IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735. Volume 4, Issue 6 (Jan. - Feb. 2013), PP 01-08
www.iosrjournals.org

Surveillance on Manycasting Over Optical Burst Switching (OBS) Networks under Secure Sparse Regeneration

C.Veera lakshmi[1], V.Kavitha[2]
[1](Second M.E.Communication Systems, M. Kumarasamy College of Engineering, Karur, India)
[2](Professor/HOD Dept of ECE, M. Kumarasamy College of Engineering, Karur, India)

Abstract : *In wavelength-routed WDM optical networks requires regeneration for few light paths, when the strength of optical signal reduced and also security and privacy are essential before Optical Burst Switching (OBS) networks can become an everyday reality. Manycasting is a communication paradigm to implement the distributed applications. In manycasting, destinations can join or exit the group, depending on the service requirements imposed on them. This dynamic movement of the destinations in the group decreases blocking effect. Each application requires its own QoS threshold attributes like physical layer properties, delay as a result of transmission and reliability of the link. If the destinations satisfy the required QoS constraints set up by the application, then only they will qualify. There are two algorithms MCM-SPT and MCM-DM required for manycasting to resolve the multiconstraint QoS drawback. For continuous burst transmission lightpath should be regenerated before it loses the information due to lack of signal strength. To recover signal strength by sparse regeneration, where OOO switches are replaced by OEO swtches. There are three algorithms 1).NDF 2).CNF 3).SQP. Sometimes, there is an opportunity for the attacker to join the group. Service provided to the attacker is restricted by providing two levels of security. Using1).RSA algorithms, data level security is provided and using 2).certificate authentication, link level security is provided.*

Keywords: *Certificate generation, Manycasting, Optical Burst Switching Networks (OBS), Sparse regeneration, Quality of Service (QoS).*

I. Introduction

Information Security has become an important issue in data communication. Encryption plays a vital role in information authentication system. This authentication mechanism uses some algorithms to scramble data into unreadable text which can be only being decoded or decrypted by party those possesses the associated key. Different types of algorithms are used to prevent malicious attack when the time of data transmission. Encryption algorithm can be divided into symmetric key (private) and asymmetric (public) key Encryption [1]. In Symmetric keys encryption or secret key encryption, the data is encrypted or decrypted using only one key. In Asymmetric keys, two keys such as private and public keys are used. Encryption is done by Public key and decryption is done by Private Key (example: RSA).

When the time of signal degradation due to long distance (>4000 km), attacker can easily hack the information. To avoid this dangerous situation at continuous wavelength routed WDM network, Sparse Regeneration is necessary. It is a technique where the light path is routed using Optical switches (OOO switches). The light path which carries the information loses its strength due to physical limitation of distance. Instead of OOO switches, OEO switches can be replaced in the network to regenerate the signal strength. The transparent optical network can be converted into opaque optical network which uses 3R (reamplification, reshaping, retiming) regeneration method at every intermediate node of a lightpath to regenerate the signal and improve transmission quality... In such a network, a lightpath can be successfully routed if and only if: 1) common wavelength should be there on each link and 2) before regeneration at an OEO switch,) the lightpath should not exceed the reach limit. Manycasting will be secured because of this protection along the information packaged at burst.

Manycast is also called *quorumcast* and was first proposed by [2] in which the groups of destinations that receive the message are to be selected instead of being given. In manycasting messages are sent to a subset of destinations (quorum pool), which are selected from set $D_{s,}$ such that $k \leq \left| D_s \right| = m$. Manycasting has caught the attention of several researchers during the recent past, due to the emergence of many distributed applications such as distributed interactive simulations (DIS), Video conference, storage area network (SAN) and grid computing.

Optical Burst Switching (OBS) [3–8] has emerged as a promising solution for satisfying the increased capacity demands of modern broadband networks. In OBS systems, the various types of data carried in the network (e.g., IP packets, ATM cells, frame relay frames, and so forth) are electronically buffered at the ingress node, aggregated into bursts, and then transmitted all-optically through the core network as a single burst. Upon

arrival at the egress node, the burst is disassembled and the individual data packets are then forwarded to the appropriate destinations.

In this paper we propose two levels of security using RSA and Certificate Authentication for restricting burst to reach attacker and consider three algorithms for regenerating the signal strength. The rest of the paper is organized as follows: we first discuss about the existing system in Section II, proposed system in Section III, Section IV discusses performance evaluation of the proposed algorithms and explain the throughput and delay constraint of existing algorithms. Finally, conclusion is given by Section V in this paper.

1.1. Existing System

Existing work focuses on selecting the best possible destinations that can meet the service demand must effectively. Destinations chosen are able to provide quality of service attributes. If a destination satisfies the service requirements of the application, it will qualify as the member of quorum pool. Algorithms implemented in the centralized way, may fail due to a link failure and resulting in poor performance due to insufficient security and lack of information due to signal strength.

(s, D_s, k) is the manycast request where s is the source node, D_s is the destination set and k is the minimum number of destinations that are required to participate in the manycast session. Path information vector Ω is generated according to the service attributes such as Noise factor η_j, Reliability γ_j, Propagation delay $\tau_{j.}$ for link j,

$$\Omega_j = \begin{pmatrix} \eta_j \\ \gamma_j \\ \tau_j \end{pmatrix} \qquad (1)$$

For the successful establishment of QoS-based manycast session, the chosen destinations must satisfy the service requirements T ($^{\theta}{}_{P}$) that is given by,

This system proposed two algorithms, MCM-shortest path tree (MCM-SPT) and MCM dynamic

$$\top^{(\theta_p)} = \begin{pmatrix} \eta_{\max}^{(\theta_p)} \\ \gamma_{\min}^{(\theta_p)} \\ \tau_{\max}^{(\theta_p)} \end{pmatrix} \qquad (2)$$

membership (MCM-DM) for evaluating the performance of manycasting with QoS constraints. Common steps involved in these algorithms,

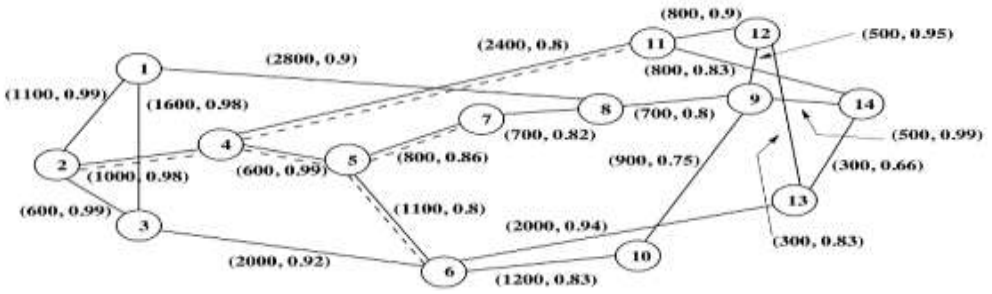


Fig.1: NSF topologies with 14 nodes. The weight represents distance (km) and corresponding reliability.

Step 1: Search the shortest path from source to all the destinations in D_S

Step 2: All the destinations in D_S are sorted in the non decreasing order according to the shortest distance from source to the destinations. Let D_S' be the new set.

Step 3: Choose the first K destinations from D_S'.

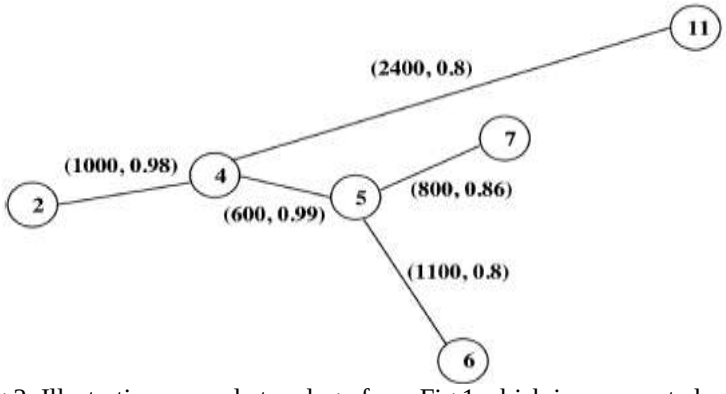


Fig.2: Illustrative example topology from Fig.1 which is represented as dotted line

In Fig.2.Consider a manycast request (s =2, D_s= {6, 7, 8}, k_0 = 2). Let the service threshold be T ($^{\theta}{}_{P}$) = [η_{th} = 6, γ_{th} = 0.6, τ_{th} = 20 ms].Then the above steps are followed. If the services are satisfied correct destinations will be found else QoS blocking will be occurred for MCM-SPT Algorithm. For MCM-DM Dynamically it calculates the destination as per the constraint they required. If it not possible QoS blocking will be occurred.

II. Proposed System

Proposed work focuses on regenerating the signal strength for lossless data communication and providing authentication for burst to reach the correct destination and protecting burst from attacker. Disadvantage of manycasting is the attacker can join into the group easily by act as a destination which satisfies the QoS service requirement and the signal degradation at long distances... To avoid this situation proposed work guaranteed two level of security such as Data level security using RSA algorithm and Link level security using Certificate authentication .Regenerating will be happened by choosing three algorithms such as 1).NDF (Nodal Degree First) 2).CNF (Centered Node First) 3).SQP (Signal Quality Prediction).

2.1. Data Level Security using RSA Algorithm

Data Security means protecting a database from destructive forces and the unwanted actions of unauthorized users. For this purpose we choose asymmetric key cryptography to provide efficient security. Public key algorithm invented in 1977 by Ron Rivest, Adi Shamir and Leonard Adleman (RSA) Supports Encryption and Digital Signatures. *Factorization* is a major problem when using RSA. The basis of security in RSA is difficulty of large number factorization Long numbers contains over the 1000 bits.

The RSA algorithm [11-13], [16] is based on the mathematical part that is easy to find two large prime numbers and multiple together, but factoring their product will be difficult. There are some important steps are involved in a RSA algorithm to solve a problem as given below:

Step 1: Assume two large prime numbers x & y.

Step 2: Compute: N = x*y Where N is the factor of two large prime number.

Step 3: Select an Encryption key (E) such that it is not a factor of (x-1)*(y-1) i.e. Ø (n) = (x-1)*(y-1) for calculating encryption exponents E, should be 1< E Ø (n) such that gcd (E, Ø (n) =1 The main purpose of calculating gcd is that E & Ø(n) should be relative prime Euler Totient Function is represented here is Ø (n) & E is the Encryption Key.

Step 4: Select the Decryption key (D), which satisfy the Equation D*E mod (p-1)*(q-1) = 1

Step 5: For Encryption:

Cipher Text= (Plain Text) E mod N (or)

CT=ME mod N

Step 6: For Decryption:

Plain Text= (Cipher Text) E mod N (or)

PT= (CT) E mod N

The main feature of RSA algorithm is the selection of large prime number (x, y) because any users or force attackers could not be able to find the capable numbers, timely to force attack is shortly non-feasible. Security will be provided by RSA depends on some parameters and its selection method.

1. Selection of large prime number(x, y):

Example: x = 5, y = 3 N= x*y = 5*3 = 15 = 1*15 = 15*1 = 3*5 = 5*

2. Selection of Encryption Key (E): During the selection Of Encryption key, selection of large prime fraction always creates impact. If the factor is high then the estimation of Encryption is infeasible.

Example: If x=7, y=17 must not be a factor of (x-1) *(y- 1) i.e. (7-1)*(17-1) = 6*16 = 96 = 2*2*2*2*2*3 So, E can be 5, 7, 11…

3. Selection of Decryption Key (D): Selection of large factors always create an effect on the Decryption key, there may be an inversely relation.

$$(E*D) \bmod (x-1)*(y-1) = 1$$

$$D \alpha 1/ [(E) (x) (y)]$$

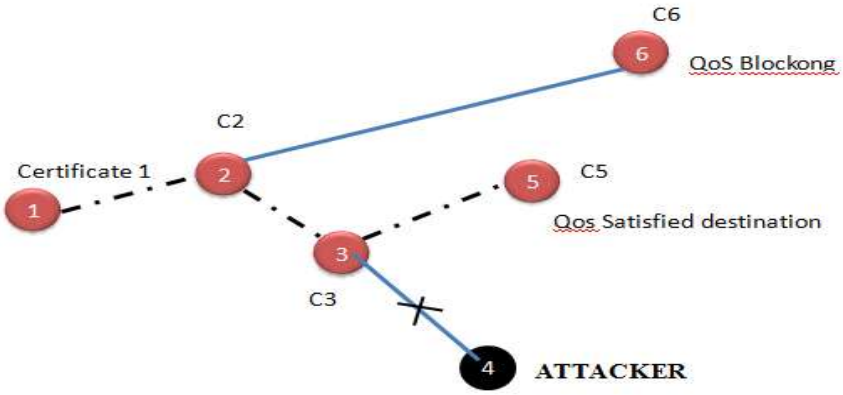


Fig.3: Restriction of attacker node by using combination of RSA and certificate authentication.

Some important points for reference are given below:

- Condition of Euler's Totient Function:
  1. Ø (1) = 0
  2. Ø (x) = x – 1 {if x is prime number}
  3. Ø (m*n) = Ø (m)* Ø (n) {if m and n is relative prime number
  4. Ø (Xe) = Xe – xe -1 {if x is a prime number}
- There is no necessity for a user to know his secret parameters x, y and Ø (n).
- The plain text or message (M) has the form of one or more positive integer M<N.
- Any user can use his private key to authenticate the communication.

RSA cryptosystem provides the facility of digital signature scheme. The message consists of letters, numbers and special characters (i.e. stop, colon, space etc.). Each character is represented by its own arrangement of Eight bits (o & 1). The most of the hardware & software products and standards that use public key technique for Encryption, Decryption etc. are based on RSA cryptosystem. There are some advantages of RSA Algorithm:

- Important benefits of Public key Cryptography is increased security and convinces.
- Second, it provides digital signature that can't be repudiated. For example, Kerberos secret-key authentication system involves central database that keeps copies of secret key.
- Public key authentication prevents the type of repudiation and each user has its own responsibility for protecting his own private key.
- We can select large prime numbers for enhancement of security of keys.
- Public key cryptography may be used with secret key cryptography.

2.2. Link Level Security using Certificate Authentication

Link encryption (sometimes called *link level* or *link layer encryption*) is the data security process of encrypting information at the data link level as it is transmitted between two points within a network. Link encryption takes place in the lowest protocol layers (layers 1 and 2 in the OSI model). because the process protects the message in transit. Link encryption is very useful in situations where the uncertainty of security. Dangerous issues can occur at a link when the message must be transmitted between hosts, the message is decrypted at each host in the transmission path, which are not known to be secure. Link encryption has been used successfully in military, where the assurance of security is provided for each link. It isn't work over the Internet, because intermediate links are neither accessible nor secure.

This can be done using certificate authentication. The most common form of trusted authentication between parties in the wide world of Web commerce is the exchange of certificates [14]. A certificate is a digital document (DN) and an associated public key. The certificate is signed digitally by a trusted third party known as the Certificate Authority (CA). The CA vouches for the authenticity of the certificate consumer. Each principal in the transaction presents a certificate as its credentials. The recipient then validates the certificate's signature against its cache of known and trusted CA certificates. A "personal certificate" identifies an end user in a transaction; a "server certificate" identifies the service provider .Commonly the certificate consists of source address, destination address, MAC address, public and private key.

Example: Certificate of Node 0
IP address - 190.06.30.8
MAC address - 25EFDHY5898
Public key - 3
Private Key - 3

2.3. Sparse Regeneration

Next generation optical networks are expected to be combination of all optical-Cross Connects (OXC) and OEO cross connects. In core network OEO cross connects are used for wavelength conversion and regeneration. Sometimes for long distance communication (>4000 km), light path needs to regenerate its strength for avoiding data loss. This problem is known as Sparse OEO Placement (SOEOP) problem [17].

A light path is a travelling path of light between two OXC's to carry the information. A typical light path is given by Fig.4. When the Optical Signal to Noise Ratio (OSNR) will be reduced, at that time regeneration will be needed. OEO OXC used for regenerating the signal strength.

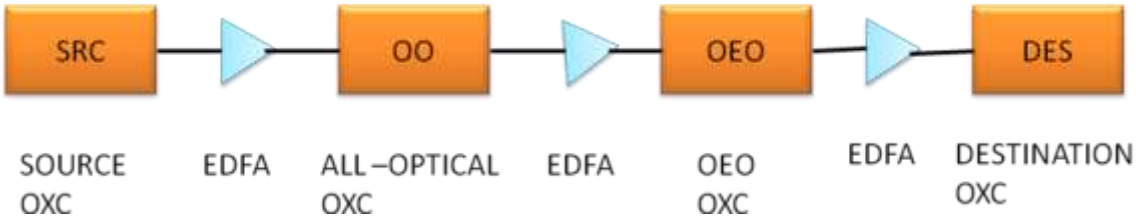


Fig.4: Light path from Transmission perspective

2.3.1. Sparse Regeneration Node Model

The 3R OEO regeneration can be performed by 3R OEO regenerator as shown in Fig.5.It consists of T-R pair with electronic processing to regenerate the optical signal through 3R processing [18]. The Sparse regeneration node model is given by Fig.6.

The wavelengths on incoming fiber links are demultiplexed at each node at OXC module. A node can add/drop the wavelength when it is needed. The regeneration capable node will be detected in the arrays of transmitter and receiver pair. At regeneration capable node, the regeneration demand is happened by the light path for say λ_1. It directs the light path for the available receiver in R_R. Then it converts optical signal into electrical signal using O/E conversion .Electrical signal is processed by 3R regenerator. After regenerating, it is converted into optical signal through E/O converter at corresponding T_R as a wavelength λ_2.This regenerated optical signal is transmit through output fiber link.

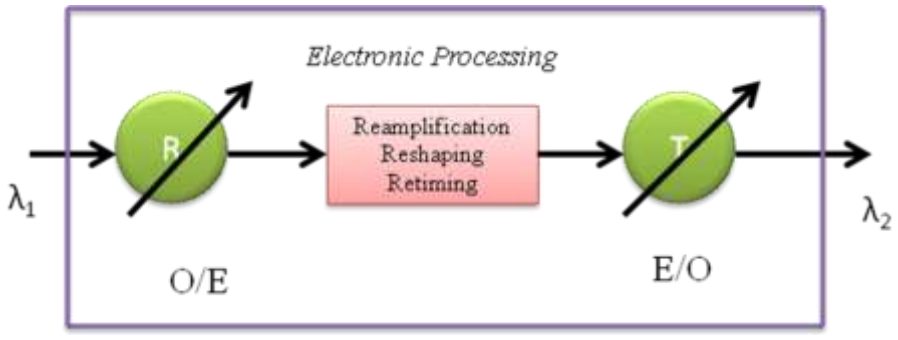


Fig.5:3R OEO Regenerator

There are three algorithms to regenerating the optical signal with high OSNR and high signal quality.

- **Nodal Degree First (NDF)**
- **Centered Node First (CNF)**
- **Signal Quality Prediction (SQP)**

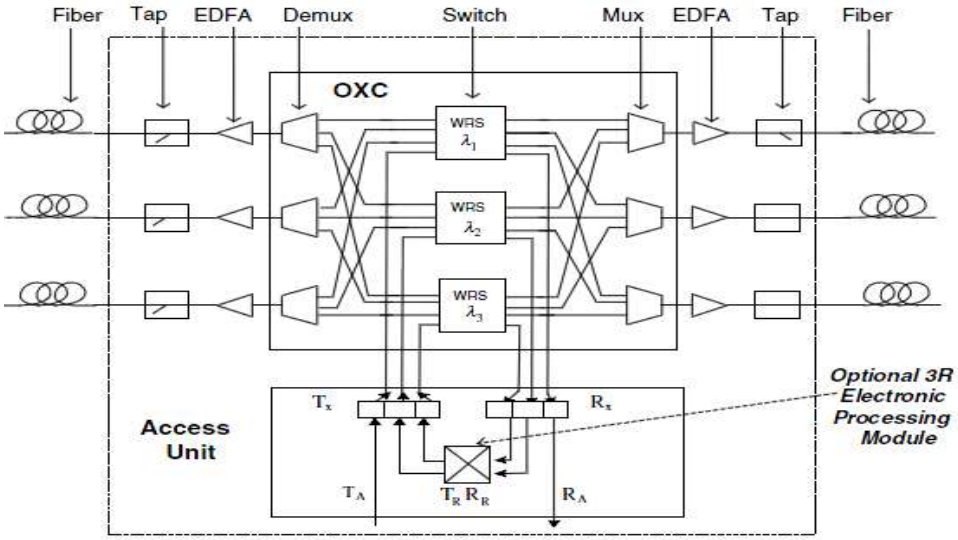


Fig.6: Sparse Regeneration Node

3.3.2 ***Explanation of Algorithms***

Algorithm 1. Nodal Degree First (NDF) regenerated placement.

Step 1: Specify each node by number equal to nodal degree.

Step 2: Choose a node with high regenerated capability. If more nodes exist choose randomly one node regenerator node and then reduce the degree by 1.

Step 3: Repeat the step until N groups of regenerators are placed.

Algorithm 2. Centered Node First (CNF) regenerator placement

Step 1: Specify large number for center node .The nodes equally centered are assigned number in random order.

Step 2: Choose N nodes as a regenerator nodes and place regenerators at each node.

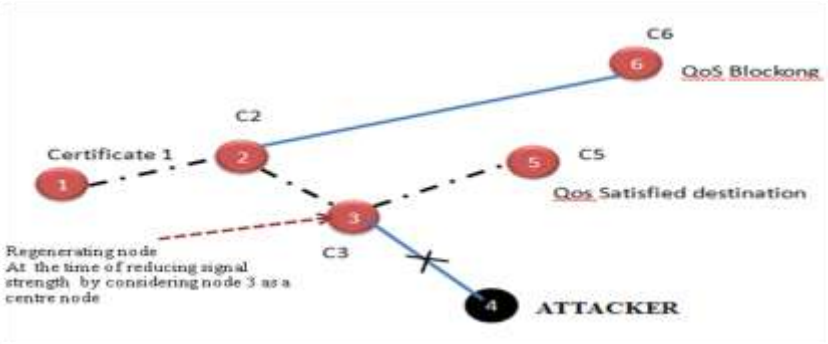


Fig.7: Using CNF regenerator placement (Node 3)

Algorithm 3. Signal Quality Prediction (SQP) based regenerator placement

Step 1: Specify each node j by number Cj, initialized to be zero.

Step 2: Workout a predefined wavelength routing algorithm on lightpaths demands which are randomly generated by following a predicted traffic pattern.

Step 3: Arrange all nodes in non-increasing order of Ci.

Step 4: Choose the first N nodes as regeneration capable nodes and place regenerators at each node.

Table 1: Comparison of Regeneration algorithms

Features	**NDF**	**CNF**	**SQP**
Blocking Probability	Moderate	High	Low
Loss of data	Moderate	High	No loss
Reachability of regenerated signal quality	Moderate	Very low	High

III. Performance Evaluation

These two authentication algorithms provides efficient security, when it finds the node is attacker at that time instantly it mark the node is attacker and the burst transmission for that node will be restricted. The certificate is not generated for that particular node. Fig.8. Describes which node is attacker and it is denoted as 0.Certificate generated nodes are denoted as 1. The comparison result of parameters such as throughput and delay also taken between MCM-SPT and MCM- DM as shown in Fig.9 and Fig.10. This simulation results show that MCM-shortest path tree (MCM-SPT) algorithm performs better than MCM-dynamic membership (MCM-DM) for delay constrained services and real time service, where as data services can be better provisioned using MCM-DM algorithm.

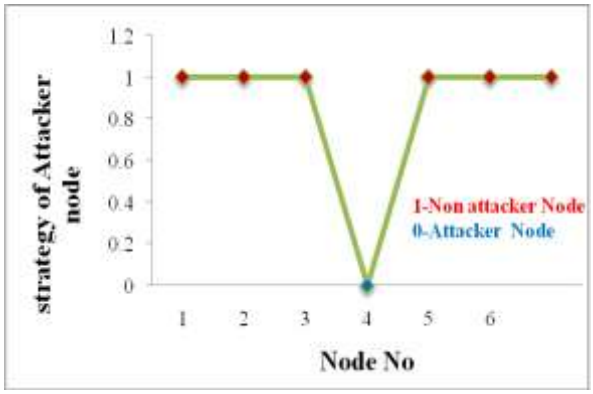


Fig.8: Certificate generation strategy of nodes

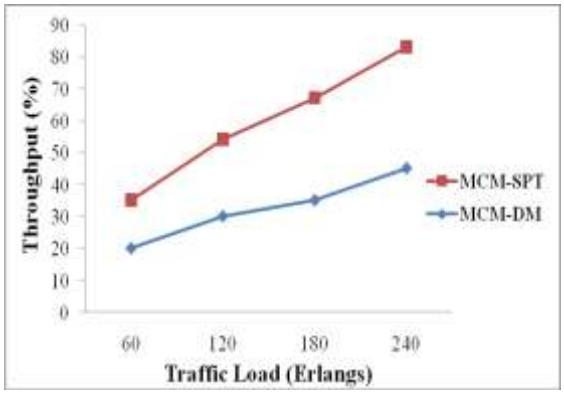


Fig.9: Traffic Load Vs Throughput

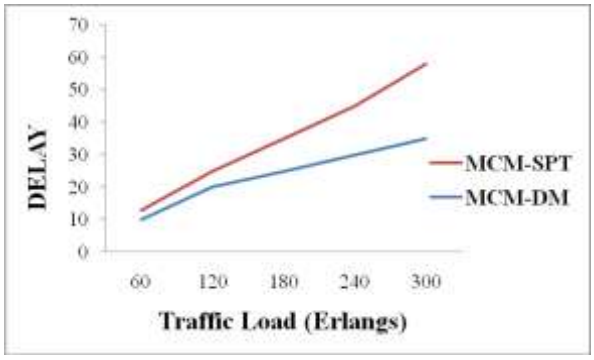


Fig.10: Traffic Load Vs Delay

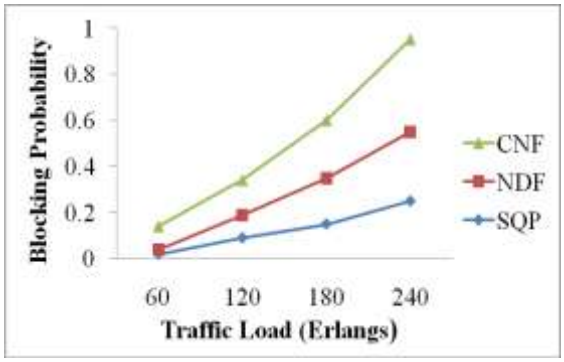


Fig.11: Traffic Load Vs Blocking Probability

One of the important features of the Sparse regeneration is the blocking probability as shown in Fig.11. it will be highest when using CNF algorithm because of choosing center node for placing regenerator. If the distance between center and Tr/Rr node is large, it will loss on the way, even it regenerate the signal. This difficulty is overcoming by NDF and SQP.

IV. Conclusion

In this paper, algorithms are proposed to support authentication and regeneration purpose in QoS constraint manycasting over Optical Burst Switching (OBS) networks. RSA and Certificate generation based cryptography supports source to fine the attacker by generating certificate with the help of public and private key, which are generated by RSA algorithm. Simulation results that multiconstrained manycast dynamic membership algorithm is suited for data service because it provides higher throughput and multiconstrained manycast shortest path tree algorithm for real-time service because it provides lower delay. NDF,CNF and SQP algorithms are used for regenerating the optical signal when the OSNR is decreased. The simulation result of Sparse regeneration said that blocking probability is higher than remaining both algorithms. We also evaluated the performance of our algorithms by strategies of attacker nodes, if it is attacker, the node doesn't generate certificate (0).Our work can be further extended by considering the manycastiing over Optical Burst Chain Switching (OBCS) networks with sparse wavelength regeneration.

Acknowledgements

Our sincere thanks to Professor V. Kavitha M.E., (Ph.D)., Head of the Department of Electronics and Communication Engineering, M.kumarasamy College of Engineering, Karur, who have contributed towards development of work. Also the author would like to thank reviewers for their valuable comments to improve the quality of the paper.

REFERENCES

[1] Diaasalama Abd Elminaam, Hatem Mohamad Abdual Kader, Mohly Mohamed Hadhoud , Evalution of the Performance of Symmetric Encryption Algorithms, *International journal of network security,vol. 10, No.3, pp, 216-222, May 201.*
[2] B. G. Balagangadhar, *QoS Aware Quorumcasting Over Optical Burst Switched Networks*, doctoral diss., Indian Institute of Science, Bangalore, India, Electrical and Communication Engineering July 2008.
[3] Chen.Y, Qiao. C, Yu and X, Optical burst switching a new area in optical networking research, *IEEE Netw.* 18, 16– 23, 2004.
[4] Qiao. C, Labeled optical burst switching for IP-over -WDM integration, *IEEE Commun. Mag.* 38(9), 104–114 (2000).
[5] Qiao, C, Yoo.M, Optical Burst Switching (OBS) — A new Paradigm for an optical internet, *J. High Speed Netw.* 8(1), 69–84, 1999.
[6] Turner.J, Terabit burst switching High Speed, *J. High Speed Netw* 8(1), 3–16, (1999).
[7] Yoo, M. Qiao, C, Dixit. S,Optical Burst switching for Service Differentiation in the next - generation optical internet, *IEEE Commun. Mag.* 39(2), 98–104, 2001.
[8] X. Huang, Q. She, V. M.Vokkarane and J. P. Jue, Manycasting over optical burst - switched (OBS) networks, *Proc. IEEE ICC, Glasgow,* Scotland, May 2007, pp. 2353–2358.
[9] A. Kaheel, T. Khattab A. Mohamed and H. Alnuweiri, Quality - of service Mechanismsin IP –over networks, *IEEE Commun. Mag. vol. 40, no. 12, pp. 38–43, Dec. 2002.*
[10] B. G. Bathula, V. M. Vokkarane and R. R. C. Bikram, Impairment aware manycasting over optical Burst - switched (OBS) networks, *Proc. IEEE ICC,* Bejing, China, May 2008, pp. 5234–5238.
[11] Anoop MS, Public key Cryptography (Applications Algorithm and Mathematical Explanations), Tata Elxsi Ltd, India.
[12] Andrea Pellegrini, Valeria Bertacco, Todd Austin, Fault-Based attack of RSA authentication, *University of Michigan*.
[13] Shashi Mehrotra Seth, Rajan Mishra, Comparative Analysis of Encryption Algorithms for Data Communication, *IJCST Vol. 2, Issue 2,* June 2011.
[14] Jim DeRoest, Certificate- Based Authentication, Published *on SunExpert Magazine,* June 1997.
[15] M.Catherine Jenifer, P.Jayachandar "Cryptanalysis on RSA algorithm" *Published on, International Journal of Communications and Engineering Volume 03– No.3, Issue 01* March2012.
[16] Prasant Singh Yadav, Pankaj Sharma ,Dr. K. P Yadav , Implementation of RSA algorithm using Elliptic curve algorithm for security and performance enhancement, *International Journal of Scientific & Technology Research Vol 1, Issue 4*, May 2012.
[17] Hui Zang, Renxiang Huang and James Pan, Methodologies on Designing a Hybrid Shared- Mesh Protected WDM Network with Sparse Wavelength Conversion and Regeneration, *Proceedings of SPIE Vol. 4910*,2002.
[18] Xi Yangz, Byrav Ramamurthy, Sparse Regeneration in Translucent Wavelength- Routed Optical Networks Architecture, Network Design and Wavelength Routing"*Journal of Photonic Network Communications 10:1, 39–53,* 2005.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735. Volume 4, Issue 6 (Jan. - Feb. 2013), PP 09-16
www.iosrjournals.org

Improving Energy Efficiency Using LEARN Algorithm for Wireless Sensor Networks

Srigitha.S.Nath (Ph.D)[1], Suviga.K[2]
Associate Professor/ECE [1] , M.E (Computer & Communication)[2] Saveetha Engineering College, Chennai, India

Abstract : *A number of energy-aware routing protocols were proposed to seek the energy efficiency of routes in Multihop wireless networks. Among them, several geographical localized routing protocols were proposed to help making smarter routing decision using only local information and reduce the routing overhead. However, all proposed localized routing methods cannot guarantee the energy efficiency of their routes. In this article, we first give a simple localized routing algorithm, called Localized Energy-Aware Restricted Neighbourhood routing (LEARN), which can guarantee the energy efficiency of its route if it can find the route successfully. The critical transmission radius in random networks which can guarantee that LEARN routing finds a route for any source and destination pairs asymptotically almost surely. So by using LEARN algorithm the energy efficient of the source and destination pair is thus may be maintained and the attacks occurs in the wireless network can also be reduced. One can also extend the proposed routing into three-dimensional (3D) networks and derive its critical transmission radius in 3D random networks.*

Keywords*: LEARN, Localized routing, 3D network, Multihop, CRT, Energy Efficient*

I. Introduction:

A Wireless sensor network is a distributed autonomous sensor to monitor physical or environmental conditions, as temperature, sound, vibration, pressure, motion or pollutants and to cooperatively pass their data through the network to a main location. The more modern networks are bi- directional, also enabling control of sensor activity. The development of wireless sensor networks was motivated by military applications such as battlefield surveillance; today such networks are used in many industrial and consumer applications, such as industrial process monitoring and control, machine health monitoring, and so on. Energy conservation and scalability are probably two most critical issues in designing protocols for Multihop wireless networks, because wireless devices are usually powered by batteries only and have limited computing capability while the number of such devices could be large. In this paper, we focus on designing routing protocols for multihop wireless networks which can achieve both energy efficiency by carefully selecting the forwarding neighbours and high scalability by using only local information to make routing decisions. Numerous energy aware routing protocols have been proposed recently using various techniques (transmission power adjustment, adaptive sleeping, topology control, multipath routing, directional antennas, etc). Most of the proposed energy-aware routing methods take into account the energy-related metrics instead of traditional routing metrics such as delay or hop count.

The problem of energy-efficient reliable wireless communication in the presence of unreliable or loss wireless link layers in multi-hop wireless networks. Prior work has provided an optimal energy efficient solution to this problem for the case where link layers implement perfect reliability. However, a more common scenario of a link layer that is not perfectly reliable, was left as an open problem. In this paper we first present two centralized algorithms, BAMER and GAMER[1], that optimally solve the minimum energy reliable communication problem in presence of unreliable links. Subsequently we present a distributed algorithm, DAMER[2], that approximates the performance of the centralized algorithm and leads to significant performance improvement over existing single-path or multi-path based techniques. Wireless communication networks have been deployed at an increasingly fast rate, and are expected to reshape the way we live in this physical world. For example, wireless ad hoc networks in paper [2] and [3]combined with satellite data networks are able to provide global information delivery services to users in remote locations that could not be reached by traditional wired networks. There are two well-known ways to achieve end-to-end reliability on multi-hop paths. The first approach employs hop-by-hop retransmissions each link layer hop retransmits lost frames as and when necessary. The secondly [3] assumes that link layers are unreliable and retransmissions are performed end-to-end. It is also possible to consider a mix of the above as a third approach, where link layers perform a few retransmissions if necessary, but perfect reliability is only guaranteed through end-to-end mechanisms. However, link layer retransmission actually cannot guarantee reliable delivery, due to various reasons. In the end-to-end retransmission model where some link in the communication path is unreliable, we rely on TCP-like transport protocols to initiate end-to-end retransmissions.

Since in paper [3]sensor networks can be composed of a very large number of nodes, the developed protocols for these networks must be scalable. Moreover, these protocols must be designed to prolong the battery lifetime of the nodes. Typical existing routing techniques for ad hoc networks are known not to scale well. On the other hand, the so-called geographical routing algorithms are known to be scalable but their energy efficiency has never been extensively and comparatively studied. For this reason, a novel analytical framework is introduced. Ina geographical routing algorithm, the packets are forwarded by a node to its neighbour based on their respective positions is given in paper[4] and [5] . The proposed framework allows to analyze the relationship between the energy efficiency of the routing tasks and the extension of the range of the topology knowledge for each node. The leading forwarding rules for geographical routing are compared in this framework, and the energy efficiency of each of them is studied. Moreover Partial Topology Knowledge Forwarding, a new forwarding scheme, is introduced. A wider topology [6] knowledge can improve the energy efficiency of the routing tasks but can increase the cost of topology information due to signalling packets that each node must transmit and receive to acquire this information, especially in networks with high mobility. The problem of determining the optimal Knowledge Range for each node to make energy efficient geographical routing decisions is tackled by Integer Linear Programming. It is demonstrated that the problem is intrinsically localized, i.e., a limited knowledge of the topology is sufficient to take energy efficient forwarding decisions, and that the proposed forwarding scheme outperforms the others in typical application scenarios. For online solution of the problem is resulted in paper [6] and [7], a probe-based distributed protocol which allows each node to efficiently select its topology knowledge, is introduced and shown to converge to a near-optimal solution very fast.

In a localized routing algorithm, each node currently holding message makes routing decision solely based on the information about itself, its neighbours and destination given in paper [8]. In a unit graph, two nodes can communicate if and only if the distance between them is no more than the transmission radius, which is the same for each node. In paper [9],a localized routing algorithm that guarantees delivery in connected unit graphs has been described recently. Also, several power, cost and power-cost aware metrics and localized loop-free routing algorithms for wireless networks based on the exact power needed for transmission between two nodes, and/or remaining battery power at nodes were proposed. This paper[10] and [11] proposes such localized routing algorithms, aimed at minimizing total power for routing a message, or maximizing the total number of routing tasks that a network can perform before a partition. The efficiency of proposed algorithms is verified experimentally by comparing their power savings, and the number of routing tasks a network can perform before a node loses all its energy, with the corresponding shortest weighted path algorithms, and localized algorithms that use fixed transmission power at each node. Significant energy savings are obtained. In the paper [12] was to examine power consumption with respect to power non-aware protocols and provide basis for further study. Power efficient methods tend to select well-positioned neighbouring nodes in forwarding the message, while cost efficient methods favour nodes with more remaining power. The performance of internal node based routing also improves with mobility (if the cost of location updates is ignored), since mobility changes the dominating set frequently. When the nodes are static, the fixed choice of internal nodes becomes routing bottleneck, since these nodes are used more often than others in routing messages.

Greedy forward routing in wireless ad hoc networks is a localized geographic routing in which each node discards a packet if none of its neighbours is closer to the destination of the packet than itself, or otherwise forwards the packet to the neighbour closest to the destination of the packet. If all nodes have the same transmission radii, the critical transmission radius for GFR is the smallest transmission radius which ensures that packets can be delivered between any source-destination pairs. In this paper[13],[15], we study the asymptotic critical transmission radius for GFR in randomly deployed wireless ad hoc networks. We assume that the network nodes are represented by a Poisson point process of density n over a convex compact region of unit area with bounded curvature .Greedy forward routing is a localized and memory less geographic routing. However, it cannot guarantee the delivery of a packet from its source to its destination if the transmission of the nodes are not large enough. The smallest transmission radius which ensures the successful delivery of any packet is referred to as the critical transmission radius. In this paper, we provides tight bounds on the critical transmission radius when the networking nodes are represented by a Poisson point process. As a future work [10]-[13], one may investigate a number of other parameters related to GFR. These parameters include the average of one-hop progress, the expected number of hops between a source and destination, the ratio of the total length of the path to the Euclidean distance between the source and the destination. It is also interesting to study the asymptotic of other localized geographic routings.

Provide relevant information about the state of nature to a fusion centre. Set of identical binary sensors is asymptotically optimal, as observation per sensors goes to infinity. Dependence across sensors leads to the existence of multiple bits. Non identical sensors over identical binary sensors leads to error bit information. Neyman-Pearson type probability of error is fixed and can be minimized. Neyman-Pearson criterion relates both the energy consumption in routing and detection performance .Routing metric for a balance between consumed

energy and detection performance by aiming for a route with maximum mean detection probability to energy ratio. Not directly indicate the amount of energy need to achieve a detection performance. Does not reflect the lifetime of sensor node. For cross layer design is scalable with respect to network size and greatly reduce the dependence on central fusion center. Clustering infrastructure with cluster heads in distributed manner. Capable of attaining energy efficiency at different network. Through cluster heads[14] only information can be passed by consultation with immediate neighbour .No centralized computation or control for routing .The Location based routing protocol address by the distance between neighbour nodes on the basis of incoming signal strength. To save energy, some location based scheme demand the node should go to sleep state if there exist no activity. Geographic Adaptive Fidelity (GAF): Increase network lifetime as the number of nodes increases. Saves energy and packet loss. Randomly Distributed sensor network. Modified Monte-carlo integration method used for obtaining the matrix and computing the weight of edge Modified LS algorithm [15] used to find the routing path with maximum sensing coverage and K hop constraint. Only hop constraint is considered. A send/no-send scenario for transmission led to a tractable and realistic problem formulation where likelihood-ratio-based detectors are optimal. Randomization of the measurement and choice of transmit rate could be used to meet a constraint on the expected cost and to optimize detection performance. Its fail to explain the dependent observation for finding the probability of error I terms of K-L distance.

II. System Description And Problem Statement

Problem Statement

` Most of the proposed energy-aware routing methods take into account the energy related metrics instead of traditional routing metrics such as delay or hop count. To select the optimal energy route, those methods usually need the global information of the whole network, and each node needs to maintain a routing table as protocol states. In opposition to these table-driven routing protocols, several stateless routing protocols, particularly, localized geographic routing protocols have been proposed to improve the scalability.

Existing System

The problem of energy-efficient routing for signal detection under the Neyman–Pearson criterion. Considering a set of sensors which are distributed over an area to detect the possible presence of an illuminated target at a specific location through active sensing. The fusion center is assumed to have the knowledge of geographical location of each sensor, and will determine the location to be probed as well as the communications routing (i.e., centralized routing is employed). The observations at each sensor are assumed to be independent conditioned on the hypothesis. Each sensor node makes a measurement to test for a target at the predetermined location, and then transforms the measurement into a likelihood ratio. This can be regarded as one particular type of in-network processing in WSNs; that is, instead of sending the raw data, each sensor node uses its processing ability to transform its measurement into a sensor likelihood ratio, a sufficient statistic for the problems we consider (conditionally independent sensor data) that provides significant compression. After collecting all the likelihood ratios from the sensor nodes along a route, the fusion center will make a final decision.

The work exploits a synergistic approach by incorporating aspects of signal processing,
computer science, and operations re-search, and can be regarded as one special type of the cross-layer network design. To summarize, the novelty and contributions of this paper are threefold.

We formulate the problem of energy efficient routing for detection under the Neyman Pearson criterion, widely adopted for target detection and surveillance related applications. This formulation, as far as we are aware, is the first one which accounts for both the energy consumption in sensing and routing, and detection performance (in terms of detection probability and false alarm probability) at the same time.

We propose three routing metrics, and for each of them, the detection performance and the energy expenditure are considered jointly in a different but interesting way by which an appropriate trade-off between them is attained. We also formulate the routing problems under these metrics into different combinatorial optimization programs, where the objectives are to maximize or minimize a quantity of interest.

We provide algorithms for solving those formulated integer programming problems, based on state-of-the-art operations research results. In cases where the optimum solution is computationally intensive, we provide methods where we can specify the acceptable complexity and a solution that can be made arbitrarily close to the optimum solution when its complexity is increased appropriately. We also present extensive simulation results which help quantify the difference proposed routing metrics.

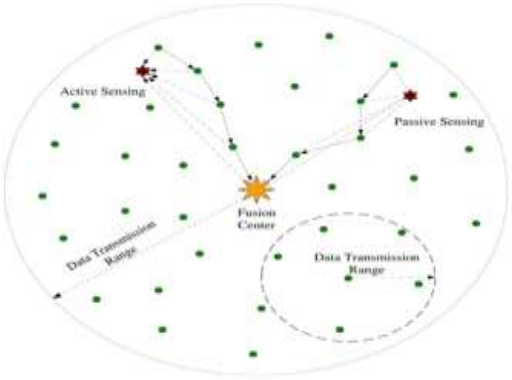


Fig. 1. Routing in WSNs for signal detection

Let us consider a scenario as depicted in Fig. 3.1, where some sensors are scattered over an area to detect the presence of an object. This scenario may correspond to monitoring the presence of people, vehicles, or military targets using radar-like sensors that emanate specific electromagnetic signals into the region of interest. For this active sensing application, the monitored space is typically divided into many range resolution cells. Each range cell could be probed sequentially in turn to deter-mine the presence of a target by using radar pulses that are possibly launched by directional antennas.

We assume the position of each sensor node is fixed. Each sensor node is assumed to be aware of its own position (e.g., by using methods based on triangulation), and has knowledge about the location of the sensors in its local neighbourhood. We assume the fusion center has information about the geographical location of each sensor, and will determine the location to be probed (for target presence) as well as the sensors that will be involved in signal detection and information transmission. This basically is a destination initiated routing, and the routing algorithm can be implemented offline, ahead of time, at the fusion center. The system will sequentially probe each location of interest, and collect the information from sensor nodes lying on a precomputed path in a predetermined order.

It is worth noting that there exist several ways for active sensors to illuminate the target. For example, in one methodology, each sensor transmits exactly the energy it needs to obtain some received signal-to-noise ratio (SNR) if a target is present at a certain location. This means that the sensing unit should be able to dynamically adjust its transmitted power. Alternatively, each radar sensor could transmit the same energy to the region of interest, so that those sensors which are closer to the target will have a better observation. This might be the situation when the sensing unit is extremely simple and does not allow power variation. For brevity, in this paper we only consider the second case, but its extension to the first case is straightforward. Therefore, let us denote etr as the amount of energy with which each sensor involved in the signal detection uses to illuminate the target. Then the energy received by an antenna at a distance of r meters away will be $\xi.etr/r\gamma$, where ξ is a known constant. Here a simple geometric path loss model [26] is assumed, and the path loss is proportional to $1/r\gamma$, where γ is the path loss exponent, an environment-dependent constant typically between 2 and 4.

A number of energy-aware routing protocols were proposed to seek the energy efficiency of routes in Multihop wireless networks. Among them, several geographical localized routing protocols were proposed to help making smarter routing decision using only local information and reduce the routing overhead. However, all proposed localized routing methods cannot guarantee the energy efficiency of their routes. Numerous energy aware routing protocols have been proposed recently using various techniques (transmission power adjustment, adaptive sleeping, topology control, multipath routing, directional antennas, etc). Most of the proposed energy-aware routing methods take into account the energy-related metrics instead of traditional routing metrics such as delay or hop count. In opposition to these table-driven routing protocols, several stateless routing protocols, particularly, localized geographical routing protocols have been proposed to improve the scalability.

Energy conservation and scalability are probably two most critical issues in designing protocols for Multihop wireless networks, because wireless devices are usually powered by batteries only and have limited computing capability while the number of such devices could be large. In those localized routing protocols, with the assumption of known position information, the routing decision is made at each node by using only local neighbourhood information. Previous localized routing protocols are not energy efficient, i.e., the total energy consumed by their route could be very large compared with the optimal. All of them cannot theoretically guarantee energy efficiency of their routes in the worst case.

Proposed System

A simple localized routing algorithm, called Localized Energy-Aware Restricted Neighbourhood routing (LEARN), which can guarantee the energy efficiency of its route if it can find the route successfully. Being focus on designing routing protocols for Multihop wireless networks which can achieve both energy efficiency by carefully selecting the forwarding neighbours and high scalability by using only local information to make routing decisions. We theoretically prove that LEARN is energy efficient, i.e., when LEARN routing finds a path from the source node to the target node, the total energy consumption of the found path is within a constant factor of the optimum.

We then theoretically study its critical transmission radius in random networks which can guarantee that LEARN routing finds a route for any source and destination pairs asymptotically almost surely. LEARN routing is the first localized routing which can theoretically guarantee the energy efficiency of its routes. In LEARN, whenever possible, the node selects the neighbour inside a restricted neighbourhood that has the largest energy mileage (i.e., the distance travelled per unit energy consumed) as the next hop node.3D wireless network has received significant attention due to its wide range of potential applications (such as underwater sensor networks).As the classical greedy routing, LEARN routing protocol can be directly applied in 3D networks.

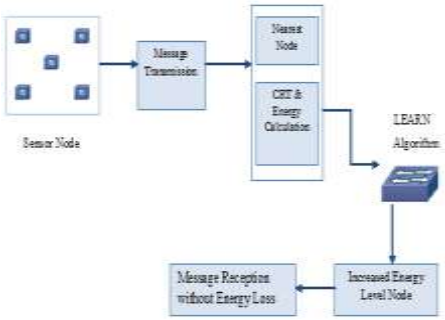


Fig 2 System Architecture

III. Localized Energy Aware Restricted Neighbourhood Algorithm (Learn)

The energy-efficient localized routing protocol for ad hoc networks. Our main contributions are follows. New Localized Routing Protocol: We propose a new localized routing protocol, called localized energy aware restricted neighbourhood routing (LEARN). In LEARN, whenever possible, the node selects the neighbour inside a restricted neighbourhood (defined by a parameter α) that has the largest energy mileage (i.e., the distance travelled per unit energy consumed) as the next hop node. If no such neighbour inside the restricted neighbourhood, it acts as greedy routing. The guarantee of delivery can be achieved by using face routing as the backup.

Power Efficiency of LEARN: We theoretically prove that LEARN is power efficient, i.e., when LEARN routing finds a path from the source node to the target node, the total energy consumption of the found path is within a constant factor of the optimum. Notice that, LEARN routing is the first localized routing which can theoretically guarantee the power efficiency of its routes. In addition, we also prove that the total Euclidean length of the found path is within a constant factor of the optimum.

Critical Transmission Range for LEARN: We theoretically prove that for a network, formed by nodes that are produced by a Poisson distribution with rate n over a compact and convex region Ω with unit area.
Simulation for LEARN: We conducted extensive simulations to study the performance of LEARN and compare it with a typical localized routing protocol (GPSR) and a global ad hoc routing protocol (DSR). The simulation results show that our LEARN routing protocol has good performances in random networks.

Since energy is a scarce resource which limits the life of the network, a number of energy efficient routing protocols have been proposed recently using a variety of techniques. Classical routing algorithm may be adapted to take into account energy-related criteria rather than classical metrics such as delay or hop distance. Most of the proposed energy-aware metrics are defined as a function of the energy required to communicate on a link or a function of the nodes remaining lifetime. However, to minimize the global consumed energy of selected route, most of minimum energy routing algorithms are centralized algorithms. In this paper, we focus on stateless localized routing. Thus, we only review the following related work about power efficient techniques for "local" routing which address how to save energy when making routing decision.

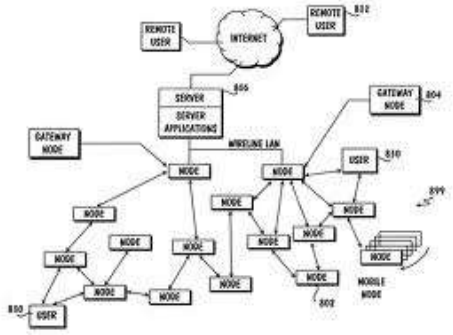


Fig.3 LEARN Algorithm model

Roosta proposed a localized routing method called probabilistic geographic routing to address power-aware routing for ad hoc and sensor networks. It selects the next hop probabilistically among a set of neighbour candidates and the neighbours assign the probability proportional to their residual energy and the link reliability. Melodia et al. proposed a partial topology knowledge forwarding for sensor network, where each node selects the shortest energy-weighted path based on local knowledge of topology. They assumed the neighbourhood discovery protocol provides each node the local knowledge of topology within certain range.

They gave a linear programming formulation to select the range which minimizes the energy expenditure of the network. Since the solution of the linear programming problem is not feasible in practice, they also proposed a distributed protocol to adjust the topology knowledge range. Stojmenovic and Lin proposed a power-aware localized routing which combining the cost metric based on remaining battery power at nodes and the power metric based on the transmission power related to distance between nodes. They proved the loop-free properties of their methods and show their efficiency by experiments. Stojmenovic and Datta further combined the above method with face routing to guarantee the delivery. However, they provide no theoretical guarantee of power efficiency for all of their methods.

Modules Description

Node Creation

When the frame is loading there is no one node has created. By clicking a mouse pointer the nodes created on the regions. Each region have five nodes, even only one region is used for the operation. When the node has created that time node name has added to the node list for selecting the source node and destination node from the regions.

Gathering Information About Nodes

In this module we have to select the source and destination node from the regions, then the transmission of the message to destination starts by clicking the transmit button, then it carefully selecting node for gather the information about each and every node of each region like neighbourhood of source node and who have highest energy in the neighbourhood, etc. Gathering information is based on the shortest or neighbourhood of the source node for calculating the Energy and Scalability. Scalability like transmission power of the nodes.

Calculating CTR, Energy And Scalability

In this module, focus only on designing routing protocols for multihop wireless networks which can achieve both energy efficiency by carefully selecting the forwarding neighbours and high scalability by using only local information to make routing decisions. Thus, to ensure that the routing is successful for every pair of possible source and destination nodes, each node in the network should have a sufficiently large transmission radius such that each intermediate node will always find a better neighbour. Routing method is successful over a network is if the routing method can find a path for any pair of source and destination nodes successfully. Thus calculate the energy of the particular node to transfer the data's and scalability of the node, and also by calculating the critical Transmission Radius(CTR) between the nodes.

Localized Routing

Only by calculating the energy and scalability of the nodes can transfer the data's and to save the energy conservation successfully. The main thing is the node should match the source and destination pairs almost surely. Only review the following related work about energy efficient techniques for localized routing which address how to save energy when making local routing decision. We propose the localized energy aware restricted neighbourhood Routing(LEARN) protocol for wireless networks. So to prove that our LEARN routing protocol is energy efficient if it can find a path. In this module Node successfully selects the source and destination with the intermediate nodes that maintains the energy at the constant high rate.

IV. Possible Attacks And The Counter Measures In The Layers Of Wsn

4.1 Physical Layer

The physical layer is concerned with transmitting raw bits of information over wired/wireless medium. It is responsible for signal detection, modulation, encoding, frequency selection and so on, and is hence the basis of network operations.

4.1.1 Attacks in the Physical Layer

Many attacks target this layer as all upper layer functionalities rely on it. In general, the following three types of attacks are categorized as physical layer attacks:

- Device Tampering
- Eavesdropping
- Jamming

4.1.2 Countermeasures in the Physical Layer

Some attacks in the physical layer are quite hard to cope with. For example, after sensors are deployed in the field, it is difficult or infeasible to prevent every single sensor from device tampering. Therefore, although there are some mechanisms that attempt to reduce the occurrences of attacks, more of them focus on protecting information from divulgence.

- Access Restriction
- Encryption

4.2 MAC Layer

Sensors rely on Medium Access Control (MAC) layer to coordinate their transmissions to share the wireless media fairly and efficiently. Node identifications are embedded in the packets to indicate senders and receivers.

4.2.1 Attacks in the MAC Layer

Due to the openness of wireless channels, the coordination's between sensors based on MAC protocols are subject to malicious manipulation. They can also forge MAC layer identifications and masquerade as other entities for various purposes.

- Traffic Manipulation
- Identity Spoofing

4.2.2 Countermeasures in the MAC Layer

To counter attacks in the MAC layer, current research focuses on detection. It allows for many kinds of further actions to stop the attacks, such as excluding the attacking nodes from interactions. There also exist some prevention approaches, which are mainly against spoofing attacks. Many solutions presented below are actually proposed for ad hoc networks. We believe they can be easily extended to wireless sensor networks.

- Misbehaviour Detection
- Identity Protection

4.3 Network Layer

In the network layer, the key issues include locating destinations and calculating the optimal path to a destination. By tampering with routing service such as modifying routing information and replicating data packets, attackers can fail the communication in WSNs.

4.3.1 Attacks in the Network Layer

As in most other networks, sensors collaborate for routing in WSNs. However, the collaboration between sensors are susceptible to malicious manipulation in WSNs.

- False Routing
    - Overflowing routing tables
    - Poisoning routing tables
    - Poisoning routing caches
- Packet Replication
- Black Hole
- Selective Forwarding

4.3.2 Countermeasures in Network Layer

Since the functionalities of the network layer require the close collaboration of many nodes, all these nodes have to be enclosed for security consideration. It is therefore relatively difficult to mitigate attacks. Nonetheless,some countermeasures are available as follows:

- Routing Access Restriction
- False Routing Information Detection
- Wormhole Detection

4.4 Application Layer

The application layer implements the services seen by users.

4.4.1 Attacks in the Application Layer

Attacks in this layer have the knowledge of data semantics, and thus can manipulate the data to change the semantics. As the result, false data are presented to applications and lead to abnormal actions.

- Clock Skewing
- Selective Message Forwarding
- Data Aggregation Distortion

4.4.2 Countermeasures in the Application Layer

Attacks in the application layer rely on application data semantics. Therefore, the counter measures focus on protecting the integrity and confidentiality of data, no matter it is for control or not.

- Data Integrity Protection
- Data Confidentiality Protection

V. Conclusion

Thus we proposed the localized energy aware restricted neighbourhood routing protocol for wireless networks, we theoretically proved that our LEARN routing protocol is energy efficient if it can find a path. We also studied its critical transmission radius for the successful packet delivery. The mathematical formulation can also extends to any routing protocol in which the region to find the next hop node by an intermediate node is compact and convex. We presented a new algorithm for energy-aware on-line routing of messages in wireless sensor networks. The algorithm uses only a single shortest path computation, and can be implemented in a distributed manner. Without making any assumptions regarding future message arrivals, we proved algorithmic competitive ratio for the case where the routing objective is network capacity maximization. Obtaining this competitive ratio result needs the network to allow admission control. However, we showed that admission control is not needed to ensure good performance in practice. We can also extend the proposed routing method into 3D networks.

References

[1] Minimum energy reliable paths using unreliable wireless links
[2] Optimal local topology knowledge for energy efficient geographical routing in sensor networks
[3] Power and cost aware localized routing with guaranteed delivery in wireless networks
[4] Asymptotic critical transmission radius for greedy forward routing in wireless ad hoc networks
[5] Decentralized detection in sensor networks
[6] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I—Fundamentals," Proc. IEEE, vol. 85, pp. 54–63, Jan. 1997.
[7] R. R. Tunney and N. R. Sandwell, "Detection with distributed sensors," IEEE Trans. Aerospace. Electron. Syst., vol. AES-17, pp. 501–510, Jul.1981.
[8] J.N.Tsitsiklis, "Decentralized detection by a large number of sensors," Math. Control, Signals, Syst., vol. 1, pp. 167–182, 1988.
[9] J.-F. Chamber land and V. V. Veeravalli, "Decentralized detection in sensor networks," IEEE Trans. Signal Process, vol. 51, no. 2, pp.407–416, Feb. 2003.
[10] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: A survey," IEEE Wireless Commun., pp. 6–28, Dec.2004.
[11] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Energy-efficient detection in sensor networks," IEEE J. Sel. Areas Commun., vol. 23, pp. 693–702, Apr. 2005.
[12] L. Yu, L. Yuan, G. Qu, and A. Ephremides, "Energy-driven detection scheme with guaranteed accuracy,"inProc.5thInt.Conf.Information Processing Sensor Networks (IPSN), Apr. 2006, pp. 284–291. [13] W.LiandH.Dai, "Energy efficient distributed detection via multi-hop transmission in sensor networks," in Proc. Military Communications Conf. (MILCOM), Washington, DC, Oct. 2006, pp. 1–7. [14] B.Chen, L.Tong, and P.K.Varshney, "Channel aware distributed detection in wireless sensor networks," IEEE Signal Process. Mag., vol.23, pp. 16–26, Jul. 2006.
[15] R. S. Blum and B. M. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," IEEE Trans. Signal Process, vol. 56, no. 7, pp. 3229–3234, Jul. 2008

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735. Volume 4, Issue 6 (Jan. - Feb. 2013), PP 17-21
www.iosrjournals.org

Printed Coupled Loop Wideband Antenna Design for Wireless Communication

Shoaib Ahmed Khan
COMSATS Institute of Information and Tech, Abbottabad, Pakistan

Abstract: *This paper explains the wideband loop antenna design for wireless communication. Printed antennas are supposed to be more attractive for mobile phone operations for achieving WWAN purposes. In mobile phones theses antenna usually cover the internal circuitry at the top or bottom of the device and they also require further isolation in order to get negligible coupling effect properly that comes due to the coupling between the radiation part of the antenna and the ground plane [1] to [3]. This will limit the internal integration of the antenna with the other circuitry of the cell phone for using it for WWAN purposes for example for achieving better SAR [4]. The proposed antenna consists of the folded loop and some patches in the ground plane to achieve desired goals. The ground plane and feeding part of the antenna are separated by FR4 substrate. Input is given through the capacitively coupled feeding port on the backplane. The antenna is designed to cover ranges of many bands such as, GSM (850/900MHz and 1800/1900MHz), WLAN (2.4GHz) and also the higher frequency bands such as ISM band. The antenna designed, gives very good return loss and has fine radiation pattern. The simulation of antenna in software called HFSS, version 11.0. I personally prefer to design antenna on this software because it gives a lot of customization options.*

I. Introduction

Nowadays the growing scientific research has made life dependent on new technology. The blessing of science has made living quite easy and flexible. First when mobile phone was invented, human became able to make contact with anyone from any corner of the world but with the invention of smart phones and laptops, we became able to carry on our daily updates related to news, weather, accessing our online accounts and even we can now carry on our business while we are on move. Now in this paper an attempt has been made to share the innovative idea in the field of antenna and telecommunication.

From current situation of growing rate of the use of smart phones, has increased the demands for evolution in antenna designing. In smart phones generally PIFA or patch antenna is used. Smart phones usually require multiple bands feature for its various functions and in case of patch antenna, generally a separate antenna is used for each band which are the integrated together to achieve multiple bands. Now in this case, a folded loop antenna has been used which will cover all the desired frequency bands.

The proposed antenna is for handheld devices and will operate on GSM/PCS/UMTS/WLAN bands. The proposed prototype of antenna will be of the folded loop structure and will be suitable to be fitted in a smart phone or in any other device depending on radio technology. The designed antenna will radiate so that its 1^{st} band will occur from 0.85GHz to 2.8GHz with the bandwidth of 1.95GHz and the 2^{nd} band from 3.6GHz to 4.9GHz occupying the bandwidth of 1.3GHz. The dimensions of the antenna are 116mm x 60mm x 10mm. All the simulated and experimental results are shown below in the further description.

II. Antenna Design

The antenna is designed for handheld devices such as smart phones or tablets. At the very bottom we will be having a ground plane, upon which FR4 substrate is used and on the other side of the FR4, coupling strip and feeding strip, which are capacitively coupled.

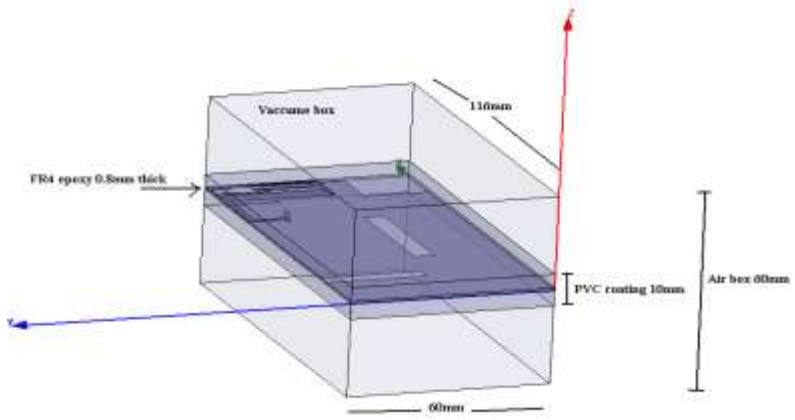


Figure 1: 3d view of the simulated antenna

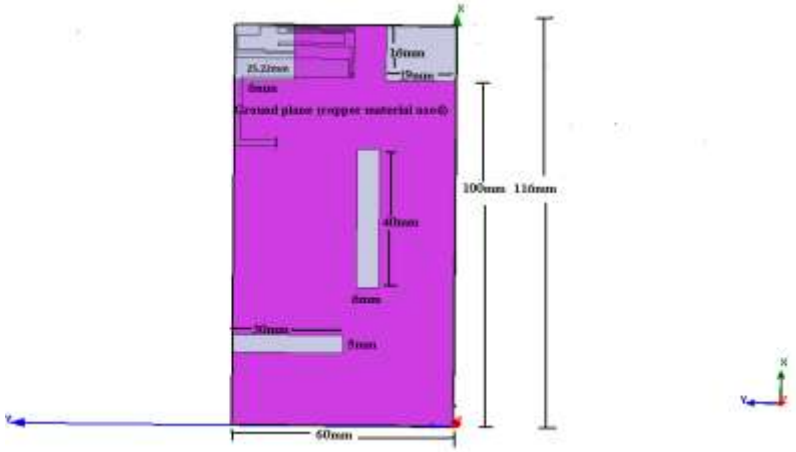


Figure 2: Ground plane

Starting from the ground plane, the total dimensions of the ground plane are 100mm x 60mm and its protruded part of the ground plane is 16mm x 15.78mm. There are also no-ground plane regions on the both sides of the protruded ground plane, one with the dimensions of 16mm x 25.22mm and other with 16mm x 19mm. The no-ground planes are defined to allow some free space for other components of antenna. The protruded ground plane improves overall efficiency of antenna. The dimensions of the FR4 substrate are 116mm x 60mm and thickness of 0.8mm with relative permittivity of 4.4. There are two strips on the antenna both are on designed over the FR4, the feeding strip and coupling strip. When input is applied at the feeding port of antenna having input resistance of 50Ω, gives better resonance specially at the lower bands of GSM[5].The coupling strip is shortened with the ground plane via cylinder. The input energy to antenna is given via feeding port located on the end of the feeding strip. In order to increase the electrical length of the antenna, the loop of the antenna has been folded many times. The antenna can also be coated with PVC of thickness 10mm to make it suitable for use in smart phone.

Figure 3: Coupling strip

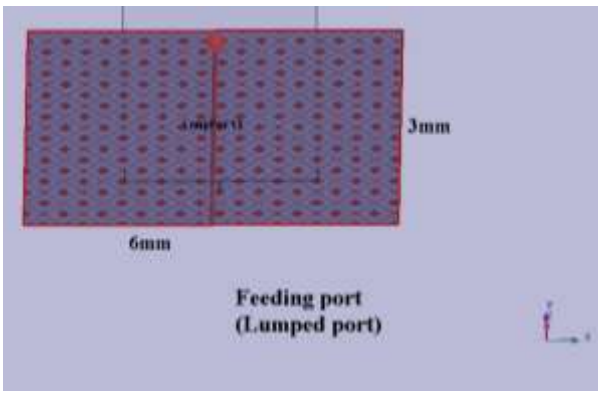


Figure 4: Feeding port

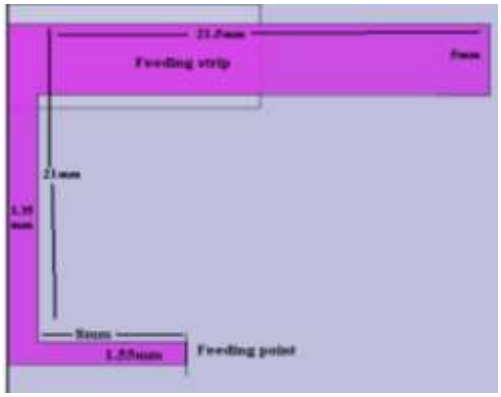


Figure 5: Coupling strip

Figure 6: Hardware image (Ground plane)

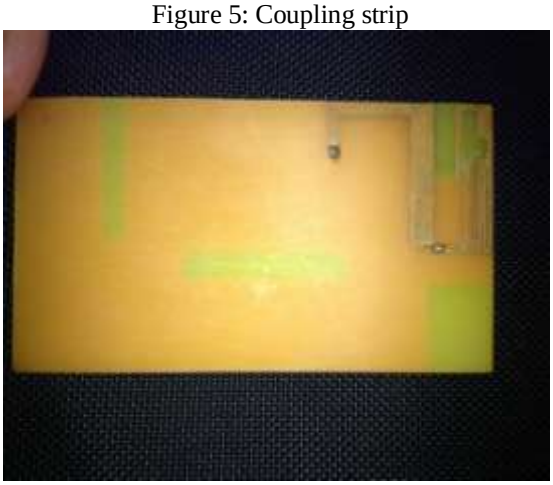


Figure 7: Resonating part of antenna

There are also two rectangular slots cut in the ground plane. One slot is with the dimensions of 40mm x 6mm and other slot with dimensions of 5mm x 30mm. These slots helped in improving the return loss and increased the electrical length of the antenna.

III. Results

The proposed antenna has been designed and simulated in Ansoft HFSS. Then after simulation, antenna was fabricated and tested in Telecommunication Lab of COMSATS Institute of Information and Tech, Abbottabad, Pakistan. And the measured results were taken using network analyzer. The simulation and the measured results are shown below in figure.

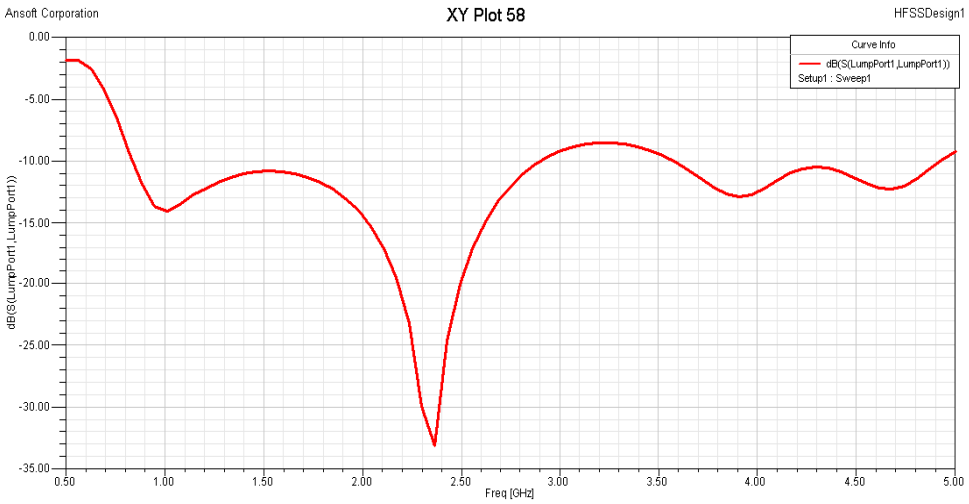


Figure 8: Results obtained from simulation in Ansoft HFSS

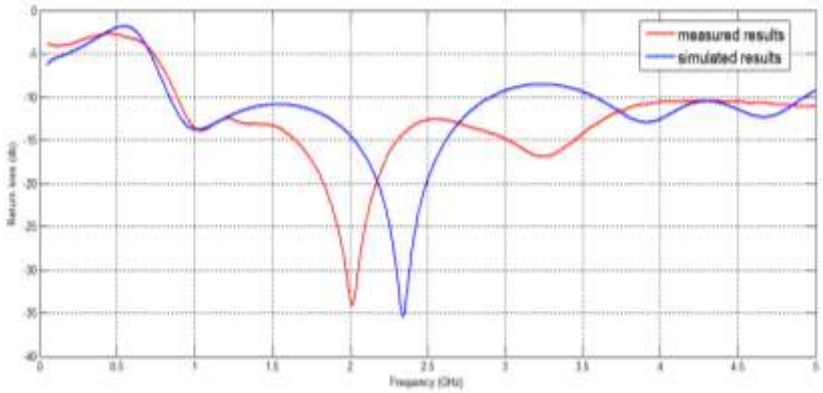


Figure 9: Comparison of both simulated results (one with blue color) and the measured results (one with red color)

In the measured results the graph has been shifted little on left, this is because of the coaxial cable used for taking readings was not properly covered, which resulted in effecting the lower frequencies. The antenna is covering GSM/UMTS/WWAN/PCS/ISM. The input impedance is set 50 ohm to achieve the mentioned goals. From results it is quite clear that the antenna is covering two frequency bands, the first one is an Ultra wide band and the second one is wide band. The 1ST band Is starting from 0.85Ghz to 2.8GHz which Is covering all GSM/UMTS/PCS/WWAN bands and the 2nd band is starting from 3.6Ghz to 4.9GHz which is basically covering ISM band. The radiation pattern and the gain of the antenna are giver below.

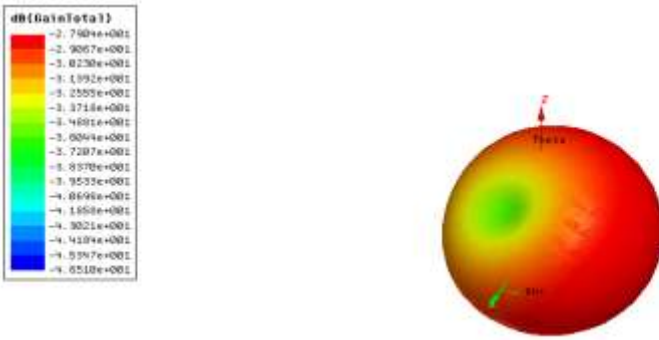


Figure 10: Total gain of antenna

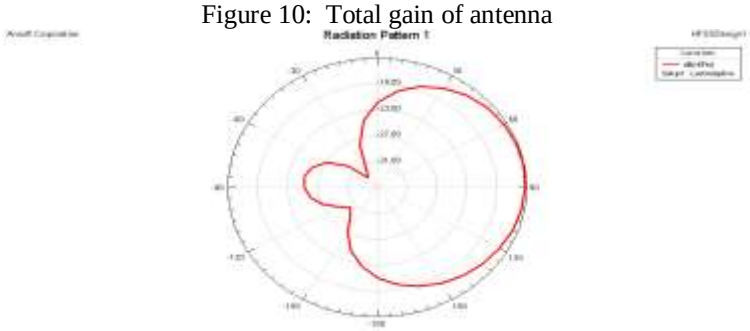


Figure 11: Radiation pattern of antenna (xy-plane)

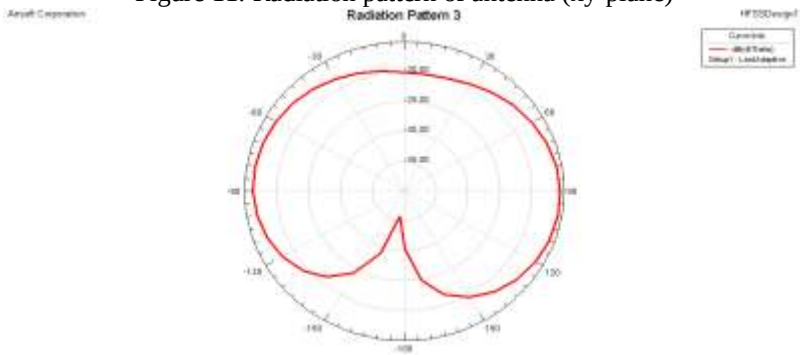


Figure 12: Radiation pattern of antenna (yz-plane)

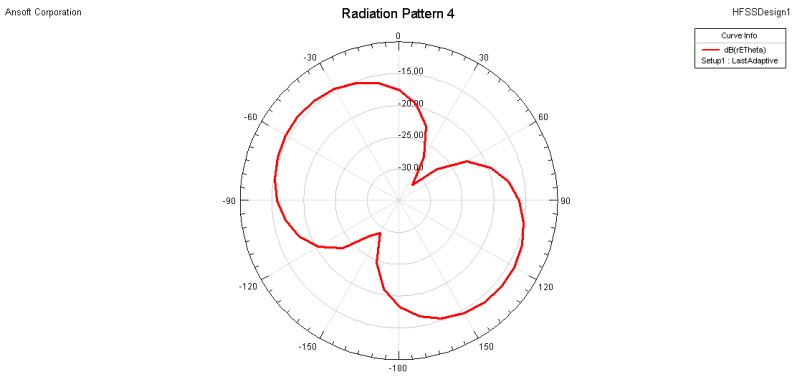


Figure 13: Radiation pattern of antenna (xz-plane)

IV. Conclusion

First of all it concluded that loops antenna is a very good choice for smart phones or other devices working in the frequency range of this antenna. By using the technique of the folding the loop of antenna, you can get more electrical length by maintaining the overall size of antenna constant. The protruded ground plane plays an important role in the getting better return loss and better gain. By cutting the parts of ground plane we can increase the electrical size of antenna and can also get the better return loss.

From results it is quite clear that this antenna can be used in almost all the wireless devices that operate on frequency range of 850MHz to 2.8GHz or from 3.6GHz to 4.9GHz. This is a cost effective antenna and suits best for the handheld devices.

References

[1] C. M. Su, K. L. Wong, C. L. Tang and S. H. Yeh, "EMC internal patch antenna for UMTS operation in a mobile device," IEEE Trans. Antennas Propagat., vol. 53, pp. 3836-3839, 2005.

[2] K. L. Wong and C. H. Chang, "Surface-mountable EMC monopole chip antenna for WLAN operation," IEEE Trans. Antennas Propagat., vol. 54, pp. 1100-1104, 2006.

[3] C. H. Wu, K. L. Wong and J. S. Row, "EMC internal GSM/DCS patch antenna for thin PDA phone application," Microwave Opt. Technol. Lett., vol. 49, pp. 403-408, 2007.

[4] American National Standards Institute (ANSI), "Safety levels with respect to human exposure to radiofrequency electromagnetic field, 3 kHz to 300 GHz," ANSI/IEEE standard C95.1, 1999.

[5] Y. W. Chi and K. L. Wong, "Quarter-wavelength printed loop antenna with an internal printed matching circuit for GSM/DCS/PCS/UMTS operation in the mobile phone," IEEE Trans. Antennas Propagat., vol. 57, pp. 2541-2547, 2009.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735. Volume 4, Issue 6 (Jan. - Feb. 2013), PP 22-27
www.iosrjournals.org

Enhancement Of Satellite Image Resolution With Moving Objects

A. Abirami[1], N. Akshaya[2], D. Poornakala[3], D. Priyanka[4], C. Ram kumar[5]
[1, 2, 3, 4] *UG Students, SNS College of Engineering, Coimbatore - 641107.*
[5]*Assistant professor, SNS College of Engineering, Coimbatore - 641107.*

Abstract: *Satellite images are being used in many applications like Meteorology, Agriculture, Geology, Forestry, Landscape, Biodiversity, Planning, Education, Regional, Seismology and oceanography. The Enhanced satellite images make diagnostic details more obvious. The Image Enhancement is the main technique for improving the resolution and visual appearance of the image. One of the major issues in Image Enhancement is Wavelet Transform. The Wavelet Transform is the method which decomposes an image into a set of basic functions called Wavelets. These basis functions are limited in duration and are inherently local. A Resolution Enhancement technique is based on the Interpolation of the high-frequency subbands obtained by Discrete Wavelet Transforms (DWT). Bicubic interpolation is used as a intermediate stage for estimating high frequency components and it is more sophisticated than the nearest neighbor and bilinear techniques. The proposed technique has the advantages of superior resolution, sharper image and smoother edges. The PSNR improvement of the proposed technique is up to 7.19dB.*
Keywords – *Image resolution, Edge detection, transforms, cycle spinning*

I. Introduction

Image enhancement is a method of improving the definition of a video picture by a computer program, which reduces the lowest gray values to black and the highest to white, for the pictures from microscopes, surveillance cameras, and scanners. Improvement of the quality of a picture, with the aid of a computer, by giving it higher contrast or making it less blurred or less noisy. Image enhancement techniques can be divided into two broad categories are Spatial domain methods, which operate directly on pixels and frequency domain methods, which operate on the Fourier transform of an image.

Interpolation in image processing is a method to increase the number of pixels in a digital image. Interpolation has been widely used in many image processing applications, such as facial reconstruction, multiple description coding, and image resolution enhancement .The interpolation-based image resolution enhancement has been used for a long time and many interpolation techniques have been developed to increase the quality of this task.

Wavelets are also playing a significant role in many image processing applications. The 2-D wavelet decomposition of an image is performed by applying the 1-D discrete wavelet transform (DWT) along the rows of the image first, and then the results are decomposed along the columns. This operation results in four decomposed sub-band images referred to low-low (LL), low-high (LH), high-low (HL), and high-high (HH). The frequency components of those sub-bands cover the full frequency spectrum of the original image. A satellite image resolution enhancement technique is based on the interpolation of the high-frequency subband obtained by discrete wavelet transform (DWT) and the input image. This resolution enhancement technique uses DWT to decompose the input image into different subband. Then, the high-frequency subband images and the input low resolution image have been interpolated, followed by combining all these images to generate a new resolution enhanced image.

II. Literature Survey

2.1 Cycle Spinning And Edge Modelling

Recursive Cycle Spinning (RCS) was employed to avoid the artifacts. This method can get the better visual effect and PSNR value compared with the methods like wavelet image de-noising using the Recursive Cycle Spinning[16]. Artifact is a classic signal interpolation problem and conventional approaches such as zero-order interpolation (sample-and-hold) cause severe pixelation impairments while bilinear and spline interpolation invariably result in undesirable levels of smoothing across salient edges. Recently several efforts in the field have utilized wavelet-domain methodologies with the intention of overcoming some of the problems associated with conventional treatment[2]. A common feature of these algorithms is the assumption that the low-resolution (LR) image to be enhanced is the low-pass filtered sub-band of a high-resolution (HR) image which has been subjected to a decimated wavelet transform.

2.2 Edge Detection For Images

Edge detection is a fundamental tool in image processing, machine vision and computer vision, particularly in the areas of feature detection and feature extraction[12]. which aim at identifying points in a digital image at which the image brightness changes sharply or, more formally, has discontinuities. Edges are significant local changes of intensity in an image[14]. Edges typically occur on the boundary between two different regions in an image.

An edge may be loosely defined as a local discontinuity in the pixel value that exceeds a given threshold. There are many steps in edge detection which are as follows:

Smoothing: Suppress as much noise as possible, without destroying the true edges.

Enhancement: Apply a filter to enhance the quality of the edges in the image (sharpening).

Detection: Determine which edge pixels should be discarded as noise and which should be retained (usually, thresholding provides the criterion used for detection).

Localization: Determine the exact location of an edge (sub-pixel resolution might be required for some applications, that is, estimate the location of an edge to better than the spacing between pixels). Edge thinning and linking are usually required in this method.

2.3 Image Resolution Enhancement Using Inter-Subband Correlation In Wavelet Domain

Conventional image resolution enhancement methods, such as bilinear and bicubic interpolation methods may generate false information and blurred images because they do not utilize any information relevant to edges in the original image [13]. Wavelet-based methods enhanced the image resolution by estimating the preserved high frequency information from the given images [7]. They were based on the assumption that the image to be enhanced was the low frequency subband among wavelet-transformed subband of the original one and the target is to estimate the high frequency subband of wavelet transform, so that a resolution enhanced image can be obtained[1] . Because the analysis filter bank used in the wavelet transform has a poor frequency characteristic such as wide transition region, some information on the high-frequency band has remained in the low frequency band. The phase - shifting matrix was first derived by XinLi, for saving the computational complexity of LBS method[10] . It showed that two sub-bands of a two-channel perfect reconstruction filter bank was linked by a unique phase-shifting matrix. Since it is customary to put the center of an odd symmetric filter at 0, for odd length biorthogonal filters, the sample positions in the low-frequency band and the high frequency band are alternative, which means even-phase for low frequency band and odd phase for high frequency band.

2.4 CYCLE SPINNING

The decimated wavelet transform is not shift-invariant and as a result, distortion of wavelet coefficients, due to quantization of coefficients in compression applications or non-exact estimation of high-frequency coefficients in resolution enhancement applications (including zero padding of coefficients as in WZP) , introduces cyclostationarity into the image which manifests itself as ringing in the neighborhood of discontinuities.

Cycle-spinning (CS) is an effective method against ringing when used for de-noising purposes in the wavelet domain and also for reducing ringing and increasing the perceptual quality of compressed images [16]. CS method aims to approximate shift-invariant statistics by averaging out the cyclostationarity[1]. CS can also be applied as a post-processing operation after decompression results in significant improvements in the framework of JPEG and JPEG2000 image compression.

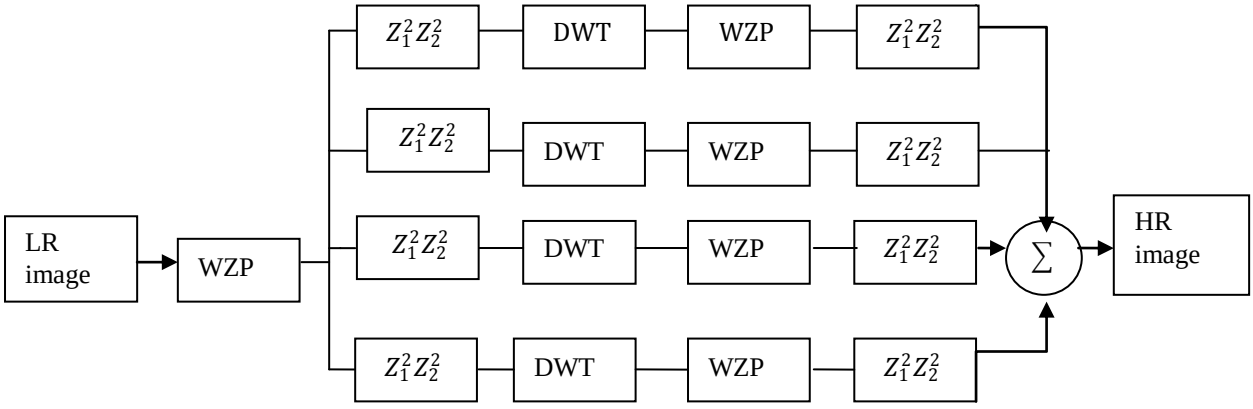


Fig 2.1 WZP- and CS-based image resolution enhancement.

III. Enhancement Of Satellite Image Resolution

A discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discretely sampled. The Wavelet Series is just a sampled version of CWT and its computation may consume a significant amount of time and resources, depending on the resolution required. The Discrete Wavelet Transform (DWT), which is based on sub-band coding, is found to yield a fast computation of Wavelet Transform. It is easy to implement and reduces the computation time and resources required [5]. As with other wavelet transforms, a key advantage it has over Fourier transforms is temporal resolution: it captures both frequency and location information.

The transform of a signal is just another form of representing the signal. It does not change the information content present in the signal. The Wavelet Transform provides a time-frequency representation of the signal[8] . It was developed to overcome the shortcoming of the Short Time Fourier Transform (STFT), which can also be used to analyze non-stationary signals.

3.1 Dwt-Based Resolution Enhancement

DWT has been employed in order to preserve the high-frequency components of the image. DWT separates the image into different sub-band images, namely, LL, LH, HL, and HH. High-frequency sub-bands contain the high frequency component of the image[3]. The interpolation can be applied to these four sub-band images. In the wavelet domain, the low-resolution image is obtained by low-pass filtering of the high-resolution image. The low resolution image (LL sub-band), without quantization is used as the input for the proposed resolution enhancement process[9].

In other words, low frequency sub-band images are the low resolution of the original image. Therefore, instead of using low-frequency sub-band images, which contains less information than the original input image, we are using this input image through the interpolation process. Hence, the input low-resolution image is interpolated with the half of the interpolation factor, $\alpha/2$, used to interpolate the high-frequency sub-bands, as shown in fig 3.2. In order to preserve more edge information, i.e., obtaining a sharper enhanced image, we have proposed an intermediate stage in high frequency sub-band interpolation process.

As shown in Fig. 3.2, the low-resolution input satellite image and the interpolated LL image with factor 2 are highly correlated. The difference between the LL sub-band image and the low-resolution input image are in their high-frequency components. Hence, this difference image can be used in the intermediate process to correct the estimated high-frequency components. This estimation is performed by interpolating the high frequency sub-bands by factor 2 and then including the difference image (which is high-frequency components of low-resolution input image) into the estimated high-frequency images, followed by another interpolation with factor $\alpha/2$ in order to reach the required size for IDWT process.

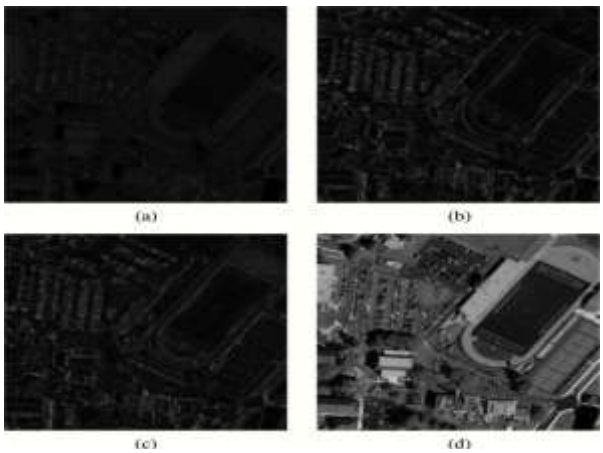


Fig.3.1 (a) The proposed enhanced image, (b) the standard bicubic interpolation, (c) The WZP- and CS-based image resolution enhancement technique and (d) the original high-resolution satellite image

Fig 3.1(a)–(c)show the difference between the high resolution images with the enhanced image by using the proposed resolution enhancement technique, the difference obtained by using bicubic interpolation directly, and the difference image with WZP and CS-based image resolution enhancement technique, respectively. Fig 3.1 (d) shows that higher-frequency components have been preserved in the proposed technique.

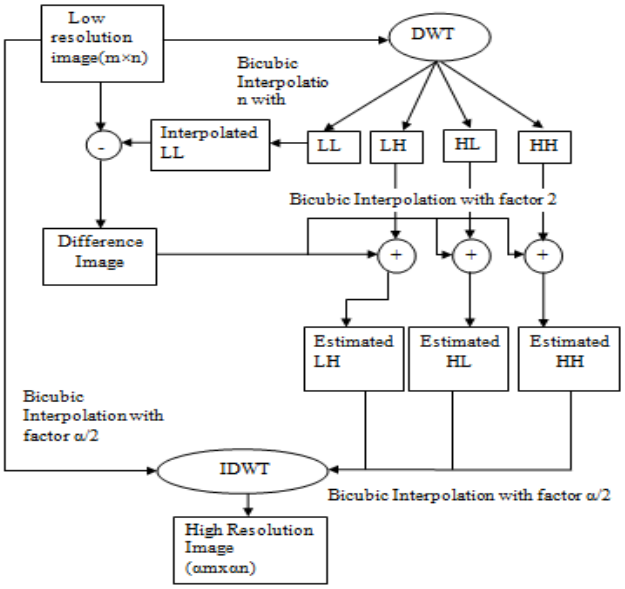


Fig 3.2 DWT resolution enhancement algorithm.

The intermediate process of adding the difference image, containing high-frequency components, generates significantly sharper and clearer final image. This sharpness is boosted by the fact that, the interpolation of isolated high-frequency components in HH, HL, and LH will preserve higher-frequency components than interpolating the low-resolution image directly.

Not only visual comparison but also quantitative comparisons is confirming the superiority of the proposed method. Peak signal-to-noise ratio (PSNR) and root mean square error (RMSE) have been implemented in order to obtain some quantitative results for comparison. PSNR can be obtained by using the following formula:

$$PSNR = 10\ log_{10}\left(\frac{R^2}{MSE}\right) \qquad (1)$$

Where *R* is the maximum fluctuation in the input image (255 in here as the images are represented by 8 bits) and MSE is representing the MSE between the given input image I_{in} and the original image I_{org} which can be obtained by the following:

$$MSE = \frac{\sum_{i,j}(I_{in}(i,j) - I_{org}(i,j))^2}{M \times N} \qquad (2)$$

Where *M* and *N* are the size of the images.

Clearly, RMSE is the square root of the MSE, hence it can be calculated by the following:

$$RMSE = \sqrt{\frac{\sum_{i,j}(I_{in}(i,j) - I_{org}(i,j))^2}{M \times N}} \qquad (3)$$

IV. Implementation

4.1 Results And Discussion

The implementation is done using MATLAB. The proposed resolution enhancement technique uses DWT to decompose the input image into different subband. Then, the high-frequency subband images and the input low-resolution image have been interpolated, followed by combining all these images to generate a new resolution-enhanced image by using inverse DWT. In order to achieve a sharper image, an intermediate stage (interpolation technique) has been used.

The discrete wavelet transform is applied to the input image and the four subband are obtained as shown in the Fig.3.3.

Bicubic interpolation of isolated high-frequency components in HH, HL, and LH will preserve higher-frequency components than interpolating the low-resolution image directly and they are shown in the Fig 3.4.

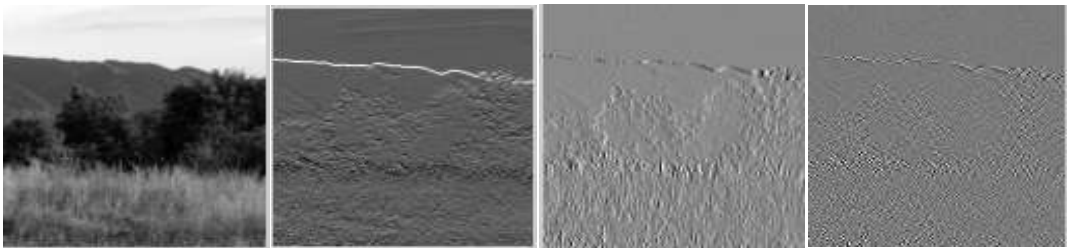


Fig 3.3 LL, LH, HL, HH subband of input image

Bicubic interpolation of isolated high frequency components in HH, HL, and LH will preserve higher-frequency components than interpolating the low-resolution image directly and they are shown in the Fig 3.4.

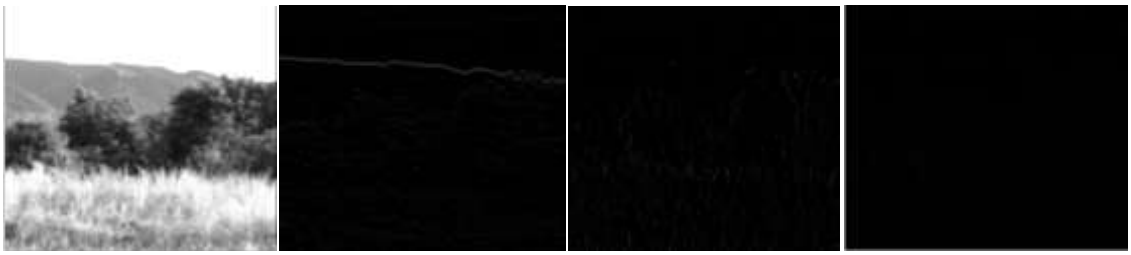


Fig 3.4 Bicubic interpolation of LL,LH,HL,HH subband.

The proposed technique has been tested on several different satellite images and they have superior resolution, smoother edges and sharper image over the conventional and state-of-art techniques from a visual point of view. The enhanced images by using bicubic interpolation, WZP and CS-based image resolution enhancement methods and the enhanced images obtained by the proposed technique are analyzed. It is clear that the resultant image, is sharper than the other techniques.

Table 1.1 PSNR, RMSE Results for DWT Method.

IMAGE	RMSE	PSNR
TEST IMAGE	3.79	37.8
MOON IMAGE	3.99	38.9
WATER IMAGE	4.01	40

The PSNR improvement of the proposed technique is up to 7.19 dB compared with the standard bicubic interpolation. The PSNR and RMSE values are tabulated in the table 1.1. As expected, highest level of information content is embedded in the original images. The main reason of having the relatively high information content level of the images generated by the DWT technique is due to the fact that the unquantized input LL subband images contain most of the information of the original high-resolution image.

4.2 SIMULATED OUTPUT

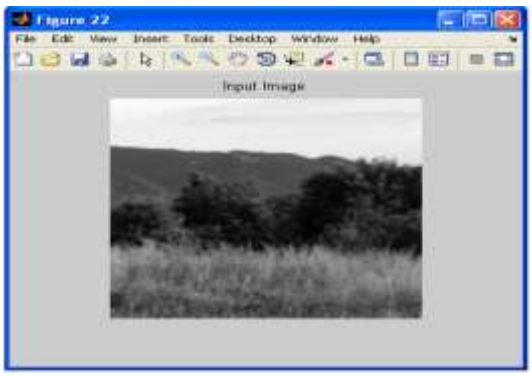


Fig4.1 Imput Image

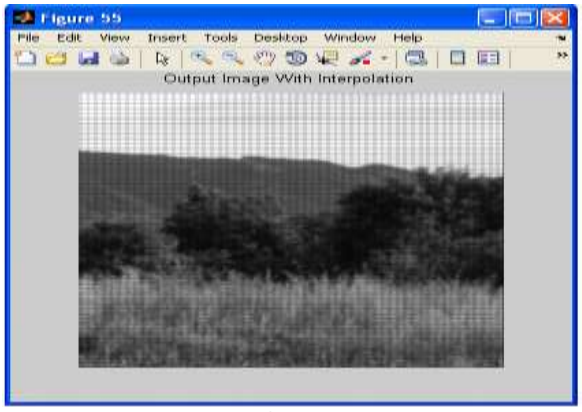


Fig 4.2 Output Image

V. Conclusion & Future Work

This paper has proposed a new resolution enhancement technique based on the interpolation of the high-frequency subband images obtained by DWT and the input image. The proposed technique has been tested on well-known benchmark images, where their PSNR, RMSE and visual results show the superiority of the proposed technique over the conventional and state-of-art image resolution enhancement techniques. The PSNR improvement of the proposed technique is up to 7.19 dB compared with the standard bicubic interpolation.

For future work of the project may include color image enhancement instead of gray scale image enhancement. It also improves the efficiency of the input image by removal of noise. Also it may include the enhancement of satellite image with moving objects.

References

[1] G. Anbarjafari and H. Demirel, "Image super resolution based on interpolation of wavelet domain high frequency subbands and the spatial domain input image," *ETRI J.*, vol. 32, no. 3, pp. 390–394, Jun. 2010.

[2] S.G Chang, Z. Cvetkovic and M. Vetterli, "Resolution enhancement of image using wavelet transform extrema ex-trapolation", *Proc.ICASSP'95*, vol.4,pp.2379-2382, May 1995

[3] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. SignalProcess.*, vol. 46, no. 4, pp. 886–902, Apr. 1998.

[4] H. Demirel, G. Anbarjafari, and S. Izadpanahi, "Improved motion-based localized super resolution technique using discrete wavelet transform for low resolution video enhancement," in *Proc. 17th EUSIPCO*, Edinburgh, U.K., Aug. 2009, pp. 1097–1101.

[5] H. Demirel and G. Anbarjafari, "Satellite image resolution enhancement using complex wavelet transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 1, pp. 123–126, Jan. 2010.

[6] Glenn, W., Daniel, J.J., Zia, U.R. and Glenn, H. (2005) 'Enhancement of Imagery in Poor Visibility Conditions', Proceedings of SPIE, Vol. 5778, No.05, pp. 889.

[7] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 2007.

[8] HasanDemirel and GholamrezaAnbarjafari,"Discrete Wavelet Transform Based Satellite Image Resolution enhancement",IEEETrans.geosci.RemoteSens.,vol.49,no.6,pp.1997-2004,june 2011.

[9] H. Demirel, G. Anbarjafari, and S.Izadpanahi, "Improved motion-based localized super resolution technique using discrete wavelet transform for low resolution video enhancement," in *Proc. 17th EUSIPCO*, Edinburgh,U.K., Aug.2009, pp. 1097–1101.

[10] K. Kinebuchi, D. D. Muresan, and T.W. Parks, "Image interpolation using wavelet based hidden Markov trees," in *Proc. IEEE ICASSP*, 2001, vol. 3, pp. 7–11.

[11] N.G. Kingsbury. The dual-tree complex wavelet transform with improved orthogonality and symmetry properties. *IEEE International Conference on Image Processing*, pages 375–378, September 2000

[12] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[13] N. Nguyen, P. Milanfar, "An efficient wavelet-based algorithmfor image superresolution", *Proc. ICIP '00*, vol.2, pp. 351-354, Sep. 2000

[14] Y. Piao, L. Shin, and H. W. Park, "Image resolution enhancement using inter-subband correlation in wavelet domain," in *Proc. IEEE ICIP*, 2007, vol. 1, pp. I-445–I-448.

[15] L. A. Ray and R. R. Adhami, "Dual tree discrete wavelet transform with application to image fusion," in *Proc. 38th Southeastern Symp. Syst.Theory*, Mar. 5–7, 2006, pp. 430–433.

[16] Temizel and T. Vlachos, "Wavelet domain image resolution enhancement using cycle-spinning," *Electron. Lett.*, vol. 41, no. 3, pp. 119–121, Feb. 3, 2005.

[17] Waheeb, A.U., Akram, M. and Ziad, A.A. (2009) ' Gray Image Reconstruction', European Journal of Scientific Research, ISSN 1450-216X, Vol. 27 ,pp.167-173.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 28-31
www.iosrjournals.org

A review on RF based intruder detection using CW Doppler radar

Lynette M. T.[1], J. M. Kudargi[2]
[1,2](EXTC Department,K. J. Somaiya CoE, University of Mumbai, India)

ABSTRACT : *In recent years, electromagnetic waves have become the medium of information exchange in free space for various applications ranging from home appliances to industrial products, from research laborites to defense applications, from mobiles to health monitoring systems. The development in the non-contact based intruder detection system via human vital sign using Doppler radar has attracted considerable interest of researchers. Due to the non contact nature and deep penetration level of radio frequencies, the vital sign detection system includes various domain of application, viz. medical, security, etc. this paper presents review of RF based intruder detection using CW Doppler radar. In this paper, vital sign detection techniques are presented.*

Keywords - *CW Doppler radar, heartbeat, intruder, radar receivers, respiration rate.*

I. INTRODUCTION

The word protection of the belongings leads to the implementation of security system. Various types of security systems are available today with the types varying with respect to the applications. For intruder detection, various types of burglar alarms are introduced today with the advantages of easy availability, cost effectiveness etc. but, these systems also have drawbacks, such as high rate of false alarms, weather effect, animal activity, air movement, machinery noise, vibration, change in temperature. In 1886, Heinrich Hertz discovered the radiation of electromagnetic waves while their propagation. This led to the discovery of radar technology which was used as intruder sensor in WWII. As the advance use of radar technology for various applications increases rapidly, the recent area of research is human vital sign detection system. In case of intruder detector, where conventional alarms has drawbacks in through the wall detection, weather conditions, the RF based vital sign detection system can provide good solution. This is due to the fact that, RF signals can pass through any weather conditions, they can pass through walls. This system is used when continuous monitoring is required. This paper gives an overview of human vital sign detection system.

II. DOPPLER RADAR THEORY

When a source generating waves moves relative to an observer, or when an observer moves relative to a source, there is an apparent shift in frequency. If the distance between the observer and the source is increasing, the frequency apparently decreases, whereas the frequency apparently increases if the distance between the observer and the source is decreasing. This relationship is called Doppler Effect (or Doppler Shift) after Austrian Physicist Christian Johann Doppler (1803-1853).[1]

III. METHODOLOGY

According to Doppler theory, a constant frequency signal reflected off an object with a periodically varying displacement will result in a reflected signal at the same frequency, but with a time varying phase $\varphi(t)$.[2]

Doppler radar motion sensing systems typically transmit a continuous wave (CW) electromagnetic signal that is reflected off a target and then demodulated in the receiver.

$$R(t) \approx \cos\left[2\pi ft - \frac{4\pi d_0}{\lambda} - \frac{4\pi x(t)}{\lambda}\right] \quad (1)$$

The received signal consists of the reflected transmitter signal, the Doppler shift d_0 with time delay related to the target distance, where λ is wavelength.[3] A stationary person's chest can be considered as the target of CW radar since it has a periodic movement with no net velocity and hence will therefore receive a signal similar to the transmitted signal, with its phase modulated by the time-varying chest position, as shown in Figure 1. Demodulating the phase will then provide a signal directly proportional to the chest position, which contains information about movement due to heartbeat and respiration, from which heart and respiration rates can be determined. Non-contact heart and respiration monitors have been developed based on this principle [2].

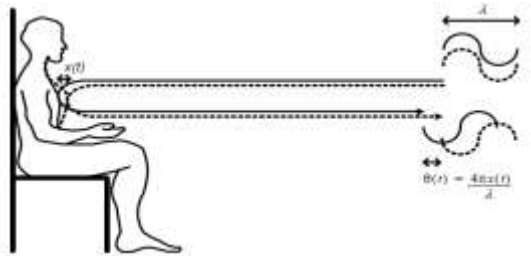


Figure 1: Block diagram of continuous wave radar for measurement of physiological motion. The phase of the reflected signal, θ(t), is directly proportional to the chest motion, x(t), and is scaled by the wavelength, λ.

IV. SYSTEM DESIGN

The various types of the vital sign monitoring system architecture can be classified on the basis of their receiver design. Receiver used in CW Doppler radar can be direct conversion receiver or quadratuer Doppler receiver. According to the Doppler shift, if the target surface has periodic movement, it is then characterized as phase shift w.r.t to the displacement. Also, if this movement is very small, then at the receiver side, mixer produces an output of low frequency component proportional to the displacement. Fig (2) shows this concept. It is a single channel, direct conversion, CW radar. A major limitation of the single channel configuration is detection sensitivity to target position due to a periodic phase relationship between the received signal and local oscillator.[1,3]

The received signal can be given as:

$$R(t) = A_R \cos\left[2\pi ft - \frac{4\pi d_0}{\lambda} - \frac{4\pi x(t)}{\lambda} + \phi\left(t - \frac{2d_0}{c}\right) + \theta\right] \tag{2}$$

Received signal which is the RF input to the mixer is

$$R(t) = A_R \cos\left[2\pi ft - \frac{4\pi d_0}{\lambda} - \frac{4\pi x(t)}{\lambda} - \frac{4\pi y(t)}{\lambda} + \phi\left(t - \frac{2d_0}{c}\right) + \theta\right] \tag{3}$$

Here, *x(t)* is the signal due to respiration and *y(t)* is the signal due to the heart beat. LO signal to be given to the mixer is

$$C(t) = A_L \cos(2\pi ft + \phi(t))$$

Hence the low frequency signal components obtained after filtering at the output of the mixer is

$$D(t) = \cos\left[\theta + \frac{4\pi x(t)}{\lambda} + \Delta\phi(t)\right] \tag{4}$$

Here, $\Delta\varphi(t)$ is the residual phase noise in the baseband demodulated signal

$$\Delta\phi(t) = \left[\phi(t) - \phi\left(t - \frac{2d_0}{c}\right)\right] \tag{5}$$

θ is the constant phase shift dependant on the nominal distance to the target *d0*

$$\theta = \left[\frac{4\pi d_0}{\lambda} + \theta_0\right] \tag{6} [4]$$

One major disadvantages of using a homodyne receiver for vital signs monitoring is the amount of DC offset introduced by the system; this makes the signal demodulation difficult. This is due to the vital signal occurs at the low frequency, which is very close to the DC value of the signal.[4]

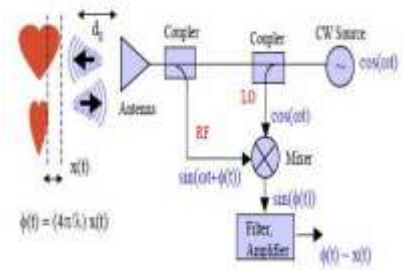


Figure 2: Vital signs remote monitoring Doppler radar system (from [1,3])

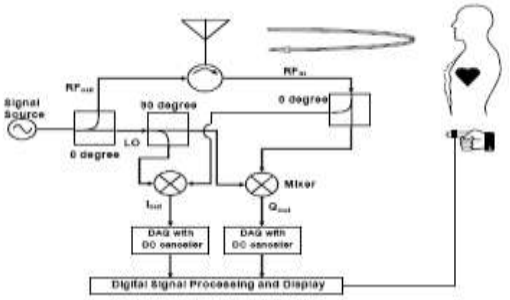


Figure 3: Block diagram of a quadrature Doppler radar system with dc canceller. (from [5])

Fig. (3) Shows the quadreture bio radar sensor. This architecture consists of two mixers which provide output on I channel and its 90^0 phase shift on Q channel. A channel selector is used to select the optimum output from both the channels. This architecture has advantage of DC cancellation as compared to the single channel receiver.[3] the baseband signal is filtered, amplified and then processed using DSP tools. While quadrature Doppler receiver has more accurate results than direct conversion receiver, the main drawback of this system is large circuitry due to two mixers, phase shifters. Two demodulation methods of quadrature outputs have been proposed: linear (complex) and non-linear (arctangent). Linear demodulation essentially rotates the data to the optimum position, and at lower frequencies, and for small displacements this technique yields accurate results. The arctangent demodulation overcomes the sensitivity to the target position, while also extending the reliable phase deviation detection range that is limited by small angle approximation in single channel receivers.[3]
The quadrature baseband output is given as

$$B(t) = A_R \exp[\theta + 4\pi\Delta x(t)/\lambda] \tag{7}$$

Where θ is the constant phase shift related to the phase change at the surface of a target and the phase delay between the mixer and antenna.
By arctangent demodulation technique

$$\phi(t) = \arctan\left(\frac{B_Q(t)}{B_I(t)}\right) = \arctan\left(\frac{V_Q + A_R \sin(\theta + 4\pi\Delta x(t)/\lambda)}{V_I + A_R \cos(\theta + 4\pi\Delta x(t)/\lambda)}\right) \tag{8}$$

Where V_Q and V_I and refer to the dc offsets of each channel resulting from the finite port to port isolation of the transceiver as well as from clutter reflections. The dc signal contains this dc offset as well as the dc information associated with target's position required for accurate demodulation.[3,5]

V. RESULTS

The microwave CW radar vital sign monitoring system can be operated at various frequencies and at various conditions. The higher the operating frequency, the more improved detection sensitivity we obtain.[6,7] The fft results of the signal obtained from fig. (4) gives the peaks of respiration frequency at (0.2 Hz to 0.5 Hz) and the peaks of heartbeat frequency at 0.8 Hz to 2.5Hz. Figure below shows the Fast Fourier Transform of the time domain signal obtained.[1]

Fig 5(a) below shows the Doppler output with reference signal taken from finger. It clearly shows that heartbeat and respiration rates can be obtained using Doppler radar architecture. Fig. 5(b) shows the calculated heartbeat and respiration rates obtained from I and Q channels with arctangent and linear demodulation techniques.[3]

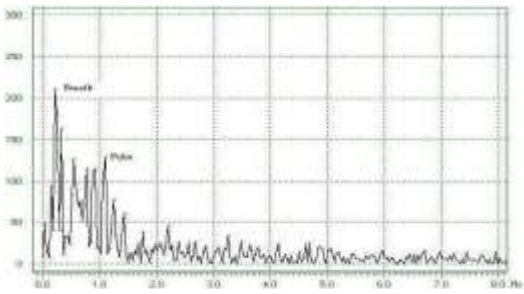


Figure 4: Frequency spectrum of breathing and heartbeat, D=1m (from [1])

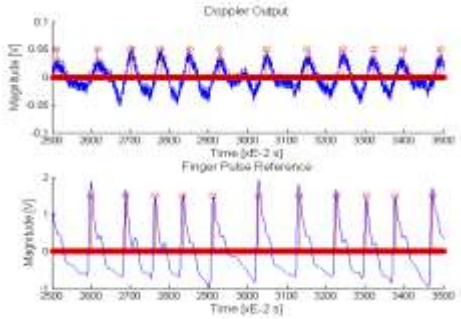


Figure: 5(a)

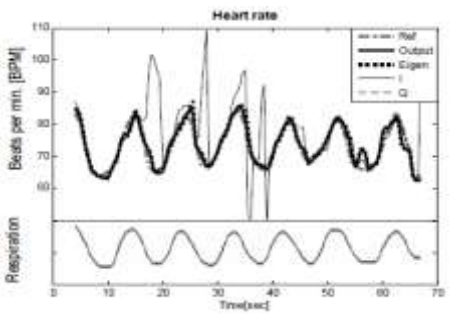


Figure: 5(b)

Figure 5: Doppler radar output signal (top) and finger pulse reference (bottom) obtained using a quadratuer Doppler radar (a) and extracted heart rate versus time compared to respiratory signal (b) (from [3])

VI. CONCLUSION

From the review of RF based vital sign detection using CW Doppler radar, it is clearly concluded that, The concept of detecting human vital signs is successfully implemented. It can be stated as quadratuer Doppler receiver gives more accurate results than direct conversion receiver. The drawbacks in direct conversion receiver such as, DC offset can be reduced by the use of quadratuer Doppler receiver. It has the scope of future advancements in various domains for industrial approach such as, clutter cancellation techniques, antenna design with frequency and power constraints, effect on the signal w.r.t. body movement, more than one human are present etc.

References

[1] Miss Zade Gauri N., Mr. Badnerkar S.S., "A Modern Microwave Life Detection System For Human Being Buried Under Rubble", International Journal Of Advanced Engineering Research And Studies, E-Issn2249 – 8974, Ijaers/Vol. I/ Issue I/October-December, 2011/69-77.

[2] Lohman, B.; Boric-Lubecke, O.; Lubecke, V.M.; Ong, P.W.; Sondhi, M.M.; , "A Digital Signal Processor For Doppler Radar Sensing Of Vital Signs," Engineering In Medicine And Biology Magazine, Ieee , Vol.21, No.5, Pp.161-164, Sept.-Oct. 2002 Doi: 10.1109/Memb.2002.1044188

[3] Olga Boric-Lubecke, Victor M. Lubecke, Isar Mostafanezhad1, Byung-Kwon Park, Wansuree Massagram, Branka Jokanovic, "Doppler Radar Architectures And Signal Processing For Heart Rate Extraction", Mikrotalasna Revija, Decembar 2009.

[4] Ichapurapu, R.; Jain, S.; John, G.; Monday, T.; Lie, D.Y.C.; Banister, R.; Griswold, J.; , "A 2.4ghz Non-Contact Biosensor System For Continuous Vital-Signs Monitoring," Wireless And Microwave Technology Conference, 2009. Wamicon '09. Ieee 10th Annual, Vol., No., Pp.1-3, 20-21 April 2009 Doi: 10.1109/Wamicon.2009.5207236.

[5] Byung-Kwon Park, Alex Vergara, Olga Boric-Lubecke, Victor M. Lubecke, Anders Høst-Madsen, "Quadrature Demodulation With Dc Cancellation For A Doppler Radar Motion Detector",

[6] Dany Obeid1, Sawsan Sadek2, Gheorghe Zaharia1, Ghaïs El Zein1, "Feasibility Study For Non-Contact Heartbeat Detection At 2.4 Ghz And 60 Ghz", Xxix General Assembly Of The International Union Of Radio Science (Ursi), Chicago: United States (2008),

[7] Obeid, D.; Sadek, S.; Zaharia, G.; El Zein, G.;, "Touch-Less Heartbeat Detection And Measurement-Based Cardiopulmonary Modeling," Engineering In Medicine And Biology Society (Embc), 2010 Annual International Conference Of The Ieee , Vol., No., Pp.658-661, Aug. 31 2010-Sept.4 2010Doi: 10.1109/Iembs.2010.5627214

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735. Volume 4, Issue 6 (Jan. - Feb. 2013), PP 32-36
www.iosrjournals.org

Modeling and Performance analysis of Metallic CNT Interconnects for VLSI Applications

[1]Y.Amar Babu, [2]G.M.V.Prasad
[1] *L.B.R. College of Engineering Mylavaram-521230 Andhrapradesh, India*
[2]*B.V.C. Institute of Technology & Science Batlapalem-533221 Andhrapradesh, India*

Abstract: *THE SEMICONDUCTOR INDUSTRY is confronting an acute problem in the interconnect area as IC feature sizes continually scale below 32 nm. When the cross sectional dimension of copper wires approach their mean free path (about 40 nm at room temperature), they suffer significant size effects because of increasing surface scattering, grain boundary scattering, and the presence of a highly resistive diffusion barrier layer, resulting in a sharp rise in copper resistivity. According to the 2011 International Technology Roadmap for Semiconductors (http://public.itrs.net), copper's resistivity could be more than three times higher than its bulk value at the 22-nm technology node. This steep rise in resistivity will adversely impact both performance and reliability in terms of circuit delay, chip temperature, and current-carrying capacity. This limitation of copper interconnects is driving research for alternative interconnect materials and technologies for next-generation ICs. In this research, carbon nanomaterials, with their many attractive properties, are emerging as the frontrunners to potentially replace copper for interconnects and passive devices in ICs, including vias and through-silicon vias (TSVs), horizontal wires (local, intermediate, and global levels) and off-chip interconnects.*

I. Introduction

In this paper, we discuss various carbon nanomaterials, along with their prospects for next-generation interconnects and passive devices. We also provide a comparative analysis of these nanomaterials, and we illustrate why carbon nanomaterials constitute the ideal interconnect technology choice for next-generation ICs. Unique properties of carbon nanomaterials Low-dimensional allotropes of carbon, known as carbon nanomaterials, have extraordinary physical properties because of their unique structure. In particular,researchers have extensively investigated their 1D forms, carbon nanotubes (CNTs) and grapheme nanoribbons (GNRs), because of their exciting prospects in various applications, including interconnects and passive devices in the nanoelectronics area[1].

II. Atomic structure and properties

To understand the physics of CNTs and GNRs, we need to study their atomic structure. Both CNTs and GNRs can be formed from a single-layer grapheme sheet. The carbon atoms in a graphene sheet are arranged in a 2D honeycomb lattice structure.

A GNR can be formed by cutting a ribbon out of grapheme following an edge shape (either armchair or zigzag). A CNT can be formed by rolling up a ribbon along circumferential vector *C (see Figure 1a). The CNT's chirality is defined using *C (the roll-up direction), which is a combination of lattice vectors

$$\vec{C} = n \cdot \hat{a}_1 + m \cdot \hat{a}_2$$

where n and m are a pair of integers known as the chiral indices. Depending on the circumferential edge shape, a CNT can be armchair (n = m), zigzag (n = 0 or m = 0), as shown in Figure 1a, or chiral (other shapes). On the other hand, a GNR's chirality is defined by the edge shape, which is opposite to that of a CNT. Thus, starting from the same graphene strip, the CNT shown in Figure 1a is armchair, whereas the GNR is zigzag. The diameter D of the CNTs is also determined by the chiral indices:

$$D = \left|\vec{C}\right| / \pi = \frac{a}{\pi}\sqrt{n^2 + m^2 + mn}$$

where a is the lattice constant of grapheme (0.246 nm). Depending on the number of concentrically rolled up shells, CNTs can be classified as single-walled CNTs(SWCNTs), double-walled CNTs (DWCNTs), or multiwalled CNTs (MWCNTs). Similarly, GNRs can be
classified as monolayer, bilayer, or multilayer. The ideal interval distance between different shells in DWCNTs and MWCNTs, or between adjacent layers in multilayer GNRs is the Van der Waal's gap (about 0.34 nm). The unique properties of CNTs and GNRs are largely due to the unique band structure of graphene, in which the E-k (energy vs. wave vector) relation is linear for the low energies near the six corners of the 2D hexagonal Brillouin zone (as shown in Figure 1b), leading to zero effective mass for electrons and holes.1 Because of this

linear dispersion relation at low energies, electrons and holes near these six points behave like relativistic particles described by the Dirac equation for spin-1/2 particles.

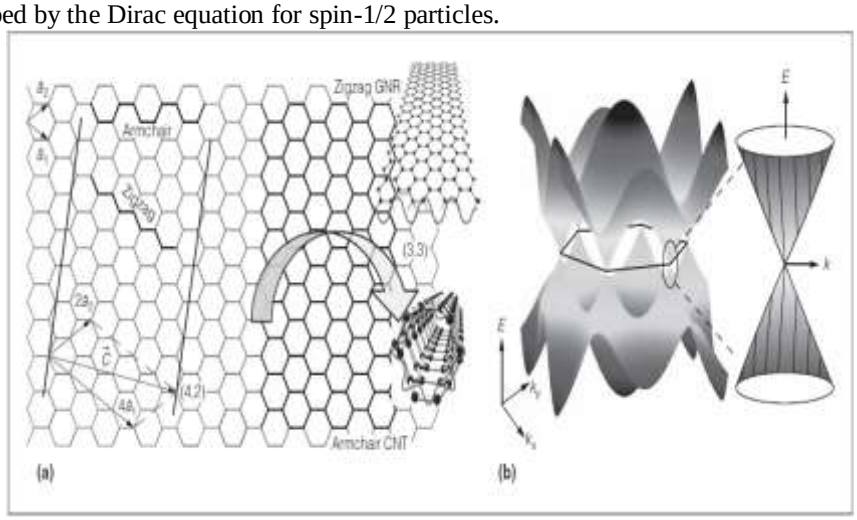


Figure 1. The atomic structure of CNT and GNR derived from a graphene sheet (a), and the band structure of graphene, where the conduction band and valence band meet at the six conical vertices (Dirac points) (b). The left side of (a) shows the chirality definition in terms of lattice vectors ^a1 and ^a2; the (4, 2) vector is shown as an example. The right side of (a) shows a zigzag-GNR and a (3, 3) armchair CNT based on the same ribbon, along with the schematic view of their wave function quantization. The zoomed-in conical shape on the right side of (b) shows a linear E-k relationship, where E is the energy and k is the wave vector. The wave function quantization leads to the formation of a set of slice cuts on the graphene band structure (the vertical lines on the right conical). If a slice cuts at one conical intersection point, there is no band gap (metallic); otherwise, there is a band gap (semiconducting). The band gap is inversely proportional to the GNR width or the CNT diameter.

Hence, the electrons and holes are called massless Dirac fermions, and the six corners of the Brillouin zone are called the Dirac points. Moreover, the sp2 bonding in these materials (where s and p are the atomic orbitals), which is stronger than the sp3 bonds in diamond, makes graphene the strongest material ever measured. CNTs and GNRs also have very high current- carrying capability (at least two orders of magnitude higher than that of copper). In addition, CNTs and GNRs have long mean free paths (MFPs) at low bias because of weak acoustic phonon scattering and suppressed optical phonon scattering at room temperature. Table 1 summarizes the key electrical, thermal, and mechanical properties of carbon nanomaterials relevant to their interconnect and passive applications.1

III. Metallicity of CNTs and GNRs

For interconnect applications, the metallicity of CNTs and GNRs is the main concern. It's important to understand their electronic structure, which again originates from graphene's band structure, as shown in Figure 1b. The band structure of CNTs and GNRs can be studied from that of graphene by considering their structural confinement.

The difference in the confinement conditions between CNTs and GNRs is that the wave function along the CNT circumference is periodic, whereas the wave function along the GNR width vanishes at the two edges (Figure 1a). These confinements can be reflected into the band structure as slice cuts (Figure 1b), each representing one subband. Depending on the position of the slice cuts due to chirality and diameter, the resulting band structure could be without band gap (metallic) or with band gap (semiconducting). Li et al. have explained that the condition for achieving a metallic CNT can be expressed as n -m =3i, where i is an integer.1 Hence, armchair CNTs are always metallic (see Figure 1a), whereas zigzag CNTs could be metallic or semiconducting, depending on the chiral indices (n, m). Statistically, a natural mix of CNTs will have one-third metallic and two-thirds semiconducting chirality. On the other hand, zigzag GNRs are always metallic (a small band gap will be induced in zigzag GNRs once electron spin is considered, because of the staggered sublattice potential from magnetic ordering1). Armchair GNRs can be either metallic or semiconducting, depending on the number (N) of atoms across the width: metallic when N ¼ 3i _ 1, and semiconducting when N ¼ 3i or 3i þ 1.

Table 1. Properties of carbon nanomaterials relevant to VLSI interconnects and passive applications.

Property	W	Cu	SWCNT	MWCNT	Graphene or GNR
Maximum current density (A/cm^2)	10^8	10^7	$>10^9$	$>10^9$	$>10^8$
Melting point (K)	3695	1357		3800 (graphite)	
Density (g/cm^3)	19.25	8.94	1.3–1.4	1.75–2.1	2.09–2.33 (graphite)
Tensile strength (GPa)	1.51	0.22	22.2 ± 2.2	11 to 63	
Thermal conductivity ($\times 10^3$ W/m·K)	0.173	0.385	1.75 to 5.8	3	3 to 5
Temperature coefficient of resistance ($\times 10^{-3}$/K)	4.5	4	<1.1	−1.37	−1.47
Mean free path (nm) at room temperature	33	40	$>10^3$	2.5×10^4	1×10^3

* CNT: carbon nanotube; Cu: copper; GNR: graphene nanoribbon; MWCNT: multiwalled CNT; SWCNT: single-walled CNT; W: tungsten.

** All the references for these values reported in this table are listed in Li et al.[1] But these values are valid only for certain CNTs or graphene cases, because they depend on many other parameters, such as CNT diameter and GNR width, or even on different fabrication processes.

IV. Modeling and performance analysis

As interest in carbon-nanomaterial-based interconnects gains momentum, a realistic analysis of these interconnects is necessary to evaluate their performance and identify the domains of on-chip interconnections (local, intermediate, and global interconnects or vias) where this novel interconnect technology has the potential to replace copper. Here, we introduce circuitmodels for varius types of CNTs and GNRs, and analyze their performance in various applications.

RLC model

Figure 2 shows the proposed RLC-equivalent circuit models for SWCNTs and GNRs. For DWCNT and MWCNT interconnects, a more complicated equivalent circuit is employed. (Li et al. presented a detailed model for MWCNT interconnects, which is also valid for DWCNT interconnects. They also provided a detailed discussion about other CNT and GNR modeling work.1) We can derive the conductance of CNTs and GNRs using the linear-response Landauer formula, 1 which takes into account the large quantum contact resistance RQ (about 12.9 kO per conducting channel) between 1D conductors and 3D materials. Although both CNTs and GNRs are considered 1D materials, a significant difference between them is the existence of edge scattering in GNRs. CNTs are seamless tubular structures, which allow electrons to travel around. GNRs, however, have abrupt edges and induce scattering because of edge roughness. Figure 3a compares the resistances of various CNTs (SWCNTs, DWCNTs, and MWCNTs) and GNRs (monolayer GNRs, neutral multilayer GNRs, and doped multilayer GNRs). For short lengths, due to large quantum contact resistance, the resistance per unit length of CNTs and GNRs are large but decrease with increasing length, and they become stable after 10 mm. For longer lengths, all types of CNTs could offer lower resistance than copper, whereas only AsF5 (arsenic pentafluoride) intercalation-doped multilayer GNRs with high specularity (or very smooth edges) can provide lower resistance than that of copper. (Intercalation involves inserting a dopant layer between adjacent graphene layers.) Besides quantum contact resistance RQ, the RLC model in Figure 2 has two other new circuit elements: quantum capacitance CQ and kinetic inductance LK. Both of these arise mainly because CNTs and GNRs are low dimensional materials with low density of states near the Fermi level. In materials with low density of states, adding charges to the system not only requires certain electrostatic energy, but also requires non negligible additional energy to occupy the higher energy states. This additional energy can be modeled as the quantum capacitance CQ (about 193 aF/mm per channel) in series with electrostatic capacitance CE.

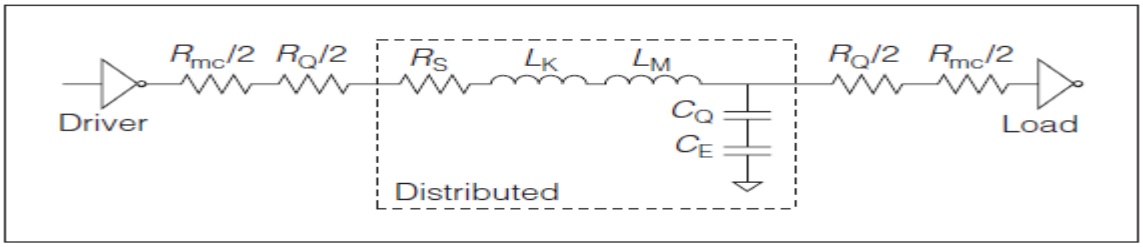


Figure 2. The equivalent distributed RLC circuit model of an SWCNT or GNR interconnect. Rmc is the imperfect contact

resistance between the CNT and the metal; RQ is the quantum contact resistance associated with the contact from the 1D conductor to the 3D conductor; RS is the scattering-induced resistance, depending on mean free path; LK is the kinetic inductance; LM is the magnetic inductance associated with the CNTs or GNRs; and CQ and CE are the quantum and electrostatic capacitances.

Similarly, to establish a current in a system, the number of charges moving in one direction must exceed the charges moving in the other direction. These excess moving charges also require excess energy to occupy higher energy states, and the excess energy associated with the moving charges can be modeled as the kinetic inductance LK (about 8 nH/mm per channel) in series with the conventional magnetic inductance LM. The value of kinetic inductance is about three orders larger than that of magnetic inductance in a single CNT or monolayer GNR. (Other experimental work on the high-frequency characterization of CNTs has verified the existence of kinetic inductance; see the work of Li et al. for details.1)

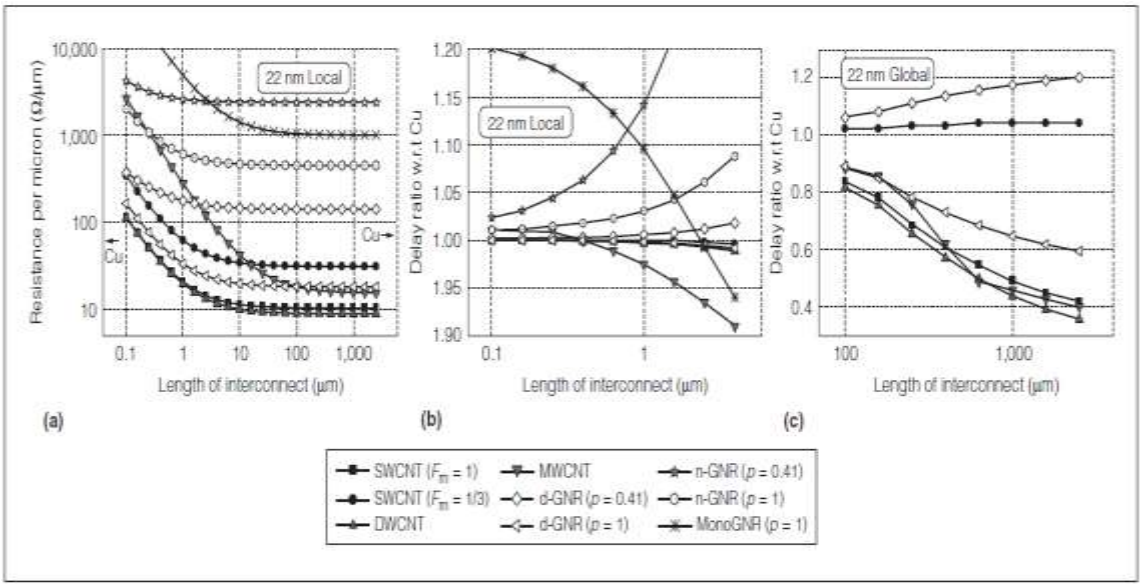


Figure 3. Resistance per unit length comparison for different types of CNT and GNR interconnects for the dimensions

predicted in the 2008 International Technology Roadmap for Semiconductors (http://public.itrs.net) for local interconnects at the 22-nm technology node (22-nm width and 44-nm height) (a). Signal delay ratios (with respect to copper) of SWCNT, DWCNT, MWCNT, and different types of GNRs at the 22-nm technology node for local (b) and global (c) interconnects. The driver size is assumed to be 2 in (b) but is set to 100 in (c). The global-interconnect width is set to be five times the minimum width, as predicted by the ITRS. Fm indicates the fraction of metallic SWCNTs in the SWCNT bundle; n-GNR and d-GNR represent neutral multilayer GNRs and AsF5 (arsenic pentafluoride) intercalationdoped GNRs, respectively; and p indicates the specularity of a GNR edge: p _ 1 indicates no edge scattering (best case), and p _ 0 implies a completely diffusive edge (worst case). For the DWCNT, the diameter is set to 1.5 nm, and metallic fraction Fm _ 1. The diameter of the MWCNT and the width of the GNR are equal to the wire width (W).

V. Conclusion:

We've compared the performance of different carbon-nanomaterial_based interconnects with that of copper interconnects on the basis of the model in Figure 2. Figures 3b and 3c show the delay ratios of CNT and GNR interconnects with respect to copper for local- and global-level interconnects at the 22-nm technology node. Metallic CNT interconnects best for global interconnect in VLSI applications.

References

[1] H. Li et al., ''Carbon Nanomaterials for Next-Generation Interconnects and Passives: Physics, Status, and Prospects,'' IEEE Trans. Electron Devices, vol. 56, no. 9, 2009, pp. 1799-1821.

[2] T. Xu et al., ''Aligned Carbon Nanotubes for Through- Wafer Interconnects,'' Applied Physics Letters, vol. 91, no. 4, 2007, article 042108.

[3] C. Xu et al., ''Compact AC Modeling and Analysis of Cu, W, and CNT Based Through Silicon Vias (TSVs) in 3-D ICs,'' Proc. IEEE Int'l Electron Device Meeting (IEDM 09), IEEE Press, 2009, pp. 521-524.

[4] J. Li et al., ''Bottom-up Approach for Carbon Nanotube Interconnects,'' Applied Physics Letters, vol. 82, no. 15, 2003, pp. 2491-2493.

[5] T. Iwai et al., ''Thermal and Source Bumps Utilizing Carbon Nanotubes for Flip-Chip High Power Amplifiers,'' Proc. IEEE Int'l Electron Device Meeting (IEDM 05), IEEE Press, 2005, pp. 257-260.

[6] K. Korda´s et al., ‘‘Chip Cooling with Integrated Carbon Nanotube Microfin Architectures,’’ Applied Physics Letters, vol. 90, no. 12, 2007, article 123105.
[7] M. Nihei et al., ‘‘Low-Resistance Multi-walled Carbon Nanotube Vias with Parallel Channel Conduction of Inner Shells,’’ Proc. IEEE Int’l Interconnect Technology Conf. (IITC 05), IEEE Press, 2005, pp. 234-236.
[8] M. Katagiri et al., ‘‘Fabrication of 70-nm-Diameter Carbon Nanotube Via Interconnects by Remote Plasma- Enhanced Chemical Vapor Deposition and Their Electrical Properties,’’ Proc. IEEE Int’l Interconnect Technology Conf. (IITC 09), IEEE Press, 2009, pp. 44-46.
[9] J. Dijon et al., ‘‘Carbon Nanotubes for Interconnects in Future Integrated Circuits: The Challenge of the Density,’’ Diamond and Related Materials, vol. 19, nos. 5-6, 2010, pp. 382-388.
[10] D.N. Futaba et al., ‘‘Shape-Engineerable and Highly Densely Packed Single-Walled Carbon Nanotubes and Their Application as Super-Capacitor Electrodes,’’ Nature Materials, vol. 5, no. 12, 2006, pp. 987-994.
[11] A.R. Harutyunyan et al., ‘‘Preferential Growth of Single- Walled Carbon Nanotubes with Metallic Conductivity,’’ Science, vol. 326, no. 5949, 2009, pp. 116-120.
[12] A. Kawabata et al., ‘‘Robustness of CNT Via Interconnect Fabricated by Low Temperature Process over a High- Density Current,’’ Proc. Int’l Interconnect Technology Conf. (IITC 08), IEEE Press, 2008, pp. 237-239.
[13] Q. Yu et al., ‘‘Graphene Segregated on Ni Surfaces and Transferred to Insulators,’’ Applied Physics Letters, vol. 93, no. 11, 2008, article 113103.
[14] X. Li et al., ‘‘Large-Area Synthesis of High-Quality and Uniform Graphene Films on Copper Foils,’’ Science, June 2009, pp. 1314-1314.
[15] D. Kondo, S. Sato, and Y. Awano, ‘‘Self-Organization of Novel Carbon Composite Structure: Graphene Multilayers Combined Perpendicularly with Aligned Carbon Nanotubes,’’ Applied Physics Express, vol. 1, no. 7, 2008, article 074003.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 37-41
www.iosrjournals.org

Skin Infection Recognition using Curvelet

Manisha[1], Nitika[2,] Mr Jaspal singh[3]
[1]M.Tech Student, [2]Assistant Professor Department of Electronics and Communication Engineering Rayat Institute of Engineering and Information Technology, Punjab Technical University,Punjab (India)

Abstract:-*This thesis identifies a novel space to address the problem of analyzing magnitude of skin infection from still images. This is based on the PCA space of the features extracted by a new multiresolution analysis tool called curvelet Transform. Curvelet Transform has better directional and edge representation abilities than widely used wavelet transform. Inspired by these attractive attributes of curve lets , we introduce the idea of decomposing images into its curvelet sub bands and applying PCA (Principal Component Analysis) on the selected subbands in order to create a representation feature set.A comparative study with wavelet – based and traditional PCA technique is also presented High accuracy rate achieved by the proposed method for two well-known database indicates the potential of this curvelet based feature extraction method.*
Keywords:- *Curvelet,Wavelet,Pca,Eczima, Image Processing*

I. Introduction

The skin infection detection has been studied extensively for than 20 year. Since the beginning of 90's the subject has became a major issue; mainly due to its importance in real world applications areas like Diagnosis of bone,joints,soft tissue[1] .Multiresolution analysis tools notably wavelet have been found quite useful for analyzing the information content of images; hence they enjoyed wide – spread popularity in areas like image processing , pattern recognition and computer version. Over the past two decades , following wavelets , other multi resolution tools like contourlets , ridelets etc were developed.'Curvelet Transformer' is a recent addition to the list of multiscale transforms. It has already been used to resolve image processing problems like image compression [2],texture classification [3],image denoising [4]; but not much work has been done to explore potential of curvelet transform to solve pattern recognition problems. In some recent work majumdar shared that curvelets can serve the basis for pattern recognition problems like charater recognition [5].The focus of the study is on the use of image processing in medical area. The presented work is about the skin infection detection in infected images. There are different kind of infection as well as different formats of medical images.

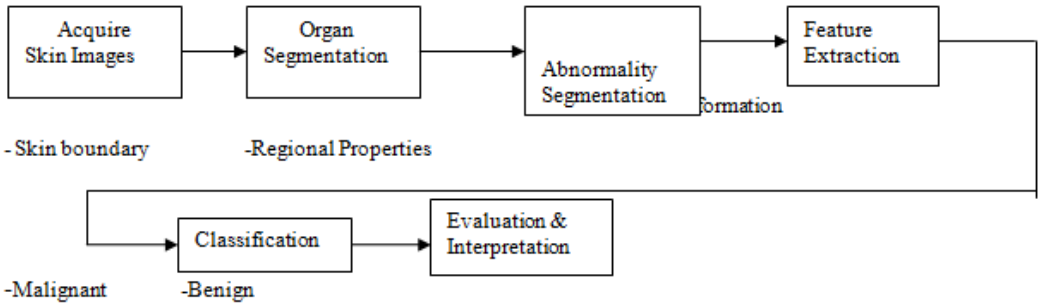


Figure 1.Basic steps of Image Processing

Digital image processing refers to the processing of digital images by means of digital computers. This process, whose inputs are images and outputs, may be images or extracted features of different attributes of images including reorganization of individual objects. Fig 1 shows, basic steps involved in the processing of digital images. Image Acquisition, Pre-processing, Image Segmentation, Wavelet, PCA,Curvelet . Due to limited scope of this paper, we are unable to delve into mathematical details of curvelet transform.

II. Related Research

In our previous work curvelet transform has been employed to extract features from bit quantized facial images and we showed that curvelets can indeed supersede the performance of wavelets.The development of Curvelet Transform by Candes and Donoho in 1999 was motivated by the need of image analysis . Curvelets present highly anisotropic behavior as it has both variable length and width. At fine scale the relationship between width and length can be expressed as width= length2; anisotropy increases with decreasing scale, in

keeping with power law. Second generation curvelet transform has two different digital implementations: curvelets via USFFT (Unequally Spaced Fast Fourier Transform) and curvelets via Wrapping. These new discrete curvelet transforms are simpler, faster and less redundant compared to their first generation version. The curvelet transform has been proposed as a multiscale geometric analysis tool, which can show the image features both at each scale and different directions. It takes no time to realize the features of the face images include curves, which form the curved singularities of the face images. Hence the use of curvelet transform for facial feature extraction is reasonable.Tanaya Mandal and Q. M. Jonathan introduce the idea of decomposing images into its curve let sub bands and applying PCA (Principal Component Analysis) on the selected sub bands in order to create a representative feature set. A comparative study with wavelet-based and traditional PCA techniques is also presented. We have introduced a new feature extraction technique from still images using PCA on curvelet domain which has been evaluated on two well-known databases.

III. Literature Review

Many researchers have done work in skin infection detection in different human organs.In previous work [6] Curvelet transform has been employed to extract Features from bit quantized facial images and we showed that curvelets can indeed suppressed the performance of wavelets.

Mrs.N.G.Chitaliya , Prof.A.I.Trivedi proposes feature extraction and classification method using Wavelet. The feature images constructed from Wavelet Coefficients are used as a feature vector for the further process. The Principal Component Analysis (PCA) is used to reduce the dimensionality of the feature vector. It also provides low dimensionality to reproduce and compare the result.The method is very fast and suitable for real time application for motion analysis. Wavelet transform is used for image compression and cleaning (noise and blur reduction). The wavelet transform of the image is computed, the wavelet representation is then modified.PCA was invented in 1901 by Karl Pearson Now it is mostly used as a tool in exploratory data analysis and for making predictive models. PCA can be done by eigen value decomposition of a data covariance matrix or singular value decomposition of a data matrix, usually after mean centering the data for each attribute. The results of a PCA are usually discussed in terms of component scores .

IV. Proposed Approach

In the present work we have improved the accuracy by using curve let based on PCA.The following steps are proposed to achieve the above mentioned objectives:

1. Study of Skin Infection Recognition System.
2. Study of file format for storing skin images.
3. Design and implementation of color detection algorithm.
4. Study and implementation of color to gray conversion, gray histogram equalization algorithm.
5. Study and implementation of laplacian edge detection algorithm, thinning & smoothing algorithm.
6. Design and implementation of Skin Infection localization using wavelet based approaches
7. Check the developed Skin Infection recognition system with different skin images.
8. Deducing conclusion at end.

The proposed work is about the detection of Skin infection in the Medical Images. The complete work is divided in four main phases.

Pre-Processing

Infection Detection

Post Processing

Analysis

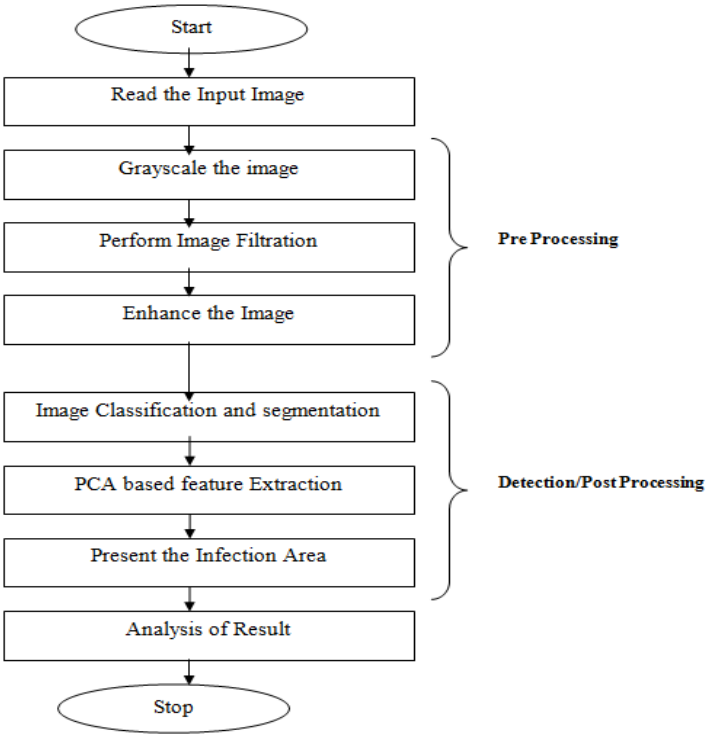
In pre-processing stage different filters are used in combination so that the resulting image is better for the use of segmentation. There are totally five filters that are used in pre-processing stage.

Next step in the pipeline is use of region of interest . The next step involved is image contrast enhancement using adaptive median filter.

The last step in the pre-processing stage is use of Sobel and Prewitt edge detection filters. To achieve better results in the segmentation process it is really important to retaining the edges. Edge detection refers to the process of identifying and locating sharp discontinuities in an image

When the input data to an algorithm is too large to be processed and it is suspected to be redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features. Transforming the input data into the set of features is called feature extraction .

FLOWCHART. 1



V. Results and Discussions

Here fig.2 is showing the Source image. The image is showing the Ezima area and the inner part of the Skin area is showing the infection area. To perform the infection detection using proposed approach. In this image fig.3, the result image after the extraction process. Here the extraction is performed using Wavelet based approach. Here fig.4 is showing the Source image. The image is showing the Ezima area and the inner part of the Skin area is showing the infection area. To perform the infection detection using proposed approach.

Result Wavelet Based

Figure.2 Input Image

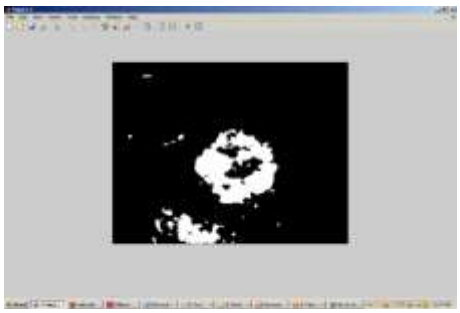


Figure.3 Eczima Extracted Image

Result curvelet Based

Figure.4 Input Image

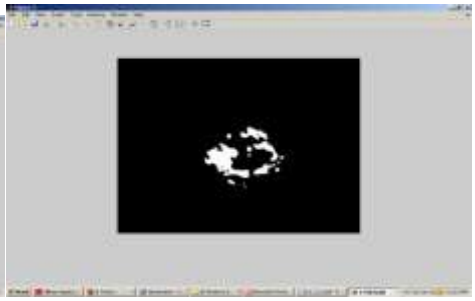
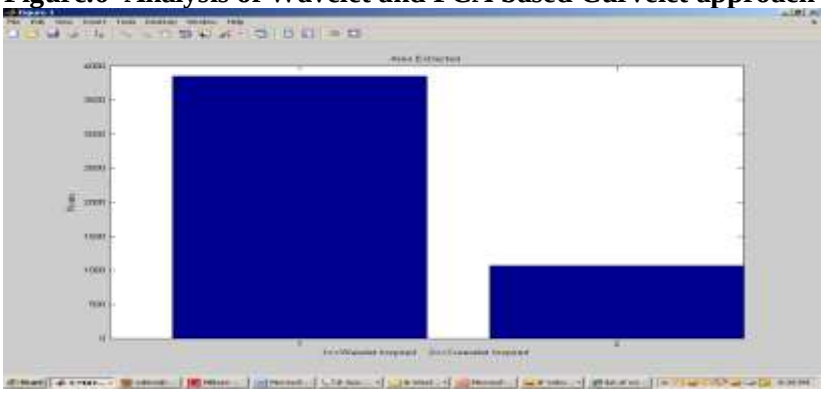


Figure.5 Eczima Extracted Image

In this image fig.5, the result image after the extraction process. Here the extraction is performed using Curvelet based approach. Here we in fig 6, Here we can see the output of the work in terms of size of area extracted from the image as the ezima part. Here we find the wavelet inspired approach has extracted about 3851 pixel count and in case of curvelet based image the extracted pixel count is 1070 pixels

Figure.6 Analysis of Wavelet and PCA based Curvelet approach



Results (Wavelet Inspired)

Here fig.7 is showing the Source image. The image is showing the Ezima area and the inner part of the Skin area is showing the infection area. To perform the infection detection using proposed approach. In this image fig.8, the result image after the extraction process. Here the extraction is performed using Wavelet based approach. Here fig.9 is showing the Source image. The image is showing the Ezima area and the inner part of the Skin area is showing the infection area. To perform the infection detection using proposed approach.

Figure.7 Input Image

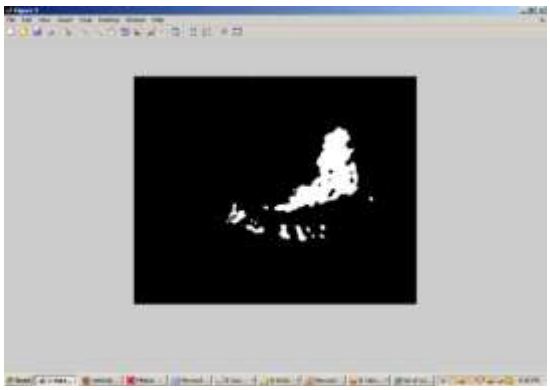


Figure.8 Eczima Extracted Image

Results (Curve let Inspired)

In this image fig.10, the result image after the extraction process. Here the extraction is performed using Cuvelet based approach.

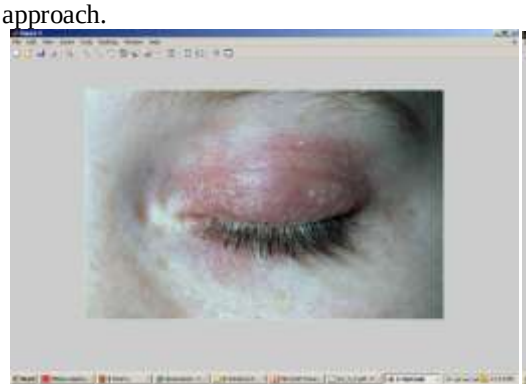
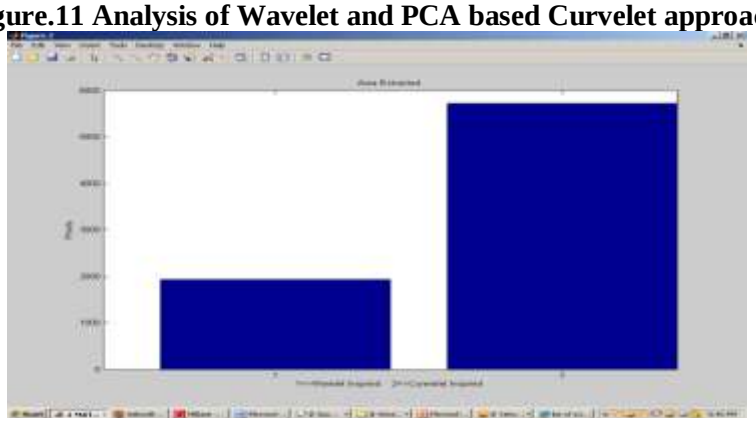


Figure.9 Input Image

Figure.10 Eczima Extracted Image

Here we can see the output of the work in terms of size of area extracted from the image as the ezima part. Here we find the wavelet inspired approach has extracted about 1924 pixel count and in case of curvelet based image the extracted pixel count is 5725 pixels. As we can see the obtained results in case of curvelet the ezima area is extracted in better way.

Figure.11 Analysis of Wavelet and PCA based Curvelet approach



VI. Conclusion

We have introduced a new feature extraction technique from still images using PCA on curvlet domain which has been evaluated on two well-known databases. Our technique has been found to be robust against extreme expression variation as it works efficiently. The subjects in this dataset make grimaces, which form edges in the eczema images and curvelet transform captures this crucial edges information. The proposed method also seems to work well for ORL databases, which shows significant variation in illumination nd infection details. The promising result indicate that curvelet transform can emerge as an effective solution to skin infection recognition problems in future. We have investigated the possibility of curvelet transformation to be used in combination with one linear analysis tool. Futher work is suggested towards the infection detection is performed on 2D images, the work can be implanted for the 3D images. We can also enhance the work by performing the implementation of sequence images

References

[1] W. Zhao, R. Chellapa , A.Rosenfeld, P. J Philips , Face Recognition.A Literature Survey , ACM Computing Survey , pp399-458,2003.

[2] M . Manikandan , A.Saravanan , K.B Bagan , Curvelet Transform Based Embedded Lossy Image Compression, Proc . of ICSCN , pp274 – 276 A , 2007.

[3] S Arivazhagan , L. Ganesan , T.G . Subhash Kumar ; Texture classification using curvelet statistical and co – occurrence features , ICPR , vol 2 , pp938 – 941 , 2003.

[4] J. L. Starch , E. J Candes , D. L. Donoho , The Curvelet Transform for Image Denoising , IEEE Trans on Image Processing , vol . 11 , Issue 6 , pp 670 – 684 , 2002.

[5] G. C Feng , P . C . Yuen , D . Q Dai , Human Face Recognition using PCA on W avelet Subband , Journal of Electronic Imaging , vol . 9 , Issue 2 , pp 226 – 233 , 2000.

[6] T. Mandal , A . Majumdar , Q . M . J . Wu Face Recognition by Curvelet Based Feature Extraction , Proc of ICIAR , Vol . 4633 , pp 806 – 817 , 2007

[7] Candes , E and Donoho , D (1999) Curvelets ; a surprisingly effective non adoptive representation for objects with edges in curves and surfaces .

[8] Tanaya Mandal and Q . M Jonathan Wu . Face Recognition using Curvelet Based on PCA . 978 – 1 – 4244 – 2175 2008.

[8] Pal, N., and Pal, S., "A Review on Image Segmentation Techniques", *Pattern Recognition*, 1993, 26, pp. 1277-1294.

[9] Haralick, R.M., and Shapiro, L.G. "SURVEY: image segmentation techniques", *Computer Vision Graphics Image Processing*, 1985, 29, pp. 100-132.

[10] Lee, J.S.L., Haralick, R., and Shapiro, L., "Morphologic Edge Detection", *Proc. 8th International Conference Pattern Recognition*, Paris (France), 1986, pp. 369-373.

[11] J. B. T. M. Roerdink and A. Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies," *Fundamenta Informaticae*, vol. 41, pp. 187–228, 2000.

[12] http:eswww.essex.ac.uk/mv/allfaces/grimace.zip.

[13] http://www.cl.cam.ac.uk/Research/DTG/attachive.

[14] J .L .Stack , " Image Processing by the curvelet Transform " , PPT.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 42-50
www.iosrjournals.org

Traffic Modelling For Capacity Analysis of CDMA Networks Using Lognormal Approximation Method

J.J. Biebuma and B.O. Omijeh
[1,2,]*Department of Electrical/Electronic Engineering University of Port Harcourt Nigeria*

Abstract: *This paper presents teletraffic modelling in cellular networks operating with Code Division Multiple Access (CDMA) to assist in capacity analysis of such systems using statistical means. Since the introduction of CDMA, owing to capacity edge over other multiple access techniques, there is the need for tools that will assist in ensuring quality of service and proper network dimensioning. This work produces a model useful for capacity analysis. This is achieved by modelling telephone traffic using lognormal assumptions to generate a CDMA blocking probability that is adapted into Erlang B formula for capacity calculations. MATLAB is used to realise the blocking probability formula so developed such that numerical and graphical results are outputted. Results show the effect of variations in network parameters on CDMA. Thus the model is applicable in network planning and dimensioning.*
Keywords: *Teletraffic, Cellular network, CDMA, Erlang, Modelling, Lognormal*

I. Introduction

Code Division Multiple Access (CDMA) is the preferred access technique for mobile communications mainly for its capacity reasons [1]. Its advantage over other multiple access schemes (like FDMA and TDMA) include higher spectral reuse efficiency, greater immunity to multipath fading, more robust handoff procedures, gradual overload capability and voice activity effects.

In CDMA however, the separation between traffic and transmission issues is not clear with capacity being determined by interference caused by all the transmitters in the network [2]. This calls for more researches to understand its traffic behaviour

Cooper in his work (3) carried out analysis of cell interference in spread spectrum. An extension of this which includes the effects of shadowing and voice activity monitoring, is found in the work of Gilhousen in 1991. In the paper [4] by Fapojuwo in 1993, a computationally intensive procedure is presented for the evaluation of the teletraffic capacity of both forward and reverse links in a CDMA cellular system. Kim in 1993, did a very similar analysis with the exception that the fixing of the PLE at 4 leads to analytic expressions for the interference from the circular cells[5]. This is extended to an analytic result for variance [6]. A standard hexagonal cellular layout is assumed with the propagation model in the paper [7] by Kohno et al in 1995, which includes lognormal shadowing taken to be independent on distinct paths .An extension of the reverse link analysis of the work by Gilhousen et al is discussed in a paper by Viterbi et al [8]. Robert, in 1996 carried out an investigation of the effect of using the actual distance of users when calculating the capacity of a CDMA network [9][10]. Simulations were carried out for twenty-seven cell CDMA network. The simulation results show that for a uniform user distribution, the difference in capacity determined using relative actual interference and relative average interference is too small to warrant the incursion of heavy computational load involved in the former case. This discovered advantage is explored in the capacity analysis in this paper .Blocking of subscribers trying to make calls occurs when the reverse link multiple access interference power reaches a predetermined level that is set to maintain acceptable signal quality .

II. Traffic Modelling Of Mobile Cellular Systems

The traffic offered to a network is a measure of the number of call initiations and the duration of calls. It is defined as a product of the mean arrival rate (calls/sec) and the mean call holding time (sec). The dimensionless unit of traffic is the Erlang.

A. Traffic Load and Trunk Size

The amount of traffic being "offered" to the switch that routes calls over the links is called the ***traffic load*** and has the measure
Load = Call rate x Call duration

$$M=A/A_0 \qquad (1)$$

B. Erlang B Statistics

The arrival of calls from different users at the switch under these conditions is assumed to be a Poisson random process, with the average rate of arriving calls λcalls/sec. The duration (holding time) of a call is assumed to have an exponential probability distribution, with an average call length of T seconds. This distribution is valid for a call departure or completion process that is Poisson, with average departure rate of 1/T = μ calls/sec. As the random call traffic arrives and departs, the number of lines occupied by ongoing calls can vary from 0 to N. The number of lines occupied is restricted to 0<k < N. Where k is the active call per time.
The state of having all lines occupied may be momentary, so it is referred to as "time congestion." If calls are rejected when all ***N*** lines are occupied, then P_k for the case of k = N is the probability that a call is rejected or "blocked":

$$B=P_N=\frac{(\lambda/\mu)^N/N!}{\sum_{k=0}^{N}(\lambda/\mu)^k/k!}=\frac{A^N/N!}{\sum_{k=0}^{N}A^k/k!},A=\frac{\lambda}{\mu} \qquad (2)$$

This expression for the blocking probability is known as the ***Erlang* B formula**. The blocking probability is also a function of the dimensionless quantity.

$$\frac{\lambda\,(call/sec)}{\mu\,(calls/sec)} = \text{A (traffic "load" in Erlangs)} \qquad (3)$$

For a finite number of users (M), the load in Erlangs can also be expressed in Terms of ρ**,** the fraction of time that each user occupies a telephone line. Thus

$$A = M\rho. \qquad (4)$$

The offered load *A* can be interpreted as the average number of lines occupied when the blocking probability is small.

III. CDMA Blocking Probability Modeling

The determination of Erlang capacity depends on the assumptions about the probability distributions of the call traffic and user interference. In this paper, CDMA Erlang capacities are determined using lognormal approximations.

A. Formulation of the Blocking Probability

First we consider a single, isolated CDMA cell with M active users. The total reverse link signal-plus-noise power received at the base station can be written as

$$\underbrace{\alpha_{r1}P_1+\alpha_{r2}P_2+\ldots\ldots+\alpha_{rM}P_M}_{M\text{ reverse link signals}}+\underbrace{(N_0W)_c}_{\text{noise power}} \qquad (5)$$

Where

- The $\{\alpha_{ri}\}$ are random variables representing the reverse link voice activity, which have the experimental values given as

  $E\{\alpha_{ri}\} = \overline{\propto_r} = 0.4$ and $E\{\propto_{ri}^2\}= \overline{\propto_r^2} = 0.31$
- The $\{P_i\}$ are the random signal powers for the M active users.
- The number of signals M is itself an RV, assumed to have a Poisson distribution, so that $E\{M\}=\bar{M} = Var\{M\}$

To a potential $(M + 1)$st reverse link user, the total power for the *M* active users and the thermal noise is interference power. Thus,
Where is the power spectral density level for the total received interference power.
Normalized by I'_0R_b, where R_b is the data bit rate, the total interference is characterized by the quantity

$$I' \equiv I'_0 W = \propto_{r1} P_1 + \propto_{r1} P_2 + \ldots + \propto_{rM} P_M + (N_0 W)_c \quad (6)$$

$$\frac{I'}{I'_0 R_b} = \frac{1}{I} \qquad I'_0 = \frac{I'}{W} = \propto_{r1} P_1 + \propto_{r1} P_2 + \ldots + \propto_{rM} P_M + N_0 \quad (7)$$

$$\frac{I'}{I'_0 R_b} = \frac{W}{R_b} = \propto_{r1} \frac{E_{b1}}{I'_0} + \propto_{r2} \frac{E_{b2}}{I'_0} + \cdots + \propto_{rM} \frac{E_{bM}}{I'_0} + \frac{N_0}{I'_0} \cdot \frac{W}{R_b} \quad (8)$$

Where $P=E_bR_b$

$$\triangleq \sum_{i=1}^{M} \propto_{r1} \rho_i = \frac{W}{R_b}(1-\eta) \quad (9)$$

$$\rho_i \triangleq E_{bi}/I'_0 \quad (10)$$

$$\eta \triangleq \frac{N_0}{I'_0} \quad \text{(thermal noise)} \quad (11)$$

Where $P=E_bR_b$ is a parameter indicating the loading of the CDMA system and W/R_b is the spread-spectrum processing gain. Given the value of η, the quality of the channel that is available to the *(M + 1)* st mobile user is characterized by the value of the random variable *Z;* if *Z* exceeds some threshold value, then the channel is effectively unavailable (blocked) to the *(M + 1)* st user. In terms of the distribution of the random variable *Z*, the probability that the *(M + 1)* st mobile CDMA user will be blocked is the probability that *Z* exceeds some threshold value Z_0, as a function of a threshold value of the interference parameter η_0.

$$B_{CDMA} = \Pr\{Z > Z_0 = \frac{W}{R_b}(1-\eta_0)\} \quad (12)$$

$$= \Pr\{\sum_{i=1}^{M} \alpha_{ri}\rho_i > \frac{W}{Rb}(1-\eta_0)\} \quad (13)$$

If a probability density function $p_z(x)$ is known or assumed for Z, then the evaluation of B_{CDMA} is simply a matter of integrating that pdf over the region defined by $Z > Z_0$:

$$B_{CDMA} = \int_{Z_0}^{\infty} dx p_z(x) \quad (14)$$

The exact *pdf* of *Z* is not known, however, so an approximation is needed to compute B_{CDMA}.
The CDMA blocking probability can be manipulated to the form

$$B_{CDMA} = \Pr\{Z > Z_0\} = \Pr\left\{\frac{Z - E\{Z\}}{\sqrt{Var\{Z\}}} > \frac{Z_0 - E\{Z\}}{\sqrt{Var\{Z\}}}\right\} \quad (15)$$

$$= Qz\left(\frac{Z_0 - E\{Z\}}{\sqrt{Var\{Z\}}}\right) \quad (16)$$

Where Qz(·) is notation for the complementary cumulative distribution function of the standardized (zero-mean, unit variance) version of the RV Z.

The approximation method to be considered in what follows is Lognormal approximation: Based on the fact the SNRs in the sum are lognormal, thus, *Z* itself can be approximately characterized as a lognormal variable. The approximation method is based on identifying the actual mean and variance of *Z* with the mean and variance of a lognormal RV. Next, we find the mean and variance of *Z* and specify its relations to lognormal RV.

B. Mean and Variance of *Z*

The form of the interference statistics *Z* is the weighted sum of the *M RVs* {ρi, i = 1, 2, …, *M*}:

$$Z = \propto_{r1} \frac{E_{b1}}{I_0'} + \propto_{r2} \frac{E_{b2}}{I_0'} + \ldots + \alpha_{rM} \frac{E_{bM}}{I_0'}$$

$$= \propto_{r1} \rho_1 + \propto_{r2} \rho_2 + \cdots + \propto_{rM} \rho_M \qquad (17)$$

The {ρi} above, when measured in dB units, are close to having a Gaussian probability distribution with median m_{dB} and standard deviation σ_{dB}:

$$\rho_i \text{ (dB)} = 10 \log_{10} \rho i = m_{dB} + \sigma_{dB} G_i, \qquad G_i = G(0, 1) \qquad (18)$$

Therefore, the RV ρ_i is lognormal and can be written

$$\rho_i = 10^{(m_{dB} + \rho_{dB} Gi)/10} = (e^{\ln 10})^{(m_{dB} + \rho_{dB} Gi)/10}$$

$$= e^{\beta(mdB + \sigma dB Gi)}, \qquad \text{using } \beta = (In\ 10)/10 \qquad (19)$$

The median, mean, and mean square of ρ_i are assumed to be same for all i and are obtained as follows:

- Median

$$Pr\left\{ G \leq \frac{\frac{1}{\beta} ln\, \rho_{med} - m_{dB}}{\sigma_{dB}} = G_{med} = 0 \right\} \qquad (20$$

$$\frac{1}{2} \triangleq \Pr\{\rho i \leq \rho med\} = \Pr\left\{ e^{\beta (mdB + \sigma dBG)} \leq \rho med \right\}$$

Thus $$\rho_{med} = e^{\beta m_{dB}} \qquad (21)$$

- Mean

$$E\{\rho i\} = E\left\{ e^{\beta (mdB + \sigma dBG)} \right\} = e^{\beta mdB} E\left\{ e^{\beta \sigma dBG} \right\}$$

$$= \rho med\, E\{e^{uG}\} | u = \beta \sigma dB \qquad (22)$$

Where

$$E\{e^{uG}\} = M_G(u) = e^{u^2/2} \qquad (MGF)$$

Thus

$$E\{\rho i\} = \rho_{med}\, M_G\, (\beta \sigma_{dB}) = \rho_{med} e^{\frac{(\beta \sigma dB)^2}{2}} \qquad (23)$$

Means square

The mean, mean square and variance of *Z* therefore are

$$E\{Z\} = E_M\{E\{Z|M\}\} = E_M\{\sum_{i=1}^{m} E\{\alpha_{ri}\rho_i\}\}$$

$$= E_M\{M\,E\{\alpha_{ri}\rho_i\}\} = E\{M\}\,E\{\alpha_{ri}\rho_i\}$$

$$= \bar{M}\ \underbrace{\overline{\alpha_r}\,e^{\beta dmB + \frac{1}{2}\beta^2\sigma^2 dB}}_{I\neq j \text{ terms}} \qquad (25)$$

I=j term

$$E\{Z^2\} = E_M\{E\{Z^2 \mid M\}\} = E_M\{\sum_{i=1}^{M}\sum_{j=1}^{M} E\{\alpha_{ri}\alpha_{rj}\rho_i\rho_j\}\} = E_M\{ME\{\alpha_{ri}^2\rho_i^2\} + M(M-1)[E\{\alpha_{ri}\rho_i\}]^2\}$$

$$= E_M\{M\{E\{\alpha_{ri}^2\rho_i^2\} + [E\{\alpha_{ri}\rho_i\}]^2\}\,\overline{M^2}\,E\{\alpha_{ri}\rho_i\}]^2\}$$

$$= \bar{M}Var\{\alpha_{ri}\rho_i\} + \overline{M^2}\,[E\{\alpha_{ri}\rho_i\}]^2 \qquad (26)$$

$$Var\{Z\} = E\{Z^2\} - [E\{Z\}]^2$$

$$= \bar{M}\,Var\{\alpha_{ri}\rho_i\} + \left(\overline{M^2} - \bar{M}^2\right)[E\{\alpha_{ri}\rho_i\}]^2$$

$$= \bar{M}\,Var\{\alpha_{ri}\rho_i\} + Var\{M\}\,[E\{\alpha_{ri}\rho_i\}]^2$$

$$= \bar{M}\,Var\{\alpha_{ri}\rho_i\} + [E\{\alpha_{ri}\rho_i\}]^2\}, \quad because\ Var\{M\} = \bar{M}$$

$$= \bar{M}\,E\{\alpha_{ri}^2\rho_i^2\} = \bar{M}\,\overline{\alpha_r^2}\,e^{2BmdB + 2\beta^2\sigma_{dB}^2} \qquad (27)$$

The interference due to mobiles in other cells can be accounted for by using first- and second-order *frequency reuse factors* $F = 1 + \xi$ and $F^1 = 1 + \xi'$, respectively, where

$$\xi = \frac{Total\ other\ cell\ received\ (median) power}{Total\ same-cell\ received\ (median) power}$$

and

$$\xi' = \frac{Total\ other\ cell\ mean\ square\ received\ power}{Total\ other\ cell\ mean\ square\ received\ power} \qquad (28)$$

A typical analytical value of $\xi' = 0.086$ and $\xi = \xi' = 0.55$. With this method of accounting for interference from other cells, the mean and variance for *Z* become

And

$$E\{Z\} = \bar{M}\,\overline{\alpha r}\,\rho_{med}\,e^{(\frac{1}{2})\beta^2\sigma_{dB}^2}\,.(1+\xi) \qquad (29)$$

$$Var\{Z\} = \bar{M}\,\overline{\alpha_r^2}\,\rho_{med}^2\,e^{2\beta^2\sigma_{dB}^2}\,.(1+\xi') \qquad (30)$$

Approximations for The Probability Distribution of *Z*

Because the *M* RVs $\{\rho_i, i=1, 2\ldots M\}$ are lognormal RV, the interference statistic Z is the weighted sum of lognormal RVs. One approximation for the distribution of *Z i*s based on assuming that the summing of variables to produce *Z* causes its distribution to converge to a Gaussian distribution according to the CLT. Another approach is to assume that the lognormal character of the $\{\rho_i\}$ makes *Z* have an approximately lognormal distribution.

Thus:

$$Z = \propto_{r1} \rho_1 + \propto_{r2} \rho_2 + \cdots + \propto_{rM} \rho_M$$
$$= \propto_{r1} e^{\beta(mdB + \sigma dB G_1)} + \propto_{r2} e^{\beta(mdB + \sigma dB G_2) + \cdots} + \propto_{rM} e^{\beta(mdB + \sigma dB G_M)}$$
$$\approx e^{m_M + \sigma_M G} \quad \text{Lognormal approximation} \qquad (31)$$

CDMA Blocking Probability Formula for Lognormal Assumptions

Under lognormal assumption, the mean and variance of *Z* are identified as the mean and variance of lognormal RV, ζ, where

$$\zeta = e^{m_M + \sigma m_G} \qquad (32)$$

The mean, mean square, and variance of ζ are given by

$$E\{\zeta\} = e^{m_M} E\{e^{\sigma_M G}\} = e^{m_M + \frac{1}{2}\sigma^2_M} \qquad (33)$$

$$E\{\zeta^2\} = e^{2m_M} E\{e^{2\sigma_M G}\} = e^{2m_M + 2\sigma^2_M}$$

$$\text{Var}\{\zeta\} = E\{\zeta^2\} - [E\{\zeta\}]^2 = e^{2m_M + \sigma^2_M}\left[e^{\sigma^2_M} - 1\right] \qquad (34)$$

By setting $E\{Z\} = E\{\zeta\}$ and $\text{Var}\{Z\} = \text{Var}\{\zeta\}$, where the mean and variance of *Z* are already given. solving for m_M and σ^2_M:

$$\bar{M}\,\overline{\alpha_r}\, e^{\beta m_{dB} + \frac{1}{2}\beta^2 \sigma^2_{dB}}(1+\xi) = e^{m_M + \frac{1}{2}\sigma^2_M} \quad \text{and}$$

$$\bar{M}\,\overline{\alpha_r^2}\, e^{2\beta m_{bB} + 2\beta^2 \sigma^2_{dB}}(1 + \xi^1) = e^{2m_M + \sigma^2_M}\left[e^{\sigma^{2m}} - 1\right] \qquad (35)$$

The solution is

$$\sigma^2_M = ln\left[\frac{\overline{\alpha_r^2}\,(1+\xi')e^{\beta^2 a^2_{dB}}}{\bar{M}\,(\alpha_r)^2(1+\xi)^2} + 1\right] \qquad (36)$$

$$m_M = In\,[\,\bar{M}\,\overline{\alpha_r}\,(1+\xi)] + \beta m_{dB} + 1/2\,(\beta^2\sigma^2_{dB} - \sigma^2_M) \qquad (37)$$

Using these parameters, the blocking probability formula for the lognormal approximation is

$$B_{CDMA} = \Pr\{Z > Z_O\} \approx \Pr\{e^{m+\sigma_M G} > Z_0\} = Q\left(\frac{In Z_0 - m_M}{\sigma_M}\right) \qquad (38)$$

Substituting the expressions for m_M and σ_M, we obtain general expressions for the CDMA blocking probability under the lognormal approximation for the interference statistic, given by

$$B_{CDMA} = Q\left(\frac{ln\left[\frac{W}{R_b}(1-\eta_0)\right] - ln[\bar{M}\,\overline{\alpha_r}\,(1+\xi)] - \beta m_{dB}}{\sqrt{ln\left[\frac{\overline{\alpha_r^2}(1+\xi')e^{\beta^2\sigma^2_{dB}}}{\bar{M}(\overline{\alpha})^2(1+\xi)^2} + 1\right]}}\right)$$

$$- \left. \frac{\frac{1}{2}\left\{\beta^2\sigma_{dB}^2 - ln\left[\frac{\overline{\alpha_r^2}(1+\xi')e^{\beta^2\sigma_{dB}^2}}{\overline{M(\alpha)^2}(1+\xi)^2}+1\right]\right\}}{\sqrt{ln\left[\frac{\overline{\alpha_r^2}(1+\xi')e^{\beta^2\sigma_{dB}^2}}{\overline{M(\alpha)^2}(1+\xi)^2}+1\right]}} \right) \quad (39)$$

in which the Erlang capacity is $\overline{M}$. Because the interference parameter η is η = 1 – X, we may convert the threshold η_0 into a loading threshold $X_0 = 1 - \eta_0$ and write (3.72) as

$$B_{CDMA}=Q\left(\frac{ln\left[\frac{W}{R_b}X_0\right] - ln[\overline{M}\,\overline{\alpha_r}\,(1+\xi)] - \beta m_{dB}}{\sqrt{ln\left[\frac{\overline{\alpha_r^2}(1+\xi')e^{\beta^2\sigma_{dB}^2}}{\overline{M(\alpha)^2}(1+\xi)^2}+1\right]}} - \frac{\frac{1}{2}\left\{\beta^2\sigma_{dB}^2 - ln\left[\frac{\overline{\alpha_r^2}(1+\xi')e^{\beta^2\sigma_{dB}^2}}{\overline{M(\alpha)^2}(1+\xi)^2}+1\right]\right\}}{\sqrt{ln\left[\frac{\overline{\alpha_r^2}(1+\xi')e^{\beta^2\sigma_{dB}^2}}{\overline{M(\alpha)^2}(1+\xi)^2}+1\right]}} \right) \quad (40)$$

IV. Results and Discussion

A. Graphical Plots and Analysis of CDMA Blocking Probability Formula for Lognormal Assumptions

The blocking probability formula derived with Lognormal Assumptions in the previous section is as follows:

$$B_{CDMA} = Q\left(\frac{ln\left[\frac{W}{R_b}X_0\right] - ln[\overline{M}\,\overline{\alpha_r}\,(1+\xi)] - \beta m_{dB}}{\sqrt{ln\left[\frac{\overline{\alpha_r^2}(1+\xi')e^{\beta^2\sigma_{dB}^2}}{\overline{M(\alpha)^2}(1+\xi)^2}+1\right]}} - \frac{\frac{1}{2}\left\{\beta^2\sigma_{dB}^2 - ln\left[\frac{\overline{\alpha_r^2}(1+\xi')e^{\beta^2\sigma_{dB}^2}}{\overline{M(\alpha)^2}(1+\xi)^2}+1\right]\right\}}{\sqrt{ln\left[\frac{\overline{\alpha_r^2}(1+\xi')e^{\beta^2\sigma_{dB}^2}}{\overline{M(\alpha)^2}(1+\xi)^2}+1\right]}} \right) \quad (41)$$

Plots of equation (41) for B_{CDMA} versus $\overline{M}$ are shown in Figure 1 for a single cell ($\xi = \xi' = 0$) and for multiple cells, using $\xi = \xi' = 0.55$ and the parameter values σ_{dB} = 2.5dB, W = 1.2288 MHz, R_b = 9.6 kbps, X_0 = 0.9, $\overline{\alpha_r}$ = 0.4, and $\overline{\alpha_r^2}$ = 0.31. The plots are parametric in $m_{dB} = E_b / N_0$, which takes the values 5, 6, and 7 dB.

When $B_{CDMA} = 10^{-2} = 0.01 = 1\%$, figure 1 shows for multiple cells that the corresponding (integer) value of the Erlang capacity $\overline{M}$ is 16 for E_b/N_0 = 7dB, 22 for E_b/N_0 = 6dB, and 29 for E_b/N_0 = 5 dB.

$$\overline{\alpha_r} = 0.4, \bar{\alpha} = 0.31, X_0 = 0.9, W/R_b = 128$$

In Figure 2, the cell loading threshold X_o is varied for the lognormal approximation and the case of E_b/N_0 = 6dB. Again we see that raising or lowering X_0 has a significant effect on the Erlang capacity $\overline{M}$ for a given value of the blocking probability. The amount of increase or decrease in $\overline{M}$ is greater than the amount of increase or decrease in X_o. For example, raising X_o by 20% from 0.75 to 0.9 increases the value of $\overline{M}$ at B_{CDMA}= 1% from about 16 to 22, or about 35%, indicating the high sensitivity of the Erlang capacity to the cell loading. Therefore, the threshold value of cell loading should be chosen very carefully. The sensitivity of the CDMA blocking probability to the value of the second-order reuse fraction ξ' under the lognormal approximation is shown in Figure 3, in which the first order reuse fraction ξ = 0.55 and E_b/N_0 = 5,6, and 7dB are used

$$\overline{\alpha_r} = 0.4, \overline{\alpha_r^2} = 0.31, E_b/N_0 = 6dB, W/R_b = 128$$

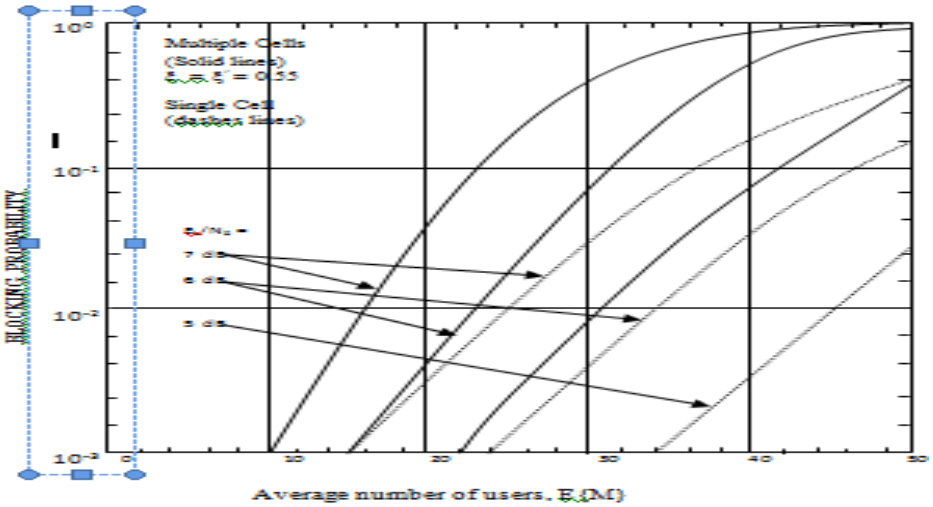


Figure 1: CDMA blocking probability (lognormal approximation) versus average number of mobile users, SNR requirement varied.

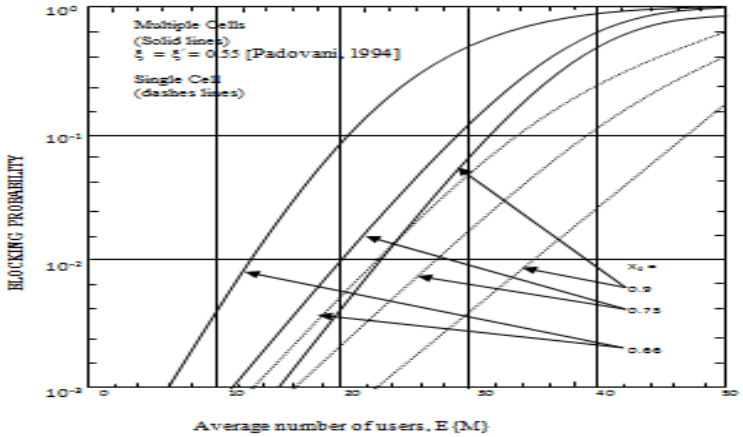


Figure 2: CDMA blocking probability (lognormal approximation) versus average number of mobile users, loading threshold varied

$$\overline{\alpha_r} = 0.4, \overline{\alpha_r^2} = 0.31, X_0 = 0.9dB, \quad W/R_b = 128$$

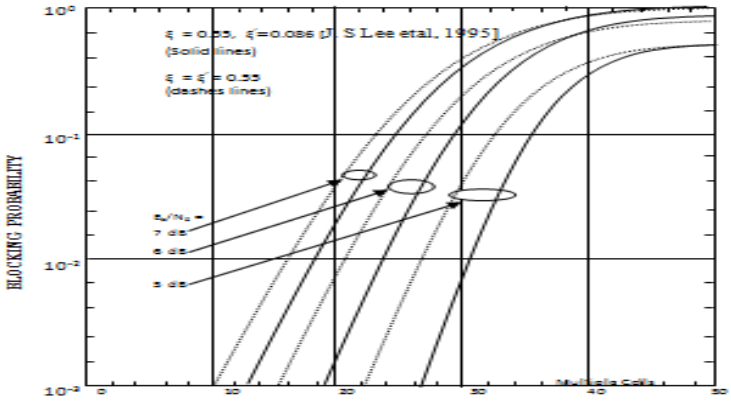


Figure 3: CDMA blocking probability (lognormal approximation) versus average number of mobile users, reuse fraction varied

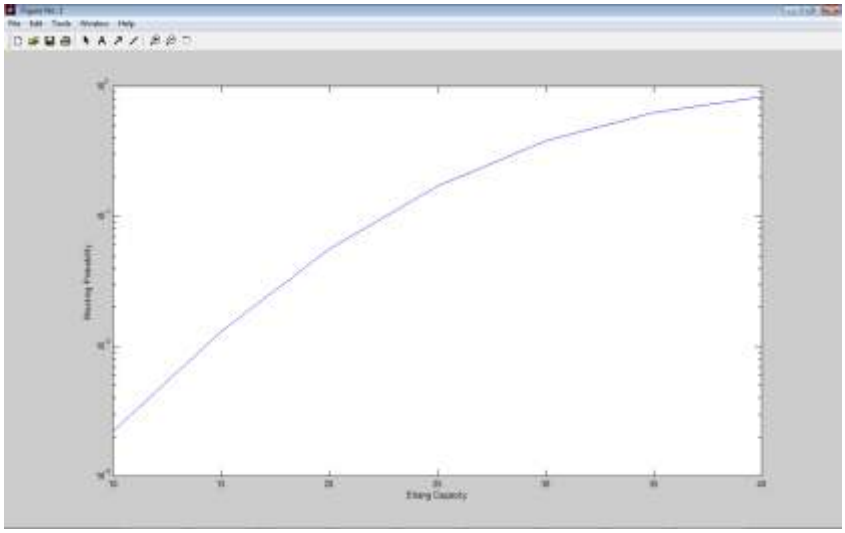


Figure 4: Erlang capacities calculated and results graphically displayed in MATLAB

V. Conclusion

We have derived Erlang capacity formulas for the CDMA cellular system under lognormal approximation, on the assumption that the sum of *M* lognormal Random Variables is also a lognormal RV. The stochastic nature of call arrivals and departures were characterized using statistical means. The interference contributed by each user was modeled as a Poison RV that summed up to statistical RV lognormal characteristics. The blocking probability formula so derived as programmed in MATLAB and Erlang capacities calculated and results graphically displayed. See Figure 4.

Blocking occurred when the reverse link multiple access interference power reached a predetermined level that is set to maintain acceptable signal quality. When the total user interference at a base station receiver exceeded the set threshold, the system blocked the next user attempting to place a call. The number of users for which the CDMA blocking probability equaled 1% as chosen was taken to be the *Erlang* capacity of the network. Thus, a new CDMA blocking probability model is developed that enabled the estimation and analysis of Erlang capacity of CDMA networks.

Graphic results for the blocking model generated showed the effect of variations in interference parameters on CDMA capacity. The Erlang capacity from the model is adaptable into Erlang B formula to estimate capacity in terms of channels.

References

[1] Lee J.S and Miller L.E (1995), "On the Erlang Capacity of CDMA cellular systems," In proc. GLOBECOM'95, Singapore, pp.1877-1883, Nov. 1995.

[2] Evans J.S and Everrit D(1995), "On the teletraffic capacity of CDMA Cellular networks" Department of Computer Engineering and Computer Science, University of California Berkelay, CA, 94720, USA, 1995.

[3] Cooper G.R and Nettleton R.W (1978), "A spread spectrum technique for high capacity mobile communications,'' IEEE Trans. Veh. Technol., VT-27, pp. 264-275, Nov. 1978.

[4] Fapojuwo A.O (1994), "Radio capacity of Direct sequence code division multiple acess mobile radio systems," in proc. IEEE Vehicular Technology Conference, Stockholm, Sweden, pp. 868-872, June 1994.

[5] Kim K.I(1993)"CDMA cellular engineering issues (1993)," IEEE Trans Veh. Technol. Vol. 42, pp. 345-350,Aug.1993.

[6] Lee J.S and Miller L.E(1998), "CDMA systems Engineering Handbook" Library of Congress Cataloging-in-Publication, Data,1998.

[7] Kohno R, Meidan R, and Milstein L.B(1995), "Spread spectrum access methods for wireless communications,'' IEEE Comm. Magazine, vol. 33, pp 58-67, Jan.1995

[8] Viterbi A.J , Viterbi A.M, and Zehavi E(1992), "Soft handoff extends CDMA cell coverage and increases reverse link capacity, " IEEE J. Select. Areas Commun., vol. 41, pp.231-242, Aug. 1992.

[9] Robert A.K.L and Parvez A (1992), "Impact of interference model on Capacity in CDMA cellular Networks," Department of Computer Science and Engineering, University of North Texas, Denton, TX, 76203.

[10] Jhong S. L. Etal (1993), "CDMA systems Engineering Handbook" Artech House mobile communications Library, 1998.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 51-56
www.iosrjournals.org

Public Transportation Management Services by Integrating GSM-GPS

Raghava Bhonagiri[1], Prof.K.Subba Rao[2], Vinod Kumar Varkala[3], Gopikrishna Siliveri[4]

[1,2]*(Electronics& Communication Engineering, DVR College of Engineering and Technology, Hyderabad, India)*
[3]*(Electronics& Communication Engineering, TRR College of Engineering and Technology, Hyderabad, India)*
[4]*(R&D Embedded Systems, CMC Limited, Hyderabad, India)*

Abstract: *This paper proposes and implements a solution for enhancing public transportation management services based on GPS and GSM in Andhra Pradesh. The system consists of four modules: BUS Station Module, In-BUS Module, BASE Station Module and BUS Stop Module. Equipped with PC and GSM modem, BUS Station Module sends the initialization information containing the bus number and license plate number to In-BUS Module and BASE Station Module using SMS. The microcontroller based In-BUS Module consisting mainly of a GPS receiver and GSM modem then starts transmitting its location and number of passengers to BASE Station Module. BASE Station Module equipped with a microcontroller unit and GSM modems interfaced to PCs is designed to keep track record of every bus, processes user request about a particular bus location out of BUS Station and updates buses location on bus stops. BUS Stop Module is installed at every bus stop and consists of a GSM modem, memory unit and dot matrix display all interfaced to a microcontroller. This module receives buses location information coming towards that stop from BASE Station module and displays the information on a dot matrix display. The results have shown that the developed system is useful for facilitating people using public transportation services.*

Keywords : *GPS;GSM;public transportation management services; Bus Station Module; In-BUS Module; BASE Station Module;BUS Stop Module;rush statistical analysis.*

I. INTRODUCTION

In this paper, a transportation management system is developed for enhancing public transportation services based on integration of GPS and GSM. GPS is used as a positioning device while GSM is used as communication link between different modules. These modules include BUS Station Module, In-Bus Module, BASE Station Module and BUS Stop Module. Bus Station Module contains a GSM engine interfaced to PC and transmits the bus index and its license plate number to BASE Station. At the same time, it turns on GPS receiver installed in the bus. The bus then starts transmitting its location to the BASE Station. The BASE Station comprises of a GSM engine interfaced to a microcontroller for processing user request of bus location as well as a number of other GSM engines interfaced to various PCs each reserved for a separate bus to update the location information of that bus. The buses location data from BASE Station is sent to each bus stop. BUS Stop Module after receiving buses location data through GSM engine displays it on dot matrix display installed at each bus stop. The block diagram of the proposed system is shown in Fig. 1.

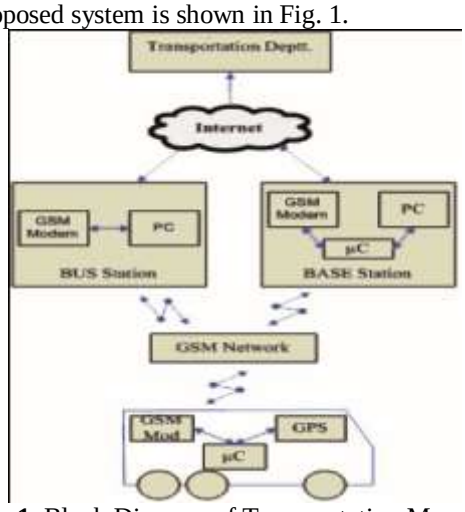


Figure 1. Block Diagram of Transportation Management System

II. HARDWARE SPECIFICATIONS

The following hardware components are used in building the entire system:

A. GPS Receiver

In order to keep track record of bus, a Garmin GPS35 receiver, powered from the bus main battery, is installed in each bus. The Garmin GPS35 is a complete GPS receiver and embedded antenna designed for a broad spectrum of OEM system applications. The GPS35 tracks up to twelve satellites at a time while providing one-second navigation updates and low power consumption. Its far-reaching capability meets the sensitivity requirements of land navigation as well as the dynamics requirements of high-performance aircraft. Internal memory backup allows the GPS35 to retain critical data such as satellite orbital parameters, last position, date, and time.

B. GSM Modem

A wireless link between the modules is provided with Nokia 12i GSM module. Nokia 12i offers advance GSM connectivity and supports EDGE/GPRS and HSCSD with automated GSM connection establishment It is equipped to provide reliable remote connections and offers application level watchdogs, inbuilt self check mechanisms and a reliable Virtual Machine (VM) for JAVA. Nokia 12i also supports reliable inbuilt internet protocols: TCP/IP for reliable data transfer, UDP/IP for audio and video streaming and HTTP for accessing web pages. The module can also be connected to an external GPS device that supports National Marine Electronics Association (NMEA) standard. The inbuilt NMEA parser can parse the location data from the output that it receives from the GPS device. External microcontroller can use AT commands to communicate with Nokia 12i and simple remote I/O applications can easily be controlled via text messages.

C. Microcontroller

AT89C52 microcontroller is selected because it is a powerful microcomputer which has low power consumption and provides a highly flexible and cost-effective solution to many embedded control applications. It has 8K bytes of in system reprogrammable flash memory, 256 bytes of internal RAM, 32 programmable I/O lines, three 16 bit timers/counters, eight interrupt sources and a programmable serial channel.

D. Memory

256K Nonvolatile RAM (NV-Ram) DS1230Y-85 is used for storing data in In-BUS Module (in case of sparse GSM coverage) and at BUS Stop Module for displaying on dot matrix display. NV-RAM is selected because it combines the best of RAM and ROM: the read and write ability of RAM and non-volatility of ROM. The DS1230 Nonvolatile SRAM is 262,144-bit, fully static, nonvolatile SRAM organized as 32,768 words by 8 bits. Each NV SRAM has a self-contained lithium energy source and control circuitry which constantly monitors VCC for an out-of-tolerance condition. When such a condition occurs, the lithium energy sources is automatically switched on and write protection is unconditionally enabled to prevent data corruption.

E. Battery Backup

In-Bus Module is provided with an internal battery so that whenever power from main battery is disconnected, microcontroller continues to transmit the location to BASE station. A message is also sent to BASE station to notify it about the disconnection of main battery. When the power is resumed, the internal battery begins to recharge.

F. Alarms

The microcontroller unit in In-BUS Module sends different alarm signals for different events to BASE Station Module.

1) On Backup Battery: When the main battery is switched off, a notification is sent to BASE station.

2) Stoppage: When the bus is stationary for more than a specified time, BASE station is informed by a stoppage alarm. In case of an accident or any other fault occurred in bus, the driver can notify the BASE station by pressing a button in bus.

3) Getting Late: When the bus is not covering a certain distance in a defined range of time, an alarm signal of getting late is sent to BASE station.

4) Route Deviation: When the bus deviates from the assigned route by a given margin, BASE station is notified.

III. SYSTEM MODULES AND NETWORK OPERATION

The entire system/network comprises of four modules: BUS Station Module, In-BUS Module, BASE Station Module and BUS Stop Module. The working and interconnection of these modules is described in this section.

A. BUS Station Module

BUS Station Module is installed at bus terminals from where the bus will depart. It contains a LASER and a GSM modem connected to a PC. When the bus enters the terminal pad, it is detected by the LASER sensor. The operator at the terminal enters the license plate number in database. A count number is then accordingly assigned to the bus e.g., bus leaving the terminal first will be assigned a number 1. The route number of bus along with the direction information, assigned count number and license plate number is sent to the BASE Station via GSM. An example of the transmitted header is of the form "33U01LZR7240" where '33' is the bus route number. Transportation Department, 'U' is upward direction of bus ('D' will be downward direction), '01' is the count number assigned to the bus and 'LZR7240' is license plate number of bus. An 'ON' signal is also transmitted to the In-BUS Module installed in the bus for initialization. The flow chart of module software is shown in Fig. 2.

B. In-BUS Module

In-BUS Module is installed inside every bus and consists of a GPS receiver, a GSM modem, a NV-RAM, infrared object counting sensors, door opening/closing sensors and an emergency button; all interfaced to AT89C52 microcontroller. After receiving the initialization signal form BUS Station Module, this module starts transmitting bus location to the BASE Station. At each stop, when the driver opens the door, an interrupt is generated and microcontroller starts counting the numbers of passengers entering and leaving the bus with the help of infrared sensors. This count value on per stop basis is transmitted to the BASE Station. In case of an emergency situation (e.g., when fault occurs in bus), driver can press the emergency button to inform BUS and BASE Station units about the location of bus. The BUS station operator can then adjust the schedule accordingly and send an additional bus for facilitating the passengers. Microcontroller present in this module continuously calculates the difference in consecutive GPS locations. If the difference remains near zero for more than a designated time, then a getting late message is transmitted to the BUS and BASE stations. In case of sparse GSM coverage, location information is stored in non-volatile RAM. After regaining the GSM network, previous locations are updated to the BASE station. The block diagram for this module is shown in Fig. 3 while flow chart of module software is shown in Fig. 4.

C. BASE Station Module

This module is the central part of the network. It accepts location information of buses through respective GSM modems and maps the information on Google Map for visualization. It also receives the number of passengers entering and leaving the bus on per stop basis from In-BUS Module for statistical analysis. The message received is of the form "20, 10, 2345.3522N, 09022.0288E". The first two strings denotes the number of passengers entering and leaving the bus respectively and next two strings denotes the location information; all separated by commas. Another GSM modem is used to get the user request of location information of a particular bus. An example of the query put by the passenger is of the form "33U" or "33U10". In first instance i.e., "33U", '33' is the route number and 'U' designates the direction flag while in second instance i.e., "33U10", additional digit '10' denotes the bus stop number where the passenger is standing. The microcontroller attached with this GSM modem passes on the user request to the PC dedicated for that route number. The PC after processing the request data sends desired location information in form of bus stop name to microcontroller. The microcontroller then transmits this information back to the user. The information that passenger will receive contains the location of all buses out of terminal in desired direction in former query while in case of later query, he will get the location of those buses which are coming towards the particular bus stop number in desired direction along with time information. The time information is embedded in message to account for any delay in processing the user request. An example of the information received by the user is of the form "N.G College, Clock Tower, V.T Colony, Srinagar colony- 12:30 P.M." where first four strings are bus stops names telling where the buses are currently followed by the time on which the location information is get from the map and message is sent to user. BASE station also monitors the emergency situations transmitted from In-BUS Module. In addition to this, the station keeps record of security issues and traffic congestion conditions and directs the driver to change the route if desired. The block diagram of the module is shown in Fig. 5 while module software is shown in Fig. 6.

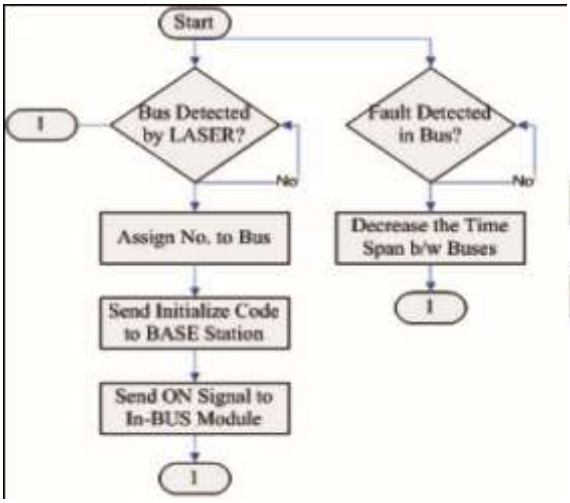


Figure 2. Flow Chart of BUS Station Module

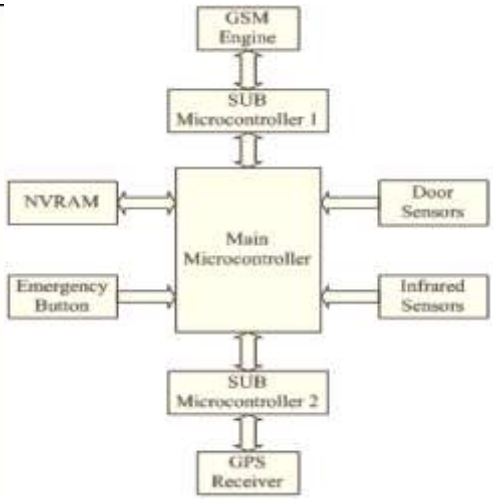


Figure 3. Block Diagram of In-BUS Module

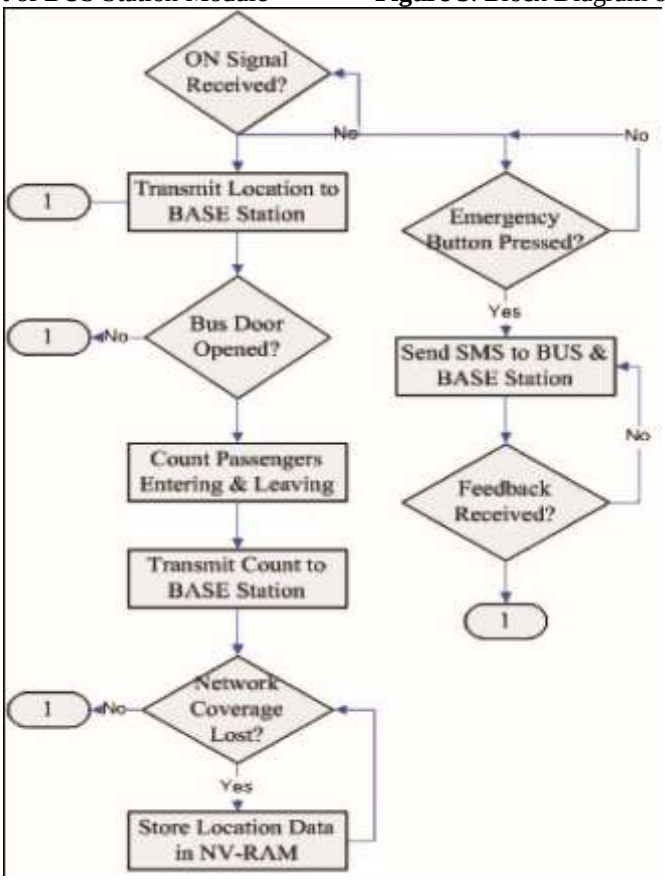


Figure 4. Flow Chart for In-BUS Module

D. BUS Stop Module

This module is installed at every bus stop to let the passenger know about the location of buses coming towards that stop. It comprises of a GSM modem, a NV-RAM and dot matrix display; all interfaced to 89C52 microcontroller. After receiving the bus location data in the form of stop names from BASE station, microcontroller stores it in non-volatile RAM. A sample message received by BUS Stop Module installed at 'N.G College Stop' is of the form "33, Clock Tower; 22, V.T Colony; 77, Srinagar colony. The message contains information of those buses only which will pass by the designated stop. First two digits of a sub-string denote the bus route number followed by the bus stop name which is the current location of bus coming towards the specified stop. Microcontroller after retrieving the stored information displays it on a 3x15 dot matrix display. The microcontroller refreshes the information with a rate of 10seconds. In case of an emergency situation, the location of next incoming bus is displayed. The block diagram of this module is shown in Fig. 7 and flow chart of module is shown in Fig 8.

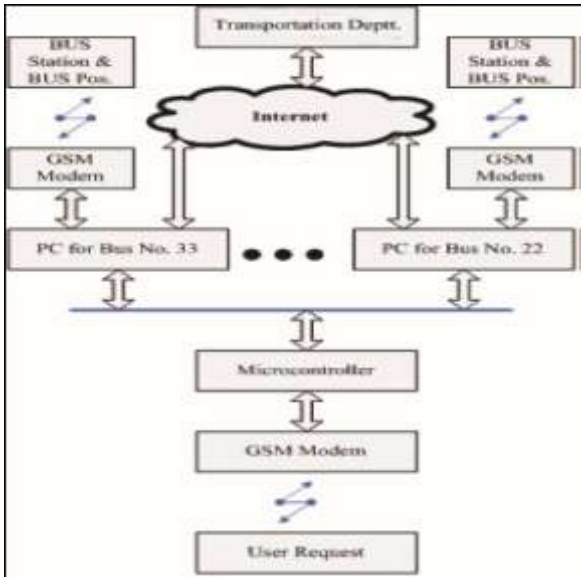


Figure 5. Block Diagram of BASE Station Module Module

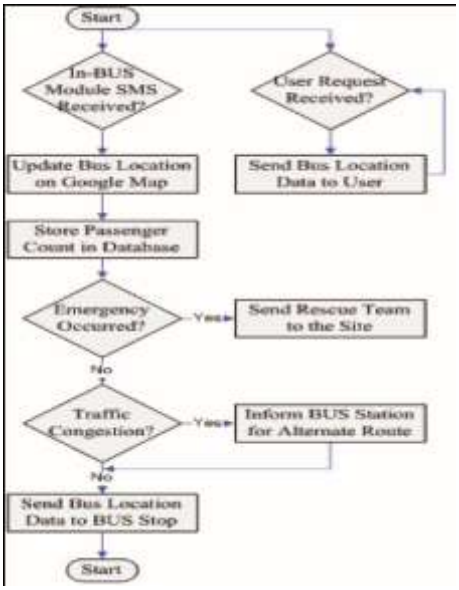


Figure 6. Flow Chart for BASE Station

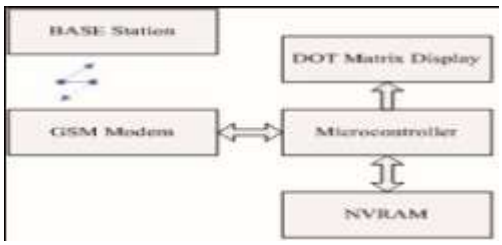


Figure 7. Block Diagram of BUS Stop Module

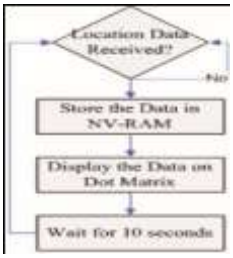
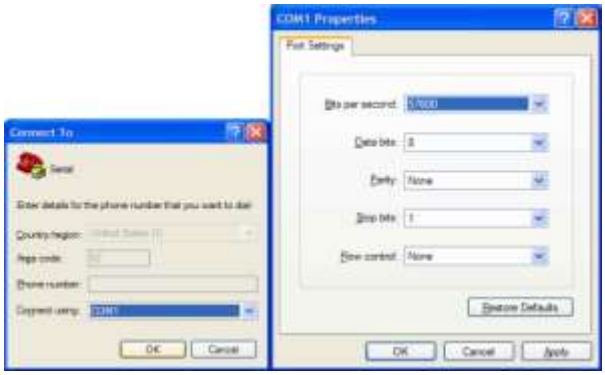


Figure 8. Flow Chart of BUS Stop Module

IV. SYSTEM SIMULATION RESULTS



Serial Terminal Program Configuration

Network Terminal Program Configuration

Windows based control panel to setup GSM modem, GPRS and send SMS

V. CONCLUSION

In this paper, design and development of a low cost transportation management system based on integration of GPS and GSM data is described. The system comprises of various modules which are wirelessly linked with GSM modems. Cost effective SMS service of GSM network is used for the transfer of data between the modules. A new service, to facilitate the people who use public transport for traveling, is introduced inside the city. The service provides the user with current location information of desired buses based on which the user can adjust his schedule accordingly. The service therefore vanishes the need of waiting at the bus stop thus saving a lot of time. For the passengers not utilizing the service, displays are installed at bus stop to let them know the buses location coming towards that stop. The system is also efficient in handling the emergency situations e.g., in case some kind of technical fault occurred in bus, the operator at bus terminal is informed and the departure time between the buses is reduced.

VI. FUTURE WORK

The system can be made automatic by installing cameras at bus terminals which can automatically read the license plate number of buses thereby eliminating the operator. An automatic route guider display can be installed in buses to better update the alternative route in case of serious road congestions. Fare collecting system can also be automated by providing another mobile service to which all the passengers using public transport are subscribed.

REFERENCES

[1] M.A.Mazdi,J.C.Mazdi,R.D.Mckinaly, The 8051 Microcontroller and Embedded Systems,Pearson education,2006
[2] Embedded Systems Design by Arnold S. Berger, 2001.
[3] The 8051 Microcontroller: Hardware, Software and Interfacing by James Stewart, Kai Mia.
[4] Internetworking with TCP/IP (VOL.I) Principles, Protocols and Architecture by Douglas Comer, 2000
[5] www.garmin.com/products/gps35
[6] www.mathworks.com
[7] www.d-d-s.nl/fotos-nokia/n12i.com
[8] www.alldatasheet.com

BIOGRAPHIES

Raghava Bhonagiri, presently working as Assistant Professor in ECE department at DVR College of engineering and technology, Hyderabad. He has got B.tech degree from Swami Ramanantha Thirtha Institute of science and Technology Nalgonda, A.P. He has got M.tech (Embedded Systems) from same Swami Ramanantha Thirtha Institute of science and Technology Nalgonda, A.P.

Prof.K. Subba Rao presently working as HOD in ECE department at DVR College of engineering and technology, Hyderabad. He has got B.tech degree from JNTU Hyderabad in1984. He has got M.tech degree from JNTU Hyderabad in1987. He has got PhD from IIT Kharagpur in1994. Previously he worked in CMC Hyderabad as Systems Engineer from 1984-1985. He worked as HOD in several colleges like Amaravathi University, Amaravathi, Auroras Engineering College, CVSR Engineering College, DRKIST in Hyderabad from 1988-2012.

Vinod Kumar Varkala, presently working as Assistant Professor in ECE department at TRR College of Engineering and Technology, Hyderabad. He got B.Tech degree in Electronics and Instrumentation from Swami Ramanantha Thirtha Institute of science and Technology Nalgonda, A.P. He got got M.Tech degree (Embedded Systems) from Swami Ramanantha Thirtha Institute of science and Technology Nalgonda, A.P.He has 3 years of teaching experience.

Gopikrishna Siliveri, presently working as IT-Engineer (Hardware/Embedded Systems) in R &D department at CMC Ltd , Hyderabad .He is having hands on experience on complete Hardware Design life cycle and development in the field of Real time embedded systems, . He has got B.tech degree in Electronic and Instrumentation from Swami Ramanantha Thirtha Institute of science and Technology Nalgonda, A.P.

.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 57-63
www.iosrjournals.org

Development of an ANFIS Neural Network System for Early Diagnosis of Critical Health Condition of a Patient

Taksala Devapriya A.[1], Deepesh B.[2]
[1](Electronics and Communication Engineering, Mountzion College of Engineering and Technology/Anna University, India)
[2](Electronics and Communication Engineering, Mountzion College of Engineering and Technology/Anna University, India)

ABSTRACT: *Aim of this work is to design and training of a ANFIS neural network based early diagnosis of a patient through a smart instrument. The data collected from the patient is used to train the Neural Network. The rules are designed on the basis of the collected data. The data collected from the patients are weight, BMI, Glucose, Creatinine, Systolic BP, and Diastolic BP. The rules for system are given according to the characteristics of the data obtained from patients. The output is mainly classified into three classes that are severe, moderately critical, and normal. Output of Neural Network is connected to LED to display the corresponding outputs. The advantage of this system is we can implement this device as portable. Hence we can easily monitor the patient's condition at anywhere. And for the normal type of application we need a personal computer to compute the inputs and outputs. Because here we are using FPGA the problem can be easily avoided and also PC will need an uninterrupted power supply. For this device just battery power is enough to work. The FPGA will take a very small power only hence we can easily use the device a long time without charging again and again. These are some of the major advantages of Smart device. The main disadvantage of using fuzzy logic is the reconfiguration problem. When we change the input parameters we should also change the rules. That is one of the important disadvantages of the fuzzy logic. This will also affect the cascaded structure of the neural network. Hence the performance of the system will effect sufficiently. The changing of rules can do only by the technical experts. This also will increase the complexity of the system. The problem can be easily solved by using ANFIS instead of fuzzy controllers. By using this we can avoid the re configuration complexity problem. Because ANFIS generates rules automatically from the training data. Hence no need for giving rules externally. Hence we can change the input data as per the user wish. Training complexity can be reduced by this method.*

Keywords: *Fuzzy Logic, ANFIS, ANN, SSA*

I. INTRODUCTION

TIn the Modern world peoples are living with different environmental condition has different physical problems because of Natural, psychological and personal duels. Medical diagnosis and decision making is a complicated and judgmental process. Diagnostic decisions made by physicians are highly variable (It may change from one physician to other) . It is not based only on medical knowledge derived from books and literatures and data obtained from various pathological tests, but also depend largely on experience, judgment and reasoning which essentially are the functions of human brain. However, in third world countries like India, doctors are scarcely available in rural areas. A recent statistical data shows that 75% of qualified consulting doctors reside in urban areas and another 23% in semi-urban areas, and only about 2% of doctors reside in rural areas, where, unfortunately, nearly 78% of Indians reside. This has created an unwarranted imbalance in patient-doctor ratio of more than 10,000 patients for one doctor in rural India. Apart from the acute scarcity of physicians, lack of availability of electricity is quite common in rural areas of third world countries like India. Such prosaic problems necessitate the use of an inexpensive, portable, low power battery operated high-speed. Recent development in the communication system is one of the big reliefs in medical side because of faster data transmission at the severe condition of a patient. It is helping to make the decision by a group of physician in different places. It is increasing the accuracy of the decision and the physicians can take turn the correct direction for the better diagnosis. Use of telecommunication to transmit health-related data to a remote centre for necessary help, has been the issue over the last several decades. Starting with a simple telephonic conversation with a doctor, the advent of internet technologies revolutionized the scenario. Using e-mail for sending patient data, sending images and physiological signals over internet by store and forward method and final video-conferencing all improved drastically tele-medicine setup for referral health services mainly. But everywhere, data, image and signals were transmitted for the ultimate decision taking by human brains. In the third world countries, an instrument with some auto decision making support can be used as a preventive device for early diagnosis of problems related to specified systems of patients' bodies. Especially in rural areas, where doctors

are scarcely available. Moreover, each doctor in these countries have to handle large number of patients, it may become handy for a doctor to keep track of the previous data of a particular patient when preservation of data and documents, like investigation reports, is poor in rural backdrop. This research focuses on the development of smart processing system that can predict the physiological state of a patient, given the past physiological data of the patient. The scheme under consideration can provide alarm to the relevant personnel in advance before the patient reaches a critical state, who would then contact the physician at the remote site. It will be very helpful to give a good diagnostic to the patient by proper medicine or surgery. The smart processing system consists of blocks for ANFIS, NNT and display unit. It can serve to handle the peripheral health screening of patients and thus screen out the few critical patients out of many who really need clinical assistance from the physician. In order to save the number of combinational logic blocks in implementing the system, the division process required for normalization of membership functions is implemented using intelligent multiplication techniques. Among other benefits are the short design cycle and the scope of re-programmability for improvement in the design without any additional cost along with the facility of desktop testing. The system can be used for diagnosis of other disease like diabetes and heart related problems.

II. FUNCTIONAL ARCHITECTURE OF SMART AGENT

2.1 The Smart Agent

Earlier diagnosis of a patient can be work out by using a smart agent. It is the combination of artificial intelligence algorithms such as ANFIS (Adaptive Neuro Fuzzy Inference System) and ANN (Artificial Neural Network). The schematic block diagram of the proposed architecture can be shown in Fig.1. It has the cascaded structure of ANFIS and Neural Network.The patients sample data is used to train the Neural Network and ANFIS. The attributes taken from the patients are given below. These six Attributes are entirely correlated with the renal failure. They are given below.

1. B.M.I (Body Mass Index)
2. Blood Glucose
3. Creatinine
4. Urea
5. Systolic Blood Pressure
6. Diastolic Blood Pressure

The data from the patient such as height or weight data cannot always be trusted as they are subjected to the quality and accuracy of measuring units and the skill of the technician.

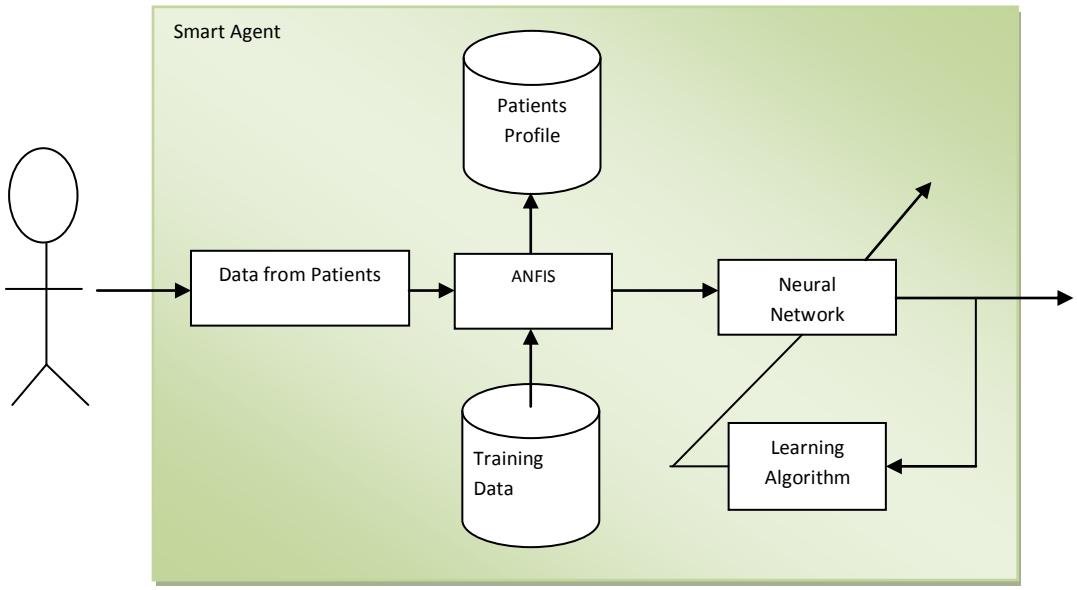


Fig.2. Block Diagram of Proposed Method

Moreover, based on a single data, it would be highly uncertain to make an accurate decision about the future physiological state of the patient. So the patient data have been fuzzified with the objective of transformation of periodic measures into likelihoods that the body mass index, blood glucose, urea, creatinine, systolic and

diastolic blood pressure of the patient is high, low or moderate. It has been established for fuzzification of patient data. These attributes are fully correlated with the renal failure. Hence the degree of accuracy is also depends upon the selected attributes.

2.2 ANFIS Training

ANFIS is trained with the patient's data samples with six attributes. Target for ANFIS is given by three classes. Three classes used to predict the system are:

1. Normal
2. Critical
3. Severe

The Dataset used to train is National renal database. In this the information collected is by 40 days interval. These data is collected and it is pre-processed i.e. aligned in a The data from the patient such as height or weight data cannot always be trusted as they are subjected to the quality and accuracy of measuring units and the skill of the technician. Moreover, based on a single data, it would be highly uncertain to make an accurate decision about the future physiological state of the patient. So the patient data have been fuzzified with the objective of transformation of periodic measures into likelihoods that the body mass index, blood glucose, urea, creatinine, systolic and diastolic blood pressure of the patient is high, low or moderate[1]. It has been established for fuzzification of patient data. These attributes are fully correlated with the renal failure. Hence the degree of accuracy is also depends upon the selected attributes.

2.3 Neural Network Training

Figure 2 shows the structure of neural network used. The left most layer is the input layer, which accepts the membership function values, which are presented to the hidden layer. The input layer consists of 60 neurons, the hidden layer consists of 18 neurons and the output layer consists of three neurons.

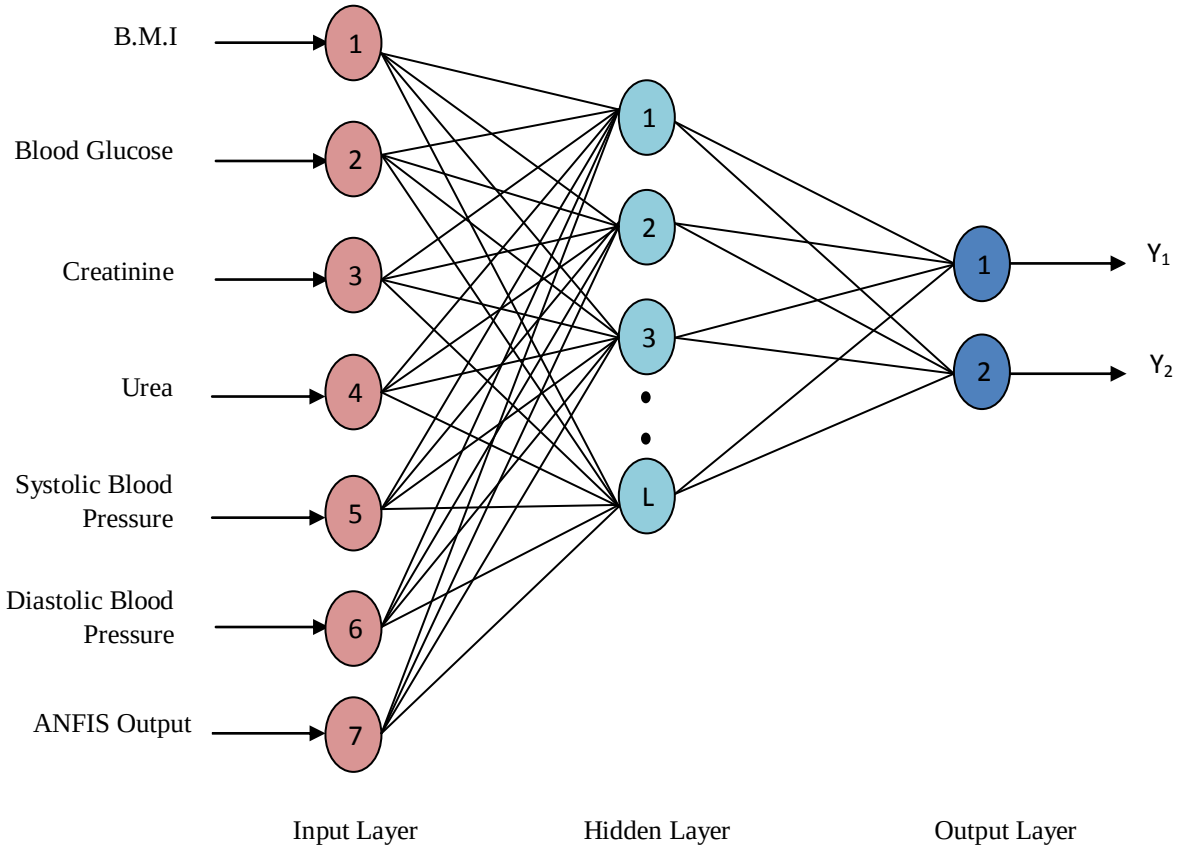


Fig.2. Neural Network Structure

The neural network calculates the possibilities of the pathophysiological parameters to below, moderate or high at the next instant of time[2]. The connection between the input layer and the hidden layer is shown as interconnection network box for clarity. There is a connection between every neuron in the input layer and every neuron in the hidden layer. The weighted sum of the possibility values from the hidden to the output layer

indicates the criticality of the current condition of the patient which is indicated as an output of the neural network.

III. INDENTATIONS AND EQUATIONS

3.1 Mathematical model of Neurons

The neural network proposed in the current work is based on inverse delayed function model of neuron[3]. The inverse delayed function model of neuron proposed is given by the following set of differential equations:

$$\tau \frac{du}{dt} = \sum w_{ij} xj + a_{ii} x_i - u_i \quad (1)$$

$$\tau_x \frac{dx_i}{dt} = u_i - g(x_i) \quad (2)$$

$$g(x_i) = f^{-1}(x_i) - Kx_i \quad (3)$$

Where w_{ij} is the synaptic weight between i^{th} and j^{th} neurons, a_{ii} is the synaptic weight of self-connection, u_i is the internal state of i^{th} neuron, x_j is the output of the j^{th} neuron, τ and τ_x are the time constant of the internal state and time constant of the neuron output, respectively. If *f(x)* is the sigmoid function used in a conventional neural network, and then $g(x_i) = f^{-1}(x_i)$ is the N-shaped inverse output function and *g(x)* can be changed with the positive parameter K.

The conversion from u to x has a very small response time compared to τ and thus $\tau_x \ll \tau_x$. However, the conversion time should take into consideration in general cases under the condition that it is much smaller than τ_x.

$$\tau_x \frac{d^2 x_i}{dt^2} = \frac{du_i}{dt} - \frac{dg(x_i)}{dx_i} \cdot \frac{dx_i}{dt} \quad (4)$$

Let

$$\phi_i = \frac{dg(x_i)}{dx_i} \cdot \frac{\tau_x}{\tau} \quad (5)$$

Putting Eq. (6) in (7) we get

$$\tau_x \frac{d^2 x_i}{dt^2} + \phi_i \frac{dx_i}{dt} - \frac{\tau_x}{\tau} \cdot \frac{dx_i}{dt} = \frac{du_i}{dt} \quad (6)$$

$$\Leftrightarrow \tau_x \frac{d^2 x_i}{dt^2} + \phi_i \frac{dx_i}{dt} = \frac{1}{\tau} \left(\sum_j w_{ij} x_j - g(x_i) \right) \quad (7)$$

$$\frac{\partial U_i}{\partial X_i} = \frac{1}{\tau} \left(g(x_i) - \sum_j w_{ij} x_j \right) \quad (8)$$

Therefore, Eq. (8) can be expressed as

$$\tau_x \frac{d^2 X_i}{dt^2} + \phi_i \frac{dx_i}{dt} = - \frac{\partial U_i}{\partial X_i} \quad (9)$$

Where

$$U_i = \frac{1}{\tau} \left(\int_0^{x_i} g(X_i) dx_i - X_i \sum_j w_{ij} X_j \right) \quad (10)$$

U_i represents the potential of the inverse delayed function model. Eq. (9) corresponds to the kinetics of a particle in presence of frictional forces. The first term in Eq. (9) represents inertia and the second term represents friction. If g

(x_i) is an N shaped function and $\frac{dg(x_i)}{dx_i}$ is $< -\tau_x/\tau$ over a certain range of x_i, (since τ_x/τ is always positive) then the inverse delayed model has the effect of negative resistance ($\phi_i < 0$). The energy function of the inverse delayed function model using Lyapnov function is given by

$$E = -\frac{1}{2\tau} \sum_i \sum_j w_{ij} x_i x_j + \frac{1}{\tau} \sum_i \int_0^{x_i} g(x_i) dx_i + \frac{\tau_x}{2} \sum_i \left(\frac{dx_i}{dt} \right)^2 \quad (11)$$

Where the self-connection between the neurons is ignored, since there is no self-connection between the neurons in our network. The last term appears in Eq.(11) because of time delay introduced in the inverse delayed function model. Differentiating both sides of Eq. (11) with respect to time t, we get

$$\frac{dE}{dt} = - \sum_i \frac{dx_i}{dt} \left\{ \frac{1}{\tau} \sum_j w_{ij} x_j - \frac{1}{\tau} g(X_i) - \tau_x \frac{d^2 x_i}{dt^2} \right\} \quad (12)$$

$$\Leftrightarrow \frac{dE}{dt} = -\sum_i \left(\frac{dg(X_i)}{dx_i} + \frac{1}{\tau}\right)\left(\frac{dX_i}{dt}\right)^2 \quad (13)$$

$$\Leftrightarrow -\sum_i \phi_i \left(\frac{dX_i}{dt}\right)^2 \quad (14)$$

From Eq.(14), we find that the time evolution of the system is a motion in the states pace that seeks out minima in E and comes to a stop at all such points if ϕ_i is positive. However, if ϕ_i is negative, the energy function does not monotonically decrease and can climb up the potential hill because of the effect to f negative resistance and can destabilizes table equilibrium points of the network within the negative resistance area. This can be made possible by harnessing the decreasing slope of N shaped characteristic of inverse a sigmoid activation function [4]. Hence, a neural network to answer an optimization problem can be constructed, as all units in the network should be at the corners of the solution space when its motion comes to as to pate global minima. On the other hand, when it comes to stop at the local minima, some of the units in the network are in the inside of the solution space. Therefore, it is expected to increase the possibility to get out from the local minima, if the network is an inverse delayed function model.

Neural Network Having 7 inputs because it is a cascaded structure of ANFIS and neural Network. One input is coming from the ANFIS and other six inputs are the six attributes. Hence the neural networks efficiency is doubled.

IV. FIGURES AND TABLES

As discussed above the ANFIS have six inputs. These six inputs are the six important attributes which are correlated with the renal diseases. Degree of prediction accuracy can be increased by increasing the number of dependent attributes. Hence the attributes should be selected in a proper manner that the increasing orders of priority. Sample patients information is given below. The following table data is used to train the ANFIS and Neural Network. The Neural Networks Accuracy is depending upon the number of Neurons in the hidden layer. The accuracy of the system is directly proposed to the Number of neurons. Even though, the system complexity will be increase if the number of neurons increases. Because the arithmetic and logical operations will increases if the number of neurons increases. Hence we should select the optimum number of neurons. It should depend upon the accuracy and the complexity of the developmental system.

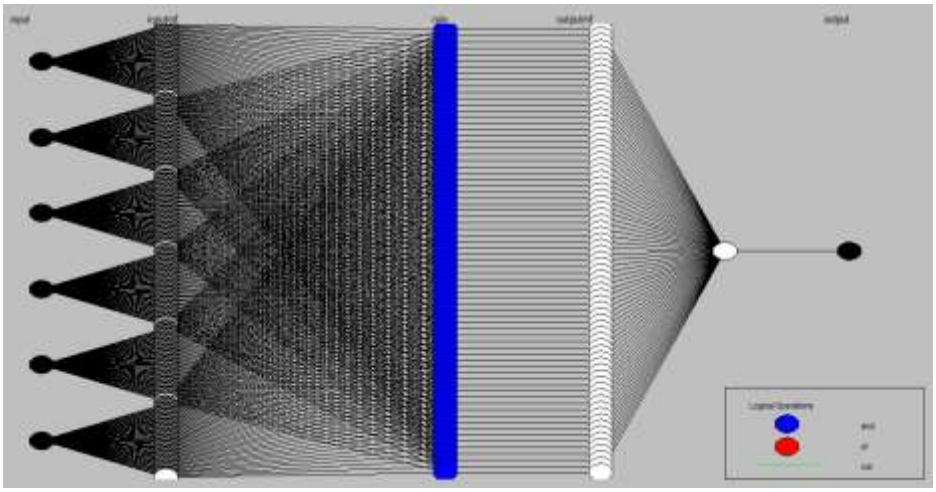


Fig.3. Structure of ANFIS

Similarly for ANFIS the training is performed on the basis of subtractive clustering algorithm. The training of ANFIS is done by ten epochs. The minimum gradient error will reach within this epoch is considering as the target. The training duration is depending upon the size of the dataset. If the size is very high training will take more time and also the ANFIS will generate a complex structure. Hence it is ensure that the number of epochs and the size of the training data should be select in an optimized way. It will increase the accuracy and reduce the complexity of the system. At the implementation view the complexity of the system should be low and also the accuracy should be very high.

The following surface plots show the membership function for the different attributes. Y-axis shows the degree of membership and X-axis and Z-axis are the corresponding attributes. The Y-axis is varies from 0 to 1. 0 represents the lower priority and 1 represent varies to the higher priority. The membership function varies in a linear fashion. The value 0 to 1 represents the Normalized value. For hardware implementation we can consider this value as 8 bit or 16 bit according to the user requirement. Normally it is the type of the processor or DSP

used. Increased number of bits will increase the resolution of the system. For the FPGA implementation we should select the number of bits according to the device used. Normally half precision floating point number is used. It has a maximum value of +1 and a minimum value of -1. Normally 16-bit is used to represent half precision fixed point number. 1 bit is used to represent the sign and remaining bits are used to represent the floating point representation

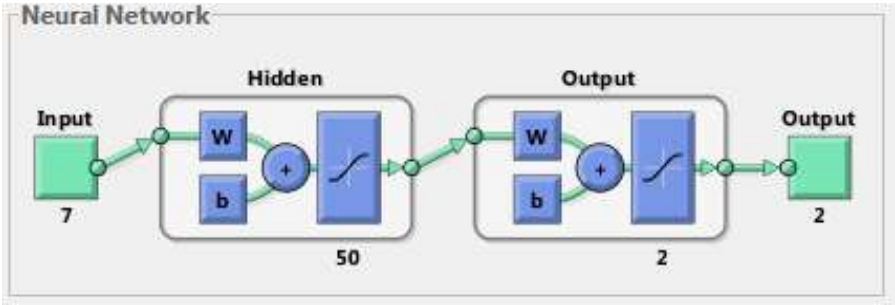


Fig.4. Structure of Neural Network

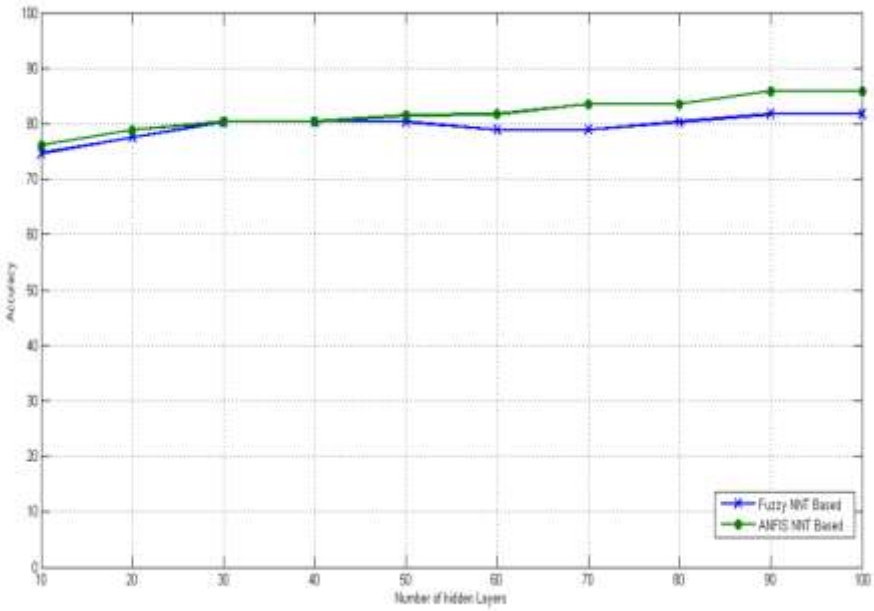


Fig.5. Performance

No of Neurons	10	20	30	40	50	60	70	80	90	100
ANFIS &NNT	76.1	78.9	80.3	80.3	81.5	81.7	83.5	83.5	85.9	85.9
Fuzzy &NNT	74.6	77.5	80.3	80.3	80.3	78.9	78.9	80.3	81.7	81.7

Table.1. Performance

V. Conclusion

The current work focuses on the application of Neuro-fuzzy synergism to detect at an early stage the probable approaching critical condition of a patient. It explains the usage of ANFIS neural networks in medical diagnosis systems and the extended example on the problem of early detection of approaching critical renal condition of patients. In order to improve the accuracy of diagnosis, optimal synaptic weights have been found out using the inverse delayed function model of the neuron. The neurons that are considered in the proposed network are devoid of self connections unlike the previously proposed networks based on inverse delayed function model of neurons where self-connected neurons have been assumed. For diagnosis purposes, body mass index (B.M.I.), glucose, urea, creatinine, systolic and diastolic blood pressures are considered as

pathophysiological parameters. The ANFIS neural network has been suitably trained and tested with real patient data to find out the correspondence with the decision being given by the fuzzy neural network and the actual pathophysiological state of the patient. A reasonably high accuracy of 85% in medical diagnostic decision-making has been obtained with the implemented system.

REFERENCES

Journal Papers:

[1] M Venkatesan, Classification of Renal Failure Using Simplified Fuzzy Adaptive Resonance Theory Map, IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.11, November 2009.

[2] D. Sheppard,Predicting cytomegalovirus disease after renal transplantation: an artificial neural network approach,International Journal of Medical Informatics 54 (1999) 55–76.

[3] ZaritaZainuddin, A Neural Network Approach in Predicting the Blood Glucose Level for Diabetic Patients, International Journal of Information and Mathematical Sciences 5:1 2009.

[4] Adenike O. Osofisan, Prediction of Kidney Failure Using Artificial Neural Networks, European Journal of Scientific Research ISSN 1450-216X Vol.61 No.4 (2011), pp.487-492

Author's Profile

Fdsfgsgdfgdfg

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 64-72
www.iosrjournals.org

Design Analysis of a Security Lock System using Pass-Code and Smart-Card

[1]Omijeh, B.O and [2]G.O.Ajabuego
[1,2] *Department of Electrical / Electronic Engineering, University of Port Harcourt*

Abstract: *In this work, Security Lock system using Pass code and smart card has been achieved. It is a micro controller based system which aims at interfacing selected hardware components with software to obtain a higher degree of security. In achieving this work, seven major stages were considered in the design methodology: Power supply unit, personal computer unit, micro controller unit, lock system unit, seven segment display, detector unit and alarm unit. The design specifications were strictly followed and the result obtained after construction and testing were very satisfactory. Indeed, security lock system using passcode and smart-card is a surer and safer form of security in this era of tremendous technological advancement.*
Keywords*: Pass-code, Smart-card, Microcontroller, Interfacing, Security Lock System.*

I. Introduction

One of the cardinal needs of man is security of life and property. Over the centuries, access control systems were put in place to prevent access to unauthorized persons. They are called locks on doors [1]. However, locks that are operated by keys have weaknesses such as the ability to be picked. Also, there are situations of unauthorized duplication of keys, the problem that a key has to be replaced if they are found missing. Further more, theft poses a very substantial problem. It is often necessary to change door locks when it is suspected that keys may have fallen into unauthorized hands [2]. In the case of a fire hazard, when one needs to get out of the building quickly, it is not advisable to waste time searching for the key to unlock the door. Hence, in a bid to handle these shortcomings, some security devices have been introduced. For example, smartcards [3][4] are now available in banks, hotels, offices and high profile establishment but not to the general populace. Recently, the use of biometrics[5] and voice automation have been introduced.

This paper, therefore, presents the design analysis of a security lock system using pass code and smartcard with the following design objectives:

- To gain access using the right code and remain locked if the right code is not entered.
- To start an alarm when the wrong code is entered more than two times.
- To retain a high level of exclusivity within the access control region.
- To reduce the number of intruders.
- To protect against unauthorized duplication of keys that operate hardware locks.

II. Building Blocks Of A Security Lock System

Fig 1 shows the block diagram for the design methodology. It consists of seven stages

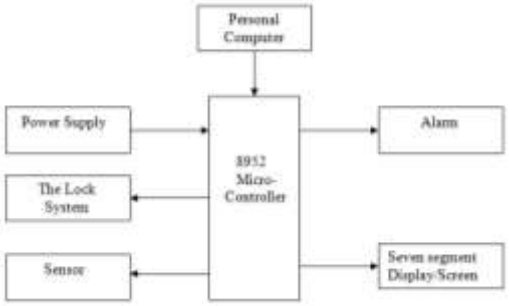


Fig 1: Block diagram of a security locks system

Power Supply Unit:

In order to operate a system or equipment there is a demand for power. The mode of application of such power will form the design of the power supply unit and its characteristic.

A 5V dc Power Supply is needed to power the microcontroller and all the other modules. Also, 12V dc power supply is used for operating the motor.

Personal Computer:

The programming language used in designing the graphical user interface is visual basic. It is a terminal emulation program. It is the interface between the personal computer and the microcontroller circuit. Programs run on the microcontroller are written in assembly language. Communication between the microcontroller and the system is through serial communication technology.

There is need for a logic level converter Max232 because the logic level of the microcontroller and the serial port of the system are different. The microcontroller uses TTL standard (Transistor Transistor Logic). Logic 1 represents 5volts; and logic 0 represents 0volts. Also, the serial port of the computer uses RS 232 standard. Logic 1 represents a range of -3 to -15volts while logic 0 represents a range of +3 to +15volts. Hence it is the logic level converter Max232 that makes it possible for the two devices to communicate.

Microcontroller

The AT89c52 microcontroller is the brain behind the locking device. It is the meeting point for all the other units. They are all connected to it. The microcontroller inputs and outputs signals (sends and receives signal). This method is known as flow control. There are two types of flow control: software flow control, hardware flow control. In this project, we used the hardware flow control because it is the most reliable. In the hardware flow control, the microcontroller uses request to send (RTS) from the computer and clear to send (CTS) from the microprocessor (Carter, 1996; Steiner, 2006).

Lock System:

Debounce push buttons are used for the Passcode lock system. When they are pressed, the logic state changes. The microcontroller detects the change in logic state and displays the number pressed. Once six digits are pressed, the microcontroller compares these digits with the preset passcode loaded from the personal computer. If they are the same, the system opens the door (Jean, 2001).

With the card lock system, a Vero board is used to construct the smart card that has binary codes. When it is inserted into the card slot, the smart card is powered from the card slot. The system reads the binary codes which are compared with a preset binary code. If the measured code is equal to the binary code, the system is activated to open the door. If not, the card is rejected.

Display unit.

The display unit (seven segment display) is used to display information. The light emitting diode is arranged properly to make up the figures on the display .

Detector Unit:

The sensor used in this project is an infrared. It is a device capable of detecting and responding to physical stimuli such as movement, light, or heat. For the purpose of this project, the infrared detects the move of a human being The body of the human being emits the infrared rays. A change in infrared rays is detected by the infrared which causes the relay to operate. This enables the motor to move either right or left so as to open or close the door [9]

Alarm unit:

In this unit, an electronic bell is used. It is programmed to ring when the wrong code is entered more than twice. It has inbuilt speakers that enable the sound of the bell to be heard loudly.

III. Design Analysis Security Lock System

In the implementation of this work, certain logical steps were taken into consideration in determining the hardware components necessary. These steps followed the model of the design. Other necessary support components were identified. The detailed circuit diagram is shown in Fig.2 [10].

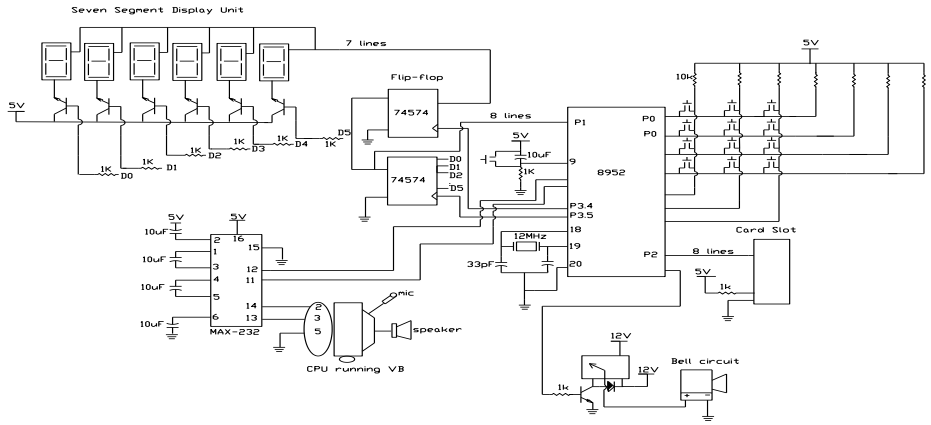


Fig 2: Detailed Circuit Diagram of Lock System

3.1 Power Supply Unit

The power used in this circuit is 5volts and 12volts dc power supply. The 5V is used to operate the microcontroller and other components, while the 12V is used for operating the motor.

3.1.1 The Transformer Unit

The transformer used in this project, is a 220volts / 12volts ac step – down transformer. It takes in 220V ac and produces 12V ac, with a load current of 300mA. It accomplishes this by electromagnetic induction and when the electric circuits are in mutual inductive influence of each other.

3.1.2 Rectification

The purpose of a rectifier is to converts ac to dc signal. The rectifier used in this work is a full wave bridge rectifier. The bridge rectifier unit is rated for a current at least as high as the transformer's secondary winding is rated for, and for a voltage at least twice as high as the RMS voltage of the transformer's output (this allows for peak voltage, plus an additional safety margin). It takes in 12 volts ac from the transformer and produces 12 volts dc. This dc has alternating current component. Thus the output is passed onto a capacitor for filtering.

3.1.3 Power Supply Filtering

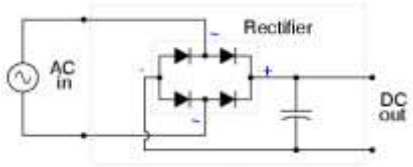


Fig 3: Schematic diagram of power supply filtering

The rectified waveforms contain ripples that have to be smoothed out in order to generate a genuine DC voltage. This we did by taking on a low-pass filter. The voltage rating of the capacitor should be greater than or equal to 1.5 the input voltage. I.e. $V \geq 1.5$ (input voltage). A capacitor of 3300uf / 50V was employed in this design.Fig 3.

This ensures small ripples, by making the time constant for discharging much longer than the time between re-charging. Also, a large capacitance value was considered okay, so long as its working voltage was high enough. However, for safety reasons, we chose a capacitor of 3300uf / 50V. Fig 3 and Fig 4 show the Schematic and the resultant waveforms after rectification.

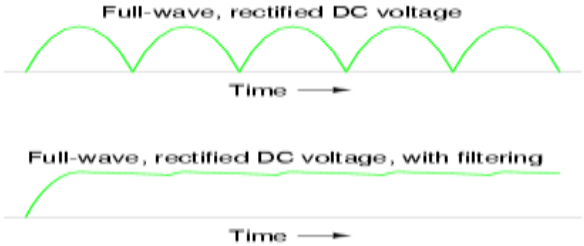


Fig 4: Waveforms after Rectification

3.1.4 Voltage Reference.

7805 IC is used as a voltage regulator to bring the voltage down to 5volts stable power source for the operation of the ICs.

3.2 Microcontroller.

The microcontroller used is the AT89c52, a variation of 8051 microcontroller architecture. The 8051 is the most popular microcontroller available today. Hence, it comes with a wide variety of features which makes it accessible to a wide range of applications.

The AT89c52 microcontroller has the following features:

1. 32 programmable I/O lines.
2. 8 Kbytes ROM.
3. 256 bytes of RAM and 256 bytes of special function register.
4. On-chip oscillator and clock circuitry.
5. Universal Asynchronous Receiver / Transmitter (UART) serial port.
6. Interrupt sources.
7. Power saving operating mode.

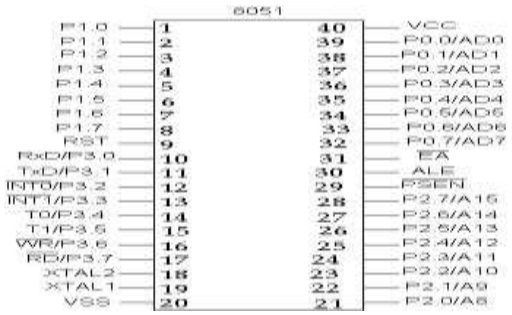


Fig 5: Pin Configuration of AT89c52

The description of the Pin Configuration of AT89c52 is as shown in Fig 5

- Vcc pin 40 : Provides supply voltage to the chip. The voltage source is +5V.
- GND pin 20: is used for earthing.
- XTAL1 and XTAL2 pins 19,18 : These 2 pins provide external clock
  - Way 1 using a quartz crystal oscillator
  - Way 2 using a TTL oscillator
- RST pin 9 is reset
  - It is a power-on reset.
  - Upon applying a high pulse to RST, the microcontroller will reset and all the values in the registers will be lost.
  - /EA - pin 31 is external access: The AT89c52 microcontroller can execute a program in external memory only if EA is low. The sign "/" means active low.
  - For this project, /EA pin is connected to Vcc because we are not using external memory i.e. external RAM
- /PSEN（pin 29）: program store enable and ALE（pin 30）: address latch enable are used for external ROM.
- There are four ports P0, P1, P2, and P3.
  - Each port uses 8 pins. They provide the user with 32 I/O lines for connecting the microcontroller to the environs.
  - All I/O pins are bi-directional.

Working Principle:

- **Machine cycle**

For the CPU to execute an instruction, it takes a certain number of block cycles. In the 8051 family, these block cycles are referred to as machine cycles.

The frequency of the crystal connected to the 8051 family varies from 4MHz to 30 MHz, depending on the chip rating and manufacturer. Very often the 11.0592 MHz crystal oscillator is used to make the 8051-based system compatible with the serial port of the IBM PC. one machine cycle lasts for 12 oscillator periods.

- **Interrupt:**

An *interrupt* is an external or internal event that interrupts the microcontroller to inform it that a device needs its service.

Steps in executing an interrupt:

I. It finishes the instruction it is executing and serves the address of the next instruction (PC) on the stack.
II. It also saves the current status of all the interrupts internally (i.e. not on the stack)

III. It jumps to a fixed location in memory called the interrupt vector table that holds the address of the interrupt service routine (ISR).
IV. The microcontroller gets the address of the ISR from the interrupt vector table and jumps to it. It starts to execute the interrupt service routine until it reaches the last instruction of the subroutine which is RETI (return from interrupt)
V. Upon executing the RETI instruction, the microcontroller returns to the place where it was interrupted. First, it gets the program counter (PC) address from the stack by popping the top two bytes of the stack into the PC. Then it starts to execute from that address.

3.3 Computer Interface unit.

The microcontroller used in this project required TTL logic levels (0 and 5V) for effective programming. Therefore the first step to connecting this device to the RS-232 port is to transform the RS-232 levels to 0 and 5 volts. This we did using the **Max-232 level converter.**

It includes a charge pump which generates +10 and -10V from a single 5V supply. It also includes two receivers and two transmitters in the same package which could be handy when only transmit and receive data lines and needed.

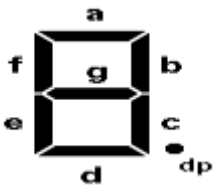
Serial Communication: This is bit by bit transmission of data on a single line. It is different from parallel communication where each bit of the data is transmitted on different lines. When interfacing a terminal to the microcontroller, parallel communication requires a lot of pin, about 8 pins and about 3 other pins for flow control. Whereas, serial communication requires 2 pins for data transfer and if desired, 2 other pins for flow control

For this design, only 5 of the DB-9 connector pins were used.

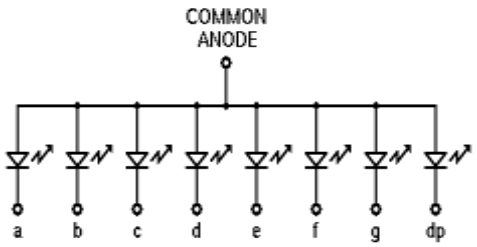
1. Received data (RxD) pin: This pin is connected to the microcontroller's serial port transmit line
2. Transmit data (TxD) pin: This pin is connected to the receive line of the microcontroller's serial port receive line.
3. Request to Send (RTS): This pin is connected to any pin of the microcontroller. Whenever the computer wishes to send data to the microcontroller it sends a low through this line.
4. Clear to Send (CTS): This pin can be connected to any pin of the microcontroller. The microcontroller sends a low when it receives a low from RTS. The computer when it receives this signal, begins the transmission
5. Signal ground: This is connected to GND of the microcontroller circuit to create a common reference for both terminals.

3.4 Seven Segment Display

It is less commonly known as a seven-segment indicator. It is an assembly of light emitting diodes which can be powered individually. For the purpose of this work, Seven Segment displays are used in displaying decimal numeric feedback on the internal operations of devices. They are arranged and labeled as shown in the diagram[12]



(a) Seven-Segment Display Layout



(b) Seven Segment LED Display Component.

Fig 6: Seven-Segment Display (a) Display Layout (b) LED Display Component.

Powering the entire segment will display the number 8. Powering a, b, c, d and g will display the number 3. Numbers 0 to 9 can be displayed. The one shown is a common anode display since all anodes are joined

together and go to the positive supply. The cathodes are connected individually to 0V. Resistors are placed in series with each diode to limit the current through each diode to a safe value.

3.5 Infrared Circuitry

IR Transmitter:

The current through the LED varies from 100mA to well over 1A. In order to get an acceptable control distance the LED currents are as high as possible. LED currents are high because the pulses driving the LEDs are very short. [11]

A simple transistor circuit is used to drive the LED. A transistor with a suitable HFE and switching speed was selected for this purpose.

Based on Ohm's law and with the knowledge of the nominal voltage drop over an IR LED, which is approximately 1.1V, the resistor value used is 11Ω.

IR Receiver

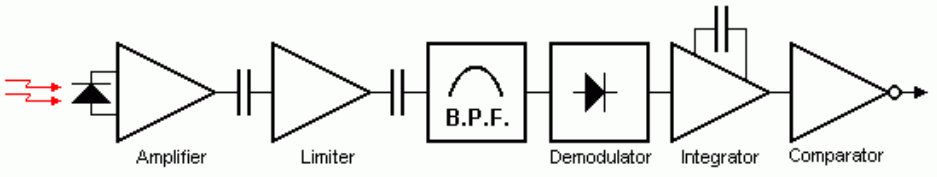


Fig 7: IR Receiver

Fig.6 is a typical block diagram of an IR receiver. Everything is built into one single electronic component. The received IR signal is picked up by the IR detection diode on the left side of the diagram. This signal is amplified and limited by the first 2 stages. The limiter acts as an AGC circuit to get a constant pulse level. Only the AC signal is sent to the Band Pass Filter. The frequency ranges from 30 kHz to 60 kHz. Fig 7

The next stages are a detector, integrator and comparator. The purpose of these three blocks is to detect the presence of the modulation frequency. If this modulation frequency is present the output of the comparator will be pulled low.

3.6 Clocked D Flip-Flop

Flip flops are generally storage devices and they have one synchronous control input. 'D' stands for data. The figure below shows the symbol for a clocked D flip flop that triggers on a positive-going transition (PGT). Fig 8 shows the waveform.

The principle of operation of the D flip flop is very simple. The level present at D is stored in the flip flop at the instant the PGT occurs. The waveform in the figure below illustrates this operation. The D type flip-flop has only one input (D for Data) apart from the clock. Thus, the INDETERMINATE state is avoided with this flip-flop. When the clock goes high, D (a 0 or a 1) is transferred to Q. When the clock goes low, Q remains unchanged. Q stores the data until the clock goes high again, when new data is available

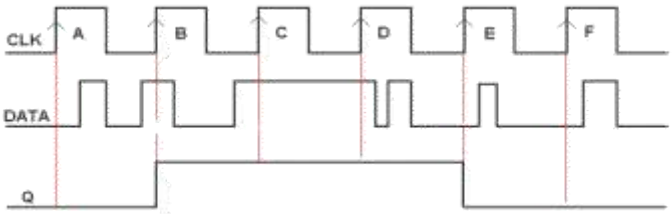


Fig 8: Waveform of Clock D Flip-Flop

Table 1: Truth table for clock D flip-flop:

D	CLK	Q
0	↑	0
1	↑	1

3.7 OR Gate

Fig 9 shows the OR gate which allows the output to be true (logic 1) if any one or more of its inputs are true.

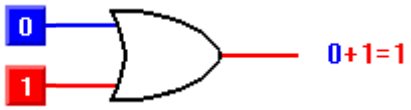


Fig 9: OR gate

Table 2: Truth table for OR gate

A	B	Y
1	x	1
x	1	1
0	0	0

3.8 Relays: Relays are electromagnetic devices that allow small electronic signals to control very large electrical signals and electronic signals. The operation of the relay depends on the voltage rating and the current rating. The voltage rating is the voltage that is applied across the terminals of the electromagnet. Current rating is the maximum current that the material can withstand. In this project, the relay rating is 12Vdc, 10A. Hence,

$$P = IV$$

the load;

$$10 \times 220 = 2{,}200VA$$

3.9 Lock Section

Password Section: Here, micro switches are used to enter the right codes which the microprocessor uses to compare the already established codes in the processor.

The shift register consists of two flip-flops connected together so that data can be transferred (shifted) along the chain from one end to the other. Calculations carried out to determine the value of the base resistors for the transistors used for the seven segment display, is as follows.

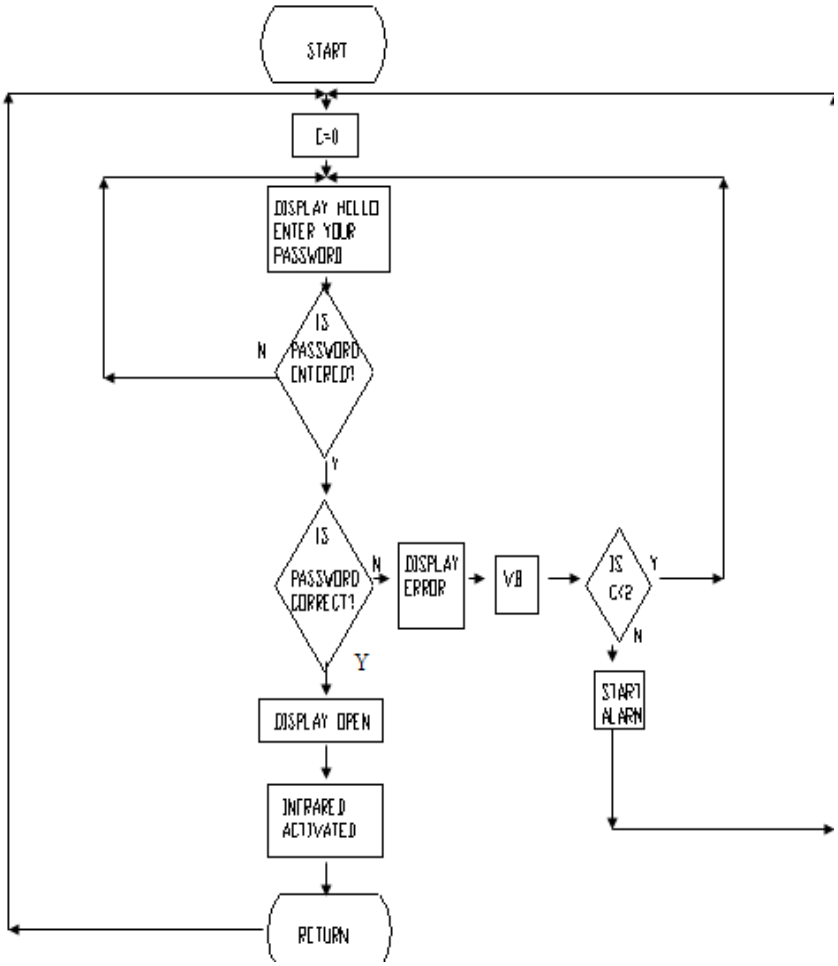


Fig 10: Flow Chart of operation process

$BIb = Ic$.....(1)
$B = 300, Ic = 0.9A$
$Ib = 0.9 \div 300 = 0.003A = 3mA$
$Ie = Ic + Ib$.....(2)
$Ie = 0.9 + 0.003 = 0.903A$
$Ve = 1/4Vcc$.....(3)
$Vcc = 5V$
$Ve = 5/4 = 1.25V$
$\mathrm{Re} = Ve \div Ie$
$\mathrm{Re} = 1.25 \div 0.903 = 1.384\Omega$
$Vb = 0.7$
$Vcc = IbRb + Vbc + Ie\,\mathrm{Re}$.....(4)
$5 = 0.003Rb + 0.7 + (0.903 \times 1.3843)$
$= 0.003Rb + 0.7 + 1.2500229$
$= 0.003Rb + 1.9500229$
$5 - 1.9500229 = 0.003Rb$
$3.04998 = 0.003Rb$
$Rb = 3.04998 \div 0.003 = 1{,}016.66\Omega$
$\approx 1000\Omega = 1k\Omega$

Smartcard Section: For the Card lock system, we used the following

- A Vero board of size $4.5 \times 3.7cm$. It is used as an interface card.
- 8 bit expansion slots; this was gotten from the system board of a PC. The Vero board (interface card) communicates with the system through the extended microprocessor buses in these slots.
- Jumper wires were soldered on the Vero board for the purpose of continuity so as to have alternate 0s and 1s e.g. 11110000 or 11111000
  The features of the 8 bit expansion slots are:
- 8 bit bidirectional data bus,
- 20 address lines for the IO channel.
- 6 interrupts channels
- Control signals for memory and IO read or write operations,
- Clock and timing signals and
- Three channels of DMA control lines.

Software Design

Fig 10 shows the Flow Chart used for the software design.

IV. Results And Discussion

The test results in Table 3 shows good response time and performance to design specifications.

Table 3: Testing and Results

Check	Action	Result	Comments
	On the VB interface, click on password and then type in the required code that is to be compared and press LOAD	Please enter your code.	This activates the system to start accepting security code.
Blinking of Red LED's	Type in the expected code on the device and press enter	Seven segment display will show open if correct code is entered.	The expected code is compared with the required code and if they match, the door will open automatically.
Door opens and remains open for about 19secs	Key in the wrong code	The display shows error.	If RED LED indication entered code is not valid then process stops. Try entering the code again.

Door will not open	Enter code more than 3 times	Alarm comes on for 14 secs.	Purpose of the alarm is to stop an intruder from trying different combinations in an attempt to break in or gain access.
	On the VB interface, click on smart card and then type in the required code that is to be compared and press LOAD	Please enter your smart card.	This activates the system to start accepting smart cards.
Blinking of Red LED's	Insert the expected smart card into the slot.	Seven segment display will show open if correct smartcard is inserted	The expected code on the smartcard is compared with the required code and if they match, the door will open automatically.
Door opens and remains open for about 19secs	Insert the wrong smartcard.	The display shows error.	If RED LED indication entered code is not valid then process stops and door will not open.

VI. Conclusion

A Security lock system using Pass-codes and Smart-card has been designed, analyzed and implemented. Its advantages cannot be over-emphasized in these modern days of technological breakthroughs. The addition of a PC interface enables flexibility in terms of manipulating the different combinations of pass codes and smart card. Security lock system using Pass-codes and Smart-card would be found very useful in applications which require strong security protection and authentication such as: Banks, Industrial Facilities, Commercial Buildings, Airports, Museums and Fine Art, Warehouses e.t.c.

Reference:

[1] Bellis M. (1972); "The history of locks" www.About.com (28th March, 2007)
[2] Abraham.D.G, GM Dolan, GP Double, JV Stevens (1991) ``Transaction Security System", in *IBM Systems Journal* v 30 no 2 (1991) pp 206-229
[3] Reid, J. & Looi, M., Making Sense of Smart Card Security Certifications, *Fourth Working Conference on Smart Card Research and Advanced Applications*, 20-22.9.2000, Bristol, UK.
[4] Ranke.W and Efing. W (2000) : Smart card hand book, 2nd Edition, New York, John Wiley AND Sons
[5] Bovenlander. E & Van Renesse,(1995) ``Smartcards and Biometrics: An Overview", in *Computer Fraud and Security Bulletin* (Dec 95) pp 8-12
[6] Carter .J (1996): Microprocessor Architecture and Microprogramming - A State-Machine Approach, Prentice-Hall, 1996.
[7] Steiner C. (2006); "8052 Instruction set" www.8052.com. (18th April, 2007).
[8] Jean-Jacques Quisquater and David Samyde (2001): ElectroMagnetic Analysis (EMA): Measures and Couter-Measures for Smard Cards. In Smart Card Programming and Security (E-smart 2001), Cannes, France, LNCS 2140, pp.200-210. September 2001.
[9] Schneize.B, A. Shostack, "Breaking Up Is Hard To Do: Modeling Security Threats for Smart Cards", USENIX Workshop on Smartcard Technology, Chicago, Illinois, USA, May 10–11, 199
[10] Kelly B. (1995); "IR Transmitter and Receiver" www.ask.com. (28th April, 2007)
[11] Odia.A (2006) – Design and Construction of a Security Lock System using Pass-Code , Smart-Card, voice Technology, Department of Electrical/Electronic Engineering, University of Port Harcourt, Port Harcourt, 2007
[12] Tocci J.,"Digital Systems Principles and Applications", Prentice-Hall Publishers., India, 2004, p.511

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 73-78
www.iosrjournals.org

Digital Audio watermarking using perceptual masking: A Review

Ms. Anupama Barai[1], Associate Prof. Rohini Deshpande[2]
[1, 2](EXTC, K J Somaiya College of Engineering/ Mumbai University, India)

Abstract : *Sharing electronic files on the internet has grown extremely fast over the last decade due to large volume of the mobile phones. These files include diverse forms of multimedia such as music, video, text and image. However digital files can be easily copied distributed and altered leading to copyright infringement. It is this ease of reproducing that causes copyright violations. Composers and distributors are more focused on implementing digital watermarking techniques to protect their material against illegal copying. Digital audio watermarking technique protects intellectual property by embedding watermark data into the audio file and recovering that information without affecting the audio quality of the original data. In this paper an overview of fundamental concepts of digital audio watermarking using perceptual masking is presented which includes a watermarking procedure to embed copyright protection into digital audio by directly modifying the audio samples. The procedure directly exploits temporal and frequency perceptual masking to guarantee that the embedded watermark is inaudible and robust. The watermark is constructed by breaking each audio clip into smaller segments and adding a perceptually shaped pseudo-random sequence. The noise-like watermark is statistically undetectable to prevent unauthorized removal.*

Keywords -*Audio watermarking, Copyright Protection, Embedding, Frequency Masking Perceptual masking, Psychoacoustic Auditory Model, Temporal Masking.*

I. INTRODUCTION

On line distribution of digital media including images, audio, video and documents has proliferated rapidly in recent years. In such environment it is convenient to get the access to various information resources. Although with the ease by which the digital formatted data can be copied and edited, copyright infringement like illegal reproduction and distribution has arisen and greatly spoilt the originator's passion for innovation. To prevent such iniquities, the enforcement of ownership management has claimed more and more attention. As a result, novel watermarking technique is introduced for copyright protection [1].

Watermarking is the process of encoding hidden copyright information in digital data by making small modifications to the data samples[3]. Unlike encryption, watermarking does not restrict access to the data. Once encrypted data is decrypted, the media is no longer protected. A watermark is designed to *permanently* reside in the host data. When the ownership of a digital work is in question, the information can be extracted to completely characterize the owner. Most schemes utilize the fact that digital media contain perceptually insignificant components which may be replaced or modified to embed copyright protection. However, the techniques do not *directly* exploit spatial/temporal and frequency masking. Thus, the watermark is not guaranteed inaudible. Furthermore, robustness is not maximized. The amount of modifications made to each coefficient to embed the watermark is estimated and not necessarily the maximum amount possible.

In this paper, we introduce a novel watermarking scheme for audio which exploits the human auditory system (HAS) to guarantee that the embedded watermark is imperceptible. As the perceptual characteristics of individual audio signals vary, the watermark adapts to and is highly dependent on the audio being watermarked[2]. The watermark is generated by filtering a pseudo-random sequence (Author id) with a filter that approximates the frequency masking characteristics of the HAS. The resulting sequence is further shaped by the temporal masking properties of the audio. Based on pseudorandom sequences, the noise-like watermark is statistically undetectable.

II. DIGITAL AUDIO WATERMARKING

Digital watermarking is the process that embeds copyright information as watermark into the multimedia object, so that the watermark can be extracted to make an assertion about the ownership. The general schematic diagram of watermarking is shown in figure 1.

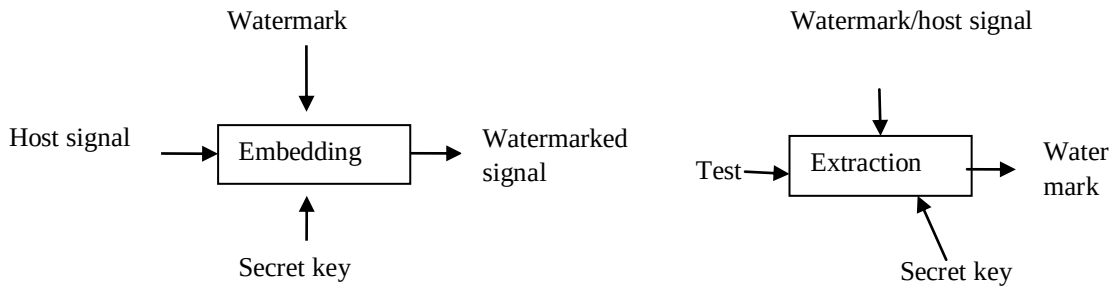


Fig.1 General Schematic diagram of watermarking

From the point of view of information hiding, watermarking is a technique of hiding messages, may be secret signatures, in the host carrier for the purpose of identification annotation and rights management. For example, a common form of hidden writing is using invisible inks. So the secret message can only be read by processed with some prescribed chemicals in a certain sort of way. Specifically, research is focused on embedding imperceptible robust and secure watermark for copyright protection. These watermarks are permanent signatures, difficult to remove without degrading the quality of the host media. When disputes happen the watermark could be extracted as reliable proofs for assuring the authentication[5]. The further subheadings cover the fundamental concepts of the basic theories Psychoacoustic masking and Spread Spectrum, watermarking and extraction.

III. Psychoacoustic Masking

Psychoacoustic explains the subjective response to everything we hear. It seeks to reconcile acoustic stimuli scientific, objective and physical properties that surround them with the physiological and psychological responses evoked by them. In simple words it is the science that studies the statistical relationships between acoustical stimuli and hearing perception, in order to explain the auditory behavioral responses of human listeners, the abilities and limitations of the human ear and the auditory complex process that occur inside the brain. Because the process of embedding the watermark is required to be imperceptible, understanding the principles of psychoacoustic and making use of its perceptual properties is helpful for watermark implementation.[6]

Hearing involves a behavioral response to the physical attributes of sound including intensity, frequency and time based characteristics that permit auditory system to find clues that determine distance, direction, loudness, pitch and tones of many individual sounds simultaneously. It is the sense by which sound is perceived hand hearing of humans is performed by the auditory system: sound as pressure waves is detected by the ear and transduced into nerve impulses that are perceived by the brain. The range of human hearing is 20 Hz to 20 kHz, where human speech mainly falls between 100 Hz and 8 kHz. The ear plays an important role in human auditory system HAS[7]. Ear is sub divided into outer, middle and inner ear. Outer ear i.e. pinna captures sound and directs through the ear canal till tympanic membrane (eardrum). Middle ear processes the sound waves into mechanical vibrations through ossicles till oval window. These vibrations are transduced into neural impulses within the cochlea of inner ear. These then enter the brain via auditory nerve fibres. The Cochlea is the main organ. It's a snail shaped fluid filled chamber separated by basilar membrane. The basilar membrane is about 32mm long and the organ of Corti the receptor rests on it. The organ of Corti contains specialized cells called hair cells including inner and outer which translate fluid motion into electrical impulses for the auditory nerve.

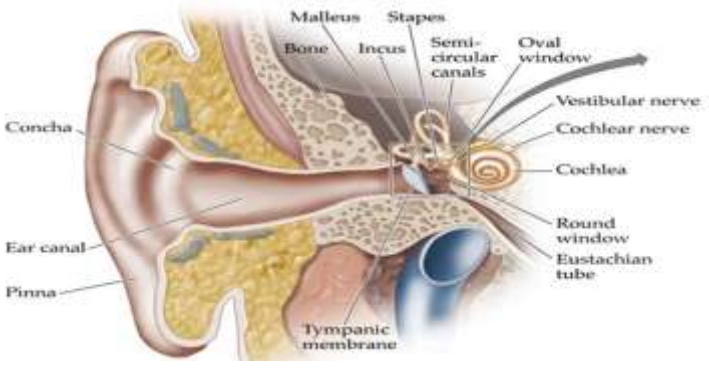


Fig 2. Diagram of Human hear

Experimental studies show that the basilar membrane is a resonant structure that has different resonant properties at different points along its length, acting as a spectral analyzer. Its motion is like travelling wave being the greatest at the point where the frequency of the incoming sound matches that of the movement of the membrane as shown figure given below. In the cochlea, low frequency signals will induce oscillations that reach maximum displacement at the apex of the BM, while high frequency signals induce oscillations that reach maximum displacement at the base of BM near oval window.

The perception of a sound is related to not only its own loudness and spectrum, but also to its neighbor components which is the effect of masking phenomenon. Masking is the phenomenon in which a very low but audible sound (known as the maskee) becomes inaudible in the presence of another loud audible sound (known as the masker). It plays an important role in hearing sensation. Masking is of two types: simultaneous and non simultaneous masking. Out these till now simultaneous masking is under study. It is also known as frequency masking which involves masking between two sounds with close frequencies but their loudness is different depicted by the figure below. There are two types of masking observed in the HAS –*frequency masking* and *temporal masking*. Watermarking procedure directly exploits both frequency and temporal masking characteristics to embed an inaudible and robust watermark

Frequency masking refers to masking between frequency components in the audio signal. If two signals, which occur simultaneously, are close together in frequency, the stronger masking signal may make the weaker signal inaudible. The masking threshold of a masker depends on the frequency, sound pressure level (SPL), and tone-like or noise like characteristics of both the masker and the masked signal . It is easier for a broadband noise to mask a tonal, than for a tonal signal to mask out a broadband noise. Moreover, higher frequency signals are more easily masked. The human ear acts as a frequency analyzer and can detect sounds with frequencies which vary from 10 to 20000 Hz. The HAS can be modeled by a set of 26 band-pass filters with bandwidths that increase with increasing frequency. The 26 bands are known as the critical bands. The critical bands are defined around a center frequency in which the noise bandwidth is increased until there is a just noticeable difference in the tone at the center frequency. Thus, if a faint tone lies in the critical band of a louder tone, the faint tone will not be perceptible. Frequency masking models are readily obtained from the current generation of high-quality audio codes.

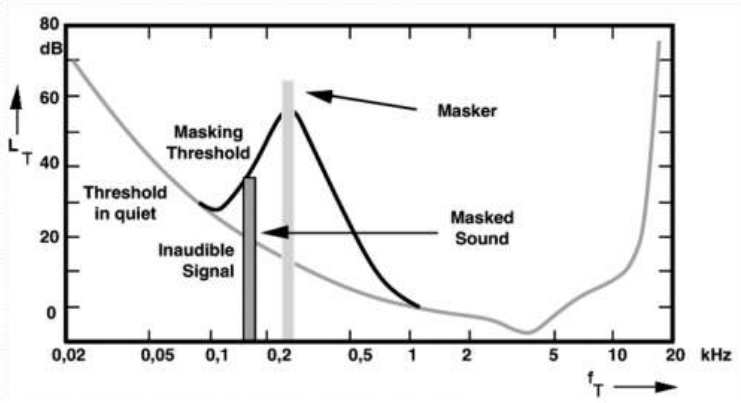


Fig. 3 Masking Effect

IV. Psychoacoustic Modeling

The implementation of the psychoacoustic model is flexible, depending on the required accuracy and the intended applications**.** The following steps describe the implementation of psychoacoustic model using ISO MPEG Audio Psychoacoustic Model 1 for Layer 1 [1]

- Calculate the power spectrum of the test signal
- Each input frame x(n) with 512 points (N) is weighted by hanning-window, h(n)[8]

$$h(n) = \frac{\sqrt{\frac{8}{3}}}{2}\left[1 - \cos\left(2\pi\frac{n}{N}\right)\right] \quad \text{.....Equation 1}$$

- So its power spectrum X(k) is computed and normalized to a reference [8]

$$X(k) = 10\log_{10}\left\{\frac{1}{N}\left[\sum_{n=0}^{N-1} x(n)h(n)\exp\left(-j2\pi\frac{nk}{N}\right)\right]^2\right\} \quad \text{.....Equation 2}$$

- Find the tone masker. Once found note the power found one index before and after and combine with the power at k to create a tone maker approximation since the tone may actually be between the frequency

samples. Determining whether a frequency component is a tone requires knowing whether it has been held constant for a period of time, as well as whether it is a sharp peak in the frequency spectrum, which indicates that it is above the ambient noise of the signal. Tonal components are special local maxima of the power spectrum. A local maximum refers to a spectral point satisfying $X(k)>X(k+1)$ and $X(k)\geq X(k-1)$. We take a local maximum as a tonal component if neighbors X(k) and X(k+j) are at greater than or equal to 7dB difference where

$j= -2,+2$ for $(2< k<63)$
$j= -3,-2,+2,+3$ for $(63< k<127)$
$j= -6,\ldots,-2,+2,\ldots,+6$ for $(127< k<250)$ ……..Equation 3

The sound pressure level of every tonal component is calculated

- Find noise-maskers and their locations within each critical band. If a signal is not a tone, it must be noise. Thus, one can take all frequency components that are not part of a tone's neighborhood and treat them like noise. Since humans have difficulty discerning signals within a critical band, the noise found within each of the bands can be combined to form one mask. Thus, the idea is to take all frequency components within a critical band that do not fit within tone neighborhoods, add them together, and place them at the geometric mean location within the critical band. Repeat this for all critical bands.
- If a masker is below the absolute threshold of hearing, it may be discarded. If two maskers are within a critical bandwidth of each other, the weaker of the two may be thrown out as well.
- Calculate the masking threshold of each mask. Sum the masking thresholds to get the overall masking threshold for all frequencies in this signal frame. The maskers which have been determined affect not only the frequencies within a critical band, but also in surrounding bands. The spreading can be described as a function that depends on the maskee location I, the masker location j, the power spectrum Ptm at j, and the difference between the masker and maskee locations in Barks (deltaz=z(i)-z(j)):

$$SF(I,j) = \begin{cases} 17deltaz - 0.4Ptm(j)+11 & -3 <= deltaz < -1 \\ (0.4Ptm(j)+6)deltaz & -1 <= deltaz < 0 \\ -17deltaz & 0 <= deltaz < 1 \\ (0.15Ptm(j)-17)deltaz - 0.15Ptm(j) & 1 <= deltaz < 8 \end{cases}$$

.....….Equation 4

There is a slight difference in the resulting mask that depends on whether the mask is a tone or noise. As a result, the masks can be modeled by the following equations, with the same variables as described above:

For tones: $Ttm(i,j) = Ptm(j) - 0.275z(j) + SF(i,j) - 6.025$ (dB SPL)……...…Equation 5
For noise: $Tnm(i,j) = Pnm(j) - 0.175z(j) + SF(i,j) - 2.025$ (dB SPL)…….……Equation 6

V. Spread Spectrum

As an effective anti-jamming technique, spread spectrum (SS) modulation has become one of the most important approaches to robust watermarking. SS-based watermarking will exhibit low embedding distortions and improved transparency when combined with a perceptual model. Spread spectrum techniques embed a narrow-band signal (the watermark) into a wide-band channel (the audio file). The process can protect watermark privacy by using a secret key to control a pseudorandom sequence generator. Pseudorandom numbers are binary numbers that have specific statistical properties such as correlations which can be used for watermark detection purpose.

Spread spectrum is a means of transmission in which the signal occupies a bandwidth in excess of the minimum necessary to send the information; the band spread is accomplished by means of a code which is independent of the data, and a synchronized reception with the code at the receiver is used for de-spreading and subsequent data recovery [9]. The process of watermark embedding can be viewed as intention jamming of the watermark signal with the music or the audio signal. In this case the signal (watermark) has much less power than the jammer (music). It is one of the problems to be overcome at the receiver end. The following chapter expresses the process of watermark generation in spread spectrum terminology. The approach selected in this algorithm is the Direct Sequence Spreading. The system modulates input the data bit-stream with the help of Pseudorandom Number sequence and modulator signal which is usually a cosine function of time with some centre frequency.

VI. Watermark Embedding and Extraction

A bit stream that represents the watermark information is used to generate a noise-like audio signal using a set of known parameters to control the spreading. These known parameters are the data bits per second *Rd*, *m* the repetition code factor, *N* is the spreading factor. $Rb = Rd*m$ is the coded bits per second, $Tb = 1/Rb$ is

the time of each coded bit, *Rc =N*Rb* is the PN sequence bits per second, *Tc=Tb/N* is the time of each PN bit or "chip".[2]

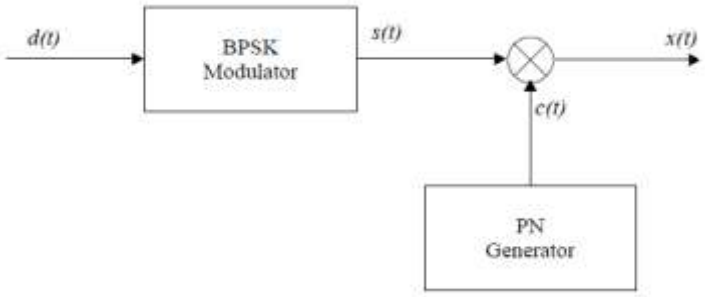


Fig 4 Watermark generation

- Select a data bit sequence some length for eg bit sequence 0110 of length 4 is is selected
- A PN sequence is also selected for spreading
- Multiplying the bit with PN sequence
- Modulating the hopped signal

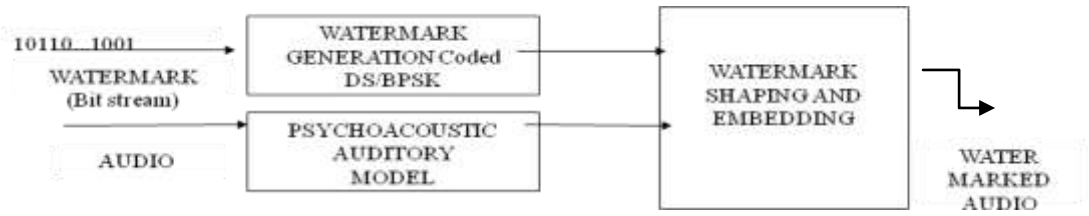


Fig 5 Watermark Embedding

At the same time, the audio (i.e. music/speech) is analyzed using a psychoacoustic auditory model. The final masking threshold information will be used to shape the watermark and embed it into the audio. The output will be a watermarked version of the original audio that can be stored.[2]

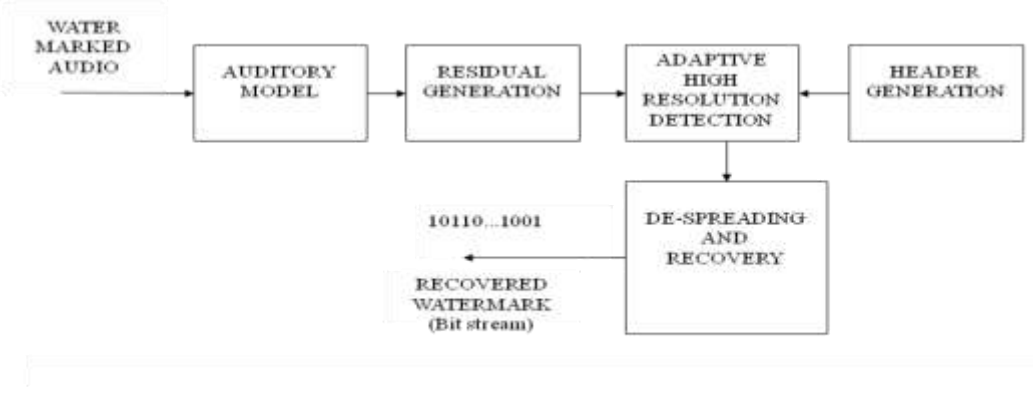


Fig 6 Watermark extraction

VII. CONCLUSION

Due to the great sensitive property of the human auditory system (HAS) compared to the human visual system (HVS), embedding watermarks into audio signal is much more challenging than inserting watermarks into image or video signals. To make the embedding process transparent and provide enjoyable high quality watermarked audio to the listeners, a psychoacoustic model is indispensable to most of the digital audio watermarking systems.

The watermarking algorithm analyzed mixes the psychoacoustic auditory model and the spread spectrum communication technique to achieve its objective. It comprised of two main steps: first, the watermark generation and embedding and second, the watermark recovery. Its robustness can be checked on the basis on

common attacks like cropping, filtering. This robustness checking is a wide area of research which can be explored further.

REFERENCES

Journal Papers:

[1] Pranab Kumar Dhar, Jong-Myon Kim, "*Digital Watermarking Scheme Based on Fast Fourier Transformation for Audio Copyright Protection",* International Journal of Security and Its Applications Vol. 5 No. 2,.pp33- 48, April, 2011

[2] Wahid Barkouti, Lotfi Salhi and Adnan Chérif, "*Digital audio watermarking using Psychoacoustic model and CDMA modulation*", Signal & Image Processing : An International Journal (SIPIJ) Vol.2, No.2, June 2011

[3] Nedeljko Cvejic, "Algorithms for audio watermarking and steganography", University of Oulu., 2004, ISBN 951-42-7383-4.

Books:

[4] Wu M & Liu B, *Multimedia Data Hiding*. Springer Verlag, 2003, New York, NY.

[5] Arnold M, Wolthusen S & Schmucker M, "Techniques and Applications of Digital Watermarking and Content Protection" 2003 Artech House, Norwood, MA.

[6] David Martin Howard, James A. S. Angus, *Acoustics and Psychoacoustics* (4th Edition, illustrated, Focal 2009).

[7] Hugo Fastl, Eberhard Zwicker, *Psychoacoustics: Facts and Models* (Edition 3, Springer 2007).

[8] Oppenheim, A.V., and Schafer, R.W., *Discrete-Time Signal Processing* (Englewood Cliffs, 1989, NJ: Prentice-Hall).

[9] Taub and Schilling, *Principles of Communication System* (Tata McGraw-Hill, New Delhi, 1995).

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 79-82
www.iosrjournals.org

A Modified ABC Algorithm & Its Application to Wireless Sensor Network Dynamic Deployment

Alina Rakhi Ajayan, Prof. S. Balaji
Dept of ECE, DMI College of Engg, Chennai, India
Dept of ECE, DMI College of Engg, Chennai, India

ABSTRACT: *Wireless Sensor Networks plays a gargantuan role in instrumenting the modern day world; hence it is crucial to develop an optimal design flow. Out of the various Swarm intelligent algorithms developed so far, Artificial Bee Colony (ABC) algorithm offers a vivid scope for this objective. First introduced for solving numerical optimization problems, ABC has now spread across almost every phase of WSN construction. In this study, an improved ABC algorithm has been proposed to match the different characteristics of wireless sensor network deployment process, which will be optimum for real time dynamic network functioning.*
KEYWORDS: *Wireless Sensor Networks, Swarm Intelligent Algorithms, scouts, foragers, onlookers, waggle dance, self organization, Opposition-based learning.*

I. INTRODUCTION

Recent advances in sensing, computing and communication technologies coupled with the need to continuously monitor physical phenomena have led to the development of Wireless Sensor Networks (WSNs). A WSN is formed by deploying sensor nodes in an application area forming an infrastructure of sensor/measurement, computation, and communication elements to instrument, observe, and react to events and phenomena in a specified environment. In most real-time deployment solutions, they demand self-organizing capabilities to keep track of the dynamic domain. The main benefits of such a configuration is the spatial diversity they provide, enabling applications such as target detection & tracking as it moves throughout the sensor field; weapon targeting and area denial Such networks are suitable for a wide variety of applications such as surveillance, precision agriculture, smart homes, automation, vehicular traffic management, habitat monitoring, disaster detection, etc. Smart disposable micro sensors can be deployed on practically any kind of environment - ground, air, under water, vehicles, living beings, buildings, & volcanoes.

Designing groups of spatially distributed autonomous sensors, connected to each other forming a network topology (ad-hoc, Star, Mesh, etc), & group wise to APs, require much precision and calculations. Most of the dynamic deployment solutions deal with optimum positioning through 1 or more metrics, starting from an initial random map. But, in such a case we might observe several demerits, like slower convergence, uncertainity in avoiding local minimums, overlap and redundancy, etc. As studied in [15], this can lead to added energy, delay and overhead expenses. Thus we resort to advanced deployment algorithms.

Bio-inspired principles have found their way into WSN R & D due to the appealing analogies between biological systems and large network. The dynamics of many biological systems and the laws governing them are based on a surprisingly small number of simple rules, yielding collaborative & effective mechanisms for resource management, task allocation, and synchronization without any central controlling element. Such Swarm Intelligent Algorithms give simple solutions to WSN Deployment problems. Artificial bee colony algorithm is one such

II. THE ARTIFICIAL BEE COLONY ALGORITHM

The Artificial Bee Colonies (ABC) is another novel optimization algorithm that comes under Swarm inspired by social behavior of natural bees. It is a complete analogy of the foraging activities of the honey bee colony, consisting of morphologically uniform individuals with different temporary specializations to represent a distributed adaptive system of smart control packets. These packets use limited computational and energy resources to explore the environment/network. They cooperate mutually & efficiently by releasing information about the discovered paths and estimated quality at the specified nodes. Due to these similarities between foraging behaviors in insect societies and network routing, in the last decade, many SI based routing protocols, specially ABC, have been developed for wired networks, satellite networks, MANETs, and, more recently, WSNs. The benefit of such organizations is increased flexibility to adapt to the changing environments.

The honey bee colonies symbolize design of optimization strategies for dynamic, time-varying, and multi-objective problems. The bee members utilize a sophisticated communication protocol with bee-to-bee signals and when required, & Stigmergic feedback cues for bee-to-group or group-to-bee interaction. The main advantage from other swarm intelligent protocols like ant colony optimization is communications are based on

visual activity & path search, unlike backtracking in many others.

Worker bees typically do brood rearing for the first week, engage in other hive maintenance duties (wax secretion, guarding, undertaking, nectar processing), when they are middle-aged switch to foraging and when about three weeks old, to colony defense. Foraging is critical task for colony survival. Foragers constantly leave the hive searching for new sources of nutrient, return with nutrient samples, and try to recruit other bees to exploit the food site by competing with other site sample bearers. Foragers announce a food source of interest to their fellow foragers by doing a dance on the dance floor inside the hive termed Waggle Dance. It is a particular figure-eight dance symbolizing the direction of the food source in terms of angular position with the sun in reference, and the distance in the duration of each waggle-run. A very short distance will be a waggle dance resembling a round dance. Foragers prefer nearer food sites over distant ones in order to increase the net energetic efficiency of the colony. Nectar foragers use Tremble Dance to show a rich food source detection, yet no free employee, upon return to the hive, after a certain threshold time, This stimulates the bees inside the hive to increase and/or to switch to nectar processing activities, and to inhibit the outside foragers from recruiting additional bees, thus keep colony's nectar processing rate matched with its nectar intake rate. Stochastic selection of food sites enables optimal choice & effective balancing between exploitation and exploration.

In ABC algorithm, colony bees are divided into three groups: employed bees, onlooker bees and scout bees. The number of employees is equal to the number of food sources i.e. number of solutions in the population.. The employee whose food source is exhausted becomes the scout bees, who also start the initial food foraging. Each food source position is equivalent to a potential solution of optimization problem, with its quality or fitness value as the base of selection, i.e. nectar amount. Whenever any bee finds the food, it signals the other bees by its dance stigmergy, the quantity and the location of the food source. This attracts a large number of bees towards good sources to search. Thus ABC algorithm model have three operational phases:

1. Scout bees do random search for the food, & find near optimal food sources, completely random.
2. Employed bees visit food source and gather information about food source location and the quality. they have memory of the places they have visited before and quality of food there, & performs the local search to try to exploit the neighboring sources to locate the best
3. Onlooker bees wait in the dance area to decide which food source is better on the basis of information provided by employees. They perform the global search for discovering the global optimum
4. The model has the following basic features to support self organization:

i) **Positive feedback:** As nectar amount of food sources increases, number of onlookers visiting them increases.
ii) **Negative feedback:** The exploitation process of poor food sources is stopped by waggle dance.
iii) **Fluctuations:** The scouts carry out a random search process for discovering new food sources.
iv) **Multiple interactions:** Bees share their information about food sources on the dance area.

III. THE PROPOSED MODEL

Generic ABC Algorithm pseudo code is given below, based random initial positioning:
1: Set swarm size S, dimensionality N, search space domain MI.
2: Create a population of S random food sources and evaluate their fitness
3: REPEAT THE FOLLOWING STEPS.
4: Send employed bees to exploit the food sources
5: For each solution i determine a neighbor k and generate a new best solution.
6: Compute the probabilities for the solutions,
7: Send onlooker bees to exploit the food sources according to probabilities.
8: For each solution i determine a neighbor k and generate a new best solution.
9: Determine abandoned solutions and send scout bees to search new food sources to replace them.
10: Update best fitness and respective food source position
11: MI = MI + 1
12: UNTIL MI = MaxMI.

As we know, this algorithm needs to be modified to fit into the optimization constraints and range. The initial positioning is made arbitrarily closer to an optimal positioning using Voronoi diagrams. The concept of 'greedy bee' helps make better selection trials. Greedy bees will always be attracted towards the flower patch having the maximum nectar or the most potential food source. This is useful in avoiding local troughs and minima.

An attractive concept that is applicable is the Opposition-based learning. It is an attempt made to generate better initial food sites for the bees, especially scouts' new locations. The search space is contracted to enable scouts to locate the food sources in a fewer number of cycles. The concept is: for a real number 'x' in the range [a, b], i.e. x Є [a, b], the opposite number "x'" of x is defined as x' = a + b – x. The bottom-line of this concept is that 50% of the computation time the current solution is further away from the optimum than its

opposite solution. Hence considering both the numbers and retaining the better of the two, chances of finding the optimum quickly is improved.

Optimization with respect to network communication paradigms also helps in improving efficiency. The wireless links, system model, etc determine the energy parameters of a WSN, which is of course the designers' primary concern. Instead of complying to minimum no. of hops, distance, graph structures etc, inclusion of minimum energy threshold, remaining energy measure, energy-distance metric, etc, keeps a full check on the throughput, and lifetime as well.

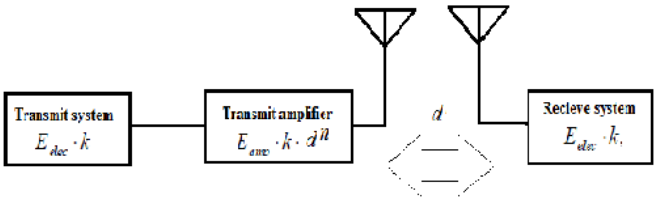


Figure 1. basic nodal transmission model

The proposed algorithm comprises of following steps:

PSUEDOCODE

1. Initialize Control Parameters of ABC Algorithm: Colony size CS, dimension of the problem D = 2, and Limit for scout L = (CS*D)/2.
2. Initialize scout bee positions as per subroutine below:

2.1. Let s be the total no. of scout bees positions.
2.2. Generate Voronoi diagram based positioning for the m nodes.
2.3. Evaluate the fitness of the positioning.

$$f_i = d(m,n) = -\left[\left[\frac{c\sqrt{P_{tx}}}{4\pi f}\right].\sqrt{[P_{rx}(m,n)]}\right].$$

For node $s_i(m,n)$, Where *m* and *n* are node position coordinates, *c*= speed of light, *Ptx* = transmitted signal power per node to reach the entire domain dimension, *Prx* = received signal strength, *f* = channel frequency

2.4. Calculate weightage probability as,

$$P_i = \left[\frac{(0.9 \times f_i)}{f_{best}}\right] + 0.1.$$

3. Set no. of iterations for optimisation, MaxIt.
4. REPEAT for i = 1 to MaxIt

4.1. Assign k employee bee nodes, best out of s scouts, as k_i.
4.2. Produce new solutions for onlooker positions based on Opposition learning, as v_{ij}, for each k_i.
4.3. Check for fitness as, use Probabilistic detection model to prevent coverage overlap.

$$c_{mn}(v_{ij}) = \begin{cases} 0, \text{if } r + r_e \le d(k_i, v_{ij}) \\ e^{\left(\frac{-\lambda_1 \alpha 1^{\beta 1}}{\alpha 2^{\beta 2} + \lambda_2}\right)}, \text{if } r - r_e \le d(k_i, v_{ij}) \le r + r_e \\ 1, \text{if } r + r_e \ge d(k_i, v_{ij}) \end{cases}$$

And

4.4. Apply greedy selection process between new nodes.

$v_{ij} = k_{best,i} + r_{ij}\ (k_{best,i} - k_i)$

4.5. Determine the probability values out of their fitness values.
4.6. Produce new solutions for v_{ij} out of this check.
4.7. Record the best solutions yet.
4.8. Check for abandonment condition according to the remaining energy model,

$$F_{ij} = \sum_i \left[\sum_j \left[d(v_{ij}, k_i)\right]^2 + \left[d(v_{ij}, k_{best,i})\right]^2\right],$$

& if satisfied replace with new solutions using Opposition based learning.

4.9. Increment i = i + 1.

4.10. UNTIL i = MaxIt.

5. Ensure deployment coverage check using

$$c_{mn}(v_{ov}) = 1 - \prod_{v_{ij}} \left(1 - c_{mn}(v_{ij})\right) \geq c_{mnth}$$

IV. CONCLUSION & FUTURE WORK

Artificial Bee Colony algorithm has opened up a vast stage for WSN protocol suite design. Like any swarm intelligent scheme, they are compatible to any stage of a WSN design and implementation, which makes them an attractive choice as the base. Though the generic version gives inferior performance in high target goals, a few adaptations to fit into our formulated problem gives colorful results. This paper makes such a proposal, which has been molded to optimize the total deployment scenario

Further goals include adapting the same algorithm with minor parameter variations for the routing scheme. Hence it can, by itself, constitute the backbone of a complete WSN network layer protocol suite. The analysis part can be done in NS2 simulator, or even Matlab.

REFERENCES

[1] Fahad S. Abu-Mouti, Mohamed E. El-Hawary, "Overview of Artificial Bee Colony (ABC) Algorithm and Its Applications", 978-1-4673-0750-5/12/ ©2012 IEEE

[2] Celal ¨OZT¨URK, Dervi¸s KARABO˘GA, Beyza G¨ORKEML˙, "Artificial bee colony algorithm for dynamic deployment of wireless sensor networks", Turk J Elec. Eng & Comp Sci., Vol.20, No.2, 2012, c_ T¨UB˙ITAK doi:10.3906/elk-1101-1030

[3] Gaige Wang 1,2, Lihong Guo 1,*, Hong Duan 3, Luo Liu 1,2 and Heqi Wang. "Dynamic Deployment of Wireless Sensor Networks by Biogeography Based Optimization Algorithm" J. Sens. Actuator Netw. 2012, 1, 86-96; doi:10.3390/jsan1020086,Journal of Sensor and Actuator NetworksISSN 2224-2708.

[4] 4 Siba K. Udgata, Samrat L. Sabat, S. Mini , "Sensor Deployment in Irregular Terrain Using Artificial Bee Colony algorithm", 978-1-4244-5612-3/09/$26.00_c 2009 IEEE

[5] Avinash Kaur, Sonu Agrawal, Department of CSE, CSVTU, Bhilai, India , " Locality Detection in Wireless Sensor Network using Population based Algorithm", International Journal of Computer Applications (0975 – 8887) Volume 51– No.4, August 2012.

[6] Zhang Dongli , Guan Xinping , Tang Yinggan, Tang Yong, "Modified artificial bee colony algorithms for numerical optimization", IEEE, 2012

[7] Dervis Karaboga1, Selcuk Okdem, and Celal Ozturk, "Cluster Based Wireless Sensor Network Routings using Artificial Bee Colony Algorithm", 978-1-4244-7107-2/10/$26.00 ©2010 IEEE

[8] Selcuk Okdem, Dervis Karaboga and Celal Ozturk, "An Application of Wireless Sensor Network Routing based on Artificial Bee Colony Algorithm",

[9] Celal Ozturk *, Dervis Karaboga and Beyza Gorkemli, " Probabilistic Dynamic Deployment of Wireless Sensor Networks by Artificial Bee Colony Algorithm", Sensors 2011, 11, 6056-6065; doi:10.3390/s110606056, ISSN 1424-8220, 3 June 2011.

[10] Tarun Kumar Sharma, Millie Pant, "Enhancing the Food Locations in an Artificial Bee Colony Algorithm ",978-1-61284-052-9/11/$26.00 ©2011 IEEE proceedings.

[11] Dervis Karaboga, Beyza Gorkemli, "A Quick Artificial Bee Colony -qABC- Algorithm for Optimization Problems", Research Fund of the Erciyes University, Project Number: FBA-11-3536, 978-1-4673-1448-0/12/$31.00 ©2012 IEEE proceedings.

[12] Bilal Babayigit, Resul Ozdemir, "A modified artificial bee colony algorithm for numerical function optimization", 978-1-4673-2713-8/12/$31.00 ©2012 IEEE proceedings.

[13] Wei-Ping Lee, Wan-Ting Cai, "A Novel Artificial Bee Colony Algorithm with Diversity Strategy", 2011 Seventh International Conference on Natural Computation, 978-1-4244-9953-3/11/$26.00 ©2011 IEEE proceedings.

[14] 15. ISSN 1943-3581, Network Protocols and Algorithms, 2010, Vol. 2, No. 2,Macro Institue, Coverage in Wireless Sensor Networks: A Survey, Raymond Mulligan, Habib M. Ammari.

[15] 16. Hosny M. Ibrahim, Nagwa M. Omar, Ali H. Ahmed, " IMPROVING WIRELESS SENSOR NETWORKS PERFORMANCE BY USING CLUSTERED VIRTUAL RINGS", International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC) Vol.3, No.3, June 2012, DOI : 10.5121/ijasuc.2012.3304.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 83-88
www.iosrjournals.org

A 3 dimensional Digital Image Skeltonization using 3x3x3 Structuring Element

P. Srinivasa Rao, Dr. M. Madhavi Latha
ECE Department, C V R College of Engineering, Hyderabad, Andhra Pradesh, India.
Professor, ECED, JNTUCEH, Hyderabad, Andhra Pradesh, India

ABSTRACT: *Image Skeletonization promises to be a powerful complexity-cutting tool for compact shape description, pattern recognition, robot vision, animation, petrography pore space fluid flow analysis, model/analysis of bone/lung/circulation, and image compression for telemedicine. The existing image thinning/skeletonization techniques using boundary erosion, distance coding, and Voronoi diagram are first overviewed to assess/compare their feasibility of extending from 2D to 3D. Previously, skeletons have been a common tool for identifying shape components in a solid object. However, obtaining skeletons of a grayscale volume poses new challenges due to the lack of a clear boundary between object and background. In this paper we propose a fast, efficient and robust algorithm to generate the skeleton of large, complex 3D images such as CT, MRI data which make use of 3X3X3 structuring elements for processing. This algorithm has been developed in the frame work of cellular logic array processing. Cellular logic array processing is a logico mathematical paradigm developed using the fundamental notions of normal algorithms and cellular automata. The algorithm provides a straightforward computation which is robust and not sensitive to noise or object boundary complexity. Because 3D skeleton may not be unique, several application-dependent skeletonization options will be explored for meeting specific quality/speed requirements.*

Keywords: *3D Image Skeleton, structuring element, cellular automata, normal algorithm.*

I. INTRODUCTION

The history of skeletonization of digital objects/ images is almost as old as digital image analysis itself. Thinning and skeletonization have numerous applications in image analysis and computer vision. The purpose is to reduce 2D discrete objects to (1D) linear representations preserving topological and geometrical information. The literature on 2D skeletonization of digital images is very rich. An outline of various thinning and skeletonization methodologies can be found in Ref. [1]. Reducing discrete structures to lower dimensions is even more desirable when dealing with (3D) volume images. The skeleton could be a promising tool for an increasing number of applications, especially in biomedical imagery. However, compared to the literature on 2D skeletonization, the articles published on 3D skeletonizations are still not very numerous, especially on 3D gray scale skeltonization. The main reason for this seems to be the difficulty to address and efficiently solve essential problems, such as topology preservation, in more than two dimensions. In fact, although many concepts such as Euler characteristics, simple points and connectivity have been studied in the past years (e.g. Refs. [2}6]), the implementation of skeletonization methods based on their use is rather complicated. The general strategy for 3D skeletonization does not differ significantly from the strategy in the 2D case. For several applications significant amount of information is lost during the process of binarization. Applying thinning directly to gray scale images is motivated by the desire of directly processing images with gray levels distributed over a range of intensity values. This will avoid shape distortions that may irremediably affect the presence of features in the binary image generated even if an optimal thresholding algorithm is used to produce the binary image. The gray skeleton is a connected subset of a gray scale pattern which consists of a network of lines and arcs centrally placed along local higher intensity regions. Unfortunately, there is no one single agreed upon definition for gray skeletons [1], [2], [8].Skeletons are classically associated as a medial axis representation that is regenerative (i.e. could be used to generate the object back exactly). Skeletons are not easily digitizable. It is not possible to have a representation that is a medial axis, preserves connectivity, preserves homotopy and exists on the square digital grid. One of these four restrictions has to be relaxed. This has opened the way to several approximations. We can group most of the published gray scale skeletonization algorithms under one of the two approaches. The first approach considers the image as a continuous surface in the 3D Euclidean space and use the first and second partial derivatives of this surface to assign the proper topographical label to each voxel [3], [15]. The second approach is based on the repeated application of a removal process that erodes the gray scale pattern until only one voxel thick object is obtained in the center of the high intensity region. The algorithm proposed in this paper is a parallel thinning algorithm that preserves connectivity and belongs to the latter family. The algorithm has been tested on a variety of images from different applications and produced satisfactory results

that proved to be useful for compression and recognition applications. Aside from this introductory section, the rest of the paper is composed of 4 more sections. In section 2 we give a quick review of previously published work in this field under the heading of literature survey. Details of the suggested algorithm are presented in section 3. Section 4 shows results of applying the algorithm to 3 different images. Out of three images , two are form standard test data set and the other is real time MRI data set. We finally conclude in section 5 with observations and recommendations for future work.

II. LITERATURE SURVEY

Thinning and skeletonization of binary images have been studied extensively since the early sixties. In the case of gray scale images, the literature does not provide a general agreement on the requirements for a gray skeleton in order to constitute a meaningful representation. The result of an algorithm is dependent on the definition of connectivity. Different definitions of local gray scale connectivity were presented. One definition, [14], states that the neighborhood of a voxel, p, is connected if the connective strength of any pair in the neighborhood of p is not less than p. A slightly different definition given in [4] considers the neighborhood of a voxel, p, to be connected if the connective strength of any pair in the neighborhood is not less than both values of the pair even if these values are less than the value of p. While in [1] the condition was imposed such that it is not less than both the pair and p. Another approach taken by [7] is to threshold the neighbors of p, then use the binary connectivity definition. In this section we will present some of the gray scale skeletonization and thinning algorithms documented in the literature almost in historical order. Levi and Montanari [8] presented a gray weighted skeleton based on the concept of gray weighted distance. The gray weighted distance is proportional to the sum of the gray levels along the path. The skeleton is the set of all points which do not belong to any minimal gray weighted path from any other point to the background. The skeleton resulting from this algorithm does not guarantee the connectivity. Dyer and Rosenfeld [4] presented a parallel algorithm. Their definition of connectivity does not provide global connectivity of the skeleton, moreover the skeleton does not lie along the high gray values but it is positioned in a central place determined by the boundary of the image. Peleg and Rosenfeld in [13] proposed a Min-Max medial axis transformation. Salari and Siy, [14], presented a two phase sequential algorithm. In the first phase they computed the contextual gray distance transformation (CGDT). In the second phase they removed the boundary voxels which satisfied a set of conditions. The algorithm required the input image to be segmented into zero and non-zero voxels. The resulting skeleton is one voxel wide positioned in the ridge areas. Abe et al., [1], point out the problems and resulting defects with the Salari and Siy algorithm, [14], due to its pure sequential processing nature and presented a combined sequential and parallel algorithm. Their algorithm is considered an extension of the Hilditch algorithm, [5]. Maragos and Ziff, [10], compute the gray skeleton by summing the skeletons of the binary images that result from thresholding the gray scale image at each value of the gray scale. Low, [9], defined gray skeleton in terms of gray mathematical morphology. Pal and Ghosh, [12], used three functions h(p), v(p) and f(p) which represent the horizontal, vertical membership functions and the degree of brightness respectively. The three functions are combined in different ways to define g(p) which denotes if the voxel belongs to the core line in the objects. Arcelli and Ramella, [2], presented a parallel thinning algorithm. They put two implementation based on the operations R1 and R2. R1 denote the sequence of four parallel operations, each operation remove voxels that satisfy certain conditions. Thinning is accomplished by repeatedly applying the sequence of the four operations to remove north, east, south and west border points respectively. The process terminates when no voxel is removed during a whole sequence. While R2 is a parallel operation that removes voxels which satisfy another set of conditions. The thinning is accomplished by repeatedly applying R2 until no further voxel is removed. They observed that the two operations R1 and R2 produce largely similar skeletons. Once difference is that the algorithm based on operation R1 is more prone to the creation of skeleton branches than the algorithm based on R2. Skeletonization algorithms based on a topographic approach were presented in [3] and [12]. These algorithms are aimed at avoiding object distortion, reducing deformation on junctions of a skeleton.

III. PROPOSED ALGORITHM

Skeletonization is the extended version of thinning. The given 3-D digital image is plane-wise raster- scanned by the seven-neighborhood window. On each move, the 3X3X3 sub image covered by this window is examined to see whether the gray-distance, say D, which is the difference between the maximum and the minimum gray value corresponding to that sub image is less than or equal to the user specified threshold value, say T. If D is less than or equal to T, then the gray-value 0 is assigned to all the cells other than the central cell and the corner cells in the given image. For D greater than T, the original values contained in these cells are left undisturbed. This procedure is continued till the entire image is scanned. The overall effect is that the boundary removed components of various 3-D solid regions in the given image that appear to be uniform, are obtained. This procedure is repeatedly applied till there is no component with boundary is detected.
The pseudo code for this algorithm is given below.

Input: 3-D image, threshold
Output: Skeletonized version of 3-D image
Steps:
Step 1: Perform image smoothing.
Step 2: Perform hollow detection and introduce a background voxel in each hollow.
Step 3: Read the 3-D data and place voxel values in a 1-D array called *input_array.*
Step 4: Copyinput_array to output_array
Step 5: Repeat sliding the 7-neighborhood window over the image(input_array) {
Step 5(a): find the maximum gray value GMax;
Step 5(b): find the minimum gray value GMin;
Step 5(c): find the difference D = GMax -GMin;
Step 5(d): if(D <= threshold)then retain the central voxel as well as corner voxels and remove the boundary voxels in output_array else slide the 7-neighborhood } Until the structuring element spans whole of the image
Step 6: Copy output_array to input_array and repeat step 3 until there is no boundary left for removal
Step 7: Pass the output_arrayto VolumeRenderer() method

3.2 Image smoothing
This step is used to remove non-significant hollows. It is accomplished by a min-max procedure. Median filters could also be used.

3.2 Hollow detection
We use the gradient operators of Sobel or Prewitt [6] to compute the partial derivatives *fx* and *fy* in the *x* and *y* directions respectively. The gradient is calculated as $g(x,y) = |fx + fy|$. The voxel p(x,y) is declared as edge point if $g(x,y)$ exceeds a threshold value *t* which is selected so that less than 5 percent of the voxels are declared as edges. Each voxel detected as an edge voxel its value is subtracted from all its 4-neighbors. The neighbor voxel which result in the maximum absolute difference is changed to zero if its gray value less than p(x,y) otherwise p(x,y) itself is changed to zero. This step is necessary to introduce a background voxel inside the hollow to start eroding from the inside out.

3.3 Boundary voxel removal
The erosion operation is an iterative procedure that removes certain border voxels, i.e. change their gray level value to zero. Rules imposed on the removal or erosion operation must guarantee that it neither destroys connectivity nor reduces the gray scale connective strength. In the proposed algorithm, connectedness is accomplished by a method similar to the method used in the binary algorithm reported in [16]. The gray connectivity strength is preserved by retaining ridge voxels. As defined above, identification of ridge voxels is based on the value of gray distance.

3.4 Threshold Value Selection
The variance of the given 3D image is chosen as the threshold. It has been known that histogram produces substantial and consistent gap between object and the background. From the histogram we will select threshold value. The histograms itself is produced by the probability (frequency) of each level of grayness.

3.5 Selection of Structuring Element
The concept of the don't care voxels makes it possible to implement the arbitrarily shaped structuring element, the point set S, with an image S with fixed size and shape (3 x3 , 5x5, 3x3x3, 5x5x5 ; . . .). The choice of structuring elements are varying in most applications using mathematical morphology, like feature extraction, edge detection, skeletonization, etc, depending heavily on the image to be processed. For example, the choice of the structuring element to extract features depends on the size of the interested features and the requirement of how accurate the boundary localization is. If the size of the structuring element is too small, noise as well as unnecessary details may be extracted. If it is too large some small features may not be extracted. At the same time shape of the structuring element is also important in geometry preserving.

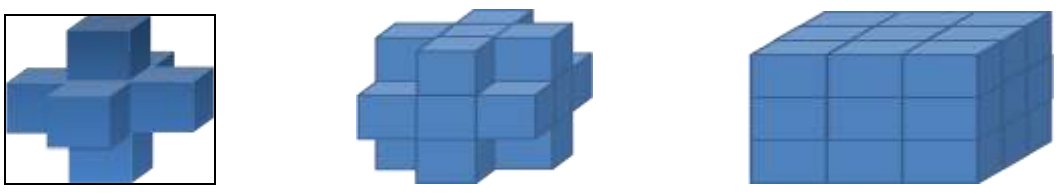


Figure (1): 3-D convex Structuring Elements

(a) 7-Neighborhood (b) 19-Neighborhood (c) 27-Neighborhood

IV. RESULTS

The algorithm has been tested on a variety of images. In this paper we report results on three test sets. The first test set is 3-D MRI image of aneurism. The second set is solid 3-D hexagonal prism. The third set is solid 3-D rectangular prism. The data set is real time MRI image where as remaining two data sets are standard test images. For the processing of these three dataset we have used seven neighborhood structuring element.

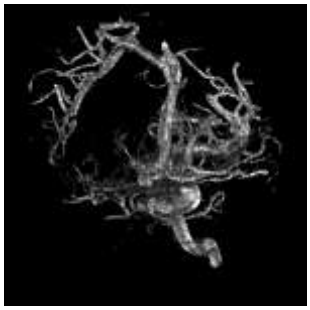
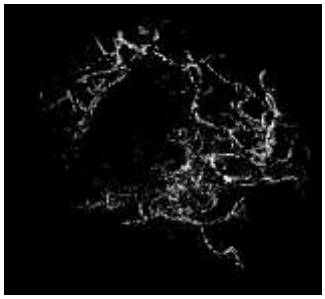



Figure2 (a) 3-D MRI image of aneurism (b) skeletonized version.

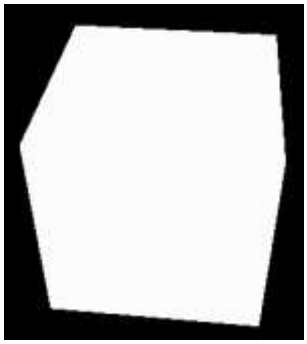
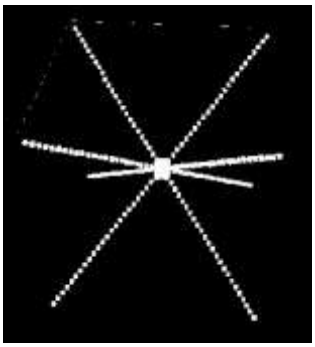



Fig 3 (a) solid 3-D Hexgonal prism (b) skeletonized version

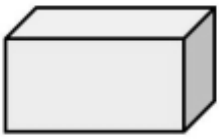
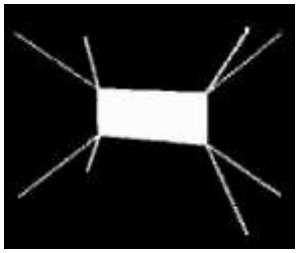



Figure 4 (a) solid 3-D rectangular prism (b) skeletonized version

V. CONCLUSION

No doubt that for several applications there are several advantages to be able to thin the gray level objects without converting the image into a binary image. In this paper we have proposed a robust parallel thinning algorithm for 3D gray scale images. The algorithm is based on eroding objects iteratively by removing certain border voxels without affecting the connectivity. The algorithm was tested on images from different domains and produced satisfactory results.

REFERENCES

[1] Abe, K., Mizutani, F. and Wang, C., "Thinning of grayscale images with combined sequential and parallel conditions for voxel removal", *IEEE Trans. on Systems Man Cybernetics*, vol. 24, no. 2, pp. 294-299, Feb. 1994.

[2] Arcelli, C. and Ramella, G., "Finding grey-skeletons by iterated voxel removal", *Image and Vision Computing*, vol.13, no. 3, pp. 159-267, Apr. 1995.

[3] Chen, S. and Shih, F., "Skeletonization for fuzzy degraded character images", *IEEE Trans. on Image Processing*, vol. 5, no. 10, pp. 1481-1485, Oct. 1996.

[4] Dyer, C. R. and Rosenfeld, A., "Thinning algorithms for grayscale pictures", *IEEE Trans. Pattern Anal and Machine Intell.*, vol. 1, no. 1, pp. 88-90, Jan. 1979.

[5] Hilditch, C., "Linear skeletons from square cupboards", *Machine Intelligence IV*, Edinburgh University Pres , pp. 403-420, 1969.

[6] Jain, A., *Fundamentals of digital image processing,*Prentice Hall, 1989.

[7] Kundu, M., Chaudhuri, B. and Dutta-Majumder, D., "A parallel greytone thinning algorithm (PGTA)", *Pattern Recognition Letters*, vol. 12, no. 8, pp. 491-494, 1991.

[8] Levi, G. and Montanari, U., "A gray-weighted skeleton", *Information and Control*, vol. 17, pp. 62-91, 1970.

[9] Low, A., *Introductory computer vision and image processing*, 1991.

[10] Maragos, P. and Ziff, R., "Threshold superposition in morphological image analysis systems", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 498-504, May 1990.

[11] Mersal, S., *Mathematical morphology techniques for gray level image processing*, Ph.D. Dissertation, Cairo University, in progress.

[12] Pal, S. and Ghosh, A., "Fuzzy geometry in image analysis", *Fuzzy Sets and Systems*, vol. 48, pp. 23-40, 1992.

[13] Peleg, S. and Rosenfeld, A., " A min max medial axis transformation", *IEEE Trans. PAMI*, vol. PAMI 3, no. 2, pp. 208-210, March 1981.

[14] Salari, E. and Siy, P. "The ridge seeking method for obtaining the skeleton of digital images", *IEEE Trans. On Systems Man and Cybernetics*, vol. SMC-14, no. 3, pp. 524-528, May/June 1984.

[15] Wang, L. and Pavlidis, T., "Direct gray-scale extraction of features for character recognition", *IEEE Trans. on Pattern Anal. And Machine Intell.*, vol. 15, no. 10, pp. 1053-1067, Oct. 1993.

[16] Zhang, Y. and Wang, P., "New parallel thinning methodology", *Intr. Journal of Pattern Recog. & Artificial Intell.*, vol. 8, no. 5, pp. 999-1011, 1994.

[17] N. Ahuja and J. Chuang. Shape representation using a generalized potential _eld model. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(2):169.176, 1997

[18] G. Bertrand. A parallel thinning algorithm for medial surfaces. *Pattern Recognition Letters*, 16(9):979.986, 1995.

[19] H. Blum. A transformation for extracting new descriptors of shape. *Models for the Perception of Speech and Visual Form*, pages 362.380, 1967

[20] A. Bonnassie, F. Peyrin, and D. Attali. Shape description of threedimensional images based on medial axis. In *ICIP*, volume 3, pages 931.934, 2001.

[21] G. Borgefors. Distance transformations in digital images. *Comput. Vision Graph. Image Process.*, 34(3):344.371, 1986.

[22] T. K. Dey and W. Zhao. Approximate medial axis as a voronoi subcomplex. In *SMA '02: Proceedings of the seventh ACM symposium on Solid modeling and applications*, pages 356.366, New York, NY, USA, 2002. ACM Press.

[23] P. Dokladal, C. Lohou, L. Perroton, and G. Bertrand. A new thinning algorithm and its application to extraction of blood vessels. *Proc. Of Biomedsim*, pages 32.37, 1999.

[24] K. Fu and J. K. Mui. A survey on image segmentation. *Pattern Recognition*, 13(1):3.16, 1981.

[25] T. Ju, M. Baker, and W. Chiu. Computing a family of skeletons of volumetric models for shape description. *Computer-Aided Design*, 39(5):352.360, 2007.

[26] T. Ju, Q.-Y. Zhou, and S.-M. Hu. Editing the topology of 3d models by sketching. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, page 42, New York, NY, USA, 2007. ACM.

[27] G. Kindlmann, X. Tricoche, and C.-F. Westin. Anisotropy creases delineate white matter structure in diffusion tensor MRI. In *Ninth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'06)*, Lecture Notes in Computer Science 4190, pages 126.133, Copenhagen, Denmark, October 2006.

[28] A. L´opez, D. Lloret, J. Serrat, and J. Villanueva. Multilocal creaseness based on the level-set extrinsic curvature. *Computer Vision and Image Understanding*, 77(2):111.144, 2000.

[29] A. L´opez, F. Lumbreras, and J. Serrat. Evaluation of methods for ridge and valley detection. *Evaluation*, 21(4):327.335, 1999.

[30] S. Mersa and A. Darwish. A new parallel thinning algorithm for gray scale images. *Proceedings of the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, pages 409.413, 1999.

[31] K. Pal´agyi and A. Kuba. A parallel 3d 12-subiteration thinning algorithm. *Graphical Models and Image Processing*, 61(4):199.221, 1999.

[32] P. Saha, B. Gomberg, and F. Wehrli. Three-dimensional digital topological characterization of cancellous bone architecture. *IJIST*, 11(1):81.90, 2000.

[33] P. K. Sahoo, S. Soltani, A. K. Wong, and Y. C. Chen. A survey of thresholding techniques. *Comput. Vision Graph. Image Process.*, 41(2):233.260, 1988.
[34] D. J. Sheehy, C. G. Armstrong, and D. J. Robinson. Shape description by medial surface construction. *IEEE Transactions on Visualization and Computer Graphics*, 2(1):62.72, 1996.
[35] S. Svensson, I. Nystrom, C. Arcelli, and G. S. di Baja. Using greylevel and distance information for medial surface representation of volume images. *icpr*, 02:20324, 2002.
[36] S. Svensson, I. Nystr¨om, and G. Sanniti di Baja. Curve skeletonization of surface-like objects in 3d images guided by voxel classi_cation. *Pattern Recognition Letters*, 23(12):1419.1426, October 2002.
[37] J. Weszka. A survey of threshold selection techniques. *Pattern Recogn*, 7:259.265, 1978.
[38] Z. Yu and C. Bajaj. A structure tensor approach for 3d image skeletonization: Applications in protein secondary structure analysis. *Image Processing, 2006 IEEE International Conference on*, pages 2513-2516, 8-11 Oct. 2006.
[39] S. Zhang, C¸ agatay Demiralp, and D. H. Laidlaw. Visualizing diffusion tensor mr images using streamtubes and streamsurfaces. *IEEE Transactions on Visualization and Computer Graphics*, 9(4):454.462, 2003.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p-ISSN: 2278-8735. Volume 4, Issue 6 (Jan. - Feb. 2013), PP 89-93
www.iosrjournals.org

Systematic Analysis on Power Consumption in Conventional Cmos, Multiplexer Based and Hybrid Adder

J.Muruga bharathi , M.Kalaiyarasi, R.Anusuyadevi, S.Kalaivani

Abstract: *There has been a rapid increase in the popularity of portable and wireless electronic devices, like laptop computers, portable video players cellular phones, which rely on embedded digital signal processors. Since the desire is to design digital systems at best performance without power sacrifices, the need for high performance and low power multipliers is inevitable. Since multiplication is one of the most critical operations in many computational systems, Multipliers are in fact complex adder arrays. In this paper we performed a comparative analysis on power consumption in three different adders, each offering different advantages and having tradeoffs in terms of circuit complexity and power consumption. The 28 Transistor full adder is the pioneer CMOS traditional adder circuit. Which is the one which consume more power when compare to other two adder, the number of transistors required for conventional CMOS adder is 28. Recently, it has been proved that the multiplexer-based multiplier outperforms the modified Booth multiplier both in speed and power dissipation by 13% to 26%, due to small internal capacitance. After analyzing the performance characteristics of conventional multiplier types, it is observed that the one designed using multiplexer-based multiplication algorithm is more advantageous, especially when the size of the multiplied numbers is small. The number of transistors required for multiplexer based adder is 16. In order to achieve optimal power savings at smaller geometry sizes, we proposed a heuristic approach known as hybrid adder models. The hybrid adder models which consume low power among all three adder and the number of transistors required is 12. As an added advantage there will be no path from one voltage level (VDD) to the other (GND).The elimination of the direct path to the ground removes the short circuit power component for the adder module. This reduces the total energy consumed in the circuit and making it an energy efficient design. The SERF adder is not only energy efficient but also area efficient due to its low transistor count.*

Key words: *Hybrid Adder, SERF Adder*

I. Introduction

Conventional CMOS, allows very efficient implementation of simple gates (e.g. NAND/NOR) having only few transistors and nodes, and a small delay due to the single inversion level. The disadvantages lie in the large PMOS transistors resulting in high input capacitances and area requirements, and the weak output driving capability caused by series transistors. The conventional CMOS adder is designed in tanner s-edit and the circuit is simulated. The power consumed by the conventional CMOS adder is high when compare to multiplexer based adder and hybrid adder because the number of transistors is high which may lead to large dynamic power dissipation. Due to this the circuit will draw more power. In multiplexer based adder parallel combination of p-mos and n-mos are used. For designing a 4:1 multiplexer 16 transistors are required. In 4:1 mux four input data port, two selection lines and a single output. Similar to conventional CMOS adder the multiplexer based adder is designed in tanner s-edit and the circuit is simulated. The power consumed by multiplexer based adder is less than conventional CMOS adder but higher than hybrid adder. Finally the hybrid adder is designed in tanner s-edit and simulated for power calculation. The power consumed by each adder circuit is tabulated. The technology-independent low-power design strategy reduces power consumption through a refined design process. An obvious method to reduce power consumption is to reduce number of transistors in a circuit. Since the dynamic power dissipation in a VLSI is usually introduced by signal transitions in the circuit, many studies have also been carried out to minimize the average power dissipation by reducing switching activities of a given logic circuit. The minimization can be achieved at technology mapping phase and logic design phase. The major logic design methods for low power include. Techniques to eliminate glitches and to reduce switching activity in normal computation. In logic transformation we can either reduce the number of transistors and glitches or to transform the logic of the circuit.

II. Adder Modules

Adders are the fundamental building blocks in all the multiplier modules. Hence employing fast and efficient full adders plays a key role in the performance of the entire system. In the following section we briefly describe the adder modules used in our design. Since addition forms the basis of many binary operations, adder circuits are of great interest in digital design. In order to fulfill the various speed, power and area requirements of implementations, a wide variety of adder circuits have been proposed in literature. As the most frequently

used block in the overall design is full-adder, now we turn our attention to design an efficient full-adder, which operates with the possible power consumption.

A. Conventional CMOS Adder:

The 28 Transistor full adder is the pioneer CMOS traditional adder circuit. The schematic of this adder is shown in Figure 2.1.This adder cell is built using equal number of N-fet and P-fet transistors. The logic for the Complimentary MOS logic was realized using the Esq.(1) and (2)

$Carry = AB+BC+AC$ (1)
$Sum = ABC+(A+B+C)\,\acute{C}$ (2)

The first 12 transistors of the circuit produce the Carry and the remaining transistors produce the Sum outputs. Therefore the delay for computing *C*out is added to the total propagation delay of the Sum output. The structure of this adder circuit is huge and thereby consumes large on-chip area.

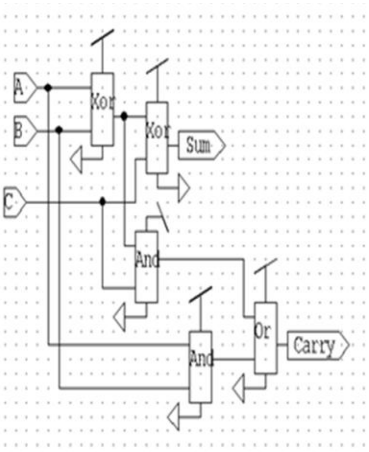


Figure 2.1 Transistor-level schematic of conventional CMOS 28-T one-bit full-adder

MY CONTRIBUTION IN THIS PAPER

B. Multiplexer Based Adder:

A multiplexer is a combinational circuit that selects binary information from one of many input lines and directs the information to a single output line. Selection of a particular input line is controlled by a set of input variables, called selection or control bits. Generally, there are 2^n input lines and n control inputs, whose bit combinations determine which input is selected. Designing an n-to-1 MUX is possible by building a tree of 2-to-1 multiplexers. Considering that the delay of 2-to-1 MUX is smaller than that of an AND gate. Since we will use 4-to-1 MUX's in this design, creating a 2-to-1 MUX tree does not bring any drawback by means of speed or power dissipation. The schematics of the cascaded block are given in Fig. 2.2

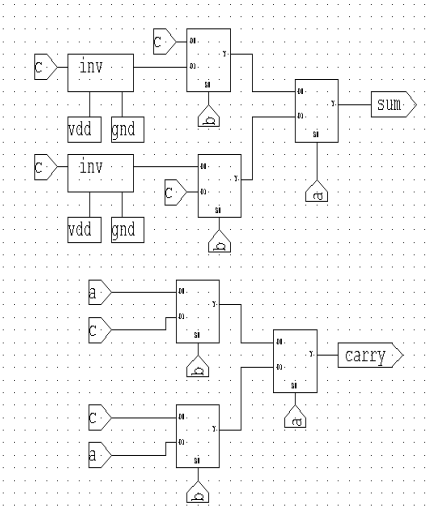


Figure 2.2 Multiplexer-based full adder

C. Hybrid Adder:

The transistor connection for a complementary switch or transmission gate consists of an n-mos and p-mos transistor connected in parallel with separate gate connection. The control signal is applied to the gate of n-mos, and its complement is applied to the gate of the p-mos device. The operation of transmission gate can be best explained by considering the characteristics of both the n-device and p-device as pass transistor individually. We will address this by treating the charging and discharging via transmission gate. n-mos pass transistor is good for transmission of '1' and poor for transmission of '0'.p-mos pass transistor is good for transmission of '0' and poor for transmission Of '1'. The circuit of the 12 pass transistor adder is shown in figure 2.3 The capacitance at the outputs of the gates is also reduced as they are not loaded with inverter. If the signal degradation at the SUM and *CARRY* is significant for deep sub-micron circuits, drivers can be used to reduce the degradation. The driver will help in generating outputs with equal rise and fall times. This results in better performance regarding speed, low power dissipation and driving capabilities. The output voltage swing will be equal to the VDD, if a driver is used at the output.

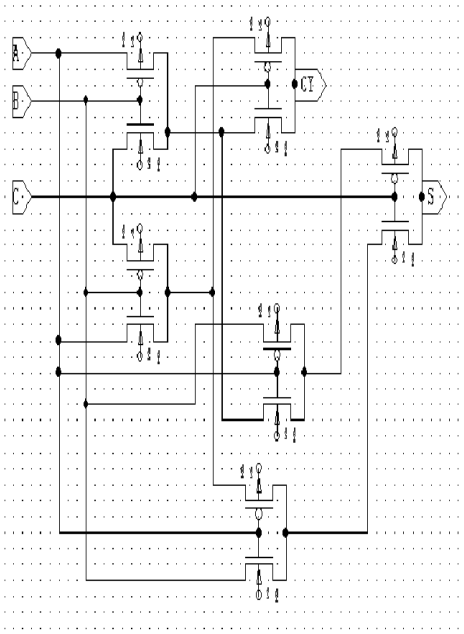


Figure 2.3 Hybrid 12 transistor adder

In this section, performance measurement of conventional CMOS 28-T one-bit full-adder, Multiplexer-based full adder and Hybrid 12 transistor adder are compared. These results were obtained from tanner s-edit simulator, the design constraints were the same for all the three adders. The results were obtained for the three circuits and the circuits were compared with respect to power consumed. For all the operand sizes, the multiplexer-based adder consumed considerably less energy compared to the conventional CMOS 28-T full-adder. The power consumed for hybrid 12 transistor adder is less as compared to multiplexer based adder and hence can be used where pass transistor logic is used.

TABLE I

ADDER	**TRANSISTORS**	**TOTAL AVERAGE POWER (mW)**
Conventional CMOS Adder	28	0.564
Multiplixer based adder	16	0.337
Hybrid adder	12	0.20

III. Simulation Results

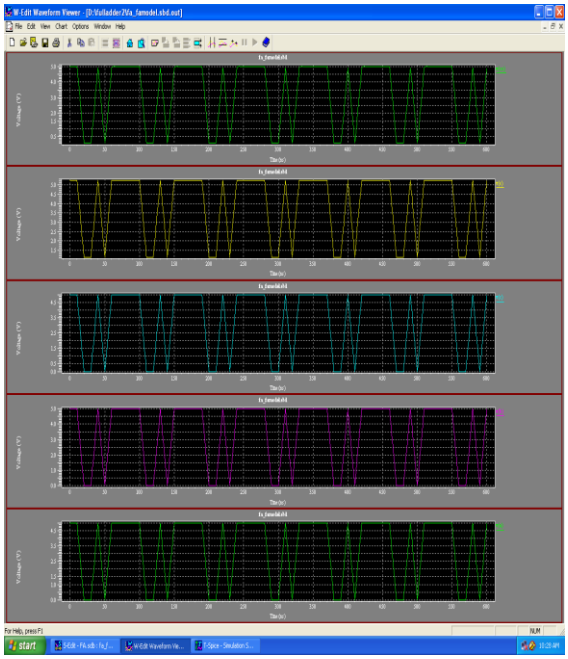


Fig 3.1 Waveform of conventional CMOS adder

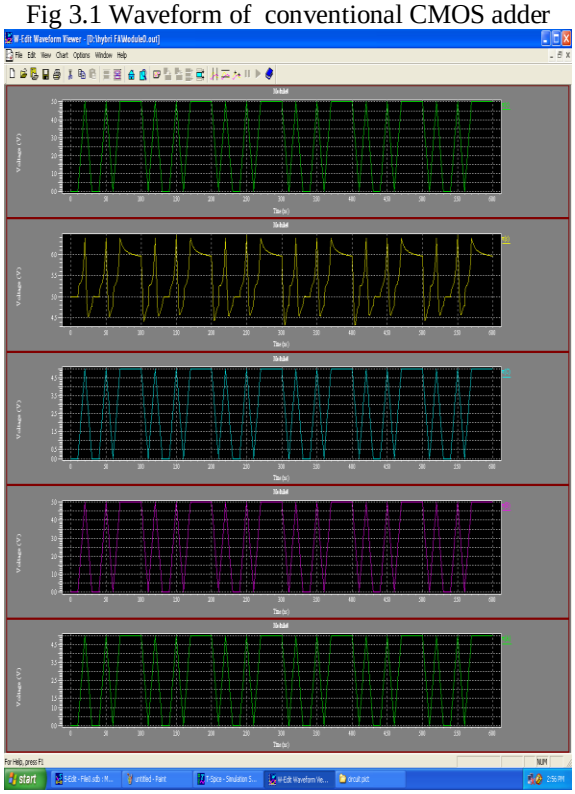


Fig 3.1 Waveform of Hybrid 12 transistor adder

IV. Conclusion

In this paper we have presented the power performance characteristics of three different adders and the transistor count is reduced. In all the adder configurations investigated, the hybrid adder exhibited better power performance compared to multiplexer based adder and conventional CMOS 28T adder based multipliers. Logic restructuring methods prove to be effective in reducing power consumption of the circuit. In the analysis conducted, only power was used in the cost function for optimization. The optimized circuit has the least power consumption of all compared circuits.

References

[1] S.Shah, A.J.Al-Khalili, and D.Al-Khalili, Comparison of 32-bit multipliers for various performance measures. Proceedingsof the 12th International Conference on Microelectronics (2000), pp.75–80.

[2] Dhireesha Kudithipudi and Eugene John Implementation of Low Power Digital Multipliers Using 10 Transistor Adder Blocks, Laboratory for Low Power Design, Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX 78249, USA

[3] Anantha P. Chandrakasan, Samuel Sheng, and Robert W. Brodersen, *Fellow, IEEE*, Low-Power CMOS Digital Design IEEE JOURNAL OF SOLID-STATE CIRCUITS. VOL 27, NO 4. APRIL 1992.

[4] A. R. Fridi, "Partial multiplication: A low-power approach for parallel multiplier," ECE729 Course Project, Dept. Electrical and Computer Eng., Univ. Waterloo, Apr. 1994.

[5] J. J. F. Cavanagh, *Computer Science Series: Digital Computer Arithmatic*. New York: McGraw-Hill, 1984.

[6] A. Bellaouar and M. I. Elmasry, *Low-Power Digital VLSI Design Circuits and Systems*. Boston: Kluwer, 1995.

[6] Y. Sasaki *et al.*, "Pass transistor based gate array architecture," in *1995 Symp. VLSI Circuits, Dig. Tech. Papers*, June 1995, pp. 123–124

[7] S. Devadas, S. Malik, "A Survey of Optimization Techniques Targeting Low Power VLSI Circuits," Annual ACM IEEE Design Automation Conference, Proceedings of the 32nd ACM/IEEE conference on Design automation, San Francisco, California, United States, pp. 242 – 247, 1995

[8] Pradhan D.K., Chatterjee M., Swarna M.V., Kunz W, "Gate-level synthesis for low-power using new transformations," Low Power Electronics and Design, 1996, International Symposium on, pp 297-300, Aug 1996

[9] Wang Q., Vrudhula S.B.K., "Multi-level logic optimization for low power using local logic transformations," Computer-Aided Design, 1996. ICCAD-96., 1996 IEEE/ACM International Conference on 10-14 Nov. 1996 Page(s):270 – 277

[10] Shih-Chieh Chang, Marek-Sadowska M, "Perturb And Simplify: Multi-level Boolean Network Optimizer," Computer-Aided Design, 1994., IEEE/ACM International Conference on November 6-10, 1994 Page(s):2 - 5

[11] R. V. Menon, S. Chennupati, N. K. Samala, D. Radhakrishnan and B. Izadi, "Power Optimized Combinational Logic Design," *Proceedings of the International Conference on Embedded Systems and Applications*, pp. 223 - 227, June 2003.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 94-101
www.iosrjournals.org

Performance Analysis of MAC Protocol for Reliable Broadcast in VANETs

Savita[1]
[1]*Electronics and Communication Engineering, Lovely Professional University, India*

Abstract : *Wireless communications are becoming the dominant form of transferring information,and the most active research field. In this dissertation, we will present one of the most applicable forms of Ad-Hoc networks; the Vehicular Ad-Hoc Networks (VANETs). VANET is the technology of building a robust Ad-Hoc network between mobile vehicles and each other, besides, between mobile vehicles and roadside units.Traffic maintenance in highway using broadcasting protocol is a new theme. It is important to find a reliable broadcasting protocol that is especially designed for an optimum performance of public-safety and data travelling related applications. Using RSU and OBU, there are four novel ideas presented in this research work, namely choosing the nearest following node as the network probe node, headway-based segmentation, non-uniform segmentation and application adaptive. The integration of these ideas results in a protocol that possesses minimum latency, minimum probability of collision in the acknowledgment messages and unique robustness at different speeds and traffic volumes.*

Keywords - *Broadcasting protocol ,DSRC , headway,MATLAB,VANET*

I. Introduction

Efforts related to traffic management in big cities led to the promising technology of building a robust wireless mobile Ad-Hoc network between vehicles (with On-Board-Units, OBUs) and roadside units (RSUs, mounted in centralized locations such as intersections, parking lots or gas stations), referred to as a Vehicular Ad-Hoc Networks (VANET). Among the main applications of VANETs, categorized as Public/Non-Public Safety (S/NS) and Vehicle-to-Vehicle/RSU (VV/VR), are co -operative collision warning (S, VV), intersection collision warning (S, VR), approaching emergency vehicle warning (S, VV) , work zone warning (S,VR), traffic management (NS, VV or VR), toll collection (NS, VR), and Internet services (NS, VR). Due to the high mobility of vehicles, the distribution of nodes within the network changes so very rapidly and unexpectedly that wireless links are established and broken down frequently and unpredictably, eliminating any usefulness of prior topology information. VANET operations in the absence of a fixed infrastructure force OBUs to organize network resources in a distributed way. So, broadcasting of messages in VANET environments plays a crucial rule in almost every application and represents a critical challenge that needs novel solutions based on the unique characteristics of VANETs. The target is to optimally develop a reliable highly distributed broadcasting protocol minimizing collisions and latency (especially in cases of public-safety related applications) without prior control messaging while considering different speeds, environments (urban and rural), and applications.

Many broadcasting algorithms have been introduced not matching the requirements of public safety applications as summarized in Sec. II. Therefore, we propose an application adaptive (multi-mode) headway-based protocol for reliable broadcasting (particular for public -safety related messages) that is robust at different speeds and traffic volumes.

II. Related Protocols

We assume that the reader is familiar with the following acronyms [2]-[7]: Ready/Clear to Send (RTS/CTS), Contention window (CW), Short Interframe Space (SIFS), Distributed Coordination Function IFS (DIFS), Network Allocation Vector (NAV), and the Hidden node problem.

Based on the IEEE 802.11 standard [3], 1- "There is no MAC-level recovery on broadcast or multicast frames. As a result, the reliability of this traffic is reduced.", 2- "The RTS/CTS mechanism cannot be used for messages with broadcast and multicast immediate destination since there are multiple recipients for the RTS, and thus potentially multiple concurrent senders of the CTS in response." Existing VANET broadcasting protocols [4]-[15] just addressed 2 points: 1-How to deliver the broadcast to nodes within a single communication range with highest possible reliability, i.e. reliable protocols? and 2- How to deliver the broadcast to the entire network, i.e. dissemination protocols?

1.1 Reliable Protocols

Reliable protocols are managed by the source node only and are used with applications related to direct

neighbors (e.g. public-safety applications). Broadcast reliability is increased through the following 3 approaches:

2.1.1 Re-broadcasting of the same message for many times:

The question is, how many times are considered practically enough? Xu [4] suggested that, re-broadcasting should be for a fixed number of times after sensing the channel as idle in each time. Yang [5] suggested re-broadcasting with a decreasing rate. Alshaer [6] proposed an adaptive algorithm where each node determines its own rebroadcast probability according to an estimate of vehicle density around it which is extracted from the periodic packets of routing management.

2.1.2 Selective ACK:

ACKing is the ultimate method of reliability, but with broadcasting we cannot let all receivers reply simultaneously. Tang [7] suggested unicasting the message to every node, one by one. Huang [8] suggested exchanging RTS/CTS with every node, then broadcasting the message once. Xie [9] proposed, on every broadcast, requesting ACK from only one receiver, on a round-robin style.

Changing transmission parameters: Balon [1] proposed decreasing collisions by changing the contention window size, based on an estimate of the current state of the network.

2.2 Dissemination Protocols

Dissemination protocols are managed by all nodes of the network, and are used with applications related to the entire network (e.g. traffic management). Here, the key design parameters are redundancy and dissemination speed. Researchers took 2 approaches to enhance the performance:

1.1.1 Flooding:

Flooding protocols are highly distributive, where it is each node's responsibility to determine whether it will re-broadcast the message or not. Ni [11] was the first to study flooding techniques in Ad-Hoc networks, and introduced the well-known "broadcast storm" problem. Then, he suggested that each node should only rebroadcast after comparing its location with the sender location and calculating the additional coverage it can provide. Heissenbüttel [12] proposed the same idea but, each node should introduces a back-off time that is shorter for greater additional areas.

1.1.2 Single relay:

We can mind single relay protocols as sequential ones, where the source node handles the responsibility of the broadcast to a next hop node. The question here is how to inform the next node of this new job. Zanella [13] proposed the Minimum Connected Dominating Set (MCDS), which is the minimum set of connected nodes that every other node in the network is one-hop connected with a node in this set. If the message was forwarded only by MCDS nodes, we would achieve the largest progress along the propagation line, while guaranteeing the coverage of all other network nodes, giving the theoretical optimal performance.

In the "Urban Multihop Broadcast Protocol (UMB)", Korkmaz [14] defined the term RTB/CTB (Ready/Clear to Broadcast), equivalent to the IEEE RTS/CTS, and suggested that the farthest node could be known by using black-burst, where its duration is longer for farther nodes. In the "The Smart Broadcasting Protocol (SB)" Fasolo [15] addressed the same idea but, using backoff time that is shorter for farther nodes. Reliable protocols care for all nodes randomly, but dissemination protocols care for the furthest node only.

III. Proposed Protocol

Giving more consideration to public-safety related applications, we propose a novel broadcasting protocol that is basically useful in emergency situations where the abnormal vehicle needs to open an instant communication channel with the vehicle(s) in the most dangerous situation.. Thus it is a case of unicast information packed in a broadcast protocol because there is not enough time for handshaking and moving to a service channel. But, it is worth emphasizing that it is still a broadcasting protocol in the sense that all surrounding vehicles within the communication range should receive and process the message while taking actions in their turn, especially if potentially probable to be affected by the danger. The question here is how to get ACK from the vehicle that is in the most dangerous situation.

In this section, we propose an application adaptive (multi-mode), headway-based protocol for reliable broadcasting (of public-safety related messages in particular) that is robust at different speeds and traffic volumes. We use the notation RTB/CTB as an equivalent to the IEEE RTS/CTS in broadcasting [3],[14]. Irrespective of the slightly increased overhead in case of short stream of data with the use of RTS/CTS, an appropriate node to reply with ACK (or CTS in case of long stream of data) is chosen.

The proposed protocol involves the following assumptions and 4 proposed concepts/approaches, namely 1- Reversing Order of Priority, 2- Headway-Based Segmentation, 3- Non-Uniform Segmentation based on naturalistic model of driver's reactions, and 4- Application Adaptive Multi-Mode schemes:

3.1 Assumptions

We assume that each vehicle involved in the protocol is at least equipped with: a high accuracy positioning device (GPS), one wireless transceiver (5.9 GHz) and a speed sensor. The broadcasted message (RTB) contains the following: source node MAC address, the coordination of the source node, current traveling speed of the source node, the message propagation direction and broadcast mode (given later).

3.2 Reversing the Order of Priority

In almost all emergency situations (e.g. co-operative collision warning), the most threatened vehicle is the nearest one running behind the source vehicle. Hence, the first proposed approach is reversing the order of priority as shown in Fig.1. With this step, the protocol chooses the nearest node with a plain uniform distance-based segmentation algorithm. Though during communication between the source vehicle and the nearest following one, there could be collisions at far range nodes due to the hidden terminal problem, this choice gives the protocol an incomparable minimum latency. As a compensation for this type of collision, we recommend that, the ACK message should be the same as the broadcast one. Hence, we include an "ACK" field in the broadcast; which should be set in the ACK message.

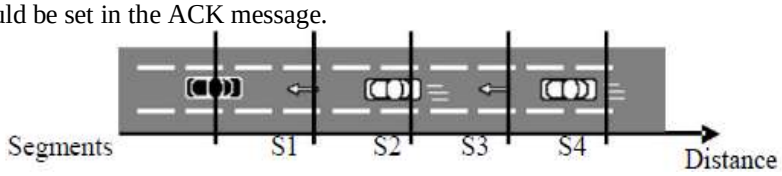


Fig. 1. Proper segmentation

3.3 Headway Based Segmentation

If vehicles are running at different speeds, distance-based segmentation logically fails in simulating the dangerous situation. Hence, the 2nd proposed approach is to include the effect of speed through time headway based segmentation to assign segment numbers. The time headway or headway for short: is the time interval between two vehicles passing a point. Fig.2 shows a 3-lane highway with three following vehicles running at different speeds, (30,60,120 Km/h) with reference to distance (meter). Fig.3 shows the same situation after calculating the headway for each vehicle to produce an imaginary calculated image. This image reveals that headway-based segmentation mimics dangerous situations better than distance -based one, as it puts the 120Km/h-vehicle in the 1st priority, consistent with the intuitive analysis of the situation. So, the algorithm elects the nearest vehicle (in time) by a plain uniform headway-based segmentation method.

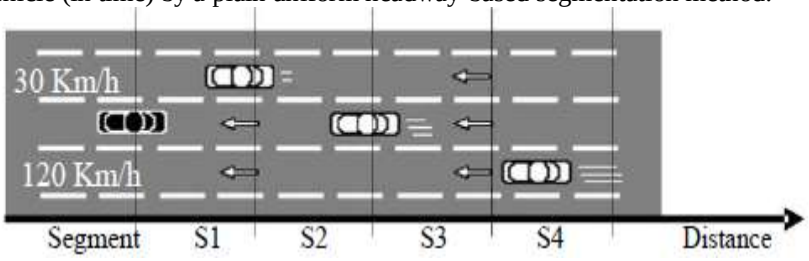


Fig.2. Distance-based segmentation

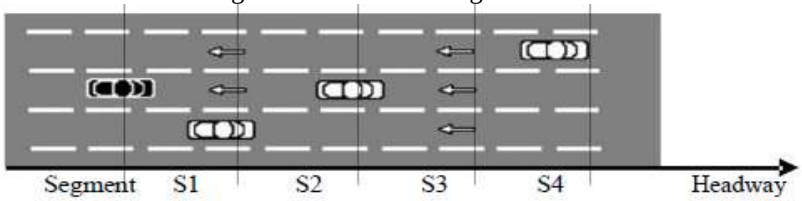


Fig.3. Time-based segmentation

3.4 Non-Uniform Segmentation (Headway Model)

We propose to let the width of each segment to be chosen according to the expected headway that drivers tend to leave apart. We adopt Semi-Poisson distribution headway model describing the average naturalistic headway that drivers tend to leave apart [16] as a basis for a non-uniform segmentation. Without loss of generality, assume only 2 vehicles in the transmission range of the source node. The headway between

the source vehicle and the first one is X_1 sec, and the headway between those two vehicles is X_2 sec. Both X_1 and X_2 are random variables with a Semi-Poisson probability distribution function. We also assume that the highway is only one lane and both CW_{min} and CW_{max} equals to one, i.e. there is no contention or random backoff. For studying the collision probability in one of the segments, we assume that the segment is in-between any arbitrary headways l_i and l_f sec. There will be a collision in the CTB message if there are more than one node in this segment. The probabilities of collision (P_C), successful broadcast (P_b), i.e. only one node in the segment, idle (P_i), and prior nodes captured the broadcast phase (P_o) are given as follows (with discretization):

$$:Pc = \sum_{x=li}^{lf} \frac{P(X=x)}{P(li<X<lf)} \times P(X < lf - x) \quad (1)$$

$$:Pb = \sum_{x=li}^{lf} \frac{P(X=x)}{P(li<X<lf)} \times P(X > lf - x) \quad (2)$$

$$:Pi = P(X > lf) \quad (3)$$

$$:Po = P(X < li) \quad (4)$$

The objective of the non-uniform segmentation is to find the best points of segmentation within a single communication range (10 sec) that results in linearly increasing P_C's with a minimum slope. There are two reasons behind minimizing the slope instead of the absolute minimum: 1- Intuitively, vehicles in 1st segments are more threatened than those in the last segments. Each vehicle is exposed to a danger that is inversely proportional to the time before collision, i.e. the headway time. 2- The other reason is a traffic concept that if there are no vehicles in the first segment, we can expect that the traffic is moderate or low, and let later segments be of a wider width.

3.5 Application Adaptive (Multi-mode) Scheme

Although the majority of VANET applications require message broadcasting, each application has its unique flavor and needs a special treatment. The main difference is which of the following vehicles should have the highest priority to respond first, either replying to the source vehicle or replaying to the following vehicles. Without loss of generality, we propose only 4 modes covering major applications:

3.5.1 Mode 0- Basic Broadcasting

The zero mode is the original basic mode, where broadcasting is omni-directional with no intended vehicle nor acknowledgment. This mode is still useful in VANET environment especially in case of the 'status message', where, as recommended by DSRC [17], every vehicle should broadcast its position, speed, direction of travel, and acceleration every 300 ms, and this transmission is intended for all vehicles within 10-sec travel time.

3.5.2 Mode 1- Furthest Following Node: The intended vehicle in this mode is the physically furthest one following the transmitting node. This mode is suitable to work as a dissemination protocol for applications like "Traffic Information", and "Work Zone Warning". So, we recommend the regular distance-based protocols (e.g. The Smart Broadcasting Protocol [15]) to be used in this mode.

3.5.3 Mode 2 - Nearest Following Node (in time): The intended vehicle is the nearest one (in time) running behind the source vehicle. This mode is suitable to work as a reliable protocol for all public-safety related applications like "Cooperative Collision Warning" and "Stop Light Assistant". Our non-uniform headway-based protocol is superior in this mode.

3.5.4 Mode 3 - Furthest Leading Node: The intended vehicle is the furthest one leading the source vehicle as in Fig.4. This mode is suitable for emergency applications like "Approaching Emergency Vehicle" either it was an ambulance or a police car. In this case, the headway is identical to distance because the speed is constant (headway is measure with reference to source node speed). However, with headway-based protocols, we can implement a non-uniform segmentation based on headway studies.

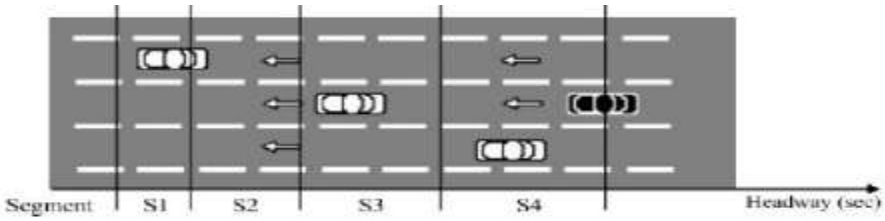


Fig. 4 Priority arrangement of mode 3.

3.6 Proposed Protocol

3.6.1 Procedure of the source vehicle:

In case of an OBU has a message to broadcast, the MAC layer of the system has to proceed with the following (Fig. 5):

1- It sends an RTB message including its MAC address, current location, mode of operation, …. etc..

2- It then waits for a valid CTB message within SIFS+N+1 time-slots (assuming N segments). If locked with a CTB, then send the unencrypted broadcast with the intended receiver as that indicated in the CTB message. Otherwise (if not), repeat from Step-1 (as long as the application requires).

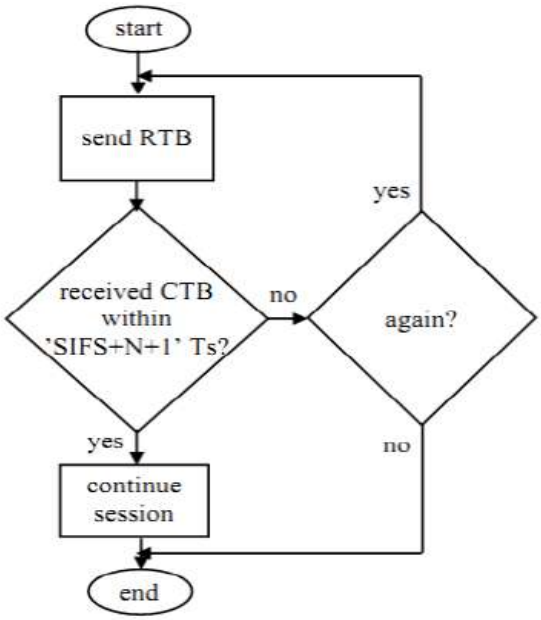


Fig.5. Actions of the source node.

3.6.2 Procedure of other vehicles:

Receiving of an RTB, other nodes proceed as follows (Fig. 6):

1- Set the NAV to be SIFS+N+2 time-slots so that no node will start a new session until the end of the current broadcast.

2- Check the broadcasting mode field.

3- Compare the geographical coordinates of the source node with their own and get its relative position. If the node is in the opposite driving direction or not in the message propagation direction, ignore it and go to end. However, if the node is in the message propagation direction, proceed to Step 4.

4 - Compute the headway in seconds (or distance in meter for mode 2), then determine its segment number.

5- If the segment number equals to Si where (i <= N) assuming 'i' is the cell number, set the back-off counter = i-1. Then, the node should wait for CTB message, if locked with a valid CTB then exit contention phase and listen for the coming broadcast. The node reaching 0 initiates the CTB including its MAC address and continues the session with the source node.

It should be noticed that in case of a lost source packet, the source sends again as long as the application requires.

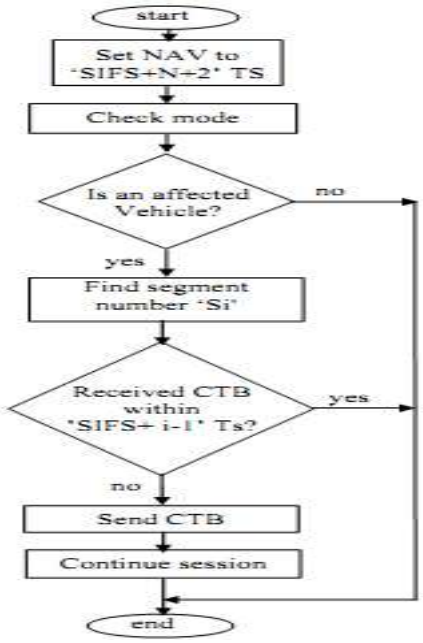


Fig.6. Actions of other nodes.

IV. Simulations and Results

Table 1 summarizes the assumptions taken during simulation, taken from the 802.11p [18] standard. Using these random variables, we conducted a simulation program for estimating the probability of collisions and the average latency within each segment. The latency is computed as typical based on [17]: contention starting time, success broadcasting time, collision time, and wait time, taking into account the MAC delay based on IEEE P802.11-REVma/D7.0 [18]. A 600 and 1600 veh/h traffic volume is considered in the headway-model.

Table 1. Simulation assumptions

Time-Slot	16 µs	CTB	14 bytes
SIFS	32 µs	Messages	512 bytes
DIFS	64 µs	ACK	512 bytes
RTB	20 bytes	Data rate	3 Mbps

Using these random variables, a simulation program was conducted for estimating the probability of collisions and the average latency within each segment of the communication range (10 sec). The width of each segment is taken according to Table 1. The probability of collision is shown in Figure 7, while the average latency is shown in Fig. 8.

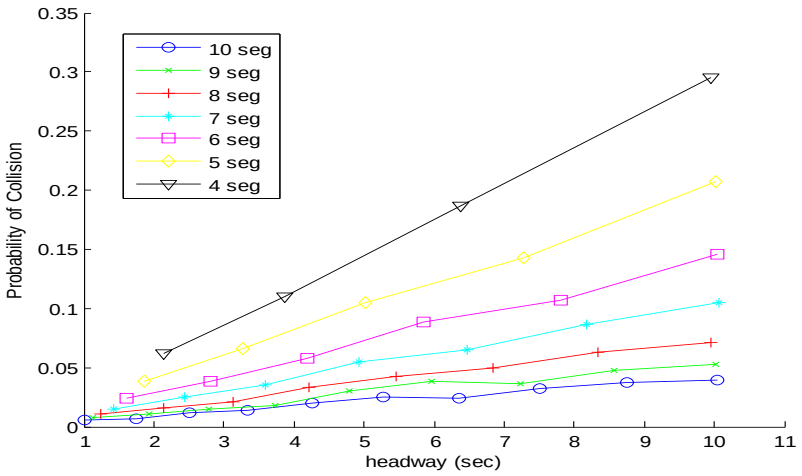


Fig. 7: Simulated calculation for *PC* for best segmentation for 600 veh/h.

In Fig. 8, the curves are close to each other. The average latency associated with each segment reveals that the case of 7-seg gives the minimum latency (best performance) before over-segmentation begins to take place with 4 segments.

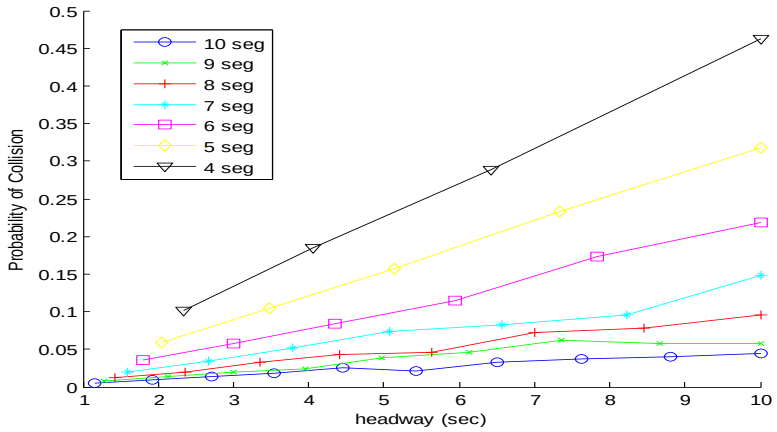


Fig. 8: Simulated calculation of latency at best segmentation for 600 veh/h.

In Fig. 9, we have seen that for 1600 vehicles/h, the probability of collision is higher than the 500 vehicles/h. The latency of the segmentation for 1600 vehicles/h has given in Fig. 10.

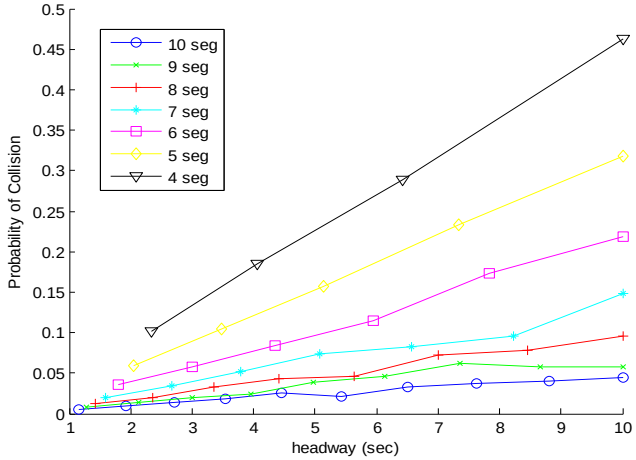


Fig. 9: Simulated calculation for *PC* for best segmentation for 1600 veh/h.

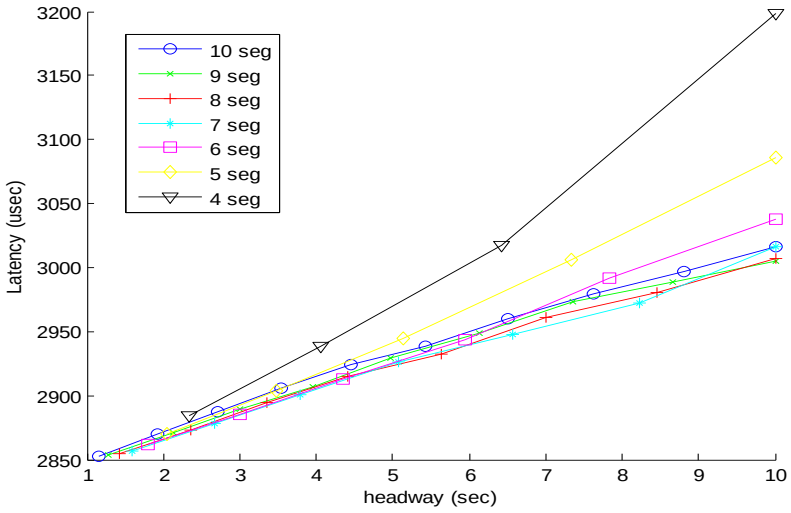


Fig. 10: Simulated calculation of latency at best segmentation for 1600 veh/h.

V. Conclusion

In this research, we introduced a novel broadcasting protocol in VANET environments with these new distinguishing features:

- The first protocol to use the concept of headway-based segmentation and to include effects of human behaviors in its design with the headway model.
- Non-uniform segmentation achieving a unique a minimum slope linearly increasing latency distribution.
- Unique robustness at different speeds and traffic volumes rooted to the headway robustness at different traffic volume variations. Superior minimum latency for public safety applications. Application adaptability with special multi-mode operations.
- Considered offering a solution to applications never discussed in literature, like "Approaching Emergency Vehicle".

Further analysis and simulation will be conducted to accommodate more complicated highway situations.

References

[1] "USA National Transportation Statistics 2007," Bureau of Transportation Statistics, USA, 2007.

[2] Y. M. Y. Hasan, M. M. I. Taha, S. Alshariff, and A. Alakshar, "Integrated intra-vehicle – VANET system for increasing the road safety," in Proc. Global Knowledge Forum NOOR-2008, Almadina, KSA, June 2008.

[3] IEEE P802.11p, "Amendment 3: wireless access in vehicular environments (WAVE)," Draft D0.26, January 2006.

[4] Q. Xu, T. Mak, J. Ko, and R.Sengupta, "Vehicle-to-vehicle safety messaging in DSRC," in Proc. of the 1st ACM Int. Workshop on Vehicular Ad Hoc Networks VANET'04, NY, USA, pp.19-28, 2004.

[5] X. Yang, L. Liu, N. H. Vaidya, and F. Zhao, "A vehicle-to-vehicle communication protocol for cooperative collision warning," in Proc. of the 1st Int. Conf. on Networking and Services, pp.114-123, 2004.

[6] H. Alshaer and E. Horlait, "An optimized adaptive broadcast scheme for inter-vehicle communication." in Proc. IEEE 61st Int. Vehicular Technology Conf. VTC'05, vol.5, pp.2840–2844, 2005.

[7] K. Tang and M. Gerla, "MAC reliable broadcast in ad hoc networks," in Proc. IEEE Military Communications Conference Communications for Network-Centric Operations: Creating the Information Force, vol.2, pp.1008-1013, vol.2, 2001.

[8] L. Huang, A. Arora, and T.H. Lai, "Reliable MAC layer multicast in IEEE 802.11 wireless networks," in Proc. of the IEEE Int. Conf. on Parallel Processing ICPP'02, Washington DC, USA, 2002.

[9] J. Xie, A. Das, S. Nandi, and A. K.Gupta, "Improving the reliability of IEEE 802.11 broadcast scheme for multicasting in mobile ad hoc networks," vol.1, pp.126-131, vol.1. 2005.

[10] N. Balon, and J. Guo, "Increasing broadcast reliability in vehicular ad hoc networks," in Proc. of the 3rd ACM Int. Workshop VANET'06, New York, USA, pp.104-105, 2006.

[11] S. Y. Ni, Y. C. Tseng, Y. S. Chen, and J. P. Sheu, "The broadcast storm problem in a mobile ad hoc network," in Proc. of the 5th ACM/IEEE int. conf. on Mobile computing and networking MobiCom'99, NY, USA, pp.151-162, 1999.

[12] M. Heissenbüttel, T. Braun, M. Wälchli, and T. Bernoulli, "Optimized stateless broadcasting in wireless multi-hop networks," in Proc. IEEE Infocom'06, Barcelona, April 2006.

[13] A. Zanella, G. Pierobon, and S. Merlin, "On the limiting performance of broadcast algorithms over unidimensional ad-hoc radio networks," in Proc. of WPMC'04, Abano Terme, Padova, Sep. 2004.

[14] G. Korkmaz, E. Ekici, F. Özgüner, and U. Özgüner, "Urban multi-hop broadcast protocol for inter-vehicle communication systems," in Proc. of the 1st ACM Int. Workshop on Vehicular Ad Hoc Networks VANET'04, NY, USA, pp. 76-85, 2004.

[15] E. Fasolo, A Zanella, and M. Zorzi, "An effective broadcast scheme for alert message propagation in vehicular ad hoc networks," in Proc. of IEEE Int. Conf. on Communications ICC'06, vol.9, pp.3960-3965, 2006.

[16] T. Luttinen, "Statistical Analysis of Vehicle Time Headways," Teknillinen korkeakoulu, pp.155-172, 1996.

[17] "Dedicated Short Range Communications (DSRC) Home," http://www.leearmstrong.com/DSRC/DSRCHomeset.htm

[18] IEEE P802.11-REVmaTM/D7.0, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications", Rev. of 802.11-1999, June 2006.

[19] M. M. I. Taha and Y. M. Y. Hasan, "VANET-DSRC Protocol for Reliable Broadcasting of Life Safety Messages," in Proc. of the 7th IEEE Int. Symp. on Signal Processing & Information Technology - ISSPIT'07, Cairo, Egypt, pp.105-110, Dec. 2007.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 102-108
www.iosrjournals.org

Real Time Implementation of Nodes Self Scheduling Approach for Maximizing Wireless Sensor Network Lifetime

E. Abarna Ignatius[1], Mrs.F.Vincy Lloyd [2]
[1](ECE, D.M.I College of Engineering / Anna University, India)
[2](Associate Prof-ECE, D.M.I College of Engineering/ Anna University, India)

ABSTRACT: *To face the short lifetime of the WSN, the objective is to optimize energy consumption while maintaining the full sensing coverage. A major technique to save the energy is to use a wake-up scheduling protocol through which some nodes stay active whereas the others enter sleep state so as to conserve their energy. This technique presents an original algorithm for node self-scheduling to decide which ones have to switch to the sleep state. The novelty is to take into account the best shortest path among the nodes in wireless sensor network .Hence, the sensed data from the network is transmitted over the best shortest path by verifying the updates in the routing table. To verify and evaluate the proposed algorithm, simulations have been conducted and have shown that it can contribute to extend the network lifetime. A comparison with existing works is also presented and the performance gains are highlighted.*

Keywords: *wireless sensor networks, simultaneous wake up scheduling, best shortest path.*

I. INTRODUCTION

Wireless sensor networks (WSNs) consist of tiny devices, which have a battery, a sensor, a microprocessor and a radio transmitter component. The application area of WSNs can be classified into two general classes: monitoring applications and tracking applications .While the rest class of applications includes habitat monitoring, building monitoring, machinery monitoring and greenhouse monitoring etc., the second class includes animal tracking, vehicle tracking and goods tracking in supply chains etc. Due to this large range of application area, performance metrics in sensor network are strictly application specific. However, 'unattended operation of the network for long time' or 'long network lifetime' can be determined as a common performance requirement for the most of the applications. In general, network lifetime can be defined as time span until the network is considered non functional. In fact, as in performance metrics, perception of no functionality is also application-specific for sensor networks.

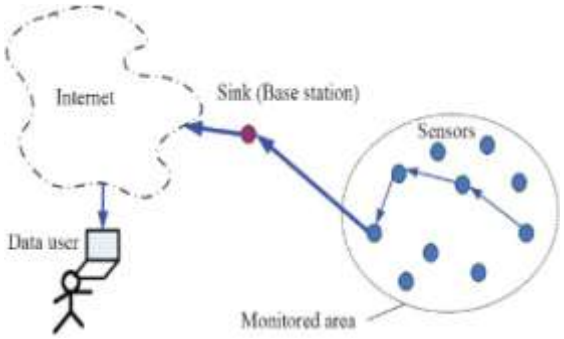


Fig 1: A Basic Wireless Sensor Network

A sensor network can be considered as non-functional if data delivery delay exceeds a threshold, or if the coverage of the monitoring area is less than the desired level, or if the network is partitioned due to the energy deficiency of some bottleneck nodes. Energy depletion of nodes can interrupt communication and, in a worse case, it could cause network partitioning which lead the interruption of monitoring. Unlike ad-hoc networks, recharging or replacing the sensors battery may be inconvenient, or even impossible in some monitoring environments. Therefore the key challenge in the design of wireless sensor network protocols is how to maximize the network lifetime, which is limited by battery energy in sensor nodes, while providing the application requirement. The data aggregation regime in the network (raw data transmission or processed data transmission, periodic or event triggered data gathering [4]); the channel access (MAC) schemes [1]; the routing protocols [3], [2]; and the energy model for transmission (based on the channel model). This paper focuses on maximizing the lifetime of a sensor network from the routing protocols point of view.

In wireless sensor networks, all nodes share common sensing tasks. This implies that not all sensors are required to perform the sensing task during the whole system lifetime. Turning off some Nodes does not affect the overall system function as long as there are enough working nodes to assure it. Therefore, if we can schedule sensors to work alternatively, the system lifetime can be prolonged correspondingly; i.e. the system lifetime is prolonged by exploiting redundancy. In this work, we present a novel node scheduling scheme, which is used to configure node work status and schedule the sensor on-duty time in large sensor networks. Our design has been driven by the following requirements: first, because it is inconvenient or impossible to manually configure sensors after they have been deployed in hostile or remote working environments, self-configuration is mandated. Second, the design has to be fully distributed and localized, because a centralized algorithm needs global synchronization overhead and is not scalable to large-populated networks. Third, the algorithm should allow as many nodes as possible to be turned off in most of the time. At the same time, it should preserve the initial sensing coverage with minimum "sensing hole", or "blind points". It is ideal if the working nodes can cover the same monitored area as the original one. Fourth, the scheduling scheme should be able to maintain the system reliability, i.e., certain redundancy is still needed. In the proposed approach, each node in the network autonomously and periodically makes decisions on whether to turn on or turn off itself only using local neighbour information. To preserve sensing coverage, a node decides to turn it off when it discovers that its neighbours (sponsors) can help it to monitor its whole working area. To avoid blind point, which may appear when two neighbouring nodes expect each other's sponsoring, a back off-based scheme is introduced to let each node delay its
decision with a random period of time.

In this work, we implement the proposed scheme to transmit the data over the sensor network by using the best shortest path by referring the routing table. During the transmission through the shortest path all the nodes are made to be in idle state except the path which transmits the data is placed in active state.

II. COVERAGE PROBLEM

2.1 Introduction

Before introducing the proposed algorithm, the coverage problem and some related works are presented. The coverage idea was introduced in [2] as a QoS. It is introduced using the fundamental question: 'How well do the nodes observe the physical space?'. Owing to the large variety of the nodes and their applications, the coverage problem is subject to a wide range of interpretations [3]. The coverage problem may be divided into three categories depending on what exactly that you are attempting to monitor. 'Area coverage' [4–6]: the overall goal is to have each location of the interest area within the sensing range of at least one node. 'Target coverage' [7, 8]: observes a fixed number of targets. 'Barrier coverage' [9, 10]: refers to the detection of movement across a barrier of sensors. This paper considers the area coverage problem. Thus, some related works are detailed in the remaining of this section to better understand the development of the proposed algorithm.

Since the sleeping state is the least energy consuming state, keeping nodes in the sleeping state is a good way to save energy. However, the full area coverage must be ensured when some nodes are sleeping.

Some work as those in [11, 12] propose to divide the nodes into disjoint sets, such that every set can individually and successively perform the area monitoring task. These sets are successively activated, and all nodes, not belonging to the active set, will be in the sleep state. Generally, such algorithms are centralized based on a full knowledge of the network topology, which increases the cost of the algorithm. Such solution is not reliable because of the failure of nodes. Indeed, when a node fails, the coverage is no more guaranteed despite the fact that the other nodes belonging to the same set remain working.

As centralized solutions are not adequate to WSNs, specially those including large numbers of nodes, because of their expensive energy cost, distributed and particularly localized algorithms were more privileged.

The algorithm presented in [13] is localized and divided into rounds. Each round begins with a self-scheduling phase followed by a sensing phase. In the self-scheduling phase, nodes investigate the off-duty eligibility rule. Based on the eligibility rule, each node verifies if its sensing area (the set of points that it reaches for sensing) is covered by the union of the sensing areas of its neighbors. Eligible nodes turn off their communication and sensing units, whereas other nodes will perform the monitoring tasks in the sensing phase. To avoid the occurrence of blind points (non-covered points), which can occur when the eligibility rule is simultaneously investigated by neighboring nodes, a back off method is used. Thus, a node that has a longer back-off time will not consider those that have already decided to be inactive (from which it has received a status advertisement message).

A localized probing-based self-scheduling algorithm is proposed in [14]. A node goes back to the sleep state if it receives at least one reply to its probation message (PRB) message otherwise it remains active. The PRB message is broadcast within a probing range r depending on the application's specificity. The problem is that such an algorithm does not guarantee the coverage of the target area.

2.1 Connectivity Problem

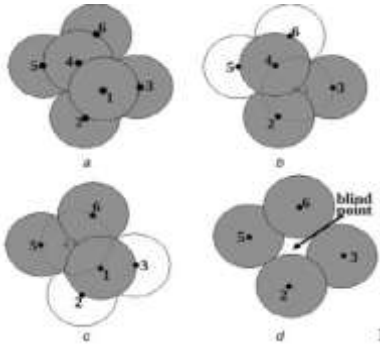
In addition to the energy preservation, connectivity is also one of the important issue related to the coverage problem. An important frequently addressed objective is to determine a minimal number of working nodes required to maintain the initial coverage and as this set is connected. An important result, proved by Zhang and Hou [6], states that if the communication range is at least twice the sensing range, a complete coverage of the area implies the connectivity of the working nodes. The authors also observed that an area is completely covered if there are at least two disks that intersect and all crossings are covered. Disk refers to a sensing area of a node and a crossing is an intersection point of the circle boundaries of two disks. Based on these results [6], the authors proposed a localized algorithm for node scheduling. A probabilistic coverage with various sensing models has also been studied in [16–18].

The development of the proposed algorithm is justified by the desire to circumvent the different challenges previously presented, as there are no algorithms, according to our knowledge, which considers all these elements when scheduling the node's activities while preserving the area coverage. Particularly, the energy constraint, which has a major impact on the WSN lifetime, must be taken into account at organizational and decisional levels. On the one hand, the best path should be made active while making all other nodes into idle state . On the other hand, decisions must be made based on local information and the exchange of many messages must be avoided, thus, reducing the overheads and as a consequence, the energy cost of the algorithm.

III. PROPOSED DESIGN

3.1 Problem Statement

In an sleep-scheduling wireless sensor network, sensors are competing only with manual sleep nodes with the usage of limited active nodes leading to coverage and connectivity problems. In most sensor networks, multiple sensors are deployed to monitor the desired area. In addition to make nodes sleep manually it is efficient to make a simultaneous wakeup scheduling approach such that only the nodes transmitting the data is made to be active were the rest of the nodes are in idle state .For transmission of next data the next best path is chosen through the routing table.



a Original sensing area covered by nodes 1–6
b Node 1 turns off itself by the eligibility rules
c Node 4 turns off itself by the eligibility rules
d Occurence of a blind point

Fig 2: Blindpoints

If multiple nodes pick the same path to transmit at the same time and it leads to network congestion, collision at receiver end; which ends up in lots of energy wastage. There is possible for the blind points to occur.

3.2 Problem Formulation

The wireless sensor network is modelled as a Nodes, and L is the set of the links between them. If two nodes n and m are connected by a directed link, they can communicate with each other. Direction of the link is determined from farther node to the closer one to the sink node. For the connection of node n with any node m, a variable is defined to represent the total flow on an outgoing links of n as fl_{nm} and this variable also represents the total flow on an incoming link of m.

For the sake of notation, let variables f1 and f2 represent the total amount of flow (fl) during the network lifetime on the outgoing (On) and incoming (In) links of a node n, respectively, and depicted as follows:

f1 = $\sum_{lnm \in On} flnm$ n, m ∈ N,L ∈ L□ ------(1)
f2== $\sum_{lnm \in In} flmn$ n, m ∈ $N, M \in L$ ------(2)

The lifetime, T_n, of the individual node n is formulated as being the sum of the time spent in transmission, reception, and idle mode of operation, and it is depicted as follow:

T_n= f1.t_{tx} + f2 . t_{rx} + t_{idle}^{n} n ,m ∈ $N, L \in L$ -----(3)

In the above equation, t_{tx} and t_{rx} are constant and represent the transmission time and the reception time of a data packet respectively, hence f1. t_{tx} Represent the total transmission time and f2. t_{rx} represent the total reception time for a node n at the end of its lifetime. Note that t_{tx} and t_{rx} will differ based on the data packet size and speed of the transceiver. t_{idle}^{n} Is a variable and represents the total idle time of the node spent in its whole lifetime.

IV. SYSTEM DESIGN

Simultaneous wakeup scheduling of nodes in transmission of the data to the destination by using the best suitable path is the main focus in this work. Concurrent usage of the same path by multiple sensor nodes would lead to collision and retransmission of data packets resulting in degradation of the network efficiency and waste of energy in sensor nodes. AODV is a protocol by which a routing table is created which has the updates of all the neighbour nodes along with the distance, energy, bandwidth of the neighbouring nodes . Decisions are made by the routing table to chose the best for data transmission. Another important function handled by the data link layer is flow control. The sender and the receiver come to an understanding on how much of data can be sent, so that it is convenient for both the sender and the receiver to optimize the bandwidth utilization, this function is called flow control. Our proposed system has been classified into three phases based on the functionalities such as initialization, data collection and communication phases.

Initial phase

1. Initially all nodes are location aware and know their distance from the sink.
2. On initialisation every sensor node broadcasts a control message (ADV) which includes source ID, residual energy and radius .
3. On receiving each node creates a neighbourhood table (routing table) and updates distance from the neighbour node and residual energy.

Data Collection Phase

1. In this phase when the data is sensed ,the nodes verify the routing table and chooses the best path for transmission.
2. Depending upon the data sensed ,based on it the best path is chosen for transmission.
3. Based on the size of the data and the updates in the routing table the first best path is chosen.

Communication phase

1. The sensed data is transmitted to the destination by referring the routing table to chose the first best routing path and turning the remaining nodes that are away from the path to idle state .
2. Since the path is organised based on the bandwidth of the nodes ,based on the size of the packet the best active path is chosen from the routing table.

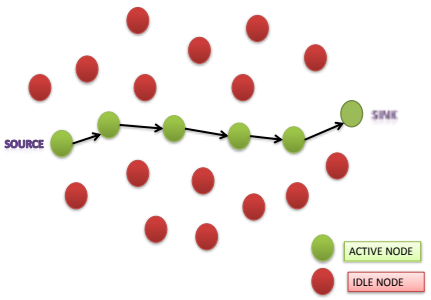


Fig 3: Data Transmission using the best shortest path

In the initialization phase, nodes in the network introduce themselves by broadcasting the ADV message which contains the source Id, location info, residual energy and best path schedule information to immediate neighbours which are within the communication radius.

Then a routing table is created by the AODV protocol based on the information obtained to chose the shortest routing path to perform the load balancing. The MAC protocol uses the sleep schedule information to control the medium access and reduce the energy wastage. The energy efficient design and implementation of MAC protocol depends on the capabilities and features of the underlying radio transceiver and microcontroller. Our MAC protocol uses the power of in-network data processing and communication tasks simultaneously to determine the total energy required to transmit the amount of sensed data. In-network data processing greatly reduce the energy consumption compared to transmitting all the raw data to the end node[5]. Proposed system integrates the shortest path routing algorithm and stores the routing table and uses the best shortest paths to route the control message and payload transmission. If the first shortest route is not feasible due to active/idle duty cycle then it selects the second best shortest path until the first best route is available.

In case of huge data transfer from multiple sensors, data packets can be converted into small packets and can be sent across using the shared transmission medium. Smaller data packet from multiple source can be placed in any order as long as there is available bandwidth to carry it. Each data packet contains the destination address and sequence number and it will help the receiver node to reconstruct the node data if required and sent to the proper recipient. Our method utilize the full extent of available network bandwidth for data transmission and it helps in reducing the waiting time of the contender nodes.

V. SIMULATION RESULTS

We have following assumptions in our simulation setup:

1) Sensor nodes are static and there is only one sink node.
2) Nodes are location aware and know their distance from the sink node.

We used the simple energy model explained in [19] for energy consumption evaluation. The energy consumption for sending *l*-bits data to distance d is ETx and energy consumption for receiving *l*-bits data is ERx The energy model used as shown in the following equations

ETx(*l*-bits, d) = *l*-bits(Edpb + ξfs * d2) -------(4)
And energy for receiving the *l*-bits data bit is equal to
ERx (*l*-bits) = *l*-bits *Edpb --------(5)

Where ξfs is constant and ξfs = 100 pJ/bit/m2 and Edpb = 50nJ/bit, Edpb is energy dissipated per bit .

The parameters involved in the process of deployment of nodes and nodes scheduling process with the details of energy consumption is given by the simulation in NS2.

Channel : Wireless channel
Propogation : Two way propogation
Antenna : Omni antenna
No. of nodes : 25
Routing protocol :AODV
Energy : 100 joules
Packet Size : 5000
Network : Phy/wireless Phy
Queue length: 5

Following simulator shows the deployment of nodes in wireless sensor network using NS2 network simulator.

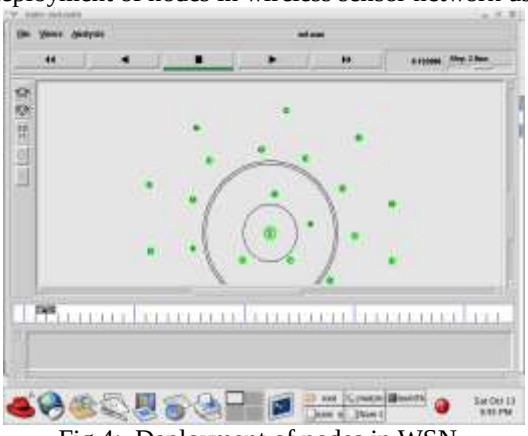


Fig 4: Deployment of nodes in WSN

Using the NS2 network simulator sleep nodes are used to conserve energy .So by manually turning the nodes to sleep state and performing the data transmission by the active nodes that stay awake.

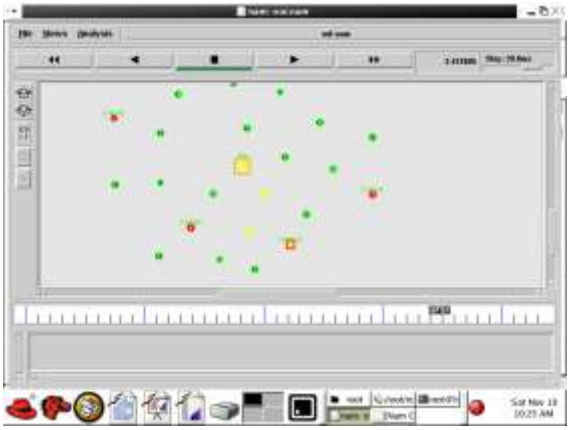


Fig 5: Data Transmission using active nodes

A final simulation using the best shortest path is performed with the representation of the graph specifying the energy, delay and analysis of the system. And also the problems of existing system are addressed.

VI. Conclusion

In this paper, the problems of energy conservation and full sensing coverage in large WSNs where nodes are randomly deployed have been addressed. An original technique has been introduced based on simultaneous wake-up scheduling concept allowing one to extend the lifetime of the WSN.

Proposed technique is based on using the routing table to choose the best active path for data transmission while engaging all other nodes to idle state. First main feature of this technique is, it is possible to choose an alternate path through the routing table based on the size of the packet. Each best shortest path is formed based on the bandwidth of each node. Second main feature is that a proper understanding is maintained between the source and destination such that the packet is divided and transmitted through the nodes based on the bandwidth of the nodes. Third feature is that if an attacker is present the IDS protocol is used to identify the attacked node and chooses the next best shortest path from the routing table. This has contributed to extend the WSN lifetime.

REFERENCES

[1] I. Demirkol, C. Ersoy, and F. Alagoz. Mac protocols for wireless sensor networks:a survey. Communications Magazine, IEEE, 44(4):115-121, april 2006.

[2] Jamal N. Al-karaki and Ahmed E. Kamal. Routing techniques in wireless sensor networks: A survey. IEEE Wireless Communications, 11:6-28, 2004.

[3] Jiageng Li, D. Cordes, and Jingyuan Zhang. Power-aware routing protocols in ad hoc wireless networks. IEEE Wireless Communications, 12:69- 81, 2005.

[4] D Niculescu. Communication paradigms for sensor networks. IEEE Communications Magazine, 43(3):116-122, 2005.

[5] Meguerdichian, S., Koushanfar, F., Potkonjak, M.,Srivastava, M.B.: 'Coverage problems in wireless ad-hoc sensor networks'. INFOCOM 2001: 20th Annual Joint Conf. IEEE Computer and Communications Societies, 2001, vol. 3, pp. 1380–1387

[6] Thai, M.T., Wang, F., Hongwei Du, D., Jia, X.:'Coverage problems in wireless sensor networks: designs and analysis', Int. J. Sens. Netw., 2008, 3,(3), pp. 191–200

[7] Ammari, H.M., Giudici, J.: 'On the connected k- coverage problem in heterogeneous sensor nets:the curse of randomness and heterogeneity'.ICDCS '09: Proc. 2009 29th IEEE Int. Conf. on Distributed Computing Systems, 2009, pp. 265–272

[8] Wang, X., Xing, G., Zhang, Y., Lu, C., Pless, R.,Gill, C.: 'Integrated coverage and connectivity configuration in wireless sensor networks', ACM Trans. Sens. Netw., 2005, 1, (1), pp. 28–39

[9] Zhang, H., Hou, J.: 'Maintaining sensing coverage and connectivity in large sensor networks', Ad Hoc Sens. Wirel. Netw., 2005, 1, (2) pp. 89–124

[10] Chang, J.-H., Tassiulas, L.: 'Energy conserving routing in wireless adhoc networks'. INFOCOM 2000: 19th Annual Joint Conf. IEEE Computer and Communications Societies, 2000, pp. 22–31

[11] Zhang, H., Wang, H., Feng, H.: 'A distributed optimum algorithm for target coverage in wireless sensor networks'. Asia-Pacific Conf. On Information Processing, 2009, pp. 144–147
[12] Chen, A.: 'Designing localized algorithms for barrier coverage'. Proc. ACM MobiCom'07,2007
[13] Osmani, A., Dehghan, M., Pourakbar, H., Emdadi, P.: 'Fuzzy-based movement-assisted sensor deployment method in wireless sensor networks'. Int. Conf. on Computational Intelligence, Communication Systems and Networks, 2009, pp. 90–95
[14] Cardei, M., Thai, M.T., Yingshu, L., Weili, W.:'Energy-efficient target coverage in wireless sensor networks'. INFOCOM 2005: 24th Annual Joint Conf. IEEE Computer and Communications Societies, 2005, vol. 3, pp. 1976–1984
[15] Slijepcevic, S., Potkonjak, M.: 'Power efficient organization of wireless sensor networks'. ICC 2001: IEEE Int. Conf. on Communications, 2001, pp. 472–476
[16] Tian, D., Georganas, N.D.: 'A coverage preserving node scheduling scheme for large wireless sensor networks'. WSNA'02: Proc. First ACM Int. Workshop on Wireless Sensor Networks and Applications, 2002, pp. 32–41
[17] Ye, F., Zhong, G., Lu, S., Zhang, L.: 'Peas: a robust energy conserving protocol for long-lived sensor networks'. Proc. of Int. Conf. on Distributed Computing Systems (ICDCS), 2002, pp. 28–37
[18] Chih-fan, H., Mingyan, L.: 'Network coverage using low duty-cycled sensors: random & coordinated sleep algorithms'. IPSN'04: Proc.Third Int. Symp. on Information Processing in Sensor Networks,2004.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834, p- ISSN: 2278-8735.Volume 4, Issue 6 (Jan. - Feb. 2013), PP 109-115
www.iosrjournals.org

Functional Link Artificial Neural Network for Denoising of Image

Charu Pandey ,Vartika Singh ,O.P.Singh ,Satish Kumar
Deptt. of Electronics & Electrical Engineering (ASET), Amity University, Uttar Pradesh, Lucknow

Abstract: *Digital image denoising is crucial part of image preprocessing. The application of denoising procesn satellite image data and also in television broadcasting. Image data sets collected by image sensors are generally contaminated by noise. Furthermore, noise can be introduced by transmission errors and compression. Thus, denoising is often a necessary and the first step to be taken before the images data is analyzed. In this paper a Modified Functional Link Artificial Neural Network (M-FLANN) is proposed which is simpler than a Multilayer Perceptron (MLP). It have been implemented for image restoration in this paper. Its computational complexity and speed and generalization ability to cancel Gaussian noise is compared with that of MLP. In the single layer functional link ANN (FLANN) the need of hidden layer is eliminated. The novelty of the FLANN structure is that it requires much less computation than that of MLP. In the presence of additive white Gaussian noise, salt and pepper noise, Random variable impulse noise and mixed noise in the image the performance of the proposed network is compared with that of MLP in this paper. The Performance of the of algorithm is evaluated for six different situations i.e. for single layer neural network, MLP and four different types of expansion in FLANN and comparison in terms of computational complexity also carried out.*

Index Terms— *MLP, FLANN, Salt and Pepper noise.*

I. Introduction

The quality of digital medical images has become an important issue today. To achieve the best possible diagnoses, it is important for medical images to be sharp, clear, and free of noise and artifacts [1]. Though, the technologies used to improve resolution and quality of noisy images remains an issue in many medical images applications. Removing noise in these digital images remains as one of the major challenges in the study of medical imaging. Denoising of ultrasound images is particularly challenging due to their peculiar texture [3]. The presence of noise will degrade the quality of image, and even conceal image details, which affects the subsequent image segmentation, feature extraction and recognition, quantitative analysis, and most importantly disease diagnosis. A computationally efficient artificial neural network (ANN) for the purpose of dynamic nonlinear system identification is proposed[6]. The major drawback of feed forward neural networks, such as multilayer perceptrons (MLPs) trained with the back propagation (BP) algorithm, is that they require a large amount of computation for learning. We propose a single-layer functional-link ANN (FLANN) in which the need for a hidden layer is eliminated by expanding the input pattern by Chebyshev polynomials. The novelty of this network is that it requires much less computation than that of a MLP. We have shown its effectiveness in the problem of nonlinear dynamic system identification. In the presence of additive Gaussian noise, the performance of the proposed network is found to be similar or superior to that of a MLP. A performance comparison in terms of computational complexity has also been carried out.

Recently artificial neural network (ANN) has emerged as a powerful learning technique to perform complex tasks in highly nonlinear environment [9]. The advantages of ANN model are due to there ability to learn based on optimization of an appropriate error function and there excellent performance for approximation of nonlinear functions.

The functional link artificial neural network (FLANN) by Pao can be used for function approximation and pattern classification with faster convergence and lesser computational complexity than a MLP network. A FLANN using sin and cos function for functional expansion for the problem of nonlinear dynamic system identification has been reported [13].A single layer orthogonal neural network using Legendre polynomials has been reported for static function approximation. In this work we proposed a FLANN structure similar to a Chebyshev polynomial-based unified model ANN for de-noising in an image corrupted with different noises[7]. Generally a linear node in its output is used in the FLANN structure reported by other researchers.

This paper tries to improve the time complexity involved in training the network, along with minimizing the number of interconnection weights and biases that are used as the network parameters which make it suitable for on-line application in comparison to MLP. This technique is able to approximate linear as well as non linear functions. This is computationally cheap and has small convergence time[3]. It de-noises image more efficiently.

II. Real-life applications of ANN

The tasks artificial neural networks are applied to tend to fall within the following broad categories:

a) Function approximation, or regression analysis, including time series prediction, fitness approximation and modeling.

b)Classification, including pattern and sequence recognition, novelty detection and sequential decision making.

c) Data processing, including filtering, clustering, blind source separation and compression.

d) Robotics, including directing manipulators, Computer numerical control.

III. Structure Of The Artificial Neural Network

Here, we briefly describe the architecture and learning algorithm for multilayer neural network and FLANN.

A. Multilayer perceptron

The MLP has a multilayer architecture with one or more hidden layers between its input and output layers. All the nodes of a lower layer are connected with all the nodes of the adjacent layer through a set of weights. All the nodes in all layers (except the input layer) of the MLP contain a nonlinear tanh( ) function. A pattern is applied to the input layer, but no computation takes place in this layer[4]. Thus the output of the nodes of this layer is the input pattern itself. The weighted sum of outputs of a lower layer is passed through the nonlinear function of a node in the upper layer to produce its output. Thus, the outputs of all the nodes of the network are computed. The outputs of the output layer are compared with a target pattern associated with the input pattern[5]. The error between the target pattern and the output layer node is used to update the weights of the network. The MSE is used as a cost function and BP algorithm attempts to minimize the cost function by updating all weights of the network[6].

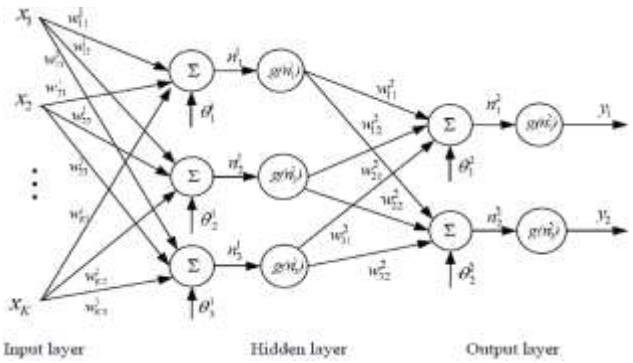


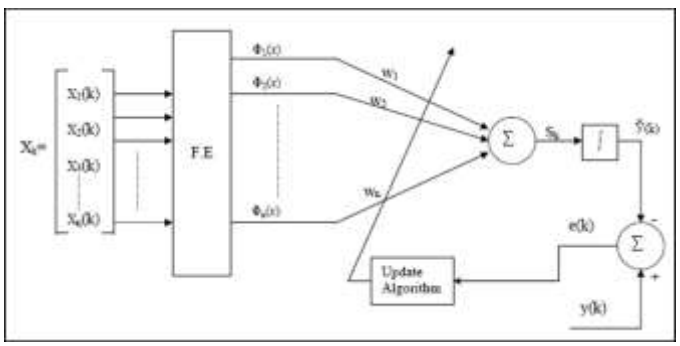
Figure 1: Multilayer perceptron neural network architecture

B. Functional link ANN

The FLANN, which is initially proposed by Pao, is a
single layer artificial neural network structure capable of performing complex decision regions by generating nonlinear decision boundaries[9]. In a FLANN the need of hidden layer is removed. In contrast to linear weighting of the input pattern produced by the linear links of a MLP, the functional link acts on the entire pattern by generating a set of linearly independent functions.

Here the functional expansion block comprises of exponential polynomials. Suppose input to this structure is X=[x1 x2]'. An enhanced pattern obtained by using functional expansion given by

X= [1x1T2(x)……x2T2(x2)……] '.



F.E: - Functional Expansion

Figure 2: Structure of FLANN with a single output

a) Learning with the FLANN

Let there be K number of input-output pattern pairs to be learned by the FLANN. Let the input pattern vector X is of dimension and for ease of understanding, let the output, y be a scalar. Each of the input patterns is passed through a functional expansion block producing a corresponding N-dimensional (N>=n) expanded vector.

In this case, the dimension of the weight matrix is of 1 x N and hence, the individual weights are represented by a single subscript[10]. Let w= [w1 w2…..wN] be the weight vector of its FLANN. The linear weighted sum, Sk is passed through the tanh (.) nonlinear function to produce the output $\breve{Y}$ (k) with the following relationship:

$$\hat{y}_k = \tanh(S_k) \;\; or \;\; S_k = \frac{1}{2}\log_e\left(\frac{1+\hat{y}_k}{1-\hat{y}_k}\right) \qquad (1)$$

b) The Learning Algorithm

Let K number of patterns be applied to the network in a sequence repeatedly. Let the training sequence be denoted by {Xk,Yk} and the weight of the network be W(k), where k is the discrete time index given by k=k+λk, for λ=0,1,2…, and k=0,1,2….,K. At kth instant, the n-dimensional input pattern and the m-dimensional FLANN output are given by

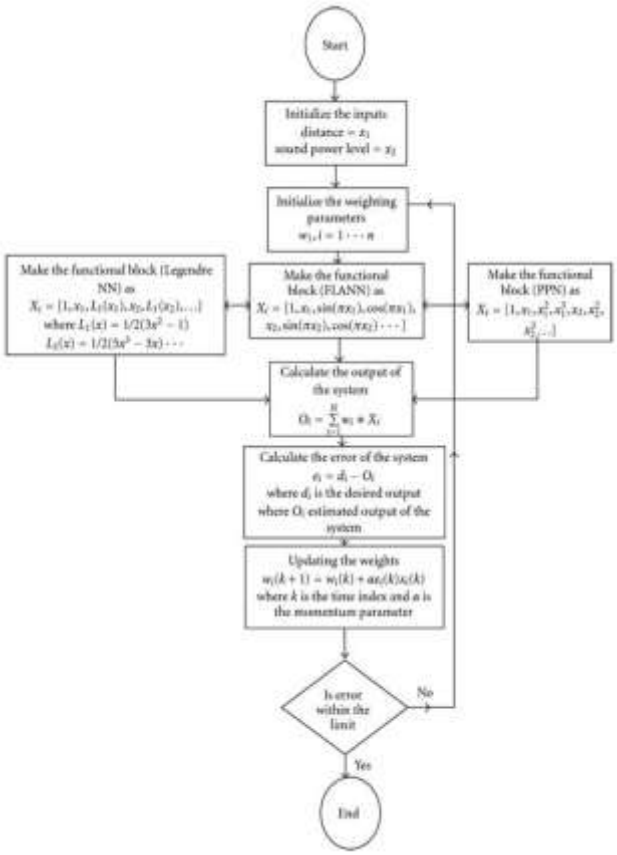


Figure 3: Systematic algorithm of functional based neural network

Xk=[x1(k) x2(k)……..xn(k)]T and

$\breve{Y}$(k) = [$\breve{Y}$1 (k) $\breve{Y}$2 (k)…….. $\breve{Y}$m (k)] T respectively.

Its corresponding target pattern is represented by

Y (k) = [y1 (k) y2 (k)…….. ym (k)]T.

The dimension of input pattern increases from n to N by a basis function Φ given by

$$\Phi (Xk) = [\Phi 1 (Xk)\ \Phi 2 (Xk)\ldots\ldots..\ \Phi N (Xk)]T$$

The (mxN) dimensional weight matrix is given by

$$W(k)=[W1(k)W2(k)\ldots\ldots..Wm(k)]T$$

Where Wj(k) is the weight vector associated with jth output and is given by

$$Wj (k) = [Wj1 (k)\ Wj2 (k)\ldots\ldots..WjN (k)]$$

The jth output of the FLANN is given by

$$\hat{y}_j(k) = \rho\left(\sum_{i=1}^{N} w_{ji}(k)\,\phi(X_k)\right)$$

for j=1, 2, 3….m

c) Trigonometric Functional Expansion

Here the functional expansion block make use of a functional model comprising of a subset of orthogonal sin and cos basis functions and the original pattern along with its outer products. For example, consider a two dimensional input pattern X=[x1x2] T, the enhanced pattern is obtained by using a trigonometric functions as x'= [x1cos (πx1) sin (πx1) cos (3πx1) sin (3πx1)]. The LMS algorithm, which is used to train the network, becomes very simple because of the absence of any hidden layer.

d) Chebyshev Expansion

The Chebyshev polynomials are a set of orthogonal polynomials defined as the solution to the Chebyshev differential equation and denoted as Tn(x). These higher the first few Chebyshev polynomials are given by

$$T_0(x) = 1.0,\ T_1(x) = x \text{ and } T_2(x) = 2x\text{^}2 - 1$$

The higher order Chebyshev polynomials may be generated by the recursive formula given by:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

The first few Chebyshev polynomials are given by

T0(x) = 1
T1(x) = x
T2(x) = 2x2-1
T3(x) = 4x3-3x
T4(x) = 8x4-8x2+1
T5(x) = 16x5-20x3+5x

2.1.7 Exponential expansion

In this study we used exponential polynomials for functional expansion as shown in figure-1. These polynomials are easy to compute than that of other polynomials. For two-dimensional input patterns, the enhanced pattern is obtained by using exponential functions as:

X1 = [x1expx1 exp2x1 exp3x1… x2expx2 exp2x2 exp3x2 exp4x2….]T.

This type of polynomial expansion needs less number of computations and is very easy to implement than other three types of polynomial expansion.

IV. Computational Complexity

Here, we present a comparison of the computational complexity between an MLP, and a FLANN having different expansions and all having the tanh(.) as their nonlinear function. In all the cases, multiplications, additions and computations of the tanh(.) are required. However in the case of FLANN, additional computations of the sine and cosine functions are needed for its functional expansion. In the training and updating of the weights of MLP, extra computations are incurred due to its hidden layer. This is due to the error propagation for the calculation of the square error derivative of each neuron in the hidden

layer. For each iteration the computation are: (1) Forward calculations to find the activation value of all the nodes of the entire network.(2) Back-error propagation for calculation of square error derivatives.(3) Updating weights of the entire network. In the case of MLP with {I-J-K}, the total number of weights is given by (I+1)J + (J+1)K. Whereas, in the case of FLANN with {D-K}, it is given by (D+1)K. The number of computation for both MLP and FLANN are shown in table:

TABLE
COMPARISON OF COMPUTATIONAL COMPLEXITY IN ONE ITERATION

Operation	MLP {I-J-K}	FLANN {D-K}
Addition	2IJ + 3JK + 3K	2K(D+1) + K
Multiplication	3IJ + 4JK + 3J + 5K	3K(D+1) + 2K
Tanh(.)	J+K	K

From this table it may be seen that the number of additions, multiplications and computation of tanh are much less in case of a FLANN than that of a MLP network. As the number of hidden layer increases the computations in a MLP increases. But due to absence of hidden layer in the FLANN its computational complexity reduces drastically.

V. Simulation Results

Experiment-1

TABLE:

Gaussian Noise Mean zero and Variance	Noise Power at Input of FLANN Filter	Noise power at output of FlANN Filter	PSNR at Input of FLANN Filter	PSNR at Output of FLANN Filter	Noise Power Attenuated By the Filter in dB
0.01	0.0429	0.00762	16.5456	26.6325	7.5
0.02	0.0339	0.00724	17.1014	25.6989	6.7
0.04	0.0281	0.00756	16.0198	23.6656	5.7

In this experiment computer simulations are carried out to evaluate the performance of Functional link neural network with exponential functional expansion for denoising of image corrupted with Salt and Pepper noise of mean zero and different variance. Here the 9 inputs of this network were the entries of3X3 window of the noisy image and the target was the corresponding pixel value of the original image.

Experiment-2

In this experiment its tried to cancel noise by using M-FLANN network. First the noisy Cameraman image is pass through the single layer neural network and the target be the original cameraman image. Image is corrupt with Gaussian noise of mean 0 and variance 0.01. After passing through M-FLANN network the image obtain will be shown here. It is clear from result that image obtain is objectively better.

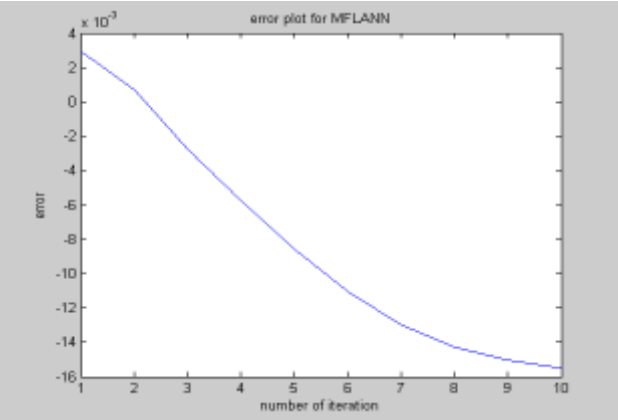
a) Original Image

b) Noisy Image

c) Image Pass Through M-FLANN

Experiment3

In this experiment computer simulations are carried out to evaluate the performance of Modified Functional link neural network with trigonometric functional expansion for denoising of image corrupted with Salt and Pepper Noise of 10 percent density.



VI. Conclusion

Simulation results indicate that the performance of FLANN is better than MLP for Salt and Pepper noisesuppression from an image.It has better convergence speed than MLP. It has no hidden layer, is computationally cheap and has small number of interconnection weights and biases, which makes it suitable for on-line applications. Also, MFLANN makes use of linear terms, just as used in FLANNs and is able to approximate linear as well as nonlinear functions. From the results obtained in the work, it is concluded that the M-FLANN is computationally cheap and has small convergence times and denoise an digital image more efficiently than MLP.M-FLANN improves upon the approximation ability of FLANN. Therefore, it is expected that M-FLANN is likely to re-ignite the interest of research community in FLANNs.

VII. Future Work

Better result will come out if we will improve the performance of the network by using efficient algorithms for color images processing in internet & wireless applications.

References

[1] Guez Allon, James L. Ellbert and Moshe Kam, "Neural networks architecture for control," IEEE Control System magazine, pp. 22-25,1998.
[2] Rivals Isabelle and Léon Personnaz, "Nonlinear internal model control using neural networks: Application to processes with delay and design issues," IEEE Trans. on Neural Networks, vol. 11, no.1, pp. 80-90, 2000.
[3] Yu Wen and Li Xiaoou, "Some new results on system identification with dynamic neural networks," IEEE Transactions on Neural Networks, vol.12, no. 2, pp. 412-417, 2001.
[4] Naira Hovakimyan, Flavio Nardi, Anthony Calise, and Nakwan Kim, "Adaptive output feedback control of uncertain nonlinear systems using single-hiddenlayer neural networks," IEEE Trans. on Neural Networks, vol. 13, no. 6, pp. 1420-1431, 2002.
[5] Sunan N Huang., K. K. Tan and T. H. Lee, "Further results on adaptive control for a class of nonlinear systems using neural networks," IEEE Trans. on Neural Networks, vol. 14, no. 3, pp. 719-722, 2003.
[6] Sam Ge Shuzhi and Cong Wang, "Adaptive neural control of uncertain MIMO nonlinear systems," IEEE Trans. on Neural Networks, vol. 15, no. 3, pp. 674- 692, 2004.
[7] S. E Lee and B. R. Holt, "Regression analysis of spectroscopic process data using a combined architecture of linear and nonlinear artificial neural networks," Neural Networks, vol. 4, pp. 549-554, 1992.
[8] Ying-Kai Zhao, Ning Cai, "DLF NEural network applied to process modelling and control," Proc. Pacific-Asian Conf. on Expert Systems PACES'95: Huangshan, PR China (1995) T.X Brown., "A high performance two-stage packet switch architecture," IEEE Transactions on Communications, vol. 47, no. 8, pp. 1792-1795, 1999.
[9] Patra.J.C, Pal.R. N, Chatterji.B.N, Panda,G, "Identification of nonlinear dynamic systems using functional link artificial neural networks" IEEE Transactions , Systems, Man and Cybernetics, Part B,Vol 29 , April-1999, pp 254 – 262.
[10] A.Namatame, and N.Ueda,"Pattern classification with Chebyshev neural networks," Ind.J.Neural Networks Vol 3,Mar. 1992, pp 23-31
[11] Patra,J.C, Pal,R.N, "Functional link artificial neural network-based adaptive channel equalization of nonlinear channels with QAM signal" IEEE International Conference, Systems, Man and Cybernetics, 1995,Vol3, Oct.-1995, pp 2081-2086
[12] R Grino, G.Cembrano, and C.Torres, "Nonlinear system Identification using additive dynamic neural networks two on line approaches."IEEE Trans Circuits SystemIvol47, Feb 2000, pp 150-165.
[13] A.R.Foruzan, B.N.Araabi, "Iterative median filtering for restoration of images with impulsive noise.'' Electronics, Circuits and Systems, 2003. ICECS 2003.Dec 2003 ,PP 14-17 Dec.
[14] L. Corbalan, G.Osella, Massa.C.Russo, L.Lanzarini,. De Giusti ''Image Recovery Using a New Nonlinear daptive Filter Based on Neural Networks''Journal of Computing and Information Technology - CIT 14, Apr.2006, pp 315– 320.