6-23-2012

Proceedings of International Conference on Computer Science, Information & Technology

Prof.Srikanta Patnaik Mentor
IRNet India, patnaik_srikanta@yahoo.co.in

Proceedings of International Conference on

Computer Science, Information & Technology

(ICSIT-2012)
23rd JUNE, 2012
AHMEDABAD, India

Interscience Research Network (IRNet)
Bhubaneswar, India

Editorial

In the quest of making this earth a better place to live we have to make a strong hold upon sustainable energy source. Sustainable energy sources include all renewable energy sources, such as hydroelectricity, solar energy, wind energy, wave power, geothermal energy, bio energy, and tidal power. It usually also includes technologies designed to improve energy efficiency. Energy efficiency and renewable energy are said to be the twin pillars of sustainable energy. Renewable energy technologies are essential contributors to sustainable energy as they generally contribute to world energy security, reducing dependence on fossil fuel resources, and providing opportunities for mitigating greenhouse gases. If we progress, the discussion will never limit because advancement is seamlessly flowing at the most efficient and state-of-the art universities and research labs like Laboratory for Advanced Systems Research, University of California. Unquestionably apex bodies like UNO, WTO and IBRD include these two disciplines in their millennium development agenda, realizing the aftermath of the various application projects like VSAT, POLNET, EDUSAT and many more. 'IT' has magnified the influence of knowledge management and congruently responding to social and industrial revolution.

Each Conference is conducted in collaboration with other institutes. We encourage and invite proposals from institutes within India to join hands to promote research in various areas of discipline. So far we have organized more than 100 conferences since our inception in 2008. There were over 5000 research papers presented in those conferences. These conferences have not only promoted the international exchange and cooperation, but have also won favorable comments from national and international participants, thus enabling us to reach out to a global network within three years time. The areas covered under the auspices of this conference are:

Algorithms	Artificial Intelligence
Automated Software Engineering	Bio-informatics
Bioinformatics and Scientific Computing	Biomedical Engineering
Compilers and Interpreters	Computational Intelligence
Computer Animation	Computer Architecture & VLSI
Computer Architecture and Embedded Systems	Computer Based Education
Computer Games	Computer Graphics & Virtual Reality
Computer Graphics and Multimedia	Computer Modeling
Computer Security	Computer Simulation
Computer Vision	Computer-aided Design/Manufacturing
Computing Ethics	Computing Practices & Applications
Control Systems	Data Communications
Data Compression	Data Encryption
Data Mining	Database Systems
Digital Library	Digital Signal and Image Processing

Digital System and Logic Design	Distributed and Parallel Processing
Distributed Systems	E-commerce and E-governance
Event Driven Programming	Expert Systems
High Performance Computing	Human Computer Interaction
Image Processing	Information Retrieval
Information Systems	Internet and Web Applications
Knowledge Data Engineering	Mobile Computing
Multimedia Applications	Natural Language Processing
Neural Networks	Parallel and Distributed Computing
Pattern Recognition	Performance Evaluation
Programming Languages	Reconfigurable Computing Systems
Robotics and Automation	Security & Cryptography
Software Engineering & CASE	System Security
Technology in Education	Technology Management
Theoretical Computer Science	Ubiquitous Computing
Wireless Sensor Networks	Wireless Communication and Mobile Computing

The conference is first of its kind and gets granted with lot of blessings. About 150 papers were received and after scrutiny 35 papers were selected for publication in the proceedings. I wish all success to the paper presenters. The papers qualifying the review process will be published in the forthcoming IOAJ journal.

The conference designed to stimulate the young minds including Research Scholars, Academicians, and Practitioners to contribute their ideas, thoughts and nobility in these disciplines of engineering. It's my pleasure to welcome all the participants, delegates and organizer to this international conference on behalf of IRNet family members. We in IRNet believe to make "**RESEARCH COOL**"**.** We received a great response from all parts of country and abroad getting more than 100 papers and peer review 19 papers got selected for the presentation and publication in the proceeding of the conference. I sincerely thank all the authors for their invaluable contribution to this conference. I am indebted towards the reviewers and Board of Editors for their generous gifts of time, energy and effort.

Editor-in-Chief

Dr. Srikanta Patnaik
Chairman, I.I.M.T., Bhubaneswar
Intersceince Campus,
At/Po.: Kantabada, Via-Janla, Dist-Khurda
Bhubaneswar, Pin:752054. Orissa, INDIA.

A Survey on Neuromorphic Engineering

Pratibha. S. Karajgi & Mohammed Riyaz Ahmed

VLSI Design and Embedded System, Department of ECE, RITM, Bangalore, India.

Abstract - Neuromorphic systems are inspired by the structure, function and plasticity of biological nervous systems. This field is evolving a new era in computing with a great promise for future medicine, healthcare delivery and industry. This paper focuses on the emerging trends in computational cognitive sciences by surveying on a new interdisciplinary field called neuromophic engineering. A complete overview starting from its origin to its applications is described. The overall process of developing neural networks and simulation of them in a form of neuromophic chip is explained, and then introduces neuromophic vision systems to demonstrate the potential of this technology. Various advancements in Neuromorphic approach by several research groups/labs are discussed. Finally, to articulate significant conclusions to a wider audience, paper concludes with open research issues.

Key Words **:** *neuromorphic, neuromorphic engineering , neuromorphic chip; neuromorphic vision systems.*

I. INTRODUCTION

The term neuromorphic was coined by Carver Mead, in the late 1980s to describe very-large-scale integration (VLSI) systems containing electronic analog circuits that mimic neurobiological architectures present in the nervous system. In recent times the term neuromorphic has been used to describe analog, digital or mixed-mode analog/digital VLSI systems that implement models of neural systems (for perception, motor control, or sensory processing) as well as software algorithms.

Neuromorphic engineering is a promising interdisciplinary discipline that takes inspiration from biology, physics, mathematics, computer science and engineering to design artificial neural systems, such as vision systems, head-eye systems, and autonomous robots, whose physical architecture and design principles are based on those of biological nervous systems [3].

Neuromorphic engineers are using garden-variety VLSI, complementary metal oxide semiconductor (CMOS) technology to achieve their goal. This effort is facilitated by similarities between VLSI hardware and neural wetware. Both technologies: Provide millions of inexpensive, poorly-matched devices. Operate in the information-maximizing low-signal-to-noise/high-band width regime [4].

II. THE PROCESS OF DEVELOPING NEUROMORPHIC CHIP

The neuromorphic engineering could be divided into neuromorphic modeling, reproducing neurophysiological phenomena to increase the understanding of the nervous systems and neuromorphic computation which uses the neuronal properties to build neuron like computing hardware. Basically, the former provides the knowledge of the biological algorithm while the latter translates the algorithm into electrical circuits. This is an iterative process, since the understanding of the biological algorithm is a very complex process. As more knowledge evolves yielding improved algorithm, the electrical circuits are revised and improved.

These circuits then pass through all the stages of developing integrated circuit (or chip), which involves the circuit layout, verification, fabrication in foundry and testing and subsequent deployment. A brief explanation of each of these steps is provided below:

- Layout design: This stage involves the translation of the circuit realized in the previous stage into silicon description through geometrical patterns aided by computer aided design (CAD) tools. This translation process follows a process rule that specifies the spacing between transistors, wire, and wire contacts and so on.

The layout is designed to represent the electrical circuit schematics obtained from the algorithm.

- Fabrication: Upon satisfactory verification of the design, the layout is sent to the foundry where it is fabricated. The process of chip fabrication is very complex. It involves many stages of oxidation, etching, photolithography, etc. Typically, the fabrication process translates the layout into silicon or any other semiconductor material that is used.
- Testing: The final stage of the chip development is called testing. Electronic equipment like oscilloscopes, probes, and electrical meters are used to measure some parameters of the chip, to verify its functionalities based on the chip specifications [2].

A. A neuromorphic approach to vision systems

Compact, efficient electronics based on the brain's neural system could yield implantable silicon retinas to restore vision, as well as robotic eyes and other smart sensors. -- By Kwabena Boahen

Today's computers can perform billions of operations per second, but they are still no match for even a young child when it comes to skills such as pattern recognition or visual processing. The human brain is also millions of times more energy-efficient and far more compact than a typical personal computer.

The efficiency of neuromorphic chips make it possible to develop fully implantable artificial retinas for people afflicted by certain types of blindness as well as better electronic sensors. Someday, neuromorphic chips could even replicate the self growing connections the brain uses to achieve its amazing functional capabilities [5].

Dr. Carver Mead, an emeritus professor of California Institute of Technology (Caltech), Pasadena, pioneered the field of neuromorphic engineering [1, 6]. He reasons that biological evolutionary trends over millions of years have produced organisms that engineers can study to develop better artificial systems. By giving senses and sensory-based behaviors to machines, these systems can possibly compete with human senses and bring an intersection between biology, computer science and electrical engineering [7].

Neuromorphic vision sensors and preprocessors are increasingly being used to implement the first steps of visual processing in artificial systems. A typical application domain is autonomous mobile systems, for which size, power consumption and speed are important parameters. While until recently research in neuromorphic vision concentrated on designing and optimizing individual circuits, part of the effort is now shifting towards multichip neuromorphic systems and hybrid systems interfacing neuromorphic circuits with other types of processors [8].

B. A neuromorphic approach to Autonomous Robots

Autonomous Robots, which achieve tasks without human operators, are required in many fields. The autonomous robots can carry out tasks in various environments by themselves like human beings. They have to be intelligent to determine their own motions in unknown environments based on sensory information.

Most people have never even heard of the term "neuromorphic"- which is a technology with a specific form ("morphic") that is based on brain ("neuro") architecture. The neural models being developed by the neuromorphics lab implement "whole brain systems", or large-scale brain models that allow virtual and robotic agents to learn on their own to interact with new environments [9].

C. Areas of Applications

Their applicability to engineering challenges is widespread, and includes biomedical, displays, hearing, imaging, language, locomotion, neural interface, power, processing, robotic, security and vision. Neuromorphic engineering proposes to fill the gap between, on the one hand, computational neuroscience, and, on the other hand, traditional engineering [13].

In the biomedical field, the neuromorphic chip can be used to stimulate locomotion in an animal (tested on a temporarily-paralyzed cat) [14].

Neuromorphic vision systems are ideal for mobile applications because they promise compact computational sensing at lower power consumption compared to the traditional imager systems [15].

Neuromorphic motion sensors are attractive for use on battery powered robots which require a low payload [17]. Because of its compact hardware and low power dissipation, the neuromorphic vision system developed in the present study is suitable to robotic vision. More interestingly, it provides insights to explore the visual function of the neuronal network of the brain, visualizing neural images inferred from physiological experiments [18].

III. THE SURVEY

Here we briefly review some recent and important work which was carried out in last two decades.

1. Telluride Neuromorphic Cognition Engineering Workshop, Colorado: Neuromorphic engineers design and fabricate artificial neural systems whose organizing principles are based on those of biological nervous systems. Over the past 16 years, this research community has focused on the understanding of low-

level sensory processing and systems infrastructure; efforts are now expanding to apply this knowledge and infrastructure to addressing higher-level problems in perception, cognition, and learning [10].

2. Hewlett-Packard (HP) Laboratories: Working in collaboration with HP, the Neuromorphics lab, has undertaken the ambitious project of creating a brain on a chip- a fundamental predecessor to the design of autonomous robotics and general intelligence [9].

3. University of Utah, Salt Lake City: The vision chip, built by Reid Harrison, is a "pixellated" light sensor that reads an image using an array of individual cells, with additional circuitry built locally into each cell to process the incoming signals. The fact that these processing circuits are local and analog is crucial to the device's operation-and is a feature that is barrowed from the biological model.

4. Dr Harrison and his supervisor at Caltech and co-founder of the Telluride summer School, Christ of Koch: Identified the various processes taking place in the so called lamina, medulla and lobular-plate cells in a fly's brain as being worth implementing in silicon. These cells form a system that allows the fly to detect motion throughout most of its visual field-letting the insect avoid obstacles and predators while compensating for its own motion.

5. Giacomo Indiveri, a researcher at the Federal Institute of Technology (ETH) in Zurich: Has been using a network of "silicon neurons" to produce a simple kind of selective visual attention. Instead of working with purely analog devices, the ETH group uses electrical circuits to simulate brain cells(neurons)that have many similarities with biological systems-displaying both analog and digital characteristics simultaneously, yet retaining all the advantages of being analog.

6. Telluride Neuromorphic Cognition Engineering Workshop: The concept of sensory feedback is a key part of another project shown at the Telluride workshop. In this case, a biologist, robotics engineer and analog-chip designer collaborated on a walking robot that used the principle of a "central pattern generator" (CPG)-a kind of flexible pacemaker that humans and other animals use for locomotion. (It is a chicken's CPG that allows it to continue running around after losing its head) [10].

7. Experimental neuroscience: There is a large number of laboratories world-wide, with a trend for the USA to lead, but Europe is very strong in a number of areas, e.g. vision, cortex synchronized oscillation in neuronal networks, hybrid (biological-plus silicon or computer simulated) neuronal networks, patch clamp and multiple single cell recording techniques, psychophysics and neurophysiology of multisensory perception by vision, the vestibular system and at a lesser degree the tactile and proprioceptive systems.

8. University of Maryland: The biological model on which the walking robot is based was developed by Avis Cohen at College Park. Dr Cohen had been studying the way that neural activity in the spinal cord of the lamprey (an eel-shaped jawless fish) allowed it to move, with the sequential contraction of muscles propelling it forward in a wave motion [11].

9. IBM's first neurosynaptic computing chips: The computing chips recreate the phenomena between spiking neurons and synapses in biological systems, such as the brain, through advanced algorithms and silicon circuitry. Its first two prototype chips have already been fabricated and are currently undergoing testing [12].

10. IBM Research centre: The Company and its university collaborators announced they have been awarded approximately $21 million in new funding from the Defense Advanced Research Projects Agency (DARPA) for Phase 2 of the Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE) project. The goal of SyNAPSE is to create a system that not only analyzes complex information from multiple sensory modalities at once, but also dynamically rewires itself as it interacts with its environment-all while rivaling the brain's compact size and low power usage [16].

IV. DISCUSSION

In this section we present a general discussion of some issues raised by our survey:

- ✓ What should the role of neuromorphic systems be? Learning about neurobiological systems by rebuilding them, or engineering new solutions to problems in sensory perception using what we know about animal sensory perception?
- ✓ How should neuromorphic systems be implemented? In software or hardware? As dedicated VLSI devices? In analog or digital hardware? In analog VLSI in the weak or strong inversion regions? Should particular methodologies or specific transistor characteristics be used?
- ✓ Can biologically-inspired techniques provide the basis for real improvements in auditory/ visual/ sensorimotor systems?
- ✓ Can cognitive science be informed by what neuromorphic systems have to tell us?
- ✓ Should biology be the guide for the evolution of silicon technologies, or should their evolution be guided by

their own properties? What do neuromorphic systems have to offer the cognitive science community? Are they just hardware implementations of neural networks, or what neuromorphic systems have to tell us about sensory perception and coding inform cognitive science?

✓ The brain is extremely complex with 100 billion neurons assemble their 100 trillion connections. Whether it is possible to rig-up such complex circuit without any explicit design or designer?

V. CONCLUSION

This paper focuses on the introduction of neuromophic engineering and neuromophic vision systems. It describes the process of developing neuromophic chip. Finally it summarizes the applications of neuromophic engineering based on neuromophic chip. Modern advancement in biology is enabling better understanding of forms, structure, and behavior of biological systems to develop algorithms implemented in analog integrated circuits. These circuits are developed in parallel distributed architectures with elements of adaptation and learning using low power and high integration density technologies.

Effective interdependence collaborations among the key areas of biology, electrical engineering, physiology and computer science are very fundamental to develop and engineer this emerging area of computational biology [2], [3].

REFERENCES

[1] C.A. Mead, "Neuromorphic electronic Systems, Proceedings of the IEEE, vol.78 (10) pp.1629-1639, 1990.

[2] PaulaBach "Neuromorphics: Biological nervous systems morphed on silicon," unpublished.

[3] Wikipedia,[online] http://en.wikipedia.org/wiki/.

[4] K. Boahen, "Point-to-Point connectivity between neuromorphic chips using address events," IEEE Trans. On Circuits and Systems Part-II, vol. 47 (5), pp. 416-434, May 2000.

[5] K. Boahen,. "Neuromorphic Microchips," Scientific American, vol. 5, 2005, pp. 56-63.

[6] C.A. Mead, "Analog VLSI and neural systems," Reading, MA: Addison-Wesley, 1989, pp. 35-40.

[7] L.Smith, A. Hamilton, "Neuromorphic systems: engineering silicon from neurobiology," World Scientific Pub Co Inc, River Edge, NJ, USA, pp.145- 158, 1998.

[8] Kramer J. and Indiveri G., "Neuromorphic Vision Sensors and Preprocessors in System Applications," Advanced Focal Plane Arrays and Electronic Cameras II, vol. 3410, 1998, pp. 134- 146.

[9] The Neuromorphics Lab, part of the National Science Foundation- (NSF) sponsored Center of Excellence for Learning in Education, Science and Technology (CELEST). http://nl.bu.edu/people/researchers/.

[10] Telluride Neuromorphic engineering Workshop.http://ine- web.org/t elluride-conference- 2011/ telluride- 2011/ index.html.

[11] Cohen, A. H. and Wallén, P. (1980). The neuronal correlateto locomotion in fish: "Fictive swimming" induced in an invitro preparation of the lamprey spinal cord. Exp. Brain Res. Vol. 41, pp. 11-18.

[12] IBM Research centre, http://www-03.ibm.com/press/us/en/pressrelease/35251.wss

[13] G. Cauwenberghs , "Neuromorphic learning VLSI systems: A survey," In: T.S. Lande, Editor, Neuromorphic systems engineering, Kluwer Academic Publishers, 1998, pp. 381–408.

[14] Sunny Bains. "Analog chip does job of spinal cord for locomotion," unpublished.

[15] Ralph Etienne-Cummings, Swati Mehta, Ralf Philipp and Viktor Gruev, "Neuromorphic vision systems for mobile applications," IEEE 2006 Custom Intergrated Circuits Conference (CICC), San Jose, CA, pp. 531– 534, Sept. 2006.

[16] IBM Research centre, www.ibm.com/research or IBM Research blog: http://ibmresearchnews. blogspot.com/.

[17] Lukas Reichel, David Liechti, Karl Presser, and Shih-Chii Liu, "Robot guidance with neuromorphic motion sensors," The 2005 IEEE international Conference on Robotics and Automation, Barcelona, Spain, pp. 3540-3544, April 2005

[18] Tetsuya Yagi and Kazuhiro Shimonomura ,"Silicon primary visual cortex designed with a mixed analog- digital architecture," Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, pp. 2723–2728, 2007.

A 0.1-2 GHz Low Power Folded RF Front-End with Merged LNA and Mixer for Software-Defined Radio Applications

Manjula. K & Prathibha. S. K

RITM Bangalore.

Abstract **-** In this paper, A Software-Defined Radio (SDR) RF front-end is presented that contains merged LNA and mixers, VGAs, and frequency synthesizer, supporting various wireless communication standards in 0.1-2 GHz while guaranteeing a power/performance trade-off at any time. The proposed low power RF front-end uses the folded and current reuse techniques. for 0.18 um RF CMOS technology with 1.8V supply voltage. In the receive path the proposed design achieves a Noise Figure of 3.8 dB at 160 MHz and 5.5 dB at 2GHz. The Output-referred 3rd-order Intercept Point (OIP3) is high up to 21.3 dBm at 800 MHz. The voltage gain of the front- end is between 16-44 dB. The phase mismatch of LO quadrature signals is lower than 3deg.It consumes 13.8 mW at the 1.7V supply.

Key Words**:** *SDR; LNA; Mixer; VGA*

I. INTRODUCTION

The demand for the integration of multiple standards into a single portable terminal is growing together with the proliferation of wireless communication standards [1]-[3]. Such handsets could be implemented with multiple dedicated front-ends integrated in parallel. However, that solution is everything but optimal for cost. A Software-Defined Radio (SDR) optimizes the functionality versus area trade-off, by programming a versatile front-end to the desired standard. The boundary conditions being that for each supported standard both performance and power consumption should be comparable to dedicated solutions. The SDR front-end should be compatible with various wireless communication standards, including DRM, DAB, DVB-H, GSM and GPS.

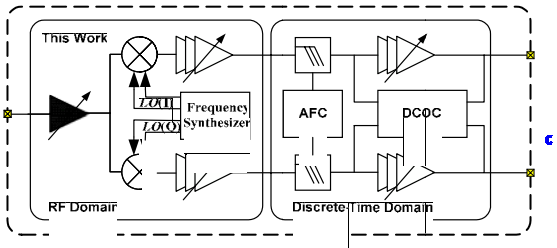


Figure 1. Simplified SDR block diagram

The system architecture is presented in Fig.1. The direct- conversion architecture is the best candidate to realize such SDR as it has the highest potential to reduce the cost, size and power. The receiver usually contains two major parts of the RF-Domain and the Discrete-Time Domain. In RF Domain, a single wideband LNA is used instead of the multi- LNA solutions. Single-ended input LNAs are preferred to save I/O pins and because antennas and RF filters usually produce single-ended signals. Compared to traditionally wireless communication applications the linearity of the proposed I/Q mixer should be high enough to prevent the receiver from being blocked by the large signals in adjacent bands. The VGAs is added to provide a moderate signal power level to the Discrete-Time Domain. In order to provide LO quadrature signals, a wideband frequency synthesizer consists of a single loop PLL and wideband high speed frequency dividers. The detail analysis and circuits design are present in section II.

II. CIRCUITS DESIGN

Referring to Fig.1, the RF front-end consists of four parts: single-ended input and differential output LNA, I/Q mixer, two VGAs and wideband frequency synthesizer. The LNA must have the lowest possible Noise Figure and high enough gain to suppress the noise contribution of the second stage. The mixer and the VGAs are designed aiming for high linearity. The VGAs should have the linear-in-dB property. Meanwhile, the frequency synthesizer should have the wide band, low phase noise and better orthogonality.

A. Low Noise Amplifier

Fig.2 shows the schematic of the proposed LNA for SDR applications. It consists of three stages. The first stage is Low Noise Stage (LNS). The LNS is

a low noise amplifier exploiting a noise-canceling technology [4]. The Noise Figure is calculated as follows:

$$NF = 1 + \frac{R_s}{R_f} + \left(\frac{R_s}{R_f}\right)^2 + \frac{\gamma}{4}\cdot\frac{R_s}{R_f}\left(\frac{2R_s}{R_f}+1\right) + \frac{KR_s^3}{8kTR_f^2\cdot f} \quad (1)$$

In Eq. (1), the parameter R_S and R_f are used to quantify the source load impedance and the feedback resistance. parameter γ is noise parameter of transistors. For a deep- submicron MOSFET, the value of γ is usually between 1 and 3. K is an experiential parameter.

The second stage is a Single to Differential Converter (SDC). The compensatory capacitor C_1 is parallel to the resistor R_1 to make the mismatch of the amplitude and the phase of the output differential signal lower than 0.3 dB and 4 degrees.

The third stage is a Differential Multiple Gate Transistor (DMGTR) [5]. The M_{11} and M_{12} are Fully Differential Amplifiers (FDA), the M_{13} and M_{14} are Pseudo-Differential Amplifiers (PDA). The FDA usually suffers from low linearity problems due to the negative g_m". Fortunately, when bias the PDA from saturation to near threshold regime, the g_m" of the PDA can be moved from negative to positive. From the above consideration, the negative value of g_m" which degrades linearity of FDA can be compensated by positive value of g_m in the PDA by adjusting the bias and transistor size of PDA. This method does not require extra power consumption, but the OIP3 can be improved above 10 dB.

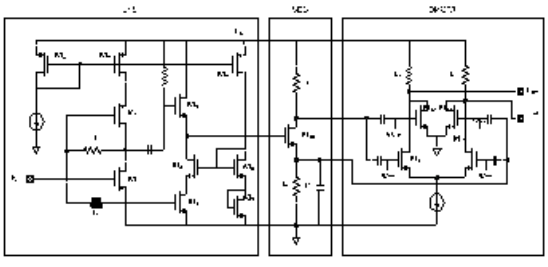


Figure 2. Schemetic of the LNA

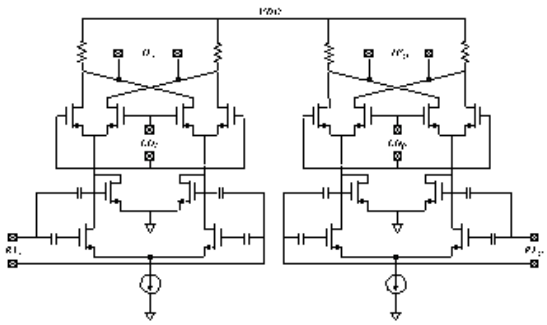


Figure 3. Schemetic of the mixer

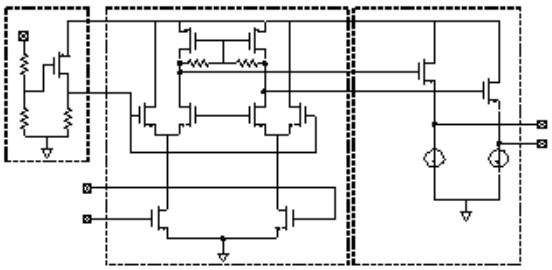


Figure 4. Schemetic of the VGA

B. Mixer

Fig.3 shows the schematic of the mixer. The mixer consists of two modified double balanced Gilbert cell. Compared to the traditional double balanced Gilbert cell, the input transconductance stage of proposed mixer uses the DMGTR topology. It can achieve high linearity but consume the same power than the FDA topology.

C. Variable Gain Amplifer

Signal-summing topology has been widely used in low- power and high-frequency VGA design. Fig. 4 shows the circuit schematic of the proposed signal-summing variable- gain stage with exponential gain control. Compared with the previous work in [6], the proposed variable-gain stage used PMOS transistors M_7 and M_8 as a load instead of a resistor resistors R_1 and R_2 are used as a common-mode feedback circuit, the tail current source is eliminated to enhance the linearity of VGA in low voltage supply process. With the same linear-in-dB gain characteristics on the control voltage, the output voltage swing is larger than the design in [6], especially at the low-gain mode when the input signal is a large signal. The reason is that the DC equivalent resistance of PMOS load in proposed VGA

is large than the resistor in [6], while the AC small-signal equivalent resistance maintains the same. The DC voltage at the drain of the PMOS M7 and M8 is around 1.3V in low-gain mode, while the DC voltage of resistor load in [6] is 1.7V. So, the total harmonic distortion in proposed VGA is much smaller than the previous signal-summing VGA design.

The exponential gain control circuit is shown in Fig. 4. The PMOS transistor M_{11} works in linear region and in common-source configurations. By combining the variable- gain stage and the exponential gain control circuit, the logarithmic current gain of the circuit becomes linear along with the control voltage as follows.

$$Gain\ (dB) = 10 \times \log K_1 + 10 \times C \times \log e + 10 \times K_2 \times \log e \times V_c \quad (2)$$

Where the K_1 and K_2 can be written as

$$K_1 = \frac{W_3}{W_4 (V_{GS4} - V_{TH4})} \quad K_2 = \frac{R_4 R_5}{R_3 + R_4} \cdot \mu_P C_{ox} \cdot \frac{W_{11}}{L_{11}} \quad (3)$$

D. Frequency Synthesizerr

Fig.5 shows the schematic of the frequency synthesizer. The frequency synthesizer consists of a phase frequency detector (PFD), a charge pump (CP), loop filters (LP), a voltage controlled oscillator (VCO), a swallow pulse frequency divider (SPFD) and a multi-mode frequency divider (MMFD). Compared to the ordinary structure, the proposed frequency synthesizer with additional MMFD can achieve better orthogonality and maintain wideband property.

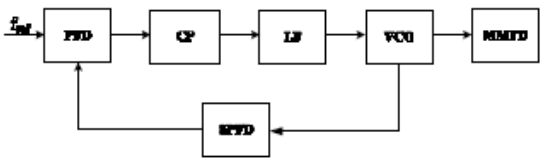


Figure 5. Schematic of the frequency synthesizer

E. Operation Principle and design of the proposed low power folded RF front end

The proposed RF front-end uses a folded mixer. Figure 6 shows the schematic of the folded Gilbert cell mixer. The use of folded technique can decrease the power consumption because the switching stage of the mixer does not need large current. High voltage headroom is also achieved. Another merit of the folded mixer is to have low flicker noise.

However, if PMOS is used at a switching stage of the mixer, flicker noise can be difference of CMOS technology is considered, the proposed RF front-end achieves the smallest power consumption and the highest gain among the other lowpower RF front-ends. The proposed folded RF front-end achieves not only the low power consumption, but also high gain compared with other RF front-ends.

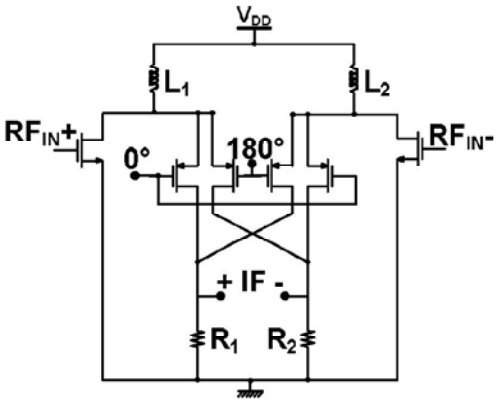


Fig. 6. The schematic of the folded Gilbert cell mixer

However, if PMOS is used at a switching stage of the mixer, flicker noise can be difference of CMOS technology is considered, the proposed RF front-end achieves the smallest power consumption and the highest gain among the other lowpower RF front-ends. The proposed folded RF front-end achieves not only the low power consumption, but also high gain compared with other RF front-ends.

Typically, two methods, folded technique and current reuse method, are used for low power operation . In folded technique, the stacked circuit is converted to a cascade circuit.

The power consumption decreases because VDD is lowered by reducing the stacks, while the chip size increases by the added inductor. This method can be applied to high gain circuits because this method increases the voltage headroom due to the reduced stacks. Current reuse method can decrease the current consumption by converting the cascade connected circuit to the stacked circuit.

By sharing current, the power consumption decreases. However, because this method can reduce the voltage headroom due to the increased stacks, its linearity is limited.

III. COMPARISON BETWEEN PROPOSED LOW POWER FOLDED RF FRONT END & GENERAL RF FRONT END

	Current(mA)	Power(mW)
Proposed Design	8.11	13.8
Conventional Design	46.35	78.8

From the above table we can summarize that the conventional design consumes six times more power than the proposed low power folded RF front end with merged LNA and Mixer The proposed RF front-end can be used in direct conversion receiver for a relatively narrow bandwidth system.

IV. EXPECTED RESULTS

A. Variable Gain Amplifier

Fig.8 shows the post-simulated voltage gain of VGA with the control voltage from 0.2 V to 1.6 V in 0.2 dB step. This VGA operates up to 700 MHz. The linear in-dB variable-gain range is from -10dB up to 18dB.

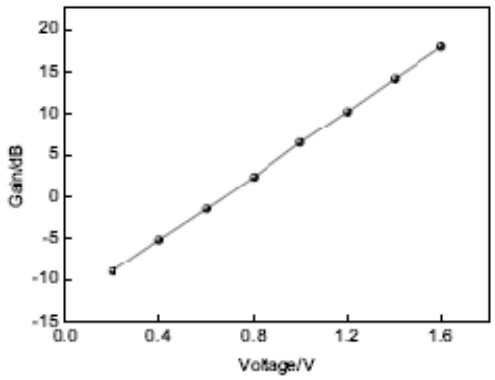


Figure 7. Post-simulated linear-in-dB characteristics of VGA

B. Frequency Synthesizer

When the control voltage is 0.9 V and the control-code is '0000', the post–simulated phase noise of VCO is shown in Fig 8.

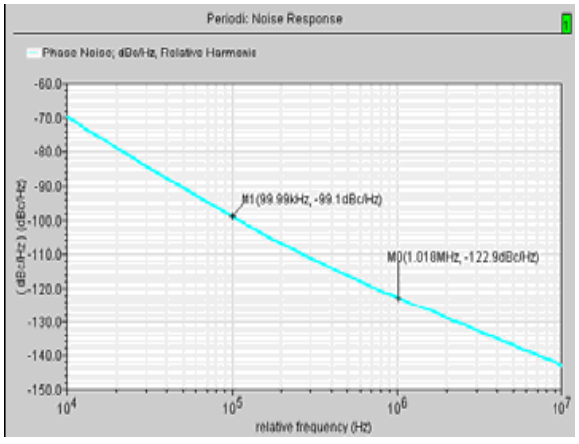


Figure 8. The post–simulated phase noise of the proposed VCO

V. CONCLUSION

In this paper, a broadband RF front-end for 0.1-2 GHz SDR radio receiver has been demonstrated in SMIC's 0.18 µm RF CMOS technology. The noise-canceling and DMGTR method breaks the trade-offs between the noise figure, source impedance matching and linearity. The proposed low power RF front-end is designed with current reuse technique and folded mixer in SIMC's 0.18 um CMOS process. The proposed RF front-end not only shows low power and high gain characteristics in comparison with those of other low power RF front-ends, The expected results of RF front-end are summarized in Table.1. Post-simulated results show that the performance of the RF front-end meets the requirement of SDR applications.

TABLE I. SUMMARY OF THE PROPOSED DESIGN

Freq. range	0.1-2 GHz
Gain. range	16-44 dB
NF (high gain)	3.8 dB@ 160MHz, 5.5 dB@2 GHz
OIP3@800 MHz	21.3 dBm
Phase noise of VCO	-122.9 @1 MHz
Phase Mismatch	<3°
Power	13.8mW@1.7V

REFERENCES

[1] J Craninckx, et al, "A fully reconfigurable Software-Defined Radio transceiver in 0.13µm CMOS", Proc of ISSCC., pp.346-347,Feb. 2007.

[2] R. Bagheri, et. al, "Software-defined radio: dream to reality," IEEE Communications Magazine, pp. 111-118, Aug 2006.

[3] A A Abidi, "The path to the software-___defined radio receiver". IEEE Journal of Solid-State Circuits, 2007, 42(5): 954-966.

[4] F. Bruccoleri, E. A. M.Klumperink and B. Nauta, "Wide-band CMOS low-noise amplifier exploiting thermal noise canceling ". IEEE J Solid-State Circuits, 2004, 39(2): 275-282.

[5] T W Kim, B Kim, "A 13-dB IIP3 improved low-power CMOS RF programmable gain amplifier using differential circuit transconductance linearization for various terrestrial mobile D-TV applications". IEEE J Solid-State Circuits, 2006, 41(4):945-953.

[6] V. Knopk and D. Belot, "0.18 um thin oxide CMOS transceiver front-end with integrated Tx/Rx commutator for low cost Bluetooth solutions," ESSCIRC 03, pp. 675-678, Sept. 2003.

Detecting and Effective Routing of Time Critical Events Using Congestion and Delay Aware Routing in WSNs

Chiranthana H R & P C Srikanth

Malnad college of engineering, Hassan

Abstract - Reliability and timeliness are two essential requirements of successful detection of critical events n Wireless Sensor Networks (WSNs). The base station (BS) is particularly interested about reliable and timely collection of data sent by the nodes close to the ongoing event, and at that time, the data sent by other nodes have little importance. In this paper, we modify Congestion and Delay Aware Routing (CODAR) protocol that tries to route data to the multiple sinks in congestion and delay aware manners. In this case every sensor communicates with the closest sink. If congestion occurs, it also mitigates congestion by utilizing an accurate data-rate adjustment. Each node collects control information from neighbours and works in a distributed manner. Experimental results show that modified CODAR protocol is capable of avoiding and mitigating congestion effectively, and performs better than similar known techniques in terms of reliable and timely event detection.

Keywords-*event reliability; congestion; timeliness; wireless sensor network.*

I. INTRODUCTION

Wireless sensor networks (WSNs) [1] have attracted increasing attention recently with the growing development of Micro-Electro-Mechanical System (MEMS). A typical WSN consists of few tens to thousands of sensor nodes that are deployed in a field and work together for a specific task. These sensors are small in size and they get their power from built-in batteries. As WSNs are frequently deployed in inaccessible areas, it is difficult to replace these batteries after depletion. Therefore, existing research works have focused mostly on energy issue [2]. But reliability remains one of the vital issues in WSNs. If a WSN is set up to detect fire in a sensitive area then we would like to get data reliably and timely from those nodes that detect higher temperatures beyond a certain threshold. Here, data delivery success rate of these critical nodes is equally important as the energy efficiency. But reliable data delivery is inherently correlated to congestion. Congestion in the network causes packet drop which reduces the reliability of data transmission. We need an effective congestion control technique to achieve the required level of reliability. In case of event sensing WSNs, time critical reliability of transmitted data is of great importance. Some event (like, fire ignition) can be controlled with minimal effort if the event is detected early. Beside the congestion control, a delay aware routing of data is necessary in these phenomena to meet the time criticality so that early detection of events is possible. In this paper, we consider applications where sensor nodes are deployed in ad-hoc manners to detect critical events in the deployed area. All sensor nodes forward their data towards a single static base station (BS). We design a routing protocol that proactively avoids congestion and meets delay requirements of transmitted data by choosing lightly loaded and low delay incurring nodes during data forwarding towards the BS. All nodes broadcast periodic control data packets describing their congestion status and delay measurements so that the neighbouring nodes can utilize these data during route selection process. The performance of the proposed scheme is highly dependent on the successful delivery of these control packets. Special effort has been made to improve the success probability of these control packets. The proposed protocol also has a congestion mitigation technique. The MAC layer always sends feedback to network layer about its achievable data forwarding rate. If the application layer has a higher traffic generation rate, the protocol suggests the application layer to lower its rate. The network layer simply drops an appropriate fraction of packets received from other nodes if the incoming rate is higher than the data forwarding rate of MAC layer. In this paper, our primary objective is to improve the reliability and the timeliness of data transmitted by the critical nodes (i.e., nodes close to the current event) through congestion avoidance and mitigation. Our contribution is three-fold: we propose (i) a simple but highly effective method to ensure desirable node density in inaccessible areas, (ii) techniques by which every node can measure the end-to-end delay of its packets reaching the BS and can route data packets in a deadline-aware manner, and (iii) a MAC layer specific approach to ensure high success probability of control data sent by different nodes in the network which are utilized by neighbouring nodes to choose appropriate routes. The rest of the paper is organized as follows:

section II provides a discussion on relevant existing works, section III thoroughly explains the design issues of the proposed protocol whereas its performance has been examined in section IV and finally section V presents conclusions on this research.

II. RELATED WORK

Most of the existing congestion-control schemes [3] in WSNs aim at mitigating congestion after its formation. Protocols presented in CODA [4], Siphon [5] and TARA [6] mitigate congestion rather than avoiding it. Packets are dropped due to congestion and thus reliability of data is reduced until congestion is mitigated. Congestion avoidance reduces such packet drops. RTMC [7] provides a reliable transport with memory consideration. A node defers transmission until it gets a node with free buffer space. Although, the authors claim to achieve congestion control, actually they avoid congestion without considering delay in data transfer. CAM [8] provides a routing protocol that tries to avoid as well as mitigate congestion to ensure successful event detection. The scheme assumes a high number of critical packets (packets sent by the nodes near an event) successfully reached the BS as the successful detection of an event. But it does not consider delay of these packets. In most cases, a critical packet reaching the BS has some significance if it reaches within a time threshold after its generation at the critical node. Moreover, the performance of the protocol relies on periodic control data broadcasted by nodes; but the scheme has no special technique to ensure successful delivery of these control data. Many recent works attempt to meet delay bound of data packets reaching the BS. The technique presented in [9], considers residual-time-aware routing where each sensor node is static and randomly duty cycled of other nodes. The scheme is applicable when only one node in the network acts as a data source. Moreover, end-to-end data delivery is not guaranteed. It cannot be used in event detection system where end-to-end data delivery from multiple critical nodes is essential. In [10], Heo *et al.* present a routing protocol where each node considers energy, delay and reliability of its neighbours to choose a suitable route. Nodes periodically broadcast beacon messages to exchange control data with neighbours. The performance of the protocol highly depends on successful transmission of these beacon messages. The protocol considers IEEE 802.11 DCF as its MAC protocol, but does not take any special measure for successful transmission of beacon messages which may be collided and the performance of the protocol might be seriously degraded. Munir *et al.* develop a mathematical model to minimize the delay in the network through the choice of suitable transmission scheduling and it is described in [11]. The model particularly fits where nodes' sampling processes are independent of their transmission schemes. The model does not fit well in an event sensing network where some nodes' sampling processes abruptly change in response to an event. Also the model is computationally expensive and therefore, can be applied to networks with few nodes which is not the case for most of event sensing WSNs. Reference [12] provides an energy aware real-time routing protocol. To choose a suitable route, every node uses energy level and hop-count of its each neighbour. Hop-count of a node means the total number of intermediate nodes involved to transmit a data packet form this node to BS. The protocol uses hop-count of a node as a measurement of delay of that node, which is not correct. Due to congestion around a particular event, each link will have a different delay associated with it. Hop-counts of neighbours are obtained only once during network setup, but energy levels of neighbours are obtained periodically through control messages sent by neighbours. Like works presented in [8, 10], this protocol does not take any special measure to improve success probability of control messages. All schemes in [9 - 12] treat all data with equal importance. But we are considering event detection in WSNs where few nodes close to the ongoing event produce very important data with a high generation rate. Our aim is to route data in congestion and delay aware manners so that the highest amount of important data can reach the BS timely which is essential for successful event detection.

III. CONGESTION AND DELAY AWARE ROUTING

Congestion and Delay Aware Routing (CODAR) protocol which tries to improve end-to-end data success rate of nodes near to an event and also tries to reduce the latency of these time critical data. It considers a static WSN with a single BS where nodes generate monitoring or regular data with a low generation rate. When some nodes sense a critical event, they generate critical data with a high generation rate. To detect the event successfully, the BS needs to receive a high number of critical data packets. Moreover, each critical data packet must reach the BS within certain time after its generation. Delay of regular data arriving at the BS is not detrimental. Critical data generating nodes are called critical nodes and other nodes are called regular nodes. When the event is no longer sensible, the critical nodes become regular nodes.

A. Congestion Avoidance

All nodes have the same fixed transmission power as scheduled MAC (like TDMA) protocols may prolong the latency of event notification at the BS, we consider contention based MAC protocols. Like CAM [8], a

utility function *f* is defined in (1). When any node *B* forwards a packet, it chooses the highest *f*-valued node among its neighbours.

$$f(k) = \alpha \times \frac{D_k}{D} + (1-\alpha) \times RSR_k$$

where *Dk* is the distance of the next node *k* towards the BS from node *B*, *D* is the maximum distance that can be covered by the transmission power of each node, *RSRk* is the relative success rate of node *k* defined as the ratio of the number of packets transmitted from MAC layer to the number of packets forwarded from network layer to MAC layer over a small period, and $0 < \alpha < 1$. To calculate *f* for each neighbour *k*, node *B* needs the values of *Dk* and *RSRk*. CODAR assumes that each node knows its location. Each node *k* can broadcast its locationand *RSR* value using control packets after receiving a fixed number of packets from other nodes or after a fixed interval whichever is earlier. Using own and *k*'s locations, *B* can calculate *Dk*. The last term in *f* helps to reduce congestion formation by choosing lightly congested nodes. In CAM [8], it is analytically shown that the distance parameter in (1) ensures high end-to-end data success rate at the BS. But this is true when nodes are deployed with uniform node density. Having uniform node density in inaccessible fields is challenging. We proposed a novel approach in section III-C for ensuring uniform node density in inaccessible areas.

B. End-to-end Delivery Delay Management

Each node has measurements of end-to-end delays of its data packets. A node can easily determine the delay of a packet 358 (queuing delay plus medium access delay) inside it. As sensor nodes are densely deployed and the communication range is very small, propagation delay can be ignored. MAC layer records the current time TC when a packet arrives from higher layer. Subsequently, the packet is transmitted and MAC layer receives acknowledgement form receiving node for this packet at time TA. Now, total delay of this packet inside this node is TA - TC – TACK – TP, where TACK is time duration of acknowledgement packet and TP is the processing time, i.e., the interval between the receipt of the packet at the receiving node and transmission of acknowledgement. Of course, TP is dependent on the particular MAC protocol in use. The nodes having direct communication to the BS get their end-to-end delays in this way. With other information, they periodically broadcast delay data using the control packets (section III-A). After receiving these control packets, other nodes can calculate their end-to-end delays by adding their own delays to their neighbours' end-to-end delays. By continuing this process, finally all nodes get their end-to-end delays. Each critical data packet has a header field that indicates its deadline by which it should reach the BS. The deadline field is set by the source node and its presence indicates that the packet is critical. All intermediate nodes check this field before forwarding the packet. If an intermediate node has end-to-end delay that cannot meet packet's deadline then the node simply drops the packet to reduce unnecessary futile packets. The routing part of CODAR is described in ALGORITHM I. It is possible that each node can estimate its end-to-end delay in case of particular MAC and routing protocols under a specified node density, and it will eliminate the necessity of measuring and broadcasting of practical delay values. But the problem with this estimation is that it is only an *average value* not an *instantaneous* measurement. When a critical event occurs, the nodes need to ensure instantaneous (not average) timeliness of critical packets. On the contrary, in case of practical instantaneous delay measurement, the performance of each node will highly depend on the successful receipt of delay measurements of its neighbours. In sub-section F, we developed a technique to ensure high success probability of control messages carrying the delay measurements.

C. Ensuring Uniform Node-density

Although planned deployments of nodes in accessible environments are possible, inaccessible areas often require aerial dropping of sensor nodes in which case uniform node density achievement is a challenging task. One novel solution is first to deploy nodes with a high density. Each node then finds its position through localization techniques [13, 14] and sends its location to the BS. After getting all nodes' locations, the BS will select the necessary nodes to be active to cover the maximum possible sensing area and will instruct other nodes to remain inactive. The network will now operate only with active nodes ensuring coverage while achieving approximately uniform node-density. Still this process may leave some void spots creating connectivity problem. Deployment with high density of nodes will solve the connectivity.

ALGORITHM I. CONGESTION AND DELAY AWARE ROUTING

```
procedure CODAR_Routing (Neighbour-list NL, Distance- list DL,
                         Success-list RSRL)

  local variables: Node R, Node k, real f

  if the packet in hand has a deadline that cannot be met by this node's
    end-to-end delay
      drop the packet
      exit procedure

  Remove all nodes from NL that cannot meet current
  packet's deadline

  for each node k in NL
      f = α × DI(k)/D + (1 − α) × RSRL(k)

  Let, node R has the highest value of f among all nodes in NL

  return R
```

problem, but it cannot eliminate it totally. There will be some spots where node-density would be insufficient causing connectivity disruption of some parts of the network to the BS. To solve this problem we can deploy nodes with communication range higher than their sensing range. This solution is also helpful in case of planned deployment where both ranges might be equal. If nodes have higher communication range then active nodes remain connected to BS after energy depletion of some intermediate nodes. Of course, the network will lose some sensing coverage. Empirical study can determine the required ratio of communication range to sensing range for different sizes of sensing fields. CODAR desires uniform node density and this technique is a part of network setup phase only. It may be noted that uniform node density is desirable but not mandatory in the operation of CODAR.

D. *Congestion Mitigation*

Network layer calculates *RSR* with the feedback from MAC layer and sends the value of *RSR* to the application layer. If the value of *RSR* is less than 1, the application layer reduces its data generation rate to *RSR* factor of the current rate. If the application layer has a lower data generation rate than its targeted rate and the value of *RSR* is 1 (which is the maximum possible value), then it increases its data generation rate by a small factor (10% of current rate). In this way, the application layer always maintains its targeted data generation rate when there is no congestion. For packets coming from other nodes, network layer simply forwards *RSR* factor of the packets to MAC layer and drops the remaining ones. While dropping packets coming from other nodes, the network layer tries to forward as many critical packets (packets sent by nodes close to the event) as possible so that the BS can get the maximum number of critical packets to be able to detect the event reliably and timely. If the network layer finds a deadline field in a packet then it understands that it is a critical packet.

E. *MAC Layer Queue Management*

The aim of CODAR is to deliver high amount of critical data within specified delays. When a critical data packet comes from network layer to MAC layer, the packet is inserted in the transmission queue according to its deadline with the packet having earliest deadline at the front. Each regular data packet is appended at the rear of the same queue. This technique will reduce the average delay of critical packets although it will increase the delay of regular packets at the same time. Each control packet (that includes location, delay and *RSR* information of a node) is placed exactly at the front of the queue.

F. *Improving the Success Probability of Control Data*

The performance of CODAR highly relies on successful receipt of control packets broadcast by different nodes. Techniques to improve the success probability of these control packets are specific to a particular MAC layer protocol. Our proposition is so far valid for any contention-based MAC protocol. But this subsection is only devoted to a widely used MAC protocol, IEEE 802.11 DCF. For other MAC protocols, similar techniques might be available. We have a very simple but highly effective technique to improve the success probability of control packets. When a node decides to broadcast a control packet, it first broadcast it using the protocol described in IEEE 802.11 DCF. After completing its transmission, the same node waits for Short Inter-Frame Space (SIFS) time and then immediately retransmits the same control packet. We can consider two neighbouring nodes *S* and *R* who have collision domains called *CS* and *CR* respectively (Fig. 1). In CODAR, nodes periodically broadcast control packets, but independently of one another. Therefore, there is a little chance that node *S* and any other node in *CS* (including *R*) will broadcast control packets at the same time. If *S* and *R* transmit control packets at the same time, then they both will retransmit packets after SIFS time and in this case, none's control packet will succeed. But there is a high probability that only node *S* in *CS* sends a control packet at a time. This first control packet may easily collide with other node's data packet. After collision, other than *S*, all nodes wait for Distributed Inter-Frame Space (DIFS) time before transmitting any data. As *S* retransmits control packet after SIFS time (SIFS < DIFS), all other nodes in *CS* can sense that transmission and they all wait until *S* finishes its transmission. So control packet of *S* will not be destroyed by other nodes' data inside *CS*. But not all neighbours of *S* will successfully receive that control packet. Node *T* is in *CR* but not in *CS*. *T* cannot sense *S*'s transmission. Therefore, it can transmit both times when *S* broadcast the control packet resulting in garbled receipt of *R* about *S*'s control packet. But this

probability is low and our simple technique will significantly improve success probability of the control packets.

IV. PERFORMANCE EVALUATION

To evaluate the performance of CODAR, we compared it with similar protocol CAM [8]. CAM provides high success rate of critical data in the presence of regular data through congestion avoidance and mitigation. But CAM considers critical data as delay insensitive which is not true in case of critical events. CODAR ensures high success rate of delay sensitive critical data.

A. Simulation Environment

We have run simulations using OPNET [15] modeler software. We placed 196 sensor nodes in an area of 345 metre × 345 metre with uniform node distribution. One BS is placed at location (345, 172). Desired node density is achieved by applying the technique described in section III-C. All nodes have a fixed transmission power. We considered ideal environment (i.e., no obstacle) and also considered energy expenditure only during transmission (as energy loss during reception is low). We employed 1 Mbps IEEE 802.11 DCF MAC protocol. An event occurs at location (50, 50). All nodes within 40-metre radius of this location generate critical data with a high generation rate. The remaining nodes generate regular data with a low generation rate. In CODAR, we set α =0.50.

B. Simulation Results

We conducted three different types of experiments. In first type of experiments, transmission range of each node and the critical data generation rate are kept fixed whereas the regular data generation rate is varied. Data generation rate is denoted by *pps* (packets per second). Transmission range of nodes and the regular data generation rate remain constant while the critical data generation rate is varied in second type of experiments. In third type of experiments, both data generation rates are constant and transmission range of nodes is varied. In each experiment, packet success rates and average packet delays of both regular and critical data, and the maximum node energy consumption are measured. As we are interested in reliable and timely event detection and we have limited space here, we showed packet success rate and average packet delay of critical data, and maximum node energy spent in the network under different conditions. Lifetime of the network is limited by the maximum energy used by any node in the network. Fig. 2(a) shows average packet success rate of critical data at BS from all critical nodes under different generation rates of regular data where data generation rate of each critical node is 15 pps and transmission range of all nodes is 80 metres. Fig. 2(b) shows average delay of critical data at BS and Fig. 2(c)shows maximum node energy spent in the network under same conditions. With the increase in regular data generation rate, the congestion status of all nodes deteriorates which causes higher number of packet drops and eventually reduces the average success rate of critical data at the BS as shown in Fig. 2(a). The average delay of critical data at the BS is also increased in Fig. 2(b) because worse congestion condition causes higher number of transmission & attempts for packets at each node. In both Fig. 2(a)

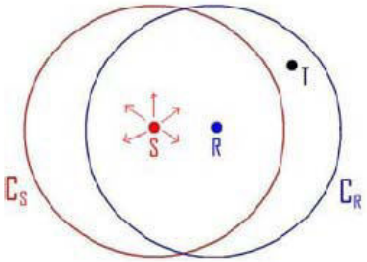
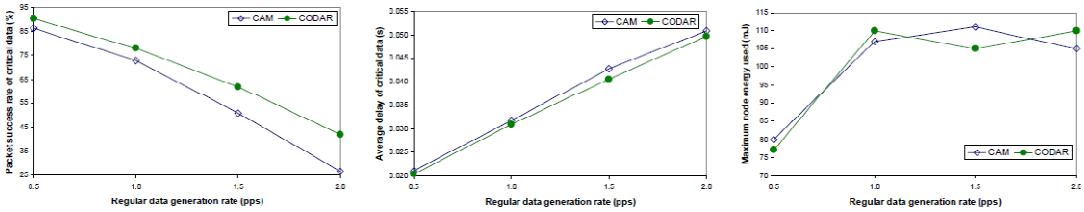


Figure 1. Node S broadcasts a control packet

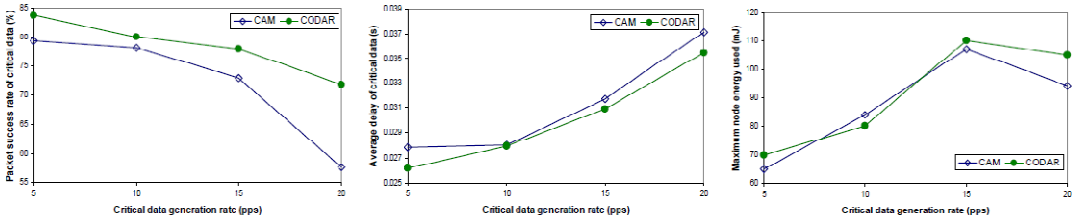
and 2(b), from each node to the BS, and it drops a particular data packet when the current node cannot meet the deadline of that packet. In this way, CODAR reduces congestion caused by unfruitful data transmission. As a result, CODAR has lower congestion than CAM and consequently it has 4.5% to 56.43% higher packet success rate and also a lower packet delay of critical data than those of CAM. Maximum node energy spent by any node in the network as shown in Fig. 2(c) increases by increased congestion in the network caused by higher generation rate of regular data. In Fig. 2(c), maximum node energies used in both CAM and CODAR are comparable. Fig. 3(a)-3(c) show same three parameters as in Fig. 2(a)-2(c) under different data generation rates of critical nodes while data generation rate of each regular node is 1 pps and transmission range of all nodes is 80 metres. As congestion in the network is increased by increased generation rate of critical data, the average success rate of critical data at the BS decreases and the average delay of critical data at the BS increases as shown in Fig. 3(a) and Fig. 3(b) respectively. Due to the congestion reduction through the dropping of unreachable packets in CODAR, it has a lower packet delay and 2.5% to 24.5% higher success rate of critical data delivery at the BS than those found in CAM. Fig. 3(c) shows that maximum node energies used in CAM and CODAR are comparable in this case also. Fig. 4(a)-4(c) show same parameters under different transmission ranges of nodes while data generation rates of regular and critical nodes are kept fixed at 1 pps and 15 pps respectively. When the transmission range of nodes is increased, the number

of nodes in collision domain of each transmitting node is also increased. This causes lower success probability and higher delay of transmission in each hop. But due to higher transmission range, the total number of hops for each data packet to reach the BS is reduced which tends to increase overall success rate and reduce total delay at the BS. Energy required to transmit a packet increases with increased transmission range of the node.



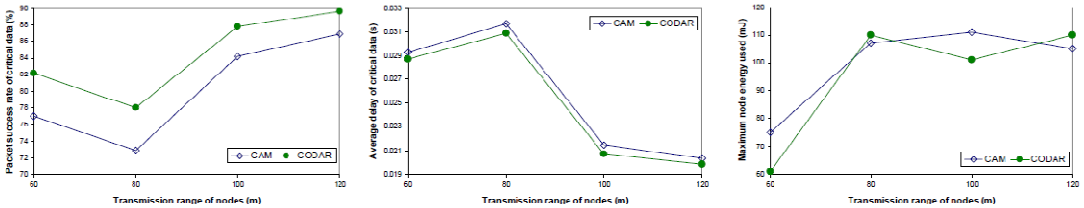
(a) (b) (c)

Figure 2: (a) Packet success rate of critical data, (b) average delay of critical data, (c) maximum node energy used, under different generation rates of regular data (critical data generation rate = 15 pps and transmission range of nodes = 80m)



(a) (b) (c)

Figure 3: (a) Packet success rate of critical data, (b) average delay of critical data, (c) maximum node energy used, under different generation rates of critical data (regular data generation rate = 1 pps and transmission range of nodes = 80m)



(a) (b) (c)

Figure 4: (a) Packet success rate of critical data, (b) average delay of critical data, (c) maximum node energy used, under different transmission ranges of nodes (critical data generation rate = 15 pps, regular data generation rate = 1 pps)

Therefore, the maximum node energy spent by any node should increase as it is found in Fig.4(c), and also the performances of CODAR and CAM is again comparable here. CODAR has a lower packet delay and 3.2% to 7.1% higher success rate of critical data at the BS than those found in CAM. But there are abnormal behaviours in Fig. 4(a) and Fig. 4(b) when transmission range changes from 60 metres to 80 metres. Regular data generation rate is 1 pps which is low compared to critical data generation rate of 15 pps. As a result, when the collision domain of a node includes more number of regular nodes, the success probability of its transmission is affected by a little. But when the collision domain of a node includes more number of critical nodes, the success probability is highly reduced. In our experiment, all critical nodes reside in a region having a diameter of 80 metres. When transmission range increased from 60 metres to 80 meters, the collision domain of each critical node contains a large number of critical nodes which reduces success probability and also increases transmission delay of each critical node significantly. Therefore, performances deteriorate in Fig. 4(a) and Fig. 4(b) when transmission range increases from 60 metres to 80 metres. But when transmission range increases above 80 metres, collision domain of each critical node adds more number of regular nodes only, which does

not affect performance significantly. Moreover, due to a lower number of transmissions required to transmit a packet from a critical node to the BS as a result of higher transmission range, end-to-end success probability and end-to-end delay of critical data at the BS is increased and reduced respectively. Therefore, higher transmission range above 80 metres gradually increases end-to-end success probability and reduces end-to-end delay of

critical data at the BS. In all experiments, CODAR achieves more number of critical data packets at the BS within lower average packet delay than that achievable in CAM. This achievement of CODAR yields more reliable and timely detection of critical events.

V. CONCLUSION

Congestion and Delay Aware Routing (CODAR) protocol presented in this paper has the potential to reduce congestion by avoiding congested nodes during route selection process and also by dropping of futile data packets. It provides high success rate by accurately adjusting data rate of a node during congestion mitigation. In achieving its successs, CODAR utilizes congestion parameters into routing decision and at the same time, it works in a distributed manner as it needs control data only from neighbouring nodes. It also endeavours to provide better success rate of control packets which increases its reliability. Simulation results show that CODAR provides significantly high success rate and low average packet delay ofcritical data which eventually results in reliable and timely event detection. Future study will focus on analysing the effect of adaptive value of the weighting factor α in the utility function *f*. As individual nodes are located at different parts of the network, different values of α for different nodes might accurately sense the node's congestion level and which in turn might help better congestion management.

ACKNOWLEDGEMENT

The successful completion of any task would be incomplete without mention of the people who made it possible and support had been a constant source of encouragement which crowned my efforts with success. My special gratitude to **Dr. P A Vijaya,** HOD, department of E&C, MCE for her guidance, constant encouragement and wholehearted support. My sincere thanks to my guide **Dr. P C Srikanth,** Professor department of E&C, MCE .And my sincere thanks to my company guide, **Mr K K Goswami.,** Project manager, CBK infotech, Bangalore.

REFERENCES

[1] Mohammad Masumuzzaman Bhuiyan, Iqbal Gondal and Joarder Kamruzzaman, "CODAR: Congestion and Delay Aware Routing to Detect Time Critical Events in WSNs" ICOIN 2011

[2] G. Pottie, "Wireless sensor networks," In *Proc. of Information Theory Workshop* (San Diego, CA), June 1998, pp. 139-140.

[3] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, pp 102-114, August 2002.

[4] C. Wang, K. Sohraby, B. Li, M. Daneshmand and Y. Hu, "A survey of transport protocols for wireless sensor networks," *IEEE Network*, vol. 20, no. 3, pp 34-40, May-June 2006.

[5] C.-Y. Wan, S.B. Eisenman and A.T. Campbell, "CODA: congestion detection and avoidance in sensor networks". In *Proc. of ACM Sensys* (Los Angeles, California, USA), November 5–7, 2003, pp 266-279.

[6] C.-Y. Wan, S. B. Eisenman, A. T. Campbell and J. Crowcroft, "Siphon: overload traffic management using multi-radio virtual sinks in sensornetworks," In *Proc. of ACM SenSys* (San Diego, CA, USA), November 2–4, 2005.

[7] J. Kang, Y. Zhang and B. Nath, "TARA: topology-aware resource adaptation to alleviate congestion in sensor networks," *IEEE Transaction on Parallel and Distributed Systems*, vol. 18, no. 7, pp 919-931. July 2007.

[8] H. Zhou, X. Guan and C. Wu, "Reliable transport with memory consideration in wireless sensor networks," In *Proc. of IEEE International Conference on Communications* (Beijing, China), May 19- 23, 2008, pp 2819-2824.

[9] M. M. Bhuiyan, I. Gondal and J. Kamruzzaman, "CAM: congestion avoidance and mitigation in wireless sensor networks," In *Proc. of IEEE Vehicular Technology Conference* (*VTC2010-Spring*), May 2010.

[10] L. Cheng, C. Chen, J. Ma, L. Shu, H. Chen and L. T. Yang, "Residual Time Aware Forwarding for Randomly Duty-Cycled Wireless Sensor Networks," In *Proc. of International Conference on Computational Science and Engineering*, vol. 2, pp 79-86, June 2009.

[11] J. Heo, J. Hong and Y. Cho, "EARQ: Energy Aware Routing for Real- Time and Reliable

Communication in Wireless Industrial Sensor Networks," *IEEE Transactions on Industrial Informatics*, vol. 5, no. 1, FEBRUARY 2009, pp 3-11, February 2009.

[12] M. F. Munir, A. A. Kherani and F. Filali, "Distributed Algorithm for Minimizing Delay in Multi-Hop Wireless Sensor Networks," In *Proc. Of 7th International Symposium on Modeling and Optimization in Mobile,Ad Hoc, and Wireless Networks*, June 2009, pp 1-9.

[13] S. Jabbar, A. A. Minhas and R. A. Akhtar, "SPERT: a stateless protocol for energy-sensitive real-time routing for wireless sensor network," In *Proc. of International Conference on Information and Communication Technologies*, August 2009, pp 117-124.

[14] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," In *Proc. of Annu. Int. Conf. Mobile Computing and Networking*, 2003, pp. 81–95.

[15] K. Yedavalli, B. Krishnamachari, and L. Venkatraman, "Fast/fair mobile localization in infrastructure wireless sensor networks," *SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 11, no. 1, pp. 29–40, 2007.

[16] (2009) The OPNET Technologies, Inc. website. [Online]. Available: http://www.opnet.com/

◈◈◈

Comparative Study of DGA Methods Using MATLAB

Pallavi Gulabrao Lone & P.S. Swami

Electrical Engg. Dept. Government Engineering college,Aurangabad, Maharashtra,India

Abstract - Dissolved Gas Analysis (DGA) is one of the primary dignostic tool widely used by utilities for transformer condition monitering. DGA is the method used for detecting the incipient fault occuring in transformer which develops slowly. The detection of certain gases generated in oil filled transformer in service is frequently the first available indication of malfunctioning that may lead to ultimate failure of a transformer, if not corrected. DGA methods detect faults by considering the ratios of specific dissolved gas concentrations in the transformer mineral oil.There are various types of DGA methods used to access the condition of a transformer. Like Key gas method, Doernenburg's ratio method, Roger's ratio method, Total dissolved gas concentration, IEC method, Duval's triangle method, *etc*. This paper involves study of DGA methods, a Matlab program on DGA methods which gives the result for condition of transformer & also comparison of this results with the experimental data results from various utilities have been made.

Keywords- *Dissolved Gas Analysis (DGA), Incipient transformer faults, DGA methods, Matlab program.*

I. INTRODUCTION

The transformer is an essential device of a transmission and distribution system. The behavior of power transformer is very important in regards to the reliability of power systems. Failures of large power transformers are problematic due to, it can cause operational problems to the system, generally large power transformers are placed in large tanks containing flammable oil and are hazardous for environment, the utilities do not have a spare one in case of breakdown, and the various types of thermal and electric stresses often age the transformer and leads to incipient faults. The life of a transformer depends directly on the life of its insulation. The deteriorating insulation and ageing are the two major causes of incipient events. The major causes of oil deterioration are moisture and oxygen coupled with heat .

Various tests used for condition monitering of transformer are- Capacitance and Tan Delta test of Bushing & winding, Frequency Response Analysis, Winding resistance test, Turns Ratio test, Magnetic Balance test, Excitation Current, IR & PI Value Measurement test, Oil test as per IS: 1866, Dissolved Gas Analysis (DGA), Furan (oil test), Domino oil test (on line test), Core Insulation test, etc. In this DGA method is one of the significant, simple and efficient test that most of the time leads to accurate solution. Dissolved gases in oil-filled electrical equipment are very complex.

DGA is one of the dignostic tool used for detecting the incipient fault occuring in oil of inservice transformer. However, there is still no universally accepted interpretation technique for fault detection. Based on this preventive maintenance can be carried out & condition can be assessed. In this paper an attempt has been made to review the various DGA methods used to find the condition of the insulation within an inservice transformer. The gas ratios and relative proportions of gases are used to diagnose the fault. The rules of the fault diagnosis are derived from the case study. DGA methods identify faults by considering the ratios of specific dissolved gas concentrations in the transformer mineral oil [3].

II. DGA METHODS

IEEE Standard C57.104-2008 describes the key gases, Doernenburg ratios, and Rogers Ratio method. IEC Standard 60599 introduces the three basic gas ratio methods and the Duval triangle method. DGA demonstrates better performance when a combination of gas ratios and relative proportions of gases are used to identify the fault.

Arcing, corona discharge, low energy sparking, overheating of insulation due to severe overloading, failure of forced cooling systems are responsible for gas generation.The detection of gases in oil filled transformer in service is frequently the first available indication of malfunctioning. These gases are hydrogen, methane, ethylene, ethane, carbonmonoxide, acetylene, carbondioxide. There are various types of DGA methods used by organizations and utilities to access the condition of a transformer [3].

A. Doernenburg's Ratio Method

This method is based on thermal degradation principles. In this method the gas concentrations in ppm (parts per millions) for H2, , CH4 , C2H2 , and C2H4 exceeds twice the value for fixed limits of each gas and for gases CO, and C2H6 exceeds thrice the value for fixed limits is used to determine the validity of the ratio procedure as shown in table I . Then the validity of the four ratios is determined, each ratio is compared with certain values and finally if all four succeeding ratios for a specific fault type fall within the predetermined values, then the suggested diagnosis is said to be valid as shown in table II . Method is a part of ANSI/IEEE Standard C57.104-1991 [8]. Implementation of this method may result in significant number of "no interpretation," due to incompleteness of the ratio ranges [3].

TABLE I. FIXED LIMITS OF GASES

Key Gas	Concentration L1 (ppm) [parts per millions]
Hydrogen(H2)	100
Methane(CH4)	120
Carbon Monoxide(CO)	350
Acetylene(C2H2)	35
Ethylene(C2H4)	50
Ethane(C2H6)	65

TABLE II. RATIOS OF GASES FOR DOERNENBURGS METHOD

Suggested Fault Diagnosis	Ratios For Key Gases			
	Main Ratio		*Auxiliary Ratio*	
	CH4/H2	*C2H2/C2H4*	*C2H6/C2H2*	*C2H2/CH4*
Thermal Decomposition	>1	<0.75	>0.4	<0.3
Corona (Low Intensity PD)	<0.1	Not Signif.	>0.4	<0.3
Arcing(High Intensity PD)	<1e>0.1	>0.75	<0.4	>0.3

B. *Roger's Ratio Method*

The Rogers ratio is based on ranges of ratios, which is used for diagnoses of faults as shown in table III] . The four conditions of the oil insulated transformer that are detectable are normal ageing, partial discharge with or without tracking, electrical and thermal faults of various degree of severity as shown in table IV . This method is also based on thermal degradation principles and is also included in ANSI/IEEE Standard C57.104-1991 [8]. There are values of ratios that do not fit into the diagnostic codes. Also for dissolved gases below the normal concentration have no consideration and due to this the exact implementation lead to many misinterpreted cases [3].

C. IEC Method

Fault diagnosis scheme recommended by International Electrotechnical Commission (IEC) originated from Roger's method except that the ratio C2H6/CH4 was dropped since it only indicated a limited temperature range of decomposition as shown in table V [9] .

TABLE III. CODE FOR ROGER'S RATIO METHOD

Ratios	VARIATION	CODE
CH4/H2	<=0.1	5
	>0.1,<1	0
	>=1,<3	1
	>=3	2
C2H6/CH4	<1	0
	>=1	1
C2H4/C2H6	<1	0
	>=1,<3	1
	>=3	2
C2H2/C2H4	<0.5	0
	>=0.5,<3	1
	>=3	2

However, no attempt is made to identify both thermal and electrical faults into more precise subtypes as shown in table VI . The typical faults in power transformers are classified as partial discharges, discharges of low and high energy, and thermal faults of three degrees of severity depending upon the temperature of the fault [2].

D. Key Gas Method

This method focuses on which of the combustible gases is the largest constituent of the TDGC (Total Dissolve Gas Concentration). This gas is labelled the key gas and historical trending has led to the association of key gases to their corresponding faults. For example if Ethylene constitutes over 60% of the TDCG, it is apparent that the likely cause of this gas generation is Thermal fault in oil "Fig.1". In this fault type gases like ethylene & methane with small quantities of hydrogen & ethane are detected. Traces of acetylene are formed if fault is severe or involves electrical contacts. Similarly by analysing all gas concentrations and relative percent in the oil sample, often one can easily determine the key gas of that sample and deduce a likely fault [8].

TABLE IV. FAULTS DIAGNOSIS BASED ON CODES FOR ROGER'S RATIO METHOD

CH4/H2	C2H6/CH4	C2H4/C2H6	C2H2/C2H4	DIAGNOSIS
0	0	0	0	Normal Deterioration

5	0	0	0	Partial Discharge
1-2	0	0	0	Thermal Fault-Low Tension 1500c
1-2	1	0	0	Thermal Fault-150 degree celcius-200 degree celcius
0	1	0	0	Thermal Fault-200 degree celcius-300 degree celcius
0	0	1	0	Overheating in Cables
1	0	1	0	Circulating Currents in Windings
1	0	2	0	Circulating Currents in Tank & Core-Overheating in Connections
0	0	0	1	Continuous Discharge
0	0	1 - 2	1-2	Arcing(High Intensity)
0	0	2	2	Low Intensity Continuous Discharge
5	0	0	1-2	Partial Discharge Involving Solid Insulation

TABLE V. CODE FOR IEC METHOD

Ratio Code	Range	code
C2H2/C2H4 = l	<0.1	0
	0.1-1	1
	1-3	1
	>3	2
CH4/H2 = i	<0.1	1
	0.1-1	0
	1-3	2
	>3	2
C2H4/C2H6 = k	<0.1	0
	0.1-1	0
	1-3	1
	>3	2

However, it should also be noted that this method does have a high tendency to return 'No result' [2].

E. Duval's Triangle Method

M. Duval developed this method in the 1960s [2]. Dissolved gases should be assessed for their normality before being interpreted using Duval triangle as shown in table I. The three sides of the Triangle are expressed in triangular coordinates (P1, P2, P3) representing the relative proportions of CH4, C2H4 and C2H2, from 0 to 100 for each gas [9]. These three gases in ppm, CH4 = g1, C2H4 = g2 and C2H2 = g3, must be transformed into triangular coordinates before being plotted onto the triangle [4]. First the sum of these three values, g1+g2+g3, should be calculated and then the relative proportion of the three gases: P1 = %CH4 = 100 x gl/(gl+g2+g3), P2 = %C2H4 = 100 x g2/(g1+g2+g3), P3 = %C2H2 = 100 x g3/(g1+g2+g3). (Fig.2) This method always provides a diagnosis, with a low percentage of wrong result, fault diagnosis is performed based on visualisation [4].

TABLE VI. FAULTS DIAGNOSIS BASED ON CODES FOR IEC RATIO METHOD

l	**i**	**k**	**Characteristic Fault**
0	0	0	Normal Ageing
Not Signif.	1	0	Partial Discharge of low Energy Density
1	1	0	Partial Discharge of high Energy Density
1-2	0	1-2	Discharge of Low Energy (Continues Sparking)
1	0	2	Discharge of high Energy (Arc with Power Flow Through)
0	0	1	Thermal Fault <150 degree celcius
0	2	0	Thermal Fault 150-300 degree celcius
0	2	1	Thermal Fault 300-700 degree celcius
0	2	2	Thermal Fault >700 degree celcius

In "Fig.2", **PD-** Partial Discharge, **D1-** Discharges of Low Eneregy, **D2-** Discharges of High Energy, **DT-** an intermediate zone DT has been attributed to mixtures of electrical and thermal faults in the transformer, **T1-** Thermal Fault ($T>300^0C$), **T2-** Thermal Fault ($300<T<700^0C$), **T3-** Thermal Fault ($T>700^0C$).

F. TDGC ANALYSIS

By examining the amounts of dissolved gas within the oil and referencing these numbers to preset limits, as shown in table VII one can determine whether there is an abnormal level of gas in the oil that suggests a fault is occurring. *IEEE* Standard provides acceptable limits for individual gas concentrations and Total Dissolved Gas concentration (TDGC) levels. If gas levels are found to

be abnormal, the engineer classifies them into conditions and obtains the most likely fault condition & corresponding actions that should be taken when a transformer is found in abnormal condition as shown in table VIII.

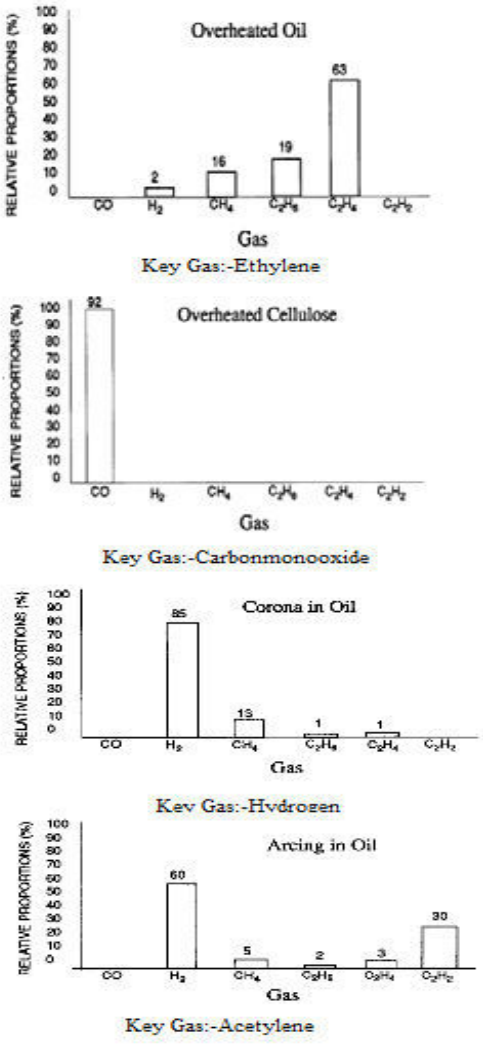


Figure 1. Key gas concentration formed during Particular fault for Key Gas Method

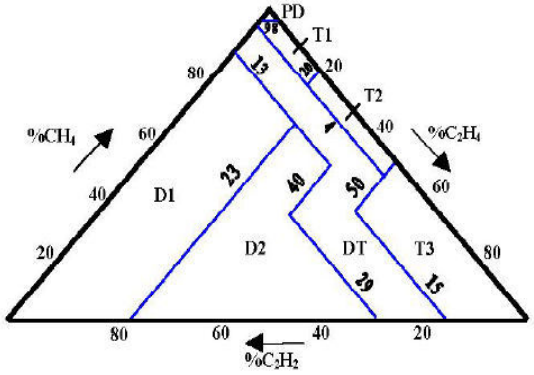


Figure 2. Duval's Triangle Method

If more than one sample has been taken from the transformer a generation rate (TDGC_rate) can be calculated in ppm/day. [7].

TABLE VII. GAS CONCENTRATION IN PPM FOR TDGC ANALYSIS

H2	CH4	C2H2	C2H4	C2H6	CO	TDCG
<100	<120	<35	<50	<65	<350	<720
101-700	121-400	36-50	51-100	66-100	351-570	721-1920
701-1800	401-1000	51-80	101-200	101-150	571-1400	1921-4630
>1800	>1000	>80	>200	>150	>1400	>4630

TABLE VIII. TDGC ANALYSIS

TDGC Levels *(ppm)*	TDGC Rates *(ppm/day)*	Sampling Intervals	Operating Procedures
>4630	>30	Daily	Consider removal from service
	10-30	Daily	
	<10	Weekly	Exercise extreme caution, analyze for indivdual gases
1921-4630	>30	Weekly	Exercise extreme caution, analyze for indivdual gases
	10-30	Weekly	
	<10	Monthly	
721-1920	>30	Monthly	Exercise extreme caution, analyze for indivdual gases
	10-30	Monthly	
	<10	Quarterly	
<=720	>30	Monthly	Exercise extreme caution, analyze for indivdual gases
	10-30	Quarterly	Continue normal operation
	<10	Annual	

III. PROGRAM STRUCTURE

I choose to write my program using the MATLAB software, as it has more inbuild functions which provide user friendly environment. The program commences by

asking for the variables. These include data such as gas levels in ppm (parts per millions), *etc.* Then this data is given to all diagnosis methods and the output is displayed.

IV. RESULT

The result of software program is compared with experimental data obtained from various utilities & displayed. Follwing table shows case studies whose results are compared. So that the recommended actions are carried out to avoid failure of a transformer as shown in table IX.

TABLE IX. CASE STUDY RESULTS

Casestudies/ variables & results	**Case1**	**Case2**	**Case3**
CH4	105	151	297
H2	98	168	99
C2H2	10	76	22
C2H4	360	99	241
C2H6	95	23	267
C0	158	351	210
TDGC rate	15	21	25
Doernenberg' Method	THERMAL DECOMP OSITION	ARCING-HIGH INTENSITY PARTIAL DISCHARGE	THERMAL DECOMPO SITION
Roger's Method	Circulating Currents in Tank & Core, Over heating Connection	Arcing (High Intensity)	Thermal Fault-Low Tension 150 degree celcius
IEC Method	Thermal Fault >700 degree celcius	Discharge of Low Energy . Discharge of high Energy (Arc with Power Flow Through)	Thermal Fault 150-300 degree celcius
Key gas Method	Ethylene is KeyGas in this sample which suggests a Thermal fault in oil.	Electrical anomaly unpredictable by KEYGAS RATIO METHOD	Electrical anomaly unpredic table by KEYGAS RATIO METHOD
Duval's triangle Method	T3-Thermal fault, T>700 °C	D2, DT	T2-Thermal fault, 300 < T < 700 °C
TDGC Analysis	sample monthly & exercise extreme caution, analyze for individual gases	sample monthly & exercise extreme caution, analyze for individual gases	sample monthly & exercise extreme caution, analyze for individual gases
Expt. Data results	core & tank circulating currents, over heated joints	Arcing	Thermal Fault

V. CONCLUSION

Analysis of results shows that program can be used to evaluate the condition of the transformer provided sufficient amount of reliable DGA data is available which has provided efficient solutions to incipient fault identification.The work will be useful to various agencies & utilities. For reliable results one should not rely on a single technique (method) when dealing with real life applications. In this paper, a comprehensive investigation of various methods for interpreting DGA results was carried out & the possibility of improving the diagnosis with aid of Matlab program was explored.

The program diagnoses the faults using the concentration levels of gases obtained from oil-filled power transformer in services. Futhermore, the rules and methods for DGA, derived based on the case study. But as we have all the interpretation, the results is likely to be incorrect for certain cases.

ACKNOWLEDGMENT

I take this opportunity to express my thanks to Mr. Dayanand; Mr.S.L. Deshmukh; Mr.Kadri; Mr.P.S.Swami; all my teachers for valuable help during this dissertation.

REFERENCES

[1] Xiaohui Li, Huaren Wu, Member, IEEE, and Danning Wu,"DGA interpretation scheme derived from caseStudy," IEEE Transaction on Power Delivery, VOL.26, NO.2, pp.1292-1293, APRIL 2011

[2] N.A. Muhamad, B.T. Phung and T.R. Blackburn,"Comparative studies and analysis of DGA methods for mineral oil using fuzzy logic," The 8th Internatinal Power Engineering Conference 2007, pp.1301, 2007

[3] Deepika Bhalla, Raj Kumar Bansal, and Hari Om Gupta, "Application of Artificial intelligence techniques for dissolved gas analysis of transformers -a review," World Academy of Science, Engineering and Technology 62 2010.

[4] A.Akbari*, A. Setayeshmehr, H. Borsi, E. Gockenbach ,"A software implementation of the duval triangle Method," IEEE Transaction 978-1-4244-2092-6/08,2008.

[5] C57.104.1991, I., IEEE Guide for Interpretation of Gases Generated in Oil-Immersed Transformer, I.

The Institute of Electrical and Electronic Engineers, Editor. 1992, The Institute of Electrical and Electronic Engineers, Inc p. 27.

[6] N. A. Muhamad, B. T. Phung, T. R. Blackburn, and K. X. Lai, "Comparative study and analysis of DGA methods for transformer mineral oil," in Proc. Power Tech, 2007, pp. 45–50..

[7] Sabina Karlsson,"A review of lifetime assessment of transformers and the use of dissolved gas analysis," XR-EE ETK 2007

[8] IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers, IEEE Standard C57.104-2008, Feb. 2009.

[9] "Interpretation of the analysis of gases in transformers and other oilfilled electrical equipment in service," IEC Publ. 60599, Mar. 1999

◈◈◈

Digital Filter Realization using Current Conveyor

Akashdeep Gautam & Tarun Kumar Rawat

Electronics & Communication, Netaji Subhas Institute of Technology. Delhi, India

Abstract - By using the concept of current conveyor, new 1st order and 2nd order digital filter sections are developed. These are then used as building blocks in a cascade synthesis. The proposed synthesis yields digital Equivalent of 1st order all-pass filter and general 2nd order filter structure.

Keywords- *Current Conveyor, wave equations, serial adaptor, parallel adaptor, all-pass filters., voltage source.*

I. INTRODUCTION

Two types of digital-filters synthesis procedures have been proposed in recent years. Direct procedures in which the discrete-time transfer function is realized directly as cascade, canonic, parallel or ladder structure and the indirect procedures in which an LC filter satisfying the desired specifications is simulated by employing the wave characterization [5-9].

Present paper describes an alternative cascade but indirect synthesis. The basis of synthesis is an analogue configuration comprising resistors, capacitors and inductors and second generation positive current conveyors. In this paper we describe the synthesis of digital equivalent of first order all pass filter and general second order filter structure, which can be used in the design of digital equivalent of LPF, HPF and BPF.

Current conveyor

Current conveyor is a three port device (Fig. 1). Three generations of current conveyor are available but the second generation positive current conveyor is extensively used in the design of analogue filters [1-4]. The second generation positive current conveyor is characterized by.

$$\begin{bmatrix} Ix \\ Vy \\ Iz \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & +1 & 0 \end{bmatrix} \begin{bmatrix} Vy \\ Ix \\ Vz \end{bmatrix} \qquad (1)$$

Digital equivalent of second generation positive current conveyor. An analogue port can be characterised by

$$A_k = V_k + I_k / G_k$$
$$B_k = V_k - I_k / G_k \qquad (2)$$

For k =1,2,…..n. where A_k and B_k are referred as the incident- and reflected-wave quantities respectively, G_k as the port conductance and V_k and I_k are the input voltage and input current at the port respectively. In the interconnection of n-ports, two interconnected ports must be assigned the same conductance so as to maintain continuity in the wave flow. Otherwise G_k can be chosen arbitrarily.

On the basis of wave characterization, digital realizations are readily available for conductances, voltage sources and wire interconnections [5,6]. With $G_1 = G_r$ in Figs. 3 and 4, B_1=0 in the first case and $A_1 = K_r V_i$ in the second. Hence, a conductance reduces to a digital sink and a voltage source of a given intensity to a digital source of same intensity [5,6]. Similarly, for a three port wire interconnection of Fig. 4 the digital realization of Fig. 5 (known as parallel adaptor) can be obtained and the multiplier co-efficient of parallel adaptor is given by

$m_k = 2G_k/(G_1 + G_2 + G_3)$ for k = 1, 2.

The same technique can be extended to second generation positive current conveyor of Fig. 1. As the second generation positive current conveyor has three set of equations as given by the matrix Eq.(1). By transforming variables A_k, B_k, V_k and I_k and then using Eq. (1) and (2), we can show that

$$B_x = 2A_y - A_x \qquad (3)$$
$$B_y = A_y \qquad (4)$$
$$Bz = A_z + (2R_z/R_x)(A_y - A_x) \qquad (5)$$

Define, $p = (2R_z/R_x)$

With $R_x = 2R_z$, Eq.(5) simplifies to

$$B_z = A_z + A_y - A_x \qquad (6)$$

And,

$$p = 1$$

Now, using the Eqs. (3), (4), (6), the digital realisation for the second generation positive current conveyor can be deduced as in Fig. 6.

II. DIGITAL REALIZATION OF ALL PASS FILTER

Considering the all pass filter of Fig. 7, having transfer function.

$$H(s) = \frac{(G - sC)}{(G + sC)} \qquad (7)$$

The individual analogue components and the corresponding ports can be identified [14] as shown in Fig. 8. On assigning the port conductances as given below.

For the first current conveyor-

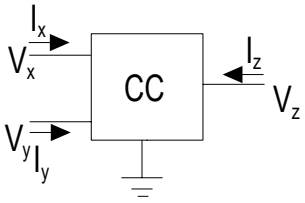


Fig.1. Current Conveyor

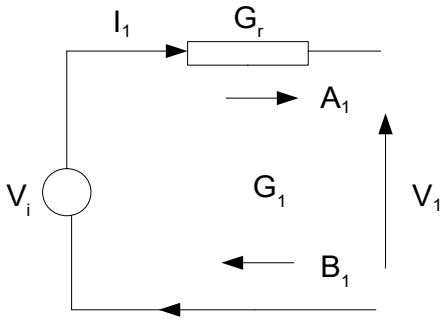


Fig.2. Conductance

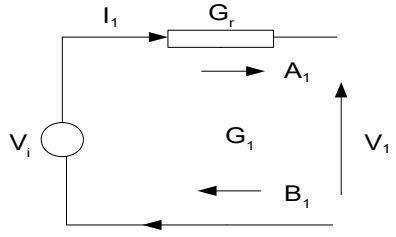


Fig.3. Voltage Source

$$G_{1x} = G + \frac{4C}{T},\ G_{1y} = G,\ G_{1z} = \frac{1}{2}(G + \frac{4C}{T}),$$

For the second current conveyor-

$$G_{2x} = 2(G + \frac{2C}{T}),\ G_{2y} = \frac{1}{2}(G + \frac{8C}{T}),\ G_{2z} = (G + \frac{2C}{T}),$$

For the first parallel adaptor-

$$G_{11} = \frac{1}{2}(G + \frac{4C}{T}),\ G_{2y} = \frac{1}{2}(G + \frac{8C}{T}),\ G_{2z} = \frac{2C}{T}$$

And the multiplier constants are

$$m_1 = \frac{(G + \frac{4C}{T})}{(G + \frac{8C}{T})},\quad m_2 = 1,$$

For the second parallel adaptor-

$$G_{21} = G,\ G_{22} = G + \frac{2C}{T},\ G_{23} = \frac{2C}{T},$$

And the multiplier constants are

$$m_1 = \frac{G}{(G + \frac{2C}{T})},\ m_2 = 1,$$

For the third parallel adaptor-

$$G_{31} = G + \frac{4C}{T},\ G_{32} = 2(G + \frac{2C}{T}),\ G_{33} = G,$$

And the multiplier constants are

$$m_1 = \frac{(G + \frac{4C}{T})}{(G + \frac{2C}{T})},\ m_2 = 1,$$

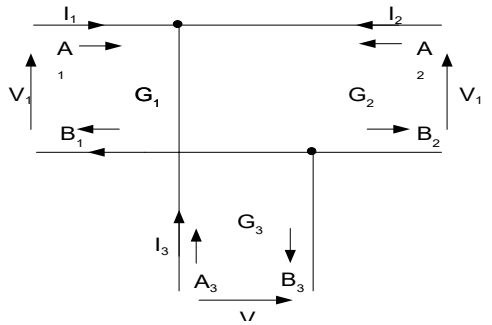


Fig.4. Parallel Adaptor

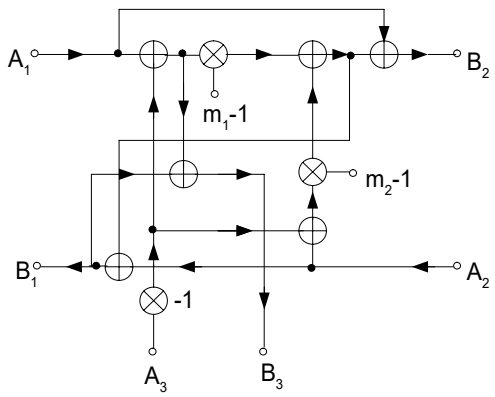


Fig.5 Digital equivalent of Parallel Adaptor

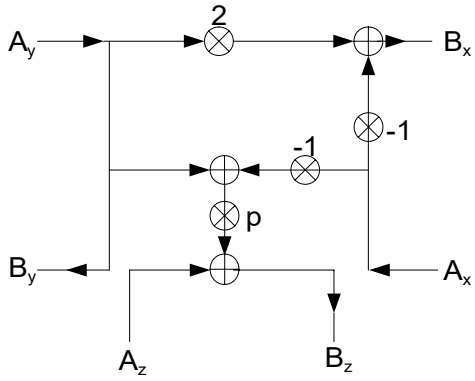


Fig.6 Digital equivalent of Current Conveyor

And now replacing the parallel wire interconnection and analogue current conveyor with its digital equivalent as given in Fig. 5 and 6, the first order digital all pass filter section is derived as shown in Fig. 9. The output proportional to V_o can be formed by using the adder at the output as shown in Fig. 8. The transfer function of the derived structure can be obtained from Eq. (2) and (8) as given by Eq. (10)

$$s \rightarrow \frac{2}{T}\frac{(z-1)}{(z+1)} \qquad (8)$$

$$H_D(z) = B_o / A_i = 2V_o / V_i \qquad (9)$$

$$H_D(z) = [2H\{s\}]_{s \rightarrow \frac{2}{T}\frac{(z-1)}{(z+1)}} \qquad (10)$$

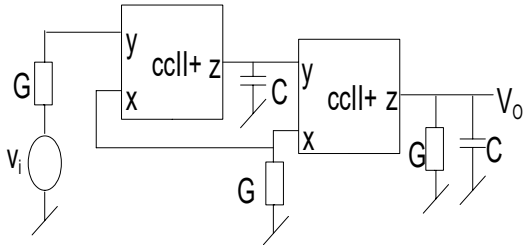


Fig.7. All pass filter of 1st order

III. DIGITAL REALIZATION OF GENERAL SECOND ORDER FILTER

We can also have the digital realization of second order filter of Fig. 10 using the method described above. The output is taken across the admittance G_4 as shown in Fig. 10,

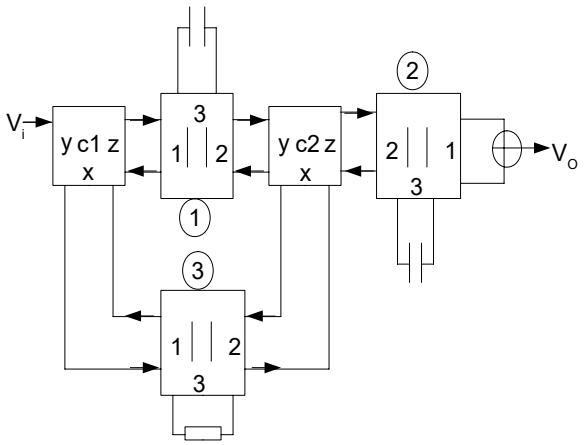


Fig.8. Port Identification of Fig.7

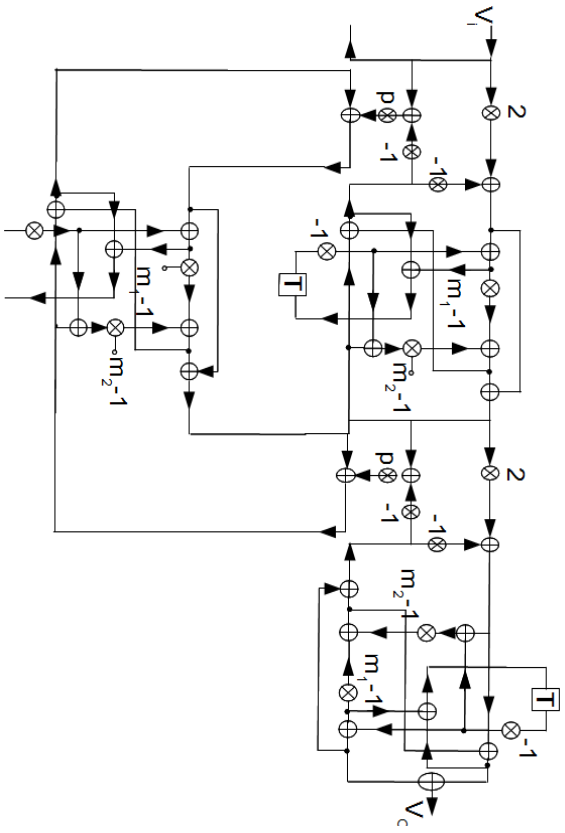


Fig.9. Digital equivalent of Fig.7

IV. DIGITAL REALIZATION OF GENERAL SECOND ORDER FILTER

We can also have the digital realization of second order filter of Fig. 10 using the method described above. The output is taken across the admittance G_4 as shown in Fig. 10, and the voltage V'_o follows the voltage V_o as described by the Eq. (1). The transfer function of the filter structure in Fig. 10 is given by

$$H(s) = V'_o / V_i = Y_1Y_2 / (Y_1Y_3+G_4Y_5) \quad (11)$$

As the voltage V_o follows the voltage V'_o so

$$H(s)=V'_o / V_i = V'_o / V_i = Y_1Y_2 / (Y_1Y_3+G_4Y_5) \quad (12)$$

By selection of different components for Y_3,Y_2,Y_1 and Y_5 we can have LPF, BPF and HPF [12]. The individual analogue components and corresponding ports can be identified [14] as shown in Fig. 11, and the port conductance of various ports are given below.

For the first current conveyor

$$G_{1x} = Y_2 + (Y_5/2),\ G_{1y} = G,\ G_{1z} = \frac{1}{2}(Y_2 + (Y_5/2)),$$

For the second current conveyor

$$G_{2x} = G_4,\ G_{2y} = Y_3 + \frac{1}{2}(Y_2 + (Y_5/2)),\ G_{2z} = (G_4/2),$$

For the third current conveyor

$$G_{3x} = Y_5,\ G_{3y} = Y_1 + (G_4/2),\ G_{3z} = (Y_5/2),$$

For the first parallel adaptor

$$G_{11} = \frac{1}{2}(Y_2 + (Y_5/2)),\ G_{12} = Y_3 + \frac{1}{2}(Y_2 + (Y_5/2)),\ G_{13} = Y_3,$$

And the multiplier constants are

$$m_1 = (\frac{1}{2}(Y_2 + (Y_5/2))) / (Y_3 + \frac{1}{2}(Y_2 + (Y_5/2))),\ m_2 = 1,$$

For the second parallel adaptor

$$G_{21} = (G_4/2),\ G_{22} = Y_1 + (G_4/2),\ G_{23} = Y_1,$$

And the multiplier constants are

$$m_1 = (G_4/2) / (Y_1 + (G_4/2)),\ m_2 = 1,$$

For the third parallel adaptor

$$G_{31} = (Y_5/2),\ G_{32} = Y_2 + (Y_5/2),\ G_{33} = (Y_5/2)$$

And the multiplier constants are

$$m_1 = (Y_5/2) / (Y_2 + (Y_5/2)),\ m_2 = 1,$$

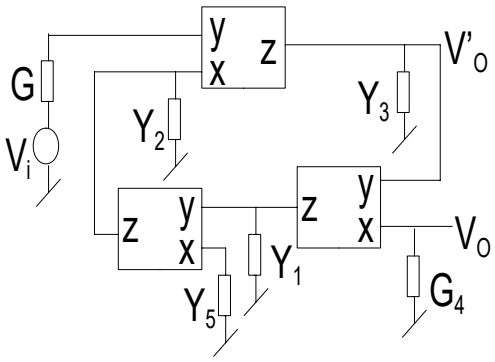


Fig.10. General second order filter structure

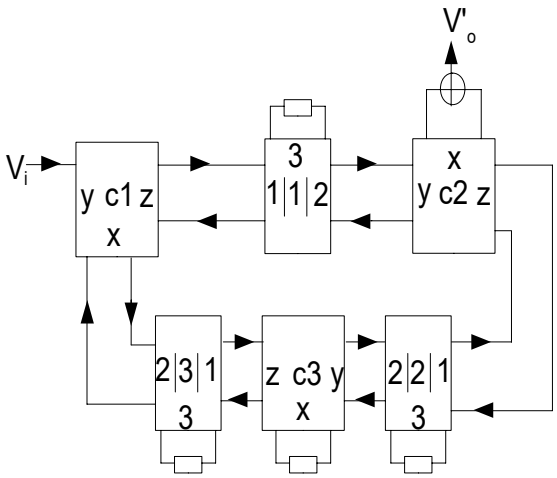


Fig.11. Port identification of Fig.10

And now replacing the parallel wire interconnection and analogue current conveyor with its digital equivalent as given in Fig. 5 and 6, the digital equivalent of general second order filter is derived as shown in Fig. 12. An output proportional to V_o can be formed by putting an adder at the output as shown in Fig. 11. The transfer function of the derived structure [14] can be obtained from the Eq. (2) and (8) given as

$$H_D(z) = [2H\{s\}]_{s\rightarrow} \frac{2}{T}\frac{(z-1)}{(z+1)} \qquad (13)$$

By selection of different components for Y_3, Y_2, G_4 and Y_5 and then replacing these components with their digital equivalent [14] we can have digital equivalent of the corresponding filter.

IV. CONCLUSION

By using the concept of second generation positive current conveyor we derived the digital equivalent of first order and second order filters. So, it can be used as a new and easy mathematical tool for the synthesis of digital filters.

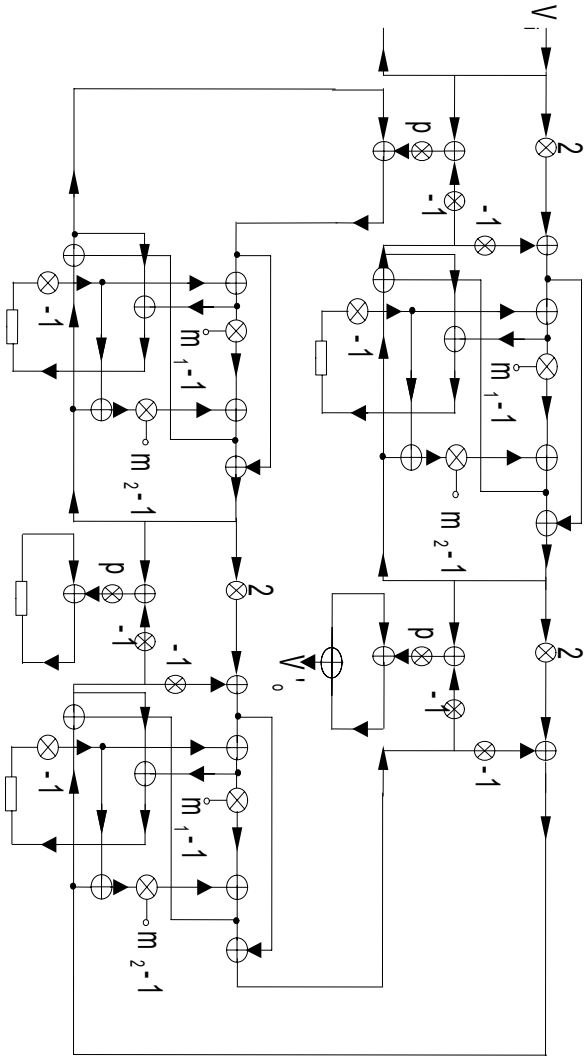


Fig.12. Digital equivalent of general secon order filter structure

REFERENCES

Capitalize only the first word in a paper title, except for proper nouns and element symbols. For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

[1] K.C.Smith and A.S.Sedra, "'The current conveyor - a new building block", Proc. IEEE, vol. 56, pp.1368-1369, Aug-1968

[2] A.S.Sedra and K.C.Smiand, "A sccond generation current conveyor and its application" , IEEE Transactions on Circuit Theory,Vol. CT. 17, pp. 132-134, Feb 1970.

[3] A.S.Sedra, ”A new approach to active network synthesi”, 'Ph.D Thesis, University of Totonto, 1969.

[4] K.C.Smith and A.S.Sedra, “Realization of chuafamily of new non-linear network elements using the current conveyor”, IEEE Transactions on circuit theory, VOl. CT.17, pp. 137-139, Feb 1970.

[5] FETTWEIS,A.: “Digital filter structures related to classical filters”, Arch., Elektron., 1971, 25, pp. 79-89.

[6] FETTWEIS,A., and SEDLMEYER,A.: ”'Digital filters with true ladder configuration “, Int. J. Circuit Theory and Applications,} 1971, 1, pp. 5-10.

[7] FETTWEIS,A.: ”'Wave digital filters with reduced number of delays”, ibid.} 1974, 2, pp. 319-330..

[8] CONSTANTINIDES, AG.: “Alternative approach to design of wave digital filters”, Electron. Lett., 1974, 10, pp. 59-60.

[9] SWAMY. M.N.S., AND THYAGARAJAN, K.S., ”A new type of wave digital filter”, Procedings of IEEE Symposium on the Circuits and systems, April 1975, pp. 174-178.D.

[10] K. Pal and R. Singh, “ Inductorless current conveyor all pass filter using grounded capacitors”, Electronics letters Vol 18, pp. 47, Jan 1982.

[11] Oguzhan Cicekoglu and H. Hakan Kuntman, “CCII+ based first order all pass filters with all grounded passive elements”,

[12] Cicekoglu, Oguzhan, and Ak, Ersin, Seref: “ Simple multifunction filter realizations with current conveyors”, December 1998.

[13] ANTONIOU, A., and Rezk, M.G., “Digital filter synthesis using concept of generalised-immitance converte”, Electronic circuits ans systems, November, 1977 vol. 1, No. 6.

[14] Digital Signal Processing by Andreas Antoniou.

◈◈◈

Fingerprint Texture Identification and Face Recognition Multimodal System

Poonam Mote & P. H. Zope

(Department of E & Tc Engineering, North Maharashtra University, Jalgaon

Abstract - A variety of biometric authentication systems are available to identify the person. Now a day for giving secure access to different systems Biometric authentication is necessary. In this paper we propose the multimodal biometric system using two traits i.e. face and fingerprint. The final decision is made by feature level fusion. Feature extraction is based on Gabor filter for fingerprint as well face. In the proposed system the stored feature dataset is update every time hence the proposed system is more reliable than the others. As well as with an accurate authentication system keep the record of login and logout time with total time spends of the user. This system is tested with the three standard data sets of face & FVC2002 &FVC2004 datasets of fingerprint. The proposed system has lower computational complexity and higher accuracy.

Keywords *– Gabor filter, Face recognition,, Fingerprint recognition, Fusion,Mutimodal biometrics.*

I. INTRODUCTION

In recent years, biometrics authentication has seen considerable improvements in reliability and accuracy, with some of the traits offering good performance. However, even the best biometric traits till date are facing numerous problems; some of them are inherent to the technology itself. In particular, biometric authentication systems generally suffer from enrollment problems due to non-universal biometric traits, susceptibility to biometric spoofing or insufficient accuracy caused by noisy data acquisition in certain environments, hence to overcome these problems is the use of multi-biometrics.

Many researchers have used faces and fingerprint, with some considering score quality when fusing results [3].A multimodal biometric system based on fusion of face and fingerprint in [11] introduced and compared different fusion methods. In this case, the fusion was using data quality information, it outperforms unimodal systems. [1] presents an efficient authentification model using multimodal face and fingerprint on space-limited tokens. The images of the face was encrypted and encoded into fingerprint images. This work proves its robustness to other attacks and proves genuineness. This system, aims at solving the problem of dynamic face authentification, it has three modules: enrolment, image enhancement and fusion, it features distinguishes customers and imposters, increases performances and presents more genericity.Our goal is to perform authentication using multimodal Biometrics, which combine multiple traits to establish identity with high accuracy.

The corresponding output obtained by using Gabor filter is good as compared to the other methods. Gabor filter have the properties of spatial localization, orientation selectivity and spatial-frequency selectivity.Therefore, Gabor filter have been applied to many fields, such as texture classification, face recognition, handwritten character recognition, fingerprint classification and fingerprint recognition. It handles sensitively the different orientations in the fingerprint image and it provide a robust representation is with respect to minor local changes thus, individuals can be recognized in spite of different facial expressions and poses.

The paper is organized as follows: in section II, we describe the steps of fingerprint preprocessing and face detection. In section III, we describe the procedure of feature extraction of fingerprint and face. In section IV, feature level fusion. Then, in section V we show the experimental results. In section VI we draw the conclusion.

II. FINGERPRINT PREPROCESSING & FACE DETECTION

2.1. Fingerprint preprocessing

Fingerprint preprocessing is necessary task before proceeding to next step for better identification. Such process increasing the clarity of ridge structure so that minutiae points can be easily and correctly extracted. The enhanced fingerprint image is binariged and thinned skeletanised image which has the ridge thickness to one pixel wide for precise location .Preprocessing removes the sensor noise and gray level background due to fingerprint pressure differences. Steps of the

fingerprint preprocessing are shown in Figure 1. Originalimage Binaryimage Thresholdimage Thinnedimage.

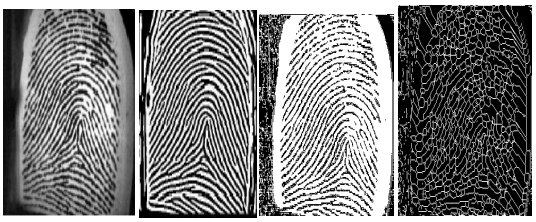


Figure.1 Steps for Fingerprint Preprocessing

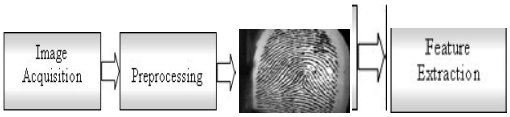


Figure.2 Image Enhancement

2.2. Feature Extraction

A fingerprint is different from the other biometric traits which possesses unique texture structure, which can be described with the orientation field of fingerprints. A fingerprint has the different orientation angle structure in different local area of the fingerprint and has a texture pattern correlation among the neighboring local areas of the fingerprint, bandwidth filter, such as the Gabor filter, can be used to emphasize ridges.

In the proposed system we first detect the core point. The core point is the special point which has the most variant changes in the directions of the lines,i.e.high curvature point of ridges. To detect the core point different techniques are used. In our paper we use Poincare Index core detection method.

Estimate the orientation field *O* using the least square orientation estimation algorithm [5]. Orientation field *O* is defined as an *M x N* image, where *O(i,j)* represents the local ridge orientation at pixel *(i,j)*. An image is divided into a set of w x w non-overlapping blocks and a single orientation is defined for each block .Initialize *A*, a label image used to indicate the core point. For each pixel *(i,j)* in *O*, compute Poincare index and assign the corresponding pixels in *A* the value of one if Poincare index is between 0.45 and 0.51.The

Poincare index at pixel *(i,j)* enclosed by a digital curve, which consists of sequence of pixels that are on or within a distance of one pixel apart from the corresponding curve, is computed as follows:

$$\text{Poincare}(i,j) = \frac{1}{2\pi} \sum_{k=0}^{Np-1} \Delta(k) \quad (1)$$

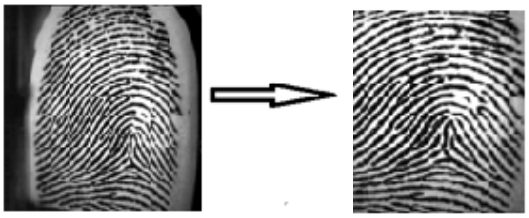
$$\Delta(k) = \begin{cases} \delta(k) & \text{if } |\delta(k)| < \pi/2 \\ \pi + \delta(k) & \text{if } \delta(k) \le -\pi/2 \\ \pi - \delta(k) & \text{otherwise} \end{cases} \quad (2)$$

$$\delta(k) = \theta(x_{(k+1) \bmod Np}, y_{(k+1) \bmod Np}) - \theta(x_k, y_k) \quad (3)$$

For our method, *Np* is selected as 8.

4) The center of block with the value of one is considered to be the center of fingerprint. If more than one block has value of one, then calculate the average of coordinates of these blocks.

After locating corepoint of finger image cropping done to get only intereste remove unwanted part of the finger for better feature extraction.



Original image Cropped image

Fig.3 Image cropping

After cropping a circular region around the core point is located and tessellated into 64 sectors. The pixel intensities in each sector are normalized to a constant mean and variance. Gabor filter is a well known technique to capture useful information in specific band pass channels. The average absolute deviation with in a sector quantifies the underlying ridge structure and is used as a feature. The feature vector (1280 values in length) is the collection of all the features, computed from all the 64 sectors, in every filtered image. The feature vector captures the local information and the ordered enumeration of the tessellation captures the invariant global relationships among the local patterns It is desirable to obtain representations for fingerprints which are translation and rotation invariant. In the proposed scheme, translation is taken care of by a reference point which is core point during the feature extraction stage and the image rotation is handled by a cyclic rotation of the feature values in the feature vector. The features are cyclically rotated to generate feature vectors corresponding to different orientations to perform the matching.

Gabor filters optimally capture both local orientation and frequency information from a fingerprint image. By tuning a Gabor filter to specific frequency

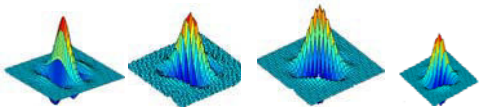
and direction, the local frequency and orientation information can be obtained. Thus, they are suited for extracting texture information from images. An even symmetric Gabor filter has the following general form in the spatial domain:

$$G(x, y ; f, \theta) = \exp\left\{\frac{-1}{2}\left[\frac{x'^2}{\delta_{x'}^2}+\frac{y'^2}{\delta_{y'}^2}\right]\right\}\cos(2\pi f x') \quad (4)$$

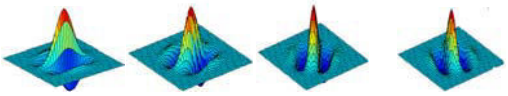
$$x'=x\sin\theta+y\cos\theta \quad (5)$$

$$y'=x\cos\theta - y\sin\theta \quad (6)$$

where f is the frequency of the sinusoidal plane wave along the direction θ from the x-axis, and x δ and y δ are the space constants of the Gaussian envelope along x and y axes, respectively. The filtering is performed in the spatial domain with a mask size of 16x16. The frequency f is the average ridge frequency (1/K), where K is the average inter ridge distance. The average inter ridge distance is approximately 10 pixels in a 500 dpi fingerprint image. Hence, the finger can examined at different orientations and this correspond to θ values.



1) 0^0 2) 45^0 3)90^0 4)135^0



5) 180^0 6)225^0 7)270^0 8)315^0

At the matching stage the gabor features of train and test image are compared and distance has been calculated, if the distance is within threshold limit the image is said to be similar.

III. FACE DETECTION AND FEATURE EXTRACTION

Face detection from cluttered images is very tough, due to the change in environment, light effects, facial expressions and different poses of the face. The most popular approaches to face recognition are based on i)the location and shape of facial attributes such as eyes,eyebrows,nose,lips and chin and there spatial reletionships,ii)the overall analysis of face image reprents a face as a weighted combinations of number of conical faces. In our proposed system we simply used the gabor filter for feature extraction from face which are used for face recognition.

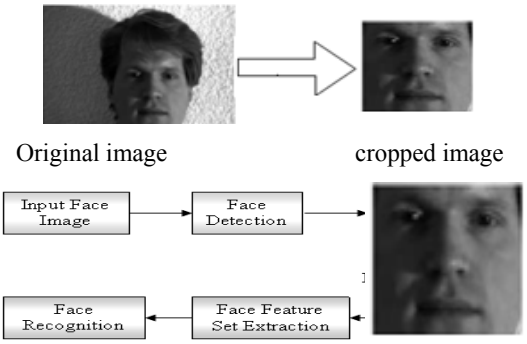


Figure.4. step of detection and Feature extraction

In proposed system,for face recognition the steps followed are shown in fig.4.First acquire the face image then haar transformation technique is applied for face face detection. Then cropped the image for removing the unwanted noise from the figure.Similar like fingerprint the center point of face image is also detected and gabor filter is applied for feature extraction.Then extracted features are stored as template.

IV. PROPOSED MULTIMODAL SYSTEM

4.1. Framework of proposed system

To overcome the problems in the unimodal biometric system .Multi-biometrics are use. With the lower hardware cost a multi biometric system uses multiple sensors for data acquisition .In unimodal system if biometric trait being sensed is noisy then matching result may be not reliable. Hence by using multiple sensors more biometric traits can captured and can get more reliable result. The block diagram of our proposed system is shown in fig.5.

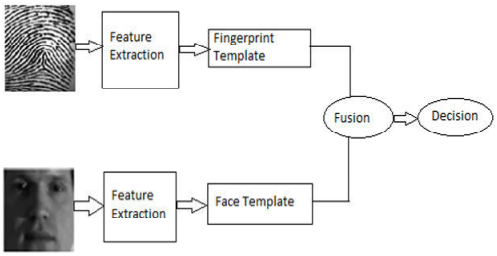


Fig. Block diagram of multimodal system

4.2. Mode of operation

A multimodal biometric system can operate in one of three different modes: serial mode, parallel mode, or hierarchical mode[4]. In our system we used serial mode of operation.In the serial mode of operation, the output of one biometric trait is typically used to narrow down

the number of possible identities before the next trait is used. This serves as an indexing scheme in an identification system. In the serial operational mode, the various biometric characteristics do not have to be acquired simultaneously. Further, a decision could be arrived at without acquiring all the traits. This reduces the overall recognition time. This is in contrast to a parallel mode of operation where information from multiple traits is used simultaneously.

4.3.Fusion

Multimodal biometric systems integrate information presented by multiple biometric indicators[4]. The information can be consolidated at various levels.

a)feature extraction level

b)matching level

c)decision level

We used fusion at feature extraction level because it is considered as a combination scheme applied as early as possible in the recognition system is more effective. i.e an integration at the feature level typically results in a better improvement than at the matching score level. This is because the feature representation conveys the richest information compared to the matching score of a matcher, while the abstract labels contain the least amount of information about the decision being made. In feature level fusion the data obtained from each biometric modality is used to compute a feature vector. If the features extracted from one biometric indicator are (somewhat) independent of those extracted from the other, it is reasonable to concatenate the two vectors into a single new vector, provided the features from different biometric indicators are in the same type of measurement scale. The new feature vector has a higher dimensionality and represents a person's identity in a different feature space.

The our system is basically divided into two parts (i)crating profile (ii)identification. In first part the the images are acquired from sensors, features are extracted using Gabor filter , extracted features are get fused then a single feature is saved as template in dataset. In the second part the fingerprint images is taken as query images again the features are extracted and single fused template is compared to the templates stored in dataset for identification. The data set is get updated every time i.e. the stored template is replaced by new extracted template at the time of next authentication .

V. EXPERIMENTAL RESULTS

The reliability of the proposed multimodal system is described with the help of experimental results. The system has been tested on three standard datasets for face and fingerprint(att,ifd,yale,FVC2004).The training database contains one fingerprint and one face image and seven images of the same person for testing of each individual.Also the system is tested on the images of fingerprints acquired using optical sensors at a resolution of 500dpi and the face image is acquired using 3-CCD camera.We implemented this method in MATLAB7.5.0(R2007b version) and processed on Pentium machine GHz.

The multimodal system has been developed at multi-classifier and multi-modal level. At first part the individual system were developed and tested for FAR,FRR and Accuracy.

Trait	Algorithm	Accuracy (%)
Finger	LBP[1]	87
	Minutia technique[7]	72
	Gabor filter	88
Face	GWN[1]	74
	PCA[8]	75
	HaarTrancform+Gabor filter	72

Table .Accuracy of unimodal system

Trait	Algorithm	FAR %	FRR %	Accuracy (%)
Finger	Gabor filter	0.1	0.17	88
Face	HaarTrancform+Gabor filter	0.4	0.23	72
Fusion	Gabor filter	0.11	0.03	97

Table2.accuracy,FRR,FAR of multimodal system

The performance of the any biometric system is represented by the ROC(Receiver operating characteristic)curve. The ROC curve plots the probability of False Acceptance rate(FAR)versus probability of False Rejection Rate(FRR) for different values of the decision threshold. To show effectiveness of proposed method, we have plotted ROC curve for Genuine Acceptance Rate(GAR)versus FRR

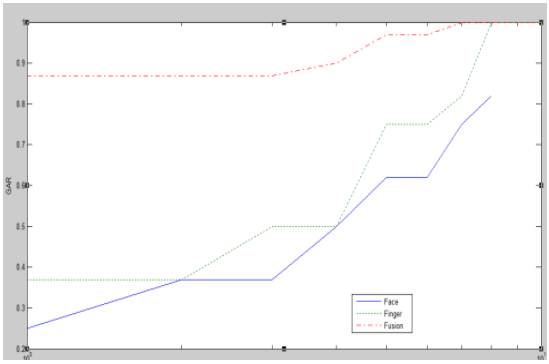


Fig.Roc curve shows the performance of the system

Face recognition (seconds)	Fingerprint verification (seconds)	Total (seconds)
1.68	1.72	4.21

Table3 shows average CPU time for one test.

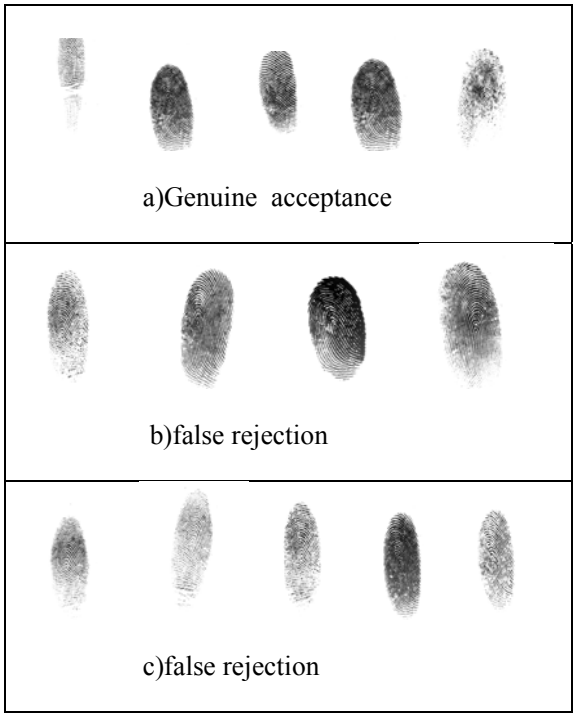


Fig.Result of tested fingerprints

Fig. Examples of tested images of Face

Fig. Examples of tested images of Finger

VI. CONCLUSION

Now a day's biometric systems are widely used to overcome the problems of traditional authentication systems.but most of the unimodal systems are fails to give results effectively due to lack of biometric information of particular trait.We presented an effective biometric multimodal system which utilizes Gabor filter for both fingerprint and face recognition. Fusion is done at feature extraction level. The performance table and accuracy curve shows that multimodal system performs better as compared to unimodal system with 97% accuracy.in future our next step will be to improve the response time of the system.

REFERENCES

[1] *Norhene Gargouri Ben Ayed,* "A New Human Identification Based on Fusion Fingerprints and faces biometrics using LBP and GWN descriptors", 2011 8th international multi-conference on systems, signals & devices.

[2] Muhammad Umer Munir and Dr. Muhammad Younas Javed" Fingerprint Matching using Gabor Filters", National Conference on Emerging Technologies 2004 147

[3] Sheetal Chaudhary , Rajender Nath," A Multimodal Biometric Recognition System Based on Fusion of Palmprint, Fingerprint and Face",2009 International Conference on Advances in Recent Technologies in Communication and Computing

[4] Anil K. Jain, Arun Ross and Salil Prabhakar" An Introduction to BiometricRecognition",IEEE Transactions on circuits and systems for Video Technology, vol. 14, no. 1, January 2004

[5] Lin Hong and Anil Jain,"Integrating Faces and Fingerprints for Personal Identification",IEEE Transactions on patternAanalysis and Machine intelligence, vol. 20, no. 12, December 1998

[6] Ravi. J, K. B. Raja, Venugopal. K. R.," Fingerprint Recognition using minutia score matching"

[7] Anil Jain, Arun Ross,Salil Prabhakar,"Fingerprint Matching using Minutiae and Texture Features",0-7803-6725-1/2001 IEEE

[8] Iftikhar Ali,Usman Ali,Abdul Malik,"Face and Fingerprint biometric Integration Model for Person Identification using Gabor Filter",1-4244-0212-3/06/2006IEEE

[9] Syed Maajid Mosin,M.Younus Javed, "Face recognition using Bank of Gabor filters",1-4244-0502/06/2006IEEE

[10] Dr.A.Wahi,R.Vinothkanna,R.Anushuya Devi,S.Bhuvaneswari,"Biometric Authentication using fingerprints: A Review",IJBB,Volume(5):Issue(1)

[11] Phalguni Gupta,Ajita Rattani,Hunny Mehrotra,Anil Kumar Kaushik"Multimodal Biometrics System for Efficient Human Recognition"

[12] Shi-Jinn Horng,Kevin Octavius Sentosal,"An Improved Score Level Fusion in Multimodal Biometric Systems",2009 international Conference on Parallel and Distributed Computing ,applications and Technologies.

[13] Arun Ross,Rohin Govindarajan,"Feature Level Fusion Using Hand and Face Biometrics",SPIE Conference on Biometric Technology for Human identification II,vol.5779.pp.196-204,march-2005

[14] Zhaomin zhu,Takashi Morimoto,Hidekazu Adachi,Osamu Kiriyama,"Multi-view Face Detection and Recognition using haar-like Features"

[15] Haiyuan W, Yukio Yoshida and Tadayoshi Shioyama,"Optimal Gabor Filters for High Speed Face Identification".

◈◈◈

Comparative Analysis of Square Rooting and Tone Injection Method for PAPR Reduction In OFDM System

Ananta S. Chavan, Akshay V. Kulkarni & Shilpa P. Metkar

Dept. of Electronics and Telecommunication, College Of Engineering Pune, Pune, India

Abstract - OFDM system is an excellent multi-carrier system for wireless communication. PAPR is one of the major critical problem in the OFDM systems. Earlier, different methods have been proposed to reduce the PAPR. In this paper, we are comparing these methods on basis of PAPR, bit error rate (BER), complimentary cumulative distribution function (CCDF), in band and out of band radiation. Later section investigates the simulation results and conclusion based on the above parameters.

Keywords: *Orthogonal frequency division multiplexin (OFDM), Peak to average power ratio(PAPR.*

I. INTRODUCTION

OFDM was first introduced in 1966, patented in 1970 and has many well known advantages, such as robustness against frequency selective fading or narrowband interference, high bandwidth efficiency, and efficient implementation[1]. OFDM is a special case of multicarrier transmission, where a single data stream is transmitted over a number of lower rate subcarriers [2]. The main reason to use OFDM is to increase the robustness against the selective fading or narrowband interference. In single carrier system if signal get fade or interfered then entire link gets failed where as in multicarrier system only a small percentage of the subcarriers will be affected. Due to OFDM system's immunity from selective fading it is widely used in high speed communication [1] devices like ADSL Modem, DAB, DVB, WLAN etc. It is also going to be used in 4G cellular communication networks. Since its discovery lot of research has been done in this area to improve OFDM communication system for future needs.

Despite many advantages, a major drawback of OFDM is its high Peak-to-Average Power Ratio (PAPR) problem, which makes system performance very sensitive to nonlinear distortions. Indeed, when the OFDM signal with high PAPR passes through a nonlinear device, the signal may suffer significant nonlinear distortions and severe power penalty which is unaffordable for battery powered portable wireless terminals. To reduce the PAPR of OFDM signals, several PAPR reduction techniques have been proposed [1][2][4]. Clipping is the most straightforward PAPR reduction technique but can lead to significant out-of-band distortion. In order to alleviate such effects filtering can be applied[6]. However this causes significant peak re-growth. Distortion less techniques such as Tone Reservation [5] can also be employed. Tone Reservation reserves a small number of data tones in the frequency domain to generate an effective cancellation signal for high peaks in the time domain. This can be computationally very burdensome and may cause a significant (up to 20%) reduction in data through-put. Selective Mapping (SLM) is implemented by generating a set of sufficiently-different candidate signals from the original data signal. The transmitter selects and submits the candidate signal having the lowest PAPR. Partial Transmit Sequencing (PTS) is a similar technique in which sub-blocks of the original signal are optimally combined at the transmitter to generate a transmitted signal with a low PAPR. Although SLM and PTS are effective at reducing the PAPR, they require the use of side information in order to decode the signal at the receiver [7].

The rest of the paper is organized as follows. Section-II describes the PAPR problem and in Section-III, we present details of the different PAPR reduction techniques. simulation results are presented in section-IV. The paper is concluded in Section-V.

II. PAPR BACKGROUND THEORY

Most radio systems employ the high power amplifier (HPA) in the transmitter to obtain sufficient transmission power. For the purpose of achieving the maximum output power efficiency, the HPA is usually operated at or near the saturation region. Moreover, the nonlinear characteristic of the HPA is very sensitive to the variation in signal amplitudes. However, the variation of OFDM signal amplitudes is very wide with high PAPR.[7]

Therefore, HPA will introduce inter-modulation between the different subcarriers and introduce additional interference into the systems due to high PAPR of OFDM signals. This additional interference

leads to an increase in BER. In order to lessen the signal distortion and keep a low BER, it requires a linear work in its linear amplifier region with a large dynamic range. However, this linear amplifier has poor efficiency and is so expensive. Power efficiency is very necessary in wireless communication as it provides adequate area coverage, saves power consumption and allows small size terminals etc [1][7].

It is therefore important to aim at a power efficient operation of the non-linear HPA with low back-off values and try to provide possible solutions to the interference problem brought about. Hence, a better solution is to try to prevent the occurrence of such interference by reducing the PAPR of the transmitted signal with some manipulations of the OFDM signal itself. Large PAPR also demands the DAC with enough dynamic range to accommodate the large peaks of the OFDM signals. Although, a high precision DAC supports high PAPR with a reasonable amount of quantization noise, but it might be very expensive for a given sampling rate of the system. Whereas, a low-precision DAC would be cheaper, but its quantization noise will be significant, and as a result it reduces the signal Signal-to-Noise Ratio (SNR) when the dynamic range of DAC is increased to support high PAPR. Furthermore, OFDM signals show Gaussian distribution for large number of subcarriers, which means the peak signal quite rarely occur and uniform quantization by the ADCs is not desirable. If clipped, it will introduce in band distortion and out-of-band radiation (adjacent channel interference) into the communication systems. [1][9] Therefore, the best solution is to reduce the PAPR before OFDM signals are transmitted into nonlinear HPA and DAC.

For an OFDM signal with *M* subcarriers and total continuous time signal $x(t)$ consisting of all the OFDM block, the PAPR is defined as,

$$\text{PAPR} = \max |x(t)|^2 / E\,|x(t)|^2 \quad (1)$$

In particular, a baseband OFDM signal with M subchannels has,

$$\text{PAPRmax} = 10\log_{10} M \quad (2)$$

From the central limit theorem, it follows that for large values of M (M>64), the real and imaginary values of $x(t)$ become Gaussian distributed. Therefore the amplitude of the OFDM signal has a Rayleigh distribution, with a cumulative distribution given by,

$$F(z) = 1 - e^{-z} \quad (3)$$

The probability that the PAPR is below some threshold level can be written as,

$$P(PAPR \le z) = (1 - e^{-z})M \quad (4)$$

In fact, the complementary cumulative distribution function (CCDF) of PAPR of an OFDM is usually used, and can be expressed as,

$$P(PAPR > z) = 1 - (1 - e^{-z})\text{M} \quad (5)$$

III. PAPR REDUCTION TECHNIQUES

There is numerous work done to reduce PAPR by lot of scholars. In this PAPR we will focus on the conventional PAPR reduction methods such as square rooting technique and tone injection method [5].

A. THE SQUARE ROOTING METHOD

Square rooting method is one of the distortion based method of PAPR reduction[3]. Some of the statistical properties of signals can be changed by applying special mathematical operations. As an example, the Rayleigh distribution of any signal will change into Gaussian distribution if the square root operation is applied to that signal. Also, the Chi-Square distribution can be transformed into Rayleigh distribution by applying the same process, square rooting. Not only the statistical distribution changes by square rooting, but the mean and variance values of the signal are also varied. This process, square rooting, is exploited in this work to realize reduction of the PAPR value. From the central limit theorem and for large number of input samples, the imaginary and real parts of the IFFT outputs will follow Gaussian distributions. Hence, the amplitude (envelope) of the complex valued OFDM symbols will have Rayleigh distribution computed by,

$$x(n) = (\text{Re}\{x[n]\})^2 + (\text{Im}\{x[n]\})^2 \quad (6)$$

x(n) **:** Amplitude values of OFDM symbols, random variable (RV) of Rayleigh distribution.

(Re{x[n]})**:**Real part of OFDM symbols, RV of Gaussian distribution.

(Im{x[n]}) **:** Imaginary part of OFDM symbols, RV of Gaussian distribution.

At the same time, the power distribution becomes a central chi-square distribution.

$$X_k = \sqrt{|x_k|}.e^{j\Phi} \quad (7)$$

where $e^{j\Phi}$ is the phase of x_k

The square rooting technique can be used to improve OFDM transmission performance. This technique is used to compand the OFDM signal before it is converted into analog waveform. The OFDM signal, after taking IFFT, is companded and quantized. After D/A conversion, the signal is transmitted through the channel. At the receiver end then the received signal is first converted into digital form and expanded in order to recover the original signal.[8]

It can increase the average transmit power to reduce the PAPR and make the signal power value easier to be near to power non-linear changes in the amplifier region, so that the signal is likely to be distorted.

B. TONE INJECTION (TI) METHOD

Tone injection is a efficient techniques to reduce the PAPR of OFDM signals. The key idea in TI is that both transmitter and receiver reserve a subset of tones for generating PAPR reduction signals ς. Note that these tones are not used for data transmission. In TI, the objective is to find the time domain signal ς to be added to the original time domain signal to reduce the PAPR. Let (ς=0,1,…N-1) denote complex symbols for tone injection at reserved tones. Thus, the data vector changes to after tone reservation processing, and this results in a new modulated OFDM signals as,

$$X = \text{IFFT}(x+\varsigma) \quad (8)$$

TI also uses an additive correction to optimize in ς. The basic idea of TI is to extend the constellation and thus the same data point corresponds to multiple possible constellation points. One option is to replicate the original shaded constellation into serval alternative ones. Therefore, is a translation vector such that ς = X mod (X). Note that TI needs not require the extra side information and the receiver only needs to know how to map the redundant constellations on the original one. An alternative strategy is to move the constellation points by applying an FFT on the clipped time signals, and the same operations are repeated till all the constellation points are within specified boundaries and the PAPR specification of the time signal is satisfied [5]. Some modifications of TI have been proposed to obtain good performance including PAPR reduction and low complexity [5]. The TI technique is more problematic than the Tone Rejection technique since the injected signal occupies the frequency band as the information bearing signals. Moreover, the alternative constellation points in TI technique have an increased energy and the implementation complexity increases for the computation the optimal translation vector[5].

IV. SIMULATION RESULTS

A. SIMULATION PARAMETERS

Table I gives the simulation parameters. The number of subcarriers, *K,* is set to 16. The number of FFT/IFFT points, *F* is set to 64, which corresponds to 4-times oversampling in the time domain. The simulations are performed for an IEEE 802.11a compliant OFDM system employing 16-QAM in which $N = 64$ subcarriers are used. MATLAB simulator is used to carry out these results.

TABLE I SIMULATION PARAMETERS

Number of subcarriers, *K*	**16**
Number of IFFT points, *F*	64
Data modulation	16-QAM
Channel model	AWGN
Maximum number of iterations in PAPR reduction	10

B. SIMULATION RESULTS

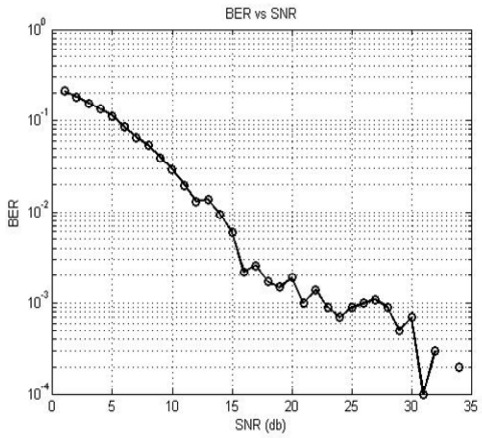


Fig1:BER vs SNR curve for Square rooting

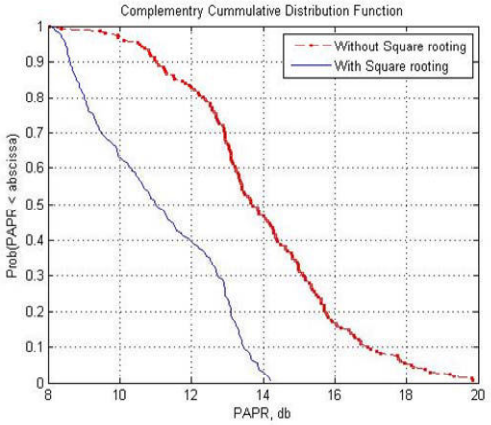


Fig 2:CCDF plot for Square rooting

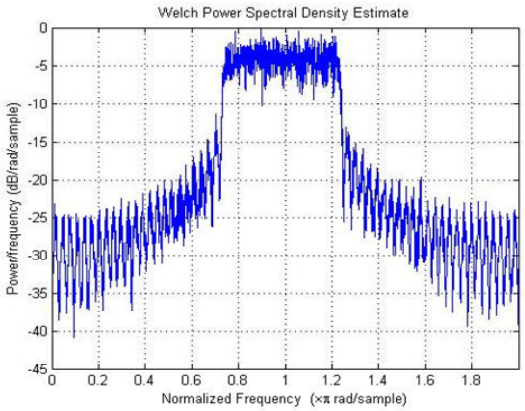


Fig 3:In-band and out of band radiation for Square rooting

Fig-1 signifies the BER vs SNR curve.It is clear that as we go on increasing the SNR,bit error rate reduces. An acceptable amount of BER 10^{-4} was achived at SNR=30db.

In Fig-2 The results are presented in terms of Complementary Cumulative Distribution Function (CCDF) for an OFDM output symbol blocks before and after processing. The difference in PAPR performance between the conventional OFDM system, and that obtained from the square rooting method are observed. The CCDF curve was obtained at SNR=10db. As could be seen, the performance of the system in term of PAPR show that maximum reduction in PAPR value of slightly more than 5dB was achieved by the square rooting method for the OFDM system parameters of 10048 samples and 16-QAM modulation scheme.

As shown in the fig-3 In-band and out of band radiation graph is investigated the performance of the technique in a communication system also depends on the effect of the resulting in-band distortion and on how much out-of-band power will result after non-ideal amplification. We can clearly observe that due to non-linear effect In-band distortion is high which effects BER of the system. Out-band distortion causes degradation in spectral efficiency. Due to non linearity effect in-band radiation is more which ultimately results in slight increase in bit error rate of the system.

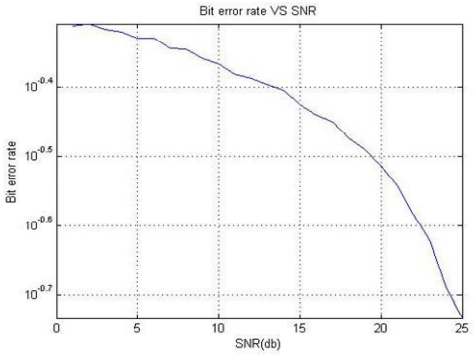


Fig4:BER vs SNR curve for TI method

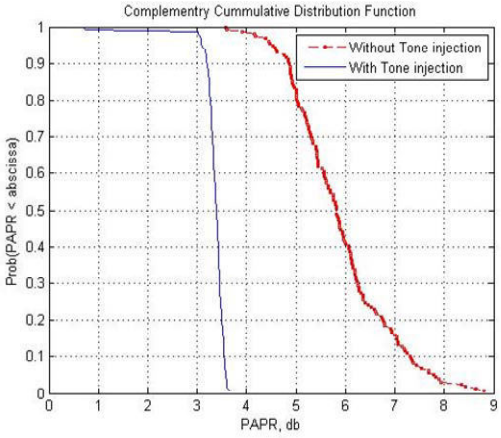


Fig 5:CCDF plot for TI Method

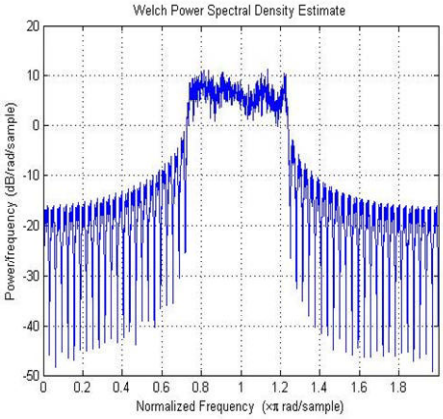


Fig 6:In-band and out of band radiation for TI method

All the simulation parameters were kept same for tone injection method. Fig-4 states that the bit error rate (BER) vs signal to noise ratio plot. Bit error rate of Tone injection method largely depends upon the side information

Fig-5 elaborates Complementary cumulative distribution function curve. CCDF curve was obtained at SNR=10db. It can be seen that an excellent PAPR reduction of 4db was achieved by TI method. This method is highly complex than square rooting technique. Method also achieves excellent PAPR reduction.

In-band and out of band radiation is investigated in Fig-6 we can conclude that out of band radiation is more than the square rooting technique method due to side information in-band radiation is less than which ultimately results in slight increase in bit error rate of the system.

V. CONCLUSION

This paper mainly focuses on one of the challenging drawback of OFDM that is high PAPR. Square rooting method is simple non-liner distortion based method, while tone injection method is complex distortion less method. The simulation proves that though TI method is affective to reduce the PAPR. It costs significant increase the out of band radiation which increases bit error rate of the system. Square rooting method achieves good PAPR reduction with less bit error rate. 30db SNR was enough to achieve desired SNR CCDF curves investigates that TI method is slightly better than clipping. We get nearly 5db and 4db reduction in TI and square rooting methods.

Square rooting method is suitable for OFDM applications that are sensitive to spectral efficiency and noise, since it allows efficient reduction in PAPR value with low out-of band radiation. TI method can be better where we require large reduction in PAPR but price is paid by reducing throughput and data rate of system. High BER is another problem which can be overcome by adding redundancy in data and coding techniques.

ACKNOWLEDGMENT

The authors are grateful to the College of engineering pune, for providing necessary support and infrastructure.

REFERENCES

[1] R.Van Nee, R. Prasad Publication by Artech House,"OFDM for wireless Multimedia Communications.", e-book.

[2] Y. Wu and W. Y. Zou, "Orthogonal frequency division multiplexing:a multi-carrier modulation scheme," *IEEE Transactions on Consumer Electron,* Aug. 1995

[3] T. Jiang and G. Zhu, "Nonlinear companding transform for reducing peak-to-average power ratio of ofdm signals," *IEEE Transactions onBroadcasting,* Sept. 2004.

[4] Horosaki B., " An orthogonally multiplexed QAM system using the DFT," IEEE Trans. Comm. Vol. COM-29, pp. 982-989, July 1981.

[5] B. S. Krongold and D. L. Jones, "PAR reduction in OFDM via activeconstellation extension," *IEEE Trans. Broadcasting*, vol. 49, no. 3, pp.

[6] 258–268, Sept. 2003.H.Ochiai and H.Imai, "Performance analysis of deliberately clipped OFDM signals," *IEEE Trans. Commun.*, vol. 50, no. 1,

[7] S. H. Han and J. H. Lee, "An overview of peak-to-average powerratio reduction techniques for multicarrie transmission," IEEE Wireless Communications, vol. 12, issue 2, pp. 56- 65, April 2005..

[8] X. Wang, T. T. Tjhung, and C. S. Ng, "Reply to the comments on 'reduction of peak-to-average power ratio of OFDM system using a companding technique'," *IEEE Trans. Broadcast.*, vol. 45, pp. 420–423, Dec.1999.

[9] P. Banelli and S. Cacopardi, "Theoretical analysis and performance of OFDM signals in nonlinear AWGN channels," *IEEE Trans Commun.*, vol. 48, pp. 430–441, Mar. 2000

◈◈◈

Ultra Wide Band Planer Antenna for Wireless Network Application

Patel maulik P, Modh Kunal K , Sheliya Rakesh L & Patel Rakesh B

Electronics And Communication Engineering, Hasmukh Goswami College Of Engineering,
Gujarat technological university ,Ahmedabad, India.

Abstract - This paper presents a planar Antenna for Ultra Wide Band frequency for cover a large bandwidth of 3.1 GHz and 13.1GHz for the resonance frequency of 6.5 and its wide application like WLAN, Wi-MAX, Medical Application, radar imaging technology, PC Peripherals, Wireless USB. The gain and directivity of the proposed antenna are presented at different UWB band frequencies. for HFSS is used to design and simulation of antenna.

Keywords— Micro strip Transmission Line, Ultra Wide Band, Planar Antenna.

I. INTRODUCTION

Application of ultra-wideband (UWB) technology on wireless communication system has increased considerably in last seven years. Because the UWB technology has great potential in the development of various modern wireless communication systems, the U.S Federal Communication Commission (FCC) authorized the unlicensed use of the ultra wideband (3.1-10.6 GHz) frequency spectrum for indoor and hand-held wireless communication since early February 2002[3]. To meet the variety of applications in UWB communication systems, many researchers around the world have been aroused on the design, research and development of UWB filter and antenna [4-5].

II. UWB ANTENNA

UWB is a Radio Frequency (RF) technology that transmits binary data, using low energy and extremely short duration impulses or bursts (in the order of picoseconds) over a wide spectrum of frequencies. It delivers data over 15 to100 meters and does not require a dedicated radio frequency so is also known as carrier-free, impulse or base-band radio.

People commonly refer to UWB as available spectrum rather than as a technology 7500 MHz of unlicensed spectrum, in the 3.1-10.6 GHz band, is currently available in the US for any Communication system that occupies more than 500MHz. fig.1 the data rates of ultra wide antenna is very high but cover distance is less in

The simulated 10 db at 3.1 to 13.1 GHz with 10Ghz large bandwidth. Low dielectric constant substrates are generally preferred for maximum radiation.

Main purpose of this antenna to use one antenna for many application like WLAN, Wi-MAX, Medical Application, radar imaging technology, PC Peripherals, Wireless USB etc..,which is the cost effective to deign one antenna for all application.

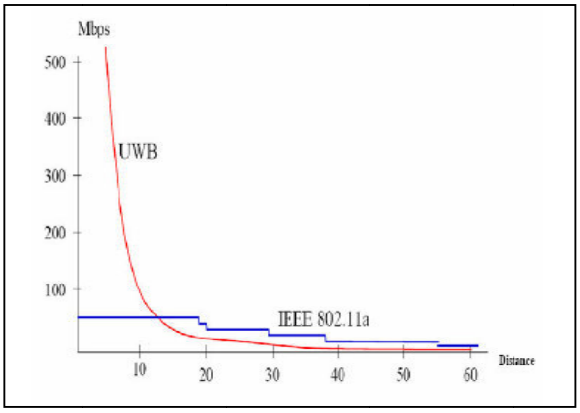


fig.1 UWB Data Rates

III. UWB ANTENNA CONFIGURATION DESIGN

In the design of this type of antennas, the width 'W' and Length 'L' plays a crucial role in determining the resonant frequency of the system. The starting values of these parameters are calculated by using the equations given in [9-10] for the substrate height (h), dielectric constant (εr) and for the lower frequency. The designed values of the antenna are optimized with HFSS tool. The optimization was performed for the best impedance bandwidth. Fig 2. shows the structure of the ultra wide band planer antenna. The antenna consists of rectangular aperture with width 'W' and length 'L' and rectangular patch with height 'H'. In this study, a dielectric

substance (FR4) with thickness of 1.55 mm with a relative permittivity of 4.4 is chosen as substrate. The CPW feed is designed for 50 Ω characteristic impedance with fixed 2.6 mm feed line width and 0.035 mm ground gap.

By properly adjusting the dimension of the antenna and feeding structure the impedance matching of the proposed antenna is improved that produces wider impedance bandwidth with satisfactory radiation pattern. The wide bandwidth and impedance matching with reduced size of the antenna is achieved by the different surface magnetic currents of the structure. [7]-[8].

Fig.2 shows the geometry and configuration of ultra wide-band (UWB) antenna. The design parameters are L=21.8mm,W=14mm, H=0.035mm, h=1.55 mm (sub.hight).

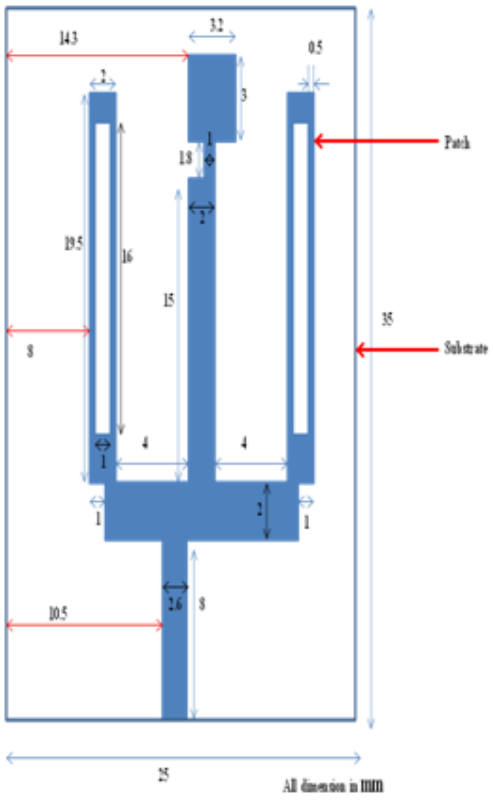


Fig. 2.Geometry and configuration of UWB planer antenna.

IV. SIMULATION RESULTS

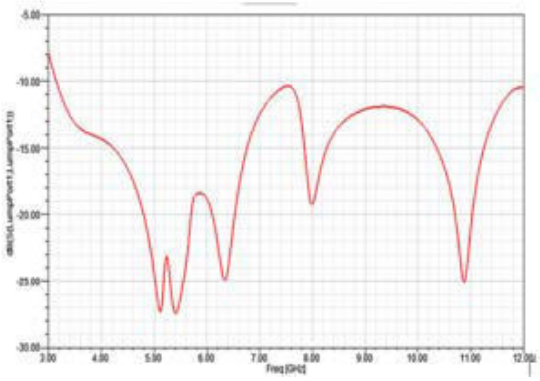


Fig. 3. Return Loss vs. frequency of UWB planer antenna

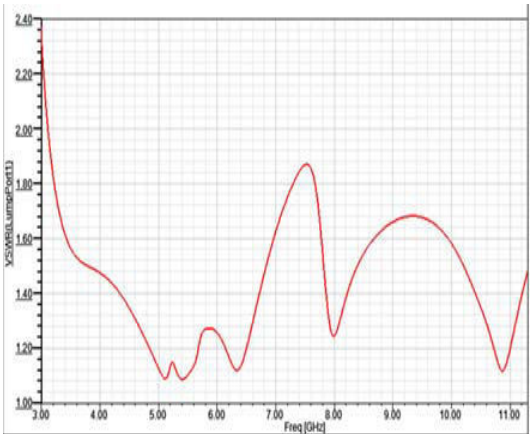


Fig. 4. VSWR vs. freq. of UWB planer antenna

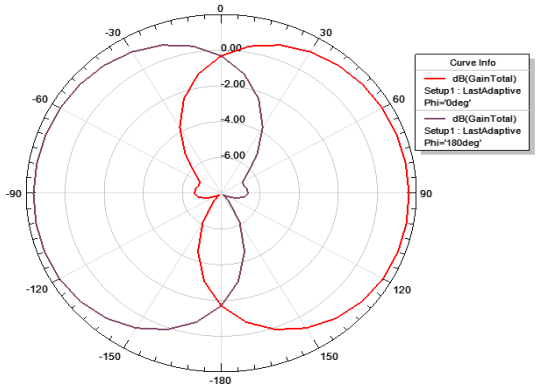


Fig. 5.Gain at 6.5 GHz.

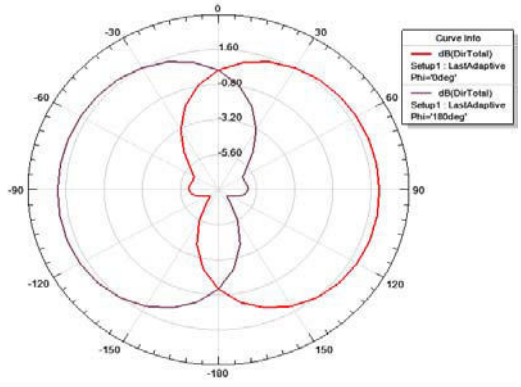


Fig. 5. Directivity at 6.5 GHz.

V. CONCLUSION

This paper investigates a planar band antenna that cover UWB bands 9GHz (3.1 to 12.1GHz) suitable for, WLAN, Wi-MAX, Medical Application, radar imaging technology, PC Peripherals, Wireless USB The antenna has a low profile and can be easily embedded into the display of a laptop computer. This simple structure made from common materials is very cost effective. The impedance and radiation performance of the antenna integrated into the lossy display of a laptop are taken into account. The SWR, maximum and average gain as well as radiation patterns of the UWB antenna are examined experimentally.

UWB and the associated networking protocol efforts are in the early stages of development, and several key deployment scenarios are being defined and evaluated.UWB complements currently deployed wireless networks in the WLAN environment, plus it extends high bit-rate, multimedia connectivity to WPANs supporting PC, CE and cellular devices. This combination will enable true convergence of computers, consumer electronics and mobile communications.

REFERENCES

[1] Kumar,G.and K. P. Ray,"Broadband Microstrip Antennas", Attach House, Boston, 2003**.**

[2] Wong, K. L., "Compactand Broadband Microstrip Antenna", John Wiley & Sons, New York, 2002.

[3] First Report and Order, "Revision of Part 15 of the Commission's Rule Regarding Ultra-Wideband Transmission systems FCC 02-48," Federal Communication Commission, 2002.

[4] Y.Yao, B.Huang and Z. Feng, "A novel ultra-wideband Microstrip-line fed wide slot antenna having frequency band-notch functions, International Conference on Microwave and Millimeter Wave Technology (ICMMT), pp.1-4, 18-21 April, Bulin, 2007.

[5] J. Young, J. Cho, K. -H Kim, D. -H Choi, S. -S Lee, and S. -O Park, " A Miniature UWB Planar Planer Antenna With 5-GHz Band-Rejection Filter and the Time-Domain Characteristics,'' IEEE Trans. Antennas Propag.,. Vol.54, no.5, pp. 1453-1460, May. 2006.

[6] Jyoti R. Panda#1, Prasadu Kakumanu and Rakhesh S. Kshetrimayum "A Wide-band Planer Antenna in Combination with a UWB Microwave Band-pass Filter for Application in UWB Communication System.

[7] Ray, K. P., ''Broadband, Dual-Frequency and Compact Microstrip Antennas,''Ph.D thesis, Indian Institute of Technology, Bombay, India, 1999.

[8] Okoshi, T., and T. Miyushi, ''The Planar Circuit An Approach to Microwave Integrated Circuitry,'' IEEE Trans. Microwave Theory Tech., Vol. 20, April 1972,pp. 245–252.

[9] J.D.krauss,"Antenna for all application", 3rd addition.TMH publication.1995

[10] C. A. Balanis, "Antenna Theory: Analysis and Design, 2nd Edition", pp. 722 – 752, John Wiley and Sons, Inc., 1997.

[11] Ansoft HFSS version 11.0.

◈◈◈

Implementation of a Baseband OFDM Transceiver Based on Xilinx Spartan 3 FPGA using VHDL

Ananta S. Chavan, Akshay V.Kulkarni & Shilpa P. Metkar

Department of Electronics and Telecommunication College of Engineering Pune, Pune, India.

Abstract - Orthogonal Frequency Division Multiplexing (OFDM) is a multicarrier communication system. It is widely used because of its immunity to frequency selective fading channels. In this paper, the design and an implementation of OFDM transceiver on FPGA is presented. The system is designed using VHDL, synthesized using high level synthesis tool and targeted on Xilinx Spartan 3e device. Presented design is simulated on ISE simulator and the results are presented. Resources utilization for transmitter and receiver is given in this paper. The design utilizes the Intellectual Property (IP) cores provided by Xilinx for floating point multiplication, addition subtraction and division. DIT radix-2 butterfly approach is used to calculate IFFT and FFT.

Keywords-*FPGA, FFT, IFFT, DIT, ISE, VHDL.*

I. INTRODUCTION

The modern digital mobile communication systems are increasingly using baseband OFDM for multi-carrier transceiver. Idea behind the high spectral efficiency of OFDM is elimination of guard bands and use of the overlapping but orthogonal subcarriers. High rate data stream is divided into a number of low rate data streams that are transmitted over a number of multiplexed orthogonal subcarriers [1]. The low rate data streams allow adding sufficient guard time between two symbols which was very small in high rate data stream. This helps in enabling the system to perform well in dispersive channel which causes the symbols to spread in time and interfere with each other called as inter symbol interference (ISI).

OFDM can be viewed as either a modulation or multiplexing technique, and its hierarchy lies in the physical and medium access layer. A basic OFDM transceiver consists of a QAM or PSK modulator/demodulator, a serial to parallel/parallel to serial converter, and an IFFT/FFT module [1]. The block diagram of basic OFDM system is shown in Figure 1. The transmitter consists of a input bit stream, serial to parallel converter, constellation mapping, IFFT, DAC. The receiver consists of ADC, FFT, parallel to serial converter, demodulation, and output bit stream.

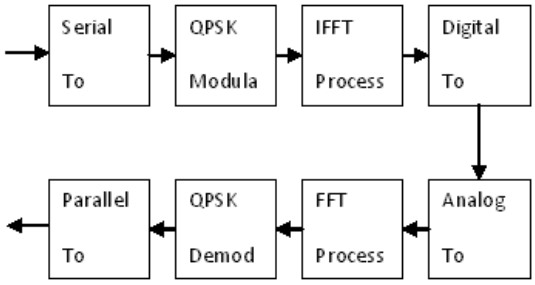


Figure 1: Block diagram of basic baseband OFDM system

This paper is organized as follows. Section II presents the design flow to implement an OFDM transceiver from system design to circuitry realization on FPGA. The implementation aspects of the transmitter and receiver are demonstrated in section 3 and 4. The results are discussed in section 5 and finally the conclusion is given in section 6.

II. DESIGN FLOW

Initially transmitter and receiver is implemented independently and tested on kit, and then both the subsystems were merged to form one system. The system is designed for two sets of subcarriers one is using 4 subcarriers and the other using 8 subcarriers. Design flow for both the system is same. It is explained as follows.

The design of system starts with the understanding of the block diagram. It gives the idea about operations

needed to be performed with help of FPGA. Then the algorithm is developed for the sequential and concurrent operations. To make the design more parallel the operations are broken in to processes and independently written in VHDL. Some blocks contain floating point complex operations, for that Intellectual Property (IP) cores provided by Xilinx are used. IP cores are used for all types of floating operations.

The system is designed completely on Xilinx Project Navigator using VHDL coding as design entry method. Then the system is simulated on ISE simulator for timing analysis. Finally the design is synthesized on FPGA Spartan 3e device using high level synthesis tool.

III. TRANSMITTER DESIGN

A. Transmitter design Flow

Input bit stream of length 16 bit is stored in the array. Serial to parallel conversion is done in independent process which simultaneously maps the parallel, grouped bits in to complex constellation for QPSK. Two different arrays are utilized to store real and imaginary parts of constellation points. Look up table is used to map grouped bits in to constellation. An algorithm is developed for IFFT calculation. Complex multiplications are involved in 8 point IFFT calculations. Twiddle factors are stored in array, which are necessarily floating point. The floating point complex multiplications are done using the algorithm. IFFT calculator is multistate process which uses different IP cores for specified operations. Two separate systems are designed using 4 subcarriers and 8 subcarriers. To have different number of subcarriers means changing the points in IFFT calculation A separate process maps the real and imaginary values in the range of DAC and transmits them through separate channels. DAC used for this purpose is LTC 2624 which is interfaced with FPGA through SPI communication. For transmitting particular OFDM symbol through DAC it must be converted in the format that is acceptable by DAC. Then this control word is transmitted serially to DAC with control signals to start the conversion.

B. Comparative analysis of transmitters

The presented system is implemented using 4 subcarriers and 8 subcarriers. The effect of increasing the number of subcarriers on computational complexity to calculate the IFFT is shown in Table I. Number of clock cycles required to perform the operation are given in Table I.

TABLE I. COMPARATIVE ANALYSIS OF TRANSMITTERS

Parameter	4 point	8 point
Number of Complex multiplications	1	5
Number of Complex additions	8	24
Number of divisions	8	16
Number of clocks for multiplications	1	95
Number of clocks for additions	2	75
Number of bits in an OFDM symbol	8	16

Table I describes the number of operations needed to be performed on a one set of data to convert it from bit stream to OFDM symbol. The comparison shows that 8 point system requires more time but it receives more number of bits per OFDM symbol.

IV. RECEIVER DESIGN

In presented system the receiver is designed on the same board. The data after the IFFT block is directly given to FFT block of receiver. The design flow of the receiver and the comparison between 4 point and 8 point receiver is presented in following sections.

A. Receiver design flow

Receiver is designed separately before connecting it to transmitter. FFT is calculated using an algorithm developed for transmitter. DIT radix-2 butterfly is used to calculate FFT and IFFT. The receiver is designed on Xilinx Project Navigator using VHDL coding. Similar to transmitter, receiver also uses the IP cores for floating point complex multiplication, additions and subtractions. Receiver operations are broken in to different processes and merged to have complete system. After FFT operation, demodulation is done for demodulation look up table approach is used. Once the bits are recovered from the received constellation, the reception is completed. For 4 point and 8 point transmitter separate receivers are designed and tested. Once the design code is ready it is simulated on ISE simulator for timing analysis and then synthesized on kit.

B. Comparative analysis of receivers

Receivers are different for 4 point transmitter and 8 point transmitter. This section presents the comparison of different receivers on the basis of computational complexity and time in terms of number of clock cycles required to perform the operation. The comparison is shown in Table II.

TABLE – II : COMPARATIVE ANALYSIS OF RECEIVERS

Parameter	**4 point**	**8 point**
Number of Complex multiplications	1	5
Number of Complex additions	8	24
Number of divisions	8	16

Number of clocks for multiplications	1	95
Number of clocks for additions	2	75
Number of bits in an OFDM symbol	8	16

Table II describes the number of operations needed to be performed on a one set of data to convert it from bit stream to OFDM symbol. The comparison shows that 8 point system requires more time than 4 point system but it receives more number of bits per OFDM symbol.

Both the transmitters and receiver are simulated on ISE before they are synthesized on FPGA. ISE results help in understanding the timing ambiguities and in resolving the conflicts between different internal signals. Next section explains the results of simulation based on that the conclusion can be made.

V. RESULTS

The designed system on Project navigator is simulated on ISE simulator. The results are tested section wise i.e. transmitter is tested first and then receiver. Then total system is simulated. Simulation results are presented here and comments are given after each result. The resources utilized by the system are also shown in this section. Resources utilization per subsystem is given.

A. Transmitter side results

Code for transmitter is simulated on ISE simulator which gives the results as shown in Figure 2.

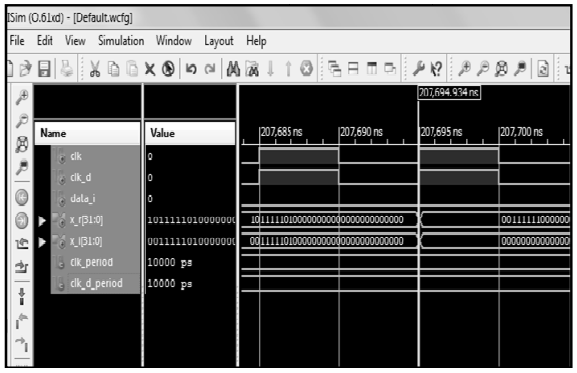


Figure – 1 : Data after IFFT operation, x_r and x_i represents the OFDM samples.

The above Figure 2 shows the output of simulator x_r and x_i represents the real and imaginary arrays that hold the complex data after IFFT is done. This data is transmitted one by one at each clock through DAC. IEEE 754 standard is used to represent the floating point results. Above fig.2 is taken from output window of simulator.

The number of clock cycles required to form the OFDM symbol can be counted from the output window. Based on timing analysis the performance of the system can be analyzed.

B. DAC Testing

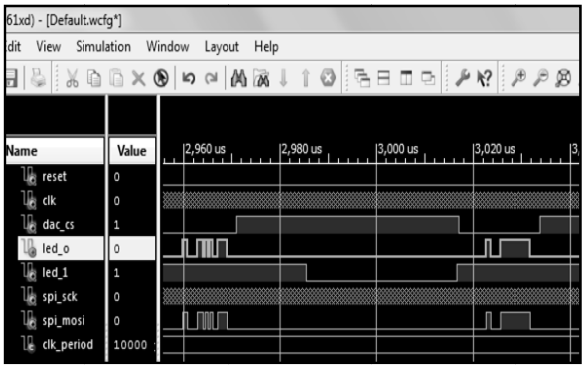


Figure 3. Serial Output from FPGA to DAC over soi_mosi pin.

Once the OFDM symbol is formed through IFFT then remaining task is to represent it in time domain, this can be done through DAC. Serial communication is used for data transfer between FPGA and DAC. The data to be transmitted serially through output pin is shown in Figure3. Synchronization between the clock, control signal and the data can be seen in the Figure 3. Receiver side results

Receiver should ideally detect all the bits it is receiving. The result of simulation shows that receiver decodes every bit correctly. The result of simulation is shown in Figure 4.

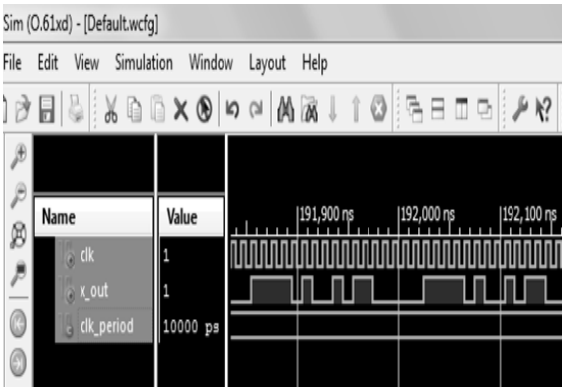


Figure 4. Output of receiverafter all operations performed on OFDM signal

Figure 4 shows the received bit stream which is exactly same as the transmitted bit stream. The received data is shown with respect to clock. Resources utilization for transmitter and receiver is shown in next section

D. Resources utilization

1) Transmitter side utilization

Xilinx synthesis tool generates the device utilization summary in synthesis report after synthesis of VHDL code. Table III shows transmitter side FPGA device utilization.

TABLE – III : DEVICE UTILIZATION SUMMARY FOR TRANSMITTER

Logic Utilization	used	present	Usage
Number of Slice Flip Flops	3195	9312	34%
Number of 4 input LUTs	7828	9312	84%
Number of occupied slices	4448	4656	95%
Number of slices containing related logic	4448	4448	100%
Number of slices containing unrelated logic	0	4448	0%
Total Number of 4 input LUTs	7924	9312	85%
Number of bonded IOBs	74	232	31%
Number of MULT 18X18 SIOs	4	20	2%

2) Receiver side utilization

Hardware resources utilization of Spartan 3 device is shown in Table IV.

TABLE-IV : DEVICE UTILIZATION SUMMARY FOR RECEIVER

Logic Utilization	used	present	Usage
Number of Slice Flip Flops	3228	9312	34%
Number of 4 input LUTs	5366	9312	57%
Number of occupied slices	4032	4656	86%
Number of slices containing related logic	4032	4032	100%
Number of slices containing unrelated logic	0	4448	0%
Total Number of 4 input LUTs	5437	9312	58%
Number of bonded IOBs	4	232	1%
Number of MULT 18X18 SIOs	4	20	2%

VI. CONCLUSION

Main aim of this paper is hardware implementation of OFDM system on Spartan 3 FPGA using VHDL language for designing the system. The system design procedure, tools and results are discussed in this paper. From the results of simulation it can be concluded that as we increase the number of subcarriers in the system, processing time required to calculate IFFT and FFT also increases. The advantage of increasing the subcarrier is the increased spectral efficiency of the system.

Results show that the system is working correctly. Device utilization of the transmitter and receiver shows that the device is utilized well bellow its capacity. Further by increasing the number of subcarriers and by making highly pipelined architecture for IFFT and FFT the system performance could be improved in terms of processing time required in transmitter and receiver.

ACKNOWLEDGMENT

The authors are thankful to the College of Engineering Pune, for the infrastructure and the support provided during the completion of this work.

REFERENCES

[1] Farzad Manavi, Yousef R. Shayan, "Implementation of an OFDM modem for the physical layer of IEEE 802.11a standard based on Xilinx Virtex-II FPGA.",0-7803-8255-2/04 2004 IEEE.

[2] Wen Fan, Chui-sing Choy, "Robust, low complexity and energy efficient downlink baseband receiver design for MB-OFDM UWB system.", IEEE transactions on circuits and system 1549-8328 2011 IEEE.

[3] Joaquin Garcia, Rene Cumplido, "On the design of an FPGA-based OFDM modulator for IEEE 802.16.", proceedings of 2005 international conference on reconfigurable computing and FPGAs.

[4] R.W. Chang, "Synthesis of band limited Orthoganal Signals for Multichannel data transmission," Bell system Technical Journal, Vol. 45, pp. 1775-1776, December1996

[5] Salzberg, B. R., "Performance of an efficient parallel data transmission system," IEEE transaction on communication, Vol. COM-15, pp. 805-813,December 1967.

[6] Weinstein, S. B., and P. M. Ebert, "Data transmission by Frequency Division Multiplexing using the DFT," IEEE Transaction on Communication, Vol. COM – 19, pp. 628-634, October 1971

[7] Horosaki B., "An orthogonally multiplexed QAM system using the DFT," IEEE Transaction on Communication Vol. COM-29, pp. 982-989, July 1981.

[8] Mosier, R. R., and R. G. Clabaugh, "Kineplex, a bandwidth efficient binary transmission system," AIEE Transaction Vol. 76. pp., 723-728, January 1958.

[9] R.Van Nee, R. Prasad ,"OFDM for wireless Multimedia Communications.", Publication by Artech House.

[10] Horosaki B., "A 19.2 kbits voice band data modem based on Orthogonality multiplexed QAM Techniques." Proceeding Of IEEE ICC'85, pp. 982-989,1985.

◈◈◈

Design and Implementation of IEEE-754 Addition and Subtraction for Floating Point Arithmetic Logic Unit

V. Vinay Chamkur & Chethana. R

VLSI DESIGN and EMBEDDED SYSTEMS,
ECE Dept, SJBIT Bengaluru-60

Abstract - This paper describes the FPGA implementation of a Decimal Floating Point (DFP) adder/subtractor using IEEE 754-2008 format. In this paper we describe an efficient implementation of an IEEE 754 single precision Standard for Binary Floating-Point Arithmetic to include specifications for decimal floating-point arithmetic. As processor support for decimal floating-point arithmetic emerges, it is important to investigate efficient algorithms and hardware designs for common decimal floating-point arithmetic algorithms. This paper presents novel designs for a decimal floating-point addition and subtraction. They are fully synthesizable hardware descriptions in VERILOG. Each one is presented for high speed computing.

Keywords- *IEEE-754 Floating Point Standard; Addition and Subtraction Algorithm.*

I. INTRODUCTION

Floating point numbers are one possible way of representing real numbers in binary format; the IEEE 754 [1] standard presents two different floating point formats, Binary interchange format and Decimal interchange format. Multiplying floating point numbers is a critical requirement for DSP applications involving large dynamic range. This paper focuses only on single precision normalized binary interchange format. Fig. 1 shows the IEEE 754 single precision binary format representation; it consists of a one bit sign (S), an eight bit exponent (E), and a twenty three bit fraction (M or Mantissa). An extra bit is added to the fraction to form what is called the significand[1]. If the exponent is greater than 0 and smaller than 255, and there is 1 in the MSB of the significand then the number is said to be a normalized number; in this case the real number is represented by (1)

The IEEE-754 standard specifies six numerical operations: addition, subtraction, multiplication, division, remainder, and square root. The standard also specifies rules for converting to and from the different floating-point formats (e.g short/integer/ long to /from single/double/quad-precision), and conversion between the different floating-point formats.

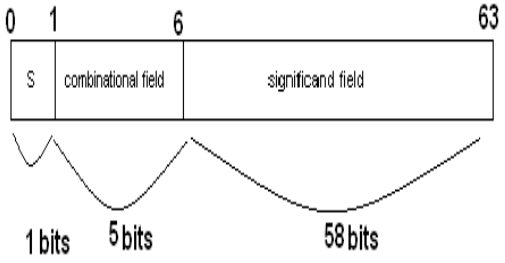


Figure 1.IEEE floating point format

$$Z = (-1^{S}) * 2^{(E - \underline{Bias})} * (1.M)_1$$

Where $M = m_{22}\, 2^{-1} + m_{21}\, 2^{-2} + m_{20}\, 2^{-3} + \ldots + m_1\, 2^{-22} + m_0\, 2^{-23}$; $\underline{Bias} = 127$.

FIG : -1

a) 1-bit sign s.

b) A w + 5 bit combination field G encoding

classification and, if the encoded datum is a finite number,the exponent q and four significand bits (1 or 3 of which are implied). The biased exponent E is a w + 2 bit quantity q + bias, where the value of the first two bits of the biased exponent taken together is either 0, 1, or 2.

c) A t-bit trailing significand field T that contains

J × 10bits and contains the bulk of the significand. J represents the number of depletes.

When this field is combined with the leading significand bits from the combination field, the format encodes a total of p = 3 × J + 1 decimal digits. The values of k, p, t, w, and bias for decimal64 interchange formats are 16, 50, 12, and 398 respectively. That means that number has p=16 decimal digits of precision in the significand, an unbiased exponent range of [383, 384], and a bias of 398.

The IEEE-754 standard specifies six numerical operations: addition, subtraction, multiplication, division, remainder, and square root. The standard also specifies rules for converting to and from the different floating-point formats (e.g short/integer/long to/from single/double/quad-precision), and conversion between the different floating-point formats, AX and BX are the significands and EAX, EBX and EX are the exponents respectively. X is a digit that denotes the outputs of different units. The symbol $(N)_Z^T$ refers to T^{th} bit of the Z^{th} digit in a number N, where the least significant bit and the least significant digit have index 0. For example, (A1)2

5 is the fifth bit of the second BCD digit in A1.

II. DECIMAL FLOATING POINT IN IEEE 754-2008:-

The primary difference between two formats, besides the radix, is the normalization of the significands (coefficient or mantissa). BFP significands are normalized with the radix point to the right of the most significant bit (MSB), while DFP mantissa are not required to be normalized and are represented as integers. The mantissa is encoded in densely packed decimal,

The exponent must be in the range [emin, emax], when biased by bias. Representations for infinity and not-a number (NaN) are also provided.

$$D = -1^{s} \times C \times 10^{q}, \quad q = E - \text{bias}$$

Where s is the sign bit, C is the non-negative integer Significand and q the exponent. The exponent q is obtained as a function of biased non-negative integer exponent E.

The mantissa is encoded in densely packed decimal [3], the exponent must be in the range [emin, emax], when biased by bias. Representations for infinity and not-a-number (NaN) are also provided. Representations of floating-point numbers in the decimal interchange formats are encoded in k bits in the following three fields (Fig1):

III. DECIMAL FLOATING-POINT ADDER/ SUBTRACTOR IMPLEMENTATION

A general overview of proposed adder/subtractor is described below. For the best performance, the design presents eight pipelined stages as is exhibited in the Fig. 2.

Arrows are used to show the direction of data flow, the dashed blocks indicate the main stages of the design, and the dotted line indicates the pipeline.

This architecture was proposed for the IEEE 754-2008 decimal64 format and can be extended for the decimal128 format. The adder/subtractor on decimal64 is carried out as follows: The decoder unit takes the two 64-bit IEEE 754-2008 operands (OP1, OP2) to generate the sign bits (SA, SB), 16-digit BCD significant (A0, B0), 10-bit biased exponents (EA, EB), the effective operation (EOP) and flags for specials values of NaN or infinity. The signal EOP defines the effective operation (EOP = 0 for effective

addition and EOP = 1 for effective subtraction), this signal is calculated as:

$$EOP = SA \text{ xor } SB \text{ xor } OP \text{-------- (2)}$$

As soon as possible the decoded significant become available, the leading zero detection unit (LZD) takes these results and computes the temporary exponents (EA1, EB1) and the normalized coefficients (A1, B1). The swapping unit swaps the operands (A1, B1) if EA1 < EB1 and Generates the BCD coefficients A2 (with higher exponent,max(EA1, EB1)) and B2 (with lower exponent, min(EA1,EB1)). In parallel with the above mentioned, this unit generates an exponent difference (Ed = |EA1 - EB1|), the exponent E2 = max(EA1, EB1), the SWAP flag if a swapping process is carried out, and the right shift amount (RSA) which indicates how many digits B2 should be right shifted in order to guarantee that both coefficients (A2, B2) have the same exponent.

The RSA is computed as follows:

```
if (Ed <= p_max)
    RSA = Ed
else RSA = p_max
```

The value p_max = 18 digits, RSA is limited to this value since B2 contains 16 digits plus two digits which will be processed to compute the guard and round digit.

Next, the Shifting unit receives as inputs the RSA, and the significand B2 generating a shifted B2 (B3) and a 2-bit signal called predicted sticky-bit (PSB) that will

predict two initials sticky bits. PSB and B3 will be utilized as inputs in the decimal addition, control signals generation and post-correction units, respectively.

The outputs above mentioned plus two signals, significand A2 and EOP, are taken as inputs in the control signals generation unit and generates the signals necessary to perform an addition or subtraction operation, these signals are described in the Sub-section 3.4 and are made up of a prior guard digit (RD2), the final partial sum (S2) and the corrected exponent (E3). digit (GD1), a prior round digit (RD1), an extra digit (ED), a signal which verifies if A2 > B3(AGTB) and a carry into (CIN).

The significant BCD (A2, B3) and the CIN are inputs the decimal addition unit generating the partial sum of magnitude |S1| = |A2 + (-1) EOP B3| and a carry out (COUT), respectively.

At once, the 16-digit decimal addition unit takes the A2, B3, EOP and CIN and computes S1 as follows:

S1 = A2 + B3 if EOP = 0, S1 = A2 + cmp9 (B3)

if EOP = 1 and A2 >= B3, and

S1 = cmp9 A2+cmp9(B3)) if EOP = 1 and A2 < B3.

The symbol cmp9 means the 9`s complement.

The post-correction unit uses as inputs the PSB, the Exponent E2, GD1, RD1, ED, the partial sum S1 and COUT to verify, correct and compute the inputs signals if only the following two cases occur: 1) COUT=1 and EOP=0 and 2) (S1)15=0 and (GD1 > 0) and (EOP=1).

The analysis is explained in the Sub-section 3.5.

This unit generates the final sticky bit (FSB), the corrected guard digit (GD2) and round Next, the Rounding unit takes the outputs of the prior unit and rounds S2 to produce the result´s significand S3 and adjusts the exponent E3 to calculate the final exponent E4. Simultaneously the overflow, underflow and sign bit signals

The final sign bit is computed as:

FS = (SA ^ ~EOP) V (EOP ^ (AGTB ⊕ SA ⊕ SWAP))

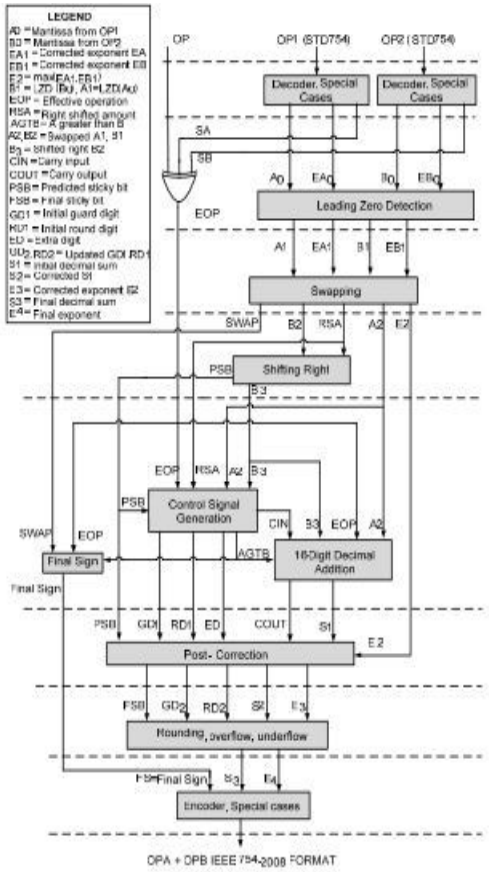


Figure 2: implementaition diagram are generated

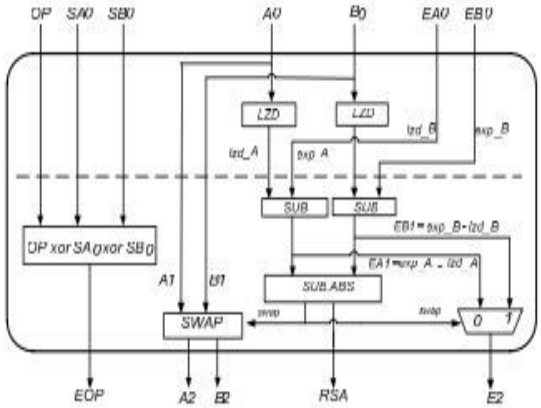


Fig. 3. Alignment and swapping unit

IV. PROBLEMS ASSOCIATED WITH FLOATING POINT ADDITION & SUBTRACTION

For the input the exponent of the number may be dissimilar. And dissimilar exponent can't be added directly. So the first problem is equalizing the exponent. To equalize the exponent the smaller number must be increased until it equals to that of the larger number. Then significant are added. Because of fixed size of mantissa and exponent of the floating-point number cause many problems to arise during addition and subtraction. The second problem associated with overflow of mantissa. It can be solved by using the rounding of the result. The third problem is associated with overflow and underflow of the exponent. The former occurs when mantissa overflow and an adjustment in the exponent is attempted the underflow can occur while normalizing a small result. Unlike the case in the fixed-point addition, an overflow in the mantissa is not disabling; simply shifting the mantissa and increasing the exponent can compensate for such an overflow. Another problem is associated with normalization of addition and subtraction. The sum or difference of two significant may be a number, which is not in normalized form. So it should be normalized before returning results.

V. ADDITION AND SUBTRACRIO ALGORITHM

Let a1 and a2 be the two numbers to be added. The notations ei and si are used for the exponent and significant of the addends ai. This means that the floating-point inputs have been unpacked and that si has an explicit leading bit. To add a1 and a2, perform these eight steps:

1. If e1 < e2, swap the operands. This ensures that the difference of the exponents satisfies

d = e1–e2 □ 0. Tentatively set the exponent of the result to e1.

2. If the sign of a1 and a2 differ, replace s2 by its two's complement.
3. Place s2 in a p-bit register and shift it d = e1-e2 places to the right (shifting in 1's if the s2 was complemented in previous step). From the bits shifted out, set g to the most- significant bit, r to the next most-significant bit, and set sticky bit s to the OR of the rest.
4. Compute a preliminary significant S = s1+s2 by adding s1 to the p-bit register containing s2. If the signs of a1 and a2 are different, the most-significant bit of S is 1, and there was no carry out then S is negative. Replace S with its two's complement. This can only happen when d = 0.
5. Shift S as follows. If the signs of a1 and a2 are same and there was a carry out in step 4, shift S right by one, filling the high order position with one (the carry out). Otherwise shift it left until it is normalized. When left shifting, on the first shift fill in the low order position with the g bit. After that, shift in zeros. Adjust the exponent of the result accordingly.
6. Adjust r and s. If S was shifted right in step 5, set r: = low order bit of S before shifting and s: = g or r or s. If there was no shift, set r: = g, s: = r. If there was a single left shift, don't change r and s. If there were two or more left shifts, set r: = 0, s: = 0. (In the last case, two or more shifts can only happen when a1 and a2 have opposite signs and the same exponent, in which case the computation s1 + s2 in step 4 will be exact.)

7. Round S using following rounding rules as in Table

Rounding	*Sign of result*	*Sign of result*
-∞		+1 if $r \vee s$
0		
Nearest	+1 if $r \wedge p_0$ or	+1 if $r \wedge p_0$ or

If a table entry is non empty, add 1 to the low order bit of S. If rounding causes carry out, shift S right and adjust the exponent. This is the significant of the result.

8. Compute the sign of the result. If a1 and a2 have the same sign, this is the sign of the result. If a1 and a2 have different signs, then the sign of the result depends on which of a1, a2 is negative, whether there was a swap in the step 1 and whether S was replaced by its two's complement in step 4. As in table below

Swap	*Complement*	*Sign (a1)*	*Sign (a2)*	*Sign (result)*
Yes	∅	+	-	-
Yes	∅	-	+	+
No	No	+	-	+
No	Yes	-	+	-
No	Yes	+	-	-
No		-	+	+

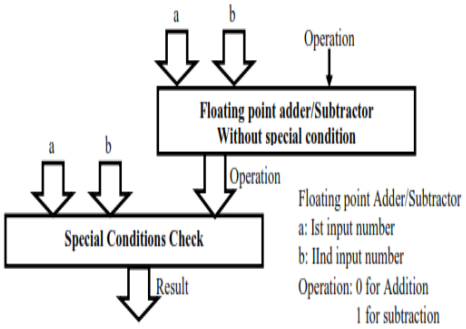
VI. SPECIAL CONDITIONS

Some special conditions are checked before processing. If any condition is met then we have no need to calculate the result by normal procedure. Results are directly calculated. So all the operations are

bypassed when any such condition is met.

1. If a1 = 0 and a2 = 0 then result will be zero.
2. If a1 = a2 and sign of a1 □ sign of a2 then result will be again zero.
3. If a1 = 0 and a2 □ 0 then result will be equal to a2.
4. If a2 = 0 and a1 □ 0 then result will be equal to a1.
5. If d = |e1 – e2| > 24 then result will be equal to larger of a1 and a2.

VII.Hardware Approach



The block diagrams of the architecture used for combinational adder is shown above in Figure 4, step by step from the lower abstract level to the higher abstract level.

B Unsigned Adder (for exponent addition)

This unsigned adder is responsible for adding the exponent of the first input to the exponent of the second input and subtracting the Bias (127) from the addition result (i.e. A_exponent + B_exponent - Bias). The result of this stage is called the intermediate exponent. The add operation is done on 8 bits, and there is no need for a quick result because most of the calculation time is spent in the significand multiplication process (multiplying 24 bits by 24 bits); thus we need a moderate exponent adder and a fast significand multiplier.

An 8-bit ripple carry adder is used to add the two input exponents. As shown in Fig. 3 a ripple carry adder is a chain of cascaded full adders and one half adder; each full adder has three inputs (A, B, C_I) and two outputs (S, C_O). The carry out (C_O) of each adder is fed to the next full adder (i.e each carry bit "ripples" to the next full adder).

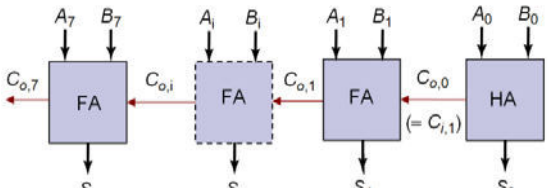


Figure 5 : - ripple carry adder

The addition process produces an 8 bit sum (S_7 to S_0) and a carry bit ($C_{O,7}$). These bits are concatenated to form a 9 bit addition result (S_8 to S_0) from which the Bias is subtracted. The Bias is subtracted using an array of ripple borrow subtractors. The above table shows one bit subtractor logic can be optimized if one of its inputs is a constant value which is our case, where the Bias is constant ($127|_{10} = 001111111|_2$).

S	T	B_i	Difference (R)	B_O
0	1	0	1	1
1	1	0	0	0
0	1	1	0	1
1	1	1	1	1

VIII.SIMULATION RESULTS

FLOATING POINT ADDER

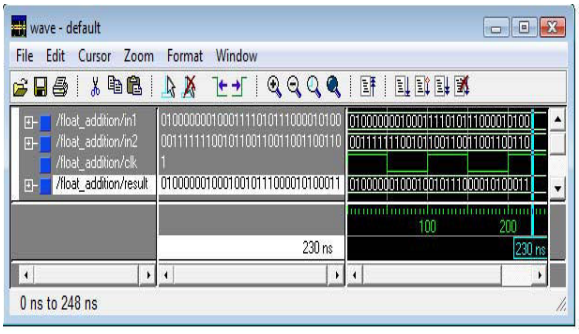
Input 1 = 01000000010001111010111000010100 (3.120 10)

Input 2 = 00111111100101100110011001100110 (1.175 10)

Required Result = 01000000100010010111000010100011

(4.295 10)

Obtained Result = 01000000100010010111000010100011 (4.295 10)

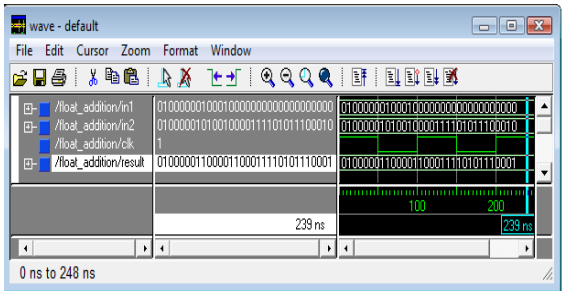


Input 1 = 01000000100010000000000000000000 (4.25 10)

Input 2 = 01000001010010000111101011100010 (12.53 10)

Required Result = 01000001100001100011110101110000 (16.78 10)

Obtained Result = 01000001100001100011110101110001 (16.78 10)



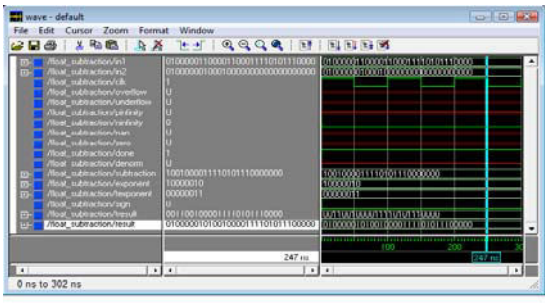
FLOATING POINT SUBTRACTOR

Input 1 = 01000001100001100011110101110000 (16.78 10)

Input 2 = 01000000100010000000000000000000 (4.25 10)

Required Result = 01000001010010000111101011100010 (12.53 10)

Obtained Result = 01000001010010000111101011100010 (12.53 10)

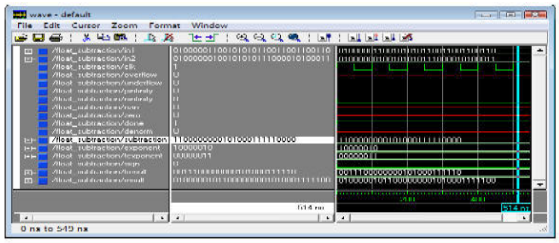


Input 1 = 01000001100101010110011001100110 (18.675 10)

Input 2 = 01000000100101010111000010100011 (4.670 10)

Required Result = 01000001011000000001010001111101 (14.005 10)

Obtained Result = 01000001011000000001010001111101 (14.005 10)



IX. IMPLEMENTATION AND TESTING

The whole adder (top unit) was tested against the Xilinx floating point adder core generated by Xilinx coregen. Xilinx core was customized to have two flags to indicate overflow and underflow, and to have a maximum latency of three cycles. Xilinx core implements the "round to nearest" rounding mode.

A testbench is used to generate the stimulus and applies it to the implemented floating point adder and to the Xilinx core then compares the results. The floating point multiplier code was also checked using DesignChecker [7]. DesignChecker is a linting tool which helps in filtering design issues like gated clocks, unused/undriven logic, and combinational loops. The design was synthesized using Precision synthesis tool [8] targeting Xilinx Virtex-5 5VFX200TFF1738 with a timing constraint of 300MHz. Post synthesis and place and route simulations were made to ensure the design functionality after synthesis and place and route. shows the resources and frequency of the implemented floating point multiplier and Xilinx core

	Technology	Clk (GHz)	Cycles	Delay (ns)	Mops/ sec
SW	Itanium2 [18]	1.4	219	156.4	6.4
	Xeon5100 [19]	3.0	133	44.3	22.6
	Xeon [18]	3.2	249	77.8	12.9
	Pentium M [21]	1.5	848	565.3	1.8
HW	Power6 [22]	5.0	17	3.4	294.1
	Z10 [23]	4.4	12	2.7	366.7
	BID 65nm [10]	1.3	3-13	10.0	100.0
	BID Virtex 5 [9]	0.16	13-18	109.8	9.1
	Proposed Virtex5	0.2	8	40	200.0

The area of Xilinx core is less than the implemented floating pointadder because the latter doesnot truncate/round the 48 bits result of the mantissa multiplier which is reflected in the amount of function generators and registers used to perform operations on the extra bits; also the speed of Xilinx core is affected by the fact that it implements the round to nearest rounding mode.

X. CONCLUSIONS AND FUTURE WORK

This paper deals with development of a Floating Point adder and subtractor for ALU in VHDL and verilog with the help of ModelSim and synthesized with Xilinx tools. Simulation results of all the designed programs have been carried out for various inputs with the help of ModelSim tool. Both are available in single cycle and pipeline architectures and fully synthesizable with performance comparable to other available high speed implementations. The design is described as graphical schematics and VHDL code.

This dual representation is very valuable as allows for easy navigation over all the components of the units, which allows for a faster understanding of their interrelationships and the different aspects of a Floating Point operation.

REFERENCES

[1] M. F. Cowlishaw, "Decimal floating-point: algorism for computers," in *Proc. 16th IEEE Symp. Computer Arithmetic*, 2003, pp. 104–111.

[2] E. M. Schwarz, J. S. Kapernick, and M. F. Cowlishaw, "Decimal floating-point support on the IBM System z10 processor," 2009, iBM Journal of Research and Development.

[3] IEEE Standard for Floating-Point Arithmetic," pp. 1–58, 2008, iEEE Std 754-2008.

[4] F. Y. Busaba, C. A. Krygowski, W. H. Li, E. M. Schwarz, and S. R. Carlough, "The IBM z900 decimal arithmetic unit," in *Proc. Conf Signals, S ystems and Computers Record of the Thirty-Fifth Asilomar Conf*, vol. 2, 2001, pp.1335–1339.

[5] W. Haller, K. Ulrich, L. Thomas, and H. Wetter, "Combined binary/decimal adder unit," in *InternationalBusiness Machines Corporation (Armonk, NY)*, 1999.

[6] G. Bohlender and T. Teufel, "BAP-SC: A Decimal Floating-Point Processors for Optimal Arithmetic," in *Computerarithmetic: Scientific Computation*

[7] J. Thompson, N. Karra, and M. J. Schulte, "A 64-bit decimal floating-point adder," in *Proc. IEEE Computer society Annual Symp. VLSI*, 2004, pp. 297–298.

[8] M. S. Cohen, T. E. Hull, and V. C. Hamacher, "CADAC: A Controlled-Precision Decimal Arithmetic Unit," no. 4, pp. 370–377, 1983.

[9] A. Farmahini-Farahani, C. Tsen, and K. Compton, "FPGAimplementation of a 64-Bit BID-based decimal floating- point adder/subtractor," in *Proc. Int. Conf. F ield- Programmable Technology FPT 2009*, 2009, pp. 518–521.

[10] C. Tsen, S. Gonzalez-Navarro, and M. Schulte, Hardware design of a Binary Integer Decimal-based floating-point adder," pp. 288–295, 2007, computer Design, 2007. ICCD 2007. 25th International Conference on.

[11] C. Minchola and G. Sutter, "A FPGA IEEE-754-2008 Decimal64 Floating-Point Multiplier," in *Proc. Int. Conf. Reconfigurable Computing and FPGAs ReConFig '09*, 2009, pp. 59–64.

[12] L.-K. Wang and M. J. Schulte, "Decimal Floating-Point Adder and Multifunction Unit with Injection-Based Rounding," in *Proc. 18th IEEE Symp. Computer Arithmetic ARITH '07*, 2007, pp. 56–68.

[13] M. Vazquez, G. Sutter, G. Bioul, and J. P Deschamps, Decimal Adders/Subtractors in FPGA: Efficient 6-input LUT Implementations," in *Proc. Int. Conf. Reconfig urable Computing and FPGAs ReConFig '09*, 2009, pp. 42–47.

[14] *Xilinx Inc. XST User G uide 12.1*, v12.1 ed., Xilinx Inc., June 2009. [Online]. Available: http://www.xilinx.com

[15] *Xilinx Inc. Xilinx ISE D esign Suite 12.1 Software Manuals*, v12.1 ed., Xilinx Inc., June 2009. [Online]. Available:http://www.xilinx.com

[16] *Xilinx Inc. Virtex-5 Li braries Guide for VHDL design*, v12.1 ed., Xilinx Inc., June 2009. [Online]. Available: http://www.xilinx.com

[17] Xilinx Inc, *DS335: Floating-Point Operator v5.0*June 2009.

[18] M. Cornea, J. Harrison, C. Anderson, P. Tang, E. Schneider, and E. Gvozdev, "A software implementation of the ieee 754r decimal floating-point arithmetic using the binary encoding format," pp.148–162, 2009, computers, IEEE Transactionson.

[19] M. Cornea, C. Anderson, J. Harrison, P. T. P. Tang, E. Schneider, and C. Tsen, "A software implementation of the ieee 754r decimal floating-point arithmetic using the binary encoding format," pp. 29–37, 2007, computer Arithmetic, 2007. ARITH '07. 18th IEEE Symposium on.

[20] *The decNumber C library*, v3.68 ed., IBM UK Laboratories, January 2010. [Online]. Available: http://- speleotrove.com/decimal/decnumber.pdf

[21] M. Cowlishaw. (2009) Decimal library performance. [Online]. Available: http://speleotrove.com/decimal/-decperf.pdf

◈◈◈

Design of Wide Band and High Gain Efficient Microstrip Antenna using Sierpinski Carpet Fractal

[1]Prabhat Ranjan Mishra, [2]Taimoor Khan and [3]Adesh Arya

[1-3] Department of Electronics and Comm. Engg, Shobhit University Meerut-250110
[2]Department of Electronics and Comm. Engg, Delhi Technological University Delhi-110042

Abstract **-** In this paper authors propose a sierpinski carpet fractal stacked rectangular microstrip antenna. By introducing sierpinski carpet fractal in the stacked rectangular microstrip antenna, the size of the antenna is reduced significantly and simultaneously the radiation parameters like gain, directivity, antenna efficiency radiation efficiency and impedance bandwidth are also improved.

I. INTRODUCTION

In modern wireless communication systems wider bandwidth and low profile antennas are in great demand for both commercial and military applications [1]. This has initiated antenna research in various directions. Recently fractal shaped antenna elements have been introduced for enhancement of different radiation characteristics. Traditionally different antennas operating at different frequencies are needed for numerous applications. This causes a space and orientation problem. In order to overcome this problem, multiband antenna can be used, where a single antenna can operate at many frequency bands [2-3].

One technique to construct a multiband antenna is by applying fractal shape into antenna geometry. It is well known that one of the most important characteristics of fractals is size reduction and space-filling. Therefore, traditional fractals have been used to design compact antennas for multiband or broadband operation [4-6]. With the introduction of fractals in antenna engineering, Hajihshemi and Abiri [7] investigated Ant with fractal shape and reported that with increase in fractal iteration the ratio of surface to volume in Ant increases and thereby enhance the Q-factor which tends to increase in antenna impedance bandwidth.

II. DESCRIPTION OF PROPOSED ANTENNA DESIGN

The geometry of proposed antenna is shown in Figure 1.

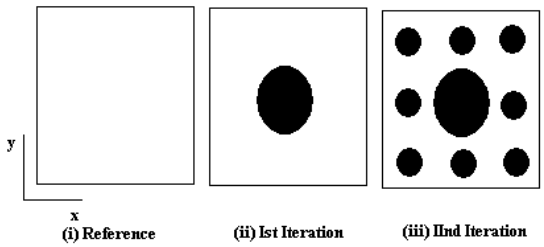


Figure (1): Geometry of Proposed Antenna Design

The study began with the design of a Square patch radiator of length L=71.32 mm on FR4 substrate with a relative dielectric constant (ε_r) of 4.4 and loss tangent of 0.02. An air gap of 1.588 mm is inserted between ground plane and substrate. A strip feed extends from the microstrip line of the same width that ends with the 50 ohm line after certain length that is used as a matching.

III. SIMULATED RESULTS

The first iteration of the fractal antenna is constructed by etching a circular slot of radius 11.8858 mm at the center of patch. In second iteration eight circular slots of 3.961945 mm radius each that follows the same procedure adopted in first iteration. And in the same fashion the third iteration is also implemented. The simulated results for reference patch, modified patch for first iteration and second iteration are given in Table1. A drastic enhancement in the performance parameters is observed by replacing the infinite ground plane by a finite ground plane of dimensions three times higher than that of patch dimensions. The variations in the S-parameters, Gain, Directivity, Antenna Efficiency and Radiation Efficiency have also been plotted in Figures (2)-(5) respectively.

IV. CONCLUSIONS

It is clear from the Table 1 that resonant frequencies have been reduced by 1.954% and 1.152% by inserting the second iteration in the reference patch on a finite ground plane. The gains corresponding to first and second frequency have been increased by 22% and 30.9%, antenna efficiencies by 41.79% and 66.94% radiation efficiencies by 41.8% and 65.837%, bandwidths by 96.87% and 278.26% only at the cost of 7.5% and 5.12% reduction in the directivities corresponding to these frequencies.

.

Table1: Variation in Parameters by Inserting Fractals

Reference (Air Gap 1.588mm)			**II[nd] Iteration**		
Frequency (GHz)	3.0700	3.4700	Frequency (GHz)	2.9700	3.3100
Return Loss (dB)	-26.400	-16.5200	Return Loss (dB)	-11.0800	-11.0000
Directivity(dBi)	6.6000	7.8000	Directivity(dBi)	6.7000	7.40000
Gain(dBi)	5.0000	5.7000	Gain(dBi)	4.6000	4.6000
Efficiency (A) (%)	70.0000	60.0000	Efficiency (A) (%)	61.0000	53.0000
Efficiency (R) (%)	70.0000	60.0000	Efficiency (R) (%)	66.0000	58.0000
Band Width (%)	3.2000	2.3000	Band Width (%)	2.3500	1.2120
I[st] Iteration			**II[nd] Iteration with Finite Ground**		
Frequency (GHz)	3.0600	3.3700	Frequency (GHz)	3.0100	3.4300
Return Loss (dB)	-24.9400	-17.0000	Return Loss (dB)	-24.1300	-17.5600
Directivity(dBi)	6.8000	7.7000	Directivity(dBi)	6.2000	7.4000
Gain(dBi)	5.1000	5.6000	Gain(dBi)	6.2000	7.2400
Efficiency (A) (%)	69.0000	60.0000	Efficiency (A) (%)	100.0000	98.5000
Efficiency (R) (%)	69.0000	61.0000	Efficiency (R) (%)	100.0000	100.0000
Band Width (%)	3.2600	3.5900	Band Width (%)	6.3000	8.7000

.

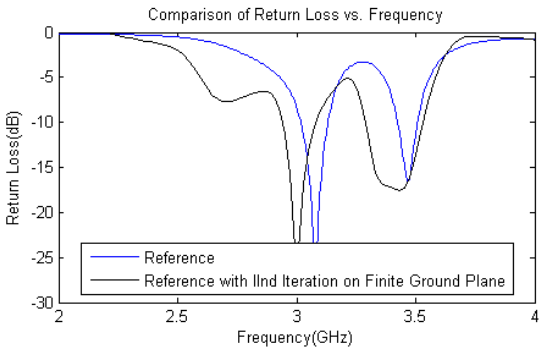


Figure (2): Variations in Return Loss vs. Frequency for Reference Patch and Reference Patch with II[nd] Iteration on finite ground plane.

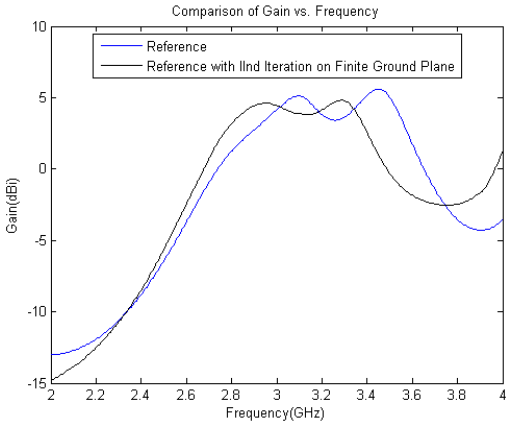


Figure (3): Variations in Gain vs. Frequency for Reference Patch, I[st] Iteration, II[nd] Iteration, and for finite ground plane.

In this paper, a printed microstrip-fed wideband high gain fractal antenna has been presented. Size

reduction and bandwidth enhancement is achieved using sierpinski carpet fractal concept. The second iteration fractal antenna is considered for implementing the antenna shape. The proposed antenna has been simulated, and parametric study has been performed using method of moments based IE3D software.

REFERENCES

[1] C.A. Balanis, "Antenna Theory", Second Edition, John Wiley & Sons, 2000.

[2] Douglas H. Werner and Suman Ganguly, "An overview of fractal antenna engineering research", IEEE Antenna and Propagation Magazine, vol. 45, no 1, pp. 38-57, February 2003.

[3] C. Puente et. al., "On behaviour of the siearpinski multiband fractal antenna," IEEE Transactions on Antenna and Propagation, pp.517-24, 1998.

[4] Kenneth Falconer, Fractal Geometry: Mathematical Foundations and Applications, 2nd edition, New York 2003.

[5] C. Puente, J. Romeu, and R. Pous *et al.*, "Small but long Koch fractal monopole," *Electron. Lett.*, vol. 34, no. 1, pp. 9–10, 1998.

[6] J. Romeu and J. Soler "On the behavior of the Sierpinski multiband fractal antenna," *IEEE Trans. Antennas Propag.*, vol. 46, no. 4, pp. 517–524, Apr. 1998.

[7] Mohammad R. Hajihashemi and Habibollah Abiri, "Parametric Study of Novel Types of Dielectric Resonator Antennas Based on Fractal Geometry", International Journal of RF and Microwave Computer-Aided Engineering vol.17, no.4, pp. 416-424, 2007.

◈◈◈

Relative Performance of BPSK & BFSK In Underground Mine Communication

M. N. Jayaram & C. R .Venugopal

Department of E&C , SJCE , Mysore570006 , Karnataka , India

Abstract **-** Wireless/Mobile signals get strongly attenuated deep inside a building/tunnel/mine. In order to study the performance parameters , underground communication channel needs to be modeled. As a first step here we have practically tried two digital modulation methods BPSK & BFSK for underground communication. Wireless link was able to support data signals. Various measurements were made up to a depth of 30m inside the mine. Beyond 30m as more than 75% of the received signal was attenuated measurements was not possible. Received signal strength, , power line noise effects are measured . SNR , Eb/No , Pe (Probability error) , penetration loss and bending loss are calculated up to 30m depth. Using these results channel modeling can be done. For the measured data we are trying to fit empirical equations .

Keywords: mine communication , BPSK ,BFSK

I. INTRODUCTION

1. The data was collected for several runs along the terrain, hence it corresponds to the average value of the signal strength measured for different runs.
2. Communication in mines is generally by TTW (Through the wire i.e line communication) using co-axial cables, twisted pairs, optical fibers, leaky waveguides / feeders etc. Only under emergency such as when there is a mine blast or fire, when line communication fails due to rubbles and other blockages TTA (Through the air i.e wireless communication) or TTE (Through the earth) communication is used.
3 Frequency selection has a great impact on signal propagation inside the mine . RF frequency band (UHF) (GSM 900) is used for establishing the link. Even though their penetration into earth is not good, UHF signals bounce round the corner's better (**Ref 1**). For better penetration we have selected higher power of 30 W . At UHF attenuation is relatively low in straight mine entries. GSM 900 band is good for line of sight communication and can turn one or two cross cuts.
4. As this is a first step towards modeling the underground communication channel, we have selected simple digital modulation techniques like BPSK & BFSK since their performance in surface communication is well known.

BPSK

Here phase of constant amplitude carrier signal is switched between two values according to the two possible signals m1 and m2 corresponding to binary 1 and 0 respectively. Normally the two phases are separated by 180 degrees. The transmitted BPSK signal is given by $V_{BPSK}(t) = m(t)\sqrt{(2E_b/T_b)}\cos(W_c t+\theta_c)$

For generating BPSK single balanced modulator and cascaded stages of power amplifiers are used to get eh requisite power level for transmission.

As 90% of the BPSK signal energy is contained within the b.w (band width), Band width = 1.6 Rb, where Rb is the data rate of transmission for rectangular input pulses. For coherent demodulation, probability error = $P_e = Q(\sqrt{2E_b/N_o})$.

BFSK

Here frequency of a constant amplitude carrier signal is switched between two values according to two states of message 0 or 1. The transmitted BPSK signal is given by $V_{BFSK}(t) = (\sqrt{2E_b/T_b})\cos(2\Pi f_c+2\Pi \square f\, t)$, $0 \leq t \leq T_b$ Where $2\Pi \square f$ is the constant offset from normal carrier frequency.

Transmission bandwidth of BFSK signals is ($2\Delta f$ +2B), where B is the bandwidth of digital base band signal. For coherent demodulation probability error = $P_e = Q(\sqrt{E_b/N_o})$

A transceiver kit fitted with loop antenna (hand held type) was used for measurement. The kit has very good sensitivity and high resolution. It can be operated in the switched mode. To improve reliability of digital microwave system protection switching and diversity system configurations are used .

Inside the mine, temperature varies from 27deg C (at mouth of mine) to 38 deg C (at a depth of 30 m) , where as humidity varies from 82 % to 94 %

II. EQUIPMENTS & SPECIFICATIONS :

To generate 30W of power with 20% efficiency in the UHF band MOSFET hybrid modules with proper heat sink is used

Transmitted power = P_T = 30 Watts = 44.77dBm

Operating frequency GSM 900 band

Transmitting antenna outside tunnel is a parabolic dish which is Cassegrain fed using a horn (obstruction free transmission and reception) , antenna gain

is 35dBi or 32.26 dBd or 316.28 (ratio), VSWR <1.4, 3dB beam width is 3.05 deg .

Height of transmitting antenna w.r.t. ground plane is = 50 m

Receiving antenna is a loop antenna (magnetic dipole)

Receiver specifications :- Receiver has sensitivity

For BPSK= - 106 dBm (minimum power detectable = 1.99 x 10exp – 14 W)

For BFSK= - 107 dBm (minimum power detectable = 2.512 x 10exp - 14 W)

Rx Band width = 280 KHz

Penetration & Deffraction loss (Bending loss) calculation

Penetration loss at the entrance :-

There was a 35 cm thick concrete brick wall at the entrance which

results in penetration loss.

For lossy dielectric medium

$\gamma^2 = (\alpha+j\beta)^2 = (\sigma+j\omega\epsilon)(j\omega\mu)$ …….. (1)

Equating real & imaginary parts we have ,

$\alpha^2-\beta^2 = -\omega^2\mu\epsilon$……. (2)

and $2\alpha\beta = \omega\mu\sigma$……..(3)

Substituting for β from eq (3) in eq (2) gives ,

$\alpha^2-(\omega\mu\sigma/2\alpha)^2 = -\omega^2\mu\epsilon$…… ……...(4)

or $\alpha\exp4+\omega^2\mu\epsilon\alpha^2-(\omega\mu\sigma/2)^2=0$……...(5)

On solving, $\alpha=\omega\sqrt{(\mu\epsilon/2*\{\sqrt{(1+(\sigma/\omega\epsilon)^2)}-1\})}$……(6)

For $\sigma/\omega\epsilon << 1$& using Binomial expansion for

$\sqrt{(1+(\sigma/\omega\epsilon)^2)} \approx 1+1/2(\sigma/\omega\epsilon)^2$…….(7)

So Attenuation constant $\alpha \approx \sigma/2\sqrt{(\mu/\epsilon)}$……...(8)

For concrete ϵr=8.9 or 9 F/m , μr=1H/m , σ =0.1 A/m² in GSM 900 band(**Ref (2), (3))**

With ϵo=8.854x10exp-12 F/m , μo=4πx10exp-7 H/m, from eq(8) , $\alpha = 60\pi\sigma/(\sqrt{\epsilon r})$ or penetration loss in dB =(20α) dB = 15.96 = 16 dB

As the electro magnetic wave (EMW) is propagating perpendicular to the earth's magnetic field with E field vector along the earth's magnetic field, the magnetic field of earth does not affect the propagation characteristics of the wave. In other words propagation constant of the wave is same as that in the absence of the earth's magnetic field (i.e. ordinary wave propagation)

Diffraction Loss :- Diffraction is bending of EMW around the corners. This allows the signals to propagate in regions that lie behind the obstructions.

Since diffracted field has sufficient field strength it can reach the Rx, get detected and establish communication link.

Approximate elevation geometry inside the mine is given in Fig 1 from which diffraction loss is calculated.

In Fig 1 , X-axis is distance inside the mine & Y-axis is gradient .

Undulations on the terrain due to digging also causes scattering, but here we assume this loss is negligible . It was observed that terrain is almost flat with negligible curvature between points A & D .

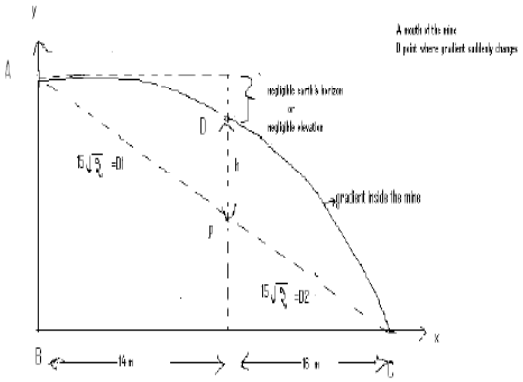


Fig1: Approximate terrain elevation inside the mine

From Fig 1 , $h \approx \sqrt{((15\sqrt{2})^2-14^2)} = 15.94$ m

Fresnel's diffraction factor is $v = -h\sqrt{[2(D1+D2)/(\lambda\, D1\, D2)]} = -12.12$

for v<- 2.4 diffraction loss or gain is given by ,

Gd(dB)=20log(-0.225/v)=-34.63dB (-ve sign indicates loss) , **(Ref (4)**)

Diffraction loss at a depth of 30 m =-34.63dB

Ground reflection attenuation : Rocky soil is common inside mine. They have

Єr = 10 ,σ = 20 mho/cm in GSM 900 band Reflection co-efficient at air earth interface

1. For vertically polarized wave is given by

$$K_V = \frac{(Єr\text{-}jx)\sin\theta - (\sqrt{(Єr\text{-}jx)\text{-}(\cos\theta)^2})}{(Єr\text{-}jx)\sin\theta + (\sqrt{(Єr\text{-}jx)\text{-}(\cos\theta)^2})} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(9)$$

Where $x=\sigma/\omega Є=18x10^3\sigma/f$ in MHz = 3.913m, θ is angle of incidence & σ = conductivity of earth

For θ =90 , Kv = 0.532 & for θ=0, Kv = 0.9106 .

2. For horizontally polarized wave the reflection coefficient is

$$K_H = \frac{\sin\theta - (\sqrt{(Єr\text{-}(\cos\theta)^2})}{\sin\theta + (\sqrt{(Єr\text{-}(\cos\theta)^2})} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(10)$$

For θ=90, |Kh|=0.52 & for θ=0, |Kh|=1.

From the above calculation it is clear that for small θ (incidence angle) , K ≈

1 and hence all the signal will be reflected without ground attenuation. Two types of trolleys are used inside the mine. One type is used for transporting the ore and other for transporting human laborers , mine

equipments etc. All measurements are made on the trolley by using a hand held transceiver kit .

II CALCULATION

Thermal Noise Power : (No)

No= kTB, where k= Boltzmann's constant = 1.381 x 10 exp-23 J/ deg K With B ≈ BN (Noise Bandwidth of Rx due to high selectivity) = 280 KHz i.e. Bandwidth of the Rx can be considered as its noise bandwidth as receiver has very good selectivity.

1. At the entrance or mouth of the mine (point A) (temp= 27deg C or 300 deg K, humidity 82 %), thermal noise power = No=1.16

x10exp-15 Watt and corresponding noise density (No/B)=4.143

x10exp-23 watt / Hz

2. At a depth of 30 m i.e at point C (temp = 38 deg C or 311deg K , humidity 94 %) No=1.2x 10exp-15 w and noise density No/B=4.286 x10exp-23 watts/Hz

3. At an intermediate point D (temp =31 deg C or 304 deg K and humidity is around 85 %) No = 1.175 x 10exp-15 w or -119.3 dbm.

SNR Calculation

Normal distance between the successive points for the reception of radio wave signal in rough mining is 20 to 35 m . In the present gold mine this distance is 12 to 14 m and beyond 30 m there was 75% loss in signal reception . Hence the SNR was calculated at 3 points A (at entrance) , D (intermediate point) and C (at a depth of 30m) along the mine terrain. These are critical points along the terrain where appreciable change in signal strength is noticed **(Ref 5).**

The SNR is same for both the modulation at the entrance (i.e point A)

BPSK

SNR at point A (mouth of the mine):

Signal strength at the entrance is 30 W or 44.77 dBm (Measured value 30W) . Due to 35 cm thick concrete wall at the entrance there is penetration loss.

This loss is already computed as 16 Db So Effective signal strength just inside the mine = 44.77 dBm -16 dB loss=28.77 dBm or 1.327 μW after penetration .

In the absence of any interference or noise this is the signal strength. But due to electrical interference noise like machinery interference noise loss, Power txn line loss actual signal strength decreases further.

For a machinery of medium area < 15 sqm area loss will be around 4 dB in GSM 900 band. For two machinery (one is used as back up or redundancy the total machinery interference noise loss is = 8dB.

Power line loss is predominant whereas the other electrical installation losses due to fan / lighting / blowers / exhaust can be neglected.

Power line noise depends on KVA of the cable, distance at which it is located from measuring equipment , height of the receiving antenna, frequency at which measurement is done, gain of receiving antenna through which it is measured. First electric field strength μV/m/MHz or dBμV /m/MHz is converted to dBm noise power and then can be converted to dB loss. Half wave dipole is used in the measurement.

RF electric field strength is 12.2 dB μV/m/MHz at point A . For a measuring instrument having Rt = 50 ohm (terminating resistance) , 0dB μV corresponds to

-107 dBm RF Power line loss =-107 dBm + 12.2 = -94.8 dB

Total electrical loss at entrance of mine = Machinery interference loss + power line loss = 8dB + 94.8 dB = 102.8 dB.

Effective signal power under the influence of noise is = signal strength – Total electrical loss= 28.77 dBm - 102.8 dB = -74.03 dBm = 3.72 x 10exp-11

W (Measurd value ≈ 32 PicoW)

SNR at point A =Effective signal power / Thermal noise power = 32028.899

Or (SNR)in dB = 45.06 dB

SNR at point D : (Intermediate Point inside the mine)

This is located at around 14m from entrance after which gradient of elevation changes abruptly .

The various losses to be considered are

1) In mines having a mouth with straight path, attenuation of signal is small. Hence up to point D signal strength variation is very small except multipath / or scattering loss and this loss is ≈ -40 dB
2) Power line interference loss : For a field strength of 27.2 dBμv/m/MHz at 940 MHz, the loss is -79.8 dB (-107 dBm + 27.2 = -79.8 dB).

   Other electrical installation losses are negligible.
3) Earth reflection loss is neglected assuming θ (angle of incidence) is small

The total loss at point D= path loss + power line interference loss = 40 dB + 79.8 dB =119.8 dB

Signal strength at point D= 28.77 dBm – 119.8 dB = -91.03 dBm= 7.94 x 10 exp-13 watts. (Measurd value ≈ 0.54 PicoW) (SNR)D = 28.3 dB

i.e. there is around 16.76 dB degradation in SNR at point D as compared to point A.

SNR at point C :- (i.e. at a depth of 30m inside the mine)

The various losses to be considered are

1. Multiple reflection scattering loss =-40dB
2. RF power line interference loss : The measured electric field strength is =49.1dB μV/m/MHz. Hence the loss is -57.9 dB

3. Diffraction loss due to bending of electromagnetic waves around the corner . This is already calculated as -34.63dB

The total loss = scattering loss +RF power line loss + bending loss

=40+57.9+34.63=132.53dB

Signal strength at point C=28.77dBm-132.53dB=-103.76 dBm or 0.0421 pW

(Measurd value ≈ 0.028 PicoW)

(SNR)C = 15.45 dB

Around 12.85 dB less than at point D

At 31 st mt depth signal strength is 0.0105pW or -109.8 dBm (i.e. more than 75 % received power is attenuated as compared to 0.0421 pW)

We know that BPSK Rx sensitivity is - 106 dBm or 2.514 x 10 exp-

14W, this power cannot be detected.

BFSK SNR Calculation :

SNR at point D : (At an intermediate point)

Signal strength at point D = 28.77 dBm – losses at point D . The various losses to considered are i)multiple reflection scattering loss = -40dB ii) power line interference loss : field strength of 29.2 μV / m /

MHz corresponds to -77.8 dB loss.

Total loss = 40 dB + 77.8 dB = 117.8 dB

The signal strength at point D = 28.77dBm – 117.8 dB = -89.03 dBm or 1.25

pW (Measurd value ≈ 0.995PicoW)

SNR at point D = 30.27 dB

i.e around 2dB better SNR as compared to BPSK

SNR at Point C :(i.e at 30 m depth)

Signal strength at C = 28.77 dBm – multi path scattering loss of 40 dB – power line interference loss 51.1 dB μV / m / MHz or 55.9 dB – bending loss

34.63 dB = -101.7 dBm or 6.67 x 10 exp-14 W(Measurd value ≈ 0.0486 PicoW)

SNR at point C = 17.45 dB

At 31st mt depth signal strength is 1.67x 10exp-14 W 0r –107.8dBm (i.e 75% of signal get attenuated). As BFSK Rx sensitivity is only – 107 dBm 0r 1.99x 10exp -14 W , this power cant be detected .

Data Transmission:

Regarding data transmission we have calculated Eb/No& Pe for the different modulation techniques , where No is thermal noise density & Pe (probability error) at a depth of 30m only .Because this is the critical point of the link ,even though it can be calculated at other points.

Raw data rate = Rb = 270 Kbps

Therefore Tb = 1/ Rb = 3.704 μsec

BPSK: Signal strength at a depth of 30 m is Prmin=0.041 pW

So Vmin =√(Prmin/2)=0.145μV

Eb=(Vmin)² Tb/4=1.95x10exp-20 J

Eb/No = 4.55 or 6.58 dB .

Pe = Q (√ 2Eb/No) = Q (3.02)

BFSK : Signal strength at a depth of 30 m is Prmin= 0.0667 pW

So Vmin = 0.183μV , Eb = 3.09x10exp-20 J

Eb/No = 7.21 or 8.58 dB .

Pe = Q (√Eb/No) = Q (2.7)

IIIConclusions :

It was noticed that SNR approximately varies from 45 dB to 16 dB based on signal strength, interference noise and temperature. This variation is also true for Gaussian Channels **(Ref2)** .

1. Even though Eb/No of BFSK is superior as compared to BPSK its error performance is poor .
2. Main limitation is high power requirement (30 W) .
3. Above model may not work in other mines .
4. With respect to SNR performance BFSK is better as compared to BPSK .
5. By improving the sensitivity of receiver depth can be increased. Also by placing reflectors at turning points depth can be increased .
6. Critical loss is penetration loss at entrance which must be minimized to improve depth .
7. It is noticed that as we go deeper power line noise strength increases
8. From above calculation it is clear that there is around 2dB better signal reception is possible in BFSK as compared to BPSK .
9. The difference between measured value and calculated value of signal strength may be due to fading , ground reflection and other losses . Here we have neglected them & hence this is only an approximate modeling .

REFERENCES :

[1]. William H. Schiff Bauer , Gary . L . Mowrey & Jurgen . F . Brune TTE & other types of communication system for under ground coal mines for disaster and normal operation .

[2]. Helchel ,IEEE transaction on antenna & propagation , vol 54 , No 12 , Dec 2006 . ‘ comparison of 900 MHz band and 1800 MHz band indoor propagation ‘ .

[3]. Shi Lihua , Xu Qiwei , Chan Bin Gao Cheng ,Measurement of the frequency dependent dielectric constant of concrete materials by TDR & wavelet modeling method at Asia pacific conference on environmental electromagnetic . CEEM 2003 , Nov 4 – 7 .

[4]. Lee , Mobile communications engineering Theory and applications 2nd ed

[5]. Gilles , Y . Delisle ,Underground mine wireless propagation modeling . School of information technology & engg , University of Ottawa , Canada .

[6]. T.S.Rappaport ,Wireless communications principles & practice 2nd ed by D.P.Agarwal , Qing An Zeng ,Wireless & mobile systems , Thmson .

[8] C. Nerguizian, C. L. Despins, S.Aff`es, and M. Djadel, “Radiochannel,characterization of an underground mine at 2.4GHz,”,*IEEE Transactions on Wireless Communications*, vol. 4, no. 5,pp. 2441–2453, 2005.

[9] Zhi, S. and I. F. Akyildiz, \Channel modeling of wireless net-works in tunnels," *IEEE Global Telecommunications Conference, (GLOBECOM'08),* New Orleans, LA, Nov. 30{Dec. 4, 2008).

[10] *Empirical formula for propagation loss in land mobile radio service,* IEEE transactions on Vehicular Technology, VOL. VT-29, NO.3, AUGUST 1980

[11] About Radio Propagation Models. http://en.wikipedia. org/wiki/Radio_ propagation model

[12] About Path Loss.

http://en.wikipedia.org/wiki/Path_loss

[13] Bertoni, H. L., *Radio Propagation for Modern Wireless Systems*, Prentice Hall, Englewood Cli®s, NJ, 2000.

[14] M. Boutin, A.Benzakour, C.Despins and S. Affes. (2008). Radio wave characterization and modeling in underground mine tunnels, *IEEE . Transaction. on Antennas and Propagation*, 2:540-549

[15] M.Lienard and P.Degauque. (2000). Natural Wave Propagation in Mine Environments, *IEEE*

Transaction on Antennas and propagation, 48(9): 1326-1339.

[16] W.E.Pakala and V.L.Chartier *Radio noise measurements on overhead power lines from 2.4to 800kv,*

[17] M.Ndoh and G.Y.Delisle. (2004). Underground mines wireless propagation modeling, *60th IEEE Vehicular Technology Conference*, Los Angeles, CA, 3584-3588.

[18] Kjeldsen, E. and M. Hopkins, \An experimental look at RF propagation in narrow tunnels," *Proc. IEEE Military CommunicationsConf. (MILCOM'06)*, Washington D.C., Oct. 23{25, 2006

[19] Li Wenfeng, Lv Yingli, Li Baiping. (2008). Experiment on characteristic of radio propagation in mine. *Journal of Xi an University of Science and Technology,* 2:327-330 (in Chinese)

◈◈◈

Virtual Platform as a Service

Mallela GuruLakshmi

VIGNAN University, Guntur, India

Abstract – Developing web applications for business with Platform-as-a-Service (PaaS) solutions is like fitting users with a tailored suit for a special occasion. As cloud computing terms begin to finally shake, loose their initial buzz factor, more and more concrete terms are emerging. PaaS, or platform as a service, currently owns a wider range in definition. PaaS can include basic application development elements that essentially equate to "hosted runtime libraries" where developers can code and run applications in the cloud. PaaS offerings can also be as comprehensive as needed to include complete architectures and frameworks that feature support for pre-built UIs, reporting, data processing, extensibility, user management, and even creating a distribution-centered marketplace to offer applications for sale. Because the platform itself is managed through the web interface, there is a large potential for end users to make "self-service" configurations that can best adapt to their work.

Keywords – *PaaS, hosted run time libraries, self service.*

I. INTRODUCTION

Cloud computing, to put it simply, means internet computing. The internet is commonly visualized as clouds; hence the term "cloud computing" for computation done through the internet. With cloud computing users can access database resources via the internet from anywhere, for as long as they need, without worrying about any maintenance or management of actual resources. Besides, databases in cloud are very dynamic and scalable. The cloud computing search request can be analysed in the fig 4.

Cloud Computing is unlike grid computing, utility computing, or autonomic computing. In fact, it is a very independent platform in terms of computing. The best example of cloud computing is Google apps where any application can be accessed using a browser and it can be deployed on thousands of computer through the internet. It also provides facilities for users to develop, deploy and manage their applications on the cloud, which entails virtualization of resources that maintains and manages itself.

A. Service models

These services are broadly divided into three categories:

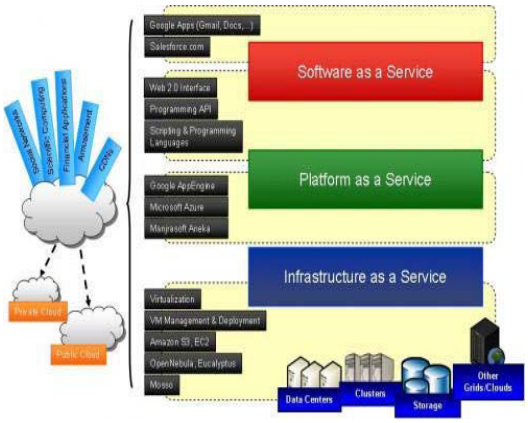


Fig : Cloud computing architecture

1. Saas (software as a service) - Saas is a model of software deployment where an application is hosted as a service provided to customers across the internet.

2. Paas (platform as a service) - The cloud provides hardware resources, typically virtual machines, which can be loaded with the users, operating system and software.

3. Iaas (infrastructure as a service) the cloud provides an infrastructure including platforms, networking, *etc.* on which applications can be placed.

We can see these sevice categories in the architecture of cloud computing clearly in fig1.

II. PROBLEM STATEMENT:

It is difficult to Develop, Maintain and Deploy the application Lot of amount required to do those Modules Sustaining efforts and Tech Support also required Cost is very High.

Traditional business applications have always been very complicated and expensive. The amount and variety of hardware and software required to run them are daunting. You need a whole team of experts to install, configure, test, run, secure, and update them.

When you multiply this effort across dozens or hundreds of apps, it's easy to see why the biggest companies with the best IT departments aren't getting the apps they need. Small and mid-sized businesses don't stand a chance.

You should consider the following issues when planning to deploy a application:

- Is this a stand-alone application that will be installed on individual machines?
- Will this application be installed on a network, with many users accessing it simultaneously?
- Which system options should be set to establish the desired application environment?
- Which windowing features should be in effect?
- When the user closes the application, where should your application return to (the display manager, another application, or the operating environment)?

The basic methodology for deploying an application should include the following steps:

1. Modify the prototype so that only the desired software windows are open when the application starts.
2. Modify the prototype so that the user is returned to the appropriate environment when the application closes.
3. Copy the catalog or catalogs that comprise your application from the testing or prototype location to the production location.
4. Prepare a custom configuration file to initialize for use by your application.
5. Create a command file or icon that launches the application.

III. PROPOSED SYSTEM:

PaaS solutions, like Virtual PAAS, will even allow for private cloud PaaS, where you can take the entire platform and install it behind the firewall or in a private hosted instance and essentially own the whole architecture. This is an important aspect as some industries require greater control beyond the openness of the cloud.

Flexible development environment including ability to build and check-in code from developer instances Ability to run on a public cloud Rapid prototyping and end-user configuraiton of applications with complete, live browser-based IDE, including database modeling, process and workflow definitions, user interface configuration and built-in reporting.

Cloud platform services or "Platform as a Service (PaaS)" deliver a computing platform and/or solution stack as a service, often consuming cloud infrastructure and sustaining cloud applications. It facilitates deployment of applications without the cost and complexity of buying and managing the underlying hardware and software layers.

Thus Platform-as-a-service in the cloud is defined as a set of software and product development tools hosted on the provider's infrastructure. Developers create applications on the provider's platform over the Internet. PaaS providers may use APIs (application program interfaces), website portals or gateway software installed on the customer's computer. Passware, (an outgrowth of Sales passware) and GoogleApps are examples of PaaS. Developers need to know that currently, there are not standards for interoperability or data portability in the cloud. Some providers will not allow software created by their customers to be moved off the provider's platform.

Host Your Applications on Pass and Reduce Your IT Expenses

As opposed to the business mantra of doing more with less, working on the Paas Cloud allows you to do more with more. By leveraging the benefits of a robust and secure computing environment, you're opening the door to a host of industry standard solutions enabling you to do more--with more agility, scalability. With the Paas Cloud, you can focus on solving pressing business issues, not tangling with technology by moving our application to the cloud as shown in fig 2.

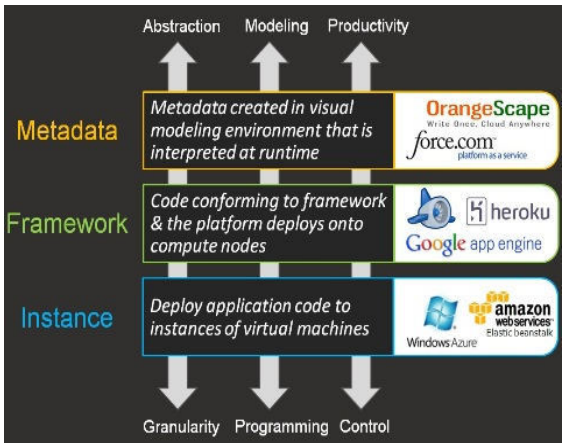


Fig : Moving applications to cloud

Multi-tenant Architecture

Forget repeating application and platform code across multiple servers. True multi-tenancy at the tenant-ensures independent UI, data modeling, processing, presentation, administration and developer configurations over a common, manageable code base. Because Paasware's multi-tenancy is also multi-level, its metadata-driven configuration enables solution providers to offer branded SaaS solutions to their clients with complete control as MSPs with separate billing and configurations all within Paasware's Cloud.

Industry-Standard Technology Stack

Built on the MySQL Enterprise Edition relational database and the Apache Tomcat Application Server, Paasware's components are vetted and highly serviceable providing greater manageability. Because the platform was built using familiar and versatile code such as Java, JSP, HTML and XML, extending the platform is easier and faster than with proprietary technologies.

Helpful Support Team

Get the answers you need fast. Paasware's dedicated support team is a natural extension of your IT department. Once a support case is filed, you can expect:

- Platform issues responded to within one hour
- Resolutions or escalations within one business day

Real-Time Data Mirroring & Disaster Recovery

Your data could not be more secure. Every transaction on Paasware is recorded in real-time on production- and mirrored-database servers to ensure instantaneous and seamless failover protection. Additionally daily server backups are made on tape and backup servers in the event of data loss due to user error, equipment failure or natural disaster.

Scalable Capacity

Paasware's automated provisioning systems can scale operations quickly and easily--rapidly deploying new servers within hours to meet demand and utilize today's state-of-the-art load balancers. We can also deploy into cloud hosting environments (Amazon EC2) for virtually unlimited scalability.

24x7x365 Response Center

Always on guard, Paasware servers and network infrastructure are monitored 24x7x365, including a series of advanced alert systems to notify hosting team members of any infrastructure issues. Most issues are resolved so quickly, they go unnoticed by users or our reseller partners. Dozens of individual system statistics are monitored to ensure peak performance such as CPU utilization, network throughput, hard disk array response times, and total system availability.

Scalable Capacity

Paasware's automated provisioning systems can scale operations quickly and easily--rapidly deploying new servers within hours to meet demand and utilize today's state-of-the-art load balancers. We can also deploy into cloud hosting environments (Amazon EC2) for virtually unlimited scalability.

Stress-Free Upgrades

One of the key benefits of Paasware's Platform-as-a-Service offering is its lack of IT involvement. Zero, to be exact, with software, infrastructure and architectural maintenance issues. Free your staff from testing, managing, and implementing software hot-fixes, updates and upgrades.

Custom Application Development: Build and Run Your Entire Business in the Cloud

The Passware platform is the fastest path to complex enterprise application development. What used to take months can now be done in days or weeks. Companies and ISVs are using Passware to build applications across the entire spectrum of business functions and industries: supply chain management,

billing, audit, tax calculation, event management, compliance tracking, brand management, pricing, accounts receivable, accounts payable, billing, HR, payment processing, employee on-boarding, claims processing, and much more.

There are more than 100,000 custom applications running on Passware, and the platform supports more than 150 million transactions per day.

Enterprise Application Development Made Easy

Traditional custom application development platforms supply only the core services of database access and containers for logic and presentation. You need other software to create your entire application as well as additional components to run the data center that supports the platform. Of course, more components mean additional expenditure for acquisition, maintenance, and integration. With Passware, you get a full enterprise application development stack—including a complete user interface to your data, comprehensive reporting and analytic capabilities, a flexible security and sharing model, and workflow and approvals. All this functionality is available and integrated from the outset. You can even leverage extended features such as built-in internationalization, full support for mobile devices, and integration with existing systems.

You access Passware through a Web browser: development and deployment both take place in the cloud. The platform itself provides everything you need for robust enterprise application development through a combination of clicks, components, and code. This unique blend makes developers much more productive: Application capabilities come from simple declarations of attributes through pre-built capabilities and extend to a fully flexible application development environment.

With clicks alone, Passware lets you include rich functionality in all your applications. You can define your data structures and relationships, expose user interfaces to interact with this data, extend internal data with logical representations, and define workflow actions and approval processes. It's simply a matter of shaping the built-in capabilities of the platform to your needs by setting the values of attributes.

Another application development advantage: The entire Passware reporting and analytics system is automatically built into all your applications. This integrated functionality lets you mine your data store for additional value.

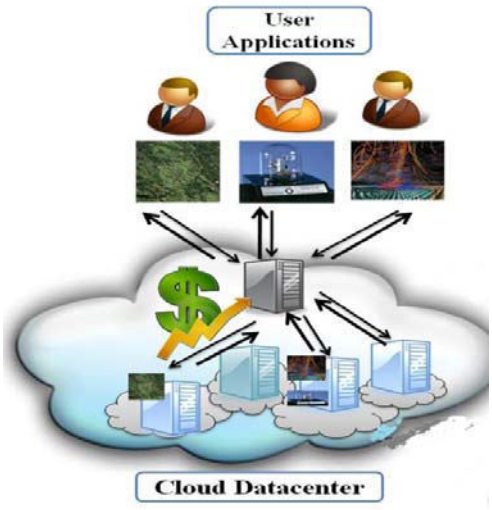


Fig : Applications in cloud

Literature Survey:

The old way, application-building that's slow and inefficient.

Over the years, the traditional way of creating and running business applications has become overly complex and cumbersome. There are too many moving parts to buy, install, configure, and maintain including hardware and software. Plus, the entire infrastructure requires constant maintenance to keep it working smoothly.

This overhead burden creates barriers to productivity in custom application development. Computing environment complexity means that every little change can trigger repercussions throughout the organization. That setup significantly reduces overall IT responsiveness, impairing the ability of a company to address constantly changing business needs. Instead, enterprise application development proceeds at a glacial pace, with long backlogs a fact of life. The end result is that business managers don't get the applications they need to run their business. Instead, they end up with a welter of unintegrated, homegrown systems on spreadsheets, personal databases, or other unsupported platforms.

IV. CONCLUSION

We first analyzed the benefits offered by Cloud computing by studying its fundamental definitions and benefits, the services it offers to end users, and its

deployment model the applications in cloud can shown as in fig 3. PaaS providers offer in general the platform services for application development. The platform facilitates the development of applications which ensures system wide energy efficiency.

It is a software energy efficiency benchmark that measures the energy required to perform an external sort. In addition, platforms itself can be designed to have various code level optimizations which can cooperate with underlying complier in energy efficient execution of applications. Other than application development, Cloud platforms also allow the deployment of user applications on Hybrid Cloud. In this case, to achieve maximum energy efficiency, the platforms profile the application and decide which portion of application or data should be processed in house and in Cloud.

ACKNOWLEDGMENT

We are very thankful to the authors who have helps us to improve the quality and performance of our paper. We are grateful to our guide who guided us to our research work.

REFERENCES

[1] Gleeson, E. 2009. Computing industry set for a shocking change. Retrieved January 10, 2010 from http://www.moneyweek.com/investment-advice/computing-industry-set-for-a-shocking-change-43226.aspx

[2] Buyya, R., Yeo, C.S. and Venugopal, S. 2008. Market-oriented Cloud computing: Vision, hype, and reality for delivering it services as computing utilities. Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications, Los Alamitos, CA, USA.

[3] Vecchiola, C., Chu, X. and Buyya, R. 2009. Aneka: A Software Platform for .NET-based Cloud Computing. In High Performance & Large Scale computing, Advances in Parallel Computing, ed. W. Gentzsch, L. Grandinetti and G. Joubert, IOS Press.

[4] Allenotor, D., Thulasiram, R. K. 2008. Grid resources pricing: A novel financial option based quality of service-profit quasi-static equilibrium.

◈◈◈

Hierarchical Clustering Approach with Hybrid Genetic Algorithm for Combinatorial Optimization Problems

M. H. Mehta & V. V. Kapadia

Computer Engineering Department, Birla Vishvakarma Mahavidyalaya Engineering College,
V. V. Nagar – 388120, India

***Abstract* –** Engineering field has inherently many combinatorial optimization problems which are hard to solve in some definite interval of time especially when input size is big. Although traditional algorithms yield most optimal answers, they need large amount of time to solve the problems. A new branch of algorithms known as evolutionary algorithms solve these problems in less time. Such algorithms have landed themselves for solving combinatorial optimization problems independently, but alone they have not proved efficient. However, these algorithms can be joined with each other and new hybrid algorithms can be designed and further analyzed. In this paper, hierarchical clustering technique is merged with IAMB-GA with Catfish-PSO algorithm, which is a hybrid genetic algorithm. Clustering is done for reducing problem into sub problems and effectively solving it. Results taken with different cluster sizes and compared with hybrid algorithm clearly show that hierarchical clustering with hybrid GA is more effective in obtaining optimal answers than hybrid GA alone.

***Keywords* –** *Genetic algorithm, Particle Swarm Optimization, Crossover, Mutation, Adaptive GA, Clustering.*

I. INTRODUCTION

A big set of engineering problems is represented by combinatorial optimization problems where problems become hardest to solve when input size becomes larger. Long-established methods [1] available for solving them take long time for obtaining results. Nevertheless, such traditional methods are guaranteed to find the optimal solution. Researchers have found out an entire new division of algorithms, which can help in solving combinatorial optimization problems known as evolutionary algorithms. Ant-Colony optimization [2], Artificial-Bee Colony optimization [3], Swarm Intelligence [4] and Genetic Algorithms (GA) [5] are evolutionary algorithms designed for optimization purpose. Evolutionary algorithms are designed by taking ideas from the nature and behaviors of birds flocking, fish schooling and bees' working method. Genetic algorithms are designed by taking inspiration from the human development process. It follows Charles Darwin's "Survival of the fittest" kind of strategy to obtain the optimal solution. Many evolutionary algorithms have proved adequate enough in getting optimal answers in some cases. But, to get optimal answers in most of the cases is not possible by any evolutionary algorithm alone but answers close to optimal can be obtained with less time in comparison of the conventional methods.

Evolutionary algorithms can be considered as a branch of approximation algorithms where much emphasis is given on heuristics and not on systematic methods. Evolutionary algorithms alone have not obtained the best results for combinatorial optimization problems. They have their own disadvantages in different terms like time and cost. However, merging two established evolutionary algorithms can yield the pioneering algorithms. Thus, these novel algorithms known as Hybrid algorithms can have the advantages of original algorithms and can eliminate deficiencies as well. Hybrid algorithms give better results than original evolutionary algorithms from which they are designed. Although, hybrid algorithms can give better results, combining them with clustering technique can further enhance these results. Clustering means to divide the whole problem into small sub problems known as clusters and then obtaining the results. Clustering is an effective technique, which can be merged with hybrid algorithms since it effectively reduces the problem size and then yields results.

In this paper, we have proposed a hierarchical clustering approach combined with hybrid genetic algorithm [6]. Hybrid genetic algorithm is a combination of genetic algorithm and catfish-PSO algorithm [7]. Genetic algorithm, which is considered in hybrid algorithm, is an adaptive version of simple original genetic algorithm. It is adaptive as it can switch between mutation operation and catfish PSO operation. The proposed hierarchical clustering algorithm is compared with Intelligent Adaptive Mutation Based – Genetic Algorithm with Catfish–PSO effect algorithm

(IAMB-GA with CF-PSO) [6]. IAMB-GA with CF-PSO is explained in detail in section III. We have taken Symmetric Traveling Salesman Problem [8] as our combinatorial optimization problem. In symmetric TSP, the cost joining one city to other remains same in forward and backward direction. TSP is a classic example of NP-Hard Combinatorial optimization problem [9]. TSP is a problem in which a salesman has a list of number of cities to visit. Every city considered in the problem is connected to every other city. Salesman's task is to visit every city exactly once starting from a city and finishing the journey returning back to start city. Every road/edge that connects cities is having some cost associated to it. The challenge for the salesman is that he has to finish the round trip such that minimum cost roundtrip is obtained by his journey. Obtaining the optimal cost trip becomes tough in TSP as number of cities increase, making it a suitable choice as a combinatorial optimization problem. The true test of TSP lies in finding most favorable solution as n (= number of cities) increases, possible tours to explore becomes (n-1)! /2. So for just 30 cities, we have 4420880996869850977271808000000 trips to study. If only one start city is considered and duplicate trips are removed in symmetric case then possible trips to explore is (n-1)! / 2. Our results show that proposed hierarchical clustering with hybrid GA algorithm's performance is definitely improved than IAMB-GA with CF-PSO effect for four classical TSPLIB [10] problems.

Remainder of this paper is organized as follows: Section II describes the GA and PSO. Section III illustrates the concept of IAMB-GA with Catfish PSO algorithm. Section IV explains hierarchical clustering technique. Section V shows complete proposed algorithm structure. Experiments and Results are discussed in Section VI. Section VII concludes the paper.

II. GENETIC ALGORITHM AND PARTICLE SWARM OPTIMISATION

In the year of 1975, J.Holland first proposed genetic algorithm. Genetic algorithm is an iterative search and optimization method, which takes inspiration for its working from the human evolution process. Any genetic algorithm has basically four steps to perform. The first step is initial population generation by choosing proper encoding method. After creating the initial population, selection operator chooses the optimal value, based upon the fitness function. In TSP, fitness function is the round trip that is traveled by the salesman and it should be minimum. Selection operator's job is to select the most optimal population(s) that can go ahead in the evolution and generate best off springs for the future generations. Many different types of selection operators are available in the literature [11]. After selecting the best population, crossover operator operates on the chromosomes that is selected population, to generate new off springs. Crossover operators actually make permutations on the chromosomes that result in new chromosomes. Different crossover operators' efficiencies and effects are different [1]. Mutation operator is included in genetic algorithm to provide population diversity. Mutation operator necessarily gets whole process out of any local minima if realized by the process. Mutation operator is necessary for genetic algorithm so that whole process of getting optimal value does not stick into some valley. Generally, crossover is given more emphasis in traditional genetic algorithm than mutation meaning that mutation is done in fewer amounts on chromosome than crossover operation. Every time, the best fitness value is considered as global minimum and the process terminates when it reaches to its stopping criteria.

Swarm Intelligence is the whole new branch of algorithms, which take motivation for their operation from the nature. Particle swarm optimization [4] is one such algorithm developed by mimicking the social behavior of birds and fishes. Eberhart and Kennedy first proposed particle swarm optimization in 1995. Standard PSO 's framework consists of individuals who fly in the search space with some velocity. The velocity of each individual is adjusted according to its own flying experience and its companions' flying experience. This velocity adjustment is dynamic in nature. The i^{th} particle is depicted by $X_i = (x_{i1},x_{i2},....,x_{in})$. The best previous position of the ith particle is calculated and represented by $P_i = (p_{i1},p_{i2},....,p_{in})$. The best particle among all the particles referred as global best is represented by symbol g. Rate of change that is velocity of particle's position is given by $V_i = (v_{i1},v_{i2},....,v_{in})$. The particles are measured based on the following equations:

$$v_{id} = w * v_{id} + c_1 * r1() * (p_{id} - x_{id}) + c_2 * r2() * (p_{gd} - x_{id}) \quad (1)$$

$$x_{id} = x_{id} + v_{id} \quad (2)$$

where c_1 and c_2 are two positive constants and r1 and r2 are two random functions in range 0 to 1 including 0 and 1. w is the weight and first part of equation 1 is dealing with previous velocity of the particle, where as second part is "cognitive" part and third part is "social" part. Two constants c_1 and c_2 are also known as individual factor and societal factor respectively.

Both genetic algorithms and particle swarm optimisation are considered as candidate algorithms for forming a hybrid algorithm which is merged with hierarchical clustering technique.

III. IAMB-GA WITH CATFISH-PSO

Simple genetic algorithms have natural deficiency of lack of intelligence. Simple genetic algorithms work in a usual way where it follows four steps of initial population by encoding, selection, crossover and mutation. However, simple GA does not give good results if more focus is given to crossover operators than mutation operators. Simple GA can be changed to intelligent GA if algorithm can switch between mutation and crossover adaptively. Whenever crossover operators stop giving more minimum values than the previous global value obtained, it is switched to mutation operator in Intelligent Adaptive Mutation Based GA (IAMB-GA)[8]. As mutation operator provides new population to the algorithm, it successfully gets the stuck process out of any local minima. So the main idea of IAMB-GA is intelligence of adaptive ness of the algorithm, whenever crossover operator is not successful in giving minimum tour than the previous one, mutation is done. The whole process of IAMB-GA for TSP starts with initial encoding of population. Here, permutation encoding is used for presenting trips. After generating initial population, selection operator is used which selects the trips from the generated population and forwards it in the process. The selected population then undergoes crossover operation. Order crossover [12] is chosen for generating children known as off springs. Order crossover operator works in a way such that the order of the string is maintained. Order crossover is shown in table I. Mutation takes place after generating children chromosomes. Mutation operator does not work on two parent strings. It works only on a solo string. Mutation alters whole population in such a way that new population is generated to get process out of any valley. Mutation and crossover operators occur adaptively in IAMB-GA process meaning that whenever crossover operator fails to get minimum tour than the previous global minimum, than mutation occurs in the process. Basic mutation Operators used in the IAMB-GA are as follows: 1) Inversion Mutation, 2) Reciprocal/Exchange Mutation, and 3) Slide Mutation [13]. In each operator, two random points are chosen as two and five. Table II, III and IV show all three operators [8]. The basic idea of IAMB-GA is to promote mutation whenever population diversity is needed to get the process out of valley, and again switch to the crossover operation when process starts generating minimum cost tour than the previous global minimum [14].

IAMB-GA uses mutation operators mainly as mutation operators have proved efficient than crossover operators. In addition of basic mutation operators, more mutation operators are included in the algorithm as various mutation operators have capacity of giving vast population diversity and variation in the new generated population forcing solution to go nearer to optimal solution. In addition of inversion, reciprocal/exchange and slide mutation, Yin-yang operators [15] are introduced in IAMB-GA. Yin-yang operators used are – 1) turnover operator and 2) mutual operator. Turnover operator is changed into two-point turnover operator and can be explained as follows: the chromosome is taken and two random points are decided. Then new chromosome derives the middle part between two random points directly. And left and right portions are reversed as shown in Table V [8](random points are 3 and 6).

TABLE I. ORDER CROSS OVER

Parent-1)	12-564-387	--→	Off spring-1)	23-564-781
Parent-2)	14-236-578	--→	Off spring-2)	54-236-871

TABLE II. INVERSION MUTATION

Chromosome	12-564-387	→	New Chromosome	12-465-387

TABLE III. RECIPROCAL/EXCHANGE MUTATION

Chromosome	12-387-564	→	New Chromosome	12-783-564

TABLE IV. SLIDE MUTATION

Chromosome	12-387-564	→	New Chromosome	12-873-564

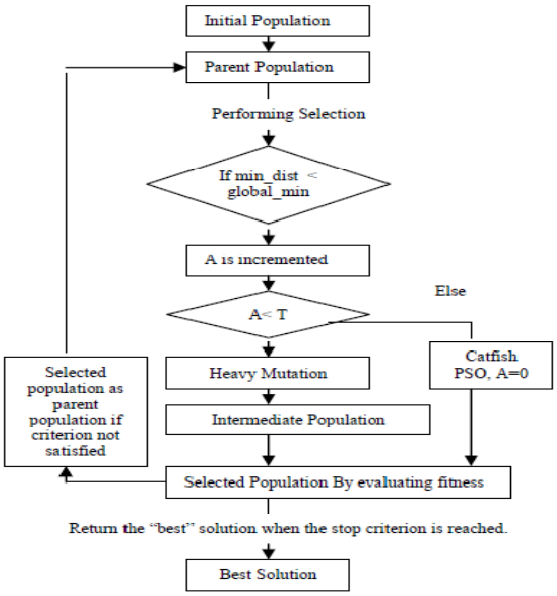


Figure 1. Intelligent Adaptive Mutation Based Genetic Algorithm with catfish PSO effect Structure [6]

Mutual operator works in following way. It maps old tour to new tour by gathering odd indexed integers from old tour and pushing them in the front in new tour or it picks all even indexed integers from old tour and

pushing them in the front direction. Its working is in Table VI [8].

Algorithm starts with initial population generation. Then it generates new population by selection operator, order-crossover operator and all illustrated mutation operators. Adaptive ness added in the IAMB-GA has the power of intelligence within. Every time evaluation is done by comparing minimum distance so far obtained with newly calculated minimum distance. Whichever is minimum is known as global minimum that is minimum distance obtained so far. Algorithm stops when it reaches the stopping condition and returns best solution thus obtained.

TABLE V. TWO-POINT TURN OVER MUTATION OPERATOR

Chromosome	123-456-789	→	New Chromosome	321-456-987

TABLE VI. MUTUAL MUTATION OPERATOR

Chromosome	12345678	→	New Chromosome	13572468

OR

Chromosome	12345678	→	New Chromosome	24681357

Catfish PSO is a novel optimisation algorithm, which is proposed by Chuang, Tsai and Cheng [7]. PSO has many available variants. Catfish PSO is an innovative algorithm, which has taken inspiration from the Norwegian fishermen's observation about catfish and sardines' behaviour. Catfish PSO starts with randomly initialised particle swarm, which can be compared with sardines. Here, every individual's best and whole swarm's best is calculated and then each particle is updated by Equations (1) and (2) of section II. Catfish PSO used however weight updating formula as well [7]. To avoid premature convergence in this process, catfish particles are introduced. These catfish particles replace 10% worst particles and give new fresh population to the whole process. Catfish particles thus necessarily get optimisation process out of any local minima state, which is of prime importance for any optimisation task.

Intelligent adaptive mutation based genetic algorithm is merged with catfish PSO idea. This IAMB-GA with CF-PSO algorithm is not having any crossover operator within it. As crossover operators were not efficient they are omitted. Heavy mutation is used in this algorithm. Algorithm starts with initial population generation and heavy mutation operation is done with every chromosome. In each iteration, the unsuccessful attempt 'A' is measured and recorded. A threshold 'T' is determined. When A crosses the determined threshold 'T', catfish PSO is introduced in the algorithm. Catfish PSO effect is introduced in IAMB-GA to get the stuck process out of local minima. Catfish PSO replace worst 10% trips and introduces new population that drives whole process to go nearer to the optimal solution. Fig.2 shows IAMB-GA with Catfish-PSO algorithm.

IV. HIERARCHICAL CLUSTERING

In hierarchical clustering technique [16], input size of TSP City problem is divided in clusters. If n cities are there, then m clusters can be created where n/m chromosomes would be involved in every cluster. Hierarchical clustering technique works by first of all creating initial population for TSP problem. Then after creating m clusters for n city problem in which n/m chromosomes are contained in every cluster as said above. After this initial cluster creation step, hybrid genetic algorithm can follow procedure. A different operation is needed to work upon these created clusters from which new optimal trip can be generated and can be forwarded.

Hierarchical Clustering with Hybrid GA (IAMB-GA with CF-PSO)

1. Generation of Initial Population
2. Generation of m Clusters with n/m chromosomes each for n city problems.
3. Tournament selection filtering to reduce the initial population

   size.
4. Minimum cost trip/chromosome is found out from every cluster.
5. Best minimum cost trips are concatenated together to produce final optimal trip.
6. Iteration starts.
7. IAMB-GA with CF-PSO algorithm yields optimal cost.

Figure 2. Proposed Algorithm

V. PROPOSED ALGORITHM

In this paper, hierarchical clustering technique is merged with intelligent adaptive mutation based genetic algorithm with catfish PSO. The proposed algorithm is not having any crossover operator within it. Algorithm starts with initial population generation, which then is followed by clusters creation. After clusters are created, these clusters are passed through tournament selection

procedure, which further reduces size of initial population. Every cluster is then passed through fitness function evaluation for every trip within it. Then after for every cluster, the minimum cost trip is obtained. Every cluster thus yields a best trip, which is having minimum cost. These all-minimum cost trips are connected to form a final optimal trip, which is forwarded to IAMB-GA with CF-PSO algorithm. Proposed algorithm is shown in Fig.2.

VI. EXPERIMENTS AND RESULTS

Implementation is done in Mat lab 7.9 on Intel Core 2 Duo processor with 3 GB Memory. In both algorithms initial population is taken as 400. After filtering by tournament selection method, finally initial population contains 100 parents to maintain similarity in both algorithms. Iterations were kept to 10000.

Four TSPLIB problems were considered for comparison. As in TSPLIB [10], the optimal solutions are given different algorithms' results can be compared with them. For every problem, five runs are taken and the best cost is finally considered.

Table VII shows results of IAMB-GA with CF-PSO for four TSPLIB cities. Table VIII to table XIII show results of proposed algorithm that is hierarchical clustering with hybrid GA. Proposed algorithm performs better than IAMB-GA with catfish effect algorithm for all cities. Different cluster sizes are taken also to see that which cluster size is optimal for which TSPLIB city. It can be observed that for city Eil51, cluster size 2 is sufficient. For Ch130, cluster size 10 is best. Where as Cluster size 15 is proven effective for Tsp225 and Rat575. In future, more runs can be taken instead of five runs for better analysis.

TABLE VII. RESULTS FOR IAMB-GA WITH CATFISH PSO EFFECT

TSP Problem	Optimal Route	Run1	Run2	Run3	Run4	Run5	Best Cost
Eil51	426	441	452	446	459	463	441
Ch130	6110	6469	6681	6617	6403	6589	6403
Tsp225	3916	4106	4046	4202	4129	4166	4046
Rat575	6773	7885	8020	7789	8046	7923	7789

TABLE XIII. RESULTS OF COMPARISON OF HIERARCHICAL CLUSTERING

TSP Problem	Optimal Route	Cluster size=2	Cluster size=5	Cluster size=10	Cluster size=15	Cluster size=25
Eil51	426	**433**	436	439	440	441
Ch130	6110	6310	6262	**6251**	6314	6403
Tsp225	3916	4105	4099	4095	**3906**	4046
Rat575	6773	7888	7812	7845	**7758**	7789

VII. CONCLUSION

In this paper, we have proposed a hierarchical clustering algorithm merged with IAMB-GA with CF-PSO effect. Different cluster sizes are taken and proposed algorithm's performance is measured with IAMB-GA with CF-PSO algorithm alone. Results definitely prove that clustering gives more enhanced results than only hybrid GA. In every TSP City problem, more optimal answers are obtained. For Tsp225, hierarchical clustering technique gives 3906 answer which is the most optimal answer even less than 3916 as given on TSPLIB site. For Rat575 also, best-cost answer obtained is 7758 which is far better than hybrid GA 's answer 12505. Results show that our proposed algorithm is efficient and thus a better choice for solving combinatorial optimization problems.

ACKNOWLEDGMENT

We thank to almighty and our family members for giving us inspiration and support. We are heartily grateful to all the persons who have helped directly or indirectly in any way to complete this work.

REFERENCES

[1] ABDOUN Otman and ABOUCHABAKA Jaafar "A comparative study of adaptive crossover operators for genetic algorithms to resolve the travelling salesman problem" IJCA (0975-8887) October 2011

[2] Carlos M. Fernandes, Agostinho C. Rosa and Vitorino Ramos "Binary Ant Algorithm" GECCO'07 July 7-11, 2007, London, ACM.

[3] Anan Banharnsakun, Tiranee Achalakul, Booncharoen Sirinaovakul "ABC-GSX: A hybrid method for solving the Traveling Salesman Problem" IEEE 2010

[4] Mei-Ping song, Guo-Chang gu "Research on particle swarm optimisation: a review" IEEE 2004

[5] David E. Goldberg "Genetic Algorithms in search, optimisation & machine learning" Pearson Education LPE

[6] M.H.Mehta and V.V.Kapadia "Intelligent Adaptive Mutation Based Genetic Algorithm With Catfish-PSO effect For Combinatorial

Optimization Problems" IEEE, RACCCT Surat, March 29-30, 2012.

[7] Li-Yeh Chuang, Sheng-Wei Tsai, and Cheng-Hong Yang "Catfish Particle Swarm Optimization" IEEE Swarm Intelligence Symposium, St. Louis MO USA, September 21-23, 2008.

[8] M.H.Mehta and V.V.Kapadia "Intelligent Adaptive Mutation Based Genetic Algorithm For Combinatorial Optimization Problems" IEEE, ICECT Kanyakumari, April 6-8, 2012.

[9] Jacek M. Zurada "Introduction to artificial neural systems"

[10] TSPLIB. http://www.tsp.gatech.edu/index.html

[11] Noraini Mohd Razali, John Geraghty "Genetic Algorithm Performance with different selection strategies in solving TSP"

[12] Monica Sehrawat and Mr. Sukhvir Singh "Modified Order Crossover (OX) Operator IJCSE – ISSN:0975-3397 Vol.3 May 2011

[13] Basima Hani Hasan and Moutaz Saleh Mustafa "Comparative Study of Mutation Operators on the Behavior of Genetic Algorithms Applied to Non-deterministic Polynomial (NP) Problems" IEEE 2011

[14] Guochu Chen "Intelligent Adaptive Genetic Algorithm and Its Application" IEEE 2011

[15] S.C.Tam, C.H.Chio and H.K.Tam "Development of a new optimization method, yin-yang algorithm, for traveling salesman problem" IEEE 2011

[16] S N S Kalyan Bharadwaj.B, Krishna Kishore.G, Srinivasa Rao.V "Solving Traveling Salesman Problem using Hierarchical Clustering and genetic algorithm" Proc. Of (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (3), 2011, 1096-1098

◈◈◈

Ant Colony Optimized Routing for Mobile Adhoc Networks (MANET)

Dweepna Garg & Parth Gohil

Dept. Of Computer Science and Engineering, Babaria Institute of Technology, Varnama, Vadodara, India

Abstract - A Mobile Ad-Hoc Network (MANET) is a collection of wireless mobile nodes forming a temporary network without using centralized access points, infrastructure, or centralized administration. Routing means the act of moving information across an internet work from a source to a destination. The biggest challenge in this kind of networks is to find a path between the communication end points, what is aggravated through the node mobility. In this paper we present a new routing algorithm for mobile, multi-hop ad-hoc networks. The protocol is based on swarm intelligence. Ant colony algorithms are a subset of swarm intelligence and consider the ability of simple ants to solve complex problems by cooperation. The introduced routing protocol is well adaptive, efficient and scalable. The main goal in the design of the protocol is to reduce the overhead for routing. We refer to the protocol as the Ant Colony Optimization Routing (ACOR).

Keywords - *Mobile ad hoc networks (MANET), Swarm Intelligence, Ant Colony Optimization Routing (ACOR).*

I. INTRODUCTION

A mobile ad-hoc network (MANET) is a set of mobile nodes which communicate over a wireless medium over single or multiple and do not need any infrastructure such as access points or base stations. Therefore, mobile ad-hoc networks are suitable for temporary communication links.

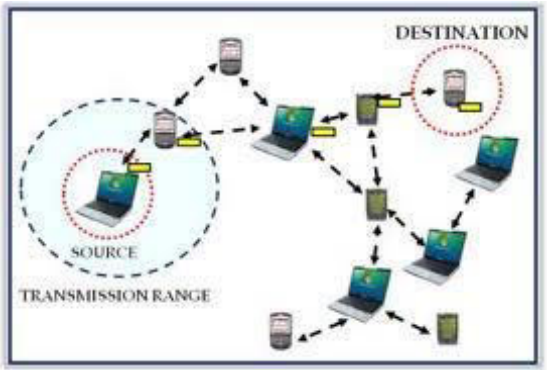


Fig. 1 : Mobile Adhoc Network

Nodes not only have to fulfil the functionality of hosts, but also each node has also to be a router, forwarding packets for other nodes. One interesting application for mobile ad-hoc networks beside the classical ones, disaster and military applications, is the deployment of mobile ad-hoc networks for multimedia applications. The biggest challenge in this kind of networks is still, the finding of a route between the communication end-points, which is aggravated through the node mobility.

This paper deals with a routing algorithm ACOR used for mobile, multi-hop ad-hoc networks to improve the performance of the existing protocol of mobile ad hoc network, which is based on swarm intelligence and especially on the ant colony based metaheuristic.

II. FUNDAMENTALS OF SWARM INTELLIGENCE

Swarm intelligence is an emerging field of biologically-inspired artificial intelligence based on the behavioural models of social insects such as ants, bees, wasps and termites. Respective algorithms are made based on ants, bees, wasps such as Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO) and Bee Swarm Optimization (BSO). It is a scientific theory about how Complex and sophisticate behaviours can emerge from social creature group. It helps in designing a framework for designing distributed algorithms which are originally derived by studying models of social insect behaviour. Swarm intelligence is a probabilistic method for building probabilistic paths between nodes based on simple rules.

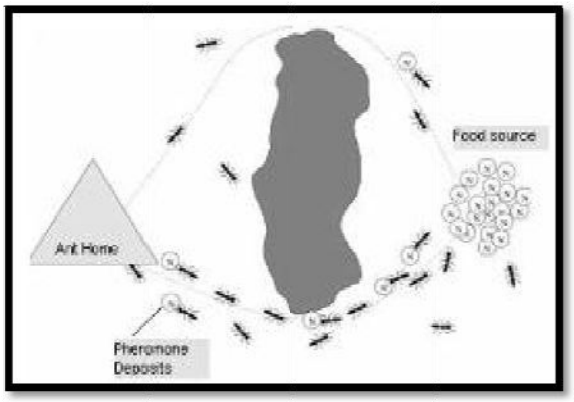


Fig. 2 Shortest Path Taken by Ant after initial searching time

In general, ant colony optimization meta-heuristic tries to solve a combinatorial problem using the collaboration of a group of simple agents called artificial ants. ACO routing algorithms establish optimum paths to the destination using a number of artificial ants that communicate indirectly with each other by *Stigmergy*. *Stigmergy* is a way of indirect communication between individuals which, in an ad-hoc network case, is done through the modification of some parameters in the nodes of the network. The Ant colony optimization is based on the foraging behaviour of ants. When ants search for food, they wander randomly and upon finding food return to their colony while laying a chemical substance called pheromone. Many ants may travel through different routes to the same food source. The ants, which travel the shortest path, reinforce the path with more pheromone that aids other ants to follow. Subsequently more ants are attracted by this pheromone trail, which reinforces the path even more. This autocatalytic behaviour quickly identifies the shortest path.

III. ROUTING IN MANETS

Routing in MANET is a Dynamic Optimization Problem as the search space changes over time. The routing policy is defined as the rule that specifies what node to take next at each decision node to reach the destination node. Due to the time varying nature of the topology of the networks, traditional routing techniques such as distance-vector and link-state algorithms that are used in fixed networks, cannot be directly applied to mobile ad hoc networks. The constraints of MANETs demand the need of specialized routing algorithms that can work in a decentralized and self-organizing way. The routing protocol of a MANET must dynamically adapt to the variations in the network topology.

The routing scheme in a MANET can be classified into two major categories – Proactive and Reactive. The proactive or table driven routing protocols maintain routes between all node pairs all the time. It uses periodic broadcast advertisements to keep routing table up-to-date. This approach suffers from problems like increased overhead, reduced scalability and lack of flexibility to respond to dynamic changes. The reactive or on-demand approach is event driven and the routing information is exchanged only when the demand arises. The routing is initiated by the source.

ANTNET and *ANTHOCNET* are two well known ant colony based routing algorithms. ANTNET is a proactive and ANTHOCNET is a reactive routing algorithm. They have a very high delivery rate and find routes whose lengths are very close to the length of the shortest path. The drawback of ANTHOCNET is the number of routing messages that needs to be sent in the network for establishing routes to the destination and the disadvantage of ANTNET is the time needed before a system of paths between the nodes of the network is established. This is referred to as the convergence time. Regarding the dynamic nature of mobile ad-hoc networks, a long convergence time is a significant drawback.

IV. ANT COLONY OPTIMIZATION

The simple ant colony optimization (ACO) meta-heuristic can be used to find the shortest path between a source node *vs* and a destination node *vd* on the graph *G*.

Let $G = (V, E)$ be a connected graph with $n = |V|$ nodes. The pheromone concentration, $\Phi i,j$ is an indication of the usage of the edge i, j. An ant located in node *vi* uses pheromone $\Phi i,j$ of node *vj* and *Ni* to compute the probability of node *vj* as next hop. *Ni* is the set of one-step neighbours of node *vi*. An ant located in node *vi* uses pheromone □*i,j* of node *vj* □ *Ni* to compute the probability of node *vj* as next hop. *Ni* is the set of one-step neighbours of node *vi*.

$$p_{i,j} = \begin{cases} \frac{\varphi_{i,j}}{\sum_{j \in N_i} \varphi_{i,j}} & \text{if } j \in N_i \\ 0 & \text{if } j \notin N_i \end{cases} \quad \text{(i)}$$

The transition probabilities *pi,j* of a node *vi* fulfil the constraint:

$$\sum_{j \in N_i} p_{ij} = 1, i \in [1, N] \quad \text{(ii)}$$

An ant changes the amount of pheromone of the edge *e*(*vi, vj*) when moving from node *vi* to node *vj* as follows:

$$\varphi_{i,j} = \varphi_{i,j} + \Delta\varphi \quad \text{(iii)}$$

Like real pheromone the artificial pheromone concentration decreases with time to inhibit a fast convergence of pheromone on the edges.

$$\varphi_{i,j} = (1-q).\varphi_{ij},\ q \in [0,1 \quad \text{(iv)}$$

A. Reasons for ACO ensembles to ad hoc networks

The reasons why ACO suits to Adhoc networks are:

a) **Dynamic topology:** This property is responsible for the bad performance of several routing algorithms in mobile multi-hop ad-hoc networks. The ant colony optimization meta-heuristic is based on agent systems and works with individual ants. This allows a high adaptation to the current topology of the network.

b) **Local work:** In contrast to other routing approaches, the ant colony optimization meta-heuristic is based only on local information, i.e., no routing tables or other information blocks have to be transmitted to neighbors or to all nodes of the network.

c) **Link quality:** It is possible to integrate the connection link quality into the computation of the pheromone concentration, especially into the evaporation process. This will improve the decision process with respect to the link quality. It is here important to notice, that the approach has to be modified so that nodes can also manipulate the pheromone concentration independent of the ants, i.e. data packets, for this a node has to monitor the link quality.

d) **Support for multi-path:** Each node has a routing table with entries for all its neighbors, which contains also the pheromone concentration. The decision rule, to select the next node, is based on the pheromone concentration on the current node, which is provided for each possible link. Thus, the approach supports multipath routing.

V. ANT BASED ALGORITHMS

A. Ant Based Control (ABC) Routing

Ant-Based Control is another stigmergy-based ant algorithm designed for telephone networks. The basic principle relies on mobile routing agents, which randomly explore the network and update the routing tables according to the current network state. The routing table, stores the probabilities instead of pheromone concentrations. ABC only uses a single class of ants (i.e. forward ants), which are initiated at regular time intervals from every source to a randomly chosen destination. After arriving at a node they immediately update the routing table entries for their source node, meaning that the pheromone pointing to the previous node is increased. When ants have reached their destination, they die. The increase in these probabilities is a decreasing function of the age of the ant, and of the original probability. The ants get delayed on parts of the system that are heavily used. Some noise can be added to avoid freezing of pheromone trails.

*B. A*nt Colony Optimized Routing (ACOR)

The ant colony optimization routing (ACOR) algorithm is a probabilistic technique for solving computational problems which can be reduced to finding good paths through graphs. This algorithm is a member of ant colony algorithms family, in swarm intelligence methods, and it constitutes some meta-heuristic optimizations. ACOR has two phases. They are: Route Discovery phase and Route Maintenance phase. Both of these phases use FAnt (Forward Ant) and BAnt (Backward Ant). A FANT is an agent which establishes the pheromone track to the source node. In contrast, a BANT establishes the pheromone track to the destination node.

1) Route Discovery

Each FANT has a unique sequence number to avoid duplicates. A node receiving a FANT for the first time creates a record [destination address, next hop, pheromone value] in its routing table. The node interprets the source address of the FANT as destination address, the address of the previous node as next hop, and computes the pheromone value depending on the number of hops the FANT needed to reach the node. Then the node relays the FANT to its neighbours. When the FANT reaches destination, it is processed in a special way. The destination node extracts the information and then destroys the FANT.

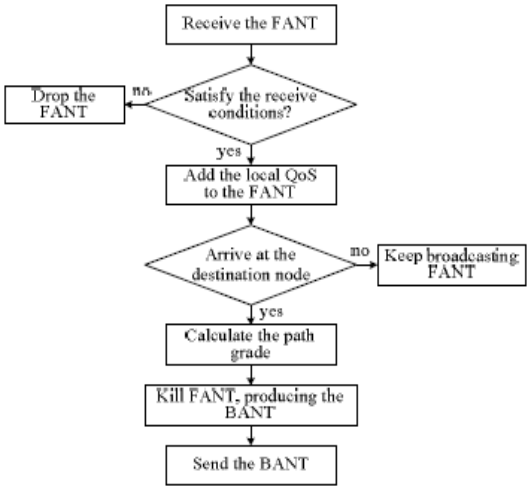


Fig. 3 The process of FANT disposing

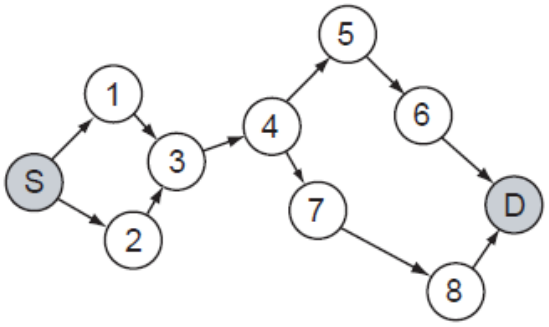


Fig. 4 A FANT is flooded from source to destination

A BANT is created and sent towards the source node. In that way, the path is established and data packets can be sent. Data packets are used to maintain the path, so no overhead is introduced. Pheromone values are changing

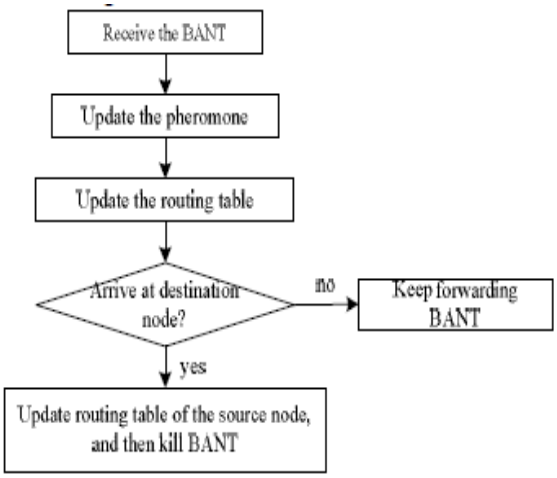


Fig. 5 The process of the BANT disposing

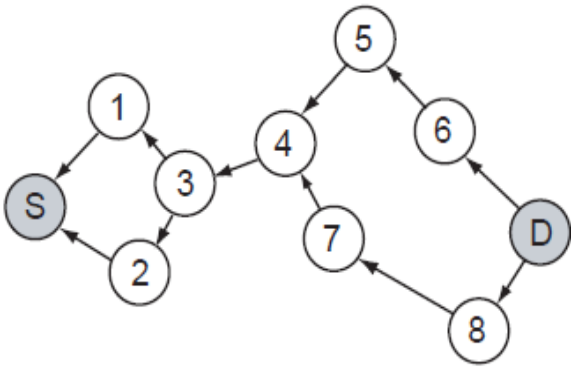


Fig. 6 A BANT takes the reverse path of FANT

2) Route Maintenance

The second phase of the routing algorithm is called route maintenance, which is responsible for the improvement of the routes during the communication. ACOR does not need any special packets for route maintenance. Once the FAnt and BAnt have established the pheromone tracks for the source and destination nodes, subsequent data packets are used to maintain the path. Similar to the nature, established paths do not keep their initial pheromone values forever. When a node (relay node) relays a data packet toward the destination (destination address*)* to a neighbour node (next hop), it increases the pheromone value of the entry (destination address, next hop, pheromone value*)* by pheromone function, i.e., the path to the destination is strengthened by the data packets. In contrast, the next hop (next hop) increases the pheromone value of the entry (source address, relay node, pheromone value) by pheromone function, i.e. the path to the source node is also strengthened. The evaporation process of the real pheromone is simulated by regular decreasing of the pheromone values. The above method for route maintenance could lead to undesired loops. ACOR prevents loops by a very simple method, which is also used during the route discovery phase. Nodes can recognize duplicate receptions of data packets, based on the source address and the sequence number. If a node receives a duplicate packet, it sets the DUPLICATE ERROR flag and sends the packet back to the previous node. The previous node deactivates the link to this node, so that data packets cannot be sending to this direction any more.

3) Route Failure Handling

ACOR handles routing failures, which are caused especially through node mobility and thus very common in mobile ad-hoc networks. ACOR recognizes a route failure through a missing acknowledgement. If a node gets a ROUTE ERROR message for a certain link, it first deactivates this link by setting the pheromone value to 0. Then the node searches for an alternative link in its routing table. If there is a second link it sends the packet via this path. Otherwise the node informs its neighbours, hoping that they can relay the packet. Either the packet can be transported to the destination node or the backtracking continues to the source node. If the packet does not reach the destination, the source has to initiate a new route discovery phase.

4) Requirements of ACOR

The requirements needed to be fulfilled for routing algorithm for mobile ad-hoc networks are:

a) **Distributed operation:** In ACOR, each node owns a set of pheromone counter^2 □*i,j* in its routing table for a link between node *vi* and *vj* . Each node controls the pheromone counter independently, when ants visit the node on route searches.

b) **Loop-free:** The nodes register the unique sequence number of route finding packets, FANT and BANT, so they do not generate loops.

c) **Demand-based operation:** Routes are established by manipulating the pheromone counter □*i,j* in the nodes. Over time, the amount of pheromone decreases to zero when ants do not visit this node. A route finding process is only run, when a sender demands.

d) **Sleep period operation:** Nodes are able to sleep when their amount of pheromone reaches a threshold. Other nodes will then not consider this node.

Additional Requirements of ACOR are:

a) **Locality:** The routing table and the statistic information block of a node are local and they are not transmitted to any other node.

b) **Multi-path:** Each node maintains several paths to a certain destination. The choice of a certain route depends on the environment, e.g., link quality to the relay node.

c) **Sleep mode:** In the sleep mode a node snoops, only packets which are destined to it are processed, thus saving power

C. Probabilistic Emergent Routing Algorithm [PERA]

Mobile ad hoc networks are infrastructure-less networks consisting of wireless, possibly mobile nodes which are organized in peer-to-peer and autonomous fashion. The highly dynamic topology, limited bandwidth availability and energy constraints make the routing problem a challenging one. In this paper we take a novel approach to the routing problem in MANETs by using swarm intelligence inspired algorithms. The proposed algorithm uses Ant-like agents to discover and maintain paths in a MANET with dynamic topology. Multiple paths are set up, but only the one with the highest pheromone value is used by data and the other paths are available for backup. The route discovery and maintenance is done by flooding the network with ants. Both forward and backward ants are used to fill the routing tables with probabilities. These probabilities reflect the likelihood that a neighbour will forward a packet to the given destination. Multiple paths between source and destination are created.

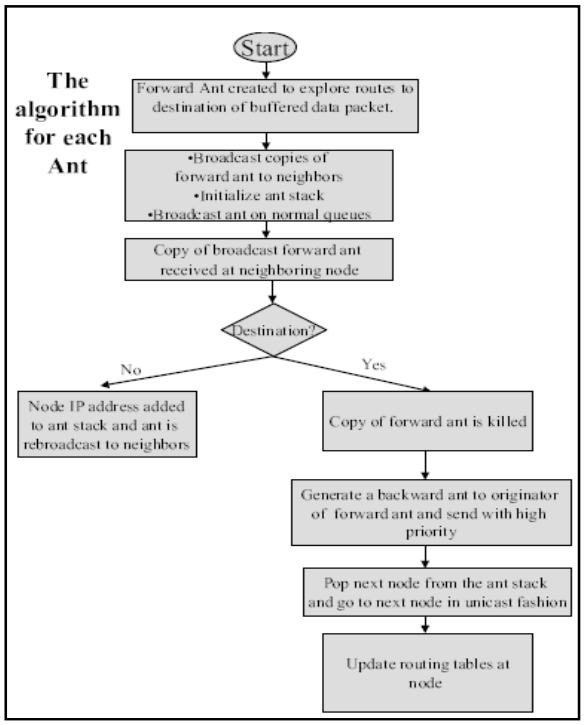


Fig. 7 Algorithm for each ant

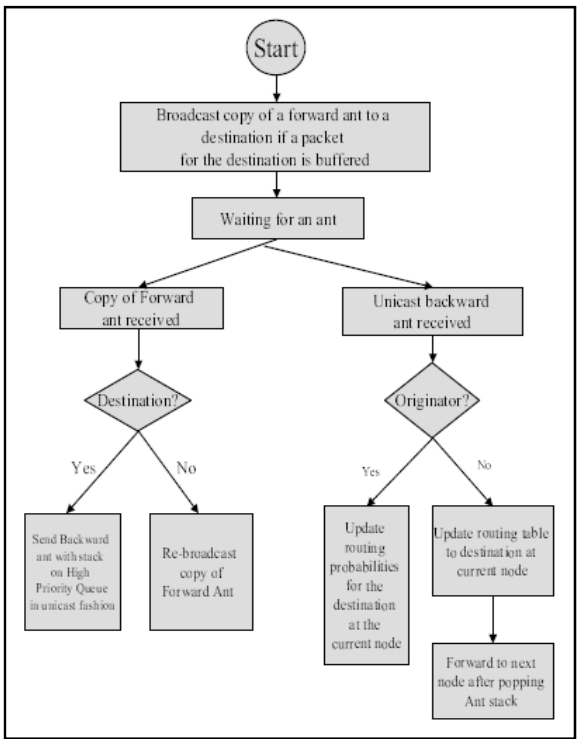


Fig. 8 Algorithm for each node

D. Ant Agents for Hybrid Multipath Routing [AntHocNet]

AntHocNet is a hybrid multipath algorithm. The algorithm consists of both reactive and proactive components. In a reactive path setup phase, multiple paths are built between the source and destination of a data session. Data are stochastically spread over the different paths, according to their estimated quality. During the course of the session, paths are continuously monitored and improved in a proactive way. Link failures are dealt locally. The algorithm makes extensive use of ant-like mobile agents which sample full paths between source and destination nodes. Backward ants are used to actually setup the route. While the data session is open, paths are monitored, maintained and improved proactively using different agents, called proactive forward ants.

E. Antnet- Ant Algorithm

In AntNet, a group of mobile agents (or artificial ants) build paths between pair of nodes; exploring the network concurrently and exchanging obtained information to update the routing tables. Ants (F-ants) are launched at regular instants from each node to randomly chosen destinations.

Ants are routed probabilistically with a probability function of:

(i) Artificial pheromone values, and

(ii) Heuristic values, maintained on the nodes

- Ants memorize visited nodes and elapsed times
- Once reached their destination nodes, ants retrace their paths backwards (B-ants), and update the pheromone tables.

AntNet is distributed and not synchronized.

1) Role of F-ants and of B-ants

F-ants collect implicit and explicit information on available paths and traffic load

- implicit information, through the arrival rate at their destinations (remember the differential length effect)
- explicit information, by storing experienced trip times

F-ants share queues with data packet. B-ants fast back propagate info collected by F-ants to visited nodes. B-ants use higher priority queues

VI CONCLUSION

In this paper we presented a on-demand routing approach for mobile multi-hop ad-hoc networks. The approach is based on swarm intelligence and especially on the ant colony optimization meta-heuristic. Routing of data packets is only through optimal path which is generated by route discovery phase as defined by ACOR. Route maintenance is done periodically to retain optimal path. This is done through data packets. Due to change in topology of adhoc network, existing routes may fail or new paths may be generated. In order to adapt for the change in topology, route refreshing is done periodically.

REFERENCES

[1] Anandamoy Sen: Swarm Intelligence based optimization of MANET cluster formation, 2006

[2] G. D. Caro, F. Ducatelle, and L. M. Gambardella. Anthocnet: An adaptive nature-inspired algorithm for routing in mobile ad hoc networks. Special Issue on Self-Organisation in Mobile Networking, 16:443–455, 2005

[3] M. Gunes, U. Sorges, and I. Bouazizi: Ara – the ant-colony based routing algorithm for manets. In Proceedings of the 2002 International Conference on Parallel Processing Workshops, pages 79–89, 2002

[4] Di Caro, G., Dorigo, M, "Antnet: Distributed stigmergetic control communications networks

[5] AntHocNet: an Ant-Based Hybrid Routing Algorithm for Mobile Ad Hoc Networks 2005

[6] Andrew S Tannenbaum, "Computer Networks", 4th Edition, Prentice- Hall of India

◈◈◈

An Analytical Study of Different Crossover Operators for De Jong's Optimization

Monika & Rakesh Kumar

Dept. of Computer Science and Application, Kurukshetra University, Kurukshetra, India

Abstract **–** Genetic is one the most efficient soft computing technique, used by many of the computer science applications to optimize the throughput. Genetic have great importance in almost every field of computer science. Different stages of Genetics include Initialization, Selection, Crossover and mutation. Here we are presenting a comparative analysis of different crossover operations with genetics. In this work we are implementing the Genetic Algorithm on De Jong's First function called Sphere Model and on De Jong's Sixth Function called Rastrigin's function. The proposed work is to optimize this model by implementing one point and two point crossover. We are also showing the analysis of different selection approaches such as ranking and Roulette Wheel The analysis is presented in the form of graph.

Keywords - *Genetics, Selection, Crossover, Optimization, DeJong, Mutation.*

I. INTRODUCTION

The genetic algorithm is a method for solving both constrained and unconstrained optimization problems that is based on natural selection, the process that drives biological evolution. The genetic algorithm repeatedly modifies a population of individual solutions. At each step, the genetic algorithm selects individuals at random from the current population to be parents and uses them to produce the children for the next generation. Over successive generations, the population "evolves" toward an optimal solution. Genetic algorithm can be applied to solve a variety of optimization problems that are not well suited for standard optimization algorithms, including problems in which the objective function is discontinuous, non differentiable, stochastic, or highly nonlinear.

The genetic algorithm uses three main types of rules at each step to create the next generation from the current population:

A genetic algorithm is modeled on genetics and Darwinian evolution, whereas a neural network is based on models of human cognition. One common application of a genetic algorithm is as a function optimizer. Another common application of a genetic algorithm is to evolve organisms that perform well in a given environment. In either application, the GA is based on the survival-of-the-fittest (natural selection) tenet of Darwinian evolution. A genetic algorithm consists of a population of individuals that reproduce (over many Generations) according to their fitness in an environment. Those individuals that are most fit are most likely to survive, mate, and bear children. Children are created by the stochastic application of genetic operators to the (parent) individuals. Individuals of the population, coupled with the genetic operators, combine to perform an efficient domain-independent search strategy that makes few assumptions about the search space. Each individual in a population is a point in the search space.

Traditionally, an individual in a GA is represented as a bit string of some length n. Each individual thus represent some point in a space of size $2n$.Given a bit string representation. These representations are all single-stranded in the sense that one piece of genetic material represents an individual. Such representations are termed *haploid*. However, natural genetics makes use of double stranded chromosomes (*diploid*) as well. For example, suppose two bit strings represent an individual:

1010001010

0010101001

These double strands can contain different and possibly conflicting information. In nature, *dominance* is the primary mechanism for conflict resolution. Supposing 1 to dominate 0, the individual *phenotype* can be expressed as:

1010101011

A genetic algorithm will produce offspring using the genetic operators' crossover (recombination) and mutation. Mutation operates at the bit level by randomly flipping bits within the current population. Mutation rates are low, generally around one bit per thousand. Here Figure 1 is showing the basic Architecture of Genetic Algorithm.

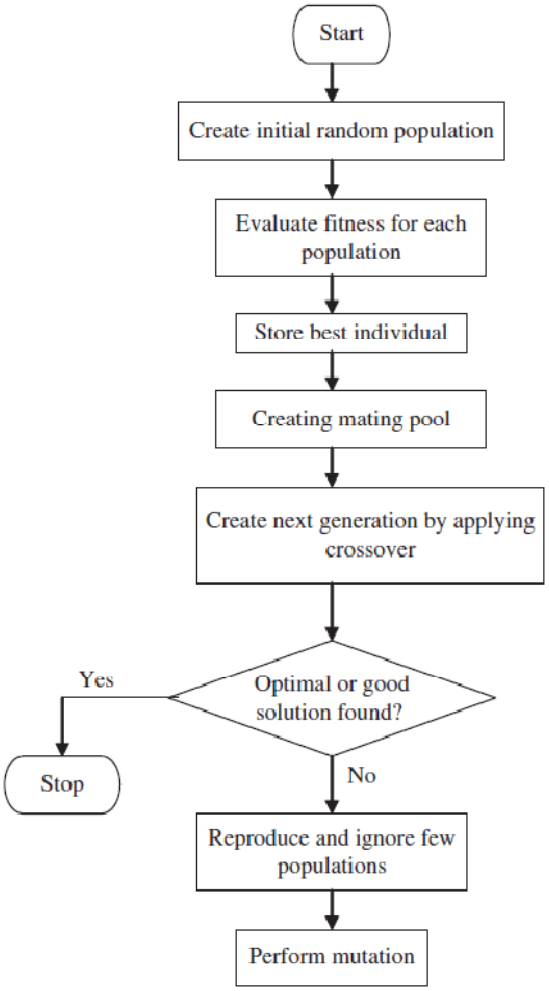


Figure 1: GA Architecture

Each stage of Genetics itself is a Research Area. Lot of work is done by many of researchers on each stage of Genetics. Number of problem are being solved and optimized by using this Genetic Algorithm. Here we presenting the comparative analysis of some different operations or algorithms performed on different stages of Genetic Algorithm.

To present the comparative analysis we are using the DeJong's First function as the problem set or the initial population set. Dejong function 1 is the simplest function provided by Dejong. It basically represents the Sphere. The working on such function is called the parametric optimization using Genetics. The De jong's Function 1 is called Sphere Model. The basic equation of Dejong Function 1 is given as

$$f(x) = \sum_{i=1}^{n} x_i^2.$$

$-5.12 <= x_i <= 5.12$

In the equation form we can write it as

F(x)= sum(x(i)^2)

Where i>=1 and i<=n

And x(i)>=-5.12 and x(i)<=5.12

We have also implemented the same work on De Jong Function 6 also, It is called Rastrigin function. This function is is based on function De Jong 1 with the addition of cosine modulation to produce many local minima. Thus, the test function is highly multimodal. However, the locations of the minima are regularly distributed.. The basic equation Dejong function 6 is

$$f(x) = 10n + \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\pi x_i)\right]$$

$$-5.12 \le x_i \le 5.12.$$

global minimum f(x)=0; x_i =0, i=1:n.

II. LITERATURE SURVEY

John Holland has explained the adaptive processes of natural systems. In order to retain some mechanisms of natural systems, they have designed some kind of artificial system software for this purpose [11, 3]. GA differs from other search algorithms in that it has a unique characteristic [11]. It only needs the input parameters of a certain problem and represent these inputs in a chromosome format. Thus, it is unaware of the problem itself. This is the reason why GAs can be applied to many types of complex problems [13, 11]. Usage of genetic algorithms began by solving academic problems such as the traveling salesman problem and the 8 Queens problem [11, 2, 3,12]. Years later, GAs grew rapidly. In a way, they increased their applications to optimize complex scheduling problems, spatial layout and many other types that are hard to efficiently maximize [2].

The Simple Standard Genetic algorithm works randomly in selecting parents. In choosing two individuals to mate together there are no constraints [3]. Many studies have been done to tackle this problem trying to overcome it, and trying to design structured population with some control on how individuals interact [3].

By Gorges-Schleuter, 1989 [3]. It is called a diffusion model. A two-dimensional Grid world is used here to arrange the individuals where these individuals interact with each other by the direct neighborhood of each individual [12]. These individuals will be distributed on a graph which is connected together; each individual connects with its neighborhood by a genetic operator. This type of GAs is designed as a probabilistic cellular automation. A self-organizing schedule is added to reproduce an operator [2]. The individual which can interact with its immediate neighbors can only be held in the cell.

According to the increasing complex problems which appear in evolutionary computation, more advanced models of evolutionary algorithms (EAs) appear. Island models are considered a family of such models [2]. Here the individuals are divided into sections. We call each section a subpopulation which is referred to as an island. These island models are able to solve problems in a better performance than standard models [3]. There is a specific relation between islands through some exchange of some individuals between islands. This process is called migration; this is what island models are famous of, and without these migrations, each island is considered as a set of separate run. Therefore migration is very important [2]

III PROPOSED WORK

In this proposed work, We are implementing different algorithms on different stages of Genetic Algorithm. Here we presenting work on two main stages called selection and Crossover.

1) Selection

Genetic Algorithm always evaluate from population, so they create the number of parents required to form the population. The most tidies job in GA is to define the fitness function, depending upon the requirement and application we have to define the function and calculate the fitness value of the entire parent, depending upon the fitness value the parent whose fitness value is low has to be replaced by the parent of higher fitness value and thus the population is called mating pool. Now how this population will be selected is the basic selection operator. Here we are defining the selection operation based on two algorithms Rank Based Selection and Roulette Wheel Selection.

1.1 Roulette Wheel Selection

The roulette Wheel selection procedure is based on the fitness values of data items. To perform the evaluation it required a real value interval. On the basis of the the real expected selection probability is defined based on some function.

$$S = \Sigma Pi$$

Here Pi defines the probability of the selection of population data item. This Pi depends on the fitness value given as

$$Pi = Fi / \Sigma Fi$$

Here Fi defines the Fitness value and Σ Fi defines the overall fitness value. The fitness value is based on the common integrals defined as the interval. To perform the selection process a random number is selected within the interval and the individual with higher span is elected as the population vector.

2) Crossover

Cross over is the most important operator of Genetics that is the main reason to generate a new child. The function of the crossover operator is to generate new child chromosomes by using two parents' chromosomes. This function will intermix the information of these two parents in some defined ways so that a new child will be extracted out of it. There are different kind of crossover exist to generate the next child. In this work we have basically worked on two main crossover types called One Point Crossover and Two Point Crossover.

2.1 Single-Point Crossover

It is one the simplest type of crossover. In this crossover type one crossover point is chosen from the set of variables present in the chromosome. All the variables in the chromosome will be chosen uniformly at random. To perform the crossover the variables will be exchanged between the individual about the single point. It will derive a new child from it. The example is given as

Parent 1	**0**	**1**	**1**	**1**	**0**	0	1	1	0	1	0
Parent 2	**1**	**0**	**1**	**0**	**1**	1	0	0	1	0	1

Now let the crossover is performed at 5th position, new generation will be

Child 1	**0**	**1**	**1**	**1**	**0**	1	0	0	1	0	1
Child 2	**1**	**0**	**1**	**0**	**1**	0	1	1	0	1	0

2.2 Multipoint Crossover

In case of Multiple Crossover more then one crossover points are chosen at random without taking any duplicate values. It will also return the values in increasing order. The values between two crossover points will be interchanged so that new child nodes can be generated. The values between the first variable and

first crossover point will not be changed. We can see the effect as given

Parent1	**0**	**1**	1	1	0	0	**1**	**1**	**0**	**1**	0
Parent2	**1**	**0**	1	0	1	1	**0**	**0**	**1**	**0**	1

As the multipoint cross over will be performed the output will be

Child1	**0**	**1**	1	0	1	1	**0**	**1**	**1**	**1**	1
Child2	**1**	**0**	1	1	0	0	**0**	**0**	**1**	**0**	0

2.3 Arithmetic Crossover

Arithmetic crossover generates two child nodes of two parents. The values of the children are set as the average values of the parent and the values of other child is given by an equation (3*p1-p2)/2 where p1 and p2 are he parent nodes. As the arithmetic crossover is applied, the resultant values created from crossover must be in the range of limited values.

IV IMPLEMENTAITON AND RESULTS

In this proposed work we have optimize the De Jong Function 1 and Function 6 using Genetics. One of the main difficulties in building a practical GA is in choosing suitable values for parameters such as population size, probability of crossover (*Pc*), and probability of mutation (*Pm*). In this experiment, we follow De Jong's guideline which is to start with a relatively high *Pc* (≥ 0.6), relatively low *Pm* (0.001-0.1), and a moderately sized population. The selections of parameter values are very depend on the problem to be solved. Noted that the larger the population size, the longer computation time it takes. We have implemented the Genetics with different Crossovers including Single Point, Multipoint and Arithmetic Crossover. The selection approach we have used is Roulette Wheel Selection. Here we are presenting the results driven from these implementations.

A. EXPERIMENTAL RESULTS

TABLE1 shows the best results obtained by applying different crossover strategy on De Jong's function 1. It is clearly shows that GA with Arithmetic crossover always converges fastly. This is then followed by Multi point and Single point crossover. Multi point and single point crossover is able to achieve optimal solution for small size instances; however the quality of solution reduces as the size of instance increase. Figure 1 shows the comparison of different crossover operator on De Jong's function 1 by bar graph.

TABLE1. RESULTS OF DIFFERENT CROSSOVER OPERATOR FOR De Jong's FUNCTION 1

Sr. No.	Number of Iterations	Single Point Crossover	Arithmetic Crossover	Multi Point Crossover
1	20	379875	201235	237813
2	40	253621	123136	167486
3	60	134562	82013	109940
4	80	94523	75326	7133.4
5	100	26324	12362	6503
6	120	13268.6	6213	4133
7	140	10316.632	3210.	3866
8	160	8632	1369.545	2142
9	180	6125.126	1324.156	1125
10	200	4139.326	1023.267	873.263
11	220	2962.632	345.63	621.6
12	240	1423.128	219.23	210.36
13	260	832.321	86.3	103.26
14	280	213.16	34.69	43.7116
15	300	131.183	19.46	35.717

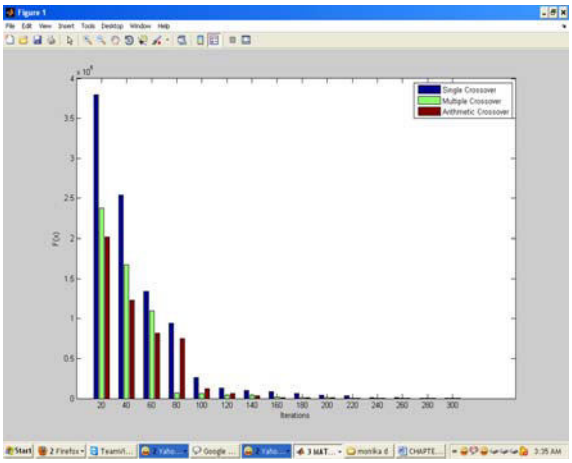


Figure2: Comparison of different crossover operator (function 1)

TABLE2 shows the best results obtained by applying different crossover strategy on De Jong's function 6. It is clearly shows that GA with Arithmetic crossover always converges fastly. This is then followed by Multi point and Single point crossover. Multi point and single point crossover is able to achieve optimal solution for small size instances. Figure 2 shows the comparison of different crossover operator on De Jong's function 6 by bar graph.

TABLE 2 . RESULTS OF DIFFERENT CROSSOVER OPERATOR FOR De Jong's FUNCTION 6

Sr. No.	Number of Iterations	Single Point Crossover	Arithmetic Crossover	Multi Point Crossover
1	20	297844	520739	515098
2	40	200409	394045	344637
3	60	129543	301812	182669
4	80	109702	225239	139315
5	100	95947.5	143477	103734
6	120	63735.2	65158.1	68140.5
7	140	49235.7	44266.4	55520.1
8	160	46606.5	17465.3	42477.2
9	180	40725.5	15165	24597.2
10	200	36966.1	13484.2	19540.9
11	220	23470.1	12191.9	15484.4
12	240	18855.8	8931.97	11836.6
13	260	2333.82	7406.7	10104.1
14	280	1625.12	2947.64	8099.73
15	300	1189.2	999.75	5901.57

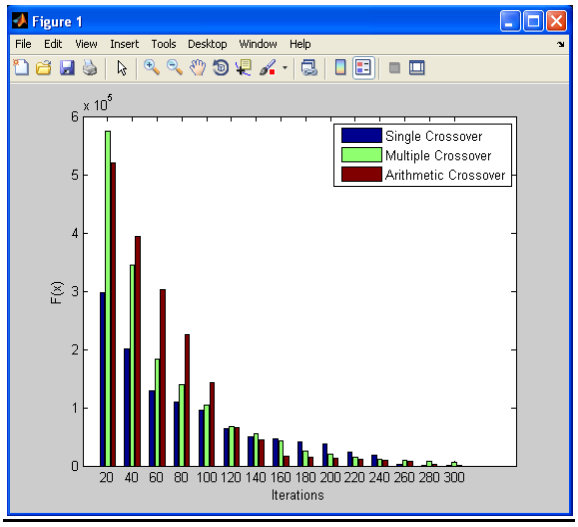


Figure 3: Comparison of different crossover operator (function 6)

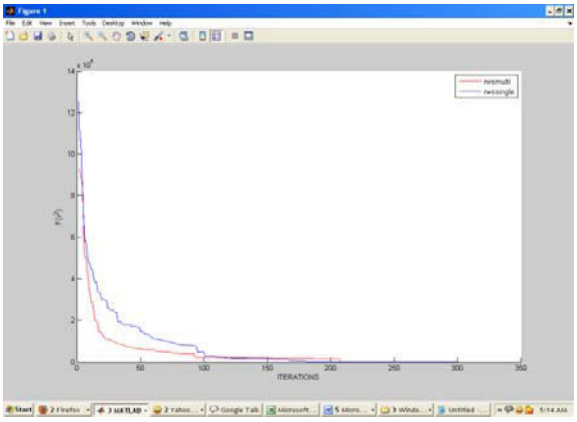


Figure 4: Single and Multipoint Crossover (De Jong Func 1)

As we can see in figure 4, We have implemented the Multipoint and Single point Crossover on De Jong function 1. Here we have used Roulette Wheel Selection approach. From this figure we can conclude that the convergence rate of multipoint crossover is faster as compared to singe point crossover. We have also implemented same crossover and selection algorithm to optimize the De Jong function 6 also. The results obtained are shown in figure 5. In this function also the convergence rate of multipoint crossover is faster. By this we can conclude that multipoint crossover always improve the convergence rate for different functions, the work is tested on De Jong function 1 and 6.

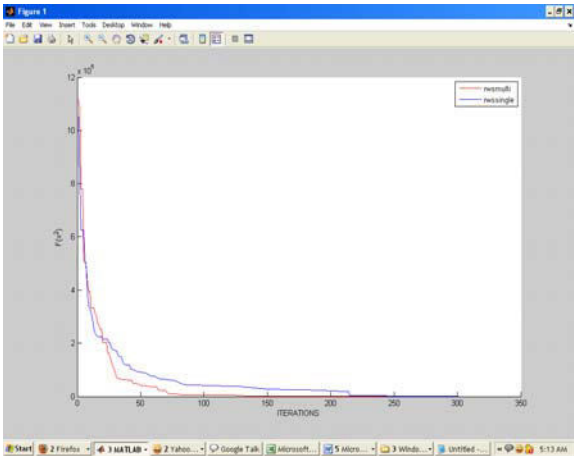


Figure 5: Single and Multipoint Crossover (De Jong Func 6)

To verify the comparative analysis of Multipoint and Single Point crossover, we also implement these crossovers with simple selection algorithm. Figure 6 and Figure 7 are showing the results. These crossover and simple selection is implemented to optimize the

Function 1 and Function 6. Figure 6 is showing the optimization results of Dejong function 1. In this figure we can conclude that the convergence rate of multipoint crossover is fast as compared to single point crossover.

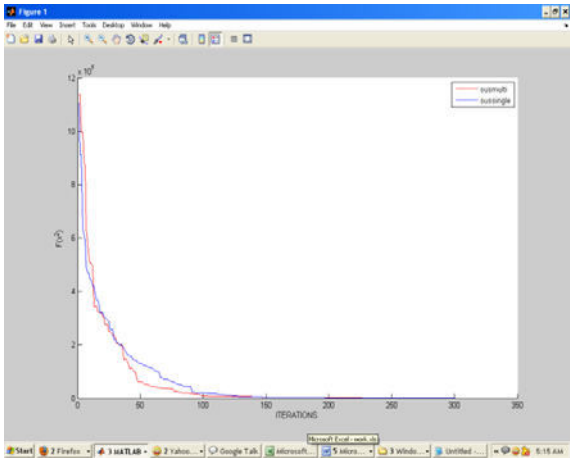


Figure 6: Single and Multipoint Crossover (De Jong Func1)

Figure 7 is showing the optimization results of Dejong function 6. In this figure we can conclude that the convergence rate of multipoint crossover is fast as compared to single point crossover. From these two figures we can conclude the multipoint crossover gives better convergence ratio with different De Jong functions as well as different selection methods.

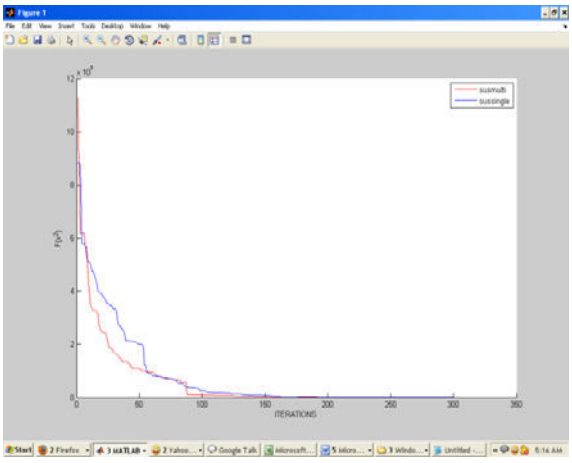


Figure 7: Single and Multipoint Crossover (De Jong Func 6)

In this work we have optimized the De Jong functions with variation of Arithmetic and single point Crossover. These crossover operations are implemented with Roulette Wheel Selection method to optimize the Dejong Function 1 and 6. Here figure 8 is showing the comparative analysis of Arithmetic and single point Crossover for dejong function 1. As we can see the convergence ratio of Arithmetic crossover is better then binary crossover.

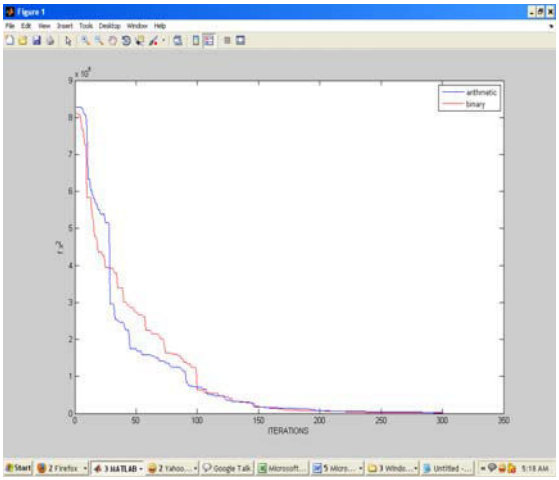


Figure 8: Single and Arithmetic Crossover (De Jong Func1)

Same work is also tested on Dejong Function 6. Here figure 9 is showing the comparative analysis of Arithmetic and Single point Crossover for Dejong function 6. As we can see the convergence ratio of Arithmetic crossover is better then Single point crossover. Here we can conclude that the converge ratio of Arithmetic crossover is better then binary crossover for different functions. Arithmetic crossover will always give results with less number of iterations so that the results obtained more quickly and the optimization will be done more efficiently.

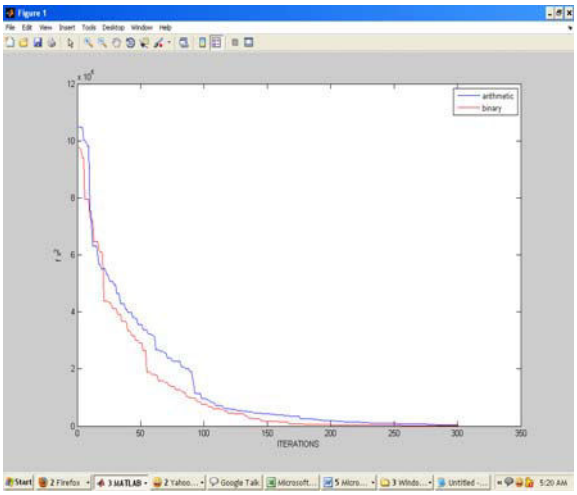


Figure 9: Single and Arithmetic Crossover (De Jong Func 6)

In this work we have focused on Arithmetic Crossover and analyze it for the different levels. The results obtained by applying the Arithmatic crossver on different levels for dejong function 6, is shown in figure 10,11 and 12.

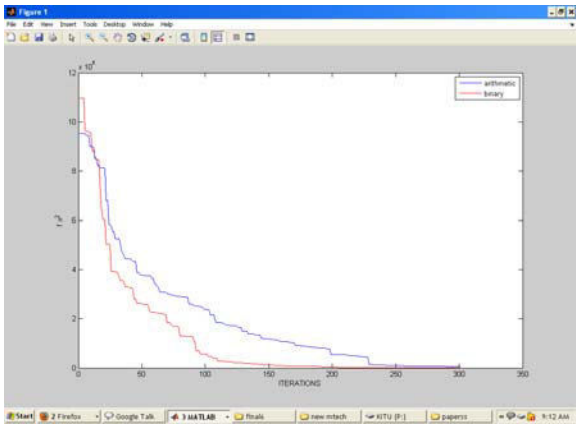


Figure 10: Arithmetic Crossover (a value .2)

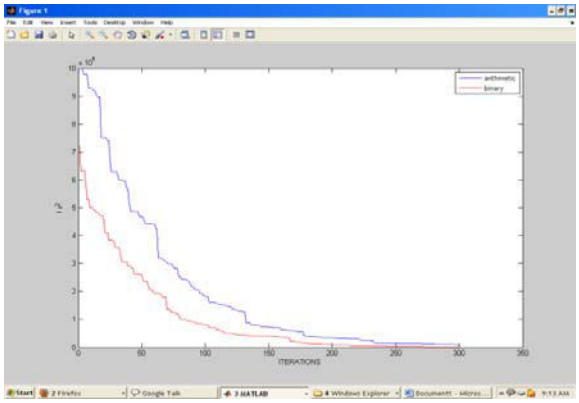


Figure 11: Arithmetic Crossover (a value .5)

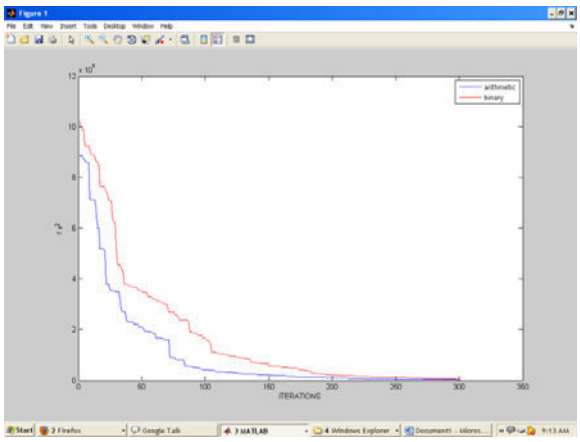


Figure 12: Arithmetic Crossover (a value .8)

From these results we can conclude that, as the value of alpha(a) increases in case of Arithmetic Crossver the convergence speed is increased and it will converge more fastly to 0.

IV CONCLUSION

In this present work we have perform the optimization on DeJong function 1 and 6 by using Genetics. The optimization comparison is also performed by implementing different crossover and selection operators. We observed that the multipoint crossover gives the better optimization ratio instead of Single point crossover. In same way to get the fast optimization results the Arithmatic crossover is better choice instead of Single point crossover. The crossover also depends on the levels also. As we increase the level the convergence ratio is also increased.

REFERENCES

[1] B. Bhanu, S. Lee, and J. Ming. Self-optimizing image segmentation system using a genetic algorithm. In R. K. Belew and L. B. Booker, editors, Proceedings of the Fourth International Conference on Genetic Algorithms and Their Applications, pages 362-369, San Mateo, CA, AJuly 1991.

[2] E. Bagheri, H. Deldari, 2006. "Dejong Function Optimization by means of a Parallel Approach to Fuzzified Genetic Algorithm" Proceedings of the 11th IEEE Symposium on Computers and Communications (ISCC'06) 0-7695-2588-1/06 2006 IEEE .

[3] J. Dalton, "Genetic Algorithms" Newcastle Engineering Design Centre, http://www. edc.ncl.ac.uk. 2007.

[4] J. H. Holland, "Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence" MIT Press, Cambridge, MA, USA, 1992

[5] K. Dejong, "an analysis of the behavior of a class of genetic adaptive systems" PhD thesis, University of Michigan, 1975

[6] L. Randy, Haupt S., Practical Genetic Algorithms, 2003, Wiley-IEEE Publication.

[7] N. A. AL-Madi, A. T. Khader, "A SOCIAL-BASED MODEL FOR GENETIC ALGORITHMS". Proceedings of the third International Conference on Information Technology (ICIT), AL-Zaytoonah Univ., Amman, Jordan, 2007.

[8] S., N. Sivanandan, S. N. Deepa, Introduction to Genetic Algorithm, Springer-Verlag Berlin Heidelberg, 2008.

[9] T. Back. Evolutionary Algorithms in Theory and Practice, Oxford, New York, 1996.

[10] T. Back. "Evolutionary Algorithms in theory and practice Evolution Strategies, Evolutionary Programming, Genetic Algorithms". Accessed 2008.

[11] www.doc.ic.ac.uk/~nd/surprise_96/journal/ vol1/hmw/ article1.html

[12] www.geatbx.com/docu/fcnindex.html #P74_1604

[13] Y. Liao, and C.T. Sun, "An Educational Genetic Algorithms Learning Tool" ©2001 IEEE

[14] Y.Zheng, S. Kiyooka,"Genetic Algorithm Applications". www.me.uvic.ca/~zdong/ourses/ mech620/GA_App.PDF/. 1999

◈◈◈

Load Balancing Protocol for Energy Accomplished Routing in WSN

Mana Vivekanand & Suma Reddy

ISE Dept., The oxford college of Engineering, Bangalore,India

Abstract **–** Wireless sensor networks (WSNs) have been considered as a promising method for reliably monitoring both civil and military environments under hazardous or dangerous conditions. Due to such environments ,the power supplies for sensors in the network are not usually rechargeable or replaceable. Therefore, the energy efficiency is critical for the lifetime and cost of WSN. Numerous mechanisms have been proposed to reduce the impact of communication protocols on the overall energy dissipation of WSN and communicating it with other nodes, moving on to the sink via transceiver. Efficiency of protocol can only be beneficial if the network is alive otherwise what to do for the novel ideas with the dead network. In this paper, Our proposed cluster based routing algorithm has exploited threshold level based load balancing and role transfer techniques along with multi-assistant cluster heads to cope with the aforementioned power hungry issues. Combination of multihop and direct routing has improved our protocol energy utilization.

Keywords - *Load balancing , sensor network, Energy efficient Routing, Cluster Head ,Assistant cluster Head, Threshold Level Exploitation.*

I. INTRODUCTION

Wireless sensor network (WSN) consists of a certain number of smart sensors which form a multi-hop Ad Hoc network by radio communications in sensor field. It aims to apperceive in collaborative mode, gather, deal with and send information to observer in network areas. Sensor, sensing object and observer form the three factors in WSN [1]. WSN protocol stack contains physical layer, data link layer, network layer, transport layer and application layer.. This idiosyncratic technology has its application in Glacier monitoring [2], volcano monitoring and tunnel monitoring and rescue, sniper localization [3], ocean water and bed monitoring, rescue of avalanche victims [4], tracking vehicles, wildlife monitoring [5], cattle herding, vital sign monitoring [6] and cold chain monitoring [7]

Apart from all these inseparable involvement, less computing power, stringent constraint energy and limited bandwidth circumscribed WSN's application as well as hiring the existing protocols from its ancestor: Adhoc and wireless technology. So an replaced protocol having opinion of above mentioned limitations is highly appreciated in WSN.

There are also other challenges that influence the design of routing protocol: deployment strategy, deployment architecture and data reporting models are among those important parameters. Regarding deployment strategies, Flat, hierarchical, and location based are three main sensor nodes' deployment architectures. We use Hierarchical-based routing in our paper.

Hierarchical-based routing aims at clustering e.g. LEACH (Low-Energy Adaptive Clustering Hierarchy) [8],. Figure 1 shows the cluster based and layered based node deployment architectures.

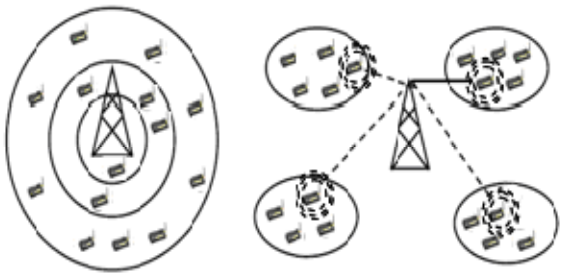


Fig. 2: Cluster based architecture (Right),Layered Based Architecture(Left).

In deployment architecture, clustering is more useful method to obtain the better results. In clustered structure, CH performs more responsibilities than any other cluster nodes. Clustering has more benefits that include balancing of load, energy consumption is less, resources can be reused and network life time can be increased. By knowing these factors, our protocol which we will be designing is based on clustering technique with multi-assistant cluster heads (ACH) which not only helps in reducing the energy consumption but also balancing the load, hence improves the network life time. More-over, ACH working in inter-cluster and intra-cluster routing also helps in fault tolerance.

In WSN, node deployment and data reporting models are application specific so as the designed protocols. The deployment can be either deterministic or stochastic. Various types of models are used for data reporting . It can be time-driven (continuous), event-driven (discrete), query-driven, or hybrid. In time-driven delivery model periodic data monitoring is done. In event-driven and query-driven models, sensor nodes response immediately when drastic changes occur in sensed attribute due to some abnormal condition or a query is generated by the BS. In our defined scenario, we assumed deterministic deployed and event driven model. Routing has been a field of great interest for the researchers resulting in large no of routing techniques empowering one or the other aspect of routing parameters and network scenarios. When BS is accessed directly from the sensor node is called one-hop model. This type of can be better used in small networks which dont have scalability problems. On the other hand accessing the BS in a multihop fashion of communication via transit nodes is categorized as multihop model. In Cluster based model whole network is partitioned into clusters. Each cluster has a cluster head (CH) that acquires the sensed data from cluster member nodes, aggregates and forwards it to other cluster heads or to the base station. Figure 3 shows two scenarios of multihop and a possibility of direct routing. LEACH (Low-Energy Adaptive Clustering Hierarchy) by Heinzelman et al. has introduced a CH selection and Rotation technique. They have proposed two layered architecture, one-layer for intra-cluster communication and other for inter-cluster communication. Through empirical results, it has been proved that the network life time increases by the rotation of cluster head as well as better management of load-balancing issue.

In [9] Ma et al. has proposed a dynamic positioning technique for designating the cluster head. Results show the better location of CH comes up with the balanced network and also prolonging the network lifetime.

Irfan et al. [10] has introduced the idea of temporary cluster heads which performs better as compared to LEACH and enhanced version of LEACH in load balancing and efficient energy utilization. EECR, TEEN, APTEEN and PEGASIS are also presents cluster based routing solutions. Rest of the paper is organized as follows. In section two, proposed solution with the working of TLPER is discussed in detail. Simulation and result discussion is in third section. Concluding remarks and Acknowledgement ends up the paper.

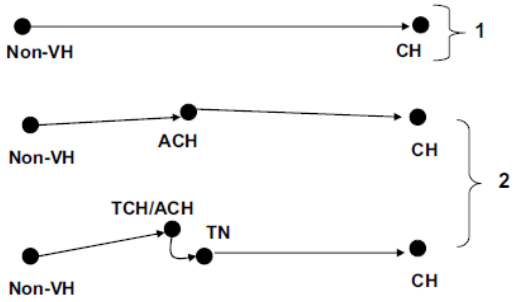


Fig. 3 : 1 – Direct Hop, 2 - Multihop

II. PROPOSED SOLUTION

In WSN, routing is really a challenging issue. A sensible and foresightedly planned routing protocol can have a vital role to add life to the network. Our proposed solution specifically target the scalable, fault tolerant and load balancing feature by synergastic mating of multihop and direct routing, energy efficient, load balancing and role transfer threshold and multi Assistant Cluster Heads (ACH). Homogenous sensor nodes, same initial energy level, deterministic deployment, centrally preselected cluster heads (CH) and preselected ACHs are the characteristics of our assumed scenario.

A. Cluster formation

Exploiting the self organizing capability of sensor nodes, each node may know its neighboring nodes as well as its Vicinity Head (VH). The term Vicinity Head applies to CH and ACH. As the deployment is deterministic, so at initial stage, selection of CH and ACH is on hand. Due to the homogenous nature of nodes, the node having the more neighbors is designated as Vicinity Head (VH). Cluster Head is the main head of the vicinity of cluster while ACH is the assisting head of sub vicinity of the same cluster. In initial and later on, rotation of VHs (CH and ACH) communicate their designation to neighboring nodes. Each node attaches itself to the VH on the basis of received signal strength (RSSI). If a node receives invitation from more than one Vicinity Heads then the following criteria is followed:

$$Sweight_i > Sweight_j \qquad \forall_j$$

Where *Sweight* is weight or strength of received signal of the invitee VH.

If $Sweight_i = Sweight_j \qquad \forall_j$

then the selection is on the basis of

$$Eweight_i > Eweight_j \qquad \forall_j$$

Where *Eweight* is the weight of energy level of invitee

VH. If $Eweight_i > Eweight_j \quad \forall_j$

Then a random selection is made.

More-over, for CH_i ,

We may have $ACH_{i1}, ACH_{i2}, ACH_{i3}, ACH_{i4}$

Communication and processing factors deplete the node's energy gradually which emerges the dynamicity of network with respect to the rotation esp. of vicinity heads (CH, ACH). Here we have introduced parallel rotational strategy of vicinity heads to cope with such network dynamicity aspect

B. Parallel Rotational Strategy of Vicinity Heads

One of main energy consumption factor is rotation of vicinity heads. Finding the next best replacement of the current vicinity head and then propagating its designation to the neighboring nodes, not only add its role in lessening the network life time but also introduce more end-to-end delay (E2E delay). In this paper, We have introduced parallel rotation strategy that if not maximize but in a little extend contribute in adding more life to the network and lessening E2E delay. Figure 4 and Figure 7 demonstrate the load balancing support to the network by its differential features which ultimately comes up with increasing the network life time as well as fault tolerance to the network.

i. Threshold Level Exploitation

Setting up checks on working levels is exploitated in parallel rotational strategy of vicinity heads (VH). The upper level check providing the load balancing capability to the network is termed as Load Balancing Threshold (LBT). The lower level check assist parallel rotation of VHs and Cluster Heads (CHs) is named as Role Transfer Threshold (RTT). Due to the deterministic deployment strategy and self organizing capability of WS nodes, each node may know its vicinity head.

ii. ACH Rotation

On reaching the LBT, ACH establishes a communication link with the most energy carrying node and designate it as a transit node for communication with forwarding Node (i.e CH of same cluster or ACH of neighboring Cluster) or destination node. Now on occurance of LBT, ACH keeps on communicating with the forwarding nodes/destination node via transit node until RTT. On reaching RTT, ACH then broadcast an updating status message of designating TN as a ACH. This saves the network partitioning issue. More-over it not only maximizes the network usability but also the energy of X-ACH will remain to that extend to atleast participate in communication and sensing process. Figure 8 shows this complete process of parallel rotation of Vicinity Heads. Node's uninformed and suddent death is also a possibility that is not considered here.

iii. CH Rotation

More or less same strategy is adopted by CH as of ACH for the rotation of its designation. Let CH_i is ith cluster head and N_{jk} are neighboring nodes of ith cluster head. The node which fulfills the condition

$$\lfloor EN_k \rfloor \leq \lfloor EN_j \rfloor \qquad \forall_k$$

will be designated as the cluster head in subsequent turns. But how to come up with the knowledge of maximum energy carrying node in the neighbor of CH. Threshold based Updated Info Communication (TUIC) strategy is proposed to minimize the beacon exchange and hence saving constraint factor of energy.

iv. How TUIC Strategy works?

On the basis of Figure 6, we can have the idea of tentative minimization in energy of neigboring nodes with the ratio of vicinity head. Here ratio between energy consumption of Non-Vicinity Head (NVH) and Assistant Cluster Head (VH) is 1:5 and between ACH and Cluster Head (CH) is about 1:2.3. It would be a better strategy to some what applying the unsupervised machine learning to train the network for the said purpose. For the safe calculation and to prevent from re-requesting for the updated info energy levels, CH request for the energy levels info from the neighboring nodes that fulfill the threshold energy level criteria (estimated by the prior training of network or from above mentioned calculation graph). So only those minimu nodes will reply which have this maximum energy Level.

C. Forwarding Node Selection in Inter-Cluster and Intra-Cluster Routing

The introduced strategy of ACH and TN assists in energy efficient cluster based routing along with load balancing feature resulting in better network utilization and its life. Based on the receiving node there are three possibilities for ACH in selection of forwarding node/Destination:

1-Base Station

2-cluster Head of its cluster

3-ACH of Neighboring cluster

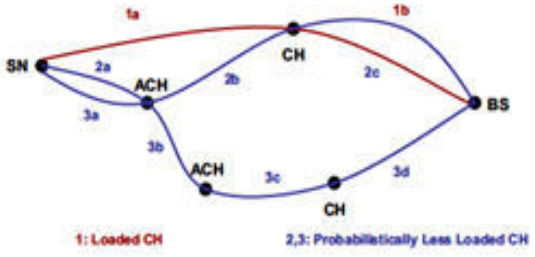


Fig. 4 : Comparison of Load Balancing Support in loaded CH (i) and loaded CH with Assistant Cluster Head

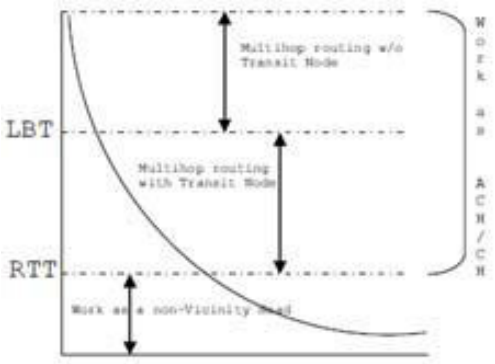


Fig. 5: Threshold Level Exploitation

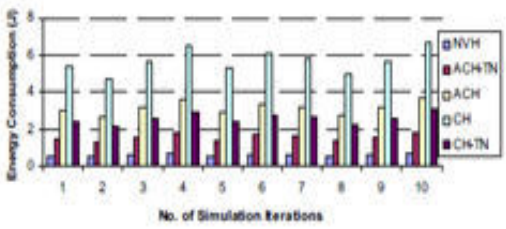


Fig. 6: Energy Consumption Ratio Finding (Graph)

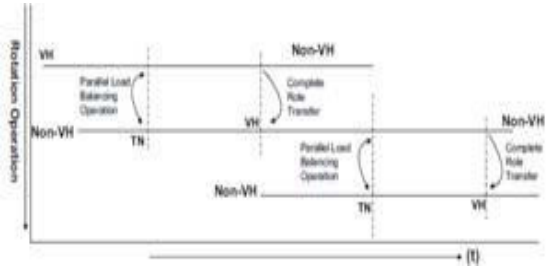


Fig. 7: Load Balancing Support (Parallel Rotation Strategy)

and two possiblities for CH in selection of forwarding node/Destination:

1-Base Station

2-ACH of its Cluster

Above mentioned routing strategy is in the normal routing process. When the LBT reaches, the routing strategy would be different then. On occurance of any event, node senses the environmental physical quantity and forward the packet to its vicinity head, ACH. Assistant Cluster Head then sends the packet to the cluster head, it then sends the packet to the ACH in the direction of destination. This ACH transmits the received packet to the ACH of neighoring Cluster. On reaching the LBT, the communication between ACH to CH and vice versa is happened via transit node and on reaching RTT, transit node take over the control of Vicinity Head and itself act as a Vicinity Head. X-ACH and X-CH then function as a normal node. One issue that can be apparently seen in LBPER is addressed as follows:

During rotation of VHs, the possibility of maximum nodes utilization is there but at the same time, the centralized management of VH's may snail from left to right and from top to bottom and vice versa. This makes boundary area nodes difficult to access the vacinity heads. This issue may arise with the boundary cluster nodes as the inner cluster's central positions creep along with relavent nodes. So, to cope with the situation arises with the boundary cluster nodes, here we introduce Pioneer Old ship Exploitation (POS) technique in conjunction with TUIC strategy. The first time selected ACH will take over the charge and elect fresh ACH among the neighbors which satisfy the condition of

III. RESULTS AND DISCUSSION

We have simulated our proposed algorithm, LBPER by making a java simulator in netbeans i.e JAVASIM, to calculate its performance. Results have been compiled and compared in comparison with Low-Energy Adaptive Clustering Hierarchy (LEACH). Figure 6, Figure 8, Figure 9, Figure 10, Figure 11, Figure 12 and Figure 13 illustrate some of the initial results drawn from simulation of our proposed Load Balancing Protocol for Energy Accomplished Routing. For the simulation an area of 80x80 meter is considered with node density of 100. 10m Node to Node distance, 20 Joule Initial energy of node and the MAC type is SMAC, are the simulation parameters. Network life time is calculated on the death of first node.

Performance Metrics: The performance metrics which we will be calculating .

- Energy Consumption
  - Per Node

- Cluster Head
- Assistant Cluster Head

- Network Utilization
- Load Balancing Effect on Energy Consumption

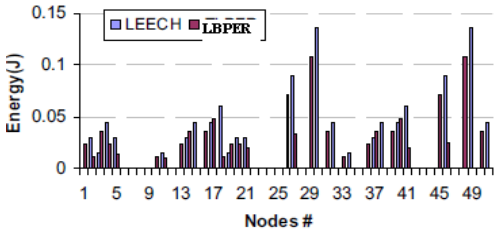


Fig. 8: Energy Consumption per Node in one Simulation

Figure 8 and Figure 13 show the energy consumption and residual energy of LEACH and LBPER on per node basis. Figure 9 dictates the lesser total energy consumption of LBPER compared to LEACH. Energy consumption of former is higher especially that of Cluster Head as compared to later because it has to bear all the load arrived from communicating nodes. But at the same time, the combined energy consumption of TN and CH in LBPER is also to be considered in comparison to energy consumption of CH in LEACH. It is intuited from Figure 12 that the proposed algorithm also perform better if evaluated on the said criteria. More-over, total energy of proposed algorithm is comparatively lower to that of competitive algorithm in a typical simulation. On the other hand, outperform working of LBPER is also apparent from Figure 11 regarding total packets entertained by CH-TN + CH in different number of Simulation Iterations. Hence, load balancing has its important effect on overall energy consumption is the concluding statement derived from the results in Figure 10 and Figure 12. Network utilization is another yard stick for efficiency of a routing protocol regarding its load balancing and energy consumption. Figure 11 demonstrates network utilization chart based on energy consumption per node and it is apparent that LBPER perform better than LEACH in this regards

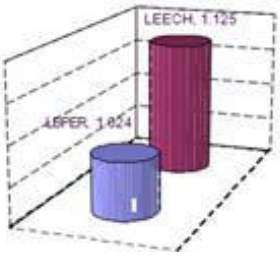


Fig 9: Total Energy Consumption in one typical round

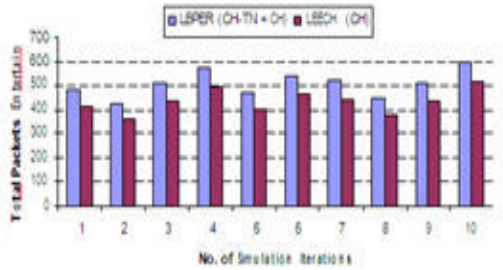


Fig. 10: Effect of Load-Balancing on Total Number of Packets Entertained by VH

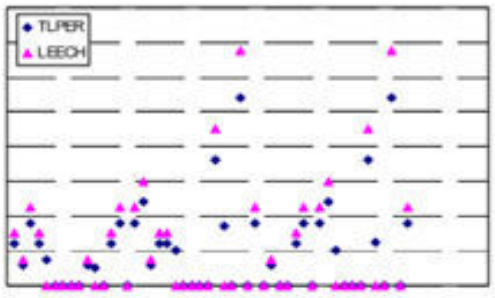


Fig. 11:Effect of Load-Balancing on Total Energy Consumption in a different aspect.

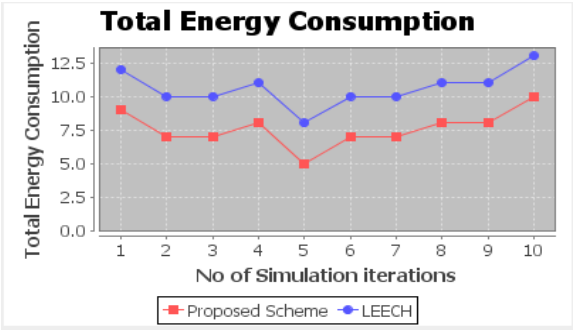


Fig. 12: Effect of Load-Balancing on Total Energy Consumption ina different aspect

IV. CONCLUSION

The proposed cluster based routing algorithm has been considered as one of the effective communication protocols in WSN and has exploited threshold level based load balancing and role transfer techniques along with multi-assistant cluster heads to cope with the power hungry issues. Both multihop and direct routing are embedded in LBPER. It has been intuited from the results that LBPER gives better network utilization and lesser per node energy consumption resulting in prolonging the network life time as compared to LEACH. Such improvements consequently assures the improvement of overall WSN lifetime.

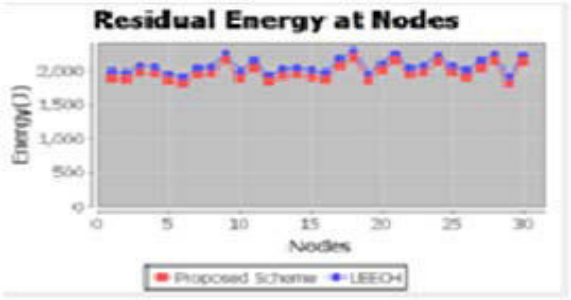


Fig. 13: Residual Energy per Node in one Simulation

REFERENCES

[1] Sun Limin, Li Jianzhong, Chen Yu, Wireless Sensor Networks, Tsinghua publishing company, Beijing, 2005.

[2] R. O. J. K. H. K. Martinez and J. Stefanov,"Glacsweb: A sensor web for glaciers." Berlin Germany: EWSN 2004, January 2004

[3] G. Simon and M. Maroti, "Sensor network-based countersniper system." Baltimore, USA: SenSys, November 2004.

[4] A. S. F. Michahelles, P. Matter and B. Sciele, "Applying wearable sensors to avalanche rescue," no. 27(6), 2003.

[5] P. Juang, H. Oki and D. Rubenstein, "Energyefficient computing for wildlife tracking: Design tradeoffs and early experiences with zebranet," ASPLOS X. San Jose, USA: IEEE, October 2002.

[6] H. Baldus and G. Muesch, "Reliable setup of body sensor networks."Berlin, Germany: EWSN2004, January 2004.

[7] R. Riem-Vis, "Cold chain management using an ultra low power wireless sensor network." Boston, USA: WAMES 2004, June 2004.

[8] Wendi Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan, "Energy-Efficient Communication Protocols for Wireless Microsensor Networks", Proc. Hawaaian Int'l Conf. on Systems Science, January 2000

[9] M. Ma and Y. Yang, "Clustering and load balancing in hybrid sensor networks with mobile cluster heads," ACM Third International Conference in Quality of Service in Heterogeneous Wired/Wireless Networks. USA: ACM, 2006.

[10] I. Nauman, A. Irfan, "MCLB: Multihop Clustering Algorithm for Load Balancing In Wireless Sensor Network", I. J. of simulation Vol. 8 No. 1, 2009.

◈◈◈

Gender Classification based on Hand Geometry

Shubhra Shivani & Suneeta Agarwal

Computer Science and Engineering Dept., Motilal Nehru National Institute of Technology, Allahabad, India

Abstract – Hand Geometry is a biometric tool that is widely used for the verification of subjects enrolled under a system. It covers features of hand such as length and width of palm and fingers. In this paper effort has been made to utilize the geometrical measurements of the hand in order to achieve gender classification. Inspired by the anthropological studies, the hand dimensions were extracted from the image of the hand. Various ratios of the linear measurements of hand constitute the feature vector representing the hand of a subject. The feature vector was then classified as of male or female using Support Vector Machine with linear kernel, achieving classification success rate of 95.6%. Variance analysis was used to determine the contribution of each component of the feature vector in correct classification. The analysis of the results revealed the effect of age on hand geometry with respect to gender.

Keywords - *Hand geometry, gender classification, SVM, ANOVA.*

I. INTRODUCTION

The term biometrics is derived from the Greek words bio (life) and metric (to measure). It is the science of determining the identity of a person based on his or her physical, chemical or behavioral features. Biometrics is widely used for security as it serves as an entity which can provide authentication and identification. It is used as a password which can neither be stolen nor forgotten. Various biometric tools include iris, hand geometry, palm print, palm vein, gait, speech, etc. Among these, hand geometry has an advantage of being non invasive in nature and lack of connection with criminal investigations, leading to high user acceptance rate [1]. But due to lack of uniqueness, its use is more in verification than in identification.

Gender classification is an important social issue. This classification problem has been dealt with using speech, face and gait as the basis of classification, but not much work has been done on gender classification using hand geometry. Gender can also act as a soft biometric trait. Soft biometric traits like gender, height, weight, age, and ethnicity factor to strengthen the identification system by providing support to the decision. Although they lack the distinctiveness and permanence to identify an individual uniquely and reliably, they provide some evidence about the user identity that could be beneficial.

Anthropological studies reveal the influence of gender on the geometry of the hand. Various inter digit ratios are found to be different in both the genders. Due to high user acceptance and low cost setup, hand geometry must be studied in perspective of gender classification problem.

In this work the aim was to use hand geometry for gender classification, making use of various image processing tools. Various features of the hand geometry are being used for identification, such as, length and width of the fingers, length and width of the palm, & angles between roots of fingers. Out of these only length of fingers (except the thumb) and length and width of palm were used to accomplish gender classification. Two separate methods were used for feature extraction and the results were analyzed to test their robustness against stretch ability of the hand. The superiority of one method over other was determined on the basis of classification rate achieved from the two methods. By analyzing the components of the feature vector an attempt has been made to reduce its size as much as possible.

The following sections discuss the approach in detail. Section II gives the overview of the prior work done related to the topic. Section III describes the two feature extraction methods used. Section IV briefs about the statistical methods used in analysis, such as SVM and ANOVA. Section V discusses the results of the experiment and Section VI concludes the paper.

II. PRIOR WORK

Extracting gender information from human hands has been studied since long time in the fields such as anthropology and psychology. Anthropological studies reveal that following three relations between the lengths

of index finger and ring finger can be observed in a human hand: i) the index finger is relatively shorter than the ring finger, ii) the index finger is equal in length to the ring finger; and iii) the index finger is relatively longer than the ring finger. All three relations occur in both sexes, however, a relatively long index finger is more frequent as reported in [2], [3].

McFadden and Shubel [4] studied all six possible ratios between the index, middle, ring and little fingers in both genders. Their results indicate that the largest gender difference was depicted by the relative lengths of the index and ring fingers. Agnihotri et al. [5] found that the average hand index in males was more than 0.44 while the average index in females was less than 0.44, where hand index is the ratio of hand breadth over hand length. Based on these results, they suggested that this value can be used as a threshold for determining gender by hand dimensions. Amayeh et al. [6] have classified hand images on basis of gender using Zernike and Fourier descriptors.

However, till date there has been no research providing evidence as to which features of hand geometry determine gender robustly. Therefore, it can be concluded that gender cannot be determined using a single feature of the hand, rather it requires the combination of multiple features.

III. FEATURE EXTRACTION

The image was first binarised (Fig. 2). As all images were taken using the same black background, a constant threshold of 0.02 was used. Then lengths were calculated using two different methods. The various ratios among these lengths formed the feature vector containing 16 components (Fig. 1). These components are:

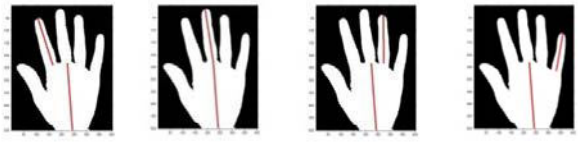
i) 4 ratios of each finger to width of the palm

ii) 4 ratios of each finger to length of the palm

iii) 1 ratio of palm width and length

iv) 1 ratio of palm width and hand length (i.e. length of palm + length of middle finger)

v) 3+2+1 ratios of fingers with each other



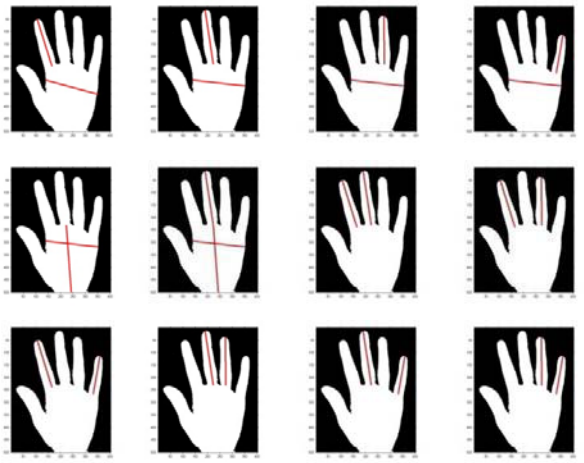
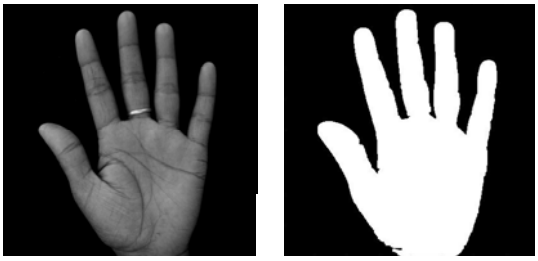


Fig. 1. Ratios used to construct the feature vector

As all the components are in form of ratios, there was no requirement of determining the real length of the hand. Here the length of the thumb was not considered due to the following reason. The image acquisition environment was unconstrained. Also the stretch ability and flexibility of the thumb is much greater than the other fingers. This may result in curved profile of the thumb causing error in measurements.



(a) Original image (b) Binary Image

Fig. 2. Original hand image and its binary image

Method I

In this method the lengths were measured from the intact hand image using the tips and roots of the fingers as reference points for calculations [7]. The steps followed were as follows:

1) *Boundary Extraction:* Boundary tracing function was used to obtain an array of boundary pixels. These pixels determine the silhouette of the hand.

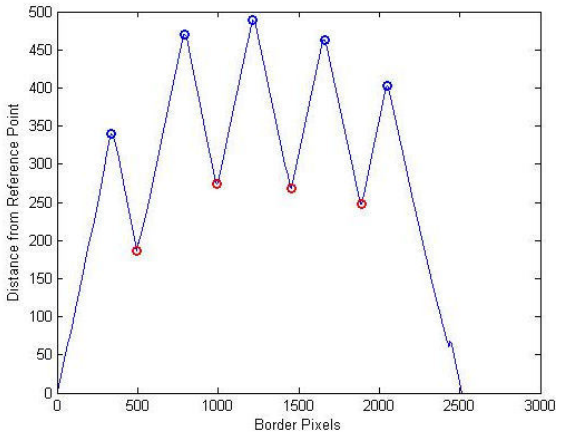


Fig. 3. Euclidean distance graph

2) *Detection of tip and root of fingers :* A point in the bottom row of the hand silhouette was selected as the reference point. Euclidean distance of all the boundary points with this reference point was calculated (Fig. 3). The Euclidean distance vector was then traversed to detect all the local maxima, which correspond to the tips of the fingers labeled as points 2, 5, 7, 9 and 11 as shown in Fig. 4. Similarly local minima were detected which correspond to the roots of the fingers, giving points 3, 6, 8 and 20. Three more points 1, 4 and 12 were required to be calculated to represent the root points of thumb, index finger and little finger respectively. This was done by tracing the hand boundary and selecting the points which were at same distance from the fingertips corresponding to the finger for which roots are to be determined. Therefore 1, 4 and 12 were determined using 3, 6 and 20 as reference points.

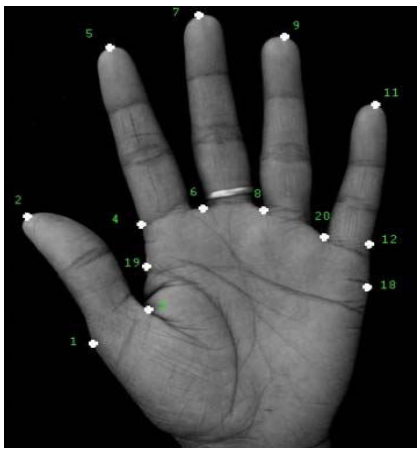
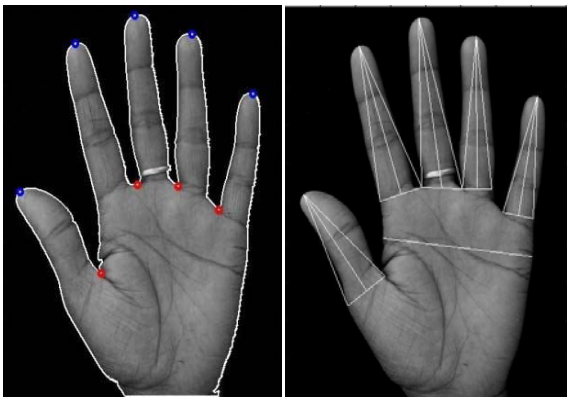


Fig. 4. Reference points on hand

3) *Calculation of lengths:* Midpoint of each line joining the successive roots of the fingers was calculated. The length of the finger was measured as the Euclidean distance between the midpoint and the corresponding finger tip. Width of the palm was measured as the distance between points 18 and 19. Midpoint between 3 and 4 was marked as 19. Point 18 is at distance 'd' from point 12, where 'd' is half distance between 3 and 4. Length of the palm was calculated as the distance of point 15 (which is the midpoint of line joining 8 and 6) from the last row of the image. Total length of hand was calculated as the sum of length of palm and length of the middle finger.



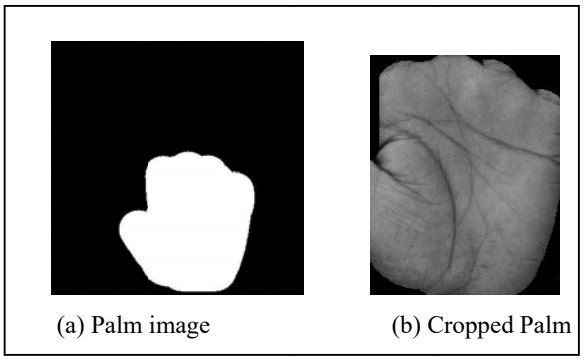
(a) Finger tips and roots (b) Length measurements

Fig. 5. Method I

Method II

In this method the components of hand, which are fingers, thumb and palm were extracted separately using morphological operators and then measured to get the lengths [6], [8]. The steps were as follows:

1) *Palm Extraction :* Using a disc structuring element, opening was performed on the binary image of the hand. The radius of the disk was determined to be 30 pixels, as this value was found to be greater than the widest finger in the database. This resulted in removal of the fingers and thumb giving a binary image containing only the palm. Processing was required as the extracted palm contained some portion of the thumb, giving rise to protruding regions. Therefore, the binary palm image was eroded using a vertical line structuring element equal to two third of the height of the image. This resulted in removal of the protruding portions in horizontal direction. Next the image was cropped in the form of a bounding box of the palm (Fig. 6(b)).



(a) Palm image (b) Cropped Palm

Fig. 6. Palm Extraction

2) *Fingers Extraction:* Subtracting the binary image containing only palm from the original image, a new image containing only the fingers was obtained (Fig. 7). This image was then used to extract the fingers separately using the concept of connected components. The image containing fingers can be segregated into five regions containing a finger each. After extraction of the fingers they were rotated to align their mid axis vertically. This was done by approximating the finger as an ellipse and locating its major axis to represent the mid axis of the finger. In order to remove the pointed and jutting edges created by the opening operation, a square structuring element of size three was used to erode the image. Then the bounding box for fingers was selected to get the final separated finger images (Fig. 8).

3) *Calculation of lengths:* Width and length of palm was calculated as the maximum width and the mean length of the palm. Length of the fingers was calculated as length of the bounding box.

Fig. 7. Hand image without palm

Fig. 8. Extracted fingers

IV. CLASSIFICATION AND ANALYSIS

The classification was done using Support Vector Machine (SVM). The results were also analyzed using Analysis of Variance (ANOVA) to determine the contribution of various components of the feature vector.

A. Support Vector Machine

SVM aims to create a model based on the training data given to it. Using this model where each data is projected in a multidimensional space, a plane was constructed to separate the points belonging to two different classes. When new test data is given, the SVM uses this model to calculate the distance of test data point from the separating plane and thus predicting its class. SVM has an advantage that the complexity does not depend upon the dimensionality of the feature space.

Given the set of training data, each belonging to one of the two classes, the SVM constructs an optimal hyper plane that maximizes the margin of separation between the two classes. In other words, the plane is at maximal distance from the closest members of both the classes. The subset of the data points, which determine the maximum-margin separator are called "support vectors". It will be computationally useful if only a small fraction of the data points are support vectors, because the support vectors are used to decide which side of the separator a test case lies.

B. Analysis of Variance (ANOVA)

Variance analysis of the data distribution determines the ratio of inter class variance to intra class variance with respect to each variable when the number of groups and their members are predefined. For a variable to have discriminative value with respect to the given classes, its inter class variance must be high and intra class variance must be low. Also it must have low probability value on the statistical distribution curve.

V. EXPERIMENTAL RESULTS

A. Experimental Setup

All the algorithms have been implemented using MATLAB on a PC with Intel Core2 duo 2.0 GHz processor and 2.0 GB of main memory, running WindowsXP.

The database contains images of both left and right hands of 165 subjects, taken against a black background. The subjects belong to two different age groups. 150 subjects are aged 12-75 years while remaining aged 5-11 years. Out of the 150, 68 are female and 82 are male. Out of the remaining 15 (having age below 12), 8 are female and 7 are male (Table 1).

The images have been captured using unconstrained and contact-free setup, which may have resulted in variations in stretching of the hands. Support Vector Machine (SVM) using linear kernel has been used for classification. Two SVM's were trained: first with the feature vectors constructed using method I, and second with feature vectors constructed using method II.

TABLE I

AGE AND SEX DISTRIBUTION IN DATABASE

Database	Age group	Female	Male
Database I	5-11 years	8	7
Database II	12-75 years	68	82
I and II	5-75 years	76	89

B. Classification Results

Initially, the experiments were done using the whole database. For the data generated using Method I, 60 feature vectors of left hand (chosen randomly out of 165) were used as training data set. The remaining 105 feature vectors formed the test data set and were classified with 85.7% success rate. For the data obtained using Method II, a success rate of 90.47% was obtained using same number of vectors for training and testing as taken in former method. To test the accuracy of both the methods used for length measurements, the lengths of the hand components were also measured manually using the image tool of MATLAB (Fig. 9). Using this data 91.42% success rate was obtained (Table 2). Similar results were obtained for right hand images.

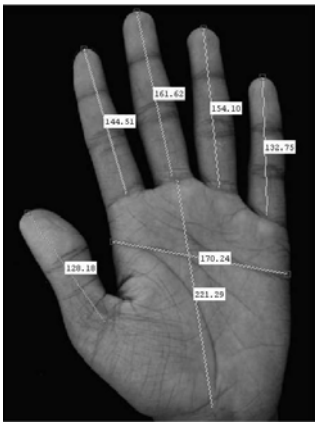


Fig. 9. Manual Measurements

TABLE II

CLASSIFICATION RESULTS FOR 5-75 YEARS

Method	Right Hand Rate	Left Hand Rate
Method I	84.76%	85.70%
Method II	90.47%	91.42%

To identify the source of error in classification, the results were analyzed taking into consideration the age of the subjects. It was observed that more than 60% of the misclassified hands were aged below 12 years of age. Also for this age group all the females were classified as males. The database was then refined to only 150 images for each hand, aging 12-75 years. Now 60 feature vectors for left hand images were used as training data set and the remaining 90 formed the test data set. The classification rates increased to 91.1%, 95.6% and 96.6% for data obtained from Method I, Method II and manual measurements, respectively (Table 3). Similar results were obtained for right hand images.

TABLE III

CLASSIFICATION RESULTS FOR 12-75 YEARS

Method	Right Hand Rate	Left Hand Rate
Method I	91.10%	90.00%
Method II	95.60%	95.60%

The results obtained help to deduce the following:

- The second method provides more robustness against unconstrained hand positions. As using this method the hand components were extracted separately, it gives more accurate measurements.
- On the basis of classification results using manual measurements, we can conclude that the error in classification is not due to incorrect measurements of lengths. Instead, they are due to lack of uniqueness in hand geometry.
- Age also has influence on the hand shape and geometry as classification results are different for different age groups.
- The dexterity of the hands with reference to geometry is proved as similar results were obtained for both the hands.

C. Data Analysis

To analyze the contribution of each component of the feature vector separately, Analysis of Variance (ANOVA) was used to analyze the feature vectors obtained from Database II using Method II (Fig. 10). The analysis revealed that eight components of the feature vector were more significant than others. The significant components in decreasing order of significance are:

- ratio of ring finger to index finger
- ratio of ring finger to middle finger
- ratio of little finger to index finger
- ratio of middle finger to index finger
- ratio of index finger to width of the palm
- ratio of little finger to ring finger

- ratio of index finger to length of the palm
- ratio of ring finger to length of the palm

This shows that in the case of gender classification, the ratios between the fingers are more significant than the hand index (i.e. ratio of hand width to length).

Also, gender influence is not only on the ratio of index finger and ring finger, but also on ratios between lengths of other fingers.

Using only these eight variables, which were found to be most significant, the classification rate remained unchanged at 95.6%. Thus the size of feature vector has been reduced to half its length.

Variable	SS Effect	df Effect	MS Effect	SS Error	df Error	MS Error	F	p
Var10	0.021007	1	0.021007	0.302674	148	0.002045	10.27191	0.001655
Var11	0.129144	1	0.129144	0.273026	148	0.001845	70.00552	0.000000
Var12	0.035184	1	0.035184	0.361835	148	0.002445	14.39099	0.000216
Var13	0.041483	1	0.041483	0.118188	148	0.000799	51.94721	0.000000
Var14	0.005110	1	0.005110	0.258813	148	0.001749	2.92225	0.089462
Var15	0.009267	1	0.009267	0.219918	148	0.001486	6.23632	0.013610
Var16	0.030335	1	0.030335	0.699146	148	0.004724	6.42151	0.012315
Var17	0.007014	1	0.007014	0.672959	148	0.004547	1.54256	0.216202
Var18	0.010751	1	0.010751	0.543986	148	0.003676	2.92506	0.089310
Var19	0.000005	1	0.000005	0.471208	148	0.003184	0.00143	0.969898
Var20	0.017150	1	0.017150	0.407179	148	0.002751	6.23361	0.013630
Var21	0.002757	1	0.002757	0.430035	148	0.002906	0.94880	0.331615
Var22	0.009634	1	0.009634	0.355797	148	0.002404	4.00741	0.047128
Var23	0.000149	1	0.000149	0.265403	148	0.001793	0.08321	0.773402
Var24	0.000741	1	0.000741	0.279489	148	0.001888	0.39254	0.531932
Var25	0.000969	1	0.000969	0.094611	148	0.000639	1.51635	0.220126

Fig. 10. Analysis of Variance on data from Method II

VI. CONCLUSION

Influence of gender on hand geometry has been determined. To provide robustness against unconstrained image acquisition, various ratios of hand dimensions were used instead of absolute lengths. Experimental analysis reveals that the discriminative power of hand geometry varies with age of the subject and the inter-digit ratios play important role in gender based discrimination of hands. The results show that with sufficient and reliable training data, classification of the hand as of male or female can be done correctly up to 95.6%.

For future work, the measurement of thumb length can also be incorporated in the feature vector in order to analyze the influence of gender on the dimensions of the thumb. This would require the use of measurement methods which are more robust to stretching.

REFERENCES

[1] N. Duta, "A survey of biometric technology based on hand shape," Pattern Recognition, vol. 42, no. 11, pp. 2797–2806, 2009.

[2] W. M. Brown, M. Hines, B. A. Fane, and S. Breedlove, "Masculinized finger length patterns in human males and females with congenital adrenal hyperplasia," Hormones and Behavior, vol. 42, no. 4, pp. 380–386, 2002.

[3] R. George, "Human finger types," The Anatomical Record, vol. 46, no. 2, pp. 199–204, 1930.

[4] D. McFadden and E. Shubel, "Relative lengths of fingers and toes in human males and females," Hormones and Behavior, vol. 42, no. 4, pp. 492–500, 2002.

[5] A. K. Agnihotri, B. Purwar, N. Jeebun, and S. Agnihotri, "Determination of sex by hand dimensions," The Internet Journal of Forensic Science, vol. 1, no. 2, 2006.

[6] G. Amayeh, G. Bebis, and M. Nicolescu, "Gender classification from hand shape," in Computer Vision and Pattern Recognition Workshops, 2008 (CVPRW '08). IEEE Computer Society Conference on, June 2008, pp. 1–7.

[7] R. Sanchez-Reillo, C. Sanchez-Avila, and A. Gonzalez-Marcos, "Biometric identification through hand geometry measurements," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 22, no. 10, pp. 1168–1171, Oct 2000.

[8] G. Amayeh, G. Bebis, A. Erol, and M. Nicolescu, "A component based approach to hand verification," in Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on, June 2007, pp. 1–8.

◈◈◈

Trainable Fuzzy Logic Based Image Spam Filtering Using Histogram

Gopal Benakanawari & K. Vinaykumar

Department of Computer Science & Engineering,
National Institute of Technology, Karnataka, Surathkal, Mangalore, India

Abstract **-** Image Spam is a recent variant of spam where the message text of the spam is presented as a picture in an image file to bypass the spam filter. Spam continues to be a headache for administrators and end-users because spammers are constantly trying to stay one step ahead of anti-spam software vendors. In this work we are proposing Trainable Fuzzy logic based Image spam filtering which utilizes external properties of image and histogram(grey, RGB and HSV), and it is a fast method because of not extracting text and analyzing the content of the image. The paper carried out an experiment on a personal dataset and Princeton benchmark dataset for spam images, on Liang spam database for text spam and Corel database for legitimate images with java eclipse platform, and got a high accuracy and recall rate.

Keywords*-componen Image spam; Fuzzy logic; Spam filter, RGB Histogram; HSV Histogram.*

I. INTRODUCTION

Email is most commonly used network application and it has become popular way of communication but many people or companies use spams for malicious purpose by sending large number of messages to the group of users who have not requested these types of messages are called spams. E-mail spam has continued to increase at a very fast rate over the last couple of years. It has become a major threat for business users, network administrators and even normal users [1]. A study in July 1997 reported that spam messages constituted approximately 10% of the incoming messages to a corporate network.

Message Labs [9] stated in its 2006 Annual Security Report that spam activity has increased significantly in 2006 with levels that reach 86.2% of the e-mail traffic. Based on projections of current analysis and trends there is a prediction that by year 2015 spam will exceed 95% of all e-mail traffic [10]. Although these figures might not be accurate enough, what can be concluded is that spam volume is dramatically increasing over years.

Until a few years ago the content of spam emails was only of textual kind and anti spammers introduced spam filter which classifies emails as legitimate and spam successfully then spammers took their tactic to the next level by introducing image spam where the text of spam is embedded in an image. In this case spammers succeeded in bypassing the filter and nowadays image spam ruling the network. Anti-spammers introduced OCR based image spam filtering techniques in which embedded text is extracted from images to filter. Spammers are constantly trying to stay one step ahead from anti-spammers by introducing new tactics. Spammers undermined OCR based image spam filter by making image background wild and noisy.

In this work we are proposing fuzzy logic based image spam filtering which uses external properties of image and histograms.

The rest of the paper is organized as follows: Section II describes existing methods; Section III is the proposed method; Section IV contains experimental results and analysis.

II. RELATED WORK

There are two types of existing styles for image spam filter first one is to extract and analyze the embedded text from image and second one is to use external features of image.

OCR based image spam filtering works based on extracting the text embedded in an attached images, then apply the traditional approaches used in spam filters to analyze emails body text is used [2], which are keyword detection and text categorization techniques. The main might of this approach is the Optical Character Recognition tool which is vulnerable to obfuscation techniques and some randomization techniques like rotate, wavy text etc.

These drawbacks led anti spammers to introduce techniques based on low level external features of

image[3] and later introduced combination of OCR and low level external features of image[4,5,6]. Sven Krasser et.al.,[7] proposed a fast and low-cost feature extraction and classification method for image spams where the features used are width, height, Area, Aspect ratio, file size, compressibility. jong et.al. [8] Proposed fuzzy inference system for spam filtering whereas existing Bayesian based spam filtering needs to be more accurate and vulnerable to Bayesian poisoning. The main drawback of second style filtering techniques is low recall ratio.

III. PROPOSED METHOD

A. System Architecture

Architecture of our spam filter mainly consists of three models explained below ;

Preprocessing: This model Extracts external properties of incoming images namely width, height, Area, Aspect Ratio, File size, Compressibility and Bit depth.

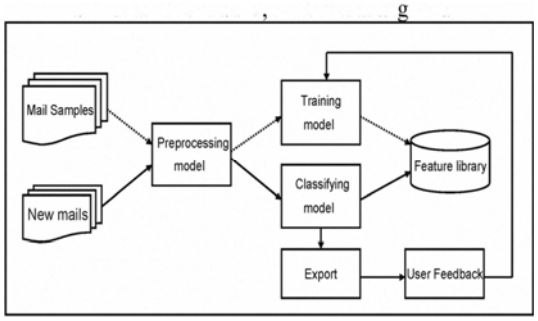


Figure1. Shows System Architecture

1. Training model: This model Maintains two feature libraries *ham* to store external features of legitimate image and *spam* to store external features of image spam. In training model we should train these two feature libraries enough to get good accuracy. After getting external features of incoming image, training model counts occurrences of each feature in each feature library and calculates membership degree for each feature. Membership degree can be calculated as

$$Md\ (t_i, c_j) = f_i,c_j / (f_{i,\ ham} + f_{i,spam})$$

where ti - token

cj - category

fi,ham - frequency of token ti in ham category

fi,spam - frequency of token ti in spam category

2. Classifying model: In classifying model rule evaluation, Defuzzification and histogram intersection methods takes place. Defuzzification outputs whether the incoming image is spam or suspected spam or legitimate. If the incoming image is suspected spam, we are in confused state and go to histogram intersection method which declares suspected spam as image spam or legitimate image. After declaring incoming image, classifying model exports external features into corresponding feature library in training model.

B. *Algorithm*

1. Read Image properties (image size, width, height, bit depth, aspect ratio, Compressibility, Area).
2. Separate as tokens
3. Compare suspected Image property with ham and spam feature database.
4. Forms fuzzy sets
5. Applies fuzzy rule
6. Declare ham or spam or suspected spam
7. If the Image is suspected spam then distinguish between image spam and legitimate image by calculating the similarity between the modal image histogram and test image histogram.

C. *Image Histogram*

Image histogram is the statistical representation of an image. Grey histogram reveals the characteristics of texts in an image for our image spam filter approach. Color histograms of legitimate images tend to be continuous whereas color histogram of image spams tends to have some isolated peaks which property is used to classify legitimate images and image spams. HSV histogram reveals hue, saturation and value which can be used for distinguishing between legitimate image and image spam.

D. *Histogram intersection*

The main objective of using histogram intersection is to compare two images (grey image or RGB image).

Given a pair of histograms P and Q, Intersection of two histograms can be defined as

$$I(P,Q) = \sum_{i=1}^{N} \min\ (P_i\ , Qi)$$

Similarity between two histograms of images can be calculated as follows

$$S = \frac{I(P,Q)}{\sum_{i=1}^{N} Q_i}$$

Histogram intersection works on the principle that if the two images are almost identical, then the sum of the minimums will be almost the same as the sum of the

smaller of the two histograms. This means that 1.0 signifies identity, and the larger the deviation from that, the more dissimilar the image.

IV. ANALYSIS OF TEST RESULTS

We have conducted an experiment on our spam filter using Princeton benchmark dataset for image spams where we can get 1071 images belong to 178 different batches and Corel database for legitimate images where we can get 1000 legitimate images with java eclipse platform. In this experiment we have folded each 25 images into one batch and one batch is used to train feature library and another batch is used for testing. There are no overlapping of trained images and tested images in analyzing results.

In this experiment we have used Java Advanced Image processing (JAI) API for extracting external features of incoming image, ImageJ (IJ) API for some image processing functions and JFuzzyLogic package for Fuzzy Inference System.

Accuracy, precision and recall ratio are well known spam filter performance measures and can be calculated as below

$$Accuracy(A) = \frac{TP + TN}{TP + TN + FN + FP}$$

$$Precision(P) = \frac{TP}{TP + FP}$$

$$Recall(R) = \frac{TP}{TP + FN}$$

TP is the number of e-mail that is spam and correctly predicted as spam; FP is the number of e-mail that is legitimate but predicted as spam; TN is the number of e-mail that is legitimate and is truly predicted as legitimate (ham); and FN the number of e-mail that is spam but predicted as legitimate. Recall indicates that the false positive is high.

No. of images trained	Accuracy	Precision	Recall
50	0.86	0.9	0.8
100	0.912	0.923	0.87
500	0.992	0.966	0.961

Table1. Shows Accuracy, Precision and Recall

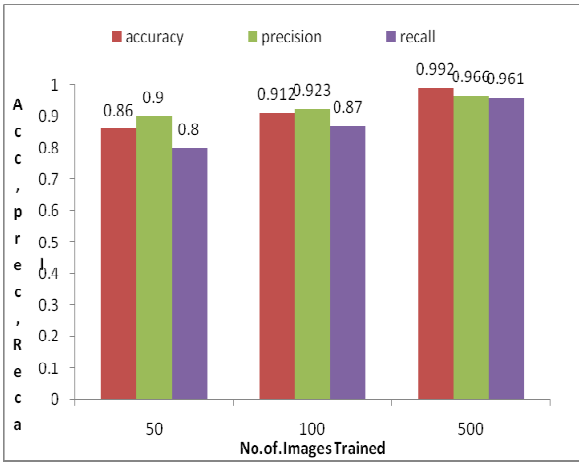


Figure 2. Bar diagram showing filter perforamance measures

V. CONCLUSION

We have designed and analyzed a fuzzy logic based image spam filter which utilizes external properties of image and histogram and we can say this is fast method because of not extracting and analyzing embedded text in an image. Experiment shows high accuracy 0.992, precision 0.966 and recall rate 0.961.

Future work is to implement video spam filter in which each video frame is captured with certain interval of a time then each frame is like an image and apply proposed method to classify legitimate and spam video.

REFERENCES

[1]. How to keep spam off your network, 2007, http://www.gfi.com/whitepapers

[2]. I. P. Giorgio Fumera, Fabio Roli "Spam Filtering Based On The Analysis Of Text Information Embedded Into Images," J. Mach. Learn. Res., vol.7, pp. 2699-2720, 2006.

[3]. B. Biggio, G. Fumera, I. Pillai, and F. Roli, "Image Spam Filtering by Content Obscuring Detection," In Proceeding Fourth Conference on Email and Anti-Spam (CEAS 2007), Mountain View, California, 2007.

[4]. G. Fumera, I. Pillai, F. Roli, and B. Biggio, "Image Spam Filtering Using Textual and Visual Information," In proceeding the MIT Spam Conference 2007, Cambridge, MA, USA 2007.

[5]. P. Klangpraphant and P. Bhattarakosol, "PIMSI: A Partial Image Spam Inspector," in Future Information Technology (FutureTech), 2010 5th International Conference on, 2010, pp. 16.

[6]. F. Gargiulo and C. Sansone, "Combining Visual and Textual Features for Filtering Spam Emails," in Pattern Recognition, 2008. ICPR 2008. 19th International Conference on, 2008, pp. 1-4.

[7]. Sven Krasser, Yuchun Tang, Jeremy Gould, Dmitri Alperovitch, Paul Judge, "Identifying Image Spam based on Header and File Properties using C4.5 Decision Trees and Support Vector Machine Learning". Proc. IEEE Workshop Information Assurance and Security (IAW 07). IEEE Press, Jun. 2007, pp. 255-261, doi: 10.1109/IAW.2007.381941.

[8]. Jong-Wan Kim, Sin-Jae Kang and Byeong Man Kim "A Fuzzy Inference Method for Spam-Mail Filtering", Springer-Verlag Berlin Heideberg 2005, LNAI 3809, pp 1112-1115.

[10]. Message Labs, "Message Labs Intelligence": 2006 Annual Security Report"

[11]. A survey of learning based techniques of email spam filtering eprints.biblio.unitn.it/ archive/00001070/01/056.pdf

[12]. M. Soranamageswari and Dr. C. Meena," Histogram based Image Spam Detection using Back propagation Neural Networks", Vol. 9 Issue 5 (Ver 2.0), January 2010.

◈◈◈

Genetic Algorithm as an Advanced Heuristic Method for finding Optimum Value Using MATLAB Coding

Bharat Soni & V. Nagaraj

Dept. of Electrical and Electronics, Manipal Institute of Technology, Manipal, Karnataka, India

Abstract – Optimization has always been a major research area in Science, Technology, Commerce, Mathematic and Business. Many Classical methods to find optimum value have failed because of their limitations and method of approaching towards optimum region. After that some Heuristic Methods were developed which helped for finding optimum region.

Genetic Algorithm is one the heuristic Methods which converge to best optimum region for any type of function and for wider domain too. This works on principle of Survival of Fittest. In this paper, MATLAB code is developed for multi-variable function in given domain to get maximum value.

Keywords – *Optimization, Classical Method, Heuristic Method, Genetic Algorithm and MATLAB.*

I. INTRODUCTION

Genetic Algorithms are search Algorithms based on the mechanics of natural selection and natural genetics. In every generation, a new set of artificial creatures (strings) is created using bits and pieces of the fittest of an old, an occasional new part is tried for good measure. It is an example of a search procedure that uses random choice as a tool to guide a highly exploitative search through a coding of a parameter space.[1]

The mechanics of simple Genetic Algorithm are surprisingly simple, involving nothing more complex than copying strings and swapping partial strings. Simplicity of Operation and Power of Effect are two of the main attractions of the Genetic Algorithm approach. A simple Genetic Algorithm that yields good results in many practical problems is composed of three operators: Reproduction, Crossover and Mutation.

This Paper represents a Simple MATLAB Code to find the Maximum value of Multi Variable Function(s) within given Domain. Some Assumptions are also made while developing code for simplicity.

II. WORKING PRINCIPLE OF GENETIC ALGORITHM

As the name suggest, Genetic Algorithms (GAs) borrow their working principle from natural genetics. Genetic Algorithm works on the principle of Survival of Fittest. Some fundamental ideas of genetics are borrowed and used artificially to construct search algorithms that are robust and require minimal problem information.

The Working Principle of GAs is very different from that of most classical optimization techniques. Basic steps of Genetic Programming are given below:

1. *Representing a Solution* : First the length of bits l is assumed for binary conversion of strings which are given by users. These strings are varying from 0 to 2^l-1 which are also called chromosomes in genetics. Population of strings P is also defined initially which is fixed for all iterations. For given lower and upper bound of problem, variable values are counted by following formula.

   $$X_i = X_i^{min} + [(X_i^{max} - X_i^{min})/ (2^l -1)]*(\text{string value})$$

   After calculation of all values for population, function values are calculated by putting X_i values.[2]

2. *Assigning Fitness to a Solution* : It is important to reiterate that binary GAs work with strings representing the decision variables, instead of decision variables themselves. Once a string (or a solution) is created by genetic operators, it is necessary to evaluate the solution, particularly in the context of the underlying objective and constraint functions. In the absence of constraints, the fitness of a string is assigned a value which is a function of the solution's objective function value. Formulation of Fitness Function is user depended and also depends on whether maximum value is needed or minimum value of function.[2]

3. *Reproduction or Selection Operator :* The primary objective of the reproduction operator is to make

duplicates of good solutions and eliminate bad solutions in a population, while keeping the population size constant. This is achieved by: first identify good solutions in population, then make multiple copies of good solution and then eliminate bad solutions so that multiple copies of good solution can be placed in the population. There exists a number of ways to achieve the above tasks. Some common methods are tournament selection, proportionate selection and ranking selection.[2]

4. *Crossover Operator :* A Crossover Operator is applied next to the strings of the matting pool, as Reproduction Operator cannot produce new solutions in the population. The creation of new solutions is performed by crossover and mutation operator. For Crossover, first crossover site is chosen, right hand side of crossover site get interchanged between two strings and other just copied. This is the simple Single Point Crossover. Like this, there are double crossovers, custom crossover and random crossover are also there, but they are not used in this paper. Like Reproduction, Crossover site are also decided by Roulette Wheel Selection, Flipping of coin and Random Number Generation with cross over probability P_c. But here we are just defining fix crossover sites in MATLAB code for simplicity.[1][3]

   If $x=(x_1,x_2,x_3,\ldots\ldots,x_n)$; $y=(y_1,y_2,y_3,\ldots\ldots,y_n)$

   Then crossover at i^{th} position is defined as

   $$X' = (x_1,\ldots\ldots, x_i, y_{i+1},\ldots\ldots,y_n)$$

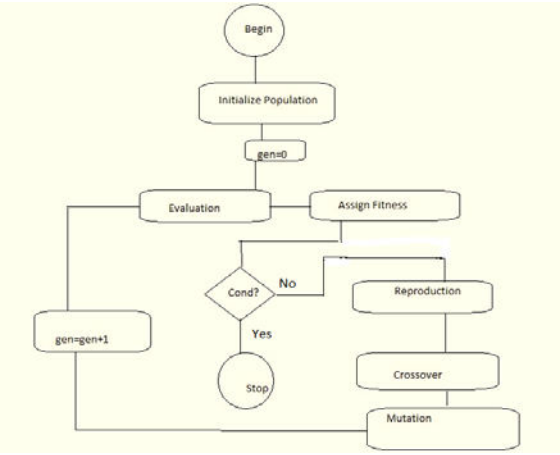
   $$Y' = (y_1,\ldots\ldots, y_i, x_{i+1},\ldots\ldots,x_n)$$

5. *Mutation Operator:* The crossover operator is mainly responsible for the search aspect of genetic algorithms, even though the mutation operator is also used for this purpose. The bit wise mutation operator changes a 1 to 0, and vice-versa, with mutation probability P_m. The need for mutation is to keep diversity in the population. The bit-wise mutation operator requires the creation of a random number for every bit. In order to reduce computational complexity, *Goldberg* suggested a *mutation clock operator,* where after a bit is mutated, the location of next mutated bit is determined by an exponential distribution.[1][3]

   0011 0101(3, 5) → 0011 0001(3, 1)

The Three Operators- selection, crossover and mutation are simple and straightforward. The reproduction operator selects good strings, while the crossover operator recombines together good substrings to hopefully form a better substring. The mutation operator alters a string locally to hopefully create a better string. Since none of these operations are performed deterministically, these claims are not guaranteed, nor explicitly tested, during GA operation. However, it is expected that is bad strings are created they will be eliminated by the reproduction operator in subsequent generations and if good strings are created, they will be emphasized.[2]

Flow chart of Genetic programming



III. MATLAB CODE EXPLANATION AND RESULT ANALYSIS

Simple MATLAB code is developed to find the maximum value of given function within domain. Here population size 8 is chosen and bit length is 6, so decimal values of strings in population can be given in between 0 to 63 by users. The above code is for 2 variables but can be change into many variables by just little change in code. Function is written in func.m file which is to be optimized by Genetic Algorithm. Here Mutation process is not considered only crossover is considered with probability 80%. The crossover sites are fixed for all iterations and they are in following manner: 3^{rd} crossover site for 1^{st} and 5^{th} population string, 4^{th} crossover site for 2^{nd} and 3^{rd} population string, 1^{st} crossover site for 4^{th} and 8^{th} population string and 5^{th} crossover site for 6^{th} and 7^{th} population string. Again these are programmer depended, they can be changed and but take care criterion of crossover probability. There is no specific reason for choosing that specific crossover sites for those strings. To make code better, binary length bit l and population size can be increased.

Function	Domain of X1;X2	Genetic Code O/P	Actual O/P

$X1^2 + 4X1 + 6X2 + 5$	[-1, 3];[-1,1]	30.3653	32
$X1^2 + X2^2$	[2, 7];[1, 4]	63.5165	65
$-X1^2 + 3X2 - 5$	[-9, 2];[0,1]	-3.1927	-2
$X1^2 + 1/X2$	[0, 5];[0,5]	Infinity	Infinity
Sin(X1+X2)	[-1, 1];[-2,2]	0.7863	1

**All these outputs are for particular set of population strings, Result may be change for another set.

IV. CONCLUSIONS

Binary GAs works with a coding of decision variables, instead of the variables themselves. They work with a discrete search space, even though function may be continuous. On the other hand, since functions values at various discrete solutions are required, a discrete or discontinuous function may be handled by using GAs. This allows GAs to be applied to a wide variety of problem domains. One of the drawbacks of using a coding is that a suitable coding must be chosen for proper working of GA. Although it is difficult to know beforehand what coding is suitable for a problem, a number of experimental studies suggest that a coding which respects the underlying building block processing must be used.

ACKNOWLEDGMENT

I express my admiration and convey sincere thanks to Mr V. Nagaraj for guiding me about Genetic Algorithm work. I also thank to my college Manipal Institute of Technology, Manipal for helping me.

REFERENCES

[1] Genetic Algorithm in Search, Optimization & Machine Learning, David E. Goldberg, Low Price Edition, Pearson Education, ISBN 81-7808-130-X .

[2] Multi-Objective Optimization using Evolutionary Algorithm, Kalyanmoy Deb, Wiley India Edition, ISBN 978-81-265-2804-2.

[3] Principles of Soft Computing, S.N. Sivanandam and S.N. Deepa, Wiley India Edition, ISBN 10: 81-265-1075-7.

[4] Essential MATLAB for Engineers and Scientist, Brian H. Hahn and Denial T. Valentine, Elsevier Publications, ISBN 978-0-12-374883-6.

◈◈◈

Simulator Platform for Benchmarking Live Streaming in P2P Systems

Gopala.V & Vinodha.K

Information Science and Engineering, Dept of ISE, The Oxford College of Engineering, Bangalore, India

Abstract **–** Video streaming over the Internet can be very difficult under the traditional client-server model. Peer-to-peer (P2P) systems, in which each participating peer contributes its upload bandwidth to other peers while it downloads data, have been successful in file-sharing applications, and they appear to be promising in delivering video contents, too. A P2PTV system over internet allows users to watch live video streams redistributed by other users via a peer-to-peer (P2P)network. In an ideal world, each peer in a P2P network would be able to redistribute more bytes than it receives. A P2PTV system built from such peers can support a virtually unlimited number of peers; with only a single copy of content stream injected into the network, it can redistribute the content to all peers. A number of commercial peer-to-peer systems for live streaming have been introduced in recent years. The behaviour of these popular systems has been extensively studied in several measurement papers. Although such studies are useful to compare different systems from end-user's perspective, it is difficult to intuitively understand the observed properties without fully reverse-engineering the underlying systems.

Keywords - Live streaming, peer to peer live streaming .Network architecture.

I. INTRODUCTION

There is an emerging market for IPTV. Numerous commercial systems now offer services over the Internet that are similar to traditional over-the-air, cable, or satellite TV. Live television, time-shifted programming, and content-on demand are all presently available over the Internet. Increased broadband speed, growth of broadband subscription base, and improved video compression technologies have contributed to the emergence of these IPTV services. We draw a distinction between three uses of peer-to-peer (P2P) networks: delay tolerant file download of archival material, delay sensitive progressive download (streaming) of archival material, and real-time live streaming. In the first case, the completion of download is elastic, depending on available bandwidth in the P2P network. The application buffer receives data as it trickles in and informs the user upon the completion of download. The user can then start playing back the file for viewing in the case of a video file. Bit torrent and variants are example of delay-tolerant file download systems. In the second case, video playback starts soon as the application assesses it has sufficient data buffered that, given the estimated download rate and the playback rate, it will not deplete the buffer before the end of file. If this assessment is wrong, the application would have to either pause playback rebuffed, or slow down playback. While users would like playback to start as soon as possible, the application has some degree of freedom in trading off playback start time against estimated network capacity. Most video-on demand systems are examples of delay-sensitive progressive download application. The third case, real-time live streaming has the most stringent delay requirement. While progressive download may tolerate initial buffering of tens of seconds or even minutes, live streaming generally cannot tolerate more than a few seconds of buffering. Taking into account the delay introduced by signal ingest and encoding, and network transmission and propagation, the live streaming system can introduce only a few seconds of buffering time end to-end and still be considered "live" [1]. The Zattoo peer-to-peer live streaming system was a free to- use network serving over 3 million registered users in eight European countries at the time of study, with a maximum of over 60,000 concurrent users on a single channel. The system delivers live streams using a *receiver-based, peer division multiplexing* scheme as described in Section II. To ensure real-time performance when peer uplink capacity is below requirement, Zattoo subsidizes the network's bandwidth requirement, as described in Section III. After delving into Zattoo's architecture in detail, we study in Sections IV and V large-scale measurements collected during the live broadcast of the UEFA European Football Championship, one of the most popular one-time events in Europe, in June, 2008 [2]. During the course of the month of June 2008, Zattoo served more than 35 million sessions to more than one million distinct users. Drawing from these measurements, we report on the

operational scalability of Zattoo's live streaming system along Several key issues.

II. SYSTEM ARCHITECTURE

The Zattoo system rebroadcasts live TV, captured from satellites, onto the Internet. The system carries each TV channel on a separate peer-to-peer delivery network and is not limited in the number of TV channels it can carry. Although a peer can freely switch from one TV channel to another, and thereby departing and joining different peer-to-peer networks, it can only join one peer-to-peer network at any one time. We henceforth limit our description of the Zattoo delivery network as it pertains to carrying one TV channel. Fig. 1 shows a typical setup of a single TV channel carried on the Zattoo network. TV signal captured from satellite is encoded into H.264/AAC streams, encrypted, and sent onto the Zattoo network. The encoding server may be physically separate from the server delivering the encoded content onto the Zattoo network.

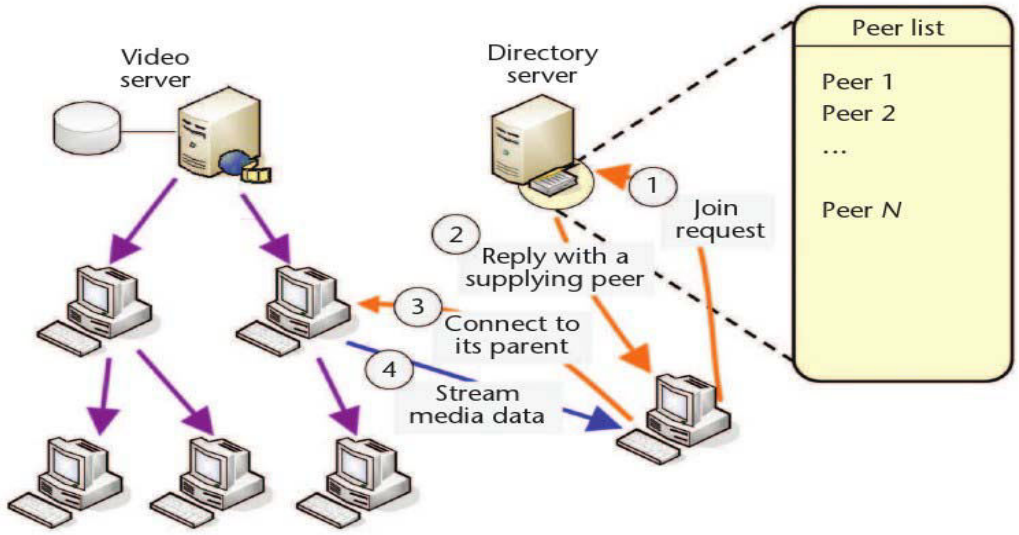


Fig. 1 : Network architecture.

For ease of exposition, we will consider the two as logically co-located on an Encoding Server. Users are required to register themselves at the Zattoo website to download a free copy of the Zattoo player application. To receive the signal of a channel, the user first authenticates itself to the Zattoo Authentication Server. Upon authentication, the user is granted a ticket with limited lifetime.

The user then presents this ticket, along with the identity of the TV channel of interest, to the Zattoo Rendezvous Server. If the ticket specifies that the user is authorized to receive signal of the said TV channel, the Rendezvous Server returns to the user a list of peers currently joined to the P2P network carrying the channel, together with a signed channel ticket.

If the user is the first peer to join a channel, the list of peers it receives contains only the Encoding Server.

The user joins the channel by contacting the peers returned by the Rendezvous Server, presenting its channel ticket, and obtaining the live stream of the channel from them. Each live stream is sent out by the Encoding Server as n logical sub-streams. The signal received from satellite is encoded into a variable-bit rate stream. During periods of source quiescence, no data is generated. During source busy periods, generated data is packetized into a packet stream, with each packet limited to a maximum size. The Encoding Server multiplexes this packet stream onto the Zattoo network as n logical sub-streams.

Thus the first packet generated is considered part of the first sub-stream, the second packet that of the second sub-stream, the n-th packet that of the n-th sub stream. The n+1-th packet cycles back to the first sub-stream, etc. such that the i-th sub-stream carries the mn+i-th packets. Being the packet index, i also serves as the sub-

stream index. The number mn + i is carried in each packet as its sequence number.

Zattoo uses the Reed-Solomon (RS) error correcting code (ECC) for forward error correction. The RS code is a systematic code: of the n packets sent per segment, k < n packets carry the live stream data while the remainder carries the redundant data [3, Section 7.3]. Due to the variable-bit rate nature of the data stream, the time period covered by a segment is variable, and a packet may be of size less than the maximum packet size. A packet smaller than the maximum packet size is zero-padded to the maximum packet size for the purposes of computing the (shortened) RS code, but is transmitted in its original size. Once a peer has received k packets per segment, it can reconstruct the remaining n – k packets. We do not differentiate between streaming data and redundant data in our discussion in the remainder of this paper.

When a new peer requests to join an existing peer, it specifies the sub-stream(s) it would like to receive from the existing peer. These sub-streams do not have to be consecutive. Contingent upon availability of bandwidth at existing peers, the receiving peer decides how to multiplex a stream onto its set of neighboring peers, giving rise to our description of the Zattoo live streaming protocol as a receiver-based, peer division multiplexing protocol. The details of peer-division multiplexing is described in Section II-A while the details of how a peer manages sub-stream forwarding and stream reconstruction is described in Section II-B. Receiver-based peer division multiplexing has also been used by the latest version of Cool Streaming peer-to-peer protocol though it differs from Zattoo in its stream management (Section II-B) and adaptive behavior.

A. Peer-Division Multiplexing

To minimize per-packet processing time of a stream, the Zattoo protocol sets up a virtual circuit with multiple fan outs at each peer. When a peer joins a TV channel, it establishes a peer-division multiplexing (PDM) scheme amongst a set of neighboring peers, by building a virtual circuit to each of the neighboring peers. Baring departure or performance degradation of a neighbor peer, the virtual circuits are maintained until the joining peer switches to another TV channel. With the virtual circuits set up, each packet is forwarded without further per-packet handshaking between peers. We describe the PDM boot strapping mechanism in this section and the adaptive PDM mechanism to handle peer departure and performance degradation. The PDM establishment process consists of two phases:the search phase and the join phase. In the search phase, the new, joining peer determines its set of potential neighbors. In the join phase, the joining peer requests peering relationships with a subset of its potential neighbors. Upon acceptance of a peering relationship request, the peers become neighbors and a virtual circuit is formed between them. Search phase. To obtain a list of potential neighbors, a joining peer sends out a SEARCH message to a random subset of the existing peers returned by the Rendezvous Server. The SEARCH message contains the sub-stream indices for which this joining peer is looking for peering relationships. The sub stream indices are usually represented as a bitmask of n bits, where n is the number of sub-streams defined for the TV channel. In the beginning, the joining peer will be looking for peering relationships for all sub-streams and have all the bits in the bitmask turned on. In response to a SEARCH message, an existing peer replies with the number of sub-streams it can forward. From the returning SEARCH replies, the joining peer constructs a set of potential neighbors that covers the full set of sub-streams comprising the live stream of the TV channel. The joining peer continues to wait for SEARCH replies until the set of potential neighbors contains at least a minimum number of peers, or until all SEARCH replies have been received.

With each SEARCH reply, the existing peer also returns a random subset of its known peers. If a joining peer cannot form a set of potential neighbors that covers all of the sub streams of the TV channel, it initiates another SEARCH round, sending SEARCH messages to peers newly learned from the previous round. The joining peer gives up if it cannot obtain the full stream after two SEARCH rounds. To help the joining peer synchronize the sub-streams it receives from multiple peers, each existing peer also indicates for each sub-stream the latest sequence number it has received for that sub-stream, and the existence of any quality problem. The joining peer can then choose sub-streams with good quality that are closely synchronized.

Join phase : Once the set of potential neighbors is established, the joining peer sends JOIN requests to each potential neighbor. The JOIN request lists the sub-streams for which the joining peer would like to construct virtual circuit with the potential neighbor. If a joining peer has l potential neighbors, each willing to forward it the full stream of a TV channel, it would typically choose to have each forward only 1 stream, to spread out the load amongst the peers and to speed up error recovery, as described in Section II-C. In selecting which of the potential neighbors to peer with, the joining peer gives highest preference to topologically close-by peers, even if these peers have less capacity or carry lower quality sub-streams. The “topological” location of a peer is defined to be its subnet number, autonomous system (AS) number, and country code, in that order of precedence. A joining peer obtains its own topological location from the Zattoo Authentication Server as part of its authentication process. The list of

peers returned by both the Rendezvous Server and potential neighbors all come attached with topological locations. A topology-aware overlay not only allows us to be "ISP-friendly," by minimizing inter-domain traffic and thus save on transit bandwidth cost, but also helps reduce the number of physical links and metro hops traversed in the overlay network, potentially resulting in enhanced user perceived stream quality.

B. Stream Management

We represent a peer as a packet buffer, called the IOB,fed by sub-streams incoming from the PDM constructed as described in Section II-A.1 The IOB drains to (1) a local media player if one is running, (2) a local file if recording is supported, and (3) potentially other peers. Fig. 2 depicts a Zattoo player application with virtual circuits established to four peers.

As packets from each sub-stream arrive at the peer, they are stored in the IOB for reassembly to reconstruct the full stream. Portions of the stream that have been reconstructed are then played back to the user. In addition to providing a reassembly area, the IOB also allows a peer to absorb some variability's in available network bandwidth and network delay. The IOB is referenced by an input pointer, a repair pointer, and one or more output pointers. The input pointer points to the slot in the IOB where the next incoming packet with sequence number higher than the highest sequence number received so far will be stored. The repair pointer always points one slot beyond the last packet received in order and is used to regulate packet retransmission and adaptive PDM as described later. We assign an output pointer to each forwarding destination. The output pointer of a destination indicates the destination's current forwarding horizon on the IOB.

In accordance to the three types of possible forwarding destinations listed above, we have three types of output pointers: player pointer, file pointer, and peer pointer. One would typically have at most one player pointer and one file pointer but potentially multiple concurrent peer pointers, referencing an IOB. The Zattoo player application does not currently support recording. Since we maintain the IOB as a circular buffer, if the incoming packet rate is higher than the forwarding rate of a particular destination, the input pointer will overrun the output pointer of that destination. We could move the output pointer to match the input pointer so that we consistently forward the

oldest packet in the IOB to the destination. Doing so, however, requires checking the input pointer against all output pointers on every packet arrival. Instead, we have implemented the IOB as a double buffer. With the double buffer, the positions of the output pointers are checked against that of the input pointer only when the input pointer moves from one sub-buffer to the other. When the input pointer moves from sub-buffer a to sub buffer b all the output pointers still pointing to sub-buffer b are moved to the start of sub-buffer a and sub-buffer b is flushed, ready to accept new packets. When a sub-buffer is flushed while there are still output pointers referencing it, packets that have not been forwarded to the destinations associate with those pointers are lost to them, resulting in quality degradation. To minimize packet lost due to sub-buffer flushing, we would like to use large sub-buffers. However, the real-time delay requirement of live streaming limits the usefulness of late arriving packets and effectively puts a cap on the maximum size of the sub-buffers. Different peers may request for different numbers of, possibly non-consecutive, sub-streams. To accommodate the different forwarding rates and regimes required by the destinations, we associate a packet map and forwarding discipline with each output pointer. Fig. 3 shows the packet map associated with an output peer pointer where the peer has requested sub-streams 1, 4, 9, and 14. Every time a peer pointer is repositioned to the beginning of a sub-buffer of the IOB, all the packet slots of the requested sub-streams are marked needed and all the slots of the sub-streams not requested by the peer are marked SKIP. When a needed packet arrives and is stored in the IOB, its state in the packet map is changed to READY.

As the peer pointer moves along its associated packet map, READY packets are forwarded to the peer and their states changed to SENT. A slot marked needed but not READY, such as slot n + 4 in Fig. 3, indicates that the packet is lost or will arrive out-of-order and is bypassed. When an out-of order packet arrives, its slot is changed to READY and the peer pointer is reset to point to this slot. Once the out-of-order packet has been sent to the peer, the peer pointer will move forward, bypassing all SKIP, NEED, and SENT slots until it reaches the next READY slot, where it can resume sending. The player pointer behaves the same as a peer pointer except that all packets in its packet map will always start out marked Needed.. IOB consisting of a double buffer, with an input pointer, a repair pointer, and an output file pointer, an output player pointer, and two output peer pointers referencing the IOB. Each output pointer has a packet map associated with it. For the scenario depicted in the figure, the player pointer tracks the input pointer and has skipped over some lost packets. Both peer pointers are lagging the input pointer, indicating that the forwarding rates to the peers are bandwidth limited. The file pointer is pointing at the first lost packet. Archiving a live stream to file does not impose real-time delay bound on packet arrivals. To achieve the best quality recording possible, a recording peer always waits for retransmission of lost packets that

cannot be recovered by error correction. In addition to achieving lossless recording, we use retransmission to let a peer recover from transient network congestion. A peer sends out a retransmission request when the distance between the repair pointer and the input pointer has reached a threshold of R packet slots, usually spanning multiple segments. A retransmission request consists of an R- bit packet mask, with each bit representing a packet, and the sequence number of the packet corresponding to the first bit. Marked bits in the packet mask indicate that the corresponding packets need to be retransmitted. When a packet loss is detected, it could be caused by congestion on the virtual circuits forming the current PDM or congestion on the path beyond the neighboring peers. In either case, current neighbour peers will not be good sources of retransmitted packets. Hence we send our retransmission requests to r random peers that are not neighbor peers. A peer receiving a retransmission request will honour the request only if the requested packets are still in its IOB and it has sufficient left-over capacity, after serving its current peers, to transmit all the requested packets. Once a retransmission request is accepted, the peer will retransmit all the requested packets to completion.

C. Adaptive PDM

While we rely on packet retransmission to recover from transient congestions, we have two channel capacity adjustment mechanisms to handle longer-term bandwidth fluctuations. The first mechanism allows a forwarding peer to adapt the number of sub-streams it will forward give its current available bandwidth, while the second allows the receiving peer to switch provider at the sub-stream level. Peers on the Zattoo network can redistribute a highly variable number of sub-streams, reflecting the high variability in uplink bandwidth of different access network technologies. For a full-stream consisting of sixteen constant-bit rate sub streams, our prior study show that based on realistic peer characteristics measured from the Zattoo network, half of the peers can support less than half of a stream, 82% of peers can support less than a full-stream, and the remainder can support up to ten full streams (peers that can redistribute more than a full stream is conventionally known as super nodes in the literature) [5]. With variable-bit rate streams, the bandwidth carried by each sub-stream is also variable.

To increase peer bandwidth usage, without undue degradation of service, we instituted measurement-based admission control at each peer. In addition to controlling resource commitment, another goal of the measurement-based admission control module is to continually estimate the amount of available uplink bandwidth at a peer. The amount of available uplink bandwidth at a peer is initially estimated by the peer sending a pair of probe packets to Zattoo's Bandwidth Estimation Server. Once a peer starts forwarding sub-streams to other peers, it will receive from those peers quality-of-service feedbacks that inform its update of available uplink bandwidth estimate. A peer sends quality of- service feedback only if the quality of a sub-stream drops below a certain threshold.2 Upon receiving quality feedback from multiple peers, a peer first determines if the identified sub-streams are arriving in low quality. If so, the low quality of service may not be caused by limit on its own available uplink bandwidth; in which case, it ignores the low quality feedbacks. Otherwise, the peer decrements its estimate of available uplink bandwidth. If the new estimate is below the bandwidth needed to support existing number of virtual circuits, the peer closes Depending on a peer's NAT and/or firewall configuration, Zattoo uses either UDP or TCP as the underlying transport protocol. The quality of a sub stream is measured differently for UDP and TCP. A packet is considered lost under UDP if it doesn't arrive within a fixed threshold. The quality measure for UDP is computed as a function of both the packet lost rate and the burst error rate (number of contiguous packet losses).

The quality measure for TCP is defined to be how far behind a peer is, relative to other peers, in serving its sub-streams. a virtual circuit. To reduce the instability introduced into the network, a peer closes first the virtual circuit carrying the smallest number of sub-streams. A peer attempts to increase its available uplink bandwidth estimate periodically: if it has fully utilized its current estimate of available uplink bandwidth without triggering any bad quality feedback from neighboring peers. A peer doubles the estimated available uplink bandwidth if current estimate is below a threshold, switching to linear increase above the threshold, similar to how TCP maintains its congestion window size. A peer also increases its estimate of available uplink bandwidth if a neighbor peer departs the network without any bad quality feedback. When the repair pointer lags behind the input pointer by R packet slots, in addition to initiating a retransmission request, a peer also computes a loss rate over the R packets. If the loss rate is above a threshold, the peer considers the neighbour slow and attempts to reconfigure its PDM. In reconfiguring its PDM, a peer attempts to shift half of the sub-streams currently forwarded by the slow neighbor to other existing neighbors. At the same time, it searches for new peer(s) to forward these sub-streams. If new peer(s) are found, the load will be shifted from existing neighbors to the new peer(s). If sub-streams from the slow neighbor continues to suffer after the reconfiguration of the PDM, the peer will drop the neighbor completely and initiate another reconfiguration of the PDM. When a peer loses a neighbor due to

reduced available uplink bandwidth at the neighbor or due to neighbor departure, it also initiates a PDM reconfiguration. A peer may also initiate a PDM reconfiguration to switch to a topologically closer peer. Similar to the PDM establishment process, PDM reconfiguration is accomplished by peers exchanging sub stream Bitmasks in a request/response handshake, with each bit of the bitmask representing a sub-stream. During and after a PDM reconfiguration, slow neighbor detection is disabled for a short period of time to allow for the system to stabilize.

III. GLOBAL BANDWIDTH SUBSIDY SYSTEM

Each peer on the Zattoo network is assumed to serve a user through a media player, which means that each peer must receive, and can potentially forward, all n sub-streams of the TV channel the user is watching. The limited redistribution capacity of peers on the Zattoo network means that a typical client can contribute only a fraction of the sub streams that make up a channel. This shortage of bandwidth leads to a global bandwidth deficit in the peer-to-peer network. Whereas bit torrent-like delay-tolerant file downloads or the delay-sensitive progressive download of video-on-demand applications can mitigate such global bandwidth shortage by increasing download time, a live streaming system such as Zattoo's must subsidize the bandwidth shortfall to provide real-time delivery guarantee. Zattoo's Global Bandwidth Subsidy System (or simply, the Subsidy System), consists of a global bandwidth monitoring subsystem, a global bandwidth forecasting and provisioning subsystem, and a pool of Repeater nodes. The monitoring subsystem continuously monitors the global bandwidth requirement of a channel. The forecasting and provisioning 6 subsystem projects global bandwidth requirement based on measured history and allocates Repeater nodes to the channel as needed.

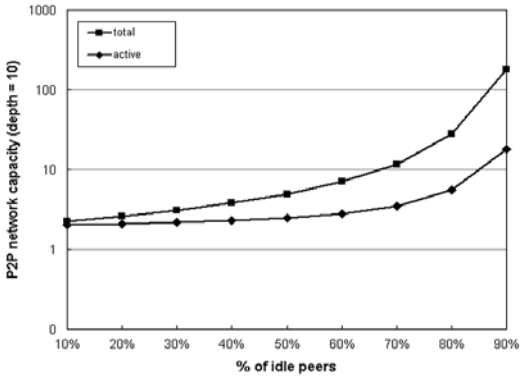
The monitoring and provisioning of global bandwidth is complicated by two highly varying parameters over time, client population size and peak streaming rate, and one varying parameter over space, available uplink bandwidth, which is network-service provider dependent. Forecasting of bandwidth requirement is a vast subject in itself. Zattoo adopted a very simple mechanism, described in Section III-B which has performed adequately in provisioning the network for both daily demand fluctuations and flash crowds scenarios (see Section IV-C). When a bandwidth shortage is projected for a channel, the Subsidy System assigns one or more Repeater nodes to the channel. Repeater nodes function as bandwidth multiplier, to amplify the amount of available bandwidth in the network. Each Repeater node serves at most one channel at a time; it joins and leaves a given channel at the behest of the Subsidy System. Repeater nodes receive and serve all n sub-streams of the channel they join, run the same PDM protocol, and are treated by actual peers like any other peers on the network; however, as bandwidth amplifiers, they are usually provisioned to contribute more uplink bandwidth than the download bandwidth they consume. The use of Repeater nodes makes the Zattoo network a hybrid P2P and content distribution network. We next describe the bandwidth monitoring subsystem of the Subsidy System, followed by design of the simple bandwidth projection and Repeater node assignment subsystem.

IV. SIMULATION OF A P2P LIVE STREAMING OVERLAY

After having studied some theoretical aspects of a P2Pstreaming overlay and having proposed enhancement mechanisms, we now present results obtained by simulation and highlight the impact of peer selection algorithms on the performance of a typical P2P live streaming system.

A. Simulation parameters and metrics

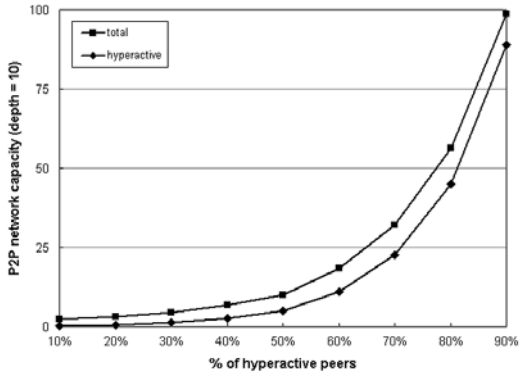
Our simulation code implements Zattoo's push-driven sub stream based streaming architecture [3], and was run on top of the network manipulator software [29]. An Internet map of 4.2k-node was used as the underlying topology [30]. We assume that a given overlay distributes one TV channel, and do not take daily channel size variations into account. Each run lasts for 12 hours and only the last 6 hours are analysed; after the 6th hour, we are in a steady state regime.



Each result value is the average of 30 runs, and the standard deviation values are provided The uplink capacity of individual peers is assigned so that the resulting distribution becomes the same as the empirical distribution. The NAT type of a peer, which determines its reachability in the overlay, is also taken from the empirical distribution of NAT types reported in [5].

B. Simulation results

This section presents our simulation results. A peer trying to connect to other peers to get all necessary sub-streams is called an orphan peer. It sends search messages to discover other peers, sends join messages to connect to available peers, and finally obtains the full stream from them. A peer who is able and willing to offer a part of or all requested sub streams for an orphan peer is called an adoptive peer. An adoptive peer has positively answered to the search message of a orphan. Once having multiple positive answers from candidate adoptive peers, an orphan has to choose to which peer it should send a join message.



V. CONCLUSION

We have presented a receiver-based, peer-division multiplexing engine to deliver live streaming content on a peer-to-peer network. The same engine can be used to transparently build a hybrid P2P/CDN delivery network by adding Repeater nodes to the network. By analyzing large amount of usage data collected on the network during one of the largest viewing event in Europe, we have shown that the resulting network can scale to a large number of users and can take good advantage of available uplink bandwidth at peers. We have also shown that error-correcting code and packet retransmission can help improve network stability by isolating packet losses and preventing transient congestion from resulting in PDM reconfigurations. We have further shown that the PDM and adaptive PDM schemes presented have small enough overhead to make our system competitive to digital satellite TV in terms of channel switch time, stream synchronization, and signal lag.

REFERENCES

[1] K. Shami, D. Magoni, H. Chang, W. Wang, and S. Jamin, "Impacts of Peer Characteristics on P2PTV Networks Scalability," in Proceedings of the 28th IEEE Conference on Computer Communications – Mini- Conference, April 2009.

[2] D. Tran, K. Hua, and T. Do, "ZIGZAG: An Efficient Peer-to-Peer Scheme for Media Streaming," in Proceedings of the IEEE Infocom, 2003.

[3] R. Rejaie and S. Stafford, "A Framework for Architecting Peer-to-Peer Receiver-Driven Overlays," in Proceedings of the ACM NOSSDAV, 2004, pp. 42–47.

[4] X. Liao, H. Jin, Y. Liu, L. Ni, and D. Deng, "Anysee: Peer-to-Peer Live Streaming," in Proceedings of the IEEE Infocom, 2006.

[5] X. Zhang, J. Liu, B. Li, and T.-S. Yum, "CoolStreaming/DONet: A Datadriven Overlay Network for Live Media Streaming," in Proceedings of the 24th IEEE Infocom, 2005, p. 21022111.

[6] F. Pianese, "PULSE: A Flexible P2P Live Streaming System," in Proceedings of the 9th IEEE Global Internet Symposium, 2006.

[7] F. Pianese, D. Perino, J. Keller, and E. Biersack, "PULSE: An Adaptive, Incentive-Based, Unstructured P2P Live Streaming System," IEEE Transactions on Multimedia, vol. 9, no. 6, 2007.

[8] K. Sripanidkulchai, B. Maggs, and H. Zhang, "An Analysis of Live Streaming Workloads on the Internet," in Proceedings of the ACM IMC, 2004, p. 4154.

[9] C. Wu, B. Li, and S. Zhao, "Magellan: Charting Large-Scale Peer-to- Peer Live Streaming Topologies," in Proceedings of the 27th International Conference on Distributed Computing Systems (ICDCS'07), 2007,p. 62.

◈◈◈

Illumination Normalization Facial Features Approach For Gender Classification

Ruchi Goel & Suneeta Agarwal
Computer Science and Engineering Department,
Motilal Nehru National Institute of Technology, Allahabad, India

Abstract **-** This paper considers the problem of gender classification by geometrically extracting the facial features (eyebrows, eyes, nose and lip) from the front view images. We have used the face features and illumination normalization features for gender classification under varying lighting conditions. The feature vector was constructed containing features. Principal component analysis (PCA) is used to reduce the features subspace. The underlying classifier, Support vector machine (SVM) is applied on a reduced features subspace. When illumination normalization features was taken, we achieved an improved classification rate of 97.14% on outdoor database of the 140 subjects, 98.00% on Caltech database of the 28 subjects.

Keywords**-**Feature Extraction; Gender Classification; SVM; PCA; Gradient domain

I. INTRODUCTION

Gender classification is one of the most biologically important tasks [1]. Visually humans are able to predict the gender fast but machine does not have this ability. Thus in computer vision and pattern recognition, gender classification of human face is one of the challenging problems. Gender classification is useful in many applications such as biometric authentication, security systems, Human Computer Interaction, and demographic research. Thus researchers have an increased interest in Gender classification for years.

In this paper, we propose an improved algorithm of Face Synthesis system [10] for extracting the facial features and a gender classification method by using two approaches. Through experiments we have analyzed that when faces are cropped from hair to chin (top to bottom) and from right ear to left ear (left to right), the eye lies in mid of cropped face.

Illumination normalization has an effect on classification performance. To handle the illumination normalization problem various methods have been proposed like histogram equalization (HE) [11], [12], logarithm transform [13]. But these methods do not provide good results for different lightening conditions. To overcome the illumination problems, we used method proposed by Zang et al. [14]. In the method proposed by Zang et al. [14] gradient faces were used for face recognition. In our approach, we first store the extracted face features (eyebrows, eyes, nose, lip) in a feature vector for classification and then, we take the gradient images of these features and store in a feature vector for classification. The best result is produced by taking illumination normalization features as gradient features.

The underlying classifier is a Support Vector Machine (SVM), which analyze data and recognize patterns, used for classification. For a classification of high-dimensional data, the complexity of SVM is managed by dimensionality reduction by Principal Component Analysis (PCA). Using intensity of features, the SVM yields promising classification performance. To further improve the classification performance, we use the gradient images of these features, which is stable shape descriptor. The use of gradient features significantly improves the classification performance. The best result is produced by combining the all gradient features into a feature

The proposed simplified model of Gender Classification system was illustrated in fig. 1.

The rest of the paper is organized as follows. In Section II, related work is described. In Section III the proposed approach is explained in detail. Dimension reduction and Classification is described in Section IV. Experimental results and implementations are presented in Section V. Finally in Section VI, conclusion is given.

II. RELATED WORK

Gender classification methods can be divided in feature based and appearance-based method. In 1990, first Gender classification method was proposed by Cottrell and Metcalfe [2] and by Golomb et al. [3]. These methods were based on appearance and used multi-layer neural network approach.

In 1997, Wiskott et al. [4] presented a system where gender classification was based on the Gabor wavelets. System was also used for face recognition. In 1996, Tamura et al. [5] experimented with very low resolution and 8*8 size face images with neural networks. They achieved 93% classification rate.

In 2000 Lyons et al. [6] used Gabor wavelets to detect a face and classify gender with linear discriminant analysis (LDA) and principal component analysis (PCA). Shakhmarovich et al. [7] used Haar-like features within an AdaBoost-based approach and achieved 79% and 79.2% recognition accuracy in gender recognition and classification, respectively.

Recently, Makinen and Raisamo [8] compared

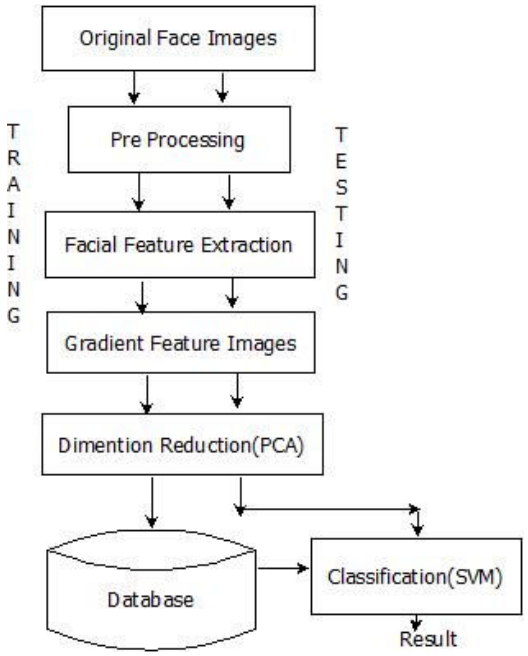


Fig. 1. Gender Classification system

appearance based, feature-based, aligned, and unaligned approaches, among others. They performed experiments and got similar performance results for feature-based AdaBoost and appearance-based SVM and RBF classifier. In another work [9] they experimented with classifier combination, and face normalizations.

III. THE PROPOSED APPROACH

A. Feature Extraction

In this process individual features of the face image are extracted and used as input in the classification process to identify the gender.

The first step is to track the position of eye which acts as a reference point to determine the position of other features like eyebrows, nose and lip. In Face Synthesis system [10], first, the face is split in two halves then the right half of face is used to extract the row where eye will lie. If the face image cropped from left ear to right ear and from hair to chin then eye will lie in the middle of the face image. Thus for calculating the eye area, we have taken 40 to 60 percent area of the height and 50 to 90 percent area of width depicted by shaded rectangular region in fig. 2. This will also result in reduction of search space complexity. In this area, we find the maximum intensity difference row. This is the row where eye will lie and then by using the symmetry property of face with respect to the eyes position, position of other features like eyebrows, nose, and lip are calculated.

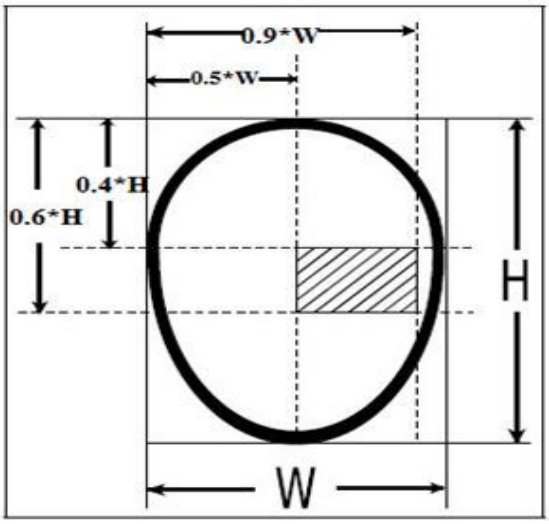


Fig. 2. Shaded rectangular region showing ROI for eye position

Algorithm for calculating the face features coordinates is as follows-

1. Eye area is located.

2. Coordinates for the corners of the area where eye will lie in the preprocessed image are calculated by$x_1 = 0.4 * H, y_1 = 0.5 * W, x_2 = 0.6 * H, y_2 = 0.9 * W$. Where W and H are width and height of the preprocessed image.

3. Maximum intensity difference row is located in this area named as $RowIndex$.

4. Coordinates for the corners of the Left Eye area are calculated by

$$x_1 = RowIndex - 0.06 * W$$

$$x_2 = RowIndex + 0.06 * W$$

$$y_1 = 0.14 * W, y_2 = \frac{w}{2}$$

5. Coordinates for the corners of the Right Eye area are calculated by

$$x_1 = RowIndex - 0.06 * W$$

$$x_2 = RowIndex + 0.06 * W$$

$$y_1 = \frac{w}{2}, y_2 = W - (0.05 * W)$$

6. Coordinates for the corners of the Left Eyebrow area are calculated by

$$x_1 = RowIndex - 0.2 * W$$

$$x_2 = RowIndex - 0.05 * W$$

$$y_1 = 0.05 * W, y_2 = \frac{w}{2}$$

7. Coordinates for the corners of the Right Eyebrow area are calculated by

$$x_1 = RowIndex - 0.2 * W$$

$$x_2 = RowIndex - 0.05 * W$$

$$y_1 = \frac{w}{2}, y_2 = W - (0.05 * W)$$

8. Coordinates for the corners of the nose area are calculated by

$$x_1 = RowIndex - 0.05 * W$$

$$x_2 = x_1 + 0.4 * W$$

$$y_1 = 0.3 * W, y_2 = y_1 + 0.4 * W$$

9. Coordinates for the corners of the lip area are calculated by

$$l_x = RowIndex + 0.35 * W$$

$$x_1 = l_x + 0.05 * W$$

$$x_2 = l_x + 0.20 * W$$

$$y_1 = 0.25 * W, y_2 = y_1 + (0.5 * W)$$

Features extracted by using above algorithm corresponding to the images are shown in fig. 3

B. *Illumination Normalization Facial Feature Images*

Most classification methods have used the intensity values of images as the input features of classifier. Edges directionality is visually a prominent feature. In our approach we have taken the gradient feature images of above extracted features for gender classification. These are images which are formed by taking ratio of gradients in horizontal direction and vertical direction. This makes them independent of the illumination conditions and also the required information is get highlighted. The use of these gradient feature images as the input of the classifier yields better classification performance.

Gradient feature images of above extracted features are obtained by using the following procedure [14]

Input: Feature image T

Output: Gradient image of T

1. Gaussian smoothing is performed to the input image by using the Gaussian function

$$K(i,j) = {}^{1}/_{2\pi\sigma^2} \exp\left(-(i^2 + j^2)/2\sigma^2\right)$$

2. Gaussian derivative filters are Constructed in x and y directions. These filters are applied to the images.

3. Gradient image is constructed by using the

$$Gradient = \arctan\left(\frac{T_y}{T_x}\right)$$

Illumination normalization facial features images by using the above algorithm is shown in fig. 4(c).

C. Approaches

We have used two approaches for gender classification. In approach I, we have used the facial feature pixel intensity information for gender classification. In approach II, we have used the gradient facial feature images as a result of illumination normalization for gender classification. Feature vectors are constructed. Facial feature vectors and illumination normalization facial feature vectors are as follows respectively:

$$x = x_{reyebrow}^T x_{leyebrow}^T x_{reye}^T x_{leye}^T x_{nose}^T x_{mouth}^T$$

$$x_g = x_{g_{reyebrow}}^T x_{g_{leyebrow}}^T x_{g_{reye}}^T x_{g_{leye}}^T x_{g_{nose}}^T x_{g_{mouth}}^T$$

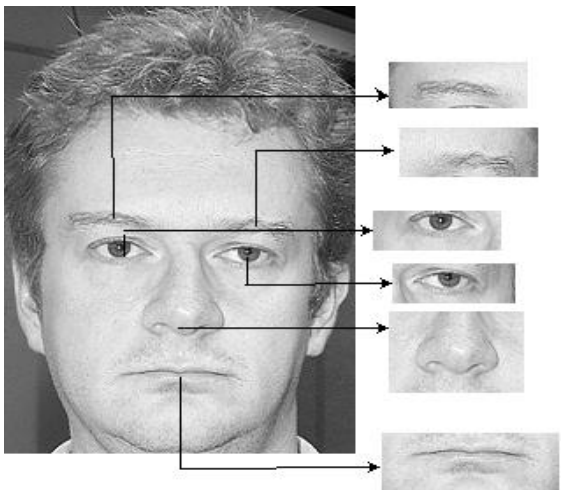


Fig. 3. Features extracted by proposed feature extraction algorithm.

IV. DIMENSION REDUCTION AND CLASSIFICATION

A. Principal Component Analysis

PCA is a technique for removal of redundant features from the image. Reducing statistical redundancy between the components of high-dimensional vector data enables a lower-dimensional representation without significant loss of information. PCA was first applied to reconstruct human faces in the context of image compression by Sirovich and kirby [15]. The Eigen face technique was developed by Turk and Pentland [16].

The size of the above two feature vectors is large therefore Principal Component Analysis technique is used to reduce the dimensionality and to enhance the classification performance.

Assume that x_r where $r = 1, 2 \ldots . n$ are n dimensional input feature vector with mean(μ)

$$\mu = {}^{1}/_{n} \sum_{j=1}^{r} x_j$$

The covariance matrix of x_r is defined by

$$CM = 1/n \sum_{j=1}^{n} (x_j - \mu)(x_j - \mu)^T$$

PCA solves the Eigen value problem of covariance matrix

$$CMv_j = \gamma_j v_j$$

Where v is eigen vector and γ is eigen value.

There is only need to compute the k eigenvectors

$$(\sum_{j=1}^{k} \gamma_j \Big/ \sum_{j=1}^{n} \gamma_j) \geq v$$

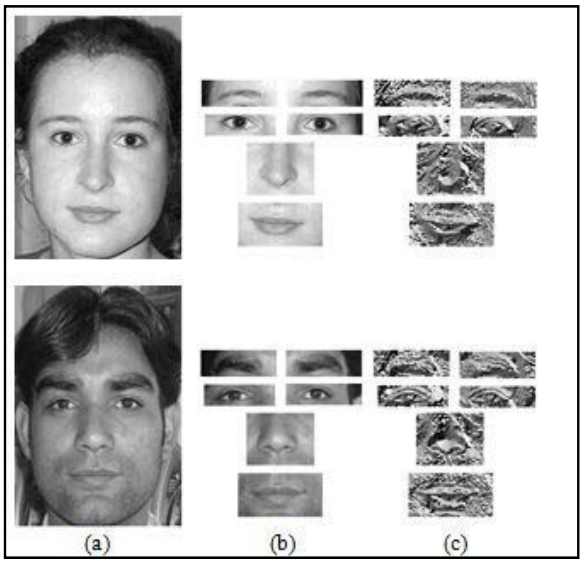


Fig. 4. Two images of which one is female and other is male. b) Extracted features. c) Illumination normalization features.

Let $V = [v_1, v_2 \ldots . . v_k]$ are k eigenvectors selected. Now, the low dimensional feature vector of a new input feature vector x is determined by

$$x_t = V^T x$$

B. Support Vector machine

SVM is one of the best classification technique used in pattern recognition, face recognition. It produces a model based on training data. This model predicts the class label of test data. Training data set is given, each belonging to one of two classes. Then it finds optimal hyper plane such that margin of separation between two classes is maximum as shown in fig. 5. This plane is at maximal distance from the closest members known as support vectors of both the classes. These are the data points that lie closest to decision surface.

Any hyper plane can be written as set of points such that

$$hyperplane = \{wx + c = 0\}$$

We define canonical hyper plane which separates the data from the hyper plane by a distance of at least 1 i.e., we consider those that satisfy

$$wx_i + c \geq +1 \text{ Where } y_i = +1$$

$$wx_i + c \leq -1 \text{ Where } y_i = -1$$

This can be rewritten as

$$y_i(wx_i + c) \geq +1, \forall\, i = 1,2 \ldots \ldots m$$

The margin M is given as $M = 2/||w||$, it is showing that minimization of a norm of a hyper plane normal weights lead to a maximization of a margin M.

To find the optimal separating hyper plane having a maximal margin, the learning problem is to minimize the $1/2\, w^T w$ subject to constrain

$$y_i(wx_i + c) \geq +1, \forall\, i = 1,2 \ldots \ldots m.$$

The optimization problem is solved by Lagrange function as follows

$$F(w, c, \vartheta) = 1/2\, ||w||^2 - \sum_{i=1}^{n} \vartheta_i (y_i(w_i x_i + c) - 1))$$

Where ϑ_i are Lagrange multipliers and we will get w and c as follows

$$w = \sum_{i=1}^{m} \vartheta_i y_i x_i \,, c = -1/2\, w[x_p + x_q]$$

Hence the weight vector is linear combination of training patterns.

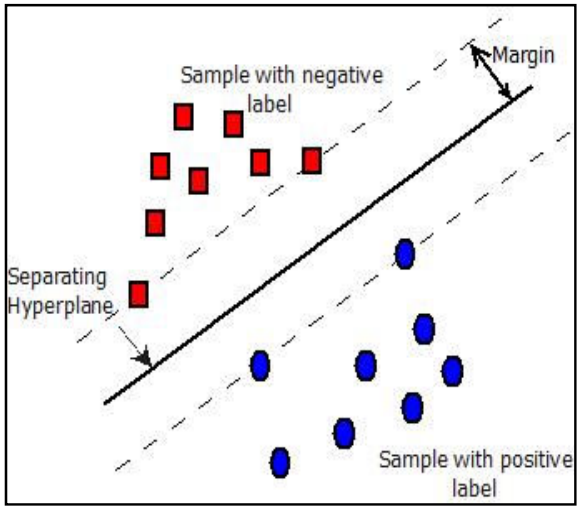


Fig. 5. Linear Classification using SVM

So this procedure leads to the determination of optimal separating hyper plane. The classification for a new data point x is, then

$$f(x) = sign(\sum_{i=1}^{m} \vartheta_i y_i x_i^T x + c)$$

V. IMPLEMENTATION AND EXPERIMENTS

To test the performance of on two approaches of gender classification, we run the experiments on outdoor database as well as CALTECH (California Institute of Technology) database. To implement our system, we use MATLAB.

A. Results on Caltech Database

The Caltech face database contains 450 color images of 26 individuals (17 males and 9 females) with different lightening conditions. This was collected by Markus Weber at California Institute of technology. There are 5 to 29 images of same person. The images are cropped and resized. Fig. 6 shows some cropped and resized images showing variation in illumination and background. We have taken 392 images comprising 241 images of male and 151 images of female that is front and straight profile images. We used two feature vectors and SVM as a classification method to conduct an experiment.

We got the high classification rate when we train the SVM by 66 images. We achieved high classification rate when we used the illumination normalization feature images as an input to the classification as shown in Table 1. We have also analyzed how much individual feature is contributing in classification. Classification rates were shown in Fig. 7 based on individual features and combination of all features. We achieved high classification rate when we took all the features (eye-brows, eyes, nose, and lip) of face.

B. Results on Outdoor Database

To study the robustness of the algorithm based on our proposed approach, we created a database named Outdoor which contains 170 images (70 males and 70 females) captured under varying illuminations. Fig. 8 shows cropped and resized images for ten subjects showing variation in illumination and background. We get the high classification rate when we train the SVM by 70 images.

Fig. 6. Resized and cropped example face images from the CalTech database.

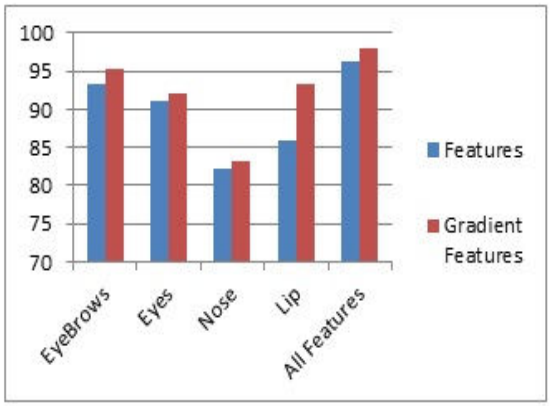


Fig. 7. Classification rates for the individual features and combined features and illumination normalization features of all these features with the Caltech Database images.

Fig. 8. Resized and cropped example face images from the Outdoor database.

Gender classification results are shown in Table 2 when we take the intensity of features as a feature vector as approach I and gradients of features as a feature vector as approach II. Higher classification rate was achieved when gradient features were used. Classification rates were shown based on individual features and combination of all features in Fig. 9. Higher classification rate was achieved when we took all the features (eye-brows, eyes, nose, and lip) of face.

TABLE I

RESULTS FOR THE CALTECH IMAGES WITH REDUCED FEATURE VECTOR AND GRADIENT FEATURE VECTOR

Approach	Classification Method	Classification rate
Approach I	LSVM	96.20%
Approach II	LSVM	**98.00%**

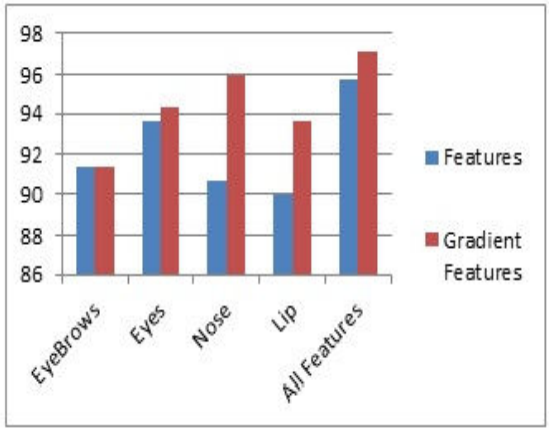


Fig. 9. Classification rates for the individual features and combined features and illumination normalization features of all these features with the Outdoor Database images

TABLE II

RESULTS FOR THE OUTDOOR IMAGES WITH REDUCED FEATURE VECTOR AND GRADIENT FEATURE VECTOR

Approach	Classification Method	Classification rate
Approach I	LSVM	95.71%
Approach II	LSVM	**97.14%**

VI. CONCLUSION

In this paper, we proposed an improved approach for facial feature extraction and these facial features are used as gender classification instead of full face image. For extracting the face features, we used the fact that eye will lie in mid-way if faces are cropped from hair to chin and left ear to right ear, then by using the symmetry property of features, other features are extracted. We compared the methods when intensity value of face features and illumination normalization features were used as input to the Linear Support Vector Machine (LSVM) classifier. We used two databases, namely the CalTech database and Outdoor image database to compare the methods. We found that when illumination normalization feature were used, the high classification rate was achieved. It justifies the appropriateness of our proposed methodology.

REFERENCES

[1] H. Abdi, D. Valentin, B. Edelman, and A. J. OToole, "More about the difference between men and women: Evidence from linear neural networks and the principal component approach, Perception", vol. 24,pp.539-562, 1995.

[2] Cottrell, G.W., Metcalfe, J., 1990. "EMPATH: Face, emotion, and gender recognition using holons". In: Lippmann, R., Moody, J.E., Touretzky, D.S. (Eds.), Proc. Advances in Neural Information Processing Systems 3 (NIPS). Morgan Kaufmann, pp. 564-571.

[3] Golomb, B.A., Lawrence, D.T., Sejnowski, T.J., 1990. "SEXNET: A neural network identifies sex from human faces". In: Lippmann, R.,Moody, J.E., Touretzky, D.S. (Eds.), Proc. Advances in Neural Information Processing Systems 3 (NIPS). Morgan Kaufmann, pp. 572-579.

[4] Wiskott, L., Fellous, J.-M., Krger, N., von der Malsburg, C., 1997. "Face recognition by elastic bunch graph matching". In: Sommer, G.,Daniilidis, K., Pauli, J. (Eds.), 7th International Conference on Computer Analysis of Images and Patterns, CAIP'97, Kiel. Springer-Verlag, Heidelberg, pp. 456-463.

[5] Tamura, S., Kawai, H., Mitsumoto, H., 1996. "Male/female identification from 8 to 6 very low resolution face images by neural network". PatternRecognition 29 (2), 331-335.

[6] Lyons, M., Budynek, J., Plante, A., Akamatsu, S., 2000. "Classifying facial attributes using a 2-d Gabor wavelet representation and discriminant analysis". In: Proc. Internat. Conf. on Automatic Face and Gesture Recognition (FG00), IEEE, Grenoble, France, pp. 202-207.

[7] G. Shakhnarovich, P.A. Viola, and B. Moghaddam, "A Unified Learning Framework for Real Time Face Detection and Classification", Proc. Int'l Conf. Automatic Face and Gesture Recognition, pp. 16-26, 2002.

[8] E. Makinen and R. Raisamo, "Evaluation of Gender Classification Methods with Automatically Detected and Aligned Faces", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30, no. 3, pp. 541-547,Mar. 2008.

[9] E. Makinen and R. Raisamo, "An Experimental Comparison of Gender Classification Methods", Pattern Recognition Letters, vol. 29, no. 10, pp.1544-1556, July 2008.

[10] S. Halder, D. Bhattacharjee, M. Nasipuri, D. K. Basu and M. Kundu "Face Synthesis (FASY) System for Determining the Characteristics of a Face Image", Proceedings of National Seminar on Recent Advances on Information Technology (RAIT-2009), ISM, Dhanbad, INDIA, February2009.

[11] S. M. Pizer and E. P. Amburn, "Adaptive histogram equalization and its variations", Comput. Vis. Graph., Image Process. vol. 39, no. 3, pp.355-368, 1987.

[12] S. Shan, W. Gao, B. Cao, and D. Zhao, "Illumination normalizationfor robust face recognition against varying lighting conditions", in Proc.IEEE Workshop on AMFG, 2003, pp. 157-164.

[13] M. Savvides and V. Kumar, "Illumination normalization using logarithm transforms for face authentication", in Proc. IAPR AVBPA, 2003, pp.549-556.

[14] T. Zhang, Y. Y. Tang, Z. Shang, and X. Liu, 2009, "Face Recognition Under Varying Illumination using Gradient faces", IEEE Transactions,vol.18, no.11, pp.2599 - 2606.

[15] L. Sirovich and M. Kirby. "Low-dimensional procedure for the characterization of human faces". Journal of Optical Society of America, 4(3):519–524, 1987.

[16] M. Turk and A. Pentland. "Eigenfaces for recognition". Journal of Cognitive Neuroscience, 3(1):71–86, 1991.

◈◈◈

Review Paper on Union-free Regular Languages

Sukhpal Singh Ghuman & Ajay Kumar

CSED, Thapar University, Patiala, Punjab, India

Abstract - Regular expressions consist of union, kleene closure and concatenation operations on alphabets. Every non-union-free regular expression can be converted into an equivalent union-free regular expression, but it may not be unique. A regular language is said to be union-free if it can be represented by the equivalent regular expression which does not contain any union operation or it can be represented by finite union of union-free regular expressions.

Keywords- Deterministic finite automata; regular expression; component; union width; union-free regular language;

I. INTRODUCTION

The regular expressions are well known in the field of computer science. They are commonly used and well-applicable in theory as well as in practice. The regular languages are used in compilers, programming languages, pattern recognition, protocol conformance testing etc. [2].

A regular language consisting of concatenation, kleene closure and union operations refers to sequential continuation, loops and branching respectively [4]. In this context, union-free regular languages represent sequences of operators that do not contains conditional transitions [4]. The union-freeness of languages accelerates the reversal operation on regular language.

II DEFINITIONS AND NOTATIONS

The regular languages can be described by regular expressions. If r is a regular expression, then the regular language corresponding to r is L(r) [3]. Union of two regular languages L_1 and L_2 consists of all the strings which are either in L_1 or L_2 [1].

Def. 1: A regular expression over input alphabets Σ can be defined as [6]:

1. Every input alphabet can be represented by itself.
2. Null language and null string also represent by themselves.
3. If r_1 and r_2 are regular expressions representing the languages l_1 and l_2 respectively, then:

   3.1 Union of r_1 and r_2 is represented by $r_1 + r_2$.

   3.2 Kleene closure of the regular expression is represented by $(r_1)^*$.

   3.3 Concatenation of r_1 and r_2 is represented by r_1r_2.

   Rule 3 can be defined recursively.

Example 1: The given regular expression a+b* denotes the set of all strings {ε, a, b, bb, bbb, bbbb, ...}.

Def. 2: A deterministic finite Automata (DFA) [6] M is defined by the five tuple (Q, Σ, δ, q_0, F), where Q is a finite non empty set of internal states, Σ is a finite non empty set of symbols called the input alphabet, δ: Q X Σ → Q is called as transition function, q_0 is the initial state, F⊆Q is a set of final states. DFA can be represented by transition diagram or transition table. Figure 1 represents the transition diagram for accepting even number of zeros over the alphabet Σ = {0,1}.

III. UNION OPERATION ON REGULAR LANGUAGES

Given two languages L_1 and L_2, the union [1] of these regular languages is the combination of the strings which are contained in both the languages, represented by $L_1 \cup L_2$. The resultant language accepts all strings from L_1 and L_2.

Component [4] of a regular expression is the individual string which is represented by an empty string or by the alphabets, concatenation operator and the star operator.

Example 2: Given a language L represented by regular expression = {b + ba* +b*b} consists of three components.

The minimum number of components which are there in the representation of a regular expression is called as the union width [4] of the regular language and

correspondingly a decomposition which may not be unique is known as minimal union-free decomposition of the regular language.

Example 3: The regular expression {ε} + a*ba* + b*ab* has three components and hence the union width is 3.

IV UNION-FREE REGULAR LANGUAGE

A regular language is said to be union-free [4,5] if it can be represented by a regular expression without the using union operation or by union of union-free regular languages

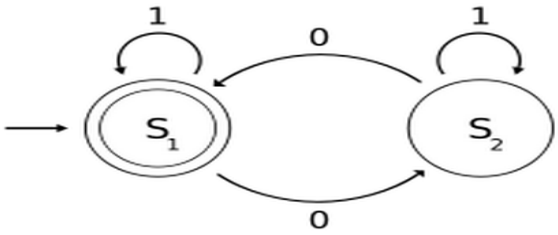


Figure 1.Transition diagram of DFA for even number of 0's over {0,1}

Example 4: The language represented by the regular expression (a + b*)*, can be converted to an equivalent union-free regular language represented by (a*b*)*.

Union-free regular languages are closed under reversal and concatenation operations [2, 4].

Union-free decomposition [5] of a non-union-free regular language corresponds to the decomposition of the language which can either be represented by the regular expression which is free from the union operation or the language consist of finite union of union-free regular languages. The resultant language may not be unique.

Example 5: The language (a + b)* can be decomposed into (a*b*)* or it can be also be written as {ε} + a*ba* + b*ab*.

V. UNION-COMPLEXITY OF REGULAR LANGUAGES

Regular expressions can be represented by expression tree (tree diagram) or by flow diagram (syntax graph) [2]. A possible normal form [2] of regular expressions can be represented using union-free regular expressions. Normal form is equivalent to the finite union of union-free regular languages. Union-complexity can be defined using the normal form.

Tree diagram in figure 2 represents a regular expression. Root and children in the tree diagram consists of the operators and terminals [2]. Tree diagram also represents corresponding union-free decomposition of regular expression. Left tree represents the regular expression (a+b)*, which is not union-free. The right tree represents the equivalent decomposition of the regular expression.

Example 6: A regular expression (a+b)* is not a union-free regular expression. It can be decomposed into an equivalent union-free regular expression (a*b*)*.

Although after decomposition, depth of the tree increases, but the union complexity decreases as the union complexity can be determined from a tree by counting the sub trees having the union operator [2].

Syntax graph used to represent regular expressions consists of nodes and edges [2]. For simpler regular expressions there are simple diagrams in which there is exactly one arrow inside the node and exactly one arrow out of the node. The graphical form of these regular operations can be seen in the figure 3. The terminals are represented by the ellipses.

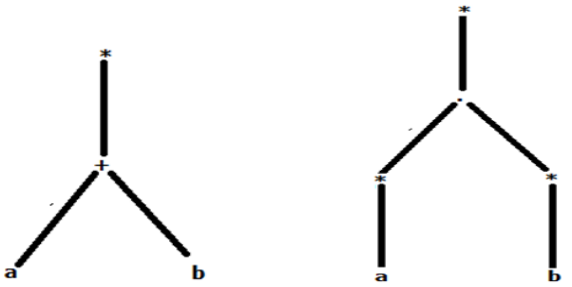


Figure 2. Rewriting of regular expressions into union-free regular expression [2].

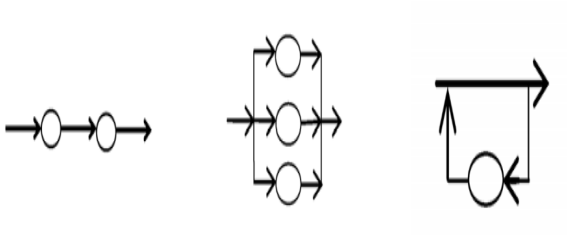


Figure 3. Three basic operations of flow diagrams: concatenation (left), alternative (middle) and iteration (right) [2].

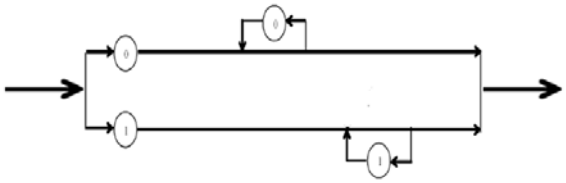


Figure 4. Syntax graph for regular expression 0(0*) + 1(1*) [2].

The first part of the figure represents concatenation operation on the symbols. The second part of the figure represents union operation on the symbols. The last part of the figure represents the iteration operation of the symbols.

Def. 3: A regular expression is said to be in normal form if it can be represented by union of union-free regular languages [2].

Example 7: A regular expression ab*+(ab)*a*+bb*b is in the normal form as it is represented by the union of union-free regular expressions.

Example 8: For the regular expression 0(0*) + 1(1*), there are two paths. Suppose a path starting with 0 is chosen. After this only the iterative part is to be traversed, depending upon the number of iterations to be traversed.

VI CONCLUSIONS AND FURTHER WORK

A regular language is said to be union-free if it can be represented by a regular expression without using the union operations or union of union-free regular languages. The regular language which is not union-free can be decomposed into an equivalent union-free regular language. The resultant decomposed language may not be unique. Union complexity of regular expressions can be determined easily by union-free regular expressions.

Following are some open problems on which work can be done regarding the union freeness of regular languages:

1. An algorithm design for determining whether a regular language is union-free or not.
2. Design an algorithm for converting a regular language into an equivalent union-free regular language.

REFERENCES

[1] Mishra K.L.P. and N. Chandrasekaran, "Theory of Computer Science (Automata Language and. Computation)", PHI, Second edition, 1998.

[2] Nagy, B, "On Union-complexity of Regular Languages", 11th IEEE International Symposium on Computational Intelligence and Informatics Hungary, November, 2010.

3] Peter Linz,"An Introduction to Formal Languages and Automata", Narosa publishers fourth edition, 2009.

[4] Sergery Afonin and Denis Golomazov, "On Minimal Union-free Decompositions of Regular languages", Third International conference, Lata, Spain, 83-93, 2009.

[5] Sinisa Crvenkovic , Igor Dolinka and Zoltan Esik, "On Equations for Union-Free Regular Languages", Information and Computation, 2001.

[6] Ullman, J., J. E. Hopcroft and R. Motwani, "Introduction to Automata Theory, Languages and Computation", Pearson Education Inc, 2001.

◈◈◈

WIRELESS EURYNOMUS
A Wireless (802.11) Probe Request Based Attack

Hitesh Choudhary & Pankaj Moolrajani

MIT (California), Wireless Researcher , Igneustech , Cyber Security Expert, RHCE Jaipur, India

Abstract - The 802.11 standard for wireless includes a procedure of sending a continuous probe request to find out if there is a nearby Access Point available or not. Most of wireless clients like laptops or smart phones are configured to auto-connect. Due to this auto connect feature any client can be compromised disregarding the fact that Access Point is nearby or not. Even if the client is connected to legal Access Point, a simple disassociation packet in 802.11 headers can be helpful for an attacker to leak your personal as well as company's top secret information.

Keyword *– probe request; probe response; beacon frames; virtual wireless routers; wireshark; wireless*.

I. INTRODUCTION

In the recent years, the proliferation of laptop computers and smart phones has caused an increase in the range of places people perform computing. At the same time, network connectivity is becoming an increasingly integral part of computing environments. As a result, wireless networks of various kinds have gained much popularity. But with the added convenience of wireless access come new problems, not the least of which are heightened security concerns. When transmissions are broadcast over radio waves, interception and masquerading becomes trivial to anyone with a radio, and so there is a need to employ additional mechanisms to protect the communications.

In this article we want to focus on some of the hidden flaws that were never taken seriously. Auto-connect is a simple and one of the most conniving facility provided by all the clients of wireless Access Point. This feature can also be used to compromise a client and the attack is counted as one of the deadly silent attacks.

II. TARGET AUDIENCE

All though this attack can affect any of the technical and non technical user of 802.11 interface. But the technical details of this attack will contain some the advance details including the usage of wireshark, little understanding of packet details over wireless and some of the details about the probe and beacon frames that are disused in this article.

III. SCOPE OF ATTACK

This attack is almost new born to the world of wireless and internet. This attack is fully capable of creating an intermediate connection between any client and attacker. Talking about the scope of this attack, it can be of wide variety. For an example if an attacker walks into a company premises and just by monitoring the air, he can easily find out the probes in air and can attack into the any laptop or he can attack to any smart phones and can collect contact details of clients. This is just a simple scenario; cases can be like T.J maxx credit card incident.

IV. HARDWARE AND SOFTWARE REQUIREMENTS

To perform this attack, we'll be needing a entire lab setup with some tools as a software requirements and some as a hardware requirements.

Hardware requirements include:

[1] Access point

[2] 2 laptop (1 as attacker and 1 as victim)

[3] Wireless card (internal or external)

[4] 1 smart phone (not an extreme requirement)

Software requirement

[1] Backtrack operating system (4-revision2 or higher version).

[2] All other required tools are preconfigured in it.

V. UNDERSTANDING PROBES AND BEACONS

When a client turns on its wireless interface, at the same moment the wireless interface starts to send many probe request to find if there is an access point is available or not. On the other hand similarly access point is also sending the beacon frames to show its presence. Once when the client get connects to an access point, there is a facility provided by different machines to remember that access point and whenever next time when clients came into the range, it automatically gets connected. This is simply because the client is continuously sending probe requests in the air to find if any saved AP is available.

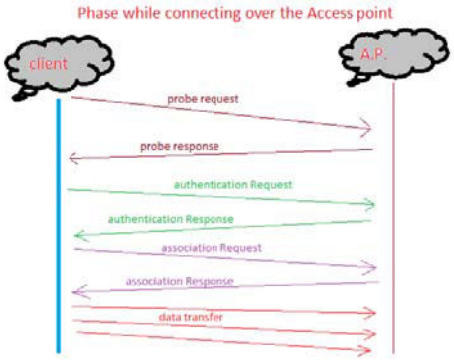
VI. TYPES OF ATTACKS

[1] IP level connectivity attacks (Metasploit based)

[2] Relay the packets to AP (MITM based attacks)

[3] Depending upon the usage, attacks can be integrated and client is still unknown.

VII.ATTACKS SCENARIO



To understand this attack, first of all scenario and little knowledge about working of the Access Point must be clear. So, what we are trying to implement is, A client who is not connected to any wireless AP and having his wireless interface up and running always transmits some probe request from its PNL i.e. Preferred Network List. It is just a sense of insecurity and shocking thing that it is independent of any AP.

First of all we'll try to make a monitor mode interface in the air, which can accept all the packets over the air regardless if the packet is destined for it or not. This is very similar to promiscuous mode over the wired network, used for the purpose of sniffing.

After finding the probes of the clients, we'll create a soft AP or known as virtual AP. A soft access point is created by a set of software which continuously sends out the beacon frames to show all nearby clients about its presence. Since the client is already attempting to connect to that access point. It will automatically connect to the attacker. Now, if a DHCP is running over the attacker it will automatically receive an IP or if there is no DHCP is running then client will receive an IP of the range 169.xxx.xxx.xxx will sent gratuitous packets. Once the IP is assigned, the tap interface created by soft AP, can have IP level connectivity with the client and the best part is, client is still unaware and will remain.

VIII. IMPLEMENTATION

We have used a BackTrack machine (attacker) and a I-Phone (victim) to implement our attack scenario. A monitor mode interface is being created at the top of a wireless interface, this monitor mode interface can be easily created by using airmon-ng set of tools. Keep noting that the wlan0 (wireless) interface is up and running.

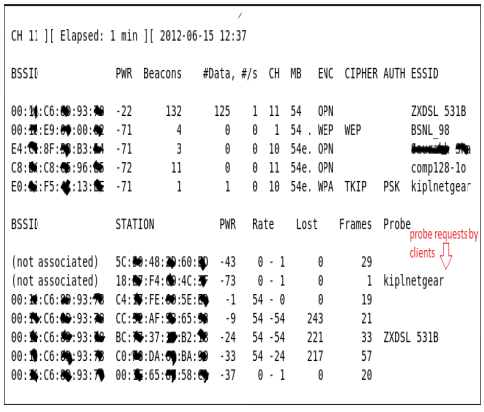
```
root@bt:~# airmon-ng start wlan0


Found 2 processes that could cause trouble.
If airodump-ng, aireplay-ng or airtun-ng stops working after
a short period of time, you may want to kill (some of) them!

PID     Name
1045    dhclient3
1656    dhclient3
Process with PID 1656 (dhclient3) is running on interface wlan0


Interface       Chipset         Driver

wlan0           Realtek RTL8187L        rtl8187 - [phy0]
                                        (monitor mode enabled on mon0)

root@bt:~#
```

airmon-ng start wlan0

Monitor mode enabled on mon0 indicates that the monitor mode has been created and now we can monitor the air.

To monitor the air, simply airodump can be used over the mon0 interface. This along with the AP will also give the details of the clients that are associated or trying to associate with the network in the surroundings

airodump-ng mon0



After finding the probe request name, attacker can easily create a soft AP or virtual access point with any of the bssid as well as any essid. Here I have used a essid of name Hitesh just for the sake of example.

airbase-ng -a <bssid> -e <essid/name> mon0

Airbase set of tools have got a lots of options, it can send response to any of the probe request that client is transmitting via its radio but for the sake of simplicity we have used this scenario .

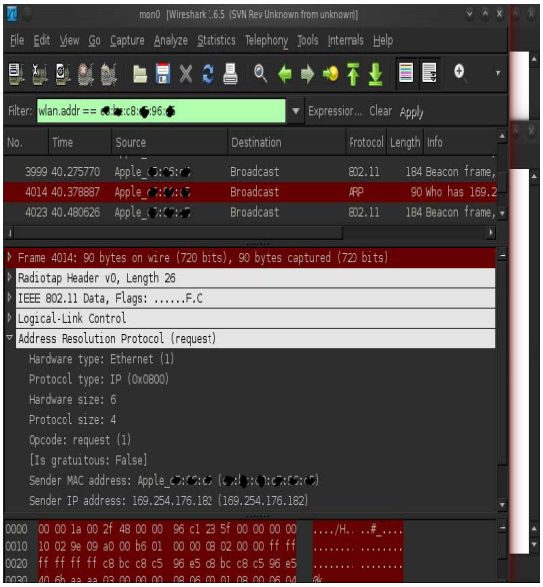
The interesting thing about this soft AP is that it also creates a tap interface. Its little basic that our access point always have 2 cards in it, one is wireless and other is for wired interface. This tap interface is the same clone of wired interface named as at0.

```
root@bt:~# airbase-ng -a 00:[illegible]:93:[illegible] -e hitesh  mon0
12:47:37  Created tap interface at0
12:47:37  Trying to set MTU on at0 to 1500
12:47:37  Access Point with BSSID 00:[illegible]:8D:[illegible] started.
```

As a result of this client will automatically get connect to this “hitesh” Ap. And since there is no DHCP is running over the attacker machine.

Client will get an IP address of the range 169.xxx.xxx.xxx and will try to send gratuitous packets. One can also use these packets as an ARP packet to sent it back to IP. So, there is can be attack at every phase.

One can also verify this by using wireshark and capturing each and every packet. These packet will show that client is again and again trying to send DHCP request and failing to that finally it is getting an IP range 0f 169.xxx.xxx.xxx.In the mean while one can also set a DHCP and can easily transfer the packets to the internet via its bridge interface and can perform Man In The Middle Attacks.



Now the final step is to just up the at0 interface and set the ip of the same range and same subnet that can be easily done with the ipconfig utility

ifconfig -a

```
root@bt:~# ifconfig -a
at0       Link encap:Ethernet  HWaddr 00:c0:ca:54:0f:c5
          BROADCAST MULTICAST  MTU:1500  Metric:1
          RX packets:98 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:500
          RX bytes:20993 (20.9 KB)  TX bytes:0 (0.0 B)
```

Finally the proof of the IP level connectivity, Post that one can easily launch some Metasploit modules or other various set of attacks.

ifconfig at0

ping 169.254.28.3

```
root@bt:~# ifconfig at0
at0       Link encap:Ethernet  HWaddr 00:c0:ca:54:0f:c5
          inet addr:169.254.28.8  Bcast:169.254.28.255  Mask:255.255.255.0
          inet6 addr: fe80::2c0:caff:fe54:fc5/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:149 errors:0 dropped:0 overruns:0 frame:0
          TX packets:15 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:500
          RX bytes:32545 (32.5 KB)  TX bytes:1014 (1.0 KB)

root@bt:~# ping 169.254.28.3
PING 169.254.28.3 (169.254.28.3) 56(84) bytes of data.
64 bytes from 169.254.28.3: icmp_seq=1 ttl=255 time=12.5 ms
64 bytes from 169.254.28.3: icmp_seq=2 ttl=255 time=7.32 ms
64 bytes from 169.254.28.3: icmp_seq=3 ttl=255 time=8.05 ms
64 bytes from 169.254.28.3: icmp_seq=4 ttl=255 time=8.05 ms
64 bytes from 169.254.28.3: icmp_seq=5 ttl=255 time=8.08 ms
64 bytes from 169.254.28.3: icmp_seq=6 ttl=255 time=7.34 ms
^C
--- 169.254.28.3 ping statistics ---
6 packets transmitted, 6 received, 0% packet loss, time 5037ms
rtt min/avg/max/mdev = 7.328/8.563/12.523/1.802 ms
```

ACKNOWLEDGMENT

Acknowledgment to Igneustech for providing appropriate equipment's and lab environment.

REFERENCES

[1] www.aircrack-ng.org

[2] www.aircrack-ng.org/doku.php?id=airodump-ng

[3] www.wireshark.org/

[4] Interception Mobile Communications, The Insecurity of 802.11 – ISAAC www.isaac.cs.berkeley.edu/isaac/mobicom.pdf

[5] An Overview of 802.11 Wireless Network Security Standards & Mechanisms By: Luis Wong (posted on January 19, 2005)

[6] Remote Access Point/IDS By: Jared Kee (posted on April 10, 2012).

◈◈◈

Empirical Model For The Fault Prediction Using Regression

Rekha Singh

Information Technology, USIT, Delhi

Abstract - In practice the software quality can only predicted with the estimate of number of defect found. This paper presents the empirical analysis of the dataset i.e. software metrics for building the model. For that correlation between the dataset calculated and analyzed. The model has been developed with the help of statistical technique stepwise regression. The dataset divided into two parts. First part is used to build the model while second part is used for the validation of the model and it is found that model build shown 5% better result over the other.

Keywords- *Emperical analysis, fault prediction, Prediction model, Software metrics.*

I. INTRODUCTION

The software quality can be measure in terms of number of defects or number of fault found in the software product. The quality software product assures the lesser number of defects found in the software after the release of the client.

The software quality can be improved by applying the model that gives estimation of the no of the faults in the software product.

The fault prediction model can be classified into two parts First one predicts the software products reliability from the design parameter perspective while another one predicts software products reliability from the test dataset.

The fault prediction model should be accurate and efficient in order to comment and predict the software quality.

Defective software module causes software failure and increase the cost of the projects. For increasing the software quality software metrics are used. It is very difficult to comment which software metrics will be best suit for the prediction of fault.

II. RELATED WORKS:

Wide range of software prediction model has been proposed. Among them some of the model predicting the software quality using Rayleigh function [1] and the following steps are followed: First collection of data from the past projects. Second using linear regression estimate of the total no of defect. Third compute the latent defects. Fourth calculate estimate defects rate by Rayleighfunction for each week. Fifth calculate the estimate defects rate phase based on the projects schedule. Sixth the Rayleigh functions for the defect prediction pattern. Seventh comparison of Rayleigh curve and actual data.

While another [2] focusing on building a fault prediction model using decision tree in combination with k-means clustering preprocessing technique.Koru and Liu [3] build a model using WEkA an open source machine learning tool and dataset collected from NASA product of the project CM1, JM1, KC1, KC2 and PC1 datain promise repository. T Nguyen and Tie N. Nguyen[4]use topic modeling to measure the issue in the code and use them as a input for machine learning-based defect prediction models.And the result of Eclipse, JDT shows that topic based metrics have high correlation to the number of bug.

III. DATA COLLECTION:

The dataset used in the study has been taken from the Promise Repository donated by Marian Jureczko and the data set collected from the open source the tomcat server project by the tool ckjm and buginfo.

IV. METHODOLOGY USED:

Linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable

Before attempting to fit a linear model to observed data, we should first determine whether or not there is a relationship between the variables of interest.

A linear regression line has an equation of the form

$$Y = a + bX,$$

where **X** is the explanatory variable and ***Y*** is the dependent variable. The slope of the line is ***b***, and ***a*** is the intercept.

V. MODEL DEVELOPMENT:

Correlation and regression analysis are related in the sense that both deal with relationships among variables. The correlation coefficient is a measure of linear association between two variables. Values of the correlation coefficient are always between -1 and +1. A correlation coefficient of +1 indicates that two variables are perfectly related in a positive linear sense a correlation coefficient of -1 indicates that two variables are perfectly related in a negative linear sense, and a correlation coefficient of 0 indicates that there is no linear relationship between the two variables.

What we want is the best subset to do the job of prediction as well as possible with as few variables as possible. Trying out all possible subsets would be tedious and time consuming, but at the end would have difficulty sorting out the best from your long list. One solution is to use a stepwise method.

Stepwise regression is a process of building a model by successively adding or removing variables based solely on the t-statistics of their estimated coefficients.

The correlation between variables is as shown in the table1:

Table 1:

		bug
Pearson Correlation	bug	1.000
	wmc	-.037
	dit	.017
	noc	-.038
	cbo	.472
	rfc	-.070
	lcom	-.043
	ca	.778
	ce	.
	npm	-.016
	lcom3	.016
	loc	-.085
	dam	.099
	moa	.101
	mfa	.186
	cam	-.087
	ic	-.084
	cbm	-.058
	amc	-.089
	max_cc	-.073
	avg_cc	-.046

The next step is model summery and ANOVA output of regression applied between independent and dependent variables shown in below table.

Table 2:

Model Summary[e]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.859[d]	.738	.716	.073	2.019

d. Predictors: (Constant), ca, lcom, mfa, moa

e. Dependent Variable: bug

The model output value is 73.80% and adjusted R2 value is 71.60 which is satisfactory.

ANOVA for the model is shown below in the table 3:

Table 3 :

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.724	4	.181	33.720	.000[e]
	Residual	.258	48	.005		
	Total	.981	52			

a. Dependent Variable: bug

e. Predictors: (Constant), ca, lcom, mfa, moa

Always look for goodness of fit lower this number better is this.

Its .000 i.e. model is statistically significant at the confidence interval of 99%.

Statistical significance of independent variables is shown in the table 4.

Table 4:

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.060	.015		-4.062	.000
	ca	.019	.002	.852	11.090	.000
	lcom	.000	.000	-.342	-4.042	.000
	mfa	.064	.021	.225	2.987	.004
	moa	.010	.005	.183	2.183	.034

a. Dependent Variable: bug

In above table the sig is less than 0.05and we can assume that the estimate in column B is asserted as true with 95% level of confidence.

Bug=-.060+ca*.019+mfa*.064+moa*.010.

VI. MODEL VALIDATION

This section of the paper is used to assess, how well the model is built in the previous section. Forthis dataset has been divided into the two parts the first part is used to build the model and the rest part of the dataset used to validate the same model.

Here is the model built with rest half of the dataset.

Table 5:

Model Summary[e]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.937[d]	.878	.868	.098	1.959

d. Predictors: (Constant), noc, ca, lcom, moa

e. Dependent Variable: bug

And here R2 value of bug is 87.8% which is better.

And the next is model built with calculated value of the bug with the equation of bug i.e.val_cal_bug.

Table 6:

Model Summary[d]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.961[c]	.924	.919	.054	2.372

c. Predictors: (Constant), ca, noc, cam

d. Dependent Variable: val_cal_bug

And here R2 value of val_cal_bug is 92.4 which is quite better.

VII. CONCLUSION AND FUTURE WORK

The paper has developed model which gives estimation of the bug related to the project which helps in fault prediction .The software metrics used in the model are moa (measure of aggregation), ca (Afferent coupling) and mfa (Measure of functional abstraction). It is found that there is approximate 5% better result shown by the model.

More sophisticated model can be developed in future with the large scale study of dataset collected from industrial projects. The future research may also focus on wider acceptance of prediction model for more generalize projects.

REFERENCES:

[1] Ana Vladu, Sergiu Iliescu, and Iona Fagarasan. "Product Defect Prediction Model".6th IEEE international symposium on computational od applied computational intelligence and informatics. May19-21,2001.

[2] A. Güneş Koru and Hongfang Liu"Building Effective Defect-Prediction Models in Practice. Vol. 22, No. 6. November/December 2005 IEEE software.

[3]Parvinder S. Sandhu,Amanpreet S. Brar,Raman Goel"A Model for Early Prediction of Faults in Software Systems". Volume 4. C. 978-1-4244-5586-7, 2010 IEEE.

[4]. Tu Honglei1, Sun Wei1, Zhang Yanan1"The Research on Software Metrics and Software Complexity Metrics". 978-0-7695-3930,2009 IEEE.

[5] T. Zimmermann, N.Nagappan, H Gall, E Giger, B. Murphy "CROSS-PROJECTDEFECT PREDICTION". August 24-28,2009,ACM 978-1-60558-001.

[6] Marian Jureczko"Significance of Different Software Metrics in Defect*Prediction*" Institute of Computer Engineering, Control and Robotics. Wrocław University of technology.Wybrzeze Wyspanskeiego 27, 50-370,Wroclaw-Poland.

[7] Marian Jureczko "Towards identifying software project clusters with regard to defect prediction". Institute of Computer Engineering, Control and Robotics. Wrocław. ACM Sep 12-13,2010.

[8] T T. Nguyen1, T N. Nguyen1, T M Phuong2."Topic-based Defect Prediction (NIER Track)".International Conference on Software Engineering (ICSE 2011).May 21-28, 2011 Waikiki, HI, USA.

◈◈◈

Performance Evaluation of Adaptive Modulation Techniques For Wimax Physical Layer with Different Cyclic Prefix

Sonam Shakya, L.D. Malviya & S.V. Charhate

Dept. of Electronics and telecommunication, Shri Govindram Seksaria Institute of Technology and Science (SGSITS) 23 Park Road Indore (M.P), India

Abstract – WiMax supports link adaptation techniques known as adaptive modulation and coding (AMC) in which the modulation scheme changes depending on channel conditions. Using adaptive modulation scheme, Wimax system can switch to the highest order modulation depending on the channel conditions. Thus AMC technique helps to reduce the time selective fading. Adaptive modulation technique uses the concept of cyclic prefix that adds additional bits at the transmitter end to the symbol transmitted. The cyclic prefix helps to minimize the inter symbol interference and to improve the bit error rate. In this paper, the physical layer performance of Wimax on the basis of bit error rate w.r.t signal to noise ratio by varying cyclic prefix for modulation techniques (QPSK-1/2, QPSK-3/4, 16-QAM-1/2, 16-QAM-3/4) is done. All these parameters are discussed with AWGN channel Model. The BER curves are used to compare the performance of different modulation techniques by changing cyclic prefix.

Keywords - *Worldwide interoperable microwave access(Wimax); Line-of-Sight (LoS); Physical layer(PHY), Quality of service(QoS), Orthogonal frequency division multiplexing(OFDM),Adaptive Modulation and coding(AMC).*

I. INTRODUCTION

Worldwide Interoperability for Microwave Access (WiMAX) is an emerging global broadband wireless system based on IEEE 802.16 standard. WiMAX is a new OFDM based technology and promises to combine high data rate services with wide area coverage in frequency range of 10 – 66 GHz (Line of sight) and 2 - 11 GHz (Non Line of Sight) and large user densities with a variety of Quality of Service (QoS) requirements .The WiMAX standard air interface includes the definition of both the medium access control (MAC) and the physical (PHY) layers for the subscriber station and base station [3].

The WiMAX/IEEE 802.16 WirelessMAN-OFDM standard supports multiple physical specifications which increase the flexibility of PHY layer and enables the system designers to tailor their system according to their requirements.The PHY specifies some mandatory features to be implemented with the system including some optional features to provide a reliable end-to-end link. WiMax PHY layer uses OFDM which supports multi users on system. In order to meet the requirements of increased spectral efficiency and higher data rate applications in both fixed and mobile environments, the IEEE 802.16d standard provides powerful tools at PHY layer level to support flexibility and efficiency over a range of different applications and environments with higher QoS. One of such tool is of Adaptive Modulation and coding technique. The basic idea of Adaptive Modulation is to adapt different order modulations that allow sending more bits per symbol and thus achieving higher throughputs or better spectral efficiencies. One of the important mechanisms is combination of modulation schemes with forward error correcting codes (FEC). Wherein, a high transmission speed is provided by a high order modulation scheme but it makes more susceptibility to interference. FEC builds redundancy into the transmission by repeating some of the information bits, so bits that are missing or in error can be corrected at the receiving end and helps to reduce latency by cutting down the retransmissions. Thus, adaptive modulation allows the system to overcome fading and other interference and improve QoS. [8][9][10]

II. PHY LAYER MODEL OF WIMAX

The PHY layer in IEEE802.16d relies on OFDM .It is a multi-carrier modulation technique that exhibits excellent robustness against the impairments introduced by the transmission channel. An OFDM symbol is a sum of multiple orthogonal carriers (the complex exponentials from equation (1)), modulated by the data symbols to be transmitted through the channel (X[*k*]) [1]:

$$s(t) = \left(\sum_{k=0}^{N-1} X[k].\, e^{j.k.2.\Delta f.t}\right) p_T(t) \qquad (1)$$

pT (*t*) is a sliding window of duration *T* which localizes the signal on the time axis and gives the duration of an

symbol ($T = 1 \Delta f$ in order to achieve subcarriers orthogonality). If the digital symbols to be transmitted $X[k]$ came from different users, then OFDM transforms into OFDMA This requires perfect synchronization of the system, but brings an additional dimension of multiple access too.

The block diagram of physical layer of Wimax is shown below[6]:

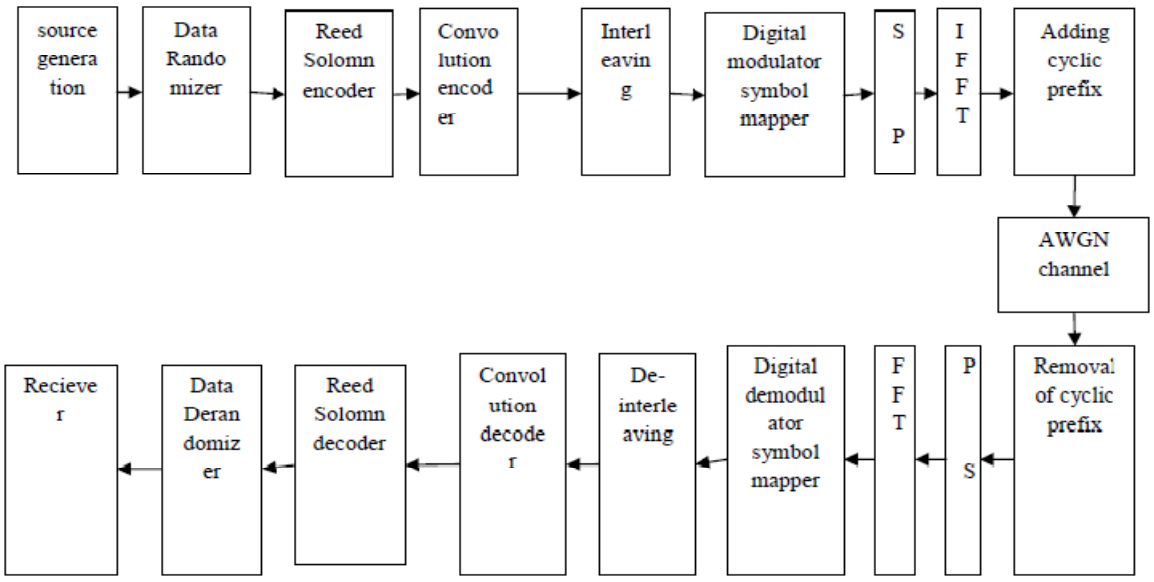


Fig. 1: OFDM based physical layer block diagram of Wimax with cyclic prefix

This structure corresponds to the physical layer of the WiMAX/IEEE 802.16 WirelessMAN-OFDM air interface. In this setup The input binary data stream obtained from a source generator .The signal is ensured against errors with forward error correction codes (FECs) and interleaved The complementary operations are applied in the reverse order at channel decoding in the receiver end.[4] The complete channel encoding setup is shown in above Figure 1.

The forward error control (FEC) consists of a Reed-Solomon (RS) outer code and a rate compatible Convolution Code (CC) inner code. A block Reed Solomon code based on the Galois field GF (28) with a symbol size of 8 bits is chosen which can correct up to 8 symbol errors calculating 16 redundant correction symbols. Reed Solomon Encoder that encapsulates the data with coding blocks and these coding blocks are helpful in dealing with the burst errors. The block formatted (Reed Solomon encoded) data stream is passed through a convolution encoder. Here a code rate

can be defined for convolution codes as well. If there are

k bits per second input to the convolution encoder and the output is n bits per second, the code rate is k/n. The redundancy is on not only the incoming k bits, but also several of the preceding k bits. Preceding k bits used in the encoding process is the constraint length m that is similar to the memory in the system. The convolutionally encoded bits are interleaved further prior to convert into each of the either of complex modulation symbols in QPSK, 16-QAM modulation and fed to an OFDM modulator for transmission[4-5].The simulated coding, modulation schemes and also noisy channel used in the present scenario is shown in Table 1.

TABLE 1

Modulation	**RS code**	**Convolution code rate**	**Channel used**
QPSK	(32,24,8)	½	AWGN
QPSK	(40,36,8)	¾	AWGN
16-QAM	(64,48,8)	½	AWGN
16-QAM	(80,72,8)	¾	AWGN

The model (figure-1) is a OFDM based Additive White Gaussian Noise (AWGN) channel model having Cyclic prefix (CP) blocks. For the purpose of elimination of the effect of intersymbol interference (ISI) cyclic prefix is added to the OFDM symbol. The cyclic prefix serves as a guard interval [2] .Also cyclic

prefix acts as a buffer region where delayed information from the previous symbols can get stored. Thus cyclic prefix is used to reduce the effect of fading in the channel and improve the performance of the channel by improving Bit Error Rate for the modulation scheme used in system. However transmission of cyclic prefix reduces the data rate as it does not carry any data information, thus system designers will have to design the system with optimal cyclic prefix duration.[7] Typically, cyclic prefix duration is determined by the expected duration of the multipath channel in the operating environment.

T_g = Guard interval or cyclic prefix duration

T_d = Multi path delay spread

Typically,

$T_g > T_d$

This function is built according to IEEE 802.16d specifications, which define 4 possible values for the ratio between the duration of the cyclic prefix and the duration of the useful OFDM symbol, i.e. $G \in \{1/4,1/8,1/16,1/32\}$.[5] In the simulation model data is modulated and then cyclic prefix is added to it to reduce the effect of fading and to give sufficient time to the receiver for storage of signal. In receiver side CP is removed and after demodulation we compare transmitted data and demodulated data to calculate Bit Error Rate.

III. RESULTS AND DISCUSSION

In this section, BER vs. SNR plots have been presented for different modulations in the standard on AWGN channel model. Simulation is done in the MATLAB 2009 version. The system parameters used are given in table 2 below:

TABLE 2

PARAMETERS	VALUE
OFDM symbol	100
FFT length	256
Data subcarriers	196
Pilot subcarriers	8
Cyclic prefix	1,1/4,1/8,1/16,1/32
Modulation	QPSK,16-QAM
Convolution code rate	½, ¾
Duplex	TDD

Figure 2, 3, 4 and 5 display the performance of QPSK-1/2, QPSK-3/4, 16-QAM-1/2 and 16-QAM-3/4 respectively on AWGN channel.

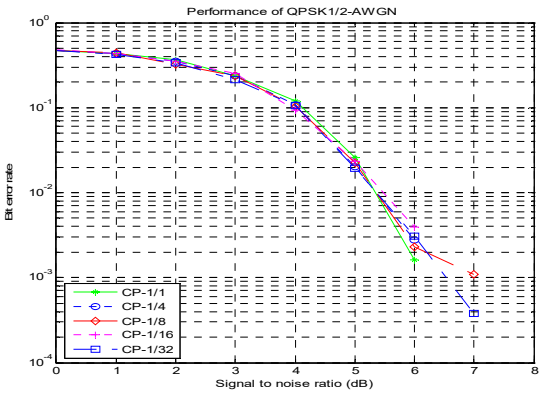


Fig. 2: Response of QPSK-1/2 modulation under AWGN channel

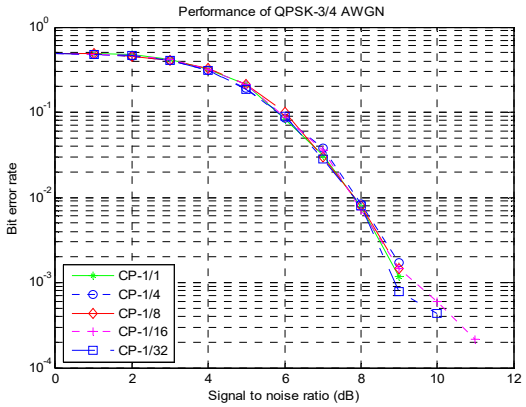


Fig. 3: Response of QPSK-3/4 modulation under AWGN channel

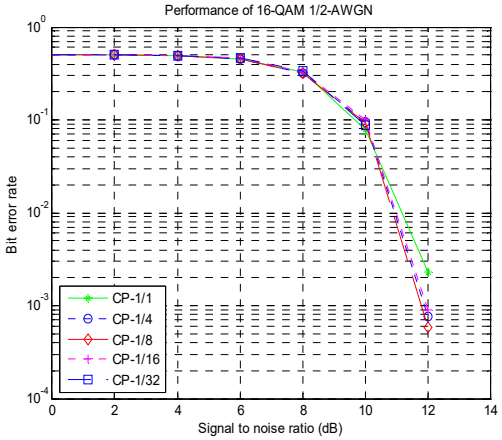


Fig. 4: Response of 16-QAM-1/2 modulation under AWGN channel

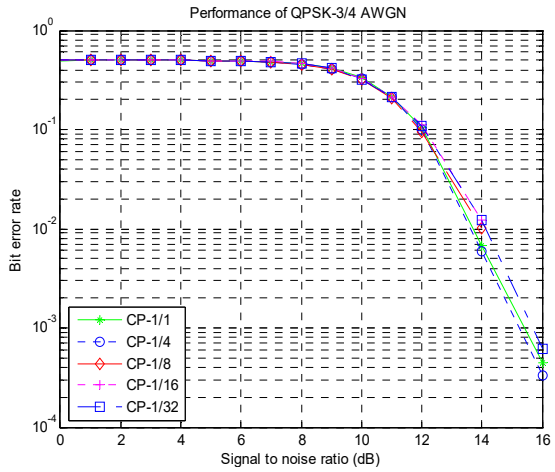


Fig. 5: Response of 16-QAM-3/4 modulation under AWGN channel

IV. CONCLUSION

The performance of the system under QPSK modulation in ½ convolution code rate is quite satisfactory as compared to QPSK-3/4, 16-QAM-1/2 and 16-QAM-3/4 modulation technique in AWGN channel for low SNR whereas for high SNR 16-QAM1/2 and 3/4 could be used. It can be noted that for different values of CP, BER could be further improved in simulation. Thus CP plays an important role in WiMAX network. CP values are compared for each modulation technique. CP limits the speed of transmission but we require it due to the multipath communication. Thus according to channel conditions CP and modulation scheme could be chosen to improve the bit error rate and thus transmission quality.

REFERENNCES

[1] Marius Oltean, Maria Kovaci, Alexandru Isar"Physical layer simulator of WiMAX",IEEE 2010 vol21 pp. 45-56.

[2] Maninder Singh"performance evaluation of adaptive modulation teccniques of wimax networks" ,International journal of advanced engineering sciences and technologies vol no. 3, issue no. 1, 034 – 038

[3] Vinit Grewal, Ajay K Sharma"On Performance Enhancements of WiMax PHY Layer with Turbo Coding for Mobile Environments", International Journal of Advanced Science and Technology Vol. 31, June, 2011

[4] Md. Ashraful Islam"Performance Evaluation of Wimax Physical Layer under Adaptive Modulation Techniques and Communication Channels",(IJCSIS) International Journal of Computer Science and Information Security,Vol. 5, No.1, 2009

[5] Mohammad Azizul Hasan "Performance Evaluation of WiMAX/IEEE 802.16 OFDM Physical Layer", helsinki university of technology2007.

[6] WiMAX Forum Applications Working Group, "System design and AWGN Results", 2008

[7] W.E.Osman, T.Abd.Rahman, "Optimization of Guard Time Length for Mobile WiMAX System over Multipath Channel", Proceedings of IMECS 2008, vol. 2, 19-21 March, Hong Kong.

[8] Yun and M. Kavehrad, "PHY/MAC Cross-Layer issues in Mobile WiMAX.", Bechtel Telecommunications Technical Journal, January 2006, pp. 45-56.

[9] H. Yaghoobi, "Scalable OFDMA Physical Layer in IEEE 802.16 WirelessMAN", Intel Technology Journal, Vol. 1, Issue 3, 2004, pp. 201-212

[10] L. Nuaymi, "WiMAX: Technology for Broadband Wireless Access," John Wiley & Sons, 2007.

◈◈◈