But despite that they still give good introduction to the advantages and disadvantages of the system. The work with E-shaped antenna has been organized in three parts (a) Single element E-shaped antenna design and study (b) Multi-element antenna design and study (c). Multi-element E-shaped antenna design and study with DGS. The schematic of E-shaped antenna element is illustrated in the Fig.1 (a) and the array of that element is shown in the Fig.1 (b)

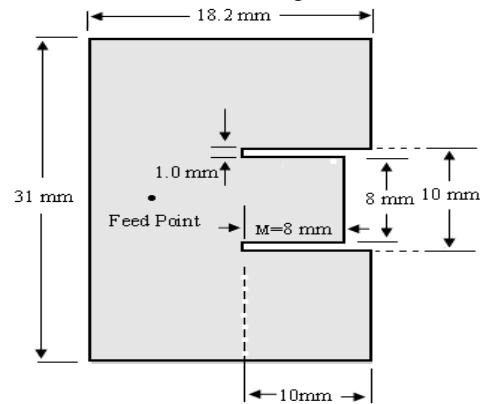


Fig.1 (a).Schematic diagram of E-shaped antenna (Single element).

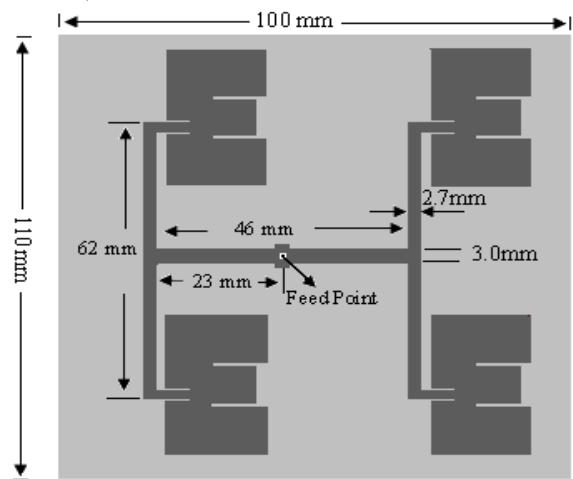


Fig.1 (b).Schematic diagram of E-shaped array antenna.

II. DESIGN PRINCIPLES

Length (L), width (W) and thickness (t) of the antenna have been calculated by using standard formulae [11]. To meet the actual design requirements i.e. operating frequency of 5.25GHz, band width and the radiation efficiency, some approximations have been taken. The calculations are based on transmission line model. The width and length of the microstrip patch have been calculated by the following formula

$$\lambda = c/f \quad (1)$$

$$W = \frac{c}{2f_0 \sqrt{\epsilon_r + 1}} \quad (2)$$

Where the required frequency $f_0=5.25$ GHz, $c=3 \times 10^{10}$ cm/s and $\epsilon_r= 2.2$. Substituting above values, the width of the patch (W) becomes 31 mm. The effective length of the patch can be calculated with the help of equations (3) and (4).

$$L_{\text{eff}} = \frac{c}{2f_0 \sqrt{\epsilon_{\text{ref}}}} \quad (3)$$

$$\epsilon_{\text{ref}} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left(\frac{1}{\sqrt{1 + 12t/W}} \right) \quad (4)$$

In this design, RT/duroid substrate has been used with substrate parameters are $\tan\delta=0.001$, thickness of the substrate, $t=1.588$ mm. Substituting $W=31$ mm, $\epsilon_r= 2.2$ in equation (4) we get $\epsilon_{\text{ref}} = 1.925$. Hence $L_{\text{eff}} = 19.88$ mm. The length extension is to be calculated by (5)

$$\Delta L = 0.412t \frac{(\epsilon_{\text{ref}} + 0.3)(W/t + 0.264)}{(\epsilon_{\text{ref}} - 0.258)(W/t + 0.8)} \quad (5)$$

Substituting $\epsilon_{\text{ref}} = 1.925$ and the values of W and t, we get $\Delta L = 0.823$ mm. Now the actual length (L) of the patch is given by $L = L_{\text{eff}} - 2\Delta L = 19.88$ mm - 1.686 mm = 18.2 mm. Apart from this approximate calculation, the dimension has been slightly adjusted in order to achieve the desired frequency. A number of feed points have been studied in this letter but the minimum RL and maximum radiation efficiency has been achieved at the feed location shown in the Fig.1(a). For the fundamental TM_{10} mode, the patch length should be slightly less than $\lambda/2$, where λ is the wavelength in the dielectric medium. Here, λ is equal to $\lambda_0/\sqrt{\epsilon_0}$, where λ_0 is the free-space wavelength and ϵ_0 is the effective dielectric constant of the patch.

High-low transmission line concept has been used in case of feed network of the array. The impedance of the patch element of the array has been matched with the feed network by using $\lambda/4$ transformer. The detail dimension of the feed network of the array is illustrated in the Fig.1 (b). The simulated result of impedance characteristic of the array antenna is shown in the Fig.2.

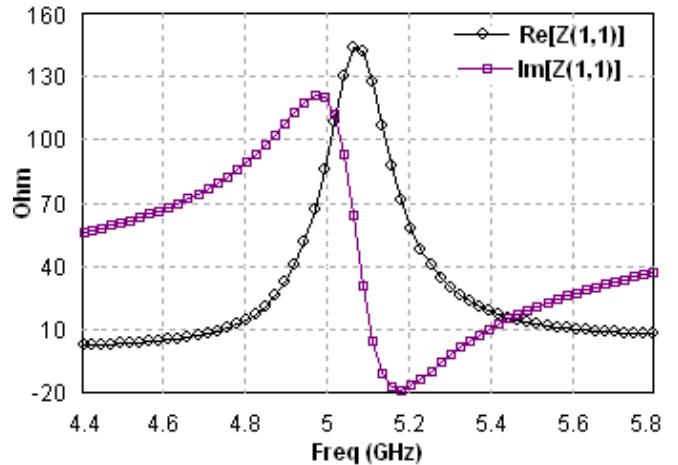


Fig.2. Simulated result of impedance characteristic of the designed antenna element.

The array is fed by a probe of diameter 1.2mm in the middle of the thicker transmission line by using SMA of impedance 50 ohms. The distance between the antenna elements has been optimized and fixed at 3cm.

III. RESULTS

The slot 'M' shown in the Fig.1 (a) is varied and optimized the performances of single element antenna. The simulated RL with different M values are illustrated in the Fig.3. The optimized performance is achieved for M=8.0 mm.

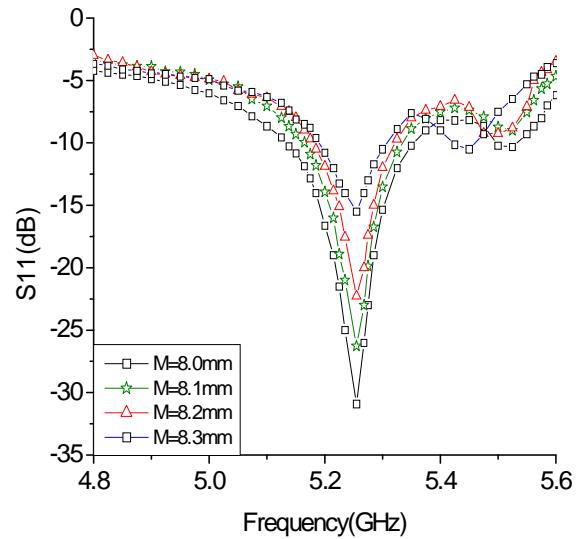


Fig.3. Simulated RL for different M values of the antenna element.

The microstrip patch antenna radiates normal to its patch surface. So the elevation pattern for $\phi=0$ and $\phi=90$ degrees are important for the patch antennas. The simulated E-plane and H-plane pattern of the single element is shown in the Fig.4. The simulated pattern views show a gain of 7.5 dBi for single element and 12.859dBi for the array. Elevation pattern view is shown in the Fig.5 (a) and its 3D pattern view is illustrated in the Fig.5 (b). The corresponding current distribution is depicted in the Fig.6.

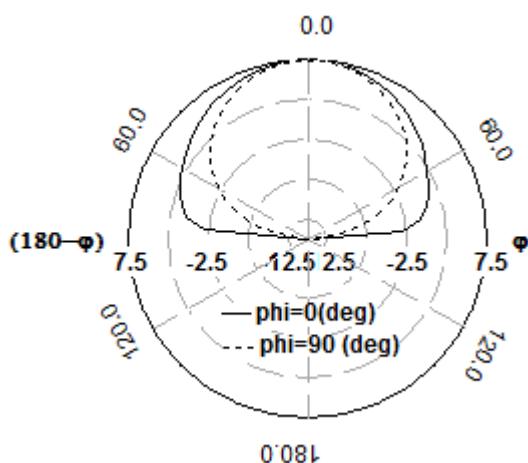


Fig.4. Elevation pattern gain for single element of the array.

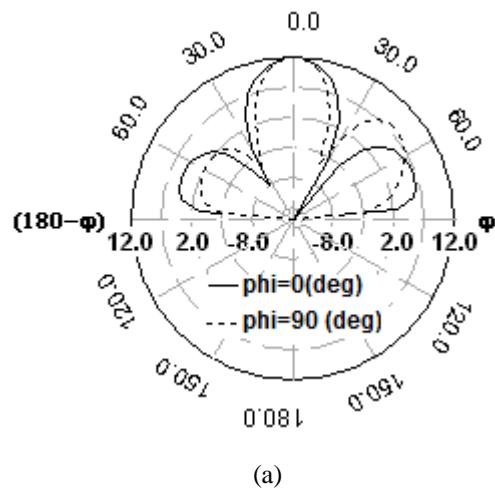


Fig.5(a). Simulated elevation pattern view of the array.

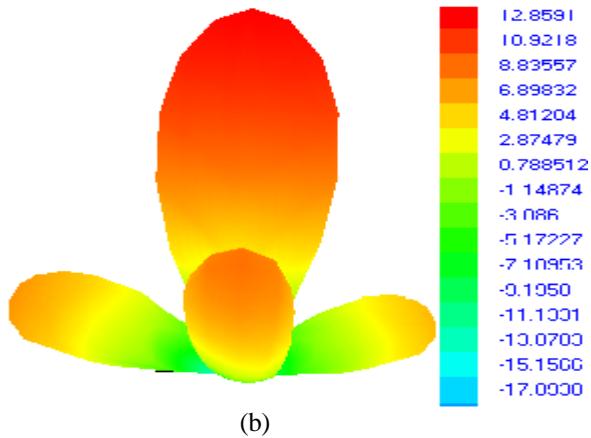


Fig.5(b). Simulated 3D pattern view of the array.

3D pattern view shows a gain of 12.8591 dBi. This gain further been improved by using rectangular shaped DGS in between two patch elements as shown in the Fig.7(a). A defected structure etched off the metallic ground plane of a microstrip antenna effectively disturbs the shield current distribution in the ground plane and thus, introduces inductance and capacitance in the ground plane. It obtains wide stop band and compact size, which meet emerging application challenges in wireless communication. Dumbbell shaped DGS was explored first time by D. Ahn offers the low pass characteristic with one finite transmission zero [12]. Recent advances of this technology

have proven that, these structures are simple solutions to the problems of surface and leaky waves and also to improve antenna performances.

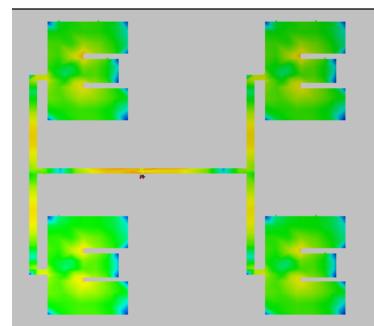


Fig.6. Simulated current distribution of the array.

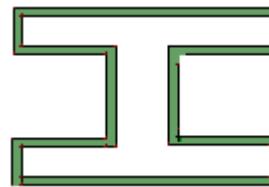


Fig.7(a). Schematic of the DGS (enlarge view).

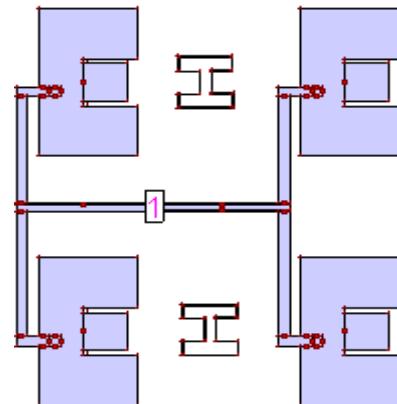


Fig.7(b). Schematic of the DGS integrated array.

Simulated RL of the Array with DGS is shown in the Fig.8

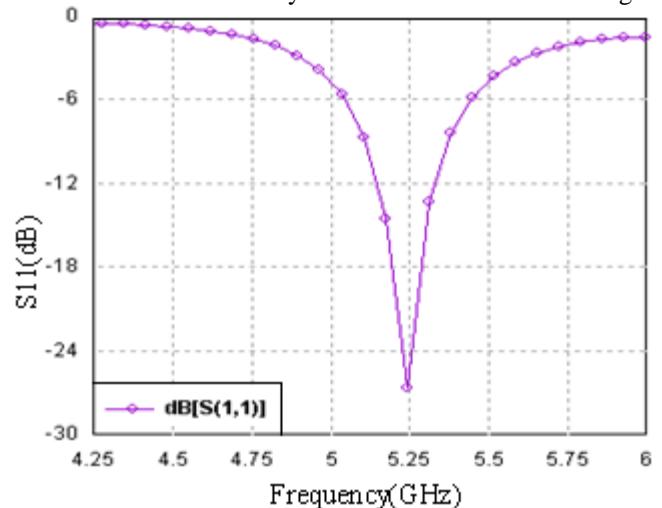


Fig.8. RL of the array with DGS.

The simulated result of radiation pattern with DGS is shown in the Fig. 9. VSWR of the array is shown in the Fig.10.

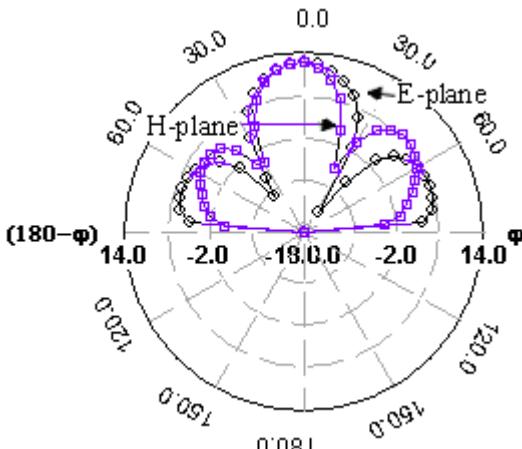


Fig.9.Simulated result of radiation pattern with DGS .

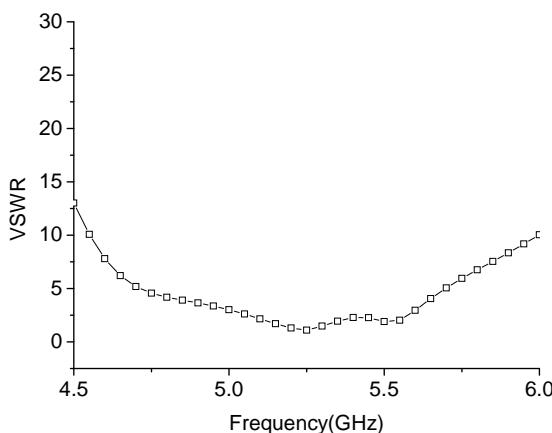


Fig.9. VSWR of the array with DGS.

REFERENCES

- [1] K.Fujimoto and J.R. James, *Mobile antenna system handbook*, 2nd edition, Artech House Inc., 2001
- [2] Gonca CAKIR,Levent SEVGİ, " Design, Simulation and Tests of a Low-cost Microstrip Patch Antenna Arrays for the Wireless Communication", *Turk J Elec. Engin*, VOL. 13. NO.1 2005.
- [3] C. N. Chiu, C. H. Kuo, and M. S. Lin, "Bandpass shielding enclosure design using multipole-slot arrays for modern portable digital devices," *IEEE Trans. Electromagn. Compat.*, vol. 50, no. 4, pp. 895–904, Nov. 2008.
- [4] F. R. Yang, K. P. Ma, Y. Qian, and T. Itoh, "A uniplanar compact hotonic-bandgap (UC-PBG) structure and its applications for microwavecircuits," *IEEE Trans. Microwave Theory Tech.*, vol. 47, no. 8, pp.1509–1514, Aug. 1999.
- [5] N. Engheta and R. W. Ziolkowski, *Metamaterials: Physics and EngineeringExplorations*. Hoboken/Piscataway, NJ: Wiley-IEEE Press,2006.
- [6] R. Chair, C. L. Mak, K. F. Lee, K. M. Luk, and A. Kishk, "Minuature wide-band half U-slot and half E-shaped patch antenna," *IEEE Trans. Antennas Propag.*, vol. 53, no. 8, p. 26452652, Aug. 2005.
- [7] Su, S. W. and J. H. Chou, "Low cost flat metal-plate dipole antenna for 2.4/5-GHz WLAN operation," *Microw. Opt. Tech.Lett.*, Vol. 50, 1686–1687, 2008.
- [8] Shun-Yun Lin and Kuang-Chih Huang, "A compact Microstrip Antenna for GP and DCS Application", *IEEE Trans. AntennasPropag.*, vol. 53, No.3, pp. 1227-1229, March. 2005.
- [9] Ramadan, A., K. Y. Kabalan, A. El-Hajj, S. Khoury, and M. Al-Husseini, \A recon\urable U-koch microstrip antenna for wirelessapplications," *Progress In Electromagnetics Research*, PIER 93,355{367, 2009.
- [10] A new reduced size microstrip patch antenna with fractal shaped defects," *Progress In Electromagnetics Research B*, Vol. 11, 29{37, 2009.*agnetics Research*, PIER 93,355{367, 2009.
- [11] R.A. Sainati CAD of micro strip antenna for wireless applications. Artech House, Inc.1996.
- [12] D. Ahn, J.S.Park, C.S.Kim, J.Kim, Y.Qian and T.Itoh, "A design of the lowpass filter using the novel microstrip defected ground structure," *IEEE Trans. on Microwave Theory and Techniques*, vol. 49, no. 1, pp. 86-93, 2001.

IV. CONCLUSION

The proposed E-shaped antenna array is electrically small and suitable to handle easily. From the simulated results it is observed that the maximum gain obtained in the broadside direction and the peak gain at design frequency for the array without DGS is 12.859 dBi for both $\varphi= 0$ degree and $\varphi= 90$ degree. In this investigation, the B.W of 240MHz and gain of the antenna array with DGS of 13.45 dBi have been achieved. It is clearly observed that the impedance bandwidth, radiation efficiency improved significantly by employing DGS in the array structure. The VSWR of 1.2 for the array with DGS is reported. To obtain high gain and directivity and also to improve the array performances, the DGS has been introduced in the design. However, it is clearly observed that the performances of the array have been improved significantly by the introduction of DGS in the array structure. The proposed antenna array is applicable to WLAN and PDA.

ACKNOWLEDGMENT

The authors like to acknowledge Bengal Engineering and Science University, Shibpur, India for providing necessary experimental support during this research work.

Hybrid Dielectric Resonator antenna with hexagonal patch for UWB applications

Kumari Padmawati⁽¹⁾, D.K.Bisoyi⁽¹⁾, S.K Behera⁽²⁾

Department of physics, N.I.T Rourkela⁽¹⁾

Department Of Electronics and communication⁽²⁾

anshu.padma@gmail.com, dkbisoyi@nitrkl.ac.in, skbehera@nitrkl.ac.in

Abstract- A hybrid dielectric resonator antenna (DRA) is proposed for ultra wideband (UWB) applications. A simple rectangular dielectric resonator is excited by a hexagonal shaped patch connected to a coplanar waveguide (CPW) feeding line. It is found that the bandwidth of the resonant modes can be expanded by using a CPW-fed ground and hexagonal-shaped patch, and thus a UWB performance can be achieved. The simulated result of this DRA shows, this antenna can work in the frequency range of 2.4 to 10.6 GHz, which is suitable for UWB applications. These results also show good performances in terms of VSWR, antenna gain and radiation pattern. With these features, the proposed antennas are expected to be good candidates in various UWB systems.

Keywords: DRA, UWB, Hybrid DRA,CPW feeding

1. INTRODUCTION

Dielectric resonator antennas (DRAs) have been subjected to many investigations since their introduction in 1983[1,2]. Dielectric resonator antennas(DRA) possess some peculiar properties which render the very promising, especially for millimeter wave applications. Their high radiation efficiency, bandwidth and polarization flexibility make them suitable than micro strip patch antennas.[3]

The dielectric resonator antennas consists of high dielectric constant material, high quality factors and mounted on a grounded substrate of lower permittivity. Dielectric resonator antennas are fabricated from low loss and high dielectric constant materials of various shapes whose resonant frequencies are function of size, shape and permittivity of the material. DRA can be in a few geometries including cylindrical, rectangular, spherical, half split cylindrical disk and

hemispherical shaped. In addition, they offer higher bandwidths and gain as compared to micro strip patch antennas [4]. Recently one major aspect of the research with DRAs has been focused on the bandwidth and many techniques have been purposed to enhance their operation bandwidth.

Different types of effective method uses like composite configuration of DRA e.g. stacked DRA another method in which DRA is made of different materials e.g. embedded DRAs. Other types of method in which hybrid feeding structure, such as Peosta et.al have purposed multisegment DRAs. Coulibaly et.al has purposed hybrid fed DRA. Liang et.al has used micro strip line with a patch to excite dielectric resonator. Meanwhile by lifting a DR above the ground plane, its Q factor can be effectively reduced and a broad bandwidth can be achieved [5].

Challenges of the UWB antenna design include the UWB characteristics of the impedance matching, the radiation stability, the compact configuration of the antenna size and the low manufacturing cost for commercial applications. The UWB antennas presented in the literature mainly focus on the slot and monopole antennas. Printed wide slot antennas have an attractive property of providing a wide operating bandwidth, especially for those having a modified tuning stub [6]. such as, the rectangular stub and the circular stub inside the wide slot. Broadband planar monopole antennas have received considerable attention because of their attractive advantages, such as ultra wide frequency band [7], good radiation characteristics, simple structure, low cost and easy fabrication. The typical shapes of these antennas are half-disc, circle, ellipse and rectangle [8].

In addition, despite the approval of the FCC for UWB to operating services that already occupy frequencies in the UWB band such as IEEE 802.11a in the USA (5.15–5.35 GHz, 5.725–5.825 GHz) and HIPERLAN/2 in Europe (5.15–

5.35 GHz, 5.47– 5.725 GHz). Furthermore, in some European and Asian countries[9], world interpretability for microwave access (Wi MAX) service from 3.3 to 3.6GHz also operate in the UWB band. It may be necessary to notch out portions of these bands in order to avoid interference.[9]

II. ANTENNA DESIGN

The proposed DRA consists of three parts: a rectangular dielectric resonator, a hexagonal-shaped patch and a coplanar waveguide with the electromagnetic coupling between the feeding patch and the dielectric resonator, a broad-band impedance bandwidth is achieved. Fig1 shows the configuration of a UWB antenna. The antenna was fabricated on Rogers 5880 substrate of dielectric constant 4.4 width and thickness of $h=0.762\text{mm}$. Fig1 a hexagonal radiator is fed by a CPW transmission line,

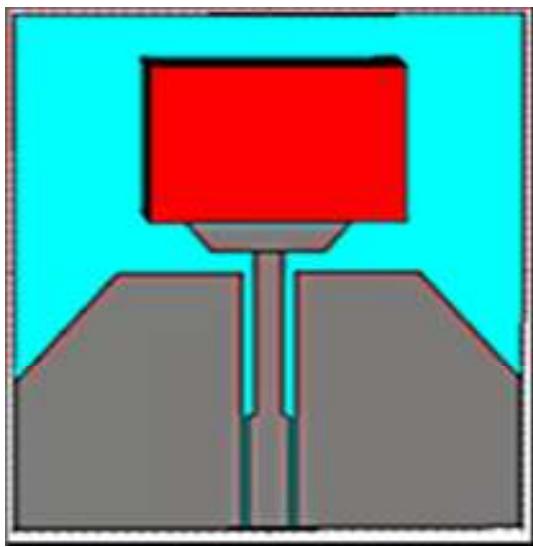


Fig 1: structure of purposed antenna

Since both the antenna and the feeding structure are implemented on the same plane only one layer of substrate with single-sided metal. The feeding patch is hexagonal shaped, which results in a smooth transition from one resonant mode to another to insure good impedance match in a wide frequency range. The parameters of the patch are denoted by $W=8$, $L=12$, four sides of the patch is extruded for getting hexagonal shape. The DR is made with a rectangular substrate with relative permittivity of 10.2 with a height of $HDR=10\text{mm}$ and side length $LDR \times WDR = 15 \times 16$. The ground is depicted by g_1 and g_2 . To design the proposed antenna, the initial dimensions of the rectangular

dielectric resonator were estimated according to the study in these values are: $LDR = 15\text{ mm}$, $WDR=18\text{mm}$ and $HDR = 9\text{mm}$.

However, the CPW fed ground is different from the conventional ground plane to further investigate the properties of the DRA, simulations of the structure are carried out by using CST microwave studio, when the structure is without DR, the antenna works from 5.7 to 10 GHz for reflection coefficient 10 dB.

When it is with DR, the lower band shifts to 2.4 GHz and shows a better impedance matching. It can be seen that the fundamental mode TE obtained at 2.97 GHz. The electrical field of this mode is plotted in by using the hexagonal patch and the CPW-fed mechanism;

it is found that the impedance bandwidth of the antenna increases. The low profile structure of DRA is preferred, so the dimension of the DR especially the height HDR, should be reduced. For this purpose, when optimization is carried out by CST.

The height of DR is reduced by reducing the dielectric constant of the material. The dielectric constant of DR is $\epsilon=10.2$.

It is noted that increasing dielectric height HDR causes the resonant frequency to decrease, and so the lower band shifts slightly to the lower side with increasing HDR.

it is found that the optimal height for the best impedance matching should be $HDR = 9\text{ mm}$. when HDR changes from 2 to 10 mm, the input impedance of the entire band is changed owing to the fundamental and the higher modes.

III.RESULT AND DISCUSSION

The purposed antenna is analyzed using CST Microwave studio 2010. The simulated plot $|S_{11}|$ of v.s frequency is shown in the Fig: 2. By varying the height of DRA the resonant frequency can be changes. The purposed antenna achieves bandwidth from 2.4 GHz to 10.6 GHz, which covers the complete ultra wide band frequency range.

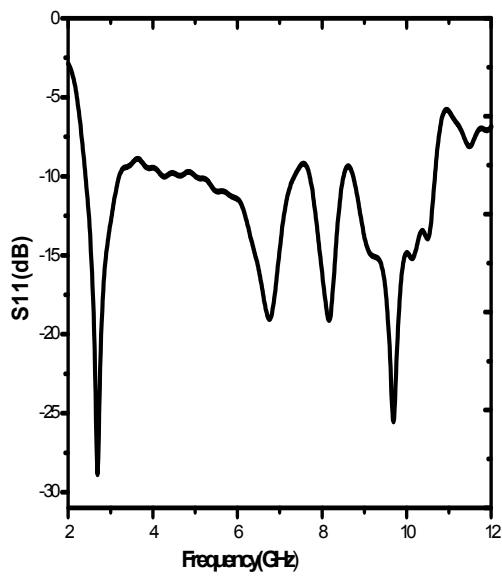


Fig:2 $|S_{11}|$ versus frequency plot

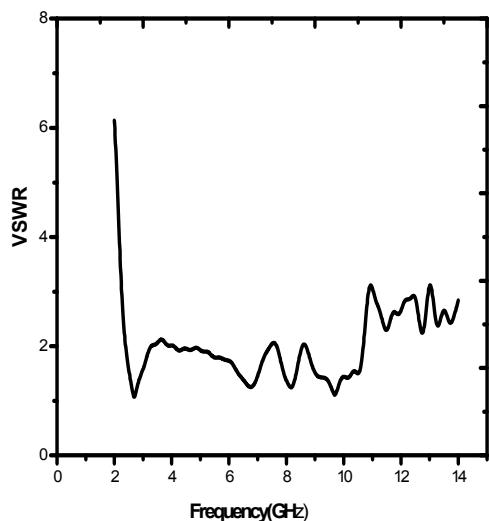


Fig:3 VSWR versus frequency plot of Purposed antenna

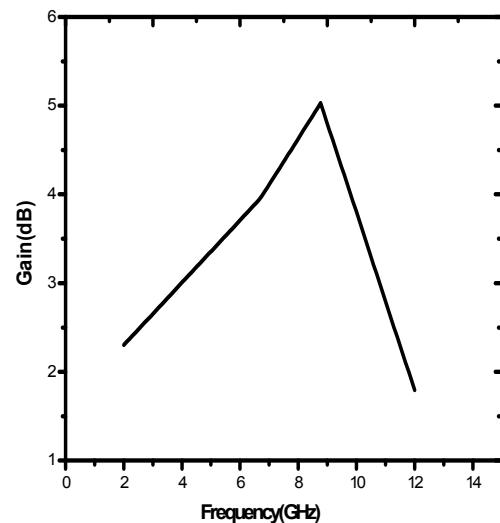
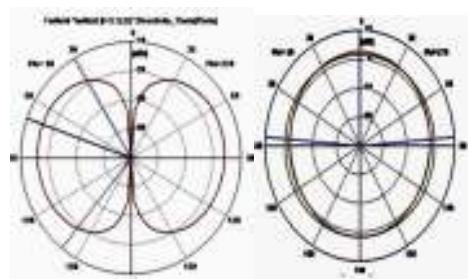


Fig:4 Gain versus frequency plot of antenna

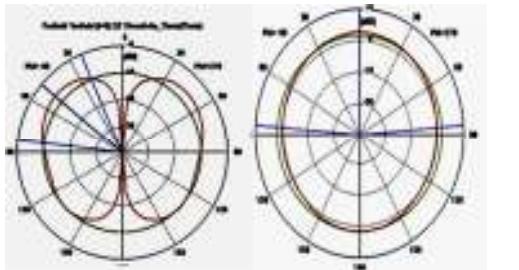
Radiation patterns characteristics

The simulated far field radiation patterns of E-Plane and H-Plane of the purposed antenna are shown in the Fig.5 at different frequencies. The simulated radiation patterns show that the E -Plane radiation pattern is in broadside direction and H- Plane is Omni-directional.

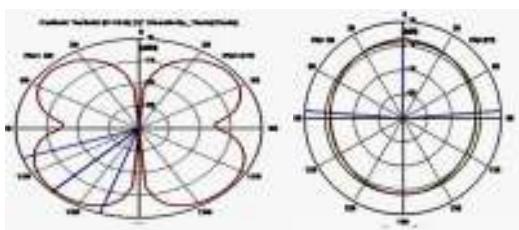


(a)

341-345



(b)



(c)

Fig: 5 simulated radiation patterns of purposed antenna at three frequencies(a) 2.7(b)5.71(c) 9.86GHz

- [4] Liang X L, Dendini T. A, " H shaped dielectric resonator antenna for wide- band applications" IEEE Antennas Wirel. Propag. Lett. 2008, 7, pp 163-166
- [5] Liang X L, Denidini T, " Cross t shape dielectric resonator antenna for wide band applications ",Electronics letters',2008 44, (20), pp 1176-1177
- [7] Almanpis G, Furneaux C, Vahldieck R; "The Trapezoidal dielectric resonator antenna ",IEEE Trans .Antennas Propag. 2008, 5 (9) ,pp.2810-2816
- [8] K. S Rye, A .A Kishk; " UWB Dielectric Resonator antenna with Low cross section polarization" IEEE Transaction 978-1-4244-4726-8/10/2010
- [9] Dendini T.A , Z Weng, " Hybrid UWB Dielectric resonator antenna and band notched designs", IET Microwave,Antennas and propagation, 2011, 4, pp 450- 458

CONCLUSION: In this paper, a novel hybrid DRA has been proposed for UWB applications. The simulated results demonstrate that the proposed DRA can provide an impedance bandwidth more than 3:1 covering the frequency range from 2.4 to 10.6GHz.Furthermore, to minimize the potential interference between the UWB system and the narrowband systems, novel compact CPW-fed UWB DRAs is simulated and discussed, The radiation efficiency of this antenna is approximately 130.6% and maximum gain is 5.453 dB.

REFERENCES:

- [1] Kishk A. A, Yin Y, Glisson A. W, "Conical dielectric resonator antennas for wide-band applications", IEEE Trans Antennas Propag, 2002, 50, 469-474
- [2] Dendini T.A, Rao Q .J, Sebak A. R, "Broadband L Shape dielectric resonator antennas ", IEEE Antennas. Wireless Propag. Lett., 2005 ,4, pp 453- 454
- [3] Coulibaly Y, Denidini T. A, Boutayeb H, "Broadband Micro-strip fed dielectric resonator for X band applications",IEEE Antennas Wirel. Propag .Letters, 2008, 7, pp

Miniaturization of a Rectangular Microstrip Patch Antenna with Slots and Defected Ground Structure

¹Mrinmoy Chakraborty ²Biswarup Rana
^{1,2}Dr.B.C. Roy Engineering College, Durgapur, India
E-mail: ¹chakraborty_wb@yahoo.co.in,²biswaruprana@gmail.com

Abstract- In this paper a rectangular microstrip antenna with slots and defected ground structure is proposed. In the absence of slots and defected ground structure (DGS), the structure found to resonant at 5.22 GHz. When the DGS and slots are introduced, a frequency shift of 5.22 GHz to 3.4 GHz is observed. The main contribution of this paper is the miniaturization of 66% which is very much encouraging

Key words: Compact rectangular microstrip antenna, Defected Ground Structure, Slots, Miniaturization.

I. INTRODUCTION

Microstrip antennas have been used widely in wireless communications due to their light weight, low profile, low cost and ease of fabrication and excellent compatibility with planar integrated circuits and even non planar surfaces. In recent years, as the demand of the small systems have increased, small size antennas at low frequency have drawn much interest from researchers[1]. Many kind of miniaturization techniques, such as using of dielectric substrate of high permittivity[2], slot on the patch, DGS at the ground plane or a combination of them have been proposed and applied to microstrip patch antennas. The other method to miniaturize the microstrip antenna is to modify its geometry using irises [3] or folded structures [4], [5] based on the perturbation effect [6]. In this paper slots and DGS are used to miniaturize the rectangular microstrip antenna. The present work deals with design and analysis of a rectangular compact microstrip antenna for wireless application. The design incorporate dumble shaped defected ground structure which is on the ground plane, which disturbs shielded current distribution in ground plane [7],[8]. Initially the antenna is designed for the resonant frequency of 5.22 GHz and the using DGS and slots the resonant frequency is brought down to 3.4 GHz. So a size reduction of 66% is achieved.

II. DESIGN PRINCIPLES

The geometry of the proposed antenna is shown on Fig. 2. The substrate Rogers RT/duriod 5880(tm) of dielectric constant 2.2 and dielectric loss tangent of 0.0009 has been taken in this design. The antenna has been designed and simulated with Ansoft Designer software. The length and width of the rectangular patch are 22 mm and 16 mm. The feed point is taken at (0,-3.5) in absence of slots and DGS

at the ground plane. For the data as mentioned above the resonant frequency is found to be 5.22 GHz. The dumble shaped DGS as shown in Fig. 3. is incorporated at the groud plane with following specifications D1 = 5 mm, D2=8 mm, L1 = 21.2 mm and L2 =14.7 mm, S1 = 1 mm and S2=2 mm. Two T shaped slots as shown in the Fig. 2. are introduced on the patch with following specification S5=1 mm, L3=0.5 mm, L4=7 mm, W2=9 mm, L5=1.4mm, S3=2mm and L7= 8.4 mm. Another rectangular slot as shown in the Fig. 2 is incorporated with the following specifications S4= 0.5 mm and L6=7 mm. The feed point is taken at (-1.6,-2.6) in presence of slots on the patch and DGS at the ground plane. Now in the presence of slots on the patch and DGS at the ground plane the resonant frequency is found to be 3.4 GHz.

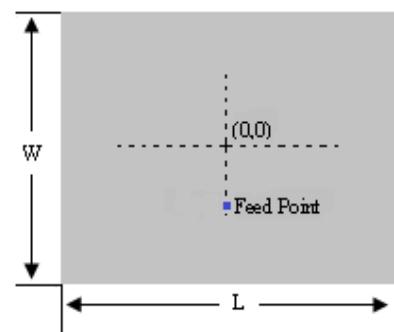


Fig. 1. Rectangular Patch

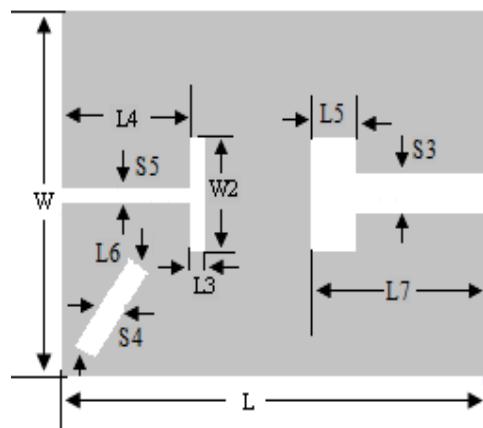


Fig. 2. Rectangular Patch with slots

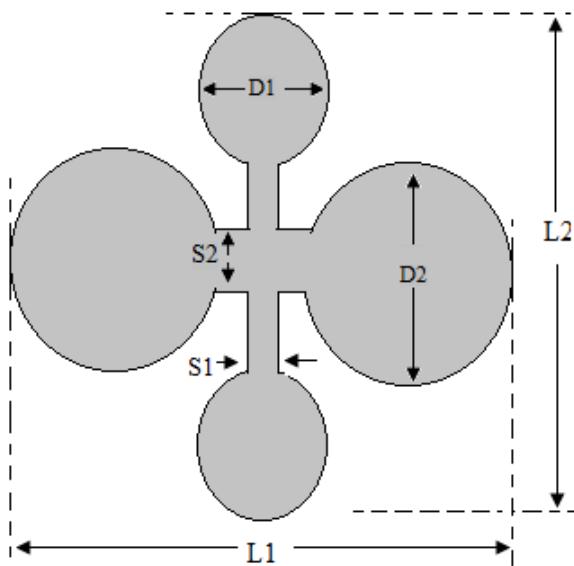


Fig. 3. DGS at the Ground Plane

III. RESULTS

Return Loss of the rectangular patch with DGS and slots and without DGS and slots is given in the Fig. 4. It is observed from Fig. 5. that return loss at 5.22 GHz is -31 dB in absence of slots and DGS at the ground plane. From Fig. 6. it is seen that return loss at 3.4 GHz is -40 dB in presence of slots and DGS at the ground plane. The VSWR without slots and DGS and with slots and DGS are shown in Fig. 7. and Fig. 8. respectively.

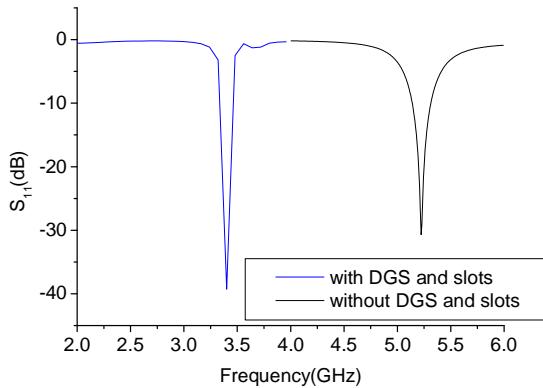


Fig. 4. RL of the Rectangular Patch without DGS and with DGS and Slot

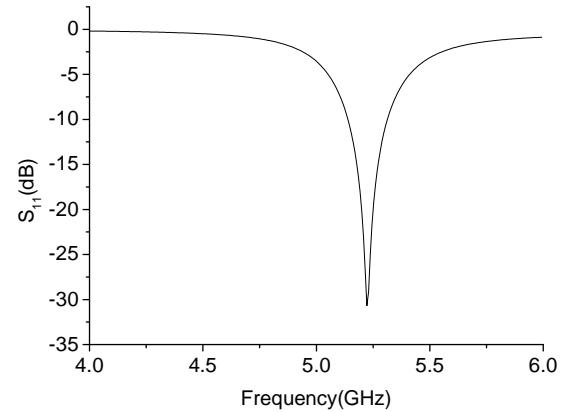


Fig. 5. RL without DGS and slots

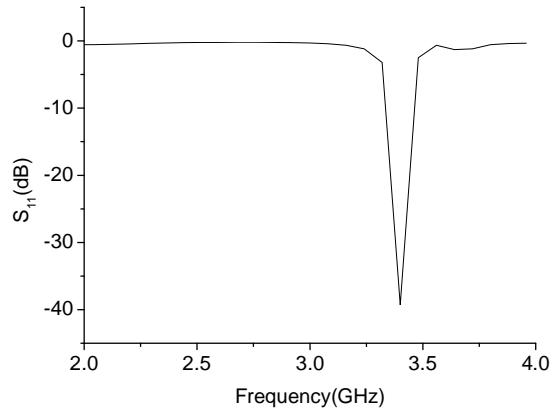


Fig. 6. Return loss with DGS and slots

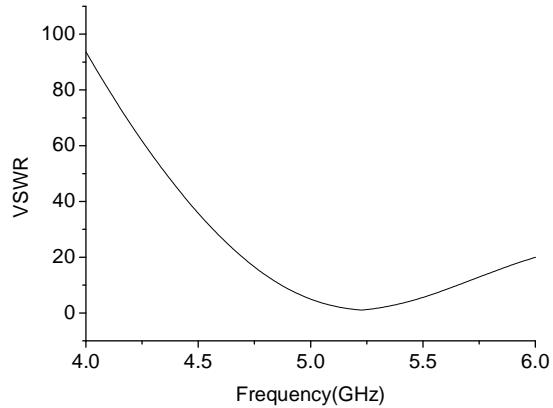


Fig. 7. VSWR without DGS and slots

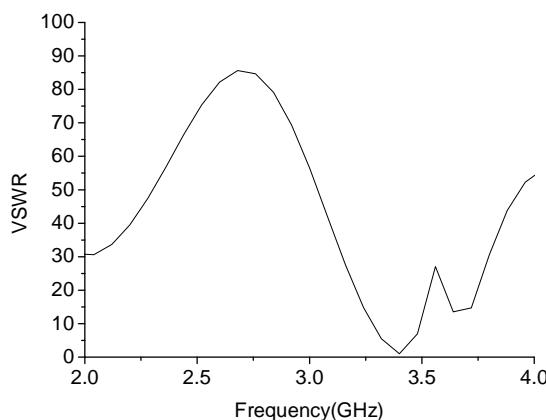


Fig. 8. VSWR with DGS and slots

The microstrip patch antenna radiates normal to its patch surface. So the elevation pattern for $\varphi = 0$ and $\varphi = 90$ degrees are important for the measurement. Fig. 9. below shows the E-plane and H plane radiation pattern at 5.22 GHz. The maximum gain is obtained at resonant frequency for the microstrip antenna without DGS at the ground plane and without slots on the patch plane is 7.2 dBi for both $\varphi = 0$ and $\varphi = 90$ degrees. Fig. 10. below shows the E-plane and H plane radiation pattern at 3.4 GHz. The maximum gain is obtained at resonant frequency for the microstrip antenna with DGS at the ground plane and slots on the patch is 5.3 dBi for both $\varphi = 0$ and $\varphi = 90$ degrees. 3D Input Gain without DGS and slots and with DGS and slots are shown in Fig. 11. and Fig. 12. respectively.

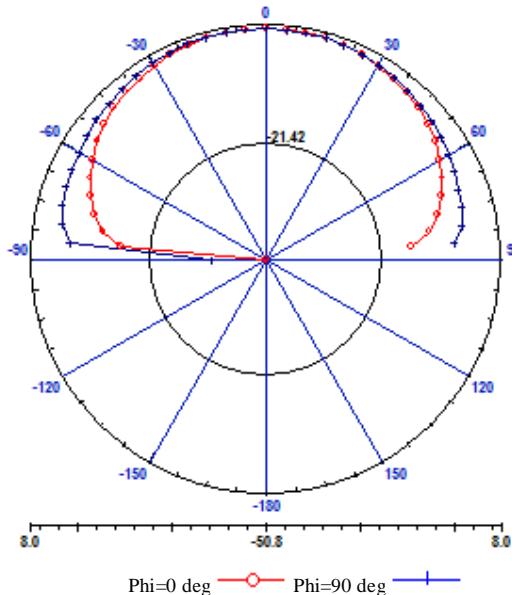


Fig. 9. Radiation Pattern without DGS and slots at 5.22 GHz

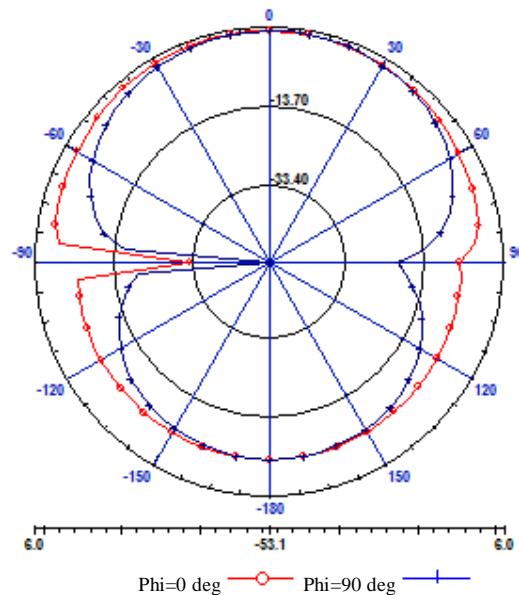


Fig. 10. Radiation Pattern with DGS and slots at 3.4 GHz

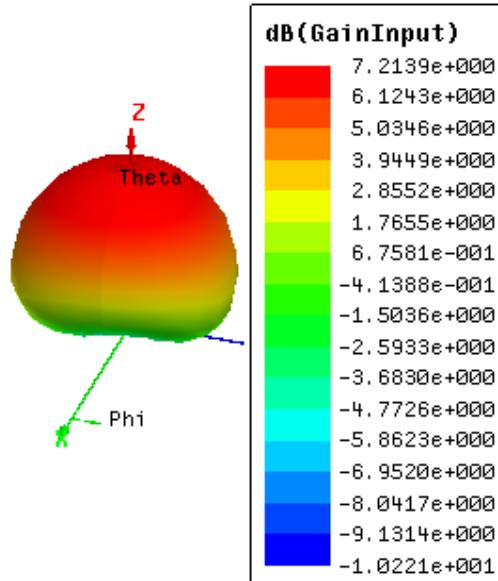


Fig. 11. 3D Input Gain without DGS and slots

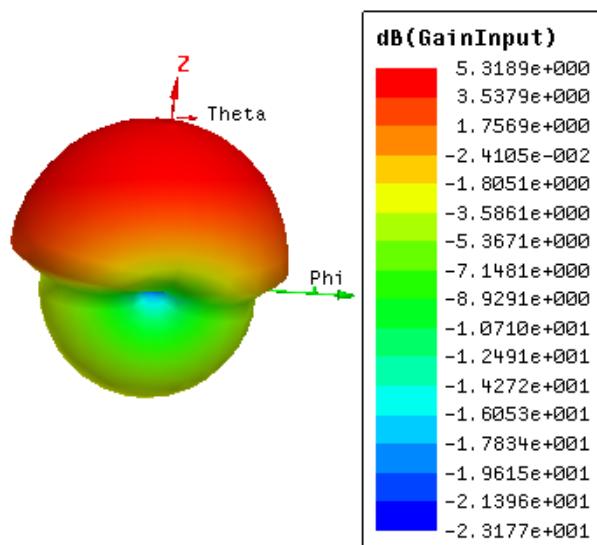


Fig. 12. 3D Input Gain with DGS and slots

IV. CONCLUSION

Design of rectangular microstrip patch antenna with DGS and slots is carried out in this work. A dumbbell shaped DGS in the ground plane found to give a size reduction of about 66% and shift the resonant frequency 5.22 GHz to 3.4 GHz with 60 MHz bandwidth and -40 dB return loss facilitating the antenna to be used for wireless application.

ACKNOWLEDGMENT

The authors like to acknowledge Dr. B.C. Roy Engineering College, Durgapur, W.B., India for providing necessary support during this research work.

REFERENCES

- [1] A.K. Skrivernik, Zurcher O. Staub and J.R. Mosig, "PCS antenna design: The challenge of miniaturization," IEEE Antenna Propagation Magazine, 43 (4), pp 12-27, August 2011M. Young, *The Technical Writers Handbook*. Mill Valley, CA: University Science, 1989.
- [2] T.K. Lo and Y.Hwang, Microstrip antennas of very high permittivity for personal communications, 1997 Asia Pacific Microwave Conference, pp. 253-256
- [3] J.S. Seo and J.M. Woo, "Miniatrization of microstrip antenna using iris," Electron Lett., vol.40 no. 12, pp.718-719, Jun.2004
- [4] K.-L. Wong, Compact and Broadband Microstrip antennas. New York: Wiley-Interscience, 2002, p.5
- [5] H. M. Heo and J. M. Woo, "Miniaturization of microstrip antenna using folded structure," in Proc. Int. Symp. Antennas Propag., Sendai, Japan, Aug. 2004, pp. 985-988.
- [6] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*. Piscataway, NJ: IEEE Press, 2001
- [7] Arya, A.K. Kartikeyan, M.V., Patnaik, A., "Efficiency enhancement of microstrip patch antennas with Defected Ground Structure," IEEE proc. Recent Advanced in Microwave theory and applications (MICROWAVE-08), pp.729-731 November 2008.
- [8] F. Y. Zulkifli, E. T. Rahardjo, and D. Hartanto, "Mutual Coupling Reduction using Dumbbell Defected Ground Structure for Multiband Microstrip antenna array" Progress In Electromagnetics Research Letters, Vol. 13, 29-40, 201

Design of a CPW-Fed Circular Shaped Slot Antenna for Wireless Application

¹Mrinmoy Chakraborty ²Biswarup Rana
^{1,2}Dr.B.C. Roy Engineering College, Durgapur, India
E-mail: ¹chakraborty_wb@yahoo.co.in, ²biswaruprana@gmail.com

Abstract— In this paper a coplanar waveguide (CPW) fed slot antenna as shown in the Fig. 1. is proposed. The physical dimension of the antenna is 60 mm × 80 mm × 2 mm. It is designed at centre frequency of 10.32 GHz with bandwidth of 1.8 GHz. This antenna can be used for wireless applications from 9.6 GHz to 11.4 GHz frequency band.

Key words: Coplanar wave guide, Slots, Microstrip antenna, Wide band.

I. INTRODUCTION

Microstrip antennas are used transmitting and receiving signals. Microstrip or printed antennas are very popular and used widely in wireless communications, mobile communications, radar applications or any other applications because of low profile, small size, light weight, ease of fabrication and excellent compatibility with planar integrated circuits and even non planar surfaces. There are several feeding technique like coaxial probe fed, microstrip line fed, edge fed, inset fed, CPW [1,2] fed. The CPW is the feeding which side-plane conductor is ground and centre strip carries the signal. The advantage of CPW fed slot antenna is its wide band characteristics. Hence CPW fed slot antenna is most effective and promising antenna for wideband wireless application. Wideband characteristic in CPW fed antenna can be achieved by using different slot geometries like bow-tie slot [3], wide rectangular slot [4], circular slot [5], and hexagonal slot [6]. The wide band characteristic in CPW fed antenna can also be achieved by using coupling techniques like inductively and capacitively coupled slots [7], dielectric resonant coupling [8] and other techniques such as using photonic band gap [9]. In this paper a circular shaped slot as shown in Fig. 1. has been taken. The simulation software used for this analysis is HFSS.

II. DESIGN PRINCIPLES

The geometry of the proposed antenna is shown in Fig. 1. The substrate Arlon CuClad 217(tm) of dielectric constant 2.17 and dielectric loss tangent of 0.0009 has been taken in this design. The antenna has been designed and simulated with HFSS software. The details dimension of this antenna is shown in the Tab. 1. The resonant frequency of this CPW-fed antenna is 10.32 GHz with band width of 1.8 GHz. This antenna can be used for wireless applications

from 9.6 GHz to 11.4 GHz frequency band.

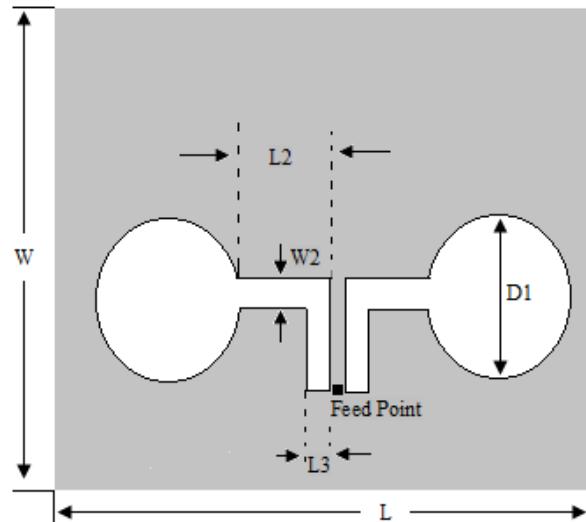


Fig. 1. Rectangular Patch

| Parameters | Value in mm |
|------------|-------------|
| L | 60 |
| W | 80 |
| L3 | 4 |
| W2 | 4 |
| L2 | 2.5 |
| D1 | 34 |

Tab. 1

III. RESULTS

Return loss of the CPW-fed slot antenna is shown in Fig. 2. It is observed from the Fig. 2. that return loss is -16 dB at 10.32 GHz. It is also seen that the band width of this proposed antenna is 1.8 GHz covering the frequency range 9.6 GHz to 11.4 GHz. The VSWR of the proposed antenna is shown in the Fig. 3. It is found from Fig. 3. that VSWR is 1.39 at 10.32 GHz.

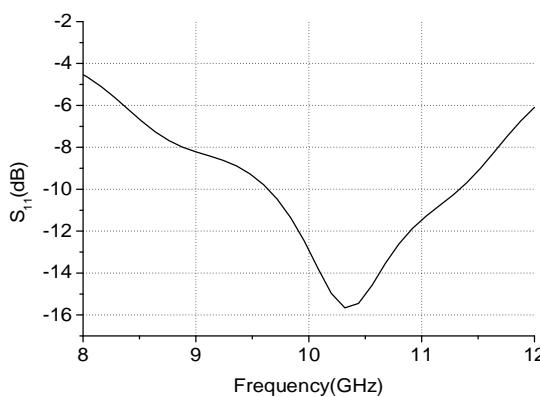


Fig. 2. Return Loss of the antenna

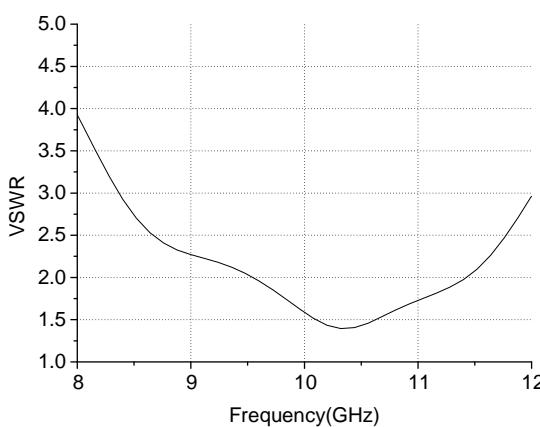


Fig. 3. VSWR of the antenna

The microstrip patch antenna radiates normal to its patch surface. So the elevation pattern for $\phi = 0$ and $\phi = 90$ degrees are important for the measurement. Fig. 4. And Fig. 5. below show the E-plane and H plane radiation pattern at 10.32 GHz.

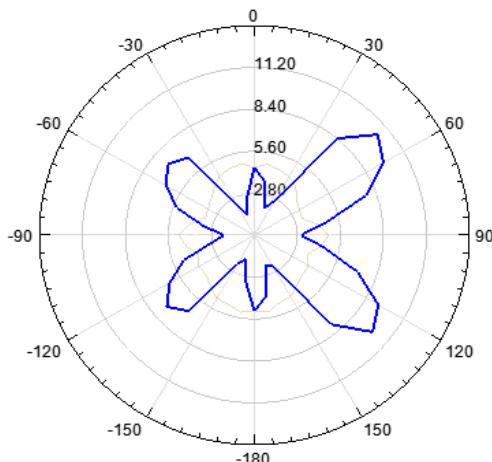


Fig.4. E –Plane radiation pattern at 10.32 GHz

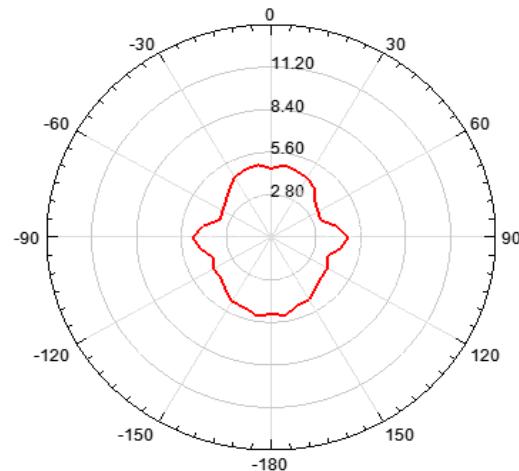


Fig.5. H –Plane radiation pattern at 10.32 GHz

IV. CONCLUSION

A CPW- fed slot antenna is carried out in this work. The CPW- fed antenna gives the bandwidth of 1.8 GHz at the resonant frequency of 10.32 GHz facilitating the antenna to be used for wireless application.

ACKNOWLEDGMENT

The authors like to acknowledge Dr. B.C. Roy Engineering College, Durgapur, W.B., India for providing necessary support during this research work.

REFERENCES

- [1] J.Y. Chiou, J.Y. Sze and K.L. Wong, "A broad-band CPW-fed striploaded square slot antenna," IEEE Trans. Antennas and Propag., Vol.51, No.4,April 2003, pp. 791-721.
- [2] H.D. Chen, "Broadband CPW-fed square slot antenna with a widened tuning stub," IEEE Trans. Antennas and Propag., Vol. 51, No 8, Aug 2003, pp. 1982-1986.
- [3] L. Marantis and Paul Brennan, "A CPW-Fed Bow-Tie Slot Antenna With Tuning Stub," Proc. Of Loughborough Antennas & Propagation Conference, Mar 2008, pp 389-393.
- [4] W. Q. Chen, G. F. Ding et.al, "Design and Simulation of Broadband Unidirectional CPW-Fed Rectangular Slot Antennas,"IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies For Wireless Communications, 2007, pp. 632-635.
- [5] E. A. Soliman, S. Brebels, E. Beyne, and G. A. E. Vandenbosch, "CPW-fed cusp antenna," Microwave Opt. Technol. Lett., Vol. 2, Aug. 1999, pp.288-290.
- [6] Kraisorn Sari-kha, Vech Vivek, and Prayoot Akkaraekthalin , "A Broadband CPW-fed Equilateral Hexagonal Slot Antenna," Proc. Of IEEE International Conference on Computer Systems and Information Technology, Sept. 2006, pp.783-786.
- [7] L.Giauffret, J. Laheurte, and A. Papiernik, "Study of various shapes of the coupling slot in CPW-fed Microstrip antennas," IEEE Trans.Antennas Propag., Vol. 45, April 1997, pp.642-647.
- [8] M.S. Al Salameh, Y.M.M Antar, and G. Seguin, "Coplanar waveguidefed slot coupled rectangular dielectric resonator antenna," IEEE Trans.Antennas Propag., Vol.50, No.10, Oct 2002, pp.1415-1419.
- [9] L.T. Wang, X.C. Lin, and J.S. Sun, "The broadband loop slot antenna with photonic bandgap," Proc. Int. Conference on Antennas and Propagation, Vol.2, 2003, pp.470-472.

Nonlinear analysis using MATLAB and Assembly Language programming to investigate possible cause of breakdown in Power system network

Swarnankur Ghosh

Prof (Dr.) Pradip Kr. Saha

Prof (Dr.) Gautam Kr. Panda

Dept. Of Electrical Engineering, Jalpaiguri Govt. Engineering College, Jalpaiguri- 73510

swarnankur.007@gmail.com

p_ksaha@rediff.com

g_panda@rediffmail.com

Abstract: In this paper possible causes of various instability and chances of system break down in a power system network are investigated based on nonlinear dynamics theory applied to a Power system network. Here a simple three bus power system model is used for the analysis. First the routes to chaotic oscillation through various oscillatory modes are completely determined. Then it is shown that chaotic oscillation eventually leads to system breakdown characterized by collapse of system voltage and high rise in Generator rotor angle (angle divergence), also known as chaos induced instability. It has been shown that chaos and chaos induced instability in Power system take place due to the variation in system parameters. With the variation in system parameter, the system undergoes various changes of state, both qualitatively and quantitatively, which is known as 'BIFURCATION'. Here a bifurcation diagram, showing changes in system state with the variation of reactive power through various kind of bifurcation which eventually leads to chaos induced instability, are thoroughly analyzed. The relation between chaotic oscillation and various system instabilities are discussed here. Using the simple power system model, here it is shown that how chaos leads to voltage collapse and angle divergence, taken place simultaneously when the stability condition of the chaotic oscillation are broken. All nonlinear analysis are implemented using MATLAB/SIMULINK and Turbo C assembly language which are user-friendly, easy to implement and here it is told that when there is a large disturbances in system parameter leading to system collapse, chaos is very likely to be an intermediate transient stage between a stable and unstable region a power system. It also indicated that there is a maximum loadability point after which the system enters into instability modes. All these studies are helpful to understand the mechanism of various instability modes and to find out effective anti-chaos strategies to prevent power system instability.

Keywords-Power system, chaos, bifurcation, Period Doubling, Voltage collapse, Angle Divergence, MATLAB/SIMULINK, Turbo C.

1. INTRODUCTION

A power system is inherently of nonlinear nature i.e. the power system dynamics is described by a set of nonlinear equations obtained from system modeling and parameters. For our practical purposes we generally intended for analyzing power system stability problem using linearized model to avoid complexity without imposing serious error. But the continuing interconnections of bulk power systems, have led to an increasingly complex system that must operate close to the limits of stability which affects the qualitative behavior of power system network. This operating environment has contributed to the growing importance of the problems associated with the dynamic stability assessment of power systems. To a large extent, this is also due to the fact that most of the major power system breakdowns are caused by problems related to the system dynamic responses. It is believed that new types of instability emerge as the system approaches the stability limit which can not be explain completely or accurately using linear analysis. So, nonlinear system theory is used to find out the explanation of that instability. Study of power system dynamics and stability through nonlinear approach is now become an important part of power system stability analysis and draws attention of many researchers. Any nonlinear dynamical system undergoes changes in state either qualitatively or quantitatively with the changes of one or more system parameters, known as **Bifurcation**. Being an inherently nonlinear system, power system also shows bifurcation with the gradual changes in parameter (like variation in load). Sometimes variation in parameter may result in complicated behavior which is random and indeterministic. This stable, bounded, random and aperiodic behavior is known as chaos. This chaotic oscillation may leads to various instability. Most common types of system instabilities, which occur when the system is heavily loaded, are voltage collapse and angle divergence. These events are characterized by a slow variation in the system operating point, due to increase in loads, in such a way that ultimately load voltage magnitudes sharply decreased to a very low value and generator rotor angular displacement rises to a very high value which eventually causes system breakdown. Voltage collapse in electric power systems has recently received significant attention by researchers. A number of physical mechanisms have been identified which possibly leading to voltage collapse. In several papers [4-10] voltage collapse is viewed as an instability which coincides with the disappearance

of the steady state operating point as a system parameter, such as a reactive power demand is quasistatically varied. In the language of bifurcation theory, these papers link voltage collapse to a fold or saddle node bifurcation of the nominal equilibrium point. Dobson and Chiang [1] first studied and analyzed voltage collapse, and concluded that this phenomenon occurs at a hypothesized static bifurcation of equilibrium points taking place as system loading is increased. The static bifurcation mechanism for voltage collapse postulated in [1] was investigated in [2] and in [3]. In the voltage collapse scenario, the disturbance may be viewed as a slow change in a system parameter, such as a power demand. The steady-state varies continuously with the changing system parameter until it disappears at a saddle node bifurcation point. It is therefore not surprising that saddle node bifurcation is being studied as a possible route to voltage collapse [4].

Another possibility is that steady state operating point loses stability before the saddle node bifurcation. If this occurs in a given system, stability of the nominal equilibrium point may be lost prior to static/saddle node bifurcation point through a Hopf bifurcation. Hopf bifurcation results in appearance of a family of small amplitude periodic solution which eventually makes system unstable. Study of Hopf bifurcation in power system was done in details in [2, 3, 12, and 13]. Except these, other type's bifurcation occurs in power system like Torus bifurcation [14], cyclic fold bifurcation, period doubling bifurcation [3, 8].

In this paper a complete and detail bifurcation analysis has been done which shows different behavior changes (Bifurcation) with the slow and gradual variation of load reactive power. It has been shown here how the stable oscillatory behavior of the power system model tends to chaotic instability through period doubling bifurcation. Period doubling bifurcation is the most important route to chaos in power system model which is analyzed here with great emphasis. Except continuation method, a detail and expanded picture of PDB has been developed which clearly shows the internal behavior changes of the proposed system which eventually leads to chaos. This paper, for the first time, proposes that Voltage collapse and angle divergence phenomenon which make the system unstable, occurs simultaneously. Also this paper gives an indication on maximum loadability point after which system tends towards instability.

In this paper a suitable power system model is set up from [2] to analyze the nonlinearity. The basic non linear equations of the proposed model of power system and load are implemented using the **MATLAB/SIMULINK** environment [15] and for the first time, a separate programme is generated using assembly language programming package (**Turbo C**) to develop bifurcation diagram. This two software are more users friendly and easy to implement compare to the compact program package for nonlinear study such as **AUTO** [16] which was used for non linear analysis of Power system in previous works. The goal of this work is to show the richness of the qualitative behaviors, which may occur near voltage collapse, and to illustrate their effect on system trajectories.

2. Brief review of Nonlinear Theory [17,18]

A typical nonlinear system with state x can often be expressed as

$$\dot{x} = f(x, \mu); x \in \mathbb{R}^n, \mu \in \mathbb{R}^p \quad (2.1)$$

The corresponding properties of such a system are:

- The solution of (2.1) is called trajectory. With initial condition $x(t_0) = x_0$, the solution is given by $x(t) = \lambda_t(x_0)$.
- Four steady state behaviors are associated with the nonlinear system- Equilibrium points, Periodic solutions, quasi periodic solution and chaos.
- The equilibrium points mean the solutions of the nonlinear equation $f(x, \lambda) = 0$. Alternatively an equilibrium point is a degenerate trajectory which stays in the equilibrium point for all time. It is asymptotically stable if all the Eigen values of its corresponding Jacobian matrix have negative real part. Power system is generally operated on a stable equilibrium point. $\Lambda_t(x^*)$ is a periodic solution if $\lambda_t(x^*) = \lambda_{t+T}(x^*)$ for all t and some minimal period $T > 0$. It represents a limit cycle which is a self sustained and bounded oscillation, and, is stable or unstable depending upon its characteristics multiplier. A quasi-periodic solution is one that can be expressed as a sum of periodic function.
- Finally, **chaos** is a random, indeterministic phenomenon exhibits stable, bounded but aperiodic behavior. While equilibrium points are zero dimensional and periodic solutions are one-dimensional, chaos is more complex and having fractional dimension.

2.1 Bifurcation Theory

Bifurcation Theory is used to interpret the way in which qualitative changes occur in the system as one or more parameters are varied. A power system is modeled in the typical form of a nonlinear dynamic system with state x :

$$\dot{x} = f(x, \mu), x \in \mathbb{R}^n, \mu \in \mathbb{R}^p \quad (2.1.1)$$

μ represents the vector of the system parameters that can be varied during the analysis. **Bifurcations** in power system mean “qualitative” changes of the asymptotic behavior of the system trajectory (2.1.1) which is obtained by varying the μ components. At a value of $\mu = \mu_c$ the vector field f loses its structural stability, is called the **Bifurcation Point** and μ_c , the Bifurcation value. This simply means that the Phase portraits for $\mu < \mu_c$ and for $\mu > \mu_c$ are different. In power system's nonlinear model there are following types of bifurcations taken place depending upon the Jacobian (J) of the system (2.1.1)-

- **Saddle-node bifurcation (SNB):-** A saddle-node bifurcation is a local bifurcation in which two fixed point (or equilibrium) of a dynamical system collide and annihilate each other, at this point the Jacobian has a zero eigenvalue and no other eigenvalue with zero real part. A saddle node bifurcation denotes the change in the no. of equilibrium points. This can happen when a stable equilibrium merges with a unstable one and disappears or when two unstable equilibrium points merges

- Hopf bifurcation point (HB):-** A Hopf-bifurcation is a local bifurcation in which a fixed point of a dynamic system loses stability as a pair of the complex conjugate Eigen values of the linearization around the fixed point cross the imaginary axis of the complex plane. If J has a pair of complex conjugate Eigen values on the imaginary axis where all other Eigen values are off the imaginary axis, then Hopf Bifurcation results the emergence of a family of periodic solution in the vicinity of μ_c . If the periodic solution is unstable then it is called “**subcritical**” and “**supercritical**” when stable.

- Period Doubling Bifurcation (PDB):-** During this kind bifurcation with the change in multiplier a new periodic solution or orbit emerges from the previous/existing solution with periodicity approx. twice that of the previous one. If there is a sequence of such bifurcation which accumulate at a critical value $\mu = \mu_c$, then period almost becomes infinite, which means we get an aperiodic but bounded solution to the system which is called “Chaos”. Actually this is one special type of **Hopf Bifurcation**.

3. Power system model for nonlinear analysis

For applying nonlinear theory in Power system network, I consider a simple power system model, first proposed in [1]. This model is widely used for nonlinear behavior study of power system [2, 3, 4, and 14]. In this power system model generator is represented by classical model Here we get a 4-D system model i.e. the system is represented by a set of 4 ordinary non linear differential equation. In this paper, classical model is implemented for a three BUS power system network as follows.

3.1 Mathematical model of a 3-bus electric power system

A simple 3-BUS power system shown in **fig.1** and its equivalent circuit in **fig.2**, this model consists of an infinite bus on the left, a load bus on the center and a generator bus on the right. $Y_0 \angle (-\theta_0 - \pi/2)$ and $Y_m \angle (-\theta_m - \pi/2)$ are the admittances of the transmission lines. One of the generator buses treated as slack bus and the other is described by swing equation. The concept of an infinite/slack bus refers to a particular node of the system with enough capacity to absorb any mismatch in the power balance equations. Thus, it can be considered as fictitious generator with constant voltage magnitude E_0 and phase δ_0 (usually $E_0 = 1$ and $\delta_0 = 0$). On the other hand, the generator has constant voltage magnitude E_m but the angle δ_m varies according to the so-called swing equation.

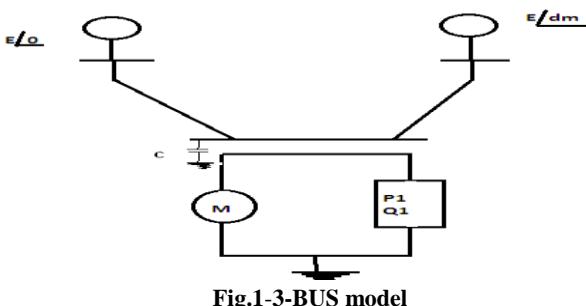


Fig.1-3-BUS model

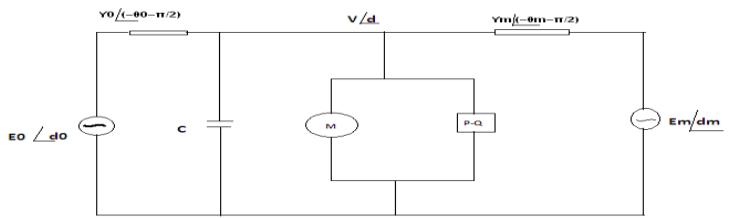


Fig.2: Equivalent circuit

So, 4-D set of dynamic equation is developed for this model as

$$\dot{\delta}m = 0 \quad (3.1.1)$$

$$M \dot{\omega} = -d_m \omega + E_m Y_m \sin(\theta_m) + E_m V Y_m \sin(\delta - \delta_m + \theta_m) \quad (3.1.2)$$

$$k_{qw} \dot{\delta} = -k_{qv} V^2 - k_{qv} V - Q_0 - Q_1 + Q(\delta_m, \delta, V) \quad (3.1.3)$$

$$Tk_{qw} k_{qv} 2 \dot{V} = k_{pw} k_{qv} V^2 + (k_{pw} k_{qv} - k_{qw} k_{pv}) V + k_{qw} [P(\delta_m, \delta, V) - P_0 - P_1] - k_{pw} [Q(\delta_m, \delta, V) - Q_0 - Q_1] \quad (3.1.4)$$

Where, δ_m = Generator rotor angle = Generator load angle, ω = Angular frequency, V = Magnitude of Generator load voltage. The load bus, with voltage magnitude V and phase δ , consists of an induction motor, a generic load $P-Q$ and a capacitor C . The dynamics of this part is derived from a power balance at the bus. Considering an empirical model for the induction motor [28] and a static load $P-Q$, the active and reactive power supplied to the load is

$$P(\delta_m, \delta, V) = P_0 + k_{pw} d \delta + k_{pv} (V + TdV) + P_1 \quad (3.1.5)$$

$$Q(\delta_m, \delta, V) = Q_0 + k_{qw} d \delta + k_{qv} V + k_{qv} V^2 + Q_1 \quad (3.1.6)$$

Where T , k_{pw} , k_{pv} , k_{qw} , k_{qv} and k_{qv2} are constants of the motor, P_0 , Q_0 and P_1 , Q_1 are the static active and reactive power drained by the motor and by the load $P-Q$, respectively.

Thus, the state vector of this model is $[\delta_m, \delta, \omega, V]^T$. Here in this model Q_1 is chosen to bifurcation parameter so that increasing Q_1 corresponds to increase in load reactive power. Therefore the model has the form

$$\dot{x} = f(x, \lambda)$$

Where $x = [\delta_m, \omega, \delta, V]^T$ is the state vector and $\lambda = [Q_1, P_1]^T$ is the parameter vector. The numerical solution which is obtained below, of the set of non-linear equations (3.1.1-3.1.4) completely describes nonlinear behavior of the power system model.

For analyzing the above mentioned classical model (eqns.3.1.1-3.1.4) for nonlinear analysis following set of Data is being used-

Table.1: Power system parameter values used in the simulation

| K _{pw} | K _{pv} | K _{qw} | K _{qv} | K _{qv2} | T | P ₀ | Q ₀ | Y ₀ | Θ ₀ |
|-----------------|-----------------|-----------------|-----------------|------------------|----------------|----------------|----------------|----------------|----------------|
| 0.4 | 0.3 | -0.03 | -2.8 | 2.1 | 8.5 | 0.6 | 1.3 | 20.0 | -5.0 |
| P ₁ | Y ₀ | Θ ₀ | E ₀ | Y _m | Θ _m | E _m | P _m | M | D |
| 0.0 | 8.0 | -12.0 | 2.5 | 5.0 | -5.0 | 1.0 | 1.0 | 0.3 | 0.05 |

So, Eqns. 3.1.1 – 3.1.6 completely specifies the nonlinear behavior of the proposed power system model.

4. Numerical Solution-Simulation and Result

Here the numerical analysis of the nonlinear Eqn. (3.1.1-3.1.4) has been carried out using **MATLAB** and **Turbo C** software package, the result of which can be sub divided into following two parts- 1) Bifurcation and Chaos, 2) Chaos induced instability and system collapse.

4.1 Bifurcation and chaos

In this paper continuous Bifurcation diagram is plotted using **MATLAB** continuation software’ **MATCONT**’ to indicate

different types of Bifurcation mentioned in **2.1** along with the respective bifurcation values in the proposed Power system model described by Eqns. **(3.1.1-3.1.6)**.

Fig.3 shows the complete bifurcation diagram using the continuation method with detail picture of various Bifurcations. Suppose the power system model describe in section **3.1** is operating at a stable equilibrium point, is a function of Q1 i.e. f (Q1), where Q1 is the reactive power demand at the load bus which acts as the main Bifurcation parameter here. **Fig.3** shows the variation of system voltage **V (p.u)** with **Q1**. Now Q1 is slowly increased while other parameters remain fixed. At this moment V is varied as Q1 is increased. Beginning from the left in Fig.3, there are two equilibrium points, one stable (where the system may operate) and the other unstable. The stable one becomes unstable at a supercritical Hopf bifurcation (S1) for **Q1=10.868**.With further increase of Q1 period doubling bifurcation occurs which result unstable limit cycle (ULC).With further increase of Q1 system enters into a series of period doubling bifurcation which eventually leads to Chaos. It can be termed as **Left side chaos**. After that it regains its stability when it enters the supercritical region between PD1 and S2. At PD1 if the value is decreased, again it undergoes another period doubling bifurcation which eventually leads to chaos, which can be termed as **Right hand side chaos**. Finally system oscillation is vanished at S2.After a short stable region beyond S2 system become unstable at S3 where both equilibrium approach each other and coalesce in a saddle-node bifurcation (SNB). If the reactive power of the load Q1 is increased beyond this value, the system does not have an operating point, and voltage collapse occurs.

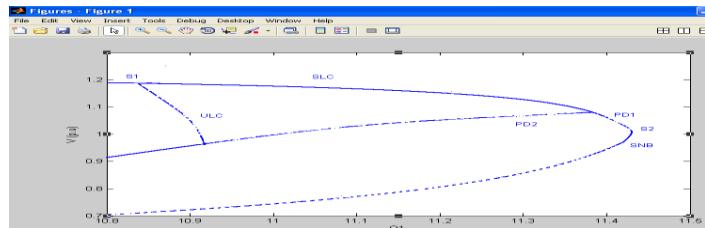


Fig.3: Bifurcation Diagram using continuous showing various zone of nonlinearity

Here the proposed system enters into Chaotic region both from Left side and Right side through a series of period doubling bifurcation as shown -**period-2 cycle** \rightarrow **period-4 cycle** \rightarrow **period-2^N cycle** \rightarrow **CHAOS**

During PDB the system which is operated at a stable equilibrium point or stable limit cycle of period 1,first enters into stable periodic oscillation of period 2,then period 4 and so on. Finally reaches to a state which is completely indeterministic and having infinite and fractional no of period or become aperiodic but exhibits a bounded oscillation. Dynamic behavior of the power system model at this region is very complex and unpredictable and there is no obvious relation between cause and effect. This region is **Chaotic region** Fig.4 below shows an elaborated and expanded bifurcation diagram relating variation of load voltage (V) in p.u with the variation of reactive power (Q1) showing PDB eventually leads to chaos (Right side chaos).This diagram is developed using **Turbo C** assembly language programming Package and **Origin 6**

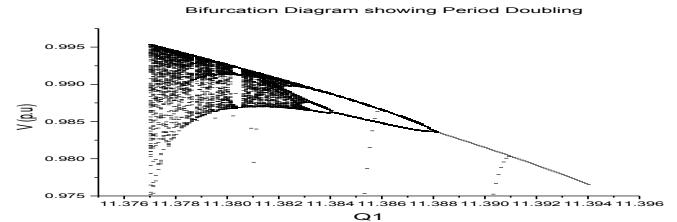


Fig.4: Bifurcation Diagram showing cascaded period doubling bifurcation (V vs. Q1)

Below Table 1 shows the values Q1 (bifurcation parameter) at different bifurcation point-

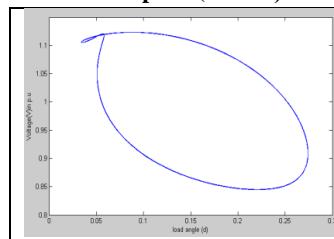
| Value of Q1 | Nature of Bifurcation |
|-------------|-------------------------------------|
| 10.946 | Subcritical Hopf Bifurcation (S1) |
| 11.407 | Supercritical Hopf Bifurcation (S2) |
| 10.870 | Period-1 LHS |
| 10.880 | Period-2 LHS |
| 10.883 | Period-4 LHS |
| 11.392 | Period 1RHS |
| 11.388 | Period 2 RHS |
| 11.384 | Period 4 RHS |

Phase Plots and Time plots:-Below the various phase plots are shown which describes the various states of the system with different values of Q1, during PDB leading to chaos for both **Left side PDB** and **Right side PDB**.Also the time plots are shown for different values of Q1.These plots are obtained from the numerical solution of the Eqn. **(3.1.1-3.1.4)** which are numerically integrated with initial condition **(0.3, 1.5, 0.2, 0.97)** and **(0.315, 0.150, 0.150)** using **Runge-Kutta method**. Here Chaos is observed for **Q1=10.894**, called **Left side chaos** and **Q1=11.383**, called **Right side chaos**.

To observe **Left side PDB** leading to **Left side chaos** phase plots relating **load voltage V (p.u)** and **load angle δ** and Time plots relating **V** and **time** are shown in **Fig.5**.

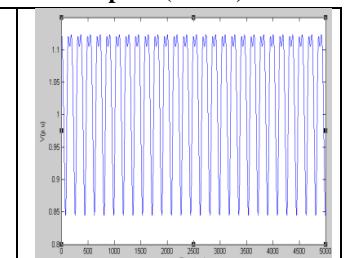
Fig.5- Phase plots (V vs. δ) and Time plots (V vs. t) showing Left side PDB leading to chaos

Phase plots (V vs. δ)

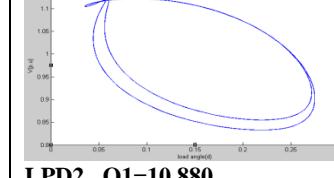


LPD1 Q1=10.870

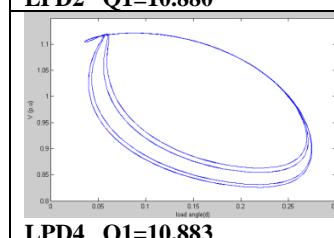
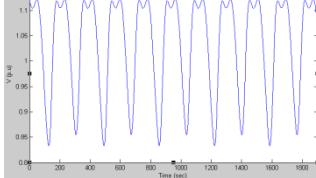
Time plots (V vs. t)



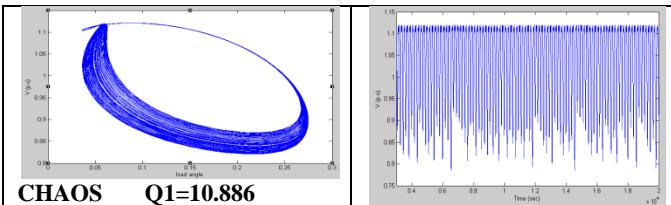
LPD2 Q1=10.880



LPD4 Q1=10.883

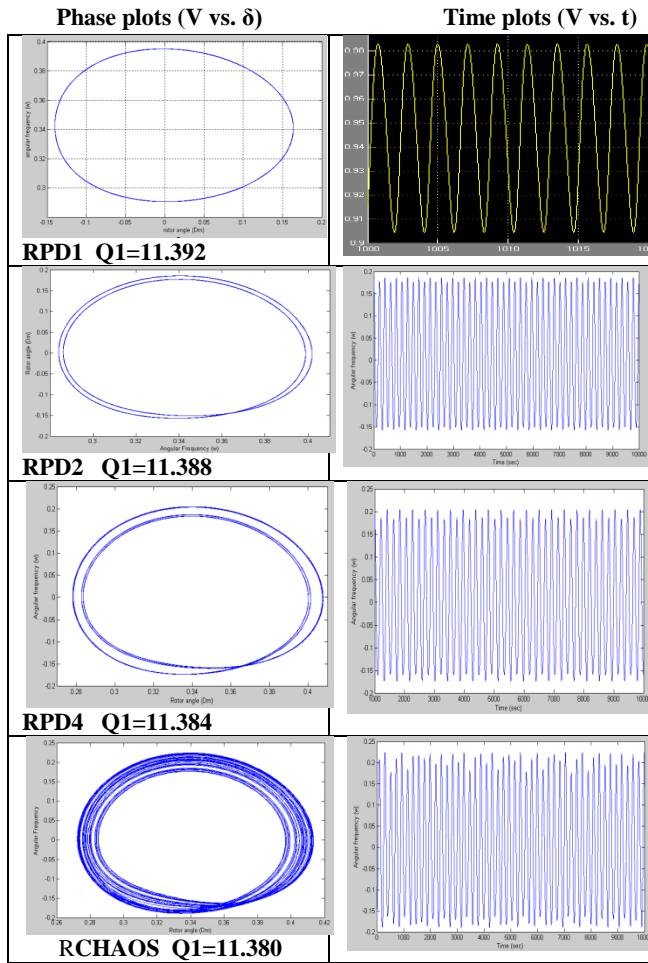


LPD4 Q1=10.883



To observe **Right side PDB** leading to **Right side chaos**, phase plots relating **Angular speed, (ω)** and **Generator load angle (δ_m)** and Time plots relating ω and **time** are shown in Fig.6.

Fig.6-Phase plots and Time plots showing Left side PDB leading to chaos



Where LPD stands for **Left side period doubling** and RPD stands for **Right side Period doubling** and 1, 2, 4 indicate periodicity.

4.2:- Chaos induced instability and system collapse

Though chaos exhibits aperiodic and random oscillatory behavior and completely indeterministic, it is a bounded and marginally stable region. But Chaos is very sensitive to initial condition and system parameter variation. Any small change to them can break their stable oscillation. Here we discuss what happens after stable chaotic oscillation is broken in power systems. It will be shown that chaos can lead to Voltage collapse and angle divergence simultaneously when value of Q1 is increased beyond the value corresponds to Left side chaos and decreased below the value corresponds to Right side chaos, which makes a stable system into complete unstable and the system breakdown takes place.

- **Voltage collapse and Angle Divergence:**-Many studies have observed this phenomenon [1-11] taking place individually.

But in this paper, voltage collapse occurring along with Angle divergence simultaneously is reported for the first time in Power system dynamic stability study. During **Voltage collapse** the system voltage sharply decline to a very low value and possibly brings the blackouts and **Angle Divergence** is the phenomenon when the generator loosing synchronism i.e. rotor angle difference is more than 2π . Here critical point at which voltage collapse and angle divergence phenomena take place simultaneously is **Q1=10.890** after **Left side chaos** and **Q1=11.377** after **Right side chaos**, using same model, parameter values and same initial condition. These are shown below in Fig.8 and 9 where time plot of Voltage and Generator angle are given for the above mentioned values of Q1. For both the values of Q1 we get almost same diagram. From the figure it is seen that voltage collapse and angle divergence appears after chaos is broken.

Fig.8: Voltage collapse with Q1=10.890 and Q1=11.377

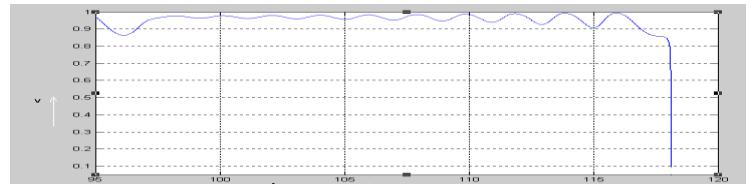
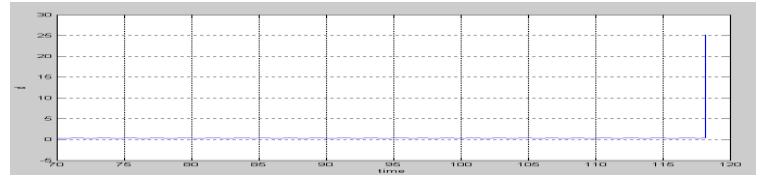


Fig.9. Angle divergence with Q1=10.890 and Q1=11.377



5. CONCLUSIONS

In this paper various nonlinearity of proposed power system model has been deeply studied using both "Matlab/Simulink" and Assembly language Programming. Here cascaded period doubling bifurcation which is one of the most important routes to chaos and system instability has been observed in details. All previous studies used continuation method and compact continuation software's such as **AUTO** for Bifurcation studies in power system network which predicted the occurrence of Period Doubling Bifurcation leading to chaotic phenomenon but were not able to show explicitly and elaborately, the structure of period doubling bifurcation leading to chaos. This limitation is overcome in this paper where bifurcation analysis is done using both MATLAB based continuation software MATCONT to get a continuation Bifurcation diagram and also get a expanded and elaborate diagram of Period doubling bifurcation which clearly shows various state changes of the proposed system with the variation of Q1 which eventually leads to chaos. This is also described by various phase plots and time plots for different values of Q1. The relationship between chaos and major instability modes in power system, such as Voltage collapse and Angle Divergence, has thoroughly been observed. Here it is shown for the first time that Chaos can induce voltage collapse and angle divergence simultaneously when stable condition is broken. So it can be conclude that chaos is an intermediate stage of the instability incident when changes in system parameter causes system breakdown .As chaos is very sensitive to initial condition and system parameters, any variation of them can make chaos to be

annihilated and breaks in into instability. In a real power system all system parameters are fluctuating with changing operating condition and disturbances. So chaos possibly exists in power system network as in prior stage of instability. When disturbance happens, power system comes into a transient stage. If the disturbance is small, HB may happen and stable oscillatory behavior follows. If the disturbance is prolonged, system may come into chaos. And, when the disturbance becomes larger, the chaos may be broken. Voltage collapse, angle instability or voltage collapse and angle divergence simultaneously may happen. If the disturbance is very large, system may directly come into the above three instability conditions over the stages of HB, chaos and chaos breaking. All this studies are helpful in understanding, how various instabilities due to the interaction between systems itself take place and the possible routes through which the system moving towards breakdown through various stable and unstable oscillatory modes and to find appropriate measure to prevent nonlinearity induced instability in power system. Also this paper gives an easy way to explain nonlinear phenomena by using user friendly and easy to implement software tools like MATLAB/SIMULINK and Turbo C.

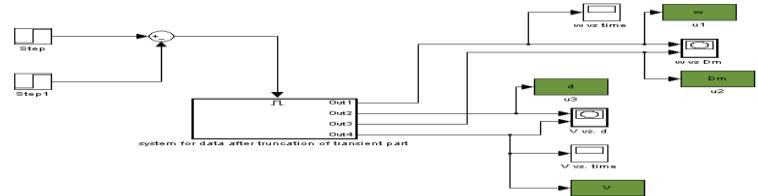
REFERENCES

- [1] I. Dobson and H.-D. Chiang, "Towards a theory of voltage collapse in electric power systems", Systems and Control Letters, Vol.13, 1989, pp. 253-262.
- [2] E. H. Abed, et al., "Dynamic bifurcations in a power system model exhibiting voltage collapse," Technical Research Report, University of Maryland, pp. 1-15, TR92-26, 25-Feb, 1992.
- [3] H. O. Wang, et al., "Bifurcations, Chaos and Crises in power system Voltage Collapse," Technical Research Report, University of Maryland, pp. 1-23, TR92-72, 25-Feb, 1992.
- [4] E. H. Abed, et al., "On bifurcations in power system models and voltage collapse," in Proc. 29th Conf. Decision and Control, pp. 3014-3015, 1990
- [5] M. M. Begovic and A. G. Phadke, "Analysis of voltage collapse by simulation," Znt. Symp. Circuits and Systems May 1989.
- [6] K. T. Vu and C. C. Liu, "Dynamic mechanisms of voltage collapse", Systems and Control Letters, vol. 15, pp. 329-338, 1990.
- [7] T. Van Cutsem, "A method to compute reactive power margins with respect to voltage collapse," LYEE Trans. Power Systems, vol. 6, pp. 145-156, Feb. 1991.
- [8] H.-D. Chiang, C.-W. Liu, P.P- Varaiya, F.F. Wu and M.G. Lauby, "Chaos in a simple power system", Paper No.92 WM 151-1 PWRS, IEEE Winter Power Meeting, 1992.
- [9] K. Walve, "Modeling of Power System Components at Severe Disturbance", CIGRE paper 38-18, International Conference on Large High Voltage Electrical Systems, 1986.
- [10] K.T. Vu and C.C. Liu "Dynamic mechanisms of voltage collapse," Systems and Control letters, Vol. 15, 1990, pp. 329-338.
- [11] Dobson and H.-D. Chiang, "A model of voltage collapse in electric power systems", IEEE Proceedings of 27th Conference on Control and Decision, Dec.1988, Austin, TX, pp. 2104-2109.
- [12] E. H. Abed and P. P. Varaiya, "Nonlinear Oscillations in power systems," Int. J. Elec. Power and Energy Sysr. vol. 6, pp. 3743, 198
- [13] C. Rajagopalan, P. W.S aw, and M.A. Pai, "Analysis of voltage control systems exhibiting Hopf bifurcation," in Proc 28th IEEE Conf. Dec.Contr.Tampa, pp., pp., pp. 332-335, 1989.
- [14] Y Yu,H Jia,P Li and J Su,"Power system instability chaos",14th PSC,IEEE conf.Sevilla,24-28 June,2002

- [15] Matlab WorksInc. Simulink (Dynamic System Simulation for MATLAB)"
- [16] E.J. Doedel, "AUTO: A program for the automatic bifurcation analysis of autonomous systems", Cong. Num., Vol.30, pp. 265-284, 1981.
- [17] NON-LINEAR DYNAMICAL SYSTEMS, Alexander Panfilov
- [18] Nonlinear Ordinary Differential Equations-An introduction for Scientists and Engineers, D.W. Jordan and P. Smith

APPENDIX:-

MATLAB/SIMULINK model



Swarnankur Ghosh receives his B.Tech in Electrical Engineering from Techno India, Salt Lake under West Bengal University of Technology (WBUT) and currently pursuing M.Tech (Final Year) in Power Electronics and Drives at Jalpaiguri Govt. Engineering College. His research interests include Power electronics, Nonlinear Dynamics and CHAOS in power system, FACTS Devices

Dr. Pradip Kr. Saha received his B.E degree in Electrical Engineering from BE College, Shibpur and M.Tech degree in Machine Drives and Power Electronics from IIT Kharagpur and Ph.D. from NBU. He is currently a professor and Head in Electrical Engineering Dept. at Jalpaiguri Govt. Engineering College. His research interests include Power Electronics, Machine Drives, and CHAOS in Power Electronics.

Dr. Gautam Kr. Panda received his B.E degree in Electrical Engineering from Jalpaiguri Govt. Eng. College and M.E.E. degree in Machine Drives and Power Electronics Jadavpur University and Ph.D. from NBU. He is currently a professor in Electrical Engineering Dept. at Jalpaiguri Govt. Engineering College. His research interests include Electrical Machines & Drives, Nonlinear Dynamics and CHAOS.

2nd Annual International Conference IEMCON 2012

A 3-D Box Parity Model for Error Detection and Correction in Digital Communication Systems

A. Lahiri

Electronics and Communication Dept.

Institute of Engineering and Management

Salt Lake, Sector V

Kolkata, India

E-mail: golu.iem@gmail.com

M. Chakraborty

Electronics and Communication Dept.

Institute of Engineering and Management

Salt Lake, Sector V

Kolkata, India

Abstract-- In this paper the drawbacks of the two dimensional parity check method are highlighted and a novel model is proposed for circumnavigating the stated problems. In the 3D box model, instead of restricting our parity generation plane along two dimensions, a three dimensional box is constructed, which allows us to perform parity checking in three dimensions. The model proposed, requires additional bits compared to that of the two dimensional parity check method, but the error detection and correction capacity also increases at the cost of increasing the number of redundant parity bits.

Keywords-- coding theory; 1-D parity checking; 2-D parity checking; 3-D parity checking, digital communication;

I. INTRODUCTION

Since Shannon's seminal papers of 1948 and 1949 [1, 2], coding theory has progressed at a notable rate [3-11]. Coding theory is a branch of mathematics concerned with transmitting data across a noisy channel and recovering the message successfully. Digital communication is the trend of the present era. With the advent of digitized networks like ISDN, digitalization of data starts from the transmitter premises only [12-16]. We assume that our message is in the form of binary digits or bits of 0 or 1. We have to transmit these bits along a noisy channel, in which error occurs randomly, but at a predictable overall rate [17-19]. To compensate for the error in channel, we need to transmit extra bits[20-23] than those present in original signal. The simplest method for detecting errors in binary data is the parity code which transmits an extra "parity" bit after every 7 bits from the source. However in this

process, only error detection is possible. A simple way to correct error in this scheme is to repeat each bit a preset number of times. The recipient sees which value (0 or 1) has been transmitted maximum times for a particular bit and takes decision accordingly. Then the idea of 2-D parity check emerged where the incoming bits are arranged in matrix form and parity bits are created for each row and each column of the matrix. Any error in the pattern is reflected in the parity bits generated, which pin points the exact coordinate of the error. But whenever there is an even number of errors along a particular column or row, then the method fails. In this paper we propose a model to overcome this difficulty and thus provide better detection facility.

II. THE TWO DIMENSIONAL MATRIX METHOD

The two dimensional matrix method of parity checking is performed as shown below in fig.1.

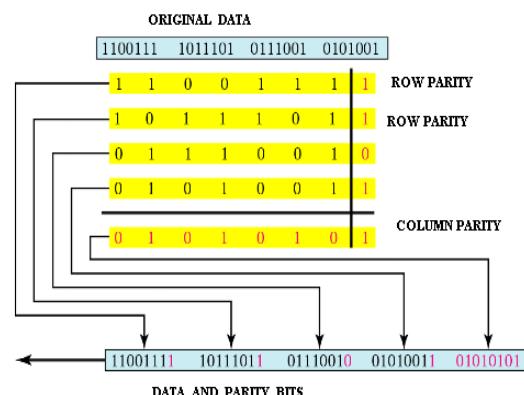


Fig. 1 Schematic representation of two dimensional parity check

A condition is considered when the fourth element of third row is corrupted by noise and it is changed to level 0. So the parity of third row is affected which denotes, that there is error in the third row. Also, the parity of fourth column is affected which isolates the fourth element of third row to be the corrupted bit and thus it is complemented for error correction. But problem lies in the fact that, if two elements of the same column are affected, we can identify only the corrupted rows but not the corrupted column. The above two points are illustrated in table (1).

TABLE I
PARITY BIT'S GENERATION FOR CORRUPTED
STREAM OF BITS

| R1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
|-----|----|----|----|----------|----|----|----------|---|
| R2 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 |
| R3 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| R4 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| C.P | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| C1 | C2 | C3 | C4 | C5 | C6 | C7 | R.P | |

C.P=column parity, R.P=row parity, red cells=corrupted bits

In the above figure, we have Original data: 1100110 1011101 0111001 0101000 Received data: 1100111 1011101 0110001 101001 The red cells denote the corrupted bits. Here though the rows R1 and R4 detect error, but correction cannot be done as C7 remains unaltered. On the other hand as both R3 and C4 is changed the error of (R3, C4) is both detected and corrected.

III. THE 3-D BOX MODEL

A. Axis Selection

Before the 3-D box can be characterized, the axis must be defined which will be helpful in addressing the various message bits later on. The selection of axis is shown in fig.2.

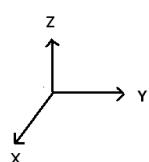


Fig. 2 Selection of Axis

B. Layout of the 3-D structure

An incoming bit pattern is considered as: 10111 1100 1100 0011. Now, the first plane will be constructed taking the most significant eight bits and the second plane with the least significant eight bits. The formation of the two planes is shown in fig.3.

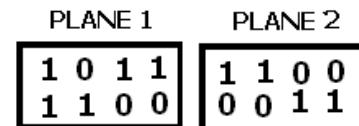


Fig. 3 Schematic representation of formation of two planes with incoming message bits

After formation of the two planes, they are placed on top of each other, analogous to opposite faces of a cube as shown in fig.4.

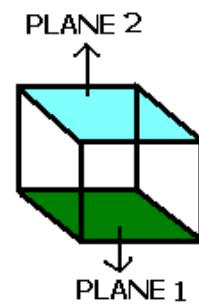


Fig. 4 Figure showing arrangement of the two planes in 3-D space

For proper identification of bits in three dimensional space, we give each of the bits an unique coordinate as shown in fig.5.

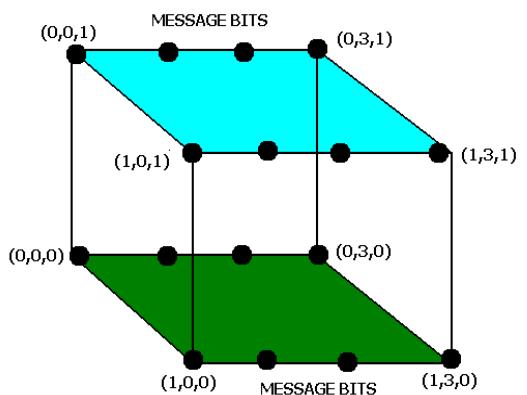


Fig.5 Representation of message bits as a function of three dimensional coordinates

C. Parity Bits Generation Scheme

After creation of the 3-D box model, we begin computation of the parity bits of the above model. Similar to the 2-D matrix method, the row and column parity bits are generated for both of the planes. Lets us take the example of Plane 1. There are two parity bits for the two rows consisting of [(0,0,0), (0,1,0), (0,2,0) and (0,3,0)] and for [(1,0,0), (1,1,0), (1,2,0) and (1,3,0)]. There are four parity bits for the four columns consisting of [(0,0,0) and (1,0,0)], [(0,1,0) and (1,1,0)], [(0,2,0) and (1,2,0)] and finally [(0,3,0) and (1,3,0)]. Similar functions are done for Plane 2 also. The idea has been depicted in fig.6. Till now, the parity generation was confined in the X-Y plane. Now, we start generating parity bits with respect to bits aligned along the Z axis. Along the Z axis parity bits are generated for the message bits for which only Z coordinate is different and both X,Y coordinates are same e.g. [(0,0,0) and (0,0,1)], [(0,1,0) and (0,1,1)] and so on. There are eight such possible cases. In general, parity bits along Z axis are created taking the two elements at a time [(i,j,0) and (i,j,1)] where $i \in [0,1]$ and $j \in [0,3]$ as shown in Table II.

D. Layout of Parity Bits

| PLANE 1 | | | | PLANE 2 | | | |
|---------|---|---|---|---------|---|---|---|
| 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| (a) | | | | (b) | | | |

Fig. 6 Figure showing method of generation of parity bits for rows and columns for each plane. Row and column parity bits are shown outside each rectangle (a) shows for Plane 1 and (b) shows for Plane 2. Transmitted data is: 1011 1100 1100 0011.

TABLE II
PARITY BITS GENERATION ALONG Z AXIS

| | | | | | | | |
|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| (0,0, 0) and (0,0, 1) | (0,1, 0) and (0,1, 1) | (0,2, 0) and (0,2, 1) | (0,3, 0) and (0,3, 1) | (1,0, 0) and (1,0, 1) | (1,1, 0) and (1,1, 1) | (1,2, 0) and (1,2, 1) | (1,3, 0) and (1,3, 1) |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

From fig. (6) and Table II, it is evident how the parity bits are generated for a given message bit pattern.

E. Superiority Over 2-D Matrix Method

The cases where the 2-D model fails are now being considered. The data transmitted is taken as 1011 1100 1100 0011. The corrupted received pattern is taken at first to be 1010 1100 1101 0011. The 2-D model for the corrupted data stream is shown first in table (3)

TABLE III.
BIT LAYOUT IN 2-D MODEL

| R1 | 1 | 0 | 1 | 0 | 0 |
|-----|----|----|----|----|-----|
| R2 | 1 | 1 | 0 | 0 | 0 |
| R3 | 1 | 1 | 0 | 1 | 1 |
| R4 | 0 | 0 | 1 | 1 | 0 |
| C.P | 1 | 0 | 1 | 0 | 0 |
| | C1 | C2 | C3 | C4 | R.P |

Data transmitted: 1011 1100 1100 0011

Data received: 1010 1100 1101 0011

It was shown earlier that due to two errors along same column, the error cannot be detected. But the situation changes in the 3-D model. The R1 of Table III is actually R1 of Plane1 of 3-D box. Similarly R2 is also R2 of Plane1. But now R3 of Table III becomes R1 of Plane 2 while R4 of table III becomes R2 of Plane 2. Now in Plane1, there is error in R1 but not in R2. As a result C4 of Plane1 is also affected. So the error coordinate is identified and corrected. Similar is the case while dealing with the second error bit located on Plane2 of 3-D model.

Second situation is considered, where the transmitted bit pattern 1011 1100 1100 0011, the corrupted received bit pattern is: 1010 1101 1101 0011. Here three bits along same column, according to the 2-D model are corrupted. In the 3-D model corrupted bits are (0,3,0), (1,3,0) and (0,3,1). Plane2 has got only one error in column 3. So the error (0, 3, and 1) is identified and corrected. Now, the utility of Z axis parity checking comes. At first parity is generated with respect to (0, 3, 0) and (0, 3, 1) which contradicts with the transmitter side. But as the bit (0, 3, 1) was corrected in the

previous step only, so the corrupted bit (0, 3, 0) is identified and corrected. At last (1, 3, 0) and (1, 3, 1) is used to generate parity bit which again contradicts with the transmitter side. As (1, 3, 1) is uncorrupted so the corrupted bit (1, 3, 0) is identified and corrected.

Another situation considered where for transmitted bit pattern 1011 1100 1100 0011, the corrupted received bit pattern is: 1101 1010 1100 0011. Here the error bits are (0,1,0), (0,2,0), (1,1,0) and (1,2,0). In this case due to two errors each along R1 and R2 on Plane1, the row errors cannot be detected. Similarly due to two errors along C1 and C2 on Plane1, the column errors cannot also be detected. Now, parity checking is done along Z axis among the elements [(0,1,0) and (0,1,1)], [(0,2,0) and (0,2,1)], [(1,1,0) and (1,1,1)] and [(1,2,0) and (1,2,1)] which helps in detection in errors on Plane1 and rectifying them as well.

F. Probability of Failure of Error Detection

In our 3-D model, suppose the element (0, j, 0) is corrupted. This corrupted bit cannot be detected and corrected if simultaneously (1, j, 0), (0, j, 1) and (1, j, 1) are also corrupted. This probability is much less than occurrence of two errors along same column or row of the 2-D matrix method. So the chance of failure to detect and correct errors in the proposed 3-D model is much smaller than the 2-D model.

III. CONCLUSION

In our paper, we thus show that the 3-D parity model performs better than the 2-D parity model. The errors which were even undetected in 2-D method can be detected and even rectified by the 3-D model. The probability of failure to detect errors in 3-D model is much less than the 2-D model. This enhanced performance is obviously achieved at the cost of extra bits to be transmitted than the 2-D method.

REFERENCES

- [1] C.Shannon and W, Weaver, "A mathematical theory of information" Bell System Tech. J. 27 (1948)
- [2] Claude E. Shannon and Warren Weaver, "The Mathematical Theory of Communication", The

University of Illinois Press, 1949.

[3] N. J. A. Sloane, "A survey of recent results in constructive coding theory", Proceedings IEEE National Telemetering Conference 1971, IEEE Press, NY, 1971, pp.18-227.

[4] F. J. MacWilliams and N. J. A. Sloane," The theory of error-correcting codes", North-Holland, Amsterdam, 1978, 762 pp., Russian translation ``Teoriya kodov, ispravlyayushchikh oshibki'' by I. I. Grushko and V. A. Zinov'ev, edited by L. A. Bassalygo, published by Svyaz, Moscow,1979.

[5] Rade Petrovi, Kanaan Jemili, Joseph M. Winograd, Ilija Stojanovi, Eric Metois, "Data hiding within audio signals", MIT Media Lab, Series: Electronics and Energetics vol. 12, No.2, pp.103-122.,1999

[6] W. Bender, W. Butera, D. Gruhl, R. Hwang, F. J. Paiz, S. Pogreb, "Techniques for data hiding", IBM Systems Journal, Volume 39 , Issue 3-4, pp. 547 – 568.,2000

[8]Poulami Dutta, Debnath Bhattacharya, Tai-hoon-Kim," Data Hiding in Audio Signal: A Review",International Journal of Database Theory and Application, Vol.2 No.2, 2000

[9] May O. Lwin, Maureen Morrin, Aradhna Krishna," Exploring the superadditive effects of scent and pictures on verbal recall: An extension of dual coding theory", Journal of Consumer Psychology , Volume 20, Issue 3, July 2010, Pages 317-326

[10] Manjusri Basu, Bandhu Prasad, "Coding theory on the m-extension of the Fibonacci p-numbers ", Chaos, Solitons & Fractals, Volume 42, Issue 4,pp.2522-2530, 30Nov, 2009

[11] Emile H.L. Aarts, Peter J.M. van Laarhoven, "Local search in coding theory ", Discrete Mathematics, Volumes 106-107, pp.11-18, 1992

[12]Tsuneyo Matsumura, Naoaki Yamanaka, Ryōichi Yamaguchi and Keiji Ishikawa, "Real-time emulation, method for ATM switching systems in broadband ISDN" Seventh IEEE International Workshop on Rapid System Prototyping(RSP'96),Greece

[13] Lester A. Gimpelson, "ISDN and value-added services in public and private networks ", Computer Networks and ISDN Systems, Volume 10, Issues 3-4, pp.147-156, 1985

[14] Cherian S. Thachenkary, "Integrated Services Digital Networks (ISDN): six case study assessments of a commercial implementation", Computer Networks and ISDN Systems, Vol.25, Issue 8, pp.921-932, 1993

[15] Jaap Van Till, "The A-ISDN proposal to bridge "personal computers" and "ISDN" , Computer Networks and ISDN Systems, Volume 17, Issue 2, pp.149-152, 1989

[16] Xingfan Du, Graham Knight, "Further study of channel management issues in a LAN-ISDN gateway ", Computer Communications, Volume 16, Issue 4, pp.251-259, 1993

[17] P.M cLane, "Error Rate Lower Bounds for Digital Communication with Multiple Interferences", IEEE, Trans. Communications-23 pp.539-543,1975

[18] Gibak Kim, Nam Ik Cho, "Time varying for getting factor for the noise estimation in multi-channel noise reduction ", Applied Acoustics, Volume 69, Issue 8,pp.751-754, 2008

- [19] François Chapeau-Blondeau, Agnès Delahais, David Rousseau, "Tsallis entropy measure of noise-aided information transmission in a binary channel ", Physics Letters A, Volume 375, Issue 23, pp.2211-2219, (2011)
- [20] NH Kuo, MP Gough, "3D parity checking models for errors in RAM memories used in space on-board computers ", Microprocessors and Microsystems, Volume 14, Issue 9, pp.599-605, 1990
- [21] A.J Poustie, K.J Blow, A.E Kelly, R.J Manning, "All-optical parity checker with bit-differential delay ", Optics Communications, Volume 162, Issues 1-3, pp.37-43, 1999
- [22] Manish Mangal, Manu Pratap Singh, "Analysis of pattern classification for the multidimensional parity-bit-checking problem with hybrid evolutionary feed-forward neural network", Neurocomputing, Volume 70, Issues 7-9, pp.1511-1524, 2007
- [23] Nam-Kyu Lee, Sung-Bong Yang, Kyoung-Woo Lee, " Efficient parity placement schemes for tolerating up to two disk failures in disk arrays ", Journal of Systems Architecture, Volume 46, Issue 15, pp.1383-1402, 2000