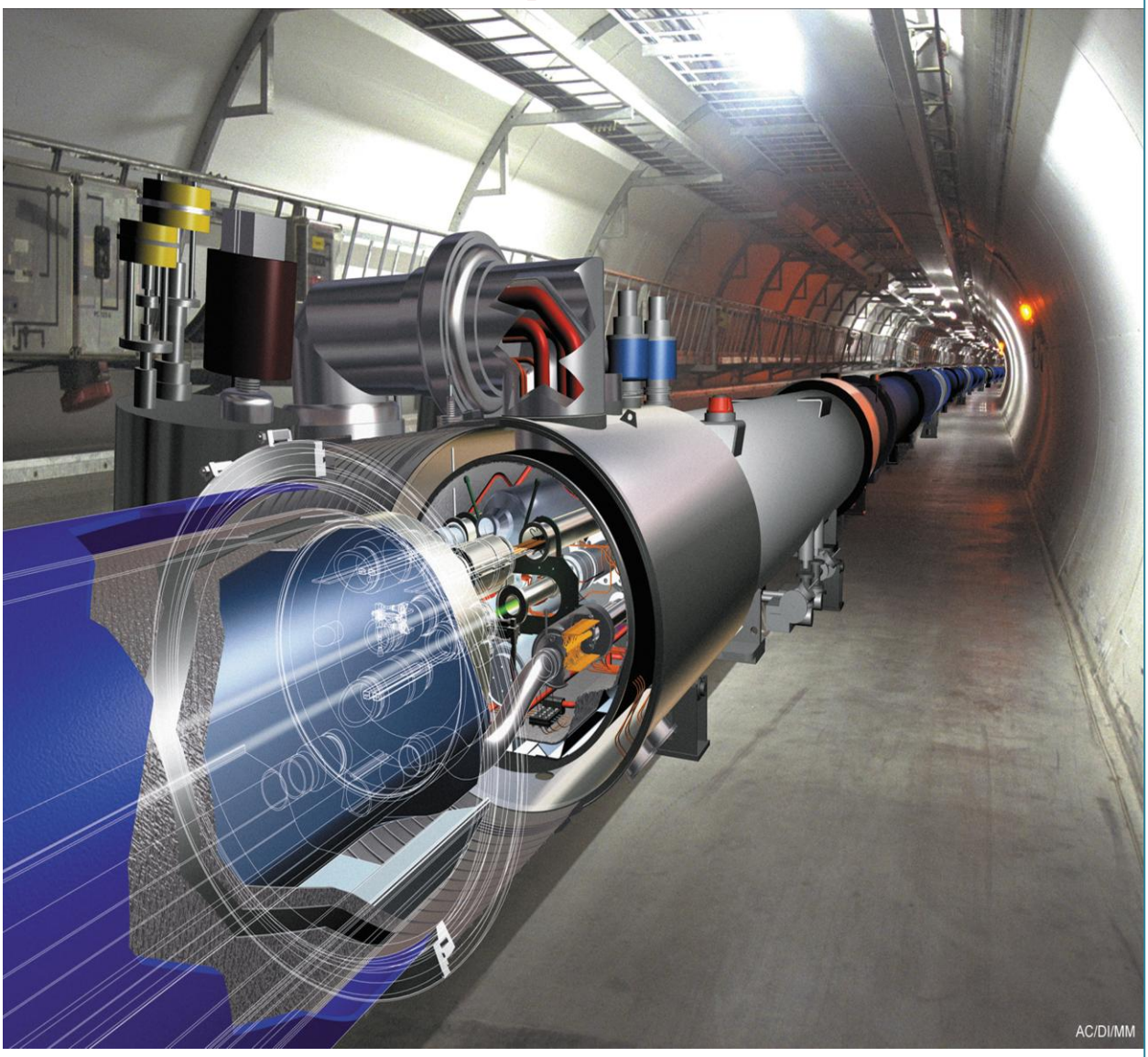
International Journal of computational Engineering Research (IJCER)

ISSN: 2250-3005

VOLUME 2 September 2012 ISSUE 5



Email: ijceronline@gmail.com **Url : www.ijceronline.com**

International Journal of computational Engineering Research (IJCER)

Editorial Board

Editor-In-Chief

Editorial Committees

CONTENTS :

Design and Analysis of HP steam turbine casing for Transient state condition

[1]J.Ramesh,[2]C.Vijaya Bhaskar Reddy, [3] Dr. B. Jayachandraiah

P.G Student[1]
Assistant Professor[2] (Sr)
Head & Vice-principal[3]
Department Of Mechanical Engineering,
Sri Kalahasteeswara Institute Of Technology, Srikalahasti.

Abstract

Transient regimes arising during start-ups, shut-downs and load changes give rise to unsteady temperature distribution with time in steam turbine casing high pressure (HP) which results in non-uniform strain and stress distribution. So that problems such as stress corrosion, cracking and fatigue of steam turbine casing, In this work the thermo mechanical analysis of steam turbine casing will be established by finite element method. In this work the temperature and stress distributions for turbine inner casing were calculated by finite element analysis. The three dimensional model of the Steam Turbine Casing was created using the CATIA software. The model was meshed using software HYPERMESH. Boundary conditions were given on the finite element model through ANSYS.

In this paper, the transient temperatures and stresses distributions within a turbine inner casing were achieved from actual operation data during cold start-up. The paper analyses the creap, centrifugal stress sub stained at high temperature in thermal stresses setup during the startup and shutdown the stream turbine and most serious thread of the rotor blades near the bore, creep cracks to initiates go to size which could results brittle fracture of the rotor blades. Due to crackness life of the steam turbine decreases.

Keywords: Transient condition, 3-D model, Hypermesh, FE model, Thermal expansion

1.Introduction

Generally turbine casings used are split horizontally and vertically. The casing houses the blades rotor, nozzles, and diaphragms. It also holds glands for steam sealing at each end for preventing leakage of steam from where the shaft passes through. The steam casing of turbine is generally arranged with centre line support i.e the support points are on the same horizontal plane as the centre line of the turbine. The steam end pedestal sits upon a flexible panting plate which provides rigidity in the vertical and lateral planes, but allows flexibility in the axial plane for casing thermal expansion.

The combined thrust and journal bearing of the turbine rotor is housed in the steam end pedestal. The rotor, therefore, is moved axially towards the steam end with the axil movement of the casing. The casing is that portion of the turbine that either supports or supported by the bearing housings. The steam ring is attached to or is a part of the casing. All casing joints have metal to metal sealing surfaces no strings or gaskets are used. All turbines manufactured by Maxwatt use multiple piece casings consisting of two or more pieces that are split at the horizontal centerline to facilitate inspection or removal of the turbine rotor. The casings are either cast, fabricated, or a combination of both depending on operating conditions. The casing can be of iron, carbon steel, carbon moly steel, or chrome moly steel.

Types Of Casings

1.1 Single Stage Turbine Casing

Single stage turbine casings are two piece casings. The casing consists of the upper half and the lower half the lower half casing houses both the inlet and exhaust connections. The bearing housings are either cast integrally with or bolted to the lower half casing. All single stage frames except the type GA and SA turbine have the steam ring cast integrally with the casing this means that the material required for the steam ring is also required for the casing. In those cases where the steam ring is not cast as the part of the casing, different materials for the steam ring and casing can be utilized. The two components are bolted together.

1.2 Multistage Turbine Casing:

Multistage turbine casing are considered to be horizontally split casing even through a vertical split may also be used. The point of juncture of the vertical and horizontal splits is called a four way joint and is the most difficult area in the whole turbine assembly to seal against steam pressure because of this Maxwatt employs a construction called barrel construction in normal construction the steam ring forms the high pressure section of the casing and is bolted directly to the intermediate and exhaust portions of the casing. This puts the four way split at the first stage which is where the highest case pressure is encountered in the multistage turbine.

2. Literature Review

[1] Has developed modern CAE tools like Hyper mesh, ANSYS and Pro-E etc. have been utilized to model and analyze existing LP casing and for redesigning it to suit the new efficient modern design of rotor. This paper presents the numerical stress analysis of the turbine casing of an aero-engine. High thermal stress gradients were found at the region of casing where fatigue cracks were detected during engine operation [2]. [3] Has analyzing the failure of a weld repaired turbine casing after 30 years of total service including 5 years after weld repair. The casing was weld repaired by a high Cr–Ni weld metal (24Cr–32Ni–4Mn–Fe). The base metal low alloy ferritic steel (1Cr–0.5 Mo steel) with ferrite–pearlite structure did not show any abnormality to indicate significant degradation. [4] Has studied about designing of complex steam turbine low pressure casing the ever growing competition in capital intensive power sector is pressing turbine manufacturer's world over to develop new efficient steam turbine designs and to update/retrofit the old steam turbine which are in service. BHEL is also putting up major design development efforts to meet the present market challenges. This paper depicts how the modern CAE tools like Hypermesh, ANSYS and Pro-E etc. have been utilized to model and analyse existing LP casing and for redesigning it to suit the new efficient modern design of rotor.

3. Modeling And Analysis

3.1 Steam Turbine Casing Model

It is very difficult to exactly model the Steam Turbine casing, in which there are still researches are going on to find out transient thermo mechanical behavior of casing during operating under higher temperature and pressure. There is always a need of some assumptions to model any complex geometry. These assumptions are made, keeping in mind the difficulties involved in the theoretical calculation and the importance of the parameters that are taken and those which are ignored. In modeling we always ignore the things that are of less importance and have little impact on the analysis. The assumptions are always made depending upon the details and accuracy required in modeling.

The assumptions which are made while modeling the process are given below

- The casing material is considered as homogeneous and isotropic.
- Inertia and body force effects are negligible during the analysis.
- The casing is stress free before the start up.
- The analysis is based on pure thermal loading and vibration and
- Thus only stress level due to the above said is done. The analysis does not determine the life of the casing.
- The thermal conductivity of the material used for the analysis is uniform throughout.

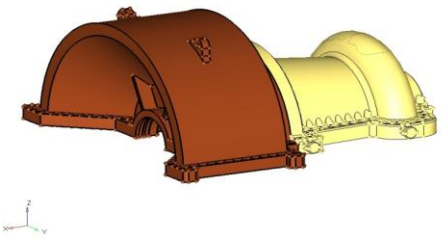


Fig 3.1 CAD model of the casing

3.2 Steps Involved in Finite Element Analysis

- Discretization of the continuum
- Selection of the displacement model
- Derivation of element stiffness matrix
- Assemblage of the algebraic equations for the overall discretized Continuum.
- Solution of the unknown displacements.

3.3 Heat Transfer Analysis

Heat transfer of a casing surface is affected by convection and temperature distribution of inner part is calculated by conduction. The boundary conditions between surface and inner area for the thermal analysis were derived from calculated heat transfer coefficient according to time Heat transfer analysis was conducted from pre-warming to steady state condition using heat transfer coefficients and a steam temperature of each location acquired from operating data. For HP casing are made from castings and the life assessment portions are corner radius, pipe inner surfaces and welds. For turbine casing, especially, the major damage occurs at the nozzle fit and disk corner of casing.

3.4 Thermal Analysis

A thermal analysis calculates the temperature distribution and related thermal quantities in steam turbine casing. Typical thermal quantities are

- The temperature distribution
- The amount of heat lost or gained
- Thermal fluxes
- Thermal gradient

Thermal simulations play an important role in the design of many engineering applications, including internal combustion engines, turbines, heat exchangers, piping systems, and electronic components. In many cases, engineers follow a thermal analysis with a stress analysis to calculate thermal stresses (that is, stresses caused by thermal expansions or contractions). The basis for thermal analysis in ANSYS is a heat balance equation obtained from the principle of conservation of energy. The finite element solution perform via ANSYS calculates nodal temperatures, and then uses the nodal temperatures to obtain other thermal quantities.

3.5 Transient Thermal Analysis

The ANSYS Multiphysics, ANSYS Mechanical, ANSYS Professional, and ANSYS FLOTRAN products support transient thermal analysis. Transient thermal analysis determines temperatures and other thermal quantities that vary over time. Engineers commonly use temperatures that a transient thermal analysis calculates as input to structural analyses for thermal stress evaluations. Many heat transfer application heat treatment problems, nozzles, engine blocks, piping systems, pressure vessels, etc. - involve transient thermal analyses. A transient thermal analysis follows basically the same procedures as a steady-state thermal analysis. The main difference is that most applied loads in a transient analysis are functions of time. To specify time-de-pendent loads, and then apply the function as a boundary condition, or you can divide the load-versus-time curve into load steps

3.5 Meshing

The goal of meshing in HYPERMESH Workbench is to provide robust, easy to use meshing tools that will simplify the mesh generation process. These tools have the benefit of being highly automated along with having a moderate to high degree of user control. In order to carry out a finite element analysis, the model using must be divided into a number of small pieces known as finite elements. Since the model is divided into a number of discrete parts, FEA can be described as a discretization technique. In simple terms, a mathematical net or mesh is required to carry out a finite element analysis. If the system under investigation is 1D in nature, use line elements to represent our geometry and to carry out our analysis. If the problem can be described in 2 dimensions, then a 2D mesh is required. Correspondingly, if the problem is complex and a 3D representation of the continuum is required, then we use a 3D mesh.

Fig 3.2: Steam Casing (HP) Meshed Model

The meshed assembly of a steam turbine casing is as shown in the Figure 3.1.Initially IGES file of a CATIA product has been imported to the HYPERMESH workbench then the meshing is carried out. In the present case we did tetra type of element has been used and detail information on meshed assembly as shown in Table 3.1

Object Name	Steam Casing
Length Unit	Millimeters
Bodies	13
Nodes	332917
Elements	1828152

Table 3.1 Detail Information about Steam Casing Meshing

Pre-Processing

The goals of the pre-processing are to develop an appropriate finite element mesh, assign suitable material properties and apply boundary condition in the form restraints and loads. The finite element mesh subdivides the geometry into elements, upon which are found nodes. The nodes which are just point location in the space are generally located at the element corner and near each mid side node there may be two-dimensional or three-dimensional elements 2D-elements can be plane stress axis-symmetric and plane strain condition s for a 3D solid analysis only one temperature or many temperature degrees of freedom exists at each node. Developing a mesh, most time consuming work in FEA. The geometry is meshed with mapping algorithm or free algorithm. The first maps a rectangular grid on to a geometric region, which must have the correct number of sides. Free meshing automatically sub-divides meshing regions into elements easy messing, disadvantage is of distorted elements.

Post- Processing

Post processing begins with a thorough check for problems that may have occurred during solution. Once the solution is verified to be free of numerical problems; the quantitative of interest may be examined.. Dynamic viewing and animation capabilities aid greatly in obtaining an understanding of the deformation pattern. Stresses being sensor qualifies, currently lack a single visualization technique, and thus derived stress quantifies are extracted and displayed. Principle stress vectors may be displayed as color-coded arrows, indicating both direction and magnitudes. Von-Misses stress may be displayed on the model as color bands. Displacements magnitudes may also be displayed by color bands, but this can lead to misinterpretation as stress plot.

4. RESULTS:

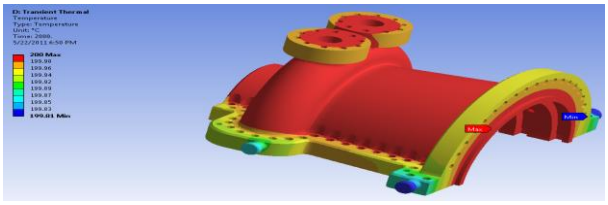


Fig4.1: Temperature Distribution in inner casing in unsteady (Transient) state condition after 2000s

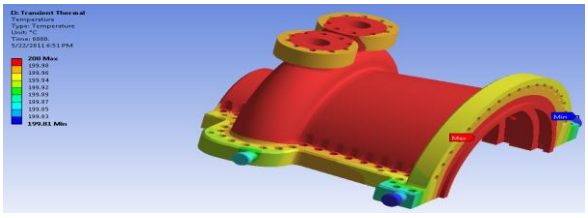


Fig4.2: Temperature Distribution in inner casing in unsteady (Transient) state condition after 6000s

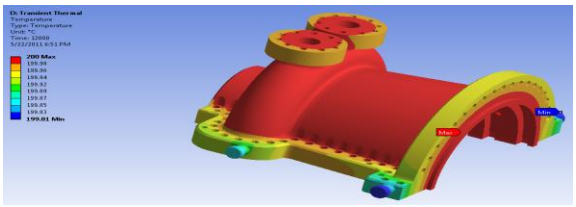


Fig43: Temperature Distribution in inner casing in unsteady (Transient) state condition after 12000s

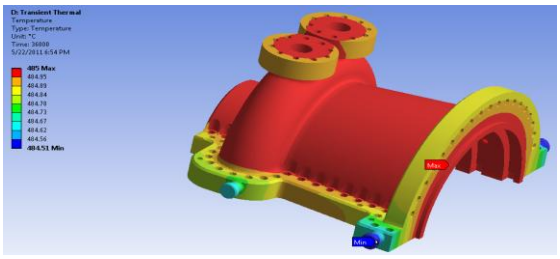


Fig4.5: Temperature Distribution in inner casing in unsteady (Transient) state condition after 36000s

5. Theoretical Calculation:

1. Thickness of Casing (t) = $\frac{P \times d}{2 \times (SE + PY)} + 1.25$

Where P is the Inlet pressure in bar, d is the diameter of casing in mm S is the Maximum allowable stress in material in (PSI) 500 Psi, E is the Joint efficiency factor 0.75 and , Y is Temperature coefficient 0.4.

2. Thermal Expansion = $\delta_t = \alpha \times \Delta t \times l$

3. Thermal Stresses ($\sigma_{Thermal}$) = $E \times \alpha \times \Delta t$

Here δ_t= Thermal expansion of casing in inch and α is Coefficient of thermal growth is 8.3×10^{-6} in/in^{0}F, Δt is the Temperature difference in ^{0}F and L be the length of casing in inch.

The Theoretical calculation for the whole HP casing is calculated as shown above and the calculations for each stage are tabulated in table

6. Comparison of Results:

	Analytical	Mm FEA-ANSYS
Total deformation mm	0.07	0.1
Stress (first stage) Pa	0.235E9	0.38 E9

Stages	Temperature		Pressure(bar)	Diameter (mm)	Length (mm)	Thermal Expansion (mm)	Thermal Stresses 10^9 N/m^2
	^{0}C	^{0}F					
1	260	500	8	688	68	0.042	0.1254
2	230	446	6	706.5	68	0.030	0.0941
3	185	365	4	725.5	74	0.049	0.1411
4	150	302	3.5	793	104	0.109	0.1098
5	105	221	3	851	163	0.029	0.1411

7. Conclusion:

To maintain a high level of availability and reliability in a fossil power plant, substantial consideration of failure by repeated thermal loading should be carried out.

- In this study, the transient temperatures and stresses distributions within a turbine inner casing were achieved from actual operation data during cold start-up.
- The maximum deformations are calculate in transient state condition within inner casing.
- Equivalent (von-Misses) Stress distribution in Transient condition.
- Total deformation and stress values are compared with analytical results calculated for 2D geometry.

If the thermal gradient is great enough, the stress at the bottom of the threads may be high enough to cause the carking. The result shows the casing develops higher stress levels in startup condition.

8. Scope for the Future Study:

Turbine facilities that operated repeatedly under transient start-up condition are affected by various damage mechanisms such as creep, fatigue, and oxidation and so on. In general, it is known that low cycle fatigue is one of main damage mechanisms to determine life time of turbine rotor or casing components.

This study can be extended further to calculate the fatigue damage by the stress analysis based procedure. This approach is based on Neuber's rule. Neuber's rule expresses the relationship between local stress and local inelastic total strain-range. Using this study, life consumption of steam turbine inner-casing can be obtained and a guideline for effective maintenance also can be proposed. The analysis can be extended further considering Elasto-Plastic analysis using non linear material properties as well as temperature dependent material properties.

References:

[1] W. S. Choi, E. Fleury, G. W. Song and J.-S. Hyun, A life assessment for steam turbine rotor subjected to thermo-mechanical loading using inelastic analysis, Key Eng. Mat. 326−328, 601−604 (2006).

[2] Lucjan Witek, Daniel Musili Ngii, thermal fatigue problems of turbine casing Vol. 1 (2009) 205-211

[3] Maneesh Batrani, BHEL Haridwar, Hypermesh an effective 3-D CAE Tool in Designing of complex steam turbine low pressure casing in 2006.

[4] T.Stubbs, the role of NDE in the life management of steam turbine rotors, Swindon, England

[5] K. Fujiyama, Development of risk based maintenance planning program for Power Plant steam turbine, Final report on the Joint Project, pp. 69−82 (2007).

[6] Kiyoshi SAITO, Akira SAKUMA and Masataka FUKUDA, "Recent Life Assessment Technology for Existing Steam Turbines", JSME International Journal Series B, Vol. 49, No. 2 (2006), pp.192-197.

[7] Development of Life Prediction System for Thermal Power Plant Based on Viscoplastic Analysis, Final report, KERPI (2007).

Performance Analysis of SPIN and LEACH Routing Protocol in WSN

[1]Geetu, [2]Sonia Juneja
Deptt. Of Comp. Sc. & Engg
HCE, Sonepat ,India

Abstract
Wireless sensor networks have emerged as a technology that are being quickly adopted due to their flexibility and use in a variety of environments. However, they consist of small, inexpensive devices or nodes that have severe constraints such as limited bandwidth, limited processing power, short battery life, small storage capability and are physically prone to external threats [1]. Sensor Network are emerging as a new tool for important application in diverse fields like military surveillance, habitat monitoring, weather, home electrical appliances and others. These sensor nodes have some constraints due to their limited energy, storage capacity and computing power. The energy efficiency is an important issue in WSN. Routing protocols makes the transmission in an efficient manner and ensures reliable delivery over multiple-hop relay in WSN. This paper analyses performance of the routing protocols.

Keywords: Wireless sensor networks LEACH and SPIN routing protocols, network structure, and energy efficiency.

I. INTRODUCTI'ON

A wireless sensor network [3], [5] with a large number of tiny sensor nodes can be used as an effective tool for gathering data in various situations. One of the major issues in wireless sensor networks is developing an energy-efficient routing protocol which has a significant impact on the overall lifetime of the sensor network. Sensor nodes measure the ambient conditions from the environment surrounding them. The applications of WSN are various from health monitoring to battle field. The practice of remote sensing has become greatly simplified by useful and affordable sensors as well as required software packages. Additionally, users can monitor and control the underlying environment from remote location. Many routing, power management, and data dissemination protocols have been specifically designed for WSNs where energy awareness is an essential design issue. Routing protocols in WSNs might differ depending on the application and network architecture. A sensor network (WSN) are highly distributed networks of small, lightweight wireless nodes, deployed in large numbers to monitor the environment or system by the measurement of physical parameters such as temperature, pressure humidity, sound, vibration, pollutants and collectively relay their sensed data to the sink node. Each node in the network connected to each other.

Each sensor node in the network consists of three subsystems:

1) The sensor subsystem which is used to sense the environment,
2) The processing subsystem which performs the local computations on the sensed data, and
3) The communication subsystem which is responsible for sharing the sensed data with the neighboring sensor nodes.

This paper aims to show analysis performance of routing protocol in wireless sensor network using data centric approach. Two Wireless Sensor Network simulator versions 2.34. Both of the routing protocols are selected from data centric routing. The result of the analysis for each protocol is compared and the best routing protocol using data centric approach is proposed for WSN. This paper examines the performance of each of routing protocols which improve the network efficiency and maximize the network lifetime.

II. Routing Protocols In Wsn

2.1 LEACH

LEACH [6](Low Energy Adaptive Clustering Hierarchy). These protocols uses cluster node for the purpose of transmission of information between the nodes. It is a self-organizing protocol and nodes organize themselves into local clusters and perform data transmission to the
Selection of cluster head node is not fixed and it depends on possibility of nodes, which possess high energy. Formation of cluster head is based on TDMA schedule for data transmission. Time Division Multiple Access (TDMA) used as a scheduling mechanism makes it prone to long delays when applied to large sensor networks. TDMA schedule prevents

data collision, among messages and preserve energy among non cluster nodes. The establishment of cluster head is as follows: Each node generates a random number between 0 and 1 and compares it with the threshold value P(n). If the number is less than the threshold value, it becomes the cluster head node. If it has been selected cluster head node in each round of cycle, the node's P (n) is set to 0 so that the node will not be re-selected as cluster head. Otherwise, the node is non-cluster head node in the current round. After the selection of cluster head, the head broadcast its own presence to all other nodes. After broadcasting the information, then all other nodes send the information to the cluster head. Together, these features allow LEACH to achieve the desired properties in the networks.

P(n)= p /1-p(r mod 1/p)
There are several desirable properties for protocol on these networks:
· Use 100's - 1000's of nodes
· Maximize the lifetime of system
· Maximize network coverage
· Use uniform, battery-operated nodes

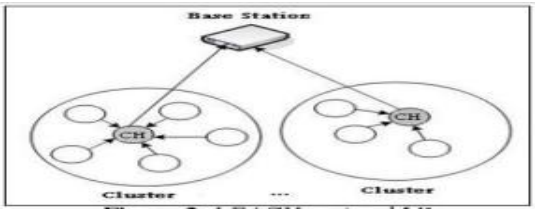


Fig 1: LEACH Protocol

As shown in fig.1, dark nodes specifies the cluster head and other non cluster head nodes send the information to cluster head on the basis of local information which in turn send the information to base station.
This protocol is divided into rounds; each round consists of two phases;

Set-up Phase
(1) Advertisement Phase
(2) Cluster Set-up Phase
Steady Phase
(1) Schedule Creation
(2) Data Transmission

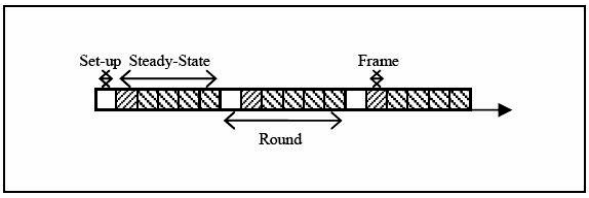


Fig.2: LEACH protocol phases [4]

Although LEACH is able to increase the network lifetime, there are still a number of issues about the assumptions used in this protocol. LEACH assumes that all nodes can transmit with enough power to reach the Base Station (BS) if needed and that each node has computational power to support different MAC protocols. Therefore, it is not applicable to networks deployed in large regions. It is not obvious how the number of the predetermined CHs (p) is going to be uniformly distributed through the network. Therefore, there is the possibility that the elected CHs will be concentrated in one part of the network. Hence, some nodes will not have any CHs in their vicinity. Furthermore, the idea of dynamic clustering brings extra overhead, e.g. head changes, advertisements etc., which may diminish the gain in energy consumption. Also, the protocol assumes that all nodes begin with the same amount of energy capacity in each election round, assuming that being a CH consumes approximately the same amount of energy for each node. The protocol should be extended to account for non-uniform energy nodes, i.e., use energy-based threshold.

2.2 SPIN

SPIN (Sensor Protocols for Information via Negotiation) Sensor Protocols for Information via Negotiation (SPIN) that disseminates all the information at each node to every node in the network assuming that all nodes in the network are potential BSs. This enables a user to query any node and get the required information immediately. These protocols make use of the property that nodes in close proximity have similar data, and hence there is a need to only distribute the data other nodes do not posses. The SPIN family of protocols uses data negotiation and resource-adaptive algorithms. Nodes running SPIN assign a high-level name to completely describe their collected data (called meta-data) and perform metadata negotiations before any data is transmitted. This ensures that there is no redundant data sent throughout the network. The semantics of the meta-data format is application-specific and not specified in SPIN. For example, sensors might use their unique IDs to report meta-data if they cover a certain known region. In addition, SPIN[5] has access to the current energy level of the node and adapts the protocol it is running based on how much energy is remaining. These protocols work in a time-driven fashion and distribute the information all over the network, even when a user does not request any data. The SPIN family is designed to address the deficiencies of classic flooding by negotiation and resource adaptation. The SPIN family of protocols is designed based on two basic ideas:

1) Sensor nodes operate more efficiently and conserve energy by sending data that describe the sensor data instead of sending all the data; for example, image and sensor nodes must monitor the changes in their energy resources.
2) Conventional protocols like flooding or gossiping-based routing protocols [2] waste energy and bandwidth when sending extra and unnecessary copies of data by sensors covering overlapping areas.

The drawbacks of flooding include implosion, which is caused by duplicate messages sent to the same node, overlap when two nodes sensing the same region send similar packets to the same neighbor, and resource blindness in consuming large amounts of energy without consideration for energy constraints. Gossiping avoids the problem of implosion by just selecting a random node to which to send the packet rather than broadcasting the packet blindly. However, this causes delays in propagation of data through the nodes.

SPIN's meta-data negotiation solves the classic problems of flooding, thus achieving a lot of energy efficiency. **SPIN is a three-stage protocol as sensor nodes use three types of messages, ADV, REQ, and DATA, to communicate. ADV is used to advertise new data, REQ to request data, and DATA is the actual message itself.** The protocol starts when a SPIN node obtains new data it is willing to share. It does so by broadcasting an ADV message containing metadata. If a neighbor is interested in the data, it sends a REQ message for the DATA and the DATA is sent to this neighbor node. The neighbor sensor node then repeats this process with its neighbors. As a result, the entire sensor area will receive a copy of the data. The SPIN family of protocols includes many protocols. The main two are called SPIN-1 and SPIN-2; they incorporate negotiation before transmitting data in order to ensure that only useful information will be transferred. Also, each node has its own resource manager that keeps track of resource consumption and is polled by the nodes before data transmission. The SPIN-1 protocol is a three-stage protocol, as described above. An extension to SPIN-1 is SPIN-2, which incorporates a threshold-based resource awareness mechanism in addition to negotiation. When energy in the nodes is abundant, SPIN-2 communicates using the three-stage protocol of SPIN1. However, when the energy in a node starts approaching a low threshold, it reduces its participation in the protocol; that is, it participates only when it believes it can complete all the other stages of the protocol without going below the low energy threshold. In conclusion, SPIN-1 and SPIN-2 are simple protocols that efficiently disseminate data while maintaining no per-neighbor state. These protocols are well suited to an environment where the sensors are mobile because they base their forwarding decisions on local neighborhood information.

One of the advantages of SPIN is that topological changes are localized since each node need know only its single-hop neighbors. SPIN provides more energy savings than flooding, and metadata negotiation almost halves the redundant data. However, SPIN's data advertisement mechanism cannot guarantee delivery of data. To see this, consider the application of intrusion detection where data should be reliably reported over periodic intervals, and assume that nodes interested in the data are located far away from the source node, and the nodes between source and destination nodes are not interested in that data; such data will not be delivered to the destination at all.

	LEACH PROTOCOL	SPIN PROTOCOL
Classification	Hierarchical	Flat

Mobility	Fixed BS	Poss.
Position Awareness	No	No
Power usage	Max	Ltd
Negotiation Based	No	Yes
Data Aggregation	Yes	Yes
Localization	Yes	No
Query Based	No	Yes
State Complexity	CHs	Low
Scalability	Good	Ltd
Multipath	No	Yes

Table1. Theoretical comparison between LEACH and SPIN Protocol

3. Performance Testing

This section discusses simulation on energy performance using Network Simulator 2.34. The simulation primarily study on routing energy usage in SPIN and LEACH.SPIN is negotiation based data dissemination protocol suitable for wireless sensor networks. Thus, it assumes that all sensor nodes can be sinks potentially. Every node uses meta-data to name their data. By using this metadata, each node can negotiate whether to deliver data or not to eliminate the redundant data transmission throughout the network. In other words, every node can make its communication decisions based on negotiations with neighbour nodes about application-specific knowledge of the data and the resource available to it. This negotiation enables sensors to distribute data efficiently with limited energy. Basically, SPIN [7] uses resource-adaptive negotiation mechanism. Before any data is really transmitted, a sensor node performs negotiations by using its meta-data. These negotiations are done by exchanging a new data advertisement message (ADV) and a request for data message (REQ) between the sender and the receiver. After the negotiation, the sender transmits its data to the receiver (DATA).SPIN assures that there is no redundant data sent throughout the sensor network. In addition, SPIN checks the current energy level of each sensor node and adapts the protocol depending on how much energy remains. In SPIN simulation testing, there is some limitation. The nodes is being designed and linked in shortest path. The negotiations are done by exchanging a new data advertisement message (ADV) will display in blue color links and a request for data message (REQ) between the sender and the receiver will display in green color links. After the negotiation, the sender transmits its data to the receiver (DATA) where will be displayed in red color links. The event will be the packets that being transmits along the gradient path. Every transmitting packet to nodes, the links will be displayed in red color. In short, SPIN simulation tests differentiate those message elements with colors. Each colors presenting different element of message in SPIN routing scheme.

LEACH is a clustering based protocol that includes the randomized adaptive self configuring cluster formation. Localized control for data transfers .Low energy media access and Application specific data processing such as data aggregation. The operation of LEACH is separated into two phases, the **setup phase** and the **steady state** phase. In the setup phase, the clusters are organized and CHs are selected and rotates this role to evenly distribute the energy load among the sensors in the network. In LEACH, the cluster head (CH) nodes compress data arriving from nodes that belong to the respective cluster, and send an aggregated packet to the base station in order to reduce the amount of information that must be transmitted to the base station. LEACH uses a TDMA/CDMA MAC to reduce inter-cluster and intra-cluster collisions. However, data collection is centralized and is performed periodically. Therefore, this protocol is most appropriate when there is a need for constant monitoring by the sensor network. A user may not need all the data immediately. Hence, periodic data transmissions are unnecessary which may drain the limited energy of the sensor nodes. After a given interval of time, a randomized rotation of the role of the CH is conducted so that uniform energy dissipation in the sensor network is obtained. LEACH assumes that all nodes can transmit with enough power to reach the Base Station (BS) if needed and that each node has computational power to support different MAC protocols. Therefore, it is not applicable to networks deployed in large regions. It is not obvious how the number of the predetermined CHs (p) is going to be uniformly distributed through the network. Therefore, there is the possibility that the elected CHs will be concentrated in one part of the network. Hence, some nodes will not have any CHs in their vicinity. Furthermore, the idea of dynamic clustering brings extra overhead, e.g. head changes, advertisements etc., which may diminish the gain in energy consumption. Also, the protocol assumes that all nodes begin with the same amount of energy capacity in each

election round, assuming that being a CH consumes approximately the same amount of energy for each node. The protocol should be extended to account for non-uniform energy nodes, i.e., use energy-based threshold.

4. Simulation Results

This section discusses simulation result performed on SPIN This section discusses simulation result performed

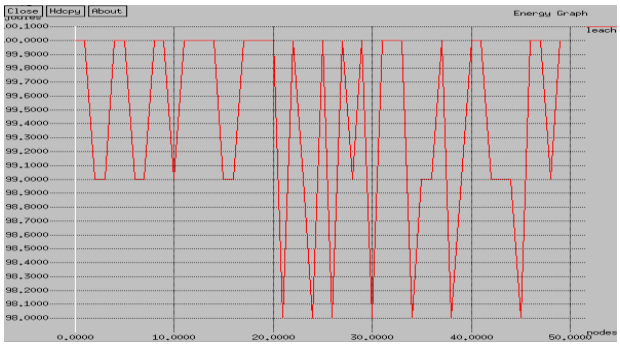


Fig 3 energy graph of LEACH

On SPIN and LEACH. There are fifty nodes being these protocols are tested on a number of factors. The main experiment is actually measure the energy of the network .When a node receives and transmit a data packet, the total energy will decrease and increase gradually. This simulation test network ran using 50 nodes for 2 minutes for both LEACH and SPIN. As observed from the below graph, SPIN fares much better than the LEACH protocol in terms of energy consumption. In LEACH as seen from the Fig 3, it will start with advertise its interest, and then waiting a request from any node before start transmitting data again. The energy of the network decreases rapidly and then increase and so on in 2 minutes. This is because the effective of transmitting and receiving data in the network. In LEACH protocol we have the limited energy. In it we have a given interval after a given interval of time; a randomized rotation of the role of the CH is conducted so that uniform energy dissipation in the sensor network is obtained. In LEACH protocol more energy is consumed because of head changes, advertisements etc.

In LEACH protocol more energy is consumed because of head changes, advertisements etc. In LEACH protocol the packet delivery ratio is more because of cluster heads. The end to end delay and dead nodes is more in LEACH because of randomized rotation of role of the cluster head as shown in figure and as the number of nodes the packet delivery ratio start decreasing.

As observed from the below graph in the LEACH end to end delay is more in LEACH Because of randomized rotation of role of the cluster head . In the LEACH protocol the packet delivery ratio is more because of cluster heads.As the number of nodes

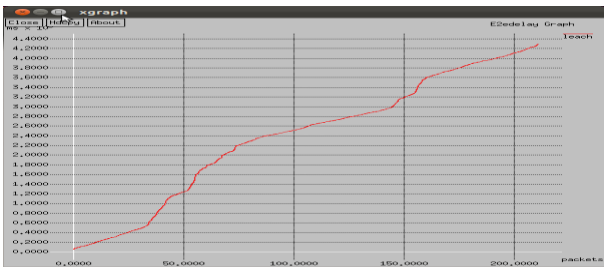


Fig 4 End to end delay graph of LEACH protocol

Increasing the packet delivery ratio start decreasing. The dead nodes in it is more due to randomized rotation of the protocol.

SPIN will start with advertise its interest, and then waiting a request from any node before start transmitting data again. SPIN
nodes negotiate with each other before transmitting data. Negotiation helps ensure that only useful information will be

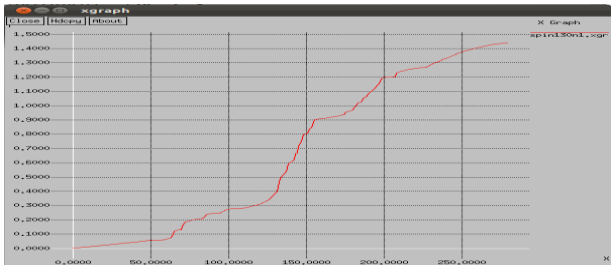


Fig 5 Energy graph of SPIN protocol

transferred. To negotiate successfully, however, nodes must be able to describe or name the data they observe. These descriptors used in SPIN negotiations are referred to as meta-data. From the result, it proves that meta-data negotiation keeps SPIN nodes from sending out even a single redundant data packet in a network .The result is based on average packet received by nodes. From Fig.5, the energy used by nodes dramatically decrease receiving packet from the first 20 seconds. This is because nodes that links to the shortest path nodes and after that the gradient links uses a lot of energy for transmitting and receiving packet. Thus, they generate overhead and reduce the life time of the nodes in the network. When this occurs, the topology and links for every node will change. The distance for for transmitting and receiving packet will be a bit less as compared to LEACH.

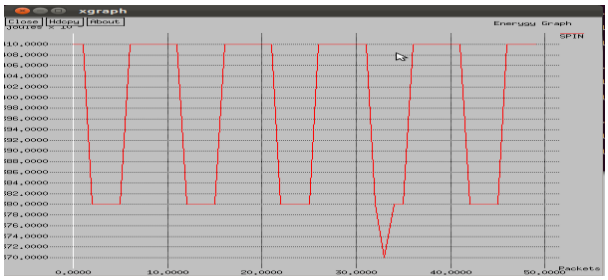


Fig 6 end to end Delay in SPIN protocol

The end to end delay in the SPIN protocol is same as that of LEACH for some starting packets but after that the delay in SPIN is less as shown in fig because of nodes link to shortest path end to end delay start increasing as number of nodes increases. SPIN operation will transmit almost zero redundant data packet and decrease the operation of sending wasted data packets.

5. Conclusion

Based on the study of these routing algorithms, it shows that some of the desirable features of a good energy efficient routing protocol for sensor network are:

If a routing algorithm can support multiple paths to a destination with low overhead, it could help in balancing the network load.

In SPIN and LEACH Protocol the LEACH has limited energy and it has the more energy consumption as compared to SPIN Protocol because of cluster head rotation .In LEACH Protocol after a given interval the cluster head are rotated and they also consume energy while rotating so it consume more energy where as SPIN uses less it do not have any cluster head .In it we first advertise the message than we send the data only those from which we get the request but this is only based on metadata negotiation only.

In LEACH the packet delivery ratio start is less as compared to SPIN .This is because of given interval of time in LEACH. In LEACH we have the limited time after that the transmission stop. But in SPIN no such time boundation so packet delivery ratio is large.

The end to end delay and dead nodes in LEACH is more as compared to SPIN. In starting end to end delay become same in both the cases after some interval there is difference .This is because the in LEACH there is cluster head rotation takes place so it have large end to end delay but in SPIN end to end delay is very less because it uses the Bellman Ford shortest path algorithm.

The dead nodes in LEACH Protocol are large because of cluster head rotation but SPIN has very less dead nodes because in it data transmission is based on the metadata negotiation of the nodes. So the LEACH protocol is most appropriate when there is a need for constant monitoring by the sensor network.

6. Acknowledgments

I take this opportunity to express my gratitude to all those people who extended their co-operation and played an important role in successful complementation of this thesis.

I express my sincere thanks to **Mrs. Sonia Juneja**, Head of IT Department and assistant professor in Computer Science & Engineering and Information Technology, for providing me invaluable support and able guidance throughout the thesis. She has been a source of inspiration for me that helped me in delivering the suggestions .Without whose support nothing could have been achieved.

References

[1] Jamal N. Al-Karaki, Ahmed E. Kamal,"ROUTING TECHNIQUES IN WIRELESS SENSOR NETWORKS: A SURVEY"1536-1284/04IEEE Wireless Communications , December 2004.

[2] Jian Wan,"A REVIEW OF ROUTING PROTOCOL IN WIRELESS SENSOR NETWORKS" IEEE Communications Magazine, vol.40, Issue:8, December 2008.

[3] Kazi Chandrima Rahman,"A Survey on Sensor Network", JCIT, ISSN 2078-5828 (PRINT), ISSN 2218-5224 (ONLINE), VOLUME 01, ISSUE 01,MANUSCRIPT CODE: 100715,2010.

[4] V.Vasanthi, "A Perspective analysis of routing protocol in wireless sensor network" (IJCSE) International Journal on Computer Science and Engineering Vol. 02.No. 08,2010,2511-2518.

[5] M. Bani Yassein ,"Improvement on LEACH Protocol of Wireless Sensor Network (VLEACH) "doi: 10.4156/jdcta.vol3.issue2.yassein.

[6] Amir Sepasi Zahmati et. al, "An Energy-Efficient Protocol with static Cluster based Routing in Wireless Sensor Networks World Academy of Science, Engineering and Technology 28 2007.

[7] Mahmood Ali**"** Real Time Support and Energy Efficiency in Wireless Sensor Networks"

[8] K. Padmanabhan, "A Study on Energy Efficient Routing Protocols in Wireless Sensor Networks" European Journal of Scientific Research ISSN 1450-216X Vol.60 No.4 (2011), pp. 499-511 © EuroJournals Publishing, Inc. 2011.

[9] A. H. Azni, Madihah Mohd Saudi, Azreen Azman, and Ariff Syah Johari,"Performance Analysis of Routing Protocol for WSN Using Data Centric Approach", World Academy of Science, Engineering and Technology 53 2009.

[10] Zeenat Rehena, Sarbani Roy, Nandini Mukherjee ,"A Modified SPIN for Wireless Sensor Networks", 978-1-4244 8953-4/11 2011 IEEE.

Shortest Path Finding Using Spatial Ranking

B.PADMAJA[1], R.SATEESH[2], K.DHANASREE[3]
Assistant Professor, DJRIET [1, 2], Associate Professor, DRKIST[3]

Abstract
The k nearest neighbor object to a point in space is the most regularly used query in finding shortest path of a given network. In this paper we present an efficient pruning method to find the nearest neighbor to a point for finding the shortest path. Finally we present the results of several experiments obtained using the implementation of our algorithm and examine the behavior of the metrics and scalability of the algorithm.

Keyword: *Spatial, Ranking, Nearest Neighbor, Shortest path, MBR, R-Tree.*

I. Introduction
The efficient implementation of Nearest Neighbor (NN) queries is of a particular interest in Geographic Information System (GIS). In this paper the shortest path in a network is obtained by finding the nearest neighbor of nearest neighbors.

Efficient processing of NN queries requires spatial data structure which capitalize on the proximity of the objects to focus the search of potential neighbors only. Finding the Nearest Neighbor of Nearest Neighbor has many applications:

I. In mobile environments, users do not have the accurate knowledge about their locations to specify the query points because all location identification methods have errors. Even if they have such knowledge, they may not want to expose these locations to the service providers for privacy reasons. RNN queries address these issues by allowing users to specify ranges rather than points for NN queries. They are particularly appealing to the large number of 2G/3G mobile subscribers whose devices are incapable of pinpointing locations. While these devices cannot support conventional NN queries, they can issue RNN queries through text messages such as "find the nearest hotels to the City Park?"

II. A user may continuously ask for nearest neighbors while moving around. It is inefficient to submit many PNN queries individually to the server. A better alternative is to submit a single RNN query around the current location to fetch all possible nearest neighbors for this area. Any PNN query issued in this area is then processed locally by a nearest-neighbor search in the prefetched set, saving both computation and communication costs

Section 2 of the paper contains the theoretical foundation for the shortest path finding using nearest neighbor search. Section 3 describes the algorithm for ordering the search and pruning during it. Section 4 has the experiments with the implementation of the algorithm.

II. Shortest Path Finding

USING R-TREES
As with Quad Tree the region is divided into successively smaller rectangles (MBRs). Rectangles need not be of the same size or number at each level. Rectangles may actually overlap. Lowest level cell has only one object, Tree maintenance algorithms similar to those for B-trees.

Leaf nodes of the R-Tree contain entries of the form (RECT, oid) where oid is an object identifier and is used as a pointer to a data object and RECT is an n-dimensional Minimal Bounding Rectangle (MBR) which bounds the corresponding object.

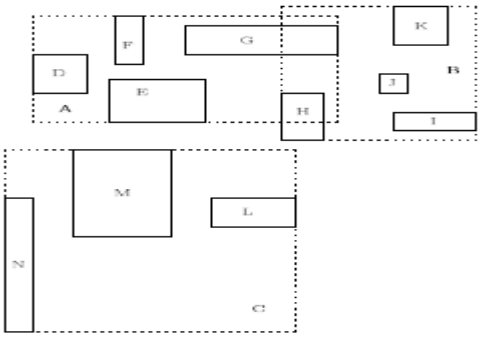


Fig 1: Collection of Rectangles

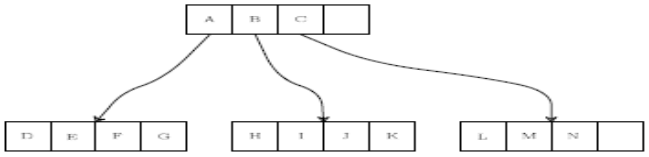


Fig 2: R-Tree Construction

Every MBR have 4 corners which are as shown below

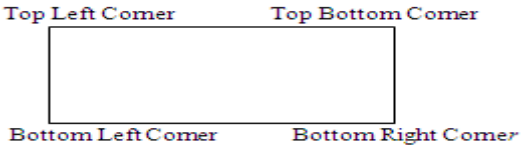


Fig 3: MBR Corners

Take Two MBR's from the source node as Bottom Right corner in both left and right direction as shown in the following figure Find the nearest neighbor of source in both left and right MBR's and find out the distance between left MBR nearest neighbor and destination node (assumed as ϵ_1), the right MBR nearest neighbor and destination node (assumed as ϵ_2). The smallest ϵ value nearest neighbor is changed as the source, the above procedure repeats until destination is reached.

When the destination reaches all the set of sources traversed becomes the shortest path.

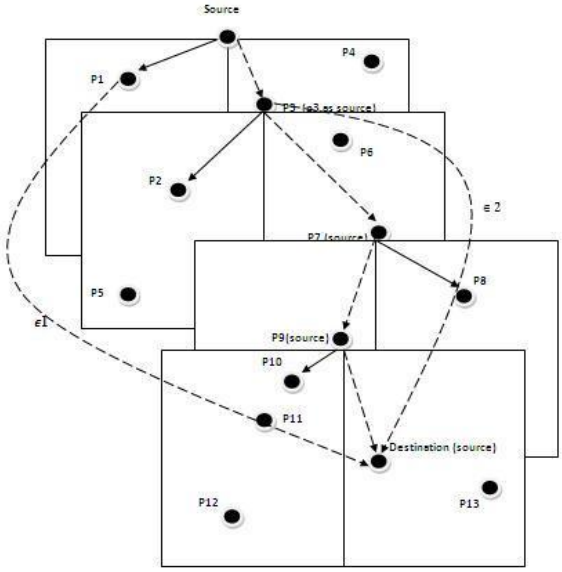


Fig 4: Shortest Path Finding

As in the above figure the Left MBR nearest neighbor to source is node P1 and the Right MBR nearest neighbor to source is node P3. Take the distance between p1 to destination as ϵ_1(Assume it is 25 units) and the distance between p2 to destination as ϵ_2 (Assume it as 20 units). Since ϵ_2 is less than ϵ_1 take the Right MBR's nearest neighbor as source i.e point p3. Now take the source as p3, take two MBR's from the bottom left and bottom right corners and repeat the above procedure. The list of all nodes taken as sources between the source node and destination nodes is the shortest path.

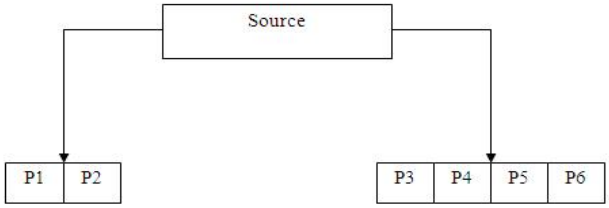


Fig 4(a)

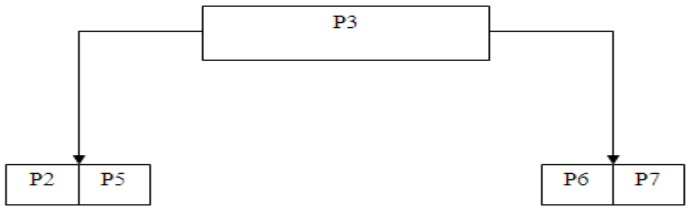


Fig 4 (b)

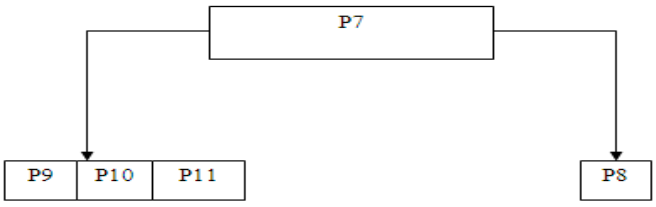


Fig 4 (c)

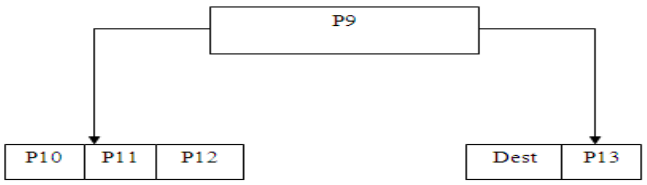


Fig 4 (d)

Above Figures shows both the Left MBR and Right MBR's Nodes.

In Figure 4(a) the root is taken as source node and the Left sub tree is taken as all the Left MBR nodes, the Right sub tree is taken as all the Right MBR nodes. The nearest neighbor to the destination in both the Left sub tree and Right sub tree of Figure 4(a) is taken as the root for the Figure 4(b). This procedure is repeated until the nearest neighbor for the root is Destination. The collection of all the root nodes becomes the shortest path.

III. Shortest Path Finding Algorithm Using R-Trees

Pseudo code for shortest path finding

Algorithm: Shortestpath_find (Source, LeftMBR, RightMBR, Set V, Value ϵ)

L1: Insert Source as root into Hc //Hc refers Min Heap
L2: Nearest Nearest
L3: int dist
L 4: **for** each left entry e' of Source **do**
L 5: **if** CN is a non-leaf node **then**
L 6: **if** $\exists e'$ Є V, mindist(e',source) ≤ ϵ **then**
L 7: dist:=objectDIST(e',Source)
L 8: **if**(dist<Nearest.dist) **then**
L 9: Nearest.dist=dist
L 10: insert e' into H α //H α refers Min Heap
L 11: **else then**
L 12: **Perform Pruning**
L 13: **for** each right entry e'' of Source **do**
L 14: **if** CN is a non-leaf node **then**
L 15: **if** $\exists e''$ Є V, mindist(e'',source) ≤ ϵ **then**
L 16: dist:=objectDIST(e'',Source)
L 17: **if**(dist<Nearest.dist) **then**
L 18: Nearest.dist=dist
L 19: Insert e'' into Hβ //Hβ refers Min Heap
L 20: **else then**
L 21: **Perform Pruning**
L 22: **while** |v| >0 and there exists a non-empty heap H α **do**
L 23: deheap an entry e' from Hα into LNR (Left Nearest Neighbor)
L 24: **while** |v| >0 and there exists a non-empty eap Hβ **do**
L 25: deheap an entry e'' from H β into RNR (Right Nearest Neighbor)
L26: ϵ_1=**dist(**LNR,destination)
L27: ϵ_2=**dist(**RNR,destination)
L28: **if** $\epsilon_1 < \epsilon_2$ **then**
L29: Change the source as LNR andCall Shortestpath_find (LNR, LeftMBR , RightMBR ,Set V,Value ϵ)
L30: **else then**
L31: Change the source as RNR and Call Shortestpath_find (RNR, LeftMBR , RightMBR ,Set V, Value ϵ)

In this paper, we assume that the object dataset is indexed by an R-tree and each feature dataset is indexed by an MIN R-tree, where each where each non-leaf entry augments the minimum quality (of features) in its sub tree.

The above algorithm shows the procedure for finding the shortest spatial path in networks. In this algorithm the Lines 1-21 are used to take three Min Heaps are used to store the Source Node, its Left MBR nodes, its Right MBR nodes respectively as Hc, H α, Hβ . In H α and Hβ all the nodes of Left MBR are stored in order of minimum distance from the source. All the nodes which are not in the Left and Right MBR's are pruned.

The functions objectDIST and dist are used to calculate the Euclidean distance between the source to destination (or) source to a specific point in MBR. Lines 22-25 shows that the deheap node from H α becomes the nearest neighbor in its Left MBR (i.e LNR), the deheap node from the H β becomes the nearest neighbor in its Right MBR (i.e RNR)

Lines 26-31 shows that the distance between the LNR, RNR and Destination node are taken into respectively ϵ_1 andϵ_2. If $\epsilon 1$ is smaller than $\epsilon 2$ then the source is taken as LNR otherwise the source is taken as RNR for the next recursive call.

IV. Experiment Results

The shortest path is finding by taking two MBR's as any one of the corner based upon the direction of destination from the source. When the destination is below or right to the source as shown in the following figure (5) the Two MBR's are taken as Top Left corner and Top Right corner. If the destination is above or left to the source then the Two MBR's are taken as Bottom Left and Bottom Right corner

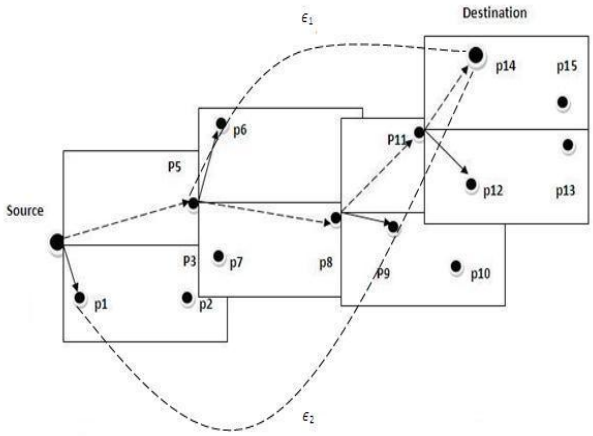


Fig (5)

Fig (6)

Fig (6) shows the results of an experiment using the data of Fig (4). From the experimental behavior, we can make two observations. First, in the Traditional system the total number of neighbors accessed to find the shortest path is at constant rate of total number of neighbor nodes. That is at any distance we need to check all the neighbors. Second, in our proposed system the total number of neighbors accessed to find the shortest path is at the rate of 20%. In figure 4, traversed two neighbors in the first 10 units, six neighbor nodes for 20, 30, 40 units, eight neighbor nodes for 50 units, traversed nine neighbor nodes for 60 and 70 units, 11neighbor nodes for 80 units. That is at a specific distance we need to check the Left MBR and Right MBR neighbors only, the remaining neighbor nodes are pruned.

V. Conclusion

In this paper, we developed a Shortest Path finding Algorithm which finds the k Nearest Neighbors of a given source point. We also introduced four corners of MBR that can be used to guide an ordered spatial search. We implemented and thoroughly tested our Shortest Path finding algorithm using k-Nearest Neighbor spatial search. The experiments on both real data sets showed that the algorithm scales up well with respect to both the number of NN requested and with size of data sets.

Further research on shortest path finding using spatial queries will focus on defining and analyzing other metrics and characterization of our algorithm in three space environment.

References

[1] Nearest Neighbor Queries by Nick Roussopoulos Stephen Kelly Frederic Vincent.
[2] Ranking Spatial Data by Quality Preferences Man Lung Yiu, Hua Lu, Nikos Mamoulis, and Michail Vaitis.
[3] Range Nearest-Neighbor Query Haibo Hu and Dik Lun Lee
[4] M. L. Yiu, X. Dai, N. Mamoulis, and M. Vaitis, "Top-k Spatial Preference Queries," in ICDE, 2007.
[5] N. Bruno, L. Gravano, and A. Marian, "Evaluating Top-k Queries over Web-accessible Databases," in ICDE, 2002.
[6] A. Guttman, "R-Trees: A Dynamic Index Structure for Spatial Searching," in SIGMOD, 1984.
[7] G. R. Hjaltason and H. Samet, "Distance Browsing in Spatial Databases," TODS, vol. 24(2), pp. 265–318, 1999.
[8] R. Weber, H.-J. Schek, and S. Blott, "A quantitative analysis and performance study for similarity-search methods in high dimensional spaces." in VLDB, 1998.

Design of Test Data Compressor/Decompressor Using Xmatchpro Method

C. Suneetha*[1],V.V.S.V.S.Ramachandram*[2]

[1]Research Scholar, Dept. of E.C.E., Pragathi Engineering College, JNTU-K, A.P. , India [2] Associate Professor, Dept.,of E.C.E., Pragathi Engineering College, E.G.Dist., A.P., India

Abstract

Higher Circuit Densities in system-on-chip (SOC) designs have led to drastic increase in test data volume. Larger Test Data size demands not only higher memory requirements, but also an increase in testing time. Test Data Compression/Decompression addresses this problem by reducing the test data volume without affecting the overall system performance. This paper presented an algorithm, XMatchPro Algorithm that combines the advantages of dictionary-based and bit mask-based techniques. Our test compression technique used the dictionary and bit mask selection methods to significantly reduce the testing time and memory requirements. We have applied our algorithm on various benchmarks and compared our results with existing test compression/Decompression techniques. Our approach results compression efficiency of 92%. Our approach also generates up to 90% improvement in decompression efficiency compared to other techniques without introducing any additional performance or area overhead.

Keywords **-** XMatchPro, decompression, test compression, dictionary selection, SOC.

I. INTRODUCTION

1.1. Objective

With the increase in silicon densities, it is becoming feasible for compression systems to be implemented in chip. A system with distributed memory architecture is based on having data compression and decompression engines working independently on different data at the same time. This data is stored in memory distributed to each processor. The objective of the project is to design a lossless data compression system which operates in high-speed to achieve high compression rate. By using the architecture of compressors, the data compression rates are significantly improved. Also inherent scalability of architecture is possible.The main parts of the system are the Xmatchpro based data compressors and the control blocks providing control signals for the Data compressors allowing appropriate control of the routing of data into and from the system. Each Data compressor can process four bytes of data

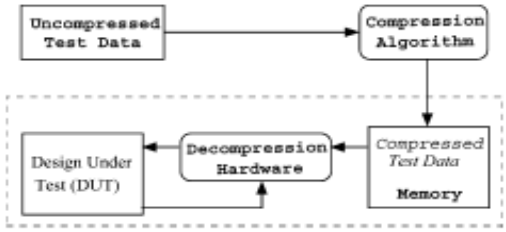


Fig. 1. Test data compression methodology.

into and from a block of data in every clock cycle. The data entering the system needs to be clocked in at a rate of 4 bytes in every clock cycle. This is to ensure that adequate data is present for all compressors to process rather than being in an idle state.

1.2. Goal of the Thesis

To achieve higher decompression rates using compression/decompression architecture with least increase in latency.

1.3. Compression Techniques

At present there is an insatiable demand for ever-greater bandwidth in communication networks and forever-greater storage capacity in computer system. This led to the need for an efficient compression technique. The compression is the process that is required either to reduce the volume of information to be transmitted – text, fax and images or reduce the bandwidth that is required for its transmission – speech, audio and video. The compression technique is first applied to the source information prior to its transmission. Compression algorithms can be classified in to two types, namely

- Lossless Compression
- Lossy Compression

1.3.1. Lossless Compression

In this type of lossless compression algorithm, the aim is to reduce the amount of source information to be transmitted in such a way that, when the

compressed information is decompressed, there is no loss of information. Lossless compression is said, therefore, to be reversible. i.e., Data is not altered or lost in the process of compression or decompression. Decompression generates an exact replica of the original object. The Various lossless Compression Techniques are,

- Packbits encoding
- CCITT Group 3 1D
- CCITT Group 3 2D
- Lempel-Ziv and Welch algorithm LZW
- Huffman
- Arithmetic

Example applications of lossless compression are transferring data over a network as a text file since, in such applications, it is normally imperative that no part of the source information is lost during either the compression or decompression operations and file storage systems (tapes, hard disk drives, solid state storage, file servers) and communication networks (LAN, WAN, wireless).

1.3.2 Lossy Compression

The aim of the Lossy compression algorithms is normally not to reproduce an exact copy of the source information after decompression but rather a version of it that is perceived by the recipient as a true copy. The Lossy compression algorithms are:

- JPEG (Joint Photographic Expert Group)
- MPEG (Moving Picture Experts Group)
- CCITT H.261 (Px64)

Example applications of lossy compression are the transfer of digitized images and audio and video streams. In such cases, the sensitivity of the human eye or ear is such that any fine details that may be missing from the original source signal after decompression are not detectable.

1.3.3 Text Compression

There are three different types of text – unformatted, formatted and hypertext and all are represented as strings of characters selected from a defined set. The compression algorithm associated with text must be lossless since the loss of just a single character could modify the meaning of a complete string. The text compression is restricted to the use of entropy encoding and in practice, statistical encoding methods. There are two types of statistical encoding methods which are used with text: one which uses single character as the basis of deriving an optimum set of code words and the other which uses variable length strings of characters. Two examples of the former are the Huffman and Arithmetic coding algorithms and an example of the latter is Lempel-Ziv (LZ) algorithm. The majority of work on hardware approaches to lossless data compression has used an adapted form of the dictionary-based Lempel-Ziv algorithm, in which a large number of simple processing elements are arranged in a systolic array [1], [2], [3], [4].

II. Previous Work On Lossless Compression Methods

A second Lempel-Ziv method used a content addressable memory (CAM) capable of performing a complete dictionary search in one clock cycle [5], [6], [7]. The search for the most common string in the dictionary (normally, the most computationally expensive operation in the Lempel-Ziv algorithm) can be performed by the CAM in a single clock cycle, while the systolic array method uses a much slower deep pipelining technique to implement its dictionary search. However, compared to the CAM solution, the systolic array method has advantages in terms of reduced hardware costs and lower power consumption, which may be more important criteria in some situations than having faster dictionary searching. In [8], the authors show that hardware main memory data compression is both feasible and worthwhile. The authors also describe the design and implementation of a novel compression method, the XMatchPro algorithm. The authors exhibit the substantial impact such memory compression has on overall system performance. The adaptation of compression code for parallel implementation is investigated by Jiang and Jones [9]. They recommended the use of a processing array arranged in a tree-like structure. Although compression can be implemented in this manner, the implementation of the decompressor's search and decode stages in hardware would greatly increase the complexity of the design and it is likely that these aspects would need to be implemented sequentially. An FPGA implementation of a binary arithmetic coding architecture that is able to process 8 bits per clock cycle compared to the standard 1 bit per cycle is described by Stefo et al [10]. Although little research has been performed on architectures involving several independent compression units working in a concurrent cooperative manner, IBM has introduced the MXT chip [11], which has four independent compression engines operating on a shared memory area. The four Lempel-Ziv compression engines are used to provide data throughput sufficient for memory compression in computer servers. Adaptation of software compression algorithms to make use of multiple CPU systems was demonstrated by research of Penhorn [12] and Simpson and Sabharwal [13].

Penhorn used two CPUs to compress data using a technique based on the Lempel-Ziv algorithm and showed that useful compression rate improvements can be achieved, but only at the cost of increasing the learning time for the dictionary. Simpson and Sabharwal described the software implementation of compression system for a multiprocessor system based on the parallel architecture developed by Gonzalez and Smith and Storer [14].

Statistical Methods

Statistical Modeling of lossless data compression system is based on assigning values to events depending on their probability. The higher the value, the higher the probability. The accuracy with which this frequency distribution reflects reality determines the efficiency of the model. In Markov modeling, predictions are done based on the symbols that precede the current symbol. Statistical Methods in hardware are restricted to simple higher order modeling using binary alphabets that limits speed, or simple multi symbol alphabets using zeroeth order models that limits compression. Binary alphabets limit speed because only a few bits (typically a single bit) are processed in each cycle while zeroeth order models limit compression because they can only provide an inexact representation of the statistical properties of the data source.

Dictionary Methods

Dictionary Methods try to replace a symbol or group of symbols by a dictionary location code. Some dictionary-based techniques use simple uniform binary codes to process the information supplied. Both software and hardware based dictionary models achieve good throughput and competitive compression. The UNIX utility 'compress' uses Lempel-Ziv-2 (LZ2) algorithm and the data compression Lempel-Ziv (DCLZ) family of compressors initially invented by Hewlett-Packard[16] and currently being developed by AHA[17],[18] also use LZ2 derivatives. Bunton and Borriello present another LZ2 implementation in [19] that improves on the Data Compression Lempel-Ziv method. It uses a tag attached to each dictionary location to identify which node should be eliminated once the dictionary becomes full.

XMatchPro Based System

The Lossless data compression system is a derivative of the XMatchPro Algorithm which originates from previous research of the authors [15] and advances in FPGA technology. The flexibility provided by using this technology is of great interest since the chip can be adapted to the requirements of a particular application easily. The drawbacks of some of the previous methods are overcome by using the XmatchPro algorithm in design. The objective is then to obtain better compression ratios and still maintain a high throughput so that the compression/decompression processes do not slow the original system down.

Usage of XMatchPro Algorithm

The Lossless Data Compression system designed uses the XMatchPro Algorithm. The XMatchPro algorithm uses a fixed-width dictionary of previously seen data and attempts to match the current data element with a match in the dictionary. It works by taking a 4-byte word and trying to match or partially match this word with past data. This past data is stored in a dictionary, which is constructed from a content addressable memory. As each entry is 4 bytes wide, several types of matches are possible. If all the bytes do not match with any data present in the dictionary they are transmitted with an additional miss bit. If all the bytes are matched then the match location and match type is coded and transmitted, this match is then moved to the front of the dictionary. The dictionary is maintained using a move to front strategy whereby a new tuple is placed at the front of the dictionary while the rest move down one position. When the dictionary becomes full the tuple placed in the last position is discarded leaving space for a new one. The coding function for a match is required to code several fields as follows:

A zero followed by:

1). Match location: It uses the binary code associated to the matching location. 2). Match type: Indicates which bytes of the incoming tuple have matched. 3). Characters that did not match transmitted in literal form. A description of the XMatchPro algorithm in pseudo-code is given in the figure below.

Pseudo Code for XMatchPro Algorithm

With the increase in silicon densities, it is becoming feasible for XMatchPros to be implemented on a single chip. A system with distributed memory architecture is based on having data compression and decompression engines working independently on different data at the same time.This data is stored in memory distributed to each processor. There are several approaches in which data can be routed to and from the compressors that will affect the speed, compression and complexity of the system. Lossless compression removes redundant information from the data while they are transmitted or before they are stored in memory. Lossless decompression reintroduces the redundant information to recover fully the original data. There are two important contributions made by the current

compression & decompression work, namely, improved compression rates and the inherent scalability. Significant improvements in data compression rates have been achieved by sharing the computational requirement between compressors without significantly compromising the contribution made by individual compressors. The scalability feature permits future bandwidth or storage demands to be met by adding additional compression engines.

The XMatchPro based Compression system

The XMatchPro algorithm uses a fixed width dictionary of previously seen data and attempts to match the current data element with a match in the dictionary. It works by taking a 4-byte word and trying to match this word with past data. This past data is stored in a dictionary, which is constructed from a content addressable memory. Initially all the entries in the dictionary are empty & 4-bytes are added to the front of the dictionary, while the rest move one position down if a full match has not occurred. The larger the dictionary, the greater the number of address bits needed to identify each memory location, reducing compression performance. Since the number of bits needed to code each location address is a function of the dictionary size greater compression is obtained in comparison to the case where a fixed size dictionary uses fixed address codes for a partially full dictionary.In the XMatchPro system, the data stream to be compressed enters the compression system, which is then partitioned and routed to the compressors.

The Main Component- Content Addressable Memory

Dictionary based schemes copy repetitive or redundant data into a lookup table (such as CAM) and output the dictionary address as a code to replace the data.

The compression architecture is based around a block of CAM to realize the dictionary. This is necessary since the search operation must be done in parallel in all the entries in the dictionary to allow high and data-independent

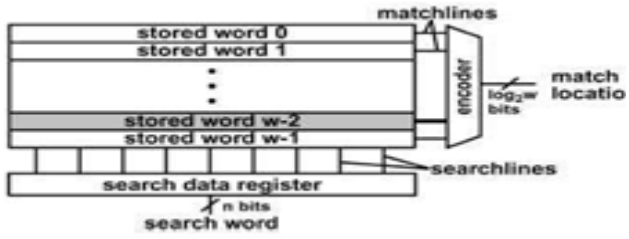


Fig..Conceptual view of CAM

The number of bits in a CAM word is usually large, with existing implementations ranging from 36 to 144 bits. A typical CAM employs a table size ranging between a few hundred entries to 32K entries, corresponding to an address space ranging from 7 bits to 15 bits. The length of the CAM varies with three possible values of 16, 32 or 64 tuples trading complexity for compression. The no. of tuples present in the dictionary has an important effect on compression. In principle, the larger the dictionary the higher the probability of having a match and improving compression. On the other hand, a bigger dictionary uses more bits to code its locations degrading compression when processing small data blocks that only use a fraction of the dictionary length available. The width of the CAM is fixed with 4bytes/word. Content Addressable Memory (CAM) compares input search data against a table of stored data, and returns the address of the matching data. CAMs have a single clock cycle throughput making them faster than other hardware and software-based search systems. The input to the system is the search word that is broadcast onto the searchlines to the table of stored data. Each stored word has a matchline that indicates whether the search word and stored word are identical (the match case) or are different (a mismatch case, or miss). The matchlines are fed to an encoder that generates a binary match location corresponding to the matchline that is in the match state. An encoder is used in systems where only a single match is expected.

The overall function of a CAM is to take a search word and return the matching memory location.

Managing Dictionary entries

Since the initialization of a compression CAM sets all words to zero, a possible input word formed by zeros will generate multiple full matches in different locations.The Xmatchpro compression system simply selects the full match closer to the top. This operational mode initializes the dictionary to a state where all the words with location address bigger than zero are declared invalid without the need for extra logic.

Iii. Xmatchpro Lossless Compression System

DESIGN METHODOLOGY

The XMatchPro algorithm is efficient at compressing the small blocks of data necessary with cache and page based memory hierarchies found in computer systems. It is suitable for high performance hardware implementation. The XMatchPro hardware achieves a throughput 2-3 times greater than other high-performance hardware implementation. The core component of the system is the XMatchPro based

Compression/ Decompression system. The XMatchPro is a high-speed lossless dictionary based data compressor. The XMatchPro algorithm works by taking an incoming four-byte tuple of data and attempting to match fully or partially match the tuple with the past data.

FUNCTIONAL DESCRIPTION

The XMatchPro algorithm maintains a dictionary of data previously seen and attempts to match the current data element with an entry in the dictionary, replacing it with a shorter code referencing the match location. Data elements that do not produce a match are transmitted in full (literally) prefixed by a single bit. Each data element is exactly 4 bytes in width and is referred to as tuple. This feature gives a guaranteed input data rate during compression and thus also guaranteed data rates during decompression, irrespective of the data mix. Also the 4-byte tuple size gives an inherently higher throughput than other algorithms, which tend to operate on a byte stream. The dictionary is maintained using move to front strategy, where by the current tuple is placed at the front of the dictionary and the other tuples move down by one location as necessary to make space. The move to front strategy aims to exploit locality in the input data. If the dictionary becomes full, the tuple occupying the last location is simply discarded.

A full match occurs when all characters in the incoming tuple fully match a dictionary entry. A partial match occurs when at least any two of the characters in the incoming tuple match exactly with a dictionary entry, with the characters that do not match being transmitted literally.

The use of partial matching improves the compression ratio when compared with allowing only 4 byte matches, but still maintains high throughput. If neither a full nor partial match occurs, then a miss is registered and a single miss bit of '1' is transmitted followed by the tuple itself in literal form. The only exception to this is the first tuple in any compression operation, which will always generate a miss as the dictionary begins in an empty state. In this case no miss bit is required to prefix the tuple.

At the beginning of each compression operation, the dictionary size is reset to zero. The dictionary then grows by one location for each incoming tuple being placed at the

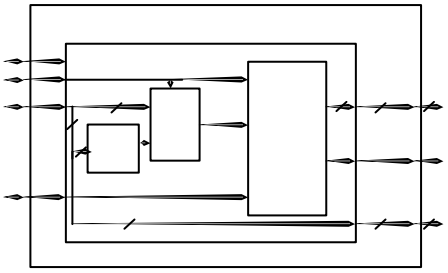
front of the dictionary and all other entries in the dictionary moving down by one location. A full match does not grow the dictionary, but the move-to-front rule is still applied. This growth of the dictionary means that code words are short during the early stages of compressing a block. Because the XMatchPro algorithm allows partial matches, a decision must be made about which of the locations provides the best overall match, with the selection criteria being the shortest possible number of output bits.

Implementation of Xmatchpro Based Compressor

The block diagram gives the details about the components of a single 32 bit Compressor. There are three components namely, COMPARATOR, ARRAY, CAMCOMPARATOR. The comparator is used to compare two 32-bit data and to set or reset the output bit as 1 for equal and 0 for unequal. The CAM COMPARATOR searches the CAM dictionary entries for a full match of the input data given. The reason for choosing a full match is to get a prototype of the high throughout Xmatchpro compressor with reduced complexity and high performance.

If a full match occurs, the match-hit signal is generated and the corresponding match location is given as output by the CAM Comparator.. If no full match occurs, the corresponding data that is given as input at the given time is got as output.

32 BIT COMPRESSIONS



Array is of length of 64X32 bit locations. This is used to store the unmatched incoming data and when a new data comes, the incoming data is compared with all the data stored in this array. If a match occurs, the corresponding match location is sent as output else the incoming data is stored in next free location of the array & is sent as output. The last component is the cam comparator and is used to send the match location of the CAM dictionary as output if a match has occurred. This is done by getting match information as input from the comparator.

Suppose the output of the comparator goes high for any input, the match is found and the corresponding address is retrieved and sent as output along with one bit to indicate that match is found. At the same time, suppose no match occurs, or no matched data is found, the incoming data is stored in the array and it is

sent as the output. These are the functions of the three components of the Compressor

IV.Design of Lossless Compression/Decompression System

DESIGN OF COMPRESSOR / DECOMPRESSOR

The block diagram gives the details about the components of a single 32-bit compressor / decompressor. There are three components namely COMPRESSOR, DECOMPRESSOR and CONTROL.

The compressor has the following components - COMPARATOR, ARRAY, and CAMCOMPARATOR.

The comparator is used to compare two 32-bit data and to set or reset the output bit as 1 for equal and 0 for unequal.

Array is of length of 64X32bit locations. This is used to store the unmatched in coming data and when the next new data comes, that data is compared with all the data stored in this array. If the incoming data matches with any of the data stored in array, the Comparator generates a match signal and sends it to Cam Comparator.

The last component is the Cam comparator and is used to send the incoming data and all the stored data in array one by one to the comparator. Suppose output of comparator goes high for any input, then the match is found and the corresponding address (match location) is retrieved and sent as output along with one bit to indicate the match is found. At the same time, suppose no match is found, then the incoming data stored in the array is sent as output. These are the functions of the three components of the XMatchPro based compressor.

The decompressor has the following components – Array and Processing Unit.

Array has the same function as that of the array unit which is used in the Compressor. It is also of the same length. Processing unit checks the incoming match hit data and if it is 0, it indicates that the data is not present in the Array, so it stores the data in the Array and if the match hit data is 1, it indicates the data is present in the Array, then it instructs to find the data from the Array with the help of the address input and sends as output to the data out.

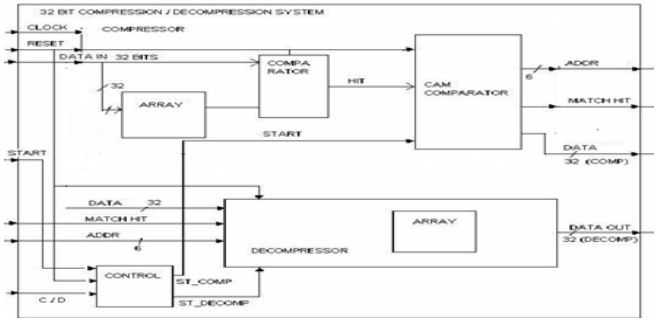


Fig. Block Diagram of 32 bit Compressor/Decompression

The Control has the input bit called C / D i.e., Compression / Decompression Indicates whether compression or decompression has to be done. If it has the value 0 then omcpressor is stared when the value is 1 decompression is done

V. Result Analysis

1 SYNTHESIS REPORT

Release 8.2i - xst I.31

Copyright (c) 1995-2006 Xilinx, Inc. All rights reserved.

--> Parameter TMPDIR set to ./xst/projnav.tmp

CPU : 0.00 / 0.17 s | Elapsed : 0.00 / 0.00 s

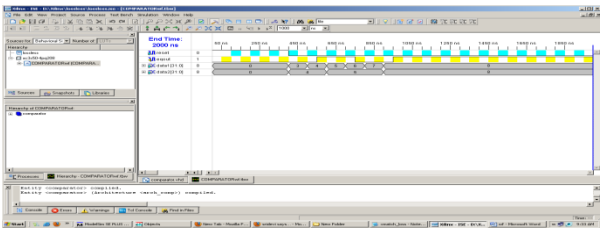
--> Parameter xsthdpdir set to ./xst

CPU : 0.00 / 0.17 s | Elapsed : 0.00 / 0.00 s

--> Reading design: ds.prj

Table of Contents

COMPARATOR



COMPARATOR waveform explanation

Whenever the reset signal is active low then only the data is compared otherwise i.e. if it is active high the eqout signal is zero. Here we are applying 3 4 5 6 7 8 as data1 and 4 5 8 as data2. So after comparing the inputs the output signal eqout raises at the data 4 and 8 instants which indicate the output signal.

COMPARATOR function

Suppose the output of the comparator goes high for any input, the match is found and the corresponding address is retrieved and sent as output along with one bit to indicate that match is found. At the same time, suppose no match occurs, or no matched data is found, the incoming data is stored in the array and it is sent as the output.

CAM

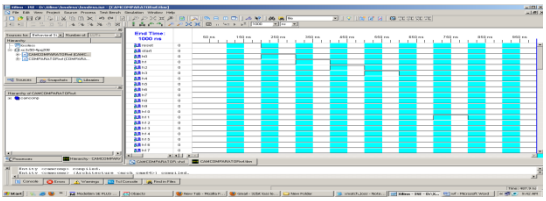


Fig. CAM data input Waveform

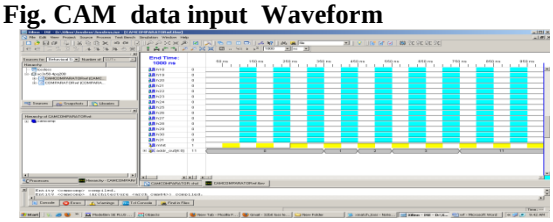


Fig. CAM data output Waveform

CAM waveform explanation

Whenever the reset signal is active low then only the cam produces the corresponding address and the match hit signal by individually raise the corresponding hit otherwise i.e. if it is active high it doesn't produces any signal. In order to obtain the outputs the start should be in active low state.

CAM function

The cam is used to send the match location Of the CAM dictionary as output if a match has occurred. This is done by getting match Information as input from the comparator.

Compression

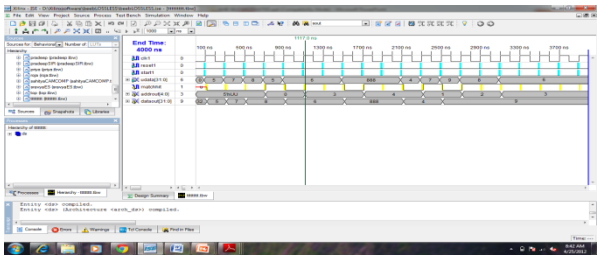


Fig. Compression of INPUT DATA waveform

Compression waveform explanation

Whenever the reset1 and start1 signal is active low then only the compression of data is occurred. Here every signal reacts at the clk1 rising edge only because it is meant for positive edge triggering. The data which has to be compressed is applied at the udata signal and the resultant data is obtained at the dataout signal which is nothing but compressed data. Whenever there is a redundancy data is there in the applied data then the matchhit signal will be in active high state by generating the corresponding address signal addrout.

Compression function

The compression engine main intension is to compress the incoming content for memory reduction purpose and to transfer the data very easily.

Decompression

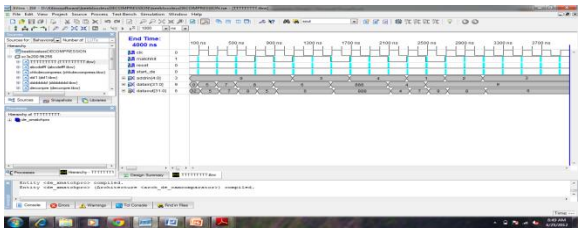


Fig. Decompression of the data

Decompression waveform explanation

Whenever the reset and start_de signal is active low then only the decompression of data is occurred. Here every signal reacts at the clk1 rising edge only because it is meant for positive edge triggering. Here for retrieving the original data the outputs of the compressor engine has to be applied as inputs for the decompression engine.

Decompression function

The decompression engine main intension is to retrieve the original content.

SCHEMATIC

Compression

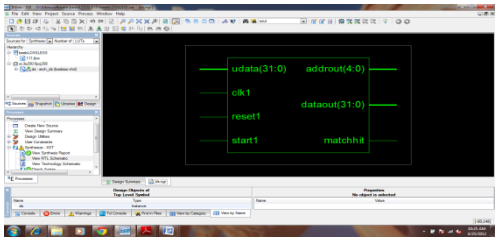


Fig. RTL Schematic of compression waveform of component level

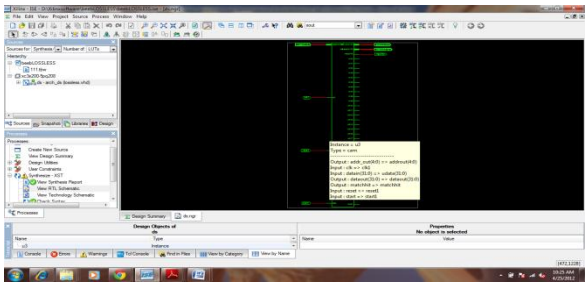


Fig. RTL Schematic of compression waveform of circuit level

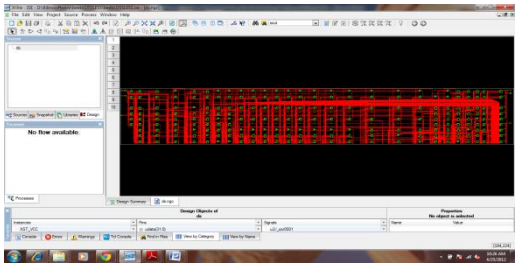


Fig.RTL Schematic of compression waveform of chip level

Decompression

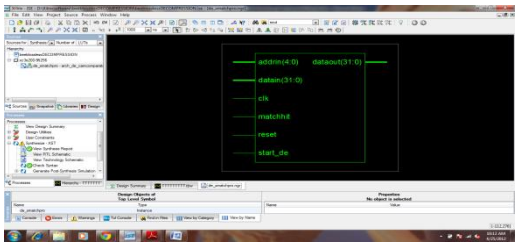


Fig RTL Schematic of decompression waveform of component level

Fig. RTL Schematic of decompression waveform of circuit level

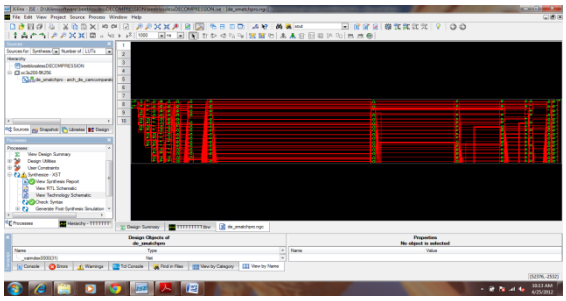


Fig. RTL Schematic of decompression waveform of chip level

VI.CONCLUSION

The various modules are designed and coded using VHDL. The source codes are simulated and the various waveforms are obtained for all the modules. Since the Compression/Decompression system uses XMatchPro algorithm, speed of compression throughput is high.

The Improved Compression ratio is achieved in Compression architecture with least increase in latency. The High speed throughput is achieved. The architecture provides inherent scalability in future.

The total time required to transmit compressed data is less than that of

Transmitting uncompressed data. This can lead to a performance benefit, as the bandwidth of a link appears greater when transmitting compressed data and hence more data can be transmitted in a given

amount of time.

Higher Circuit Densities in system-on-chip (SOC) designs have led to drastic increase in test data volume. Larger Test Data size demands not only higher memory requirements, but also an increase in testing time. Test Data Compression/Decompression addresses this problem by reducing the test data volume without affecting the overall system performance. This paper presented an algorithm, XMatchPro Algorithm that combines the advantages of dictionary-based and bit mask-based techniques. Our test compression technique used the dictionary and bit mask selection methods to significantly reduce the testing time and memory requirements. We have applied our algorithm on various benchmarks and compared our results with existing test compression/Decompression techniques. Our approach results compression efficiency of 92%. Our approach also generates up to 90% improvement in decompression efficiency compared to other techniques without introducing any additional performance or area overhead.

Vii. Future Scope

There is a potential of doubling the performance of storage / communication system by increasing the available transmission bandwidth and data capacity with minimum investment. It can be applied in Computer systems, High performance storage devices. This can be applied to 64 bit also in order to increase the data transfer rate. There is a chance to easy migration to ASIC technology which enables 3-5 times increase in performance rate. There is a chance to develop an engine which does not require any external components and supports operation on blocked data. Full-duplex operation enables simultaneous compression /decompression for a combined performance of high data rates, which also enables self-checking test mode using CRC (Cyclic Redundancy Check). High-performance coprocessor-style interface should be possible if synchronization is achieved. Chance to improve the Compression ratio compare with proposed work.

VIII. Acknowledgements

The authors would like to thank everyone, whoever remained a great source of help and inspirations in this humble presentation. The authors would like to thank Pragathi Engineering College Management for providing necessary facilities to carry out this work.

References

[1]. Li, K. Chakrabarty, and N. Touba, "Test data compression using dictionaries with selective entries and fixed-length indices," ACM Trans.Des. Autom. Electron. Syst., vol. 8, no. 4, pp. 470–490, 2003.

[2]. Wunderlich and G. Kiefer, "Bit-flipping BIST," in Proc. Int. Conf. Comput.Aided Des., 1996, pp. 337–343.

[3]. N. Touba and E. McCluskey, "Altering a pseudo-random bit sequence for scan based bist," in Proc. Int. Test Conf., 1996, pp. 167–175.

[4]. F. Hsu, K. Butler, and J. Patel, "A case study on the implementation of Illinois scan architecture," in Proc. Int. Test Conf., 2001, pp. 538–547.

[5]. M. Ros and P. Sutton, "A hamming distance based VLIW/EPIC code compression technique," in Proc. Compilers, Arch., Synth. Embed. Syst., 2004, pp. 132–139.

[6]. Seong and P. Mishra, "Bitmask-based code compression for embed systems," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst., vol. 27, no. 4, pp. 673–685, Apr. 2008.

[7]. M.-E. N. A. Jas, J. Ghosh-Dastidar, and N. Touba, "An efficient test vector compression scheme using selective Huffman coding," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst., vol. 22, no. 6, pp.797–806, Jun. 2003.

[8]. A. Jas and N. Touba, "Test vector decompression using cyclical scan chains and its application to testing core based design," in Proc. Int.Test Conf., 1998, pp. 458–464.

[9]. A. Chandra and K. Chakrabarty, "System on a chip test data compression and decompression architectures based on Golomb codes," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst., vol. 20, no. 3, pp.355–368, Mar. 2001.

[10]. X. Kavousianos, E. Kalligeros, and D. Nikolos, "Optimal selective Huffman coding for test-data compression," IEEE Trans. Computers, vol. 56, no. 8, pp. 1146–1152, Aug. 2007.

[11]. M. Nourani andM. Tehranipour, "RL-Huffman encoding for test compression and power reduction in scan applications," ACM Trans. Des.Autom. Electron. Syst., vol. 10, no. 1, pp. 91–115, 2005.

[12]. H. Hashempour, L. Schiano, and F. Lombardi, "Error-resilient test data compression using Tunstall codes," in Proc. IEEE Int. Symp. Defect Fault Tolerance VLSI Syst., 2004, pp. 316–323.

[13]. M. Knieser, F.Wolff, C. Papachristou, D.Weyer, and D. McIntyre, "A technique for high ratio LZWcompression," in Proc. Des., Autom., Test Eur., 2003, p. 10116.

[14]. M. Tehranipour, M. Nourani, and K. Chakrabarty, "Nine-coded compression technique for testing embedded cores in SOCs," IEEE Trans.Very Large Scale Integr. (VLSI) Syst., vol. 13, pp. 719–731, Jun. 2005.

[15]. L. Lingappan, S. Ravi, A. Raghunathan, N. K. Jha, and S. T.Chakradhar, "Test volume reduction in systems-on-a-chip using heterogeneous and multilevel compression techniques," IEEE Trans.Comput.-Aided Des. Integr. Circuits Syst., vol. 25, no. 10, pp.2193–2206, Oct. 2006.

[16]. A. Chandra and K. Chakrabarty, "Test data compression and test resource partitioning for system-on-a-chip using frequency-directed run-length (FDR) codes," IEEE Trans. Computers, vol. 52, no. 8, pp.1076–1088, Aug. 2003.

[17]. X. Kavousianos, E. Kalligeros, and D. Nikolos, "Multilevel-Huffman test-data compression for IP cores with multiple scan chains," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 16, no. 7, pp. 926–931,Jul. 2008.

[18]. X. Kavousianos, E. Kalligeros, and D. Nikolos, "Multilevel Huffman coding: An efficient test-data compression method for IP cores," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst., vol. 26, no. 6, pp.1070–1083, Jun. 2007.

[19]. X. Kavousianos, E. Kalligeros, and D. Nikolos, "Test data compression based on variable-to-variable Huffman encoding with codeword reusability," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,vol. 27, no. 7, pp. 1333–1338, Jul. 2008.

[20]. S. Reda and A. Orailoglu, "Reducing test application time through test data mutation encoding," in Proc. Des. Autom. Test Eur., 2002, pp.387–393.

[21]. E. Volkerink, A. Khoche, and S. Mitra, "Packet-based input test data compression techniques," in Proc. Int. Test Conf., 2002, pp. 154–163.

[22]. S. Reddy, K. Miyase, S. Kajihara, and I. Pomeranz, "On test data volume reduction for multiple scan chain design," in Proc. VLSI Test Symp., 2002, pp. 103–108.

[23]. F. Wolff and C. Papachristou, "Multiscan-based test compression and hardware decompression using LZ77," in Proc. Int. Test Conf., 2002,pp. 331–339.

[24]. A. Wurtenberger, C. Tautermann, and S. Hellebrand, "Data compression for multiple scan chains using dictionaries with corrections," in Proc. Int. Test Conf., 2004, pp. 926–935.

[25]. K. Basu and P.Mishra, "A novel test-data compression technique using application-aware bitmask and dictionary selection methods," in Proc.ACM Great Lakes Symp. VLSI, 2008, pp. 83–88.

[26]. T. Cormen, C. Leiserson, R. Rivest, and C. Stein, Introduction to Algorithms. boston, MA: MIT Press, 2001.

[27]. M. Garey and D. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness. New York: Freeman, 1979.

[28]. Hamzaoglu and J. Patel, "Test set compaction algorithm for combinational circuits," in Proc. Int. Conf. Comput.-Aided Des., 1998, pp.283–289.

Tree Based Parity Check Scheme for Data Hiding

[1] Sampath Kumar Dara, [2] Harshavardhan Awari
*1 Assoc. Professor, Department of Electronics and Communication Engineering, Vaageswari College of Engg, Karimnagar, A.P
*2 Assoc.Professor, Department of Computer Science Engineering, Vaageswari College of Engg, Karimnagar, A.P.

Abstract
The information hiding deals with distortion reduction using steganography and security enhancement using cryptography. Distortion reduction is done using Tree Based Parity Check which uses Majority vote strategy. The Tree Based Parity Check is very optimal for cloaking a message on image. The proposed majority vote strategy results in least distortion. The SHA-1 algorithm is implemented for security enhancement. The result obtained in proposed method works effectively even with large payload**.**

Key Words— Image coding, information security, Stenography.

I. I NTRODUCTION

Stenography studies the scheme to hide secrets into the communication between the sender and the receive r such that no other people can detect the existence of the secrets. A steganographic method consists of an embedding algorithm and an extraction algorithm. The embedding algorithm describes how t o hide a message into the
cover object and the extraction algorithm illustrates how to extract the message from the stego object. A commonly used strategy for steganography is to embed the message by slightly distorting the cove r object into the target stego object. If the distortion is sufficiently small, the stego object will be indistinguishable from the noisy cove r object. Therefore, reducing distortion is a crucial issue for steganographic methods. In this paper, we propose an efficient embedding scheme that uses the least number of changes over the tree-based parity check model.

The ideas of matrix embedding and defined the codes with the matrix as steganographic codes. For matrix embedding, finding the stego object with least distortion is difficult in general. In some special cases, there exist constructive and fast methods. LT Codes to improve the computational complexity of wet paper codes derive d a hash function to efficiently obtain the stego object. Proposed a scheme called tree-based parity check.

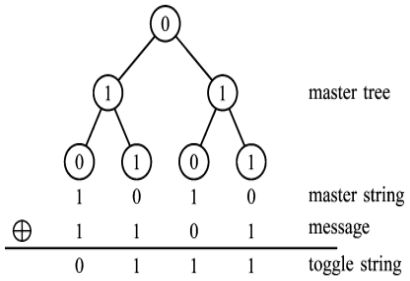


Fig. 1. Master and toggle strings of a master tree with for LSBs 0, 1, 1, 0, 1, 0, 1 o f the cove r object. (TBPC) to reduce distortion on a cover object based on a tree structure.

The embedding efficiency is defined to be the number of hidden message bits per embedding modification. Higher embedding efficiency implies better undetectability for steganographic methods. The lower embedding efficiency is defined to be the ratio of the number of hidden message bits to the maximum embedding modifications. The lower embedding efficiency is related to undetectability in the worst case. It implies steganographic security in the worst case.

II. TBPC METHOD

Before embedding and extraction, a location finding method determines a sequence of locations that point to elements in the cover object. The embedding algorithm modifies the elements in these locations to hide the message and the extraction algorithm can recover the message by inspecting the same sequence of locations. The TBPC method is a least significant bit (LSB) steganographic method. Only the LSBs of the elements pointed by the determined locations are used for embedding and

extraction. The TBPC method constructs a complete *N*-ary tree, called the master tree, to represent the LSBs of the cover object. Then it fills the nodes of the master tree with the LSBs of the cover object level by level, from top to bottom and left to right. Every node of the tree corresponds to an LSB in the cover object. Denote the number of leaves of the master tree by *L*. The TBPC embedding algorithm derives an *L*-bit binary string, called the *master string*, by performing parity check on the master tree from the root to the leaves (e.g., see Fig. 1.). The embedding algorithm hides the message by modifying the bit values of some nodes in the master tree. Assume that the length of the message is also L. Performing the bitwise exclusive-or (XOR) operation between the message and the master string, we obtain a toggle string (e.g., see Fig. 1). Then, the embedding algorithm constructs a new complete N-ary tree, called the toggle tree in the bottom-up order and fills the leaves with the bit values of the toggle string and the other nodes with 0. Then, level by level, from the bottom to the root, each nonleaf node together with its child nodes are flipped if all its child nodes have bits 1 (e.g., see Fig. 2). The embedding algorithm obtains the stego tree by performing XOR between the master tree and the toggle tree (e.g., see Fig. 3). The TBPC extraction algorithm is simple. We can extract the message by performing parity check on each root-leaf path of the stego tree from left to right.

III. Majority Vote Strategy

Two critical issues for a steganographic method are:
1) Reducing distortion on cover objects

2) Better efficiency for embedding and extraction. We give a majority vote strategy on building the toggle tree. It uses the least number of 1's under the TBPC model. Since the number of 1's in the toggle tree is the number of modifications on the master tree (i.e., the cover object), the majority vote strategy can produce a stego tree with least distortion on the master tree. The proposed method uses a set of standard measures, to capture image properties before or after the embedding process, which effect the performance of steganalysis techniques. These measures are divided into two categories. First cover image properties, and second cover-stego based distortion measures.

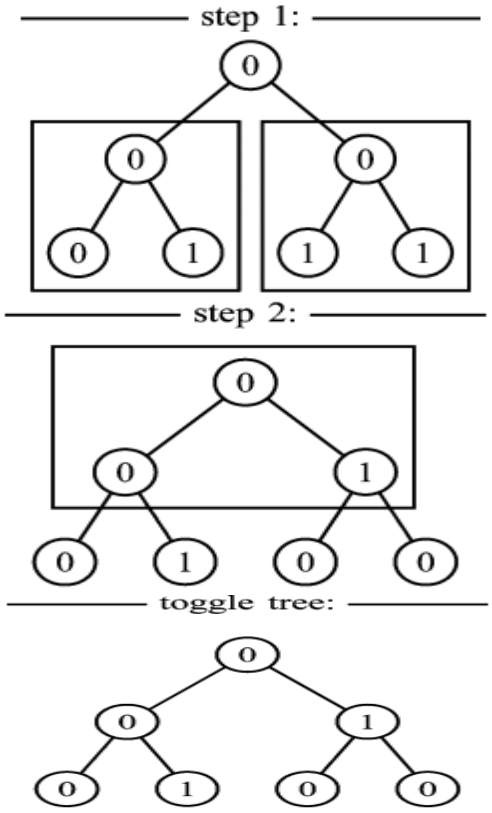


Fig. 2. Construction of a toggle tree with for toggle string 0, 1, 1, 1.

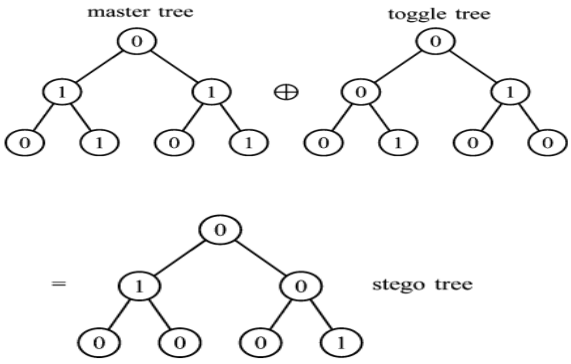


Fig. 3. Modify the master tree into the stego tree by the toggle tree constructed from the toggle string 0, 1, 1, 1.

First, index all nodes of a complete -ary tree with leave s from top to bottom and left to right. Set the bit toggle string bit by bit into the leave s from left to right and the other nodes 0. Assume that the level of the tree is. Traverse all nonleaf nodes from level 1 to. A nonleaf node and its child nodes form a simple complete subtree. For each simple complete subtree, if the majority of the child nodes hold 1, then flip the bit values of all nodes in this

subtree. Since the construction is bottom-up, the bit values of the child nodes in every simple complete subtree are set after step 3. Note that marking a node at step 4 applies only for being even. When is even, after step 3, there may exist a two level simple complete subtree with 1's in the child nodes and 1 in its root. In this case, flipping the bit values in this simple complete subtree results in one fewer node holding 1 and keeps the result of related root-leaf path parity check unchanged. Step 4 takes care of this when the condition applies, and it is done level b y level from top to bottom. Also note that for the root of the whole toggle tree, the bit value is always 0 when half of its child nodes hold 1. Thus, after step 4, the bit values of the child nodes in each simple complete subtree are determined.

IV. Experimental Results

To make it clear, we define the percentage of reduced modifications as follows:

p Reduce = Rt/Dt

Where Rt is the reduced number of 1's in the toggle tree and Dt is the number of 1's in the toggle string. The p Reduce values of both methods are shown in Fig. 4.

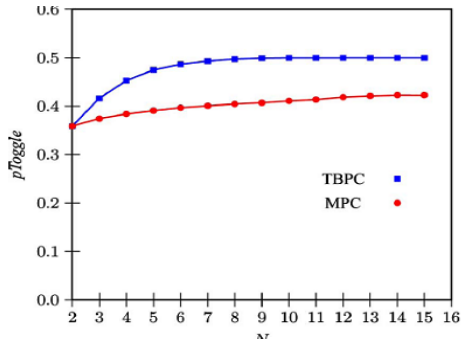


Fig.4. p Toggle comparison of MPC and TBPC.

The results show that the MPC method significantly improves previous TBPC results.

V. Conclusion

By introducing the majority vote strategy, the stego object is constructed with least distortion under the tree structure model. We also show that our method yields a binary linear stego-code and preserves the secrecy of the hidden data. In comparison with the TBPC method, the proposed MPC method significantly reduces the number of modifications on average.

References

[1] S. Chang Chun Lu, Shi-Chun Tsai, and Wen-Guey Tzeng, An Optimal Data Hiding Scheme With Tree-Based Parity Check in IEEE Trans. on Image Processing, vol. 20, no. 3, Mar 2011.

[2] J. Fridrich, M. Goljan, and D. Soukal, —Efficient wet paper codes, in Proc. 7th Int. Workshop Information Hiding (IHW 05), vol. 3727, pp. 204–218, Apr 2005.

[3] J. Fridrich and D. Soukal, —Matrix embedding for large payloads, in IEEE Trans. on Information Forensics and Security, vol. 1, no. 3, pp. 390–395, Sep. 2006.

[4] M. Khatirinejad and P. Lisonek, —Linear codes for high payload steganography, in IEEE Trans. on Discrete Applied Mathematics, vol. 157, no. 5, pp. 971–981, 2009.

[5] W. Zhang and X. Wang, Generalization of the ZZW embedding construction for Steganography, in IEEE Trans. on Information Forensics and Security, pp.564-569, Jan 2009.

[6] J. Fridrich, M. Goljan, P. Lisonek, and D. Soukal, —Writing on wet paper, in IEEE Trans. on Signal Processing., vol. 53, no. 10, pp. 3923–3935, Oct. 2005.

Speed Control of Induction Motor using Fuzzy PI Controller Based on Space Vector Pulse Width Modulation

[1]Yakala Satyanarayana, Dr.A.Srujana

Abstract

The aim of this paper is that it shows the dynamics response of speed with design the fuzzy logic controller to control a speed of motor for keeping the motor speed to be constant when the load varies. In recent years, the field oriented control of induction motor drive is widely used in high performance drive system. It is due to its unique characteristics like high efficiency, good power factor and extremely rugged. This paper presents design and implements a voltage source inverter type space vector pulse width modulation (SVPWM) for control a speed of induction motor. This paper also introduces a fuzzy logic controller to the SVPWM in order to keep the speed of the motor to be constant when the load varies. FLC is used to control the pulse width of the PWM converter used to control the speed of the motor.

Index terms — Fuzzy logic control (FLC), Fuzzy PI controller, Induction motor, Membership Function, Space Vector Pulse Width Modulation(SVPWM)

I. Introduction

In recent years, the field oriented control of induction motor drive is widely used in high performance drive system. It is due to its unique characteristics like high efficiency, good power factor and extremely rugged. Induction motor are used in many applications such as HVAC (heating, ventilation and air-conditioning), Industrial drives (motion control, robotics), Automotive control (electric vehicles), etc.. In recent years there has been a great demand in industry for adjustable speed drives.

The Space Vector Pulse Width Modulation (SVPWM) method is an advanced, computation-intensive PWM method and possibly the best among all the PWM techniques for variable frequency drive application. Because of its Superior performance characteristics, it has been finding widespread application in recent years .The PWM methods discussed so far have only considered Implementation on half bridges operated independently, giving satisfactory PWM performance. With a machine Load, the load neutral is normally isolated, which causes interaction among the phases.

Recently, Fuzzy logic control has found many applications in the past decade. Fuzzy Logic control (FLC) Has proven effective for complex, non-linear and imprecisely defined processes for which standard model based control techniques are impractical or impossible. Fuzzy Logic, deals with problems that have vagueness, uncertainty and use membership functions with values varying between 0 and1. This means that if the a reliable expert knowledge is not available or if the controlled system is too complex to derive the required decision rules, development of a fuzzy logic controller become time consuming and tedious or sometimes impossible. In the case that the expert knowledge is available, fine-tuning of the controller might be time consuming as well. Furthermore, an optimal fuzzy logic controller cannot be achieved by trial-and-error. These drawbacks have limited the application of fuzzy logic control. Some efforts have been made to solve these problems and simplify the task of tuning parameters and developing rules for the controller.

These approaches mainly use adaptation or learning techniques drawn from artificial intelligence or neural network theories. Application of fuzzy logic control for the control a speed induction motor using space vector pulse width modulation is quite new.

Uncertain systems even in the case where no mathematical model is available for the controlled system. However, there is no systematic method for designing and tuning the fuzzy logic controller.

The aim of this project is that it shows the dynamics response of speed with design the fuzzy logic controller to control a speed of motor for keeping the motor speed to be constant when the load varies. This project presents design and implements a voltage source inverter type space vector pulse width modulation (SVPWM) for control a speed of induction motor. This project also introduces a fuzzy logic controller to the SVPWM in order to keep the speed of the motor to be constant when the load varies.

II. Inverter for Ac Drives

A. *Space Vector Pulse Width Modulation*

For A.C. drive application sinusoidal voltage source are not used. They are replaced by six power IGBT's that act as on/off switches to the rectified D.C. bus voltage. Owing to the inductive nature of the phases, a pseudo-sinusoidal current is created by modulating the duty-cycle of the power switches. Fig.1. shows a three phase bridge inverter induction motor drive.

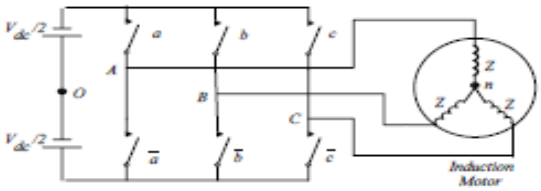


Fig. 1 .Three phase voltage source inverter

$$V_{An} = V_m \cos \omega t$$

$$V_{Bn} = V_m \cos\left(\omega t - \frac{2\pi}{3}\right)$$

$$V_{Cn} = V_m \cos\left(\omega t + \frac{2\pi}{3}\right)$$

$$\bar{V} = \frac{2}{3}\left[V_{An} + aV_{Bn} + a^2 V_{Cn}\right]$$

Van, VBn, VCn are applied to the three phase induction motor, using Equation $\bar{V}$. A three phase bridge inverter, From Figure.1, has 8 permissible switching states. Table I gives summary of the switching states and the corresponding phase-to-neutral voltage of isolated neutral machine.

For the three phase two level PWM inverter the switch function is defined as

SWi = 1, the upper switch is on and bottom switch is off.

SWi = 0, the upper switch is off and bottom switch is on.

where *i= A,B,C.*

"1" denotes V_{dc} /2 at the inverter output, "0" denotes -V_{dc} /2 at inverter output with respect to neutral point of the d.c. bus. The eight switch states S_i = (SW_A ,SW_B ,SW_C) where i=0,1,…..7 are shown in Fig. 2. There are eight voltage vectors V_0 - - - - -V_7 corresponding to the switch states S_0 - - - - - S_7 respectively. The lengths of vectors V_1 - - - - -V_6 are unity and the length of V_0 and V_7 are zero. These eight vectors form the voltage vector space as depicted in Fig. 3.

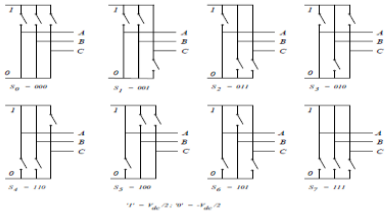


Fig. 2. Eight switching states of VSI.

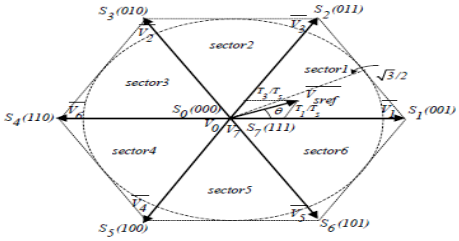


Fig. 3. Voltage space vectors.

The six non-zero voltage space vectors form a hexagonal locus. The voltage vector space is divided into six sectors. It can be seen that when the space vector moves from one corner of the hexagon to another corner, then only the state of one inverter leg has to be changed. The zero space vectors are located at the origin of the reference frame. The reference value of the stator voltage space vector V_{sref} can be located in any of the six sectors. Any desired stator voltage space vector inside the hexagon can be obtained from the weighted combination of the eight switching vectors. The goal of the space vector modulation technique is to reproduce the reference stator voltage space vector (V_{sref}) by using the appropriate switching vectors with minimum harmonic current distortion and the shortest possible cycle time. The eight permissible states are asummarized in Table I.

TABLE I :
SUMMARY OF INVERTER SWITCHING STATES

Voltage vector	SW_A	SW_B	SW_C	V_{An}	V_{Bn}	V_{Cn}
$\bar{V}_0$	0	0	0	0	0	0
$\bar{V}_1$	0	0	1	$-V_{dc}/3$	$-V_{dc}/3$	$2V_{dc}/3$
$\bar{V}_2$	0	1	0	$-V_{dc}/3$	$2V_{dc}/3$	$-V_{dc}/3$
$\bar{V}_3$	0	1	1	$-2V_{dc}/3$	$V_{dc}/3$	$V_{dc}/3$
$\bar{V}_4$	1	0	0	$2V_{dc}/3$	$-V_{dc}/3$	$-V_{dc}/3$
$\bar{V}_5$	1	0	1	$V_{dc}/3$	$-2V_{dc}/3$	$V_{dc}/3$
$\bar{V}_6$	1	1	0	$V_{dc}/3$	$V_{dc}/3$	$-2V_{dc}/3$
$\bar{V}_7$	1	1	1	0	0	0

$$\overline{V_{sref}} = \frac{T_0}{T_s}\overline{V_0} + \frac{T_1}{T_s}\overline{V_1} + \ldots\ldots + \frac{T_7}{T_s}\overline{V_7} \quad (1)$$

Where T_0 ,T_1 ,.....T_7 are the turn on time of the vectors$\overline{V_0}$,$\overline{V_1}$,.....$\overline{V_7}$ respectively and T_0 ,T_1 ,.....$T_7 \geq 0$, $\sum_{i=0}^{7} T_i = T_s$ where *Ts* is the sampling time.

In order to reduce the number of switching actions and to make full use of active turn on time for space vectors, the vector V_{sref} is split into the two nearest adjacent voltage vectors and zero vectors V_0 and V_7 in an arbitrary sector. For Sector 1 in one sampling interval, vector V_{sref} can be given as

$$\overline{V_{sref}} = \frac{T_1}{T_s}\overline{V_1} + \frac{T_3}{T_s}\overline{V_3} + \frac{T_7}{T_s}\overline{V_7} + \frac{T_0}{T_s}\overline{V_0} \quad (2)$$

where *Ts* - T_1 - T_3 = T_0 +$T_7 \geq 0$, $T_0 \geq 0$ and $T_7 \geq 0$
The length and angle of *Vsref* are determined by vectors $\overline{V_1}$,$\overline{V_2}$,.....$\overline{V_6}$ that are called active vectors and $\overline{V_0}$, $\overline{V_7}$ are called zero vectors. In general

$$\overline{V_{sref}}T_s = \overline{V_i}T_i + \overline{V_{i+1}}T_{i+1} + \overline{V_7}T_7 + \overline{V_0}T_0 \quad (3)$$

Where T_i ,T_i+*1* ,T_7 ,T_0 are respective on duration of the adjacent switching state vectors ($\overline{V_i}$,V_i+*1* ,V_7 *and*V_0) . The on durations are defined

$$T_i = mT_s Sin(60 - \theta) \quad (4)$$

$$T_{i+1} = mT_s Sin(\theta) \quad (5)$$

$$T_7 + T_0 = T_s - T_i - T_{i+1} \quad (6)$$

Asfollows:
Where *m* is modulation index defined as:

$$m = \frac{2}{\sqrt{3}}\frac{|V_{sref}|}{V_{dc}} \quad (7)$$

V_{dc} is d.c. bus voltage and θ is angle between the reference vector *Vsref* and the closest clockwise state vector as depicted in Fig. 3.

In the six step mode, the switching sequence is *S1* - S_2 - S_3 - S_4 - S_5 - S_6 - S_1. Furthermore it should be pointed out that the trajectory of voltage vector *Vsref* should be circular while maintaining sinusoidal output line to line voltage. In the linear modulation range, $\overline{V_{sref}} = \sqrt{3}/2\, V_{dc}$, the trajectory of V_{sref} becomes the inscribed circle of the hexagon as shown in the Fig. 3.

In conventional schemes, the magnitude and the phase angle of the reference voltage vector (i.e. V_{sref} and θ) are calculated at each sampling time and then substituted into (7) and (4), (5) to get the value of on duration. Due to *Sine* Function in (4) and (5) it produces a larger computing delay. Although the use of a lookup table and linear interpolation are used but it increase computation time and interpolation of non-linear function may lead to reduced accuracy and therefore contribute to the deterioration of PWM waveforms.

B. Simulink Implementation

To implement the algorithm in Simulink, we shall first assume that the three-phase voltages at the stator terminals must have the following from Equation. V_{an},V_{Bn},V_{Cn} , the frequency f and the amplitude V are variables. However, the v/f control algorithm implies that there is a relationship between the amplitude of the voltage and the frequency, i.e. the ratio between the two quantities is constant.

K=V/f [8]

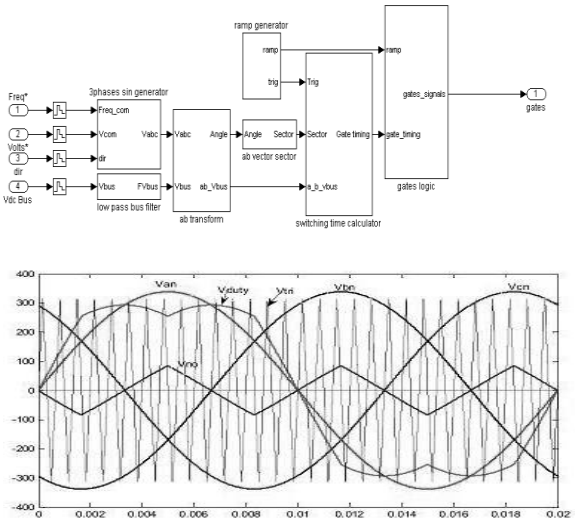


Fig. 4 (a) Simulink implementation of SVPWM, (b) Space Vector Pulse Width Modulation of v/f

III. Fuzzy Logic Controller

Fuzzy Logic control (FLC) has proven effective for complex, non-linear and imprecisely defined processes for which standard model based control techniques are impractical or impossible. The complete block diagram of the fuzzy logic controller is shown and The function of each block and its realization is explained below.

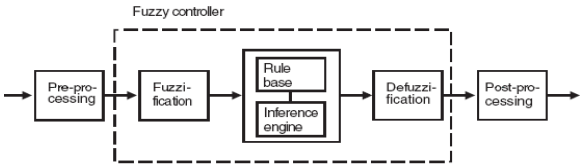


Fig.5.General Fuzzy Block Diagram

A) CONFIGURATION OF FLC:
It comprises of four principal components:
a) A fuzzification interface
b) A knowledge base
c) A decision-making logic and
d) A defuzzification interface.

a) Fuzzification
Fuzzification interface involves the following functions.
(1) Measures the values of input variable.
(2) Performs the function of fuzzification that converts input data into suitable linguistic values © 2009

b) Knowledge base
Knowledge base consist data base and a linguistic control rule base.
(1) The database provides necessary definitions, which are used to define linguistic control rules.
(2) The rule base characterized the control goals and control policy of the domain experts by means of a set of linguistic control rules.

b)Decision making\
The decision-making logic is the kernel of an FLC. It has the capability of simulating human decision making based on fuzzy concepts and of inferring fuzzy control actions employing fuzzy implication and the rules of inference in fuzzy logic.

d) Defuzzication
Defuzzification interface performs the following functions.
(1) A scale mapping, which converts the range of values of output variables into corresponding universe of discourse.
(2) Defuzzification, which yields a non-fuzzy control action from an inferred fuzzy control action.

B) Rules Creation And Inference:

In general, fuzzy systems map input fuzzy sets to output sets. Fuzzy rules are relations between input/output fuzzy sets. The modes of deriving fuzzy rules are based either of the following.

- Expert experience and control engineering knowledge.
- Operator's control actions.
- Learning from the training examples.

In this thesis the fuzzy rules are derived by learning from the training examples. The general form of the fuzzy control rules in this case is

IF x is A_i AND y is B_i THEN z =f_i (x, y)

Where x, y and z are linguistic variables representing the process state variables and the control variable respectively. A_i, B_i are the linguistic values of the linguistic variables, f_i (x, y) is a function of the process state variables x, y and the resulting fuzzy inference system (FIS) is called a first order sugeno fuzzy model.

C. Fuzzy inference engine

The function of the inference engine is to calculate the overall value of the control output variable based on the individual contributions of each rule in the rule base. (i.e.) the defuzzification process. There is no systematic procedure for choosing defuzzification. In first-order sugeno fuzzy model each rule has a crisp output and overall output is obtained as weighted average thus avoiding the time consuming process of defuzzification required in a conventional FLC.

IV. Design of Fuzzy Pi Controllr:

The basic block diagram of a PI type FLC for Induction motor speed control is shown . It is known that a FLC consists of the fuzzification process, the knowledge base and the defuzzification process.

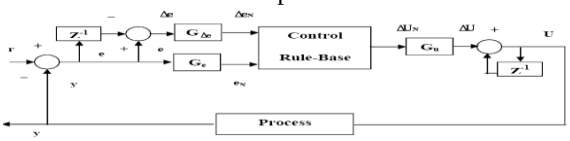


Fig.6 Block diagram of Fuzzy PI Controller

The FLC has two inputs, the error e(k) and change of error Δe(k), which are defined by e(k) = r(k)−y(k), Δe(k) =e(k) − e(k − 1), where r and y denote the applied set point input and plant output, respectively. Indices k and k−1 indicate the present state and the previous state of the system, respectively. The output of the FLC is the incremental change in the control signal Δu(k).The controller has two input variables and one output variable.

The input and output variables of fuzzy PI controller can be defined as:
E(k) = e(k).Ge ….(9)
CE(k) = ce(k).Gce ….(10)
Δi(k) =ΔI(k).*G*Δ*i* ….(11)
where e(k) is the error between reference speed and rotor speed,
ce(k) is the change of error in speed,
I(k) is the output of the fuzzy logic controller,
and Ge, Gce and Gδi are scaling factors.
If *e* is *E* and Δ*e* is Δ*E*, then Δ*u* is Δ

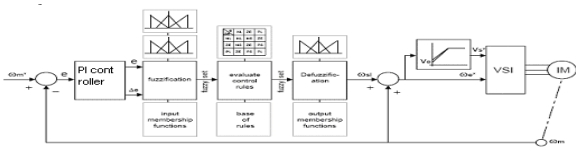


Fig.7 Speed Control of Induction Motor using Fuzzy PI

A fuzzy logic controller is proposed to control the speed of the motor to be constant when the load varies. The speed error e and the change of speed error are processed through the fuzzy logic controller whose output is the voltage command. Current error is usually processed by current regulator to produce a control frequency.

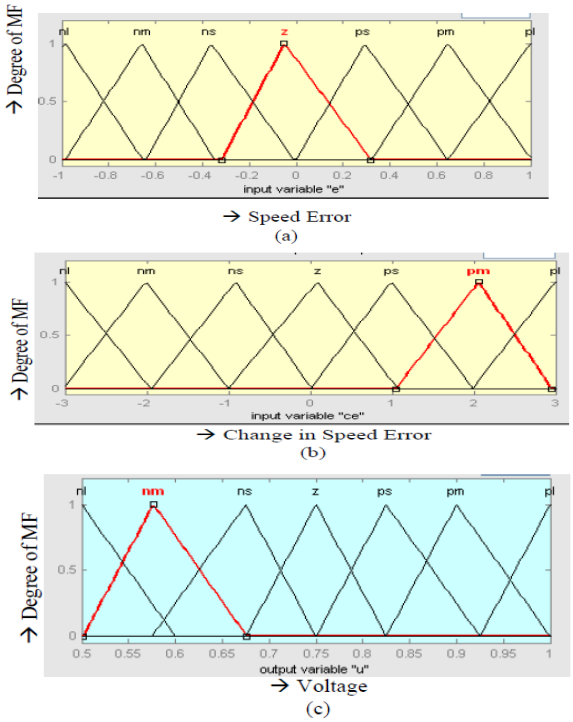


Fig.8 Membership functions

(a) MF for speed error (b) MF for change in speed error (c) MF for voltage

TABLE II

Rule Base of Fuzzy Speed and Current Control

ce/e	NB	NM	NS	Z	PS	PM	PB
NB	NB	NB	NB	NB	NM	NS	Z
NM	NB	NB	NB	NM	NS	Z	PS
NS	NB	NB	NM	NS	Z	PS	PM
Z	NB	NM	NS	Z	PS	PM	PB
PS	NM	NS	Z	PS	PM	PB	PB
PM	NS	Z	PS	PM	PB	PB	PB
PB	Z	PS	PM	PB	PB	PB	PB

V. Results And Discussions

To evaluate the performance of the system, a series of measurements has been accomplished. Fig . 9 as shown performance of the fuzzy logic controller with a fuzzy tuning rule based on Reference speed of 800 rpm with no load torque. Fig . 10 as shown performance of the fuzzy logic controller with a fuzzy tuning rule based on Reference speed of 800rpm with load torque. Fig . 11 as shown performance of the fuzzy logic controller with a fuzzy tuning rule based on Reference speed of 1200rpm with no load torque. Fig . 12 as shown performance of the fuzzy logic controller with a fuzzy tuning rule based on Reference speed of 1200rpm with load torque.

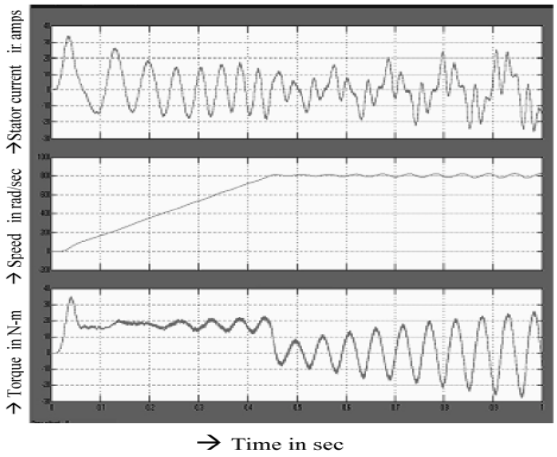


Fig. 9 Reference speed of 800 rpm with no load

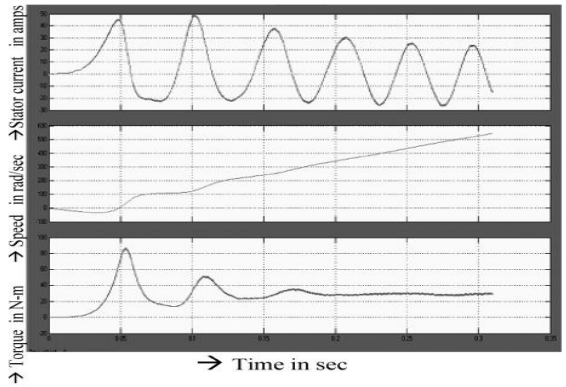


Fig. 10 Reference speed of 800 rpm with load

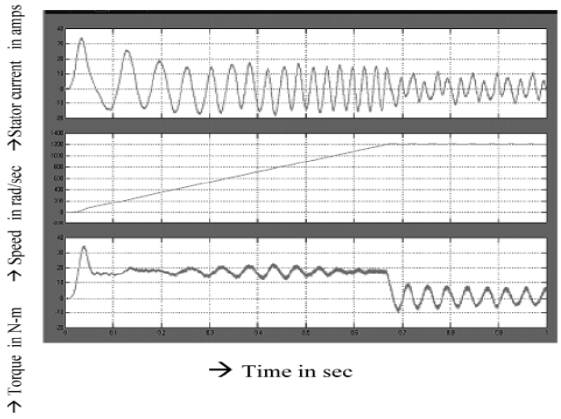


Fig. 11 Reference speed of 1200 rpm with no load

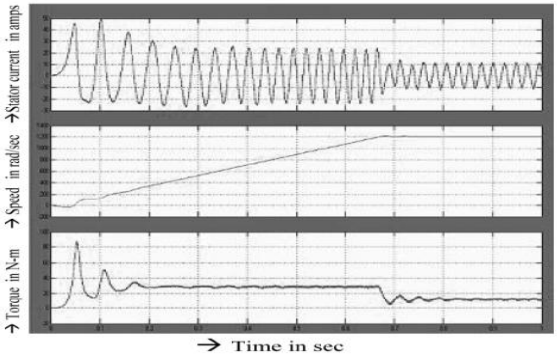


Fig. 12 Reference speed of 1200 rpm with load

From the results tested the performance of controller by a step change of the speed reference at constant load torque as shown in Figure. 11, it's found that the Rise time tr = 0.6s, Settling time t_s = 1 sec.

Vi. Conclusion

This paper presents simulation results of fuzzy logic control for speed control of induction motor. In fuzzy control it is not necessary to change the control parameters as the reference speed changes, however with the classical PI controller this does not happens. With results obtained from simulation, it is clear that for the same operation condition the induction motor speed control using fuzzy PI controller technique had better performance than the PI controller, mainly when the motor was working at lower speeds. In addition, the motor speed to be constant when the load varies.

REFERENCES

[1] W. Leonhard, Control of Electrical Drives, Springer-Verlag Berlin Heidelberg, New York, Tokyo, 1985.

[2] Bimal K.Bose, "Modern Power Electronics and AC Drives", Pearson education .

[3] R.Krishnan, " Electric motor drives", Prentice – Hall India, New Delhi 2005.

[4] A.Iqbal, "Analysis of space vector pulse width modulation for a five phase voltage source inverter"IE (I)journal-EL, Volume 89 Issue 3.September 2008 Pages 8-15.

[5] Mondal, S.K.; Bose, B.K.; Oleschuk, V.; Pinto, J.O.P, "Space vector pulse width modulation of three-level inverter extending operation into overmodulation region", IEEE Transactions on Power Electronics Volume 18, Issue 2, March 2003 Page(s):604 – 611.

[6] Peter Vas, " Artificial –Intelligence based electrical machines and drives", Oxford university 1999.

[7] Andrzej M. Trzynadlowski, "Introduction to Modern Power Electronics", copyright© 1998 by John Wiley & Sons, Inc. All rights reserved.

[8] Chuen Chien Lee, "Fuzzy Logic in Control Systems:Fuzzy Logic controller–Part 1" 1990 IEEE.

[9] Chuen Chien Lee, "Fuzzy Logic in Control Systems : Fuzzy Logic controller –Part 2" 1990 IEEE .

[10] Zdenko Kovaccic and Stjepan Bogdan, "Fuzzy Controller design Theory and Applications", © 2006 by Taylor & Francis Group. International, 2002.

[11] Hassan Baghgar Bostan Abad, Ali Yazdian Varjani, Taheri Asghar "Using Fuzzy Controller in Induction Motor Speed Control with Constant Flux "proceedings of world academy of science, engineering and technology Volume 5 April 2005 ISSN 1307-6884.

[12] Mir.S.A and Malik. E. Elbuluk, (1994), "Fuzzy controller for Inverter fed Induction Machines", IEEE Transactions on Industry Applications, Vol.30. PP. 78-84.

[13] Peter Vas, " Sensorless Vector and Direct Torque control", Oxford university press 1998.

BIOGRAPHY

Yakala Satyanarayana, II-M.Tech P.E,Department of EEE, SV Engineering College, Suryapet.

Dr.A.Srujana has obtained her B. Tech degree from KITS, Warangal, in 1998 and M. Tech degree From J.N.T.U ,Hyderabad, India, in 2002.She Has obtained her Ph.D from J.N.T.U ,Hyderabad in 2011. She has 14 years of teaching experience. Presently he is working as Professor & H.O.D E.E.E at Sri venkateswara engineering college, Suryapet.

Animal Sign language recognition using MEMS

R.BHARATH[1], DR S.A.K JILANI[2]
[1], [2](ECE, MITS, JNTUA, INDIA)

Abstract :
Mple dog's express their feelings by oscillating his tail. Similarly cows express their feelings with his feet, neck and tail. Every feeling has its distinct body movements. So by placing this model on their communicating body part we transform the animal feelings into words and phrases. This model also helps the deaf and dumb community to communicate with others in oral langage. Prototype of sign language recognition consists of ADXL335 accelerometer interfaced with PIC micro controller 16F873A. The interfacing program is written in embedded 'C' language and it is compiled with Hi-tech compiler. The accelerometer data is processed in PC using neural network pattern recognition tool available in MATLAB.

In this model we transformed six American Sign Language postures into words. We proposed two algorithms based on Euclidean distance metric and neural network pattern recognition tool with spline interpolation technique achieving an overall l efficiency of 80% and 83.3% respectively. Former algorithm is preferred because here we achieve an efficiency varying from 70% to 90% where as in Euclidean distance algorithm efficiency varies from 0% to 100%

Keywords-ASL, Euclidean distance, nprtool, MEMS accelerometer, animal sign language, spline interpolation, interpolation technique

I. Introduction

Animals cannot communicate orally with human beings. Even they make different sounds which everyone is not understandable. Every animal express its feelings in their mere way. By seeing the body language we are able to interpret the feelings of that corresponding animal. Every animal has its own distinct movements. So we have to reprogram for each animal. However the model remains same for every animal. Here we are fixing this setup to the dominant body part where animals express their feelings for example say tail, neck, head etc. by tracing this movements we are interpreting their emotions. By doing this we can build the gap between humans and animals and animals can communicate with oral language. We in this paper intended to develop a sign language interpreter which is ideally designed for animals.

The legal recognition of sign languages is one of the major concerns of the international Deaf community. There is no standard way in which such recognition can be formally or legally extended; every country has its own interpretation. In some countries, the national sign language is an official state language, whereas in others it has a protected status in certain areas such as education. However, symbolic recognition is no guarantee for an effective improvement of the life of sign language users. This model is also useful for deaf and dumb community, with this model they can communicate with outside orally.

"Figure1: Sign language interpreter used in CONGO film"

1.1 Source

The intention of this paper is to duplicate a system what is used in the CONGO film to convert 'AMMY' (a talking gorilla) hand movements into oral language as shown in the figure1. It's a science fiction movie. They made a set up and tied to arm of gorilla which is trained with sign language. If gorilla wants to communicate with his trainer it will make the signs with his hand, this setup will convert the signs into words and phrases. By duplicating this model we can use the same concept for other animals.

1.2Posters of some of the words used in this paper

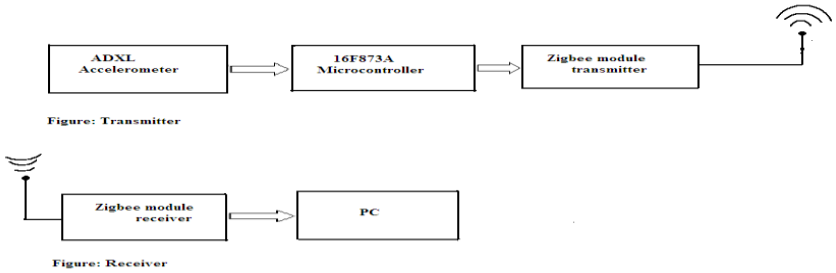
Good **Hello** **Goodbye** **Please** **Thanqu** **Hello**

"Figure2: posters of words corresponding to American sign language(ASL)"

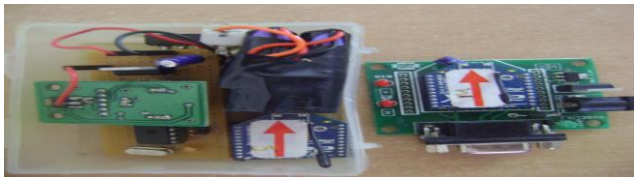
These are the posters of some of the words which we transformed into voice. These posters are belonged to ASL.

"II. Implementation"

Implementation of this model consists of two stages as shown in figure2. First stage consists of transmitting stage where accelerometer sensor ADXL335 is interfaced with PIC 16f873a and the output of microcontroller send to PC via zigbee module



"Figure3: Block Diagram implementation of model"



"Figure 4: Experimental Setup"

Adxl335, PIC, zigbee along with battery is assembled in a box as shown in figure4 and the box is tied to the wrist. This set up will trace the motion of the hand in 3-dimensional i.e., in X,Y and Z-axis. Each sample from the sensor consists of X,Yand Z values.

"III. Design Issues"

To achieve the results to maximum efficiency we have to concentrate on parameters which affect the performance of the model. All the design issues and remedies are discussed below.

3.1Hand movements

1. We should practice every single moment with precise angular velocity.

2. Palm position and orientation

These two parameters are key in practicing hand movements. We have to practice the movements to the movements which is programmed in the system. Any change in above parameters will affect the recognizing rate of the model.

3.2Number of samples

Number of samples considered for computation is key issue because we are translating the signs which have large and small hand displacements, we have to consider more number of samples for large hand displacements and we have to consider less number of samples for small hand displacements for efficient mapping. In this model we considered five samples for computation.

"IV. Results"

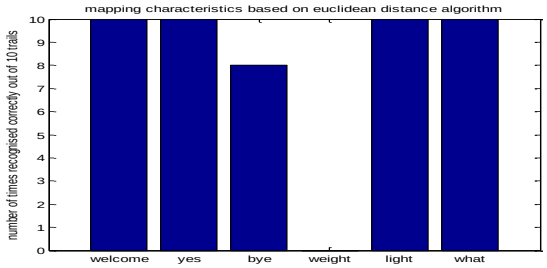
We proposed and worked three algorithms for the model and achieved different recognizing rates for each model. In this model we transformed signs of six words corresponding to American Sign Language into words. The words used in this model are YES, BYE, WELCOME, WEIGHT, LIGHT and WHAT. In each algorithm by practicing sixty trails i.e., ten trails for each word we took the statistics of how many times it recognized correctly and same is plotted in bar graphs taking words on X-axis and recognizing rate on Y-axis, recognizing rates for each word is discussed individually.

4.1 Efficiency of Euclidean distance algorithm

Table 1 reveals the statistics of the word recognition using Euclidian distance algorithm. The same is plotted as bar graph as shown in figure5, Marking words on X-axis and recognizing rate on Y-axis. In this algorithm we achieved an overall efficiency of eighty percent..

"Table1: mapping table"

	Yes	bye	welcome	weight	Light	what
Yes	**10**	**0**	**0**	**0**	**0**	**0**
Bye	**2**	**8**	**0**	**0**	**0**	**0**
Welcome	**0**	**0**	**10**	**0**	**0**	**0**
Weight	**6**	**0**	**0**	**0**	**4**	**0**
Light	**0**	**0**	**0**	**0**	**10**	**0**
What	**0**	**0**	**0**	**0**	**0**	**10**



"Table1: mapping table"

"figure5: recognition rate of each word"

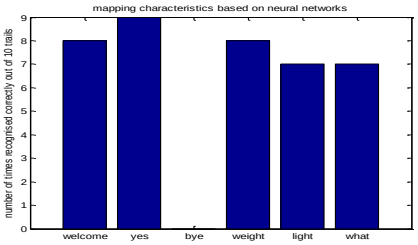
From figure5 and table 1, we reveal that four words i.e., WELCOME, YES, LIGHT, WHAT achieved 100% recognition. The word BYE achieved 80% recognition. The word WEIGHT is not at all recognized achieving 0% efficiency. Out of sixty trails forty eight times the words are recognized correctly yielding overall efficiency 80%.To improvise the recognizing rates we proposed another algorithm which is based on neural networks

4.2 Efficiency of algorithm based on neural network pattern recognition

Table 2 reveals the statistics of the word recognition using npr tool. The same is plotted as bar graph as shown in figure6, Marking words on X-axis and recognizing rate on Y-axis. In this algorithm we achieved an overall efficiency of sixty five percent. In this algorithm we considered five samples from the sensor for computation and filed the recognizing rate of each word yielding least recognizing rate. we tried for other signs corresponding to American sign language resulted almost same recognizing rate and one word is not completely recognizable.

	Welcome	Yes	Bye	weight	light	what
Welcome	**8**	0	0	0	1	1
Yes	0	**9**	0	0	0	1
Bye	1	2	**0**	5	0	1
Weight	0	1	0	**8**	0	1
Light	3	0	0	0	**7**	0
What	1	0	0	2	1	**7**

"Table2: mapping table"



"figure6: recognition rate of each word"

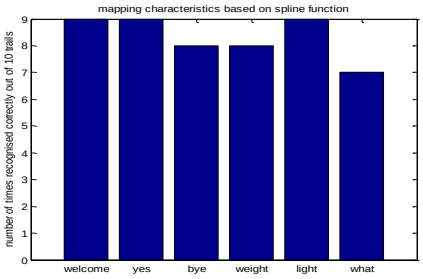
From table 2 and from figure6 we reveal that the word YES has maximally recognized nine times and the words WELCOME, WEIGHT and LIGHT, WHAT have recognized eight and seven times out of ten trails respectively. The word BYE has least recognizing rate of zero percent. In this algorithm thirty nine times out of sixty trails recognized correctly yielding a overall efficiency of sixty five percent.

To improve efficiency further SPLINE interpolation technique is employed to neural network algorithm and better recognizing rates are achieved.

4.3 Algorithm based on npr using SPLINE interpolation technique

	Yes	bye	welcome	weight	Light	what
Yes	**9**	0	0	1	0	0
Bye	1	**8**	1	0	0	0
Welcome	0	0	**9**	0	1	0
Weight	0	0	2	**8**	0	0
Light	0	0	0	0	**9**	1
What	0	3	0	0	0	**7**

"Table3: mapping table"

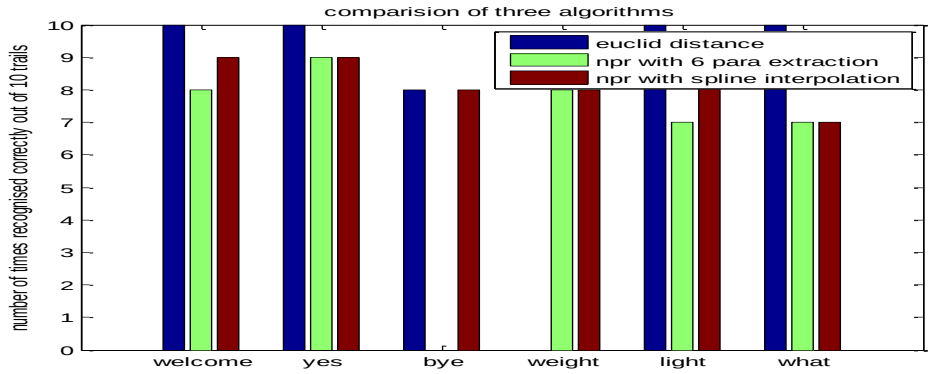


"figure7: recognition rate of each word"

From the table 3 and figure7 we infer that the least recognizing word is WHAT with seventy percent which is significant increase from zero to seventy percent when compared with previous two algorithms. The words YES, WELCOME, LIGHT has maximal recognizing rate of ninety percent and the word BYE, WEIGHT has the recognizing rate of eighty percent. These algorithm yields the overall recognizing rate of 83.3 percent. In this algorithm we achieved an increase of 3.3 percent with previous best algorithm of 80%.

4.4 Comparison of three algorithms

Comparing all the algorithms, the former algorithm is best as the efficiency range is low i.e., 70 to 90% and having overall efficiency of 83.3%, which is depicted the same in the figure8



"Figure8: comparison of three algorithms"

In first two algorithms one word is not completely recognized, which is removed with spline interpolation technique in third algorithm and we achieve an improvement of 3.3% over previous best algorithm which is based on Euclidean distance.

"V. Conclusion"

In this model, we made use of single accelerometer as we intended to transform only dominant hand movements. We proposed few algorithms which are based on Euclidian distance metric and another one based on neural network pattern recognition. We transformed six signs corresponding to ASL achieving an efficiency of 83.3%. This model is ideally suitable for animals and also for deaf and dumb community and it works in real time.

References

Books

[1] M. Birkholz, K.-E. Ehwald, P. Kulse, J. Drews, M. Fröhlich, U. Haak, M. Kaynak, E. Matthus, K. Schulz, D. Wolansky (2011). "*Ultrathin TiN Membranes as a Technology Platform for CMOS-Integrated MEMS and BioMEMS Devices*". Adv. Func. Mat. 21: 1652–

[2] M. Birkholz, K.-E. Ehwald, P. Kulse, J. Drews, M. Fröhlich, U. Haak, M. Kaynak, E. Matthus, K. Schulz, D. Wolansky (2011). "*Ultrathin TiN Membranes as a Technology Platform for CMOS-Integrated MEMS and BioMEMS Devices*".

Websites

[3] www.analog.com/static/imported-files/data_sheets/ADXL335.pdf
[4] www.microchip.com/pic16F873A
[5] category.alldatasheet.com/index.jsp?semiconductor=ZigBee/
[6] en.wikipedia.org/wiki/Euclidean distance
[7] www.mathworks.com/help/toolbox/nnet/ref/nprtool.html
[8] www.mathworks.com/matlabcentral/.../19997-neural-network-for-pat
[9] iridl.ldeo.columbia.edu/dochelp/StatTutorial/Interpolation/
[10] www.mathworks.in/help/techdoc/ref/spline.html
[11] deafness.about.com
[12] www.youtube.com/watch?v=_BF8ZH_Hut0
[13] www.animalsign.org/animalsign/kninesign-for-dogs.html

A Dynamic Filtering Algorithm to Search Approximate String

[1]A.Dasaradha, [2]P.K.Sahu

[1] Dept. of CSE, Final M.Tech Student, AITAM Engineering College, INDIA,
[2] Dept. of CSE, Associative Professor, AITAM Engineering College, INDIA

Abstract:

Recently string data management in databases has gained lot of interest in various applications such as data cleaning, query relaxation and spellchecking. Hence in this paper we provide a solution how to find similar to a query string from a given a collection of strings. The proposed solution has two phases. In the first phase three algorithms such as ScanCount, MergeSkip, and DivideSkip for answering approximate string search queries. In the second phase, we study on how to integrate various filtering techniques with the proposed merging algorithms. Several experiments have been conducted on various data sets to calculate the performance of the proposed techniques.

Introduction:

Recently string data management in databases has gained lot of interest in text mining. Hence in this paper we study a problem how to find similar to a query string "approximate string search" from a given a collection of strings. This problem can be find in various applications like data cleaning, query relaxation, and spellchecking.

***Spell Checking*:** Given an input document, a spellchecker has to find all possible mistyped words by searching similar words to them in the dictionary. We have to find matched candidates to recommend for words which are not there in the dictionary.

***Data Cleaning*:** A data collection has various inconsistencies which have to be solved before the data can be used for accurate data analysis. The process of detecting and correcting such inconsistencies is known as *data cleaning*. A common form of inconsistency arises when a real-world entity has more than one representation in the data collection; for example, the same address could be encoded using different strings in different records in the collection. Multiple representations arise due to a variety of reasons such as misspellings caused by typographic errors and different formatting conventions used by data sources.

These applications require a high real-time performance for each query to be answered. Hence it is necessary to design algorithms for answering such queries as Efficiently as possible. Many techniques have been designed, such as [1], [2], [3], [4], [5], [6], [7]. These methods assume a given similarity function to quantify the closeness between two strings. Various string-similarity functions have been studied, such as edit distance, cosine similarity and Jaccard coefficient. All these methods use the gram concept which is a substring of a string to be used as a signature of the string. These algorithms rely on inverted lists of grams to find candidate strings, and utilize the fact that similar strings should share enough common grams. Many algorithms [9], [2] [15] mainly focused on "join queries" i.e., finding similar pairs from two collections of strings. Approximate string search could be treated as a special case of join queries. It is well understood that the behavior of an algorithm for answering selection queries could be very different from that for answering join queries. We believe approximate string search is important enough to deserve a separate investigation.

In this paper the proposed solution has two phases. In the first phase, we propose three efficient algorithms for answering approximate string search queries, called ScanCount, MergeSkip, and DivideSkip. The ScanCount algorithm adopts a simple idea of scanning the inverted lists and counting candidate strings. Despite the fact that it is very naive, when combined with various filtering techniques, this algorithm can still achieve a high performance. The MergeSkip algorithm exploits the value differences among the inverted lists and the threshold on the number of common grams of similar strings to skip many irrelevant candidates on the lists. The DivideSkip algorithm combines the MergeSkip algorithm and the idea in the MergeOpt algorithm proposed in [9] that divides the lists into two groups. One group is for those long lists, and the other group is for the remaining lists. We run the MergeSkip algorithm to merge the short lists with a different threshold, and use the long lists to verify the candidates. Our experiments on three real data sets showed that the proposed algorithms could significantly improve the performance of existing algorithms.

In the second phase we study, how to integrate various filtering techniques with the proposed merging algorithms. Various filters have been proposed to eliminate strings that cannot be similar enough to a given string. Surprisingly, our experiments and analysis show that a naive solution of adopting all available filtering techniques might not achieve the best performance to merge inverted lists. Intuitively, filters can segment inverted lists to relatively shorter lists, while merging algorithms need to merge these lists. In addition, the more filters we apply, the more groups of inverted lists we need to merge, and more overhead we need to spend for processing these groups before merging their lists. Thus filters and merging algorithms need to be integrated judiciously by considering this tradeoff. Based on this analysis, we classify filters into two categories: single-signature filters and multi-signature filters. We propose a strategy to selectively choose proper filters to build an index structure and integrate them with merging algorithms. Experiments show that our strategy reduces the running time by as much as one to two orders of magnitude over approaches without filtering techniques or strategies that naively use all the filtering techniques.

The remainder of this paper is organized as follows. In section 2 discuss about related work, section 3 describes about the proposed solution, section 4 explains the experimental setup and section 5 concludes the paper.

Related Work:

Several existing algorithms assume an index of inverted lists for the grams of the strings in the collection *S* to answer approximate string queries on *S*. In the index, for each gram *g* of the strings in *S*, we have a list *lg* of the ids of the strings that include this gram, possibly with the corresponding positional information of the gram in the strings [12] [13] [14].

Heap algorithm [11]: When merging the lists, maintain the frontiers of the lists as a heap. In each step, pop the top element from the heap and increment the count of the record id corresponding to the popped frontier record. Remove this record id from this list, and reinsert the next record id on the list to the heap. Report a record id whenever its count value is at least threshold T. This algorithm time complexity is O(MlogN) and space complexity is O(N).

MergeOpt algorithm [10]: It treats the $T - 1$ longest inverted lists of $G(Q, q)$ separately. For the remaining $N - (T - 1)$ relatively short inverted lists, Use the Heap algorithm to merge them with a lower threshold i.e., 1. For each candidate string, apply binary search on each of the $T - 1$ long lists to verify if the string appears on at least T times among all the lists. This algorithm is based on the observation that a record in the answer must appear on at least one of the short lists. This algorithm is more efficient than Heap algorithm.

Proposed Solution:

Here we present our three proposed merging algorithms.

***ScanCount*:**

In this algorithm [8], we maintain an array (S) of counts for all the string ids. Then scan the inverted lists. For each string id on each list, we increment the count corresponding to the string by 1. Report the string ids that appear at least T times in the lists. The time complexity of the algorithm is O(M) for heap algorithm it is O(MlogN) and The space complexity is O(|S|), where S is the size of the string collection. ScanCount algorithm improves the Heap algorithm by eliminating the heap data structure and the corresponding operations on the heap. The algorithm is formally descried in Figure 1.

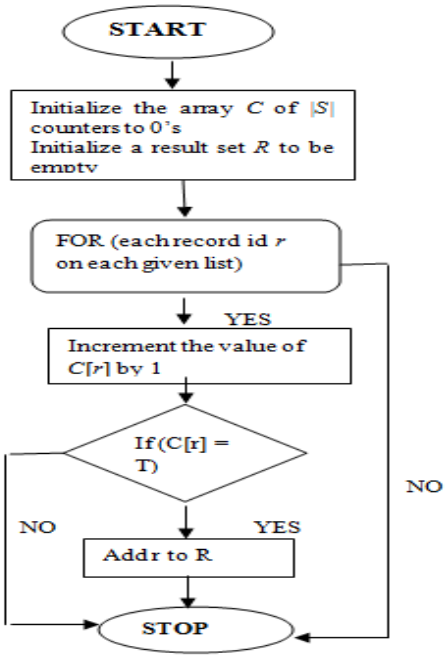


Figure 1: Flowchart for ScanCount Algorithm

MergeSkip:

The main principle of this algorithm, is to skip on the lists those record ids that cannot be in the answer to the query, by utilizing the threshold T. Similar to Heap algorithm, we also maintain a heap for the frontiers of these lists. The key difference is in each iteration, we pop those records from the heap that have the same value as the top record *t* on the heap which is described in figure2.

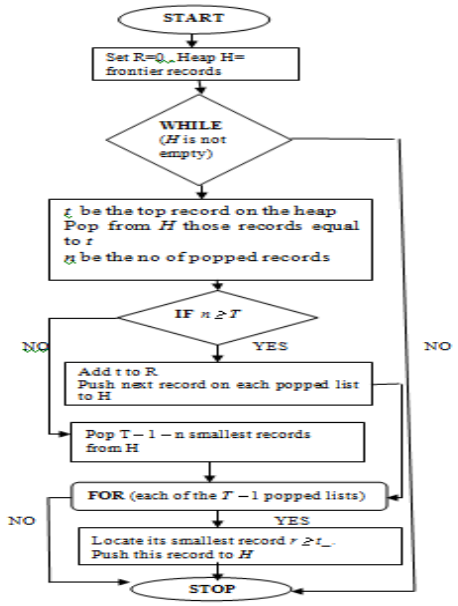


Figure 2: Flowchart for MergeSkip Algorithm

DivideSkip:

The key idea of DivideSkip algorithm is to combine MergeSkip and MergeOpt algorithms. Both the algorithms try to skip irrelevant records on the lists but using different intuitions. MergeSkip exploits the value differences among the records on the lists, while MergeOpt exploits the size differences among the lists. DivideSkip algorithm uses both differences to improve the search performance.

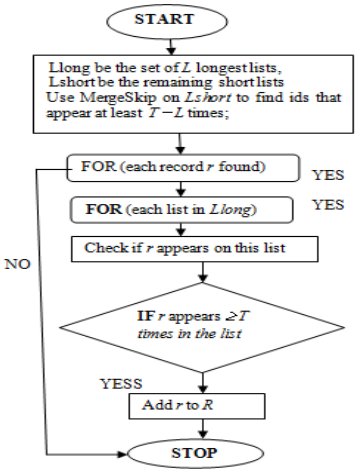


Figure 3: Flow chart for DivideSkip algorithm

Experimental Setup:

The performance of five merging algorithms such as Heap, MergeOpt, ScanCount, MergeSkip, and DivideSkip has been evaluated using DBLP dataset.

DBLP dataset: It includes paper titles downloaded from the DBLP Bibliography site1. The raw data was in an XML format, and we extracted 274,788 paper titles with a total size 17.8MB. The average size of gram inverted lists for a query was about 67, and the total number of distinct grams was 59, 940.

The gram length q was 4 for the data sets. All the algorithms were implemented using GNU C++ and run on a system with 2GB main memory, a 2.13GHz Dual Core CPU and Ubuntu operating system.

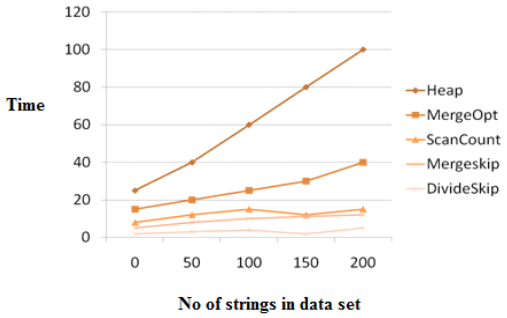


Figure 4: Average query time versus data set size.

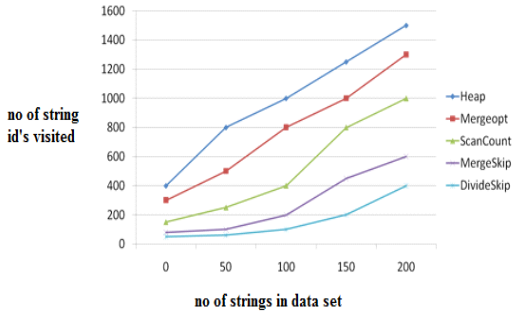


Figure 5: Number of string ids visited by the algorithms.

Classification Of Filters:

A filter generates a set of signatures for a string, such that similar strings share similar signatures, and these signatures can be used easily to build an index structure. Filters are classified into two categories. Single-signature filters generate a single signature (typically an integer or a hash code) for a string and Multi-signature filters generate multiple signatures for a string.

Length Filtering: If two strings *s*1 and *s*2 are within edit distance *k*, the difference between their lengths cannot exceed *k*. Thus, given a query string *s*1, we only need to consider strings *s*2 in the data collection such that the difference between |s1| and |s2| is not greater than *k*. This is a Single signature filter as it generates a single signature for a string.

Position Filtering: If two strings *s*1 and *s*2 are within edit distance *k*, then a *q*-gram in *s*1 cannot correspond to a *q*-gram in the other string that differs by more than *k* positions. Thus, given a positional gram (*i*1, *g*1) in the query string, we only need to consider the other corresponding gram (*i*2, *g*2) in the data set, such that $|i1-i2| \leq k$. This is a multi-signature filter because it produces a set of positional grams as signatures for a string.

Prefix Filtering [10]: Given two q-gram sets G(s1) and G(s2) for strings s1 and s2, we can fix an ordering O of the universe from which all set elements are drawn. Let p(n, s) denote the n-th prefix element in G(s) as per the ordering O. For simplicity, p(1, s) is abbreviated as ps. An important property is that, if $|G(s1) \cap G(s2)| \geq T$, then $ps2 \leq p(n, s1)$, where $n = |s1| - T + 1$.

Applying Filters Before Merging Lists:

All the existing filters can be grouped to improve the search performance of merging algorithms. One way to group them is to build a tree structure called as Filter Tree, in which each level corresponds to a filter which is described in figure 6.

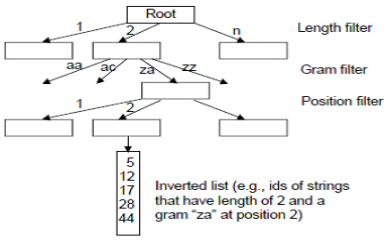


Figure 6: A FilterTree

It is very important to decide which filters should be used on which levels. To improve the performance, use single-signature filters at level 1 (close to the root) such as the length filter and the prefix filter because each string in the data set will be inserted to a single path, instead of appearing in multiple paths. During a search, for these filters we only need to traverse those paths on which the candidate strings can appear. From level 2, we can add those multi-signature ones, such as the gram filter and the position filter. Figure 7 and 8 shows the improved performance of the algorithm using filters in the DBLP dataset.

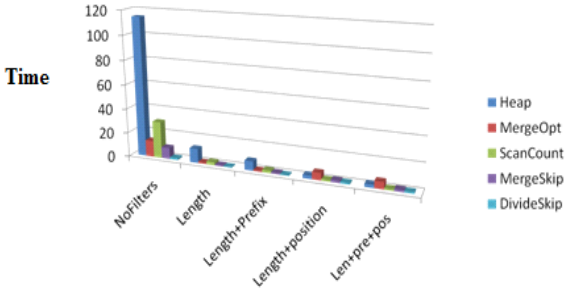


Figure 7: DBLP data set for Merge

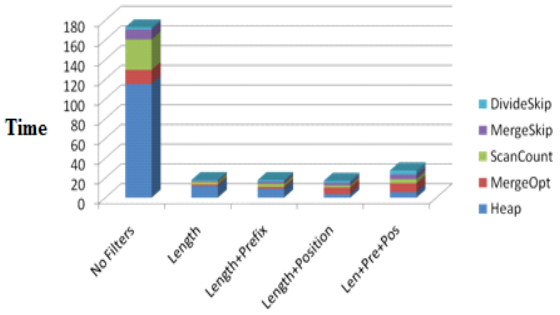


Figure 8: DBLP data set for Total

Conclusion:

In this paper we have proposed a solution for how to efficiently find a collection of strings those similar to a given string. We designed solution in two phases. In the first phase three algorithms such as ScanCount, MergeSkip and DivideSkip for answering approximate string search queries. In the second phase, we study on how to integrate various filtering techniques with the proposed merging algorithms. Several experiments have been conducted on various data sets to calculate the performance of the proposed techniques.

References:

[1] . Arasu, V. Ganti, and R. Kaushik, "Efficient Exact Set-Similarity Joins," in VLDB, 2006, pp. 918–929.

[2] R. Bayardo, Y. Ma, and R. Srikant, "Scaling up all-pairs similarity search," in WWW Conference, 2007.

[3] S. Chaudhuri, K. Ganjam, V. Ganti, and R. Motwani, "Robust and Efficient Fuzzy Match for Online Data Cleaning," in SIGMOD, 2003, pp. 313–324.

[4] L. Gravano, P. G. Ipeirotis, H. V. Jagadish, N. Koudas, S. Muthukrishnan, and D. Srivastava, "Approximate string joins in a database (almost) for free," in VLDB, 2001, pp. 491–500.

[5] C. Li, B. Wang, and X. Yang, "VGRAM: Improving performance of approximate queries on string collections using variable-length grams," in Very Large Data Bases, 2007.

[6] E. Sutinen and J. Tarhio, "On Using q-Grams Locations in Approximate String Matching," in ESA, 1995, pp. 327–340.

[7] E. Ukkonen, "Approximae String Matching with q-Grams and Maximal Matching," Theor. Comut. Sci., vol. 1, pp. 191–211, 1992.

[8] V. Levenshtein, "Binary Codes Capable of Correcting Spurious Insertions and Deletions of Ones," Profl. Inf. Transmission, vol. 1, pp. 8–17, 1965.

[9] S. Sarawagi and A. Kirpal, "Efficient set joins on similarity predicate," in ACM SIGMOD, 2004.

[10] S. Chaudhuri, V. Ganti, and R. Kaushik, "A primitive operator for similarity joins in data cleaning," in ICDE, 2006, pp. 5–16.

[11] G. Navarro, "A guided tour to approximate string matching," ACM Computing Surveys, vol. 33, no. 1, pp. 31–88, 2001.

[12] K. Ramasamy, J. M. Patel, R. Kaushik, and J. F. Naughton, "Set containment joins: The good, the bad and the ugly," in VLDB, 2000.

[13] N. Koudas, S. Sarawagi, and D. Srivastava, "Record linkage: similarity measures and algorithms," in SIGMOD Tutorial, 2005, pp. 802–803.

[14] M.-S. Kim, K.-Y. Whang, J.-G. Lee, and M.-J. Lee, "n-Gram/2L: A space and time efficient two-level n-gram inverted index structure." In VLDB, 2005, pp. 325–336.

[15] Chen Li, Jiaheng Lu, Yiming Lu, "Efficient Merging and Filtering Algorithms for Approximate String Searches", ICDE 2008, IEEE.

Explicit Study on Procedures on Content-Based Image Retrieval in Medical Imaging

[1]Mr. Shivamurthy R C, [2]Dr. B.P. Mallikarjunaswamy
[1] Research Scholar, Department of Computer Science & Engineering,
Akshaya Institute of Technology, Lingapura, Tumkur: 572106, Karnataka, India
[2] Professors, Department of Computer Science & Engg,
Sri Siddharatha Institute of Technology, Maralur, Tumkur: 572105, Karnataka, India

Abstract—the advancement in the field of medical imaging system has lead industries to conceptualize a complete automated system for the medical procedures, diagnosis, treatment and prediction. The success of such system largely depends upon the robustness, accuracy and speed of the retrieval systems. Content based image retrieval (CBIR) system is valuable in medical systems as it provides retrieval of the images from the large dataset based on similarities. There is a continuous research in the area of CBIR systems typically for medical images, which provides a successive algorithm development for achieving generalized methodologies, which could be widely used. The aim of this paper is to discuss the various techniques, the assumptions and its scope suggested by various researchers and setup a further roadmap of the research in the field of CBIR system for medical image database. This is a novel approach to provide a typical insight to the prospective researchers, which is unique of its kind.

Index Terms—Digital Images, Medical Imaging, Cbir**,**

I. INTRODUCTION

Content-based image retrieval (CBIR) is the application of computer vision techniques to the problem of digital image search in large databases. CBIR enables to retrieve the images from the databases [1, 2]. Medical images are usually fused, subject to high inconsistency and composed of different minor structures. So there is a necessity for feature extraction and classification of images for easy and efficient retrieval [3]. CBIR is an automatic retrieval of images generally based on some particular properties such as color Composition, shape and texture [4, 5]. Every day large volumes of different types of medical images such as dental, endoscopy, skull, MRI, ultrasound, radiology are produced in various hospitals as well as in various medical centres [6]. Medical image retrieval has many significant applications especially in medical diagnosis, education and research fields. Medical image retrieval for diagnostic purposes is important because the historical images of different patients in medical centres have valuable information for the upcoming diagnosis with a system which retrieves similar cases, make more accurate diagnosis and decide on appropriate treatment. The term content based image retrieval was seen in literature first time by Kato[7], while describing his experiment of image retrieval of images form a database on the basis of color and shape features. There is a significant amount of growing images databases in medical field images. It is a proven though that for supporting clinical decision making the integration of content based access method into Picture Archiving and Communication Systems (PACS) will be a mandatory need [8]. In most biomedical disciplines, digital image data is rapidly expanding in quantity and heterogeneity, and there is an increasing trend towards the formation of archives adequate to support diagnostics and preventive medicine. Exploration, exploitation, and consolidation of the immense image collections require tools to access structurally different data for research, diagnostics and teaching. Currently, image data is linked to textual descriptions, and data access is provided only via these textual additives. There are virtually no tools available to access medical images directly by their content or to cope with their structural differences. Therefore, visual based (i.e. content-based) indexing and retrieval based on information contained in the pixel data of biomedical images is expected to have a great impact on biomedical image databases. However, existing systems for content-based image retrieval (CBIR) are not applicable to the biomedical imagery special needs, and novel methodologies are urgently needed. Content-based image retrieval (CBIR) has received significant attention in the literature as a promising technique to facilitate improved image management in PACS system [9, 10]. The Image Retrieval for Medical Applications (IRMA) project [10,11] aims to provide visually rich image management through CBIR techniques applied to medical images using intensity distribution and texture measures taken globally over the entire image. This approach permits queries on a heterogeneous image collection and helps in identifying images that are similar with respect to global features. Section 2 highlights about the significance of CBIR in medical imaging followed by methods used for implementation CBIR in Section 3. The recent work done on CBIR is mentioned in Section 4. The issues or research gap from prior work is illustrated in Section 5 followed conclusion in Section 6.

II. SIGNIFICANCE OF CBIR IN MEDICAL IMAGING

There are several reasons why there is a need for additional, alternative image retrieval methods apart from the steadily growing rate of image production. It is important to explain these needs and to discuss possible technical and methodological improvements and the resulting clinical benefits. The goals of medical information systems have often been defined to deliver the needed information at the right time, the right place to the right persons in order to improve the quality and efficiency of care processes [12]. Such a goal will most likely need more than a query by patient name, series ID or study ID for images. For the clinical decision making process it can be beneficial or even important to find other images of the same modality, the same anatomic region of the same disease. Although part of this information is normally contained in the DICOM headers and many imaging devices are DICOM compliant at this time, there are still some problems. DICOM headers have proven to contain a fairly high rate of errors, for example for the field anatomical region, error rates of 16% have been reported [13]. This can hinder the correct retrieval of all wanted images. Clinical decision support techniques such as case based reasoning [14] or evidence based medicine [15,16] can even produce a stronger need to retrieve images that can be valuable for supporting certain diagnoses. It could even be imagined to have Image Based Reasoning (IBR) as a new discipline for diagnostic aid. Decision support systems in radiology [17] and computer aided diagnostics for radiological practice as demonstrated at the RSNA (Radiological Society of North America) [18] are on the rise and create a need for powerful data and metadata management and retrieval. The general clinical benefit of imaging system has also already been demonstrated in [19]. In [20] an initiative is described to identify important tasks for medical imaging based on their possible clinical benefits. It needs to be stated that the purely visual image queries as they are executed in the computer vision domain will most likely not be able to ever replace text based methods as there will always be queries for all images of a certain patient, but they have the potential to be a very good complement to text based search based on their characteristics. Still, the problems and advantages of the technology have to be stressed to obtain acceptance and use of visual and text based access methods up to their full potential. A scenario for hybrid, textual and visual queries is proposed in the CBIR system [21]. Besides diagnostics, teaching and research especially are expected to improve through the use of visual access methods as visually interesting images can be chosen and can actually be found in the existing large repositories. The inclusion of visual features into medical studies is another interesting point for several medical research domains. Visual features do not only allow the retrieval of cases with patients having similar diagnoses but also cases with visual similarity but different diagnoses. In teaching it can help lecturers as well as students to browse educational image repositories and visually inspect the results found. This can be the case for navigating in image atlases. It can also be used to cross correlate visual and textual features of the images.

III. METHODS USED FOR IMPLEMENTING CBIR

Content-based image retrieval hinges on the ability of the algorithms to extract pertinent image features and organize them in a way that represents the image content. Additionally, the algorithms should be able to quantify the similarity between the query visual and the database candidate for the image content as perceived by the viewer. Thus, there is a systemic component to CBIR and a more challenging semantic component.

- Shape Based Method: For shape based image retrieval, the image feature extracted is usually an N dimensional feature vector which can be regarded as a point in a N dimensional space. Once images are indexed into the database using the extracted feature vectors, the retrieval of images is essentially the determination of similarity between the query image and the target images in database, which is essentially the determination of distance between the feature vectors representing the images. The desirable distance measure should reflect human perception. Various similarity measures have been exploited in image retrieval. In our implementation we have used Euclidean distance for similarity measurement.
- Texture Based Method: Texture measures have an even larger variety than color measures. Some common measures for capturing the texture of images are wavelets and Gabor filters where Gabor filters perform better and correspond well to. The texture measures try to capture the characteristics of images or image parts with respect to changes in certain directions and scale of changes. This is most useful for regions or images with homogeneous texture.
- Continuous Feature Selection Method: This method deals with the “dimensionality curse” and the semantic gap problem. The method applies statistical association rule mining to relate low-level features with high-level specialist’s knowledge about the image, in order to reduce the semantic gap existing between the image representation and interpretation. These rules are employed to weigh the features according to their relevance. The dimensionality reduction is performed by discarding the irrelevant features (the ones whose weight are null). The obtained weights are used to calculate the similarity between the images during the content-based searching. Experiments performed show that the proposed method improves the query precision up to 38%. Moreover, the method is efficient, since the complexity of the query processing decreases along the dimensionality reduction of the feature vector.
- With Automatically Extracted MeSH Terms: There is still a semantic gap between the low-level visual features(textures, colors) automatically extracted and the high level concepts that users normally search for (tumors, abnormal tissue)[22]. Proposed solutions to bridge this semantic gap are the connection of visual features to known textual labels of the images

[23] or the training of a classifier based on known class labels and the use of the classifier on unknown cases [24]. Combinations of textual and visual features for medical image retrieval have as of yet rarely been applied, although medical images in the electronic patient record or case databases basically always do have text attached to them. The complementary nature of text and visual image features for retrieval promises to lead to good retrieval results.

- Using Low-Level Visual Features and The image retrieval process consists of two main phases: preprocessing phase and retrieval phase. Both phases are described as follows. The pre-processing phase is composed of two main components: a feature extraction model and a classification model. The input of the pre-processing phase is the original image database, i.e. images from the ImageCLEFmed collection, with more than 66,000 medical images. The output of the pre-processing phase is an index relating each image to its modality and a feature database.
- The Feature Extraction Model: The feature extraction model operates on the image database to produce two kinds of features: histogram features and metafeatures. Histogram features are used to build the feature database, which is used in the retrieval phase to rank similar images. Metafeatures are a set of histogram descriptors, which are used as the input to the classification model to be described later. Histogram features used in this system are:
  - Gray scale and color histogram (Gray and RGB)
  - Local Binary Partition histogram (LBP)
  - Tamura texture histogram (Tamura)
  - Sobel histogram (Sobel)
  - Invariant feature histogram (Invariant) Meta features are calculated from histogram features in order to reduce the dimensionality. These metafeatures are the four moments of the moment generating function (mean, deviation, skewness and kurtosis) and the entropy of the histogram. Each histogram has five associated metafeatures, meaning a total of 30 meta-features with information of color, texture, edges and invariants.

IV. RECENT WORK IN CBIR

Support vector machines (SVM) are extensively used to learn from relevance feedback due to their capability of effectively tackling the above difficulties. However, the performances of SVM depend on the tuning of a number of parameters. It is a different approach based on the nearest neighbor paradigm. Each image is ranked according to a relevance score depending on nearest neighbor distances. This approach allows recalling a higher percentage of images with respect to SVM-based techniques [25] there after quotient space granularity computing theory into image retrieval field, clarify the granularity thinking in image retrieval, and a novel image retrieval method is imported. Firstly, aiming at the Different behaviors under different granularities, obtain color features under different granularities, achieve different quotient spaces; secondly, do the attribute combination to the obtained quotient spaces according to the quotient space granularity combination principle; and then realize image retrieval using the combined attribute function.[26] Then a combination of three feature extraction methods namely color, texture, and edge histogram descriptor is reviewed. There is a provision to add new features in future for better retrieval efficiency. Any combination of these methods, which is more appropriate for the application, can be used for retrieval. This is provided through User Interface (UI) in the form of relevance feedback. The image properties analyzed in this work are by using computer vision and image processing algorithms.

1. Evaluating an emotional response to color images. It is mainly used for the case – base reasoning methodology, emotional evolution of color images values , and also find out fuzzy similarity relational & inter and intra similarities and used for MPEG -7 visual descriptors. [27]
2. 3D Object: The 3D objects make their efficient retrieval technology highly desired. Intelligent query methodology, multiple view and representative query view. [28]
3. Relevance Feedback: Another methodology is classify the query in text or images to relevance / irrelevance set of images to select the positive images. Reference to retrieve the relevance images from databases. [29]

V. RESEARCH GAP

There are various areas to work with for the improvement of the content based image retrieval system. It is already been discussed that the existing techniques may be used to improve the quality of image retrieval and the understanding of user intentions. An approach that combines two different approaches to image retrieval, together with active use of context information and interaction has been proposed. The problem of bridging the semantic gap between high level query which is normally in terms of an example image and low level features of an image such as color, texture, shape and object forced to apply techniques to reduce the semantic gap.

One approach to making a fair assessment of the state of the field is by comparing CBIR applications presented in the literature. However, given the large number of research domains that are included in this technology and its sensitivity As shown in table 1, major recent work that has been published are illustrated brief in accordance with the approaches used by each techniques.

Table 1. Survey on recent researches on implementation of CBIR in Medical imaging

Author	Year	Approach used
Suresh, Shanmugam	2012	Perceptual Hash (P-Hash) Algorithm
Leila, Fatima	2012	3D CBIR
Ashish, Manpreet	2012	Fourier Descriptors, Haar Wavelet, Canberra Distance
Akila, Uma	2012	Trigonometric Function Distance, wavelets
Sanjay, Trimbak	2012	High Level Feature, Low Level Features
ElBagoury, Roushdy	2012	Ranking
Gupta, Khurana	2012	structured local binary Haar pattern, Haar wavelet
Killedar, Patil, Borse	2012	Gray level co-occurrence matrix, Principal Components Analysis, Support Vector Machine
Christiyana, Rajamani, Usha Devi	2012	Two Dimensional Gray level Co-occurrence Matrix, one dimensional Gray level Co-occurrence Matrix,
Yu, Zhang, Liu, Metaxas	2012	Principal Component Analysis
Karthikeyan, Aruna	2012	K-Means Algorithm, ranking, Edge Histograms, Edge Extraction Method, Sobel
Shambharkar, Tirpude	2012	Binary partitioning tree, Minkowski difference, Wavelet
Shambharkar, Tirpude	2012	Binary tree structure, Canny edge detection Mean, Correlation.
Suryanarayana Rao Reddy Babu	2012	Pyramid-Structured Wavelet Transform Energy Level Algorithm Euclidean Distance Algorithm
Singha Hemachandran	2012	Wavelet Based Color Histogram Image Retrieval)
Khokher Talwar	2012	Minkowski-Form distance Euclidean Distance Manhattan distance
Shaila Vadivel	2012	HSV colorspace Smooth distribution, NBS distance, background complex

to the nature and content of the data, it is necessary to develop comparison methods that analyze more than the selection of particular techniques and the experimental results presented in the literature. Rather, it may be better to formally describe an idealized CBIR system and identify the shortcomings in the candidate system. These shortcomings have been labeled as "gaps" and extensively discussed in [30]. The concept of the gap is a generalization of the well-known "semantic gap" that refers to the difficulty of capturing high-level imaged content semantics from extracted low-level image features. These gaps have been broadly categorized into four types and defined below:

1. The Content Gap addresses a system's ability to foster human understanding of concepts from extracted features. In medical applications, it also refers to the extent to which the system adapts to varying modalities, context, and diagnostic protocols.
2. The Feature Gap addresses the extent to which the image features are extracted. This is measured along several dimensions: degree of automation, degree of detail captured along the content axis (object structure), use of multi-scalar techniques, the use of space and (if available) time dimension in image data, and use of all channels on each dimension.
3. The Performance Gap addresses practicalities of system implementation and acceptance. It evaluates system availability, extent of integration into the medical infrastructure, use of feature indexing techniques, and the extent to which the system was evaluated.
4. The Usability Gap measures the richness of available query features and the extent to which they can be combined, available support for comprehending the results returned by the system, and available support for query refinement.
   Addressing these aspects makes a CBIR system more usable, and may increase its acceptance into the medical (clinical, research, or education) workflow.

VI. CONCLUSION

In this work, most of systems use color and texture features, few systems use shape feature, and still less use layout features. Ontological Visual descriptor used extensively in various areas to improve the performance of the system and to achieve better results in different applications. Its integrates various features perfectly in content based image retrieval system and reflects the user's subjective requirements, the experiments achieve good performance and demonstrate the efficiency and robustness of system.. This survey also highlighting the significant contributions of content based image & information's Retrieval field.

The difficulty faced by CBIR methods in making inroads into medical applications can be attributed to a combination of several factors. Some of the leading causes can be categorized according to the "gaps" model presented above.

1. The Content Gap: It is important to consider image content in light of the context of the medical application for which a CBIR system has been optimized. Too often, we find a generic image retrieval model where the goal is to find medical

images that are similar in overall appearance. The critical factor in medical images, however, is the pathology – the primary reason for which the image was taken. This pathology may be expressed in details within the image (e.g., shape of a vertebra or texture and color of a lesion) rather than the entire image (e.g., spine x-ray or cervicographic image). In addition, there may be multiple image modalities that provide the critical information, e.g., histology slides, photographs, etc. In addition to expanding the scope of the CBIR system it is important to also consider analyzing patient histories or physician's notes for valuable information

2. The Feature Gap: Extracted features are used to define the image content. As such, decisions on the types of features, scale(s) at which the features are extracted, and their use individually or in combination determines the extent to which the system "knows" the image and, to a large extent the system capability. It is necessary for the system to support as many types of features as possible and also capture them at several scales. Medical CBIR applications are very sensitive to medical image content. So, developing toolboxes to permit user selection of features may also be very helpful in generalizing the applications and improving acceptance.
3. The Performance Gap: Benefits of medical imaging to science and healthcare have led to an explosive growth in the volume (and rate) of acquired medical images. Additionally, clinical protocols determine the acquisition of these images. There is a need for the system response to be meaningful, timely and sensitive to the image acquisition process. These requirements make linear searches of image feature data, very often presented in the literature, impractical and a significant hurdle in the inclusion of CBIR into medical applications.
4. The Usability Gap: This gap is rarely addressed during the design and development of CBIR systems. However, it is the one of most concern to the end user of the system and therefore has the greatest potential for affecting the acceptance of a new technology.

An idealized system can be designed to overcome all the above gaps, but still fall short of being accepted into the medical community for lack of (i) useful and clear querying capability; (ii) meaningful and easily understandable responses; and (iii) provision to adapt to user feedback. The opposite is also true to some extent. A technically mediocre, but promising, system may obtain valuable end user feedback, and by technical improvement may increase user acceptance with the application of usability design principles. Other than item (iii), which still needs significant research effort, the usability gap can only be bridged by keeping the end user in mind from early system development, as well as by conducting well designed usability studies with targeted users. In general, a high involvement of the user community in system design and development can significantly improve adoption and acceptance.

The preceding subsections already showed the large variability in techniques that are used for the retrieval of images. Still, several very successful techniques from the image retrieval domain have not been used for medical images as of yet. The entire discussion on relevance feedback that first improved the performance of text retrieval systems and then, 30 years later, of image retrieval systems has not at all been discussed for the medical domain.

REFERENCES

[1] M.Smeulders, Worring, and M. Santini, "Content-based image Retrieval at The End of Early Years", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22, No.12, 2000, pp. 1349-1380.

[2] V.S. Murthy, E.Vamsidhar, J.N.V.R. Swarup Kumar, and P. Sankara Rao,"Content based Image Retrieval using Hierarchical and Kmeans Clustering Techniques", International Journal of Engineering Science and Technology, Vol. 2, No. 3, 2010, pp. 209-212.

[3] B. Ramamurthy, and K.R. Chandran, "CBMIR:Shape-based Image Retrieval using Canny Edge Detection and K-means Clustering Algorithms for Medical Images", International Journal of Engineering Science and Technology, Vol. 3, No. 3, 2011, pp. 209-212.

[4] Roberto Parades, Daniel Keysers, Thomas M. Lehman, Berthold Wein, Herman Ney, and Enrique Vidal,"Classification of Medical Images Using Local Representation", Workshop Bildverarbeitung fur die Medizin, 2002, pp.171-174.

[5] Wei Zhang, Sven Dickinson, Stanley Sclaroff, Jacob Feldman, and Stanley Dunn,"Shape –Based Indexing in a Medical Image Database", Biomedical Image Analysis, 1998, pp. 221- 230.

[6] Monireh Esnaashari, S. Amirhassan Monadjami, and Gholamali Naderian,"A Content-based Retinal Image Retrieval Method for Diabetes- Related Eye Diseases Diagnosis", International Journal of Research and Reviews in Computer Science(IJRRCS), Vol. 2, No. 6, 2011, pp. 1222-1227.

[7] Kato, T., "Database architecture for content based image retrieval in Image Storage and Retrieval Systems" (Jambardino A and Niblack W eds),Proc SPIE 2185, pp 112-123, 1992.

[8] David Bandon, Christian Lovis, Antoine Geissbühler, Jean-Paul Vallée, Enterprise-wide PACS: Beyond Radiology, an Architecture to Manage All Medical Images, AUR, 2005 doi:10.1016/j.acra.2005.03.075

[9] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler," A review of content-based image retrieval systems in medical applications-Clinical benefits and future directions", International Journal of Medical Informatics, Vol. 73, No. 1, 2004, pp. 1-23.

[10] T.M. Lehmann, M.O. Guld, C Thies,B Fischer, K. Spitzer, and D. Keysers," Content-based image retrieval in medical applications", Methods of Info in Med, IOS Press , Vol. 43, No. 4, 2004, pp. 354–361.

[11] C. Thies, M.O. Guld, B Fischer, and T.M. Lehmann,"Content-based queries on the CasImage database within the IRMA framework", Lecture Notes in Computer Science,Springer 3491, 2005, pp. 781–792.

[12] A. Winter, R. Haux, A three level graph-based model for the management of hospital information systems, Methods of Information in Medicine 34 (1995) 378{396.
[13] M. O. G□uld, M. Kohnen, D. Keysers, H. Schubert, B. B. Wein, J. Bredno, T. M. Lehmann, Quality of DICOM header information for image categorization, in: International Symposium on Medical Imaging, Vol. 4685 of SPIE Proceedings, San Diego, CA, USA, 2002, pp. 280-287.
[14] C. LeBozec, M.-C. Jaulent, E. Zapletal, P. Degoulet, Uni_ed modeling language and design of a case{based retrieval system in medical imaging, in: Proceedings of the Annual Symposium of the American Society for Medical Informatics (AMIA), Nashville, TN, USA, 1998.
[15] A. A. T. Bui, R. K. Taira, J. D. N. Dionision, D. R. Aberle, S. El-Saden, H. Kangarloo, Evidence{based radiology, Academic Radiology 9 (6) (2002) 662-669.
[16] J.-P. Boissel, M. Cucherat, E. Amsallem, P. Nony, M. Fardeheb, W. Manzi, M. C. Haugh, Getting evidence to prescribers and patients or how to make EBM a reality, in: Proceedings of the Medical Informatics Europe Conference (MIE 2003), St. Malo, France, 2003.
[17] C. E. Kahn, Arti_cial intelligence in radiology: Decision support systems, RadioGraphics 14 (1994) 849-861.
[18] H. Abe, H. MacMahon, R. Engelmann, Q. Li, J. Shiraishi, S. Katsuragawa, M. Aoyama, T. Ishida, K. Ashizawa, C. E. Metz, K. Doi, Computer{aided diagnosis in chest radiography: Results of large-scale observer tests at the 1996-2001 RSNA scienti_c assemblies, RadioGraphics 23 (1) (2003) 255-265.
[19] B. Kaplan, H. P. Lundsgaarde, Toward an evaluation of an integrated clinical imaging system: Identifying clinical benefits, Methods of Information in Medicine 35 (1996) 221-229.
[20] A. Horsch, R. Thurmayr, How to identify and assess tasks and challenges of medical image processing, in: Proceedings of the Medical Informatics Europe Conference (MIE 2003), St. Malo, France, 2003.
[21] S. Antani, L. R. Long, G. R. Thoma, A biomedical information system for combined content-based retrieval of spine x-ray images and associated text information, in: Proceedings of the 3rd Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2002), Ahamdabad, India, 2002.
[22] Müller H, Ruch, P,Geissbuhler A., Enriching content-based medical image retrieval with automatically extracted MeSH.
[23] Juan C. Caicedo, Fabio A. Gonzalez and Eduardo Romero., Content-Based Medical Image Retrieval Using Low-Level Visual Features and Modality Identification.
[24] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang., Image Retrieval: Ideas, Influences, and Trends of the New Age, The Pennsylvania State University.
[25] Giorgio Giacinto "A Nearest-Neighbor Approach to Relevance Feedback in Content Based Image Retrieval"
[26] Xiangli Xu, Libiao Zhang, Zhezhou Yu, Chunguang Zhou "Image Retrieval Using Multi-Granularity Color Features" ICALIP2008 IEEE.
[27] Joonwhoan lee , eunjong park — Fuzzy similarity – based emotional classification of color images□ vol 13,no.5,oct2011.
[28] yue gao,meng wang,zheng-jun zha — less is more : efficient 3-D object retrieval with query view selection□ vol 13, no.5,oct 2011
[29] yue gao,meng wang,zheng-jun zha — less is more : efficient 3-D object retrieval with query view selection□ vol 13, no.5,oct 2011
[30] T. M. Deserno, S. Antani, L. R. Long. Ontology of Gaps in Content-Based Image Retrieval. J Digit Imaging. February 2008. DOI: 10.1007/s10278-007-9092-x.

Author Profiles:

Mr.Shivamurthy R C received the BE degree from PDA college of Engineering, Gulbarga University and received the M.Tech degree in Computer Science & Engineering from Malnad College of Engineering, Visvesvaraya Technological University, Belgaum.
He served as a Bio Medical Engineer in AIMS Hospital & Cancer Research Center. He served as Assistant Professor in B.G.S.Institute of Technology, B.G.Nagar and currently working as professor in the department of Computer Science at A.I.T, Tumkur, Karnataka, and is also a Ph.D scholar in CMJ University, India. He can be contacted at shivamurthyrc@gmail.com

Dr.B.P Mallikarjunaswamy: working as a professor in the Department of Computer Science & Engineering, Sri Siddhartha Institute of Technology, affiliated to Sri Siddhartha University. He has more than 20 years of Experience in teaching and 5 years of R & D. He is guiding many Ph.D scholars. He has published more than 30 technical papers in national and International Journals and conferences. His current research interests are in Pattern Recognition and Image Processing.

Oscillating Supersonic delta wing with Straight Leading Edges

[1]Asha Crasta, [2]S.A. Khan

1. Research Scholar, Department of Mathematics, Jain University, Bangalore,
2. Principal, Department of Mechanical Engineering, P.A College of Engineering, Mangalore

Abstract:
A Supersonic similitude has been used to obtain stability derivatives in pitch and roll of a delta wing with straight leading edge for the attached shock case. Ghosh's strip theory is been used in which strips at different span wise locations are independent of each other. This combines with the similitude to give a piston theory which gives the closed form of solution to stability derivatives in pitch and roll. Some of the results obtained have been compared with those of Hui et al ,Ghosh and Lui &Hui. Results have been obtained for supersonic flow of perfect gas over a wide range of Mach numbers, incidences and sweep angles.

Key Words: Aspect ratio, Attached shock wave, Delta wings, Damping derivative, Leading edge, Rolling derivative, Stiffness derivative, Supersonic Flow, unsteady flow

1. Introduction

Sychev's [1] large incidence hypersonic similitude is applicable to a wing provided it has an extremely small span in addition to small thickness. Cole and Brainered [2] have given a solution for a delta wing of very small span at large incidence. Messiter [3] has found a solution, in the realm of thin shock layer theory, for a steady delta wing with a detached shock; the attached shock case has been studied by Squire. Malmuth [5] obtained an analytical solution for the attached shock case at small incidence based on hypersonic small disturbance theory. Pike [4] and Hui [6]have analytically treated the steady delta wing in supersonic/hypersonic flow with an attached shock.

The role of dynamic stability at high incidence during re-entry or maneuver has been pointed out by Orlik-Ruckemann [7]. The shock-attached relatively high aspect ratio delta is often preferred (Townend) [8] for its high lift and drag ratio. Hui and Hemdan [9] have studied the unsteady shock detached case in the context of thin shock layer theory. Lui and Hui [10] have extended Hui's theory to an oscillating delta. Hui et. al. [11] has treated flat wings of arbitrary plan forms oscillating in pitch in supersonic/hypersonic flow. Ericsson [12] has used embedded Newtonian concept for unsteady flows. Ghosh [14] has developed a large incidence 2-D hypersonic similitude and piston theory; it includes lighthill's and Miles piston theories. This theory has been applied for oscillating plane ogives. Ghosh [15] has extended the large deflection similitude to non-slender cones, quasi cones and shock attached delta wings. This similitude in this paper has been extended to oscillating delta wings with straight leading edges past a supersonic flow.

2. Analysis:

A thin strip of the wing, parallel to the centerline, can be considered independent of the z dimension when the velocity component along the z direction is small. This has been discussed by Ghosh's [16]. The strip theory combined with Ghosh's large incidence similitude leads to the "piston analogy" and pressure P on the surface can be directly related to equivalent piston Mach number M_p. In this case both Mp and flow deflections are permitted to be large. Hence light hill piston theory [17] or miles strong shock piston theory cannot be used but Ghosh's piston theory will be applicable.

$$\frac{P}{P_\infty} = 1 + AM_P{}^2 + AM_P(B + M_P{}^2)^{\frac{1}{2}} \quad \text{, Where } P_\infty \text{ is free stream pressure} \tag{1}$$

Since strips at different span wise location are assumed independent of each other, the strip can be considered as a flat plate at an angle of attack. The angle of incidence is same as that of wing. Angle ϕ is the angle between the shock and the strip. A piston theory which has been used in equation (1) has been extended to supersonic flow. The expression is given below.

$$\frac{p}{p_\infty} = 1 + A\left(\frac{M_p}{\cos\phi}\right)^2 + A\left(\frac{M_p}{\cos\phi}\right)\left(B + \left(\frac{M_p}{\cos\phi}\right)^2\right)^{\frac{1}{2}} \tag{2}$$

Where p_∞ is free stream pressure, $A = \frac{(\gamma+1)}{4}, B = (4/(\gamma+1)^2, \gamma$ is the specific heat ratio and M_p = the local piston Mach number normal to the wedge surface.

2.1 Pitching moment derivatives

Let the mean incidence be α_0 for the wing oscillating in pitch with small frequency and amplitude about an axis x_0. The piston velocity and hence pressure on the windward surface remains constant on a span wise strip of length 2z at x. The pressure on the lee surface is assumed Zero. Therefore the nose up moment is

$$m = -2\int_0^c p.z(x - x_0)dx \tag{3}$$

The stiffness derivative is non-dimensionalized by dividing with the product of dynamic pressure, wing area and chord length.

$$\therefore -C_{m_\alpha} = \frac{2}{\rho_\infty U_\infty{}^2 C^3 \cot\varepsilon}\left(-\frac{\partial m}{\partial \alpha}\right)_{\substack{\alpha=\alpha_0 \\ q=0}} \tag{4}$$

The damping derivative is non-dimensionalised by dividing with the product of dynamic pressure, wing area, chord length and characteristic time factor $\left(\frac{c}{U_\infty}\right)$

$$\therefore -{}^{C}m_q = \frac{2}{\rho_\infty U_\infty c^4 (\cot\varepsilon)}\left(-\frac{\partial m}{\partial q}\right)_{\substack{\alpha=\alpha_o \\ q=0}} \tag{5}$$

The local piston Mach number normal to the wing surface is given by

$$M_p = M_\infty \sin\alpha + \frac{q}{a_\infty}(x - x_o) \tag{6}$$

Where ρ_∞ ,a_∞ are density and velocity of sound in the free stream? Combining (2) through (6), differentiation under the integral sign is performed.

Defining$x_0 = hL, S_1 = \frac{M_\infty \sin\alpha_o}{\cos\phi}$, the derivatives in pitch of a delta wing become equal to

$$-C_{m_\alpha} = \frac{\sin\alpha_o \cos\alpha_o f(S_1)}{\cos^2\phi}\left[\left(\frac{2}{3} - h\right)\right] \tag{7}$$

$$-C_{m_q} = \frac{\sin\alpha_o f(S_1)}{\cos^2\phi}\left[\left(h^2 - \frac{4}{3}h + \frac{1}{2}\right)\right] \tag{8}$$

Where $f(S_1) = \frac{(r+1}{2S_1}[2S_1 + (B + 2S_1{}^2/(B + 2S_1{}^2)^{\frac{1}{2}}]$ (9)

2.2Rolling Damping Derivative:

Let the rate of roll be $\bar{p}$ and rolling moment be L, defined according to the right hand system of reference.

$$\therefore L = 2\int_0^c \left(\int_0^{Z=f(x)} p.zdz\right)dx \tag{10}$$

The piston Mach number is given by

$$M_p = M_\infty \sin\alpha - \frac{z}{a_\infty}\bar{p} \tag{11}$$

The roll damping derivative is non-dimensionalized by dividing with the product of dynamic pressure, wing area, and span and characteristic time factor $\frac{C}{U_\infty}$

$$\therefore -C_{l_P} = \frac{1}{\rho_\infty U_\infty C^3 b \cot\varepsilon}\left(\frac{-\partial L}{\partial p}\right)_{\substack{\alpha=\alpha_0 \\ \bar{p}=0}} \qquad (12)$$

Combining through (10) to (12)

$$\therefore -C_{l_P} = \frac{\sin\alpha_o f(S_1)}{(\cos^2\phi)}\left[\frac{\cot\ \varepsilon}{12}\right] \qquad (13)$$

Where $f(S_1) = \frac{(r+1}{2S_1}[2S_1 + (B + 2S_1^2/(B + 2S_1^2)^{\frac{1}{2}}]$

3. Results and discussions:

The variation of the stability derivatives with pivot position for various Mach numbers and angle of incidence is shown in Figs. 1 to 4. The stiffness and damping derivative have been compared with Hui et al. (Fig 5. to Fig. 8).The Stiffness derivative shows good agreement. The difference in the damping derivative is attributed to the present theory being a quasi-steady one whereas Liu and Hui [13] give an unsteady theory which predicts $C_{m\dot{\theta}}$. The present work invokes strip theory arguments. Hui et al [11] also use strip theory arguments whereby the flow at any span wise station is considered equivalent to an oscillating flat plate flow; this is calculated by perturbing the known steady flat plate flow (oblique shock solution) which serves as the 'basic flow' for the theory. For a pitching wing the mean incidence is the same for all 'strips' (irrespective of span wise location) and hence there is a single 'basic flow' which Hui et al have utilized to obtain closed form expression for stiffness and damping derivatives. They have not calculated the roll damping derivative. For a rolling wing the 'strips' are at different incidences and there is no single 'basic flow'; hence it is doubtful whether approach can be extended to yield a closed form expression for roll damping derivative. Their theory is valid for supersonic as well as hypersonic flows; whereas the present theory also gives closed form expressions for Stiffness & damping derivatives in pitch as well as roll damping derivative. Liu and Hui's [10] theory is more accurate than of Hui et al [11] as far hypersonic flow is concerned.

The present theory is in good agreement with Hui et al [11] for angle of incidence up to thirty degrees and then there is no matching with the results of Hui et al [11]. This may be due to the detachment of the shock wave and stalling of the flow (Fig. 8). The present theory is simpler than both Liu and Hui [9] and Hui et al [13] and brings out the explicit dependence of the derivatives on the similarity parameters S1. (Fig. 9). Fig.10 presents the variation of damping derivatives with pivot position. There is a disagreement with Liu & Hui [9] as well as Hui [13] and the reasons for the disagreement are the same as discussed earlier. Fig. 11, Fig.12, and Fig.13 present the variation of damping derivatives with mean flow incidence for h = 0, h= 1.0 and h = 0.6. They are in disagreement with Liu & Hui [9] as well as Hui [13] for the angle of incidence more than thirty degrees. This may be due to the detachment of the shock wave and stalling of the flow. Fig.14 and Fig. 15 show the dependence of Roll damping derivative with Mach number and with the aspect ratio. The roll damping derivative decreases with Mach number initially then confirms the Mach number independence principle for large Mach numbers. Further, the roll damping derivative increases with aspect ratio of the wing.

There was an error in the formulation in Ghosh [16] for roll damping derivatives and hence in the present work the same has been corrected and implemented

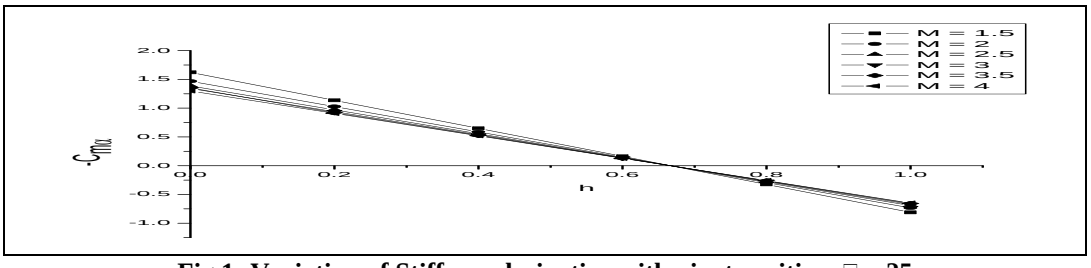


Fig.1: Variation of Stiffness derivative with pivot position □₀ =25

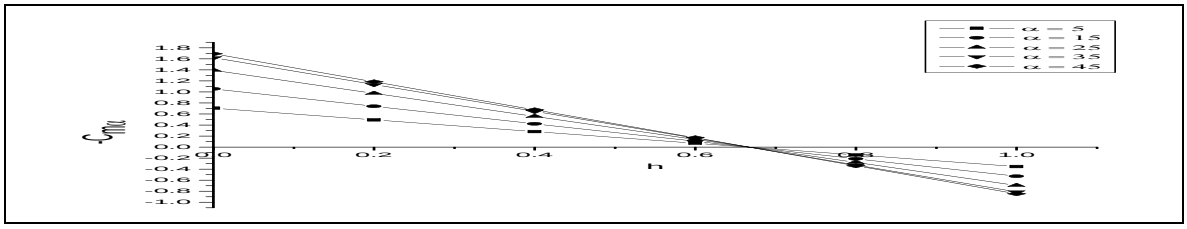


Fig. 2: Variation of stiffness derivative with the pivot position

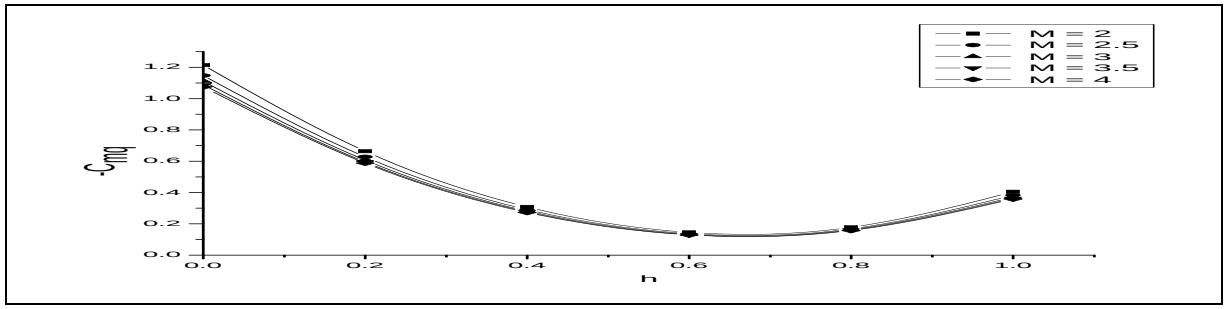


Fig. 3: Variation of Damping derivative with pivot position α_0=25

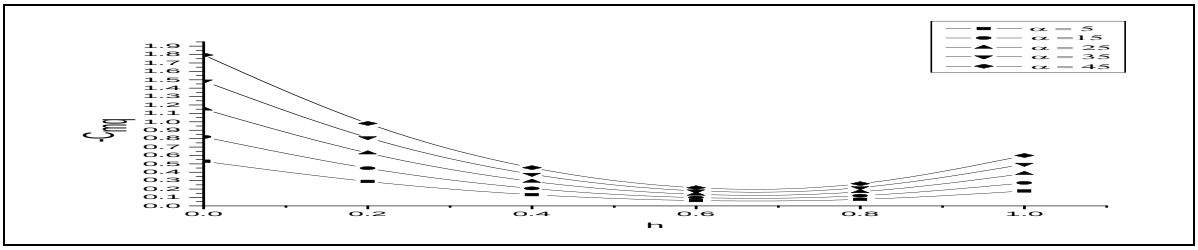


Fig. 4: Variation of damping derivative with pivot position

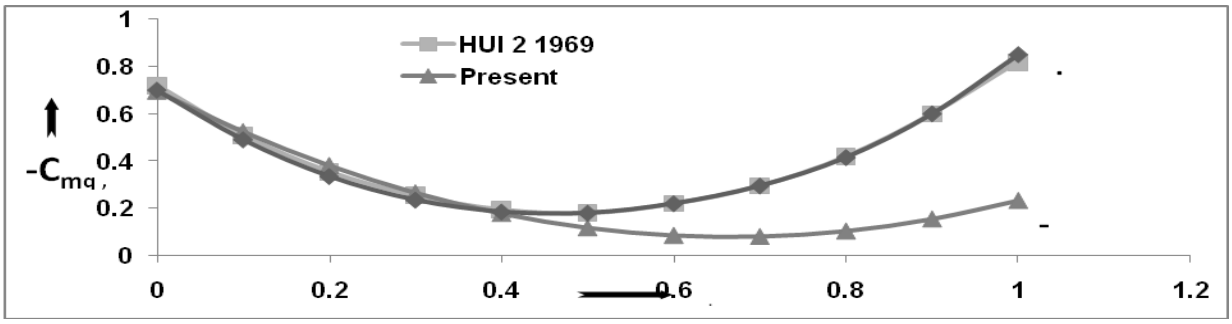


Fig. 5: Variation of Damping Derivative of a Wing with Pivot position with M=3, α_0 =10 ,γ=1.4

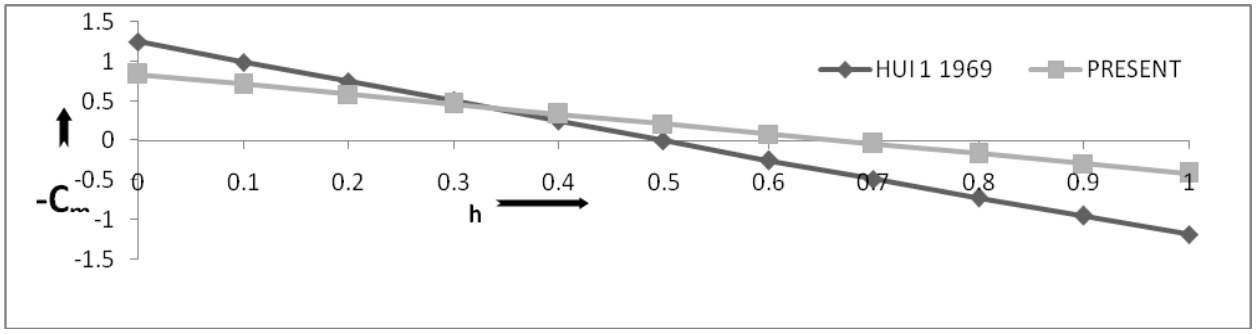


Fig. 6: Variation of Stiffness Derivative of a Wing with Pivot position with M=2.47, α_0 =6°51', γ =1.4

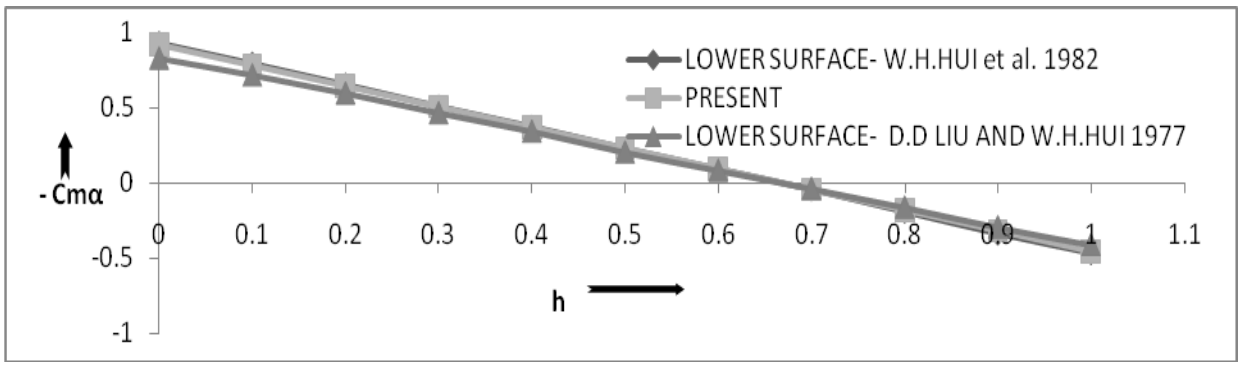


Fig. 7: Comparison of Stiffness Derivative with Theory of Liu and Hui for Triangular Wing with M=4, $\square_0 = 15^0$, $\gamma = 1.4$

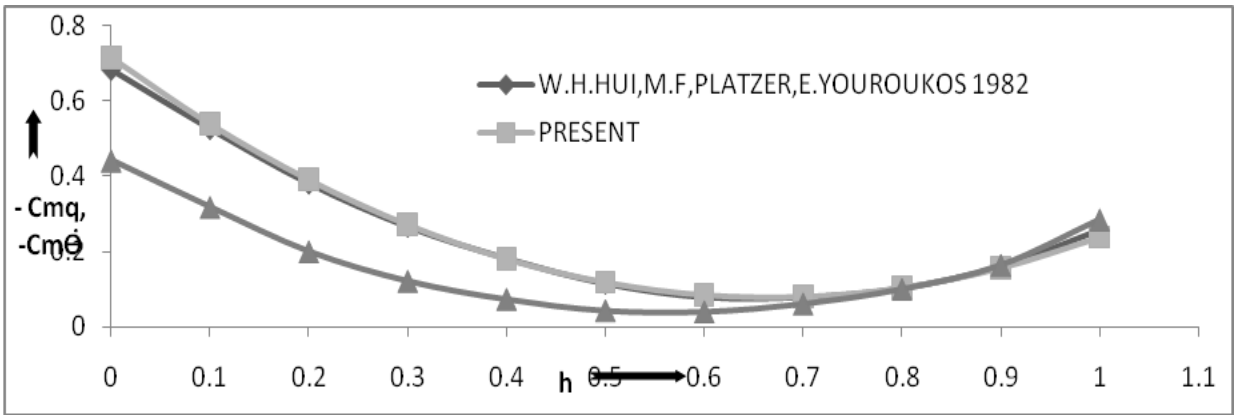


Fig. 8: Comparison of Damping Derivative of Triangular Wing with M=4, $\square_0 = 15^0$, $\gamma = 1.4$

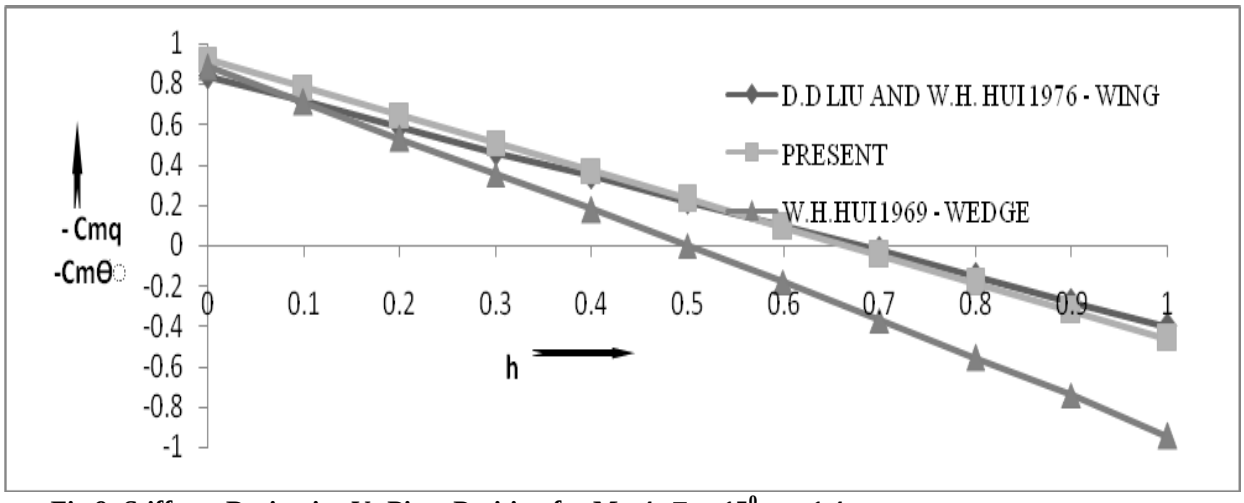


Fig.9: Stiffness Derivative Vs Pivot Position for M= 4, $\square_0 = 15^0$, $\gamma = 1.4$

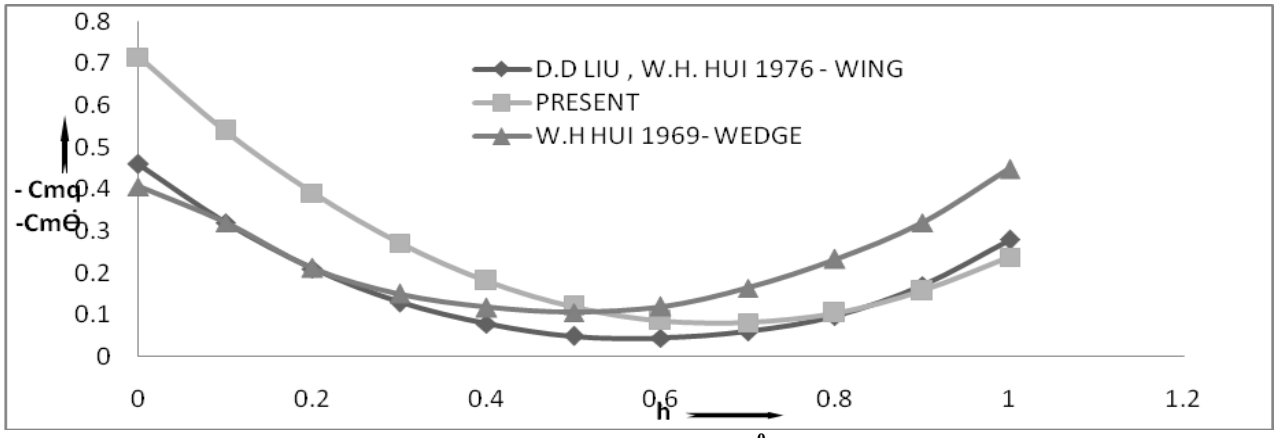


Fig. 10: Damping Derivative Vs Pivot Position for M=4, $\alpha_0 = 15^0$, γ=1.4

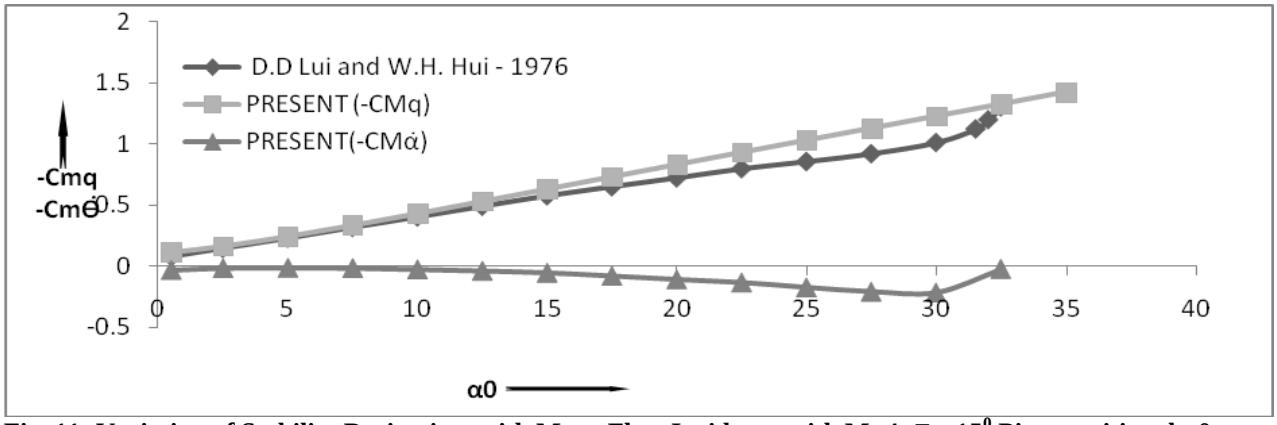


Fig. 11: Variation of Stability Derivatives with Mean Flow Incidence with M=4, α_0=15^0,Pivot position, h=0

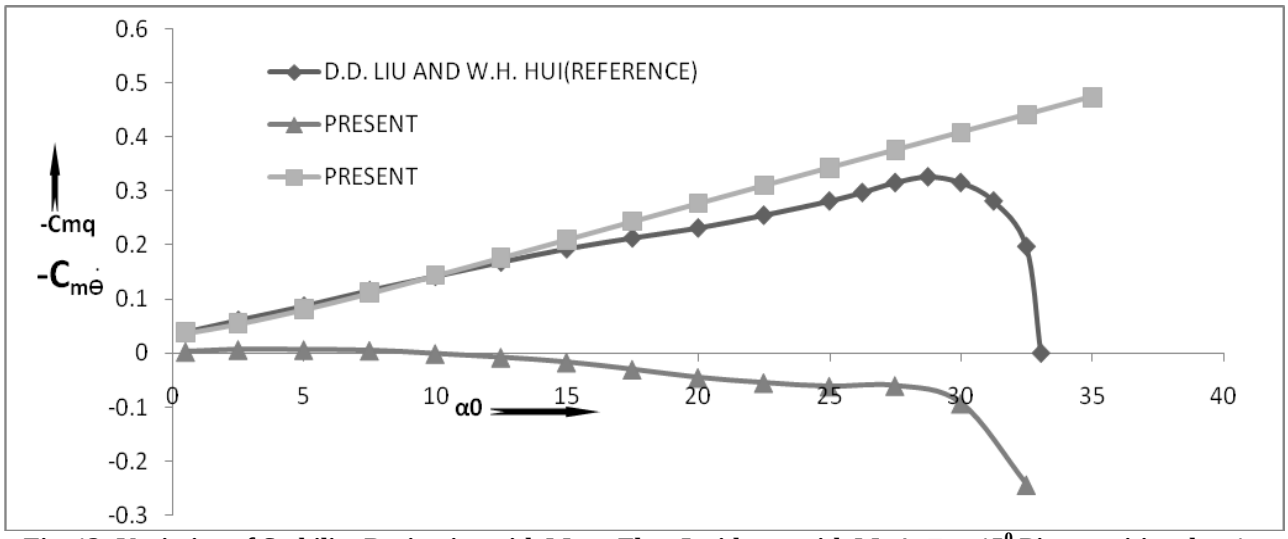


Fig. 12: Variation of Stability Derivative with Mean Flow Incidence with M=4, $\alpha_0 = 15^0$ Pivot position, h = 1

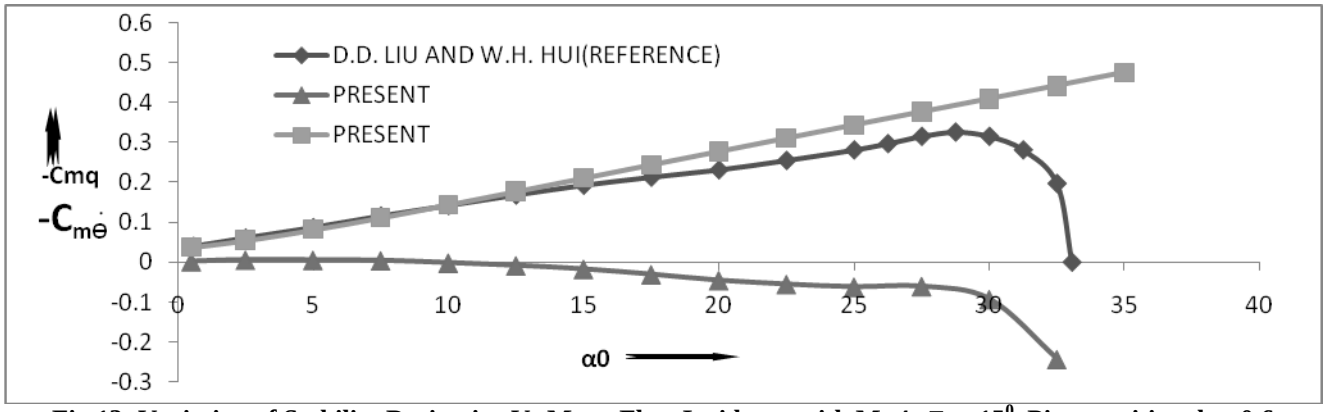


Fig.13: Variation of Stability Derivative Vs Mean Flow Incidence with M=4, $\alpha_0 = 15^0$, Pivot position, h = 0.6

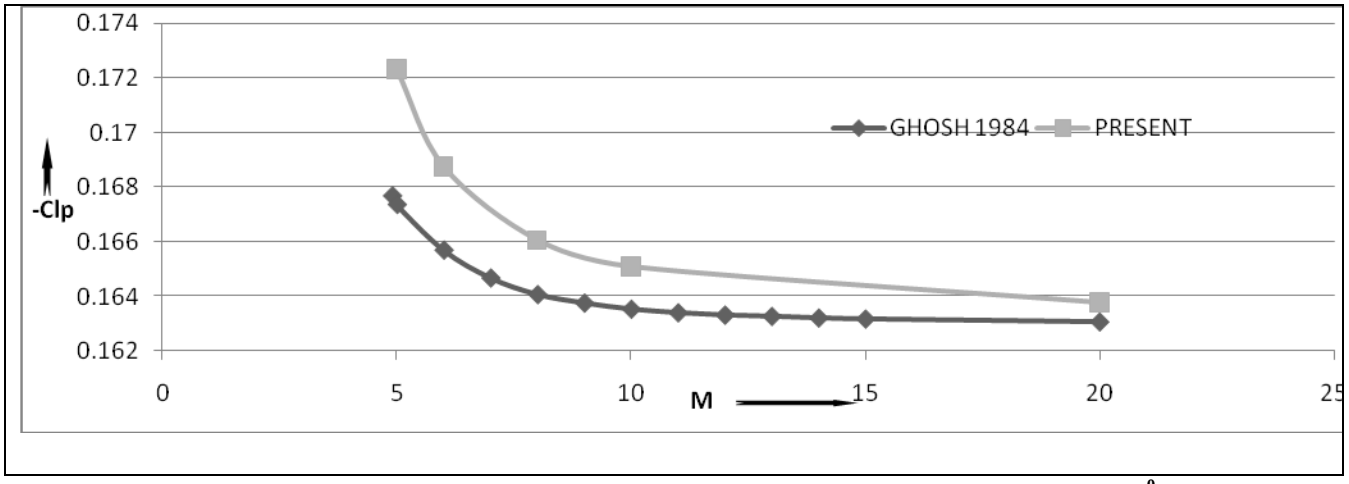


Fig.14: Rolling Moment Derivative Vs Mach number with Aspect Ratio, AR= 4.764, $\alpha_0 = 25^0$, γ=1.4

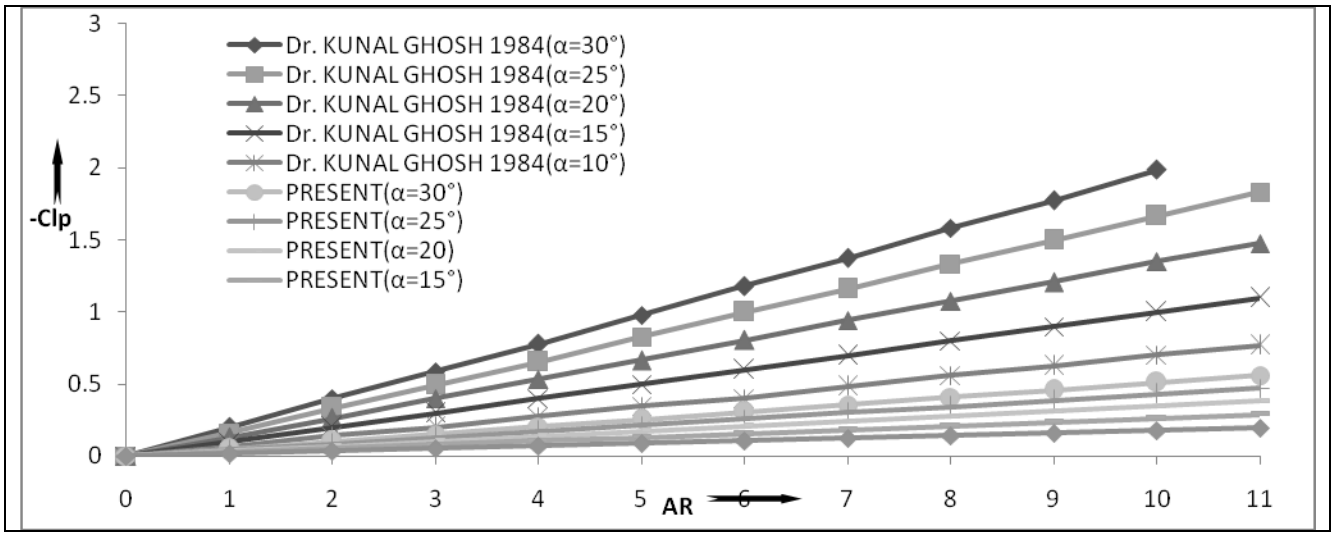


Fig.15: Roll Damping Derivative Vs Aspect Ratio of Delta Wings with M = 10

References

[1.] Sychev, V. V, "Three Dimensional Hypersonic Gas flow Past Slender Bodies at High Angles of Attack", Journal of Applied Mathematics and Mechanics", Vol. 24, Aug. 1960, pp. 296-306.

[2.] Cole, J.D. and Brainerd, J. J., "Slender wings at high angles of attack in hypersonic flow", ARS Reprint 1980-61, 1961.

[3.] Messiter, A.F., "Lift of slender delta wings according to Newtonian theory", AIAA Journal, 1963, 1, 794-802.

[4.] Pike. J.The pressure on flat and anhydral delta wings with attached shock waves, the Aeronautical Quarterly, November 1972, XXIII, Part 4, pp. 253-262.

[5.] Malmuth, N. D., "Hypersonic flow over a delta wing of moderate aspect ratio", AIAA Journal, 1966, 4, pp. 555-556.

[6.] Hui, W. H., "Supersonic and hypersonic flow with attached shock waves over delta wings", Proc. of Royal Society, London, 1971, A. 325, pp. 251-268.

[7.] Orlik-Ruckemann, K.J., "Dynamic stability testing of aircraft needs versus capabilities", Progress in the Aerospace Sciences, Academic press, N.Y., 1975, 16, pp. 431- 447.

[8.] Towend, L. H., "Some design aspects of space shuttle Orbiters'", RAE TR 70139, 1970.

[9.] Hui, W. H. and Hemdan, H. T., "Unsteady hypersonic flow over delta wings with detached shock waves", AIAA Journal, April 1976 , 14, pp. 505-511.

[10.] Lui D. D. and Hui W. H., "Oscillating delta wings with attached shock waves", AIAA Journal, June 1977, 15, 6, pp. 804-812.

[11.] Hui,W. H., Platzer, M. F. and Youroukos E., "Oscillating supersonic/hypersonic wings at high incidence", AIAA Journal, March 1982, 20, pp. 299-304.

[12.] Ericsson L. E., "Viscous and Elastic Perturbation Effects on Hypersonic Unsteady Airfoil Aerodynamics", AIAA Journal, Vol.15 Oct. 1977, pp. 1481-1490.

[13.] Hui, W. H., "Stability of Oscillating Wedges and Caret Wings in Hypersonic and Supersonic Flows", AIAA Journal, Vol. 7, August 1969, pp. 1524-1530.

[14.] Ghosh, K. and Mistry, B. K., "Large incidence hypersonic similitude and oscillating non-planar wedges", AIAA Journal, August 1980, 18, 8, 1004-1006.

[15.] Ghosh K., "Hypersonic large-deflection similitude for quasi wedges and quasi-cones", The Aeronautical Journal, March 1984, 88, 873, pp. 70-76.

[16.] G hosh K., "Hypersonic large deflection similitude for oscillating delta wings", The Aeronautical journal, Oct. 1984, pp. 357-361.

[17.] LightHill M. J., "Oscillating Aerofoil at High Mach Numbers, Journal of Aeronautical Sciences", Vol. 20, June 1953, pp. 402-406.

Securing IPv6's Neighbour Discovery, using Locally Authentication Process

[1]M. N. Doja, [2]Ravish Saggar
[1] Department of Computer Engineering, Jamia Millia Islamia, New Delhi.
[2] Research Scholar Singhania University, Faculty: BCIIT, New Delhi.

Abstract :
Internet Engineering Task Force (IETF), in IPv6, allowed nodes to Autoconfigure using neighbour discovery protocol. Neighbour Discovery (ND) and Address auto- configuration mechanisms may be protected with IPSec Authentication Header (AH). Protecting all traffic will include Address Resolution Protocol. To protect this, IPSec will need agreed Key. For Key setup, UDP packet is sent, which requires IPSec for secure communication. So IPSec requires Agreed Key and for Key setup IPSec is needed, this creates a loop. To solve this problem Locally Authentication Process is presented in this paper. This process will provide a certificate of ownership of IP address on network Interface card and Public key to provide authorization. On the other hand, it will also reduce the network load.

Keywords : Stateless Address Auto-configuration, Neighbour Discovery, Cryptographically Generated Address (CGA), Secure Neighbour Discovery (SEND), Public Key Infrastructure (PKI), Digital Certificate, Security Attacks in IPv6.

1. Introduction

The availability of IPv4 addresses is exhausting due to massive growth of the internet and the proliferation of internet-connected devices other than computers like mobile phones, PDAs etc. The used IP version 4(IPv4) was developed long time back. By the end of 2012, the number of mobile-connected devices will exceed the number of people on earth, and by 2016 there will be 1.4 mobile devices per capita[1]. IPv4 address space is of 32 bits. The theoretical limit of IPv4 addresses is 4.3 billion addresses. The aggravate problem of exhaustions of addresses, was mitigated by the introduction of Classless Inter-Domain Routing (CIDR), and reduced even more by the adoption of Network Address Translators (NAT). Other problems facing IPv4 are the lack of deployed security, and the rapid growth of the size of the routing tables. Before implementing CIDR the backbone routing table was growing at very high rate as compare to memory technology. The Internet Engineering Task Force (IETF) designed a next generation protocol Internet Protocol version 6 (IPv6) to solve these problems and eventually replacing the existing Internet Protocol, IPv4. This IPv6 was designed after having the rich experience of almost 30 years, of IPv4.

Apart from making large address space of 128 bits in IPv6, IETF added many new features. This includes address auto-configuration, host discovery, optimized header, protocol extensibility etc. In IPv4, the configuration of IP addresses is done manually by the network administrator or with the help of DHCP server. Apart from manual configuration and state full auto-configuration, using DHCP, IPv6 has stateless auto-configuration. Stateless auto-configuration does not require manual configuration of hosts, and additional servers. It allows a host to generate its own addresses using a combination of locally available information and information advertised by routers. In state-full auto-configuration, hosts obtain interface addresses and/or configuration information and parameters from a server.

2. Neighbour Discovery Protocol

The Neighbour Discovery protocol is improvement over many process defined in IPv4. New functionality has also been added. The neighbour Discovery Protocol is used for following purposes by nodes.

2.1. For Autoconfiguration of IPv6 Address.
2.2. To determine network prefix,routers and other parameters.
2.3. For Duplicate IP address detection (DAD).
2.4. To determine layer two address of nodes on the same link.
2.5. To find neighbouring routers that can forward their packet.
2.6. To keep track of which neighbours are reachable and which are not (NUD).
2.7. To detect changed link-layer addresses.

To perform all above mentioned work ND uses five ICMPv6 messages: a pair of Router Solicitation / Router Advertisement messages, a pair of Neighbour Solicitation / Neighbour Advertisement messages and an ICMP Redirect message.

3. ICMPV6 Messages

Following ICMP messages are used by Neighbour Discovery Protocol.

3.1. Router Advertisement: This message is used by Routers to inform other nodes existing on all links, to which they are connected, of its presence and other link related information. The process occurs periodically or in response to a Router Solicitation message.

3.2. Router Solicitation: Upon the enabling of an interface of a node, these messages can be used to request all routers on the same local link to send Router Advertisements immediately, rather than waiting until the next periodically scheduled advertisement.

3.3. Redirect : These messages are used by routers to tell hosts that a better on-link router exists for a given destination address.

3.4. Neighbour Solicitation: These messages have 3 main purposes. The first is to discover the link layer address of a neighbour as part of the address resolution process. This process replaces the use of ARP requests and replies in IPv4. The second purpose is to determine the reachability of a neighbour. The last is to detect the presence of duplicate IPv6 addresses during the address auto configuration process which is detailed later in this report.

3.5. Neighbour Advertisement: These messages are either in response to Neighbour Solicitations, or sent by a neighbour to announce a change in its link layer address. Upon receipt of a Neighbour Advertisement, a node will update its neighbour cache which contains mappings between IPv6 and link layer addresses of neighbours.

4. Address Auto-Configuration

Stateless Auto-configuration is the one of the most useful feature that lies in IPv6. The configuration can be done automatically without using any specific protocol such as DHCP. This very feature enables an IPv6 host to configure link-local (an address having link-only scope that can be used to reach neighboring nodes attached to the same link), site-local (an address having scope that is limited to the local site) and global addresses (an address with unlimited scope) for each of its interface. This feature is obtained by the protocol called Neighbor Discovery Protocol (NDP). This protocol includes router (a node that forwards IP packets not explicitly addressed to itself) discovery, stateless address auto-configuration, address resolution, neighbor reachability, duplicate address detection and redirection.

The address auto-configuration assumes that only the trustworthy nodes will form a network, where the communication has to take place. That is the nodes in local link know each other well. But this is not the case every time. Any malicious node or untrustworthy node can manage to reside in the local link network. This node can affect all the other nodes. This is where security factor comes in. IPv6 should make sure that no such malicious node should be able to join the network providing harm to others.

5. Address Auto-Configuration Process

The sequence of signals generated during stateless Auto-configuration is given in Figure 1.

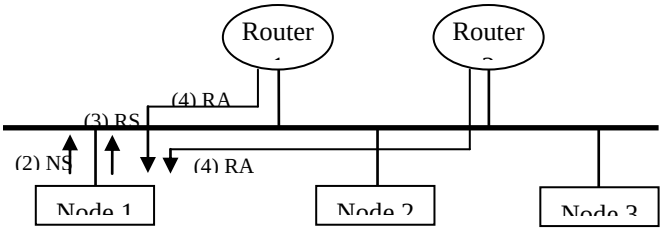


Figure 1. Process of address auto-configuration

The steps are as follows:

5.1. At the very first, a link local address is being generated by the new node. It then allocates it to one of the its interface. This link local address contains the prefix of fe80:: /64 and the 64 bit interface id as shown in Figure 2. This address is tentative address.

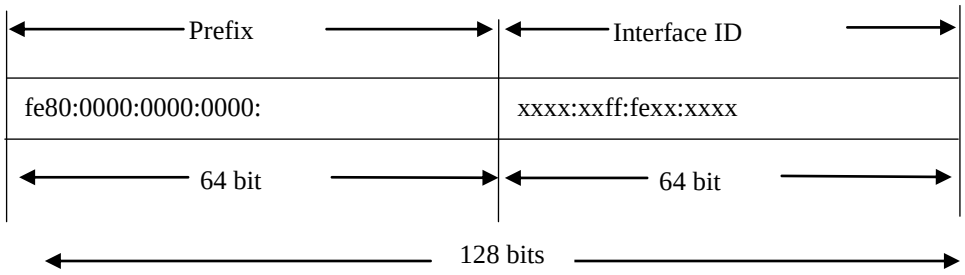


Figure 2. link-local address in ipv6

The node joins the following multicast groups. The all node multicast group(FF02::1)and the solicited-node multicast group, for the tentative address.

5.3.1. After this there is Duplicate Address Detection (DAD), which takes care that the newly generated link local address is not already present on the network. The steps are as follows:
A Neighbor Solicitation (NS) message is being transmitted by the node, with target tentative address, on the network.

5.3.2. If any another node on the network is using the same link local address, then it returns the Neighbor Advertisement (NA) message.

5.3.3. If the new node gets the NA message, no link-local will be allocated to it and the interface will terminate. Else the new node will be initialized buy its link-local address and is assumed to be unique and valid.

5.2. Now the host will send Router Solicitation (RS) messages.

5.3. After this, all those routers who received the RS message will send back the Router Advertisement (RA) message.

5.5.1. If no RA is received, then node uses a state-full address configuration to obtain its addresses.

5.5.2. Else, the node receives the RA and gets the IPv6 address prefix.

5.4. Once the prefix is obtained by the host, it can create its unique site local and global addresses.

6. Threats in Address Auto-configuration

The stateless address auto-configuration allows a host to connect to the network without registering /authenticating itself. It simply configures the address and start communicating with other nodes on the network. Since node does not have to authenticate itself, any malicious node can get access to network. It can cause various kinds of threats which are explained as follows:

6.1. Multiple Interface Address:

IPv6 interface can have multiple addresses. Attacking node, using auto-configuration, can block large number of IP addresses for its interface and deny other workstation to acquire address. This poses a possible denial of service attack.

6.2. Spoofing:

Spoofing is a way to achieve denial of service (DoS) attack, in an IPv6 network, in the Duplicate Address Detection (DAD) procedure. Attacker on the local link waits until a node sends a Neighbor Solicitation packet. The attacker falsely responds with a Neighbor Advertisement packet, informing the new node that it is already using that address. Upon receiving the Neighbor Advertisement, the new node generates another address and repeats the DAD procedure; the attacker again falsely responds with a Neighbor Advertisement packet. Finally, the new node stops initializing its interface.

6.3. Redirect Attack:

Another big threat is in Router Solicitation / Advertisement message. In Neighbor Discovery, attacker can make fake advertisement of itself as default router, causing immediately timeout of all other default routers as well as all on- link prefixes. Node received advertisement and start forwarding its packets to this particular router causes man in middle and DoS attack.

To prevent all these threats some security measures are to be taken. If not secured, it is vulnerable to various attacks. The following sections will describe some existing and proposed solutions.

7. Existing Solutions

There are few solutions to prevent these threats. These are as following:

7.1. IPSec

Internet Protocol Security is meant for protecting the communication over the IP network. It supports network-level peer authentication, data origin authentication, data integrity, and data confidentiality (encryption) and replay protection. It basically uses the cryptographic security services for protection or authentication and encrypts each IP packet of a communication session. These can be either between a pair of nodes, or between a pair of security gateways or between a security gateway and a node.

It is an open standard and makes use of the following 3 basic protocols:

- Authentication Header
  AH provides connectionless integrity and data origin authentication for IP datagram and provides protection against replay attacks. That is it can be used by the nodes to authenticate the neighbor advertisement and the router advertisement messages.
- Encapsulating Security Payloads
  ESP provides confidentiality, data origin authentication, connectionless integrity, an anti-replay service and limited traffic flow confidentiality.
- Security Associations

SA provides number of algorithms and data that provide the parameters necessary to operate the AH and/or ESP operations. It is used to make sure that the nodes in a network are trustworthy. It depends upon the addresses generated by the neighbor discovery, security keys etc. This SA needs to be set up between all the communicating nodes in advance. It can either use the manual mechanism or can be done automatically. As the networks are growing and there are more and more nodes under one network the number of SAs in a single network also increases. This large number of SAs is difficult to maintain. Another problem is that IPSec needs to use NDP for configuring security association, but NDP is requiring IPSec. And hence this is not clear approach to deploy.

7.2. SEcure Neighbor Discovery (SEND)

The SEcure Neighbor Discovery (SEND) protocol is a security extension to the Neighbor Discovery protocol in IPv6. This protocol provides an alternative way for securing NDP. It makes use of a cryptographic method for the same. This protocol came with three additional capabilities: address ownership proof, message protection and router authorization.

It defines a set of new attributes

7.2.1 Cryptographically Generated Addresses(CGA)

Cryptographically Generated Addresses are used to make sure that the sender of a Neighbor Discovery message is the "owner" of the claimed address. A public-private key pair is generated by all nodes before they can claim an address. A new NDP option, the CGA option, is used to carry the public key and associated parameters.

7.2.2 Reply Attack

In order to prevent replay attacks, two new Neighbor Discovery options, Timestamp and Nonce, are introduced. Given that Neighbor and Router Discovery messages are in some cases sent to multicast addresses, the Timestamp option offers replay protection without any previously established state or sequence numbers. When the messages are used in solicitation-advertisement pairs, they are protected with the Nonce option.

7.2.3 RSA Signature option

A new NDP option, the RSA Signature option, is used to protect all messages relating to Neighbor and Router discovery. Public key signatures protect the integrity of the messages and authenticate the identity of their sender.

7.2.4 New network discovery messages

Certification paths, anchored on trusted parties, are expected to certify the authority of routers. A host must be configured with a trust anchor to which the router has a certification path before the host can adopt the router as its default router.

The SEcure Neighbor Discovery (SEND) uses Crypto-Generated Address (CGA), to make sure that the sender of a Neighbor Discovery (ND) message is the “owner” of the claimed address. CGA is a technique whereby an IPv6 address of a node is cryptographically generated by using a one-way hash function from the node’s public key and some other parameters. This CGA is used to prevent the stealing or spoofing of existing IPv6 addresses by assuring only ownership. But the issue is that because CGAs themselves are not certified, an attacker can create a new CGA from any subnet prefix and its

own (or anyone else's) public key. However, the attacker cannot take a CGA created by someone else and send signed messages that appear to come from the owner of that address.

8. Proposed Solution

This solution envisages that only those nodes will be able to join the networks which have been authenticated by issuing valid token, issued by local trusted node. The basic purpose of token is to allow node to verify link local address and its owner ship on Public key.

8.1. The basic terminologies used are:

- Certification Authorities [CA]
  Certification Authorities are responsible for the processing of request of issue, renewal and revoking of PKCs stored in request server.
- Authentication Server [AS]
  Authentication Server is a user an trusted Anchor . There is one AS for every subnet to provide authentication of any new node joining the subnet.
- Public Key [Pu(X)(Y)]
  Pu stands for Public key. X denotes entity that has generated this key and Y identifies the entity for which it is generated. Like Pu(AS)(N) defines Public key generated by AS for node N.
- Private Key [Pr(X)(Y)]
  Pr stands for Private key. X denotes entity that has generated this key and Y identifies the entity for which it is generated. Like Pu(AS)(N) defines Private key generated by AS for node N.
  Database [DB]
  It is the place where AS stores the necessary data needed at the time of authenticating the node.
- Digital Certificate DC(X)
  Digital Certificate issued by X .
- Digital Signature DS(X)(Y)
  Digital Signature issued by X and generated using key Y.
- Message Digest MD(X)
  Message converted into fixed size encrypted message. X represents the parameters which were converted into digest number.
- Manufacturing Company [MC]
  Here Company refers to those companies which are involved in the manufacturing of NIC (Network Interface Card) of the node wishing to participate in the communication in a network.
- Tentative Address [TA]
  An IP Address Generated by node before it converted into permanent Address.
- Cryptographically Generated Address [CGA]
  Cryptographically Generated Address use to authenticate sender.
- Token [TN(X)]
  The Token is a Digital signature generated by AS using public key Pr(AS)(AS) of public key of AS Pu(AS)(AS) and TA and is issued to node X.

8.2. Working Process:

This solution assures that no node can communicate in network without having a Token, issued by AS. To protect the network from any intrusion node the following processes are used:

8.2.1. Request for Private Key and Public Key pair from CA: The Manufacturing Company requests CA to issue Private/Public Key pair Pr(CA)(MC)/Pu(CA)(MC) and digital Certificate, as shown in Fig.3. The CA issues Digital certificate DC(CA) that contains a Public Key Pu(CA)(MC) and the identity of the owner and makes it available publicly. The matching private key Pr(CA)(MC) is given to company, which keeps Private Key top secret. This certificate is a confirmation by the CA that the public key contained in the certificate belongs to company noted in the certificate.

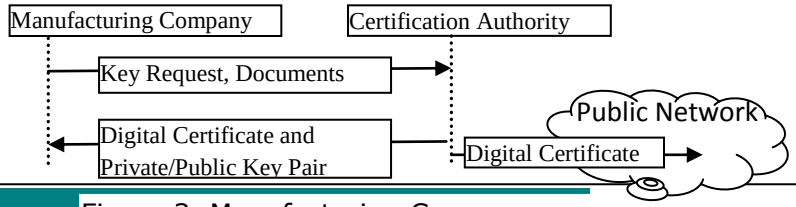


Figure 3. Manufacturing Company

8.2.2. Hard Wired Information in NIC

The company who is manufacturing the network interface card, to be installed on the nodes interested in communication over a network, generates digital signature DS(MC)(Pr(CA)(MC)) from NIC Number using private key Pr(CA)(MC), received from CA. The Digital certificate, digital signature and NIC number is hard-coded into interface card as shown in figure 4. This Hard coded information is used to verify the NIC by AS, using public key Pu(CA)(CA) provided by the CA.

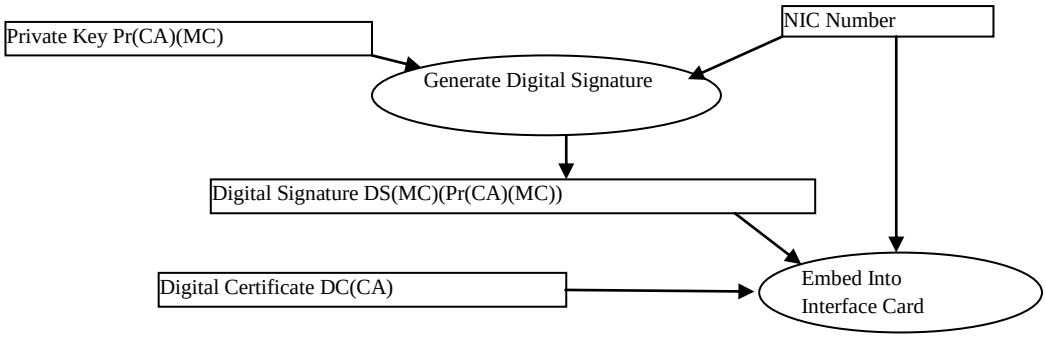


Figure 4. Generation of Information to be embedded into Interface Card

8.2.3. Digital Signature by Node

Node first generates TA and private/public key pair Pr(N)(N) and Pu(N)(N). The DC(CA), DS(MC)(Pr(CA)(MC)), TA, Public Key Pu(N)(N) and NIC number are collectively converted into fixed length message digest. This message digest is then encrypted using private key Pr(N)(N) to get digital signature DS (N)(Pr(N)(N)) as shown in figure 5.

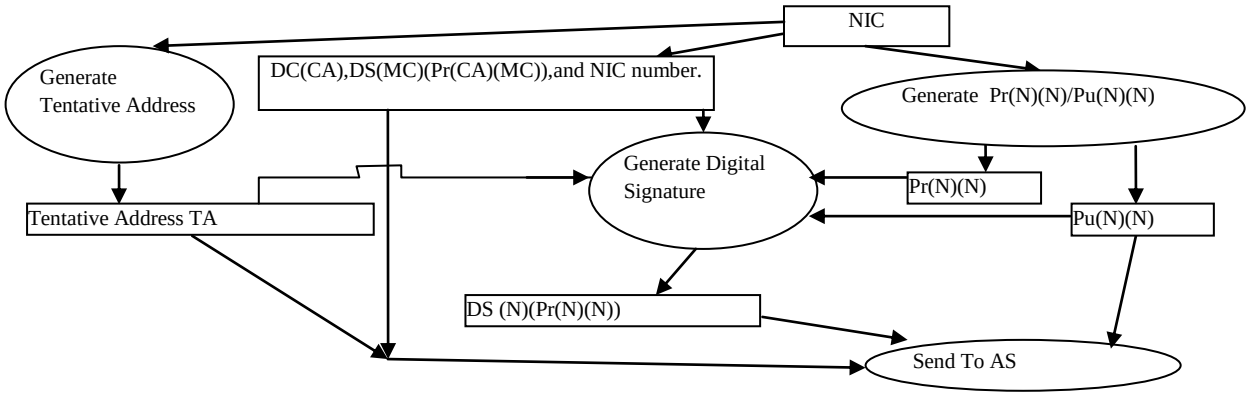


Figure 5. Node Processes

8.2.4. Verification of Certificate:

The message containing: DC(CA),DS(MC)(Pr(CA)(MC)) , Pu(N)(N), NIC number and DS (N)(Pr(N)(N)) are sent to AS. AS, then verifies Digital certificate DC(CA) by verifying public key Pu(CA)(MC) present in digital certificate with its database or from CA. However, it is not possible to verify from database, when AS does not have an entry into its database, of this particular company. Then AS sends request to the CA for verification of public key Pu(CA)(MC), present in Digital Certificate DC(CA). Once, CA verifies Pu(CA)(MC), the Company details and corresponding Pu(CA)(MC) are stored in database of AS, so that, in future, company can be verified locally. This process is shown in figure 6.

8.2.5. Verification of NIC

This process is used to Verify NIC. After verification of Pu(CA)(MC), AS extract NIC Number from Digital Signature DS(MC)(Pr(CA)(MC)),using Pu(CA)(MC), and compares it with NIC Number present in message. The matching of NIC number, confirms that NIC number is not fake.

8.2.6. Authentication of node:

Along with Digital signature DS (N)(Pr(N)(N)) the actual parameters DC(CA),DS(MC)(Pr(CA)(MC)), Pu(N)(N) and NIC number are also sent to AS. AS after confirming that NIC is not fake, AS verifies NIC number and corresponding TA, and for this purpose , AS creates message digest MD**(**DC(CA), DS(MC)(Pr(CA)(MC)) , Pu(N)(N) and NIC number and TA from message**)**. AS then decrypts the DS (N)(Pr(N)(N)) using Pu(N)(N) to get message digest. AS then compares both message digests as shown in figure 6. The matching of digests proves the owner ship of TA on Public key Pu(N)(N) key which authenticates the sender and integration of packet.

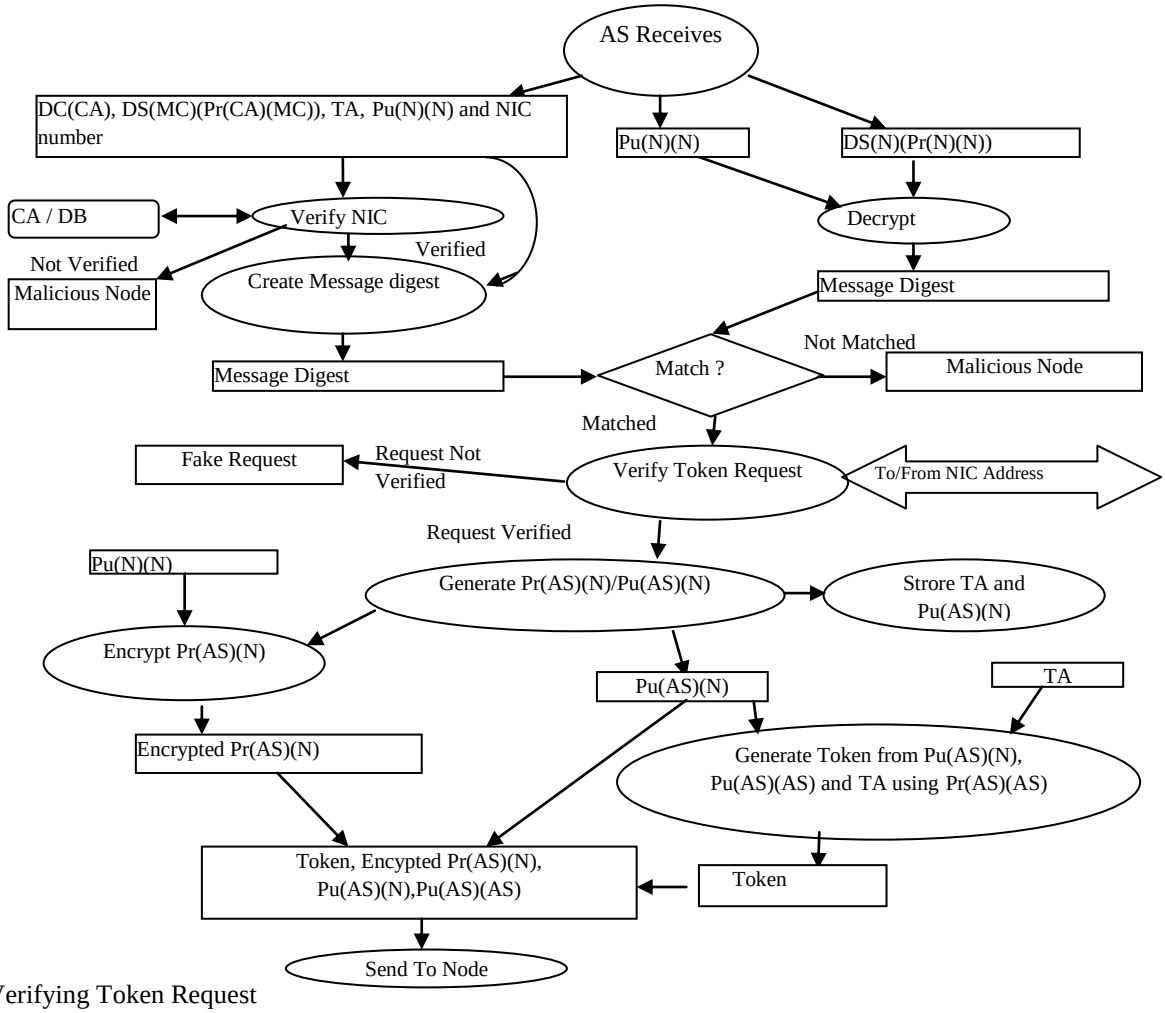


Figure 6. Working of AS

8.2.7. Verifying Token Request

It is essential to verify that ow t. AS then generates random number and encrypt it with Pr(AS)(AS). This encrypted number is again encrypted with public key of requester node Pu(N)(N) and sent to requester's NIC number address along with Pu(AS)(AS). After receiving this message from AS, requester node decrypts this message with its private key Pr(N)(N) and again decrypts it with Public key of AS Pu(AS)(AS) , to get actual random number sent by AS. This random number is now encrypted firstly with public key Pu(AS)(AS) of AS and then with private key of requester Pr(N)(N) and sent to AS. AS Decrypts encrypted number firstly with Public key Pu(N)(N) and then with Private key Pr(AS)(AS),to get the number sent by requester. The Matching of Number sent by AS and number received from requester validates that request is authentic. The message passing in this process is shown in figure 7.

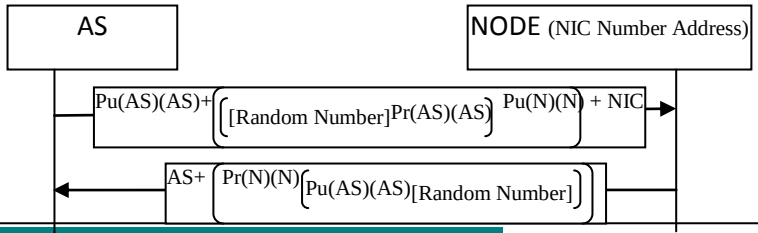


Figure 7. Authentication of Token

8.2.8. Registered Private and Public key for node

After the authentication of node and verification of token request, AS then generates Private/Public key pair Pr(AS)(N) and Pu(AS)(N) for node. The Pu(AS)(N), along with TA are stored into AS, as shown in figure 6. This information is stored to reply any request made by any node for verification of owner ship of Pu(AS)(N) of TA.

8.2.9. Issuance of Token

The Token is like a Digital signature created by AS of Public Key Pu(AS)(N), Pu(AS)(AS) and TA using Pr(AS)(AS).
The basic purpose of token is to allow node to verify TA and its owner ship on Pu(AS)(N). This is done by comparing message digest from decrypting token with message digest from TA, Pu(AS)(N) and Pu(AS)(AS). This verification of TA and corresponding certified Pu(AS)(N) restrict the node to go to AS for verification of sender every time. This reduces network load.

8.2.10. Message to claimant node from AS

The Token, Pu(AS)(N),Pu(AS)(AS), private Key Pr(AS)(N), encrypted with public key Pu(N)(N) are send to Node as shown in figure 8.

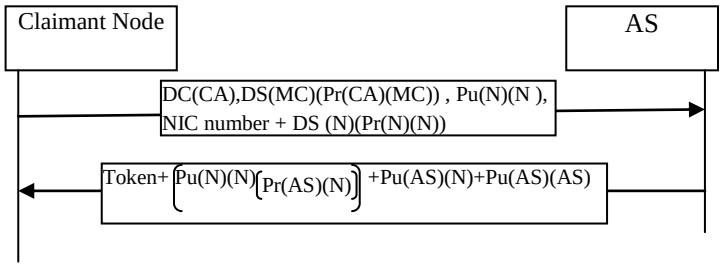


Figure 8. Message between Node and AS

8.2.11. DAD on Tentative address:

After receiving Token and other parameters from AS, AS then performs the DAD operation on tentative address. Nodes receiving DAD message performs the authentication of sender process using Token and other parameter. If any node replies DAD, it sends its token and other parameters to enquiring node. Node, then, performs authentication of sender, as explained above. If node receives message from authentic node, node again generates new TA. The node sends modification request with new TA, old TA and Token issues against old TA to AS. AS will verify node and modify its database accordingly. A new token is created to send to node again.

8.2.12. Setup of Symmetric Keys between nodes

Once the node gets the AS Digital Certificate as token, node can proceed to setup Symmetric keys between nodes for secure communication. The token issued by AS is passed to receiver node, by sender's node. The Token containing TA and corresponding AS's certified public key is used for authentication of sender. After confirmation, receiver sends its token, other parameters and secret key, encrypted with public key Pu(AS)(N) of sender as shown in figure 9. Sender validates receiver in same way as receiver has validated sender and keeps this key secure. Now both the nodes have agreed on secret key and can communicate encrypted data using this secret key.

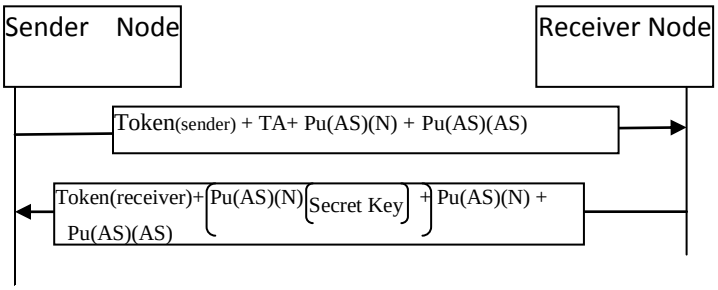


Figure 9. Message passing between Nodes

9. Conclusion

The Neighbour Discovery protocol was introduced to facilitate the node to configure itself. But if ND protocol is not protected it can open flood gate for threats. To protect from threats SEND was introduced which uses CGA address[4]. The missing part in Cryptographically Generated Address is that CGAs themselves are not certified, an attacker can create a new CGA from any subnet prefix and its own or anyone else's public key[5]. This paper presented a method wherein no new node is allowed to enter the network until and unless it proves to be a non-malicious node. Further, the scheme presented, in this paper, ensures that owner of NIC number and its corresponding IP Address has sent the message. This provides message authentication to receiver. The Public-key mechanism is used to exchange secret key. This secret key is used to encrypt the message, to provide confidentiality. The message digest of encrypted message is used to provide integrity of message. Further, the verification of TA and corresponding certified Pu(AS)(N), restrict the node to go to AS for verification of sender every time. This will also increase the efficiency of network.

References

1) http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html
2) Ahmad AlSa'deh and Christoph Meinel;"Security Neighbor Discovery ";IEEE Security & Privacy MagazineCopublished by the IEEE Computer and Reliability Societies 1540-7993/12 ,July/August 2012. pp. 26-34
3) ByungGoo Choi1, JaeHyun Ryu 2, ChoongSeon Hong3, DongJin Kwak; International Conference on Computational Sciences and Its Applications ICCSA 2008; "Enhanced SEND Protocol for Secure Data Transmission in Mobile IPv6 Environment "
4) T. Aura; Request for Comments: 3972; March 2005; Cryptographically Generated Addresses (CGA)
5) S. Thomson, T. Narten and T. Jinmei; Request for Comments: 4862; September 2007; "IPv6 Stateless Address Autoconfiguration "
6) Stefano M.Faccin and Franck Le ; "A Secure and Efficient solution to the IPv6 address ownership problem"; 0-7803-7605-6/02 ,2002 IEE Page : 162 -166.
7) Amirhossein Moravejosharieh, Hero Modares and Rosli Salleh; "Overview of Mobile IPv6 Security " ;2012 Third International Conference on Intelligent Systems Modelling and Simulation; 978-0-7695-4668-1/12,2012 IEEE.
8) [8] Joseph Davies; Published By: Prentice Hall of India - 2008; "Understanding IPv6"; Second Edition; ISBN: 978-81-203-3463-2.

Performance Analysis of Optimization Tool for Speech Recognition Using LPC & DSK TMS3206711/13 Using Simulink & Matlab

[1]Kadam V.K, [2]Dr.R.C.Thool

[1] Associate Professor, Department of Electronics, P.E.S College of Engineering, Nagsenvan, Aurangabad-431002.
[2]Professor & Head, Department of Information Technology, SGGS Institute of Engineering Technology, Nanded 431606.

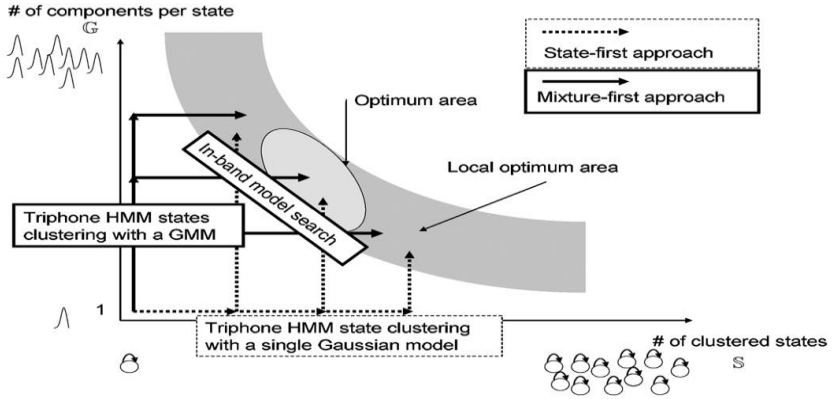
Abstract:
In this we have used various analysis tools like impulse response, frequency response, phase response etc. to get the performance analysis of said system. Here we implement a speech compression technique known as Linear Prediction Coding (LPC) using DSP System Toolbox from Matlab functionality available at the MATLAB command line. In this system we use the Levinson-Durbin and Time-Varying Lattice Filter blocks for low-bandwidth transmission of speech using linear predictive coding. We have changed the various order of the filter for various speeches signal & optimized that. We have received which is reflected in the figures given below.

Keywords: Introduction, System Implementation, Performance Analysis, Figures and Graphs, Conclusion, References, Lpc

1. Introduction

A number of noisy speech enhancement algorithms are experimentally compared in terms of linear predictive coding (LPC) perturbations. The enhancement algorithms considered are simple spectral subtraction, spectral over subtraction with use of a spectral floor, spectral subtraction with residual noise removal, and time-domain and frequency-domain adaptive minimum mean-square-error filtering. LPC perturbations considered are LPC cepstral distance, log likelihood r Linear Predictive Coding (LPC) has been used to compress and encode speech signals for digital transmission at a low bit rate. LPC determines a FIR system that predicts a speech sample from the past samples by minimizing the squared error between the actual occurrence and the estimated. The coefficients of the FIR system are encoded and sent. At the receiving end, the inverse system called AR model is excited by a random signal to reproduce the encoded speech. The use of LPC can be extended to speech recognition since the FIR coefficients are the condensed information of a speech signal of typically 10ms -30ms. PARCOR parameter associated with LPC that represents a vocal tract model based on a lattice filter structure is considered for speech recognition. The use of FIR coefficients and the frequency response of AR model were previously investigated. [1] In this we have taken the method to detect a limited number of phonemes from a continuous stream of speech. A system being developed slides a time window of 16 ms and calculates the PARCOR parameters continuously, feeding them to a classifier. A classifier is a supervised classifier that requires training. The classifier uses the Maximum Likelihood Decision Rule. The training uses TIMIT speech database, which contains the recordings of 20 speakers of 8 major dialects of American English. The classification results of some typical vowel and consonant phonemes segmented from the continuous speech are listed. The vowel and consonant correct classification rate are 65.22% and 93.51%. Overall, they indicate that the PARCOR parameters have the potential capability to characterize the phonemes.atios, and weighted likelihood ratio. [2] A communication system was built and tested to operate in the land mobile VHF band (150-174 MHz) at a channel separation of only 6 kHz. The audio source was digitally encoded at 2.4 kbits/s using linear predictive coding (LPC). The speech data stream was transmitted by frequency shift keying (FSK) which allowed the use of class-C transmitters and discriminator detection in the receiver. Baseband filtering of the NRZ data resulted in a narrow transmitter spectrum. The receiver had a 3 dB bandwidth of 2.4 kHz which allowed data transmission with minimal intersymbol interference and frequency offset degradation. A 58 percent eye opening was found. Bit error rate (BER) performance was measured with simulated Rayleigh fading at typical 150 MHz rates. Additional tests included capture, ignition noise susceptibility, adjacent channel protection, degradation from frequency offset, and bit error effects upon speech quality. A field test was conducted to compare the speech quality of the digital radio to that of a conventional 5 kHz deviation FM mobile radio. [3] In this, we try to use some part of propose a speech-model based method using the linear predictive (LP) residual of the speech signal and the maximum-likelihood (ML) estimator proposed in “Blind estimation of reverberation time,” (R. Ratnam, J. Acoust. Soc. Amer., 2004) to blindly estimate the reverberation time (RT60). The input speech is passed through a low order linear predictive coding (LPC) filter to obtain the LP residual signal. It is proven that the unbiased autocorrelation function of this LP residual has the required properties to be used as an input to the ML estimator. It is shown that this method can successfully estimate the reverberation time with less data than existing blind methods. Experiments show that the proposed method can produce

better estimates of RT60, even in highly reverberant rooms. This is because the entire input speech data is used in the estimation process. The proposed method is not sensitive to the type of input data (voiced, unvoiced), number of gaps, or window length. In addition, evaluation using white Gaussian noise and recorded babble noise shows that it can estimate RT60 in the presence of (moderate) background noise. [4]



"Figure 1: Optimum model search for an acoustic model"

2. System Implementation

2.1. Introduction

In this paper, a comparative study between two speech coders have been reported, considering their performances in simulation and in real-time. The speech coders taken for study are Linear Predictive Coder (LPC) and Cepstral coder. The simulation models are made on SIMULINK® and the real-time models can be implemented on TMS320C6713® DSK. For simulation, a comparison between synthesized speech signals using both the speech coders is given. For real-time implementation, some important parameters like memory consumption and execution time for these coders have been calculated. [5] In this system we implement LPC analysis and synthesis (LPC coding) of a speech signal. This process consists of two steps; analysis and synthesis. In the analysis section, we extract the reflection coefficients from the signal and use it to compute the residual signal. In the synthesis section, we reconstruct the signal using the residual signal and reflection coefficients. The residual signal and reflection coefficients require less number of bits to code than the original speech signal. The block diagram below shows the system we will implement. In this simulation, the speech signal is divided into frames of size 320 samples, with an overlap of 160 samples. Each frame is windowed using a Hamming window. Twelfth-order autocorrelation coefficients are found, and then the reflection coefficients are calculated from the autocorrelation coefficients using the Levinson-Durbin algorithm. The original speech signal is passed through an analysis filter, which is an all-zero filter with coefficients as the reflection coefficients obtained above. The output of the filter is the residual signal. This residual signal is passed through a synthesis filter which is the inverse of the analysis filter. The output of the synthesis filter is the original optimized signal. [6], [7], [8], [9] The Optimum Reflection Coefficients for the Lattice Forward and Backward Predictors in Section we derived the set of linear equations which provide the predictor coefficients that minimize the mean-square value of the prediction error. In this section we consider the problem of optimizing the reflection coefficients in the lattice predictor and expressing the reflection coefficients in terms of the forward and backward prediction errors. The forward prediction error in the lattice -filter is expressed as the reflection coefficient K, yields the result

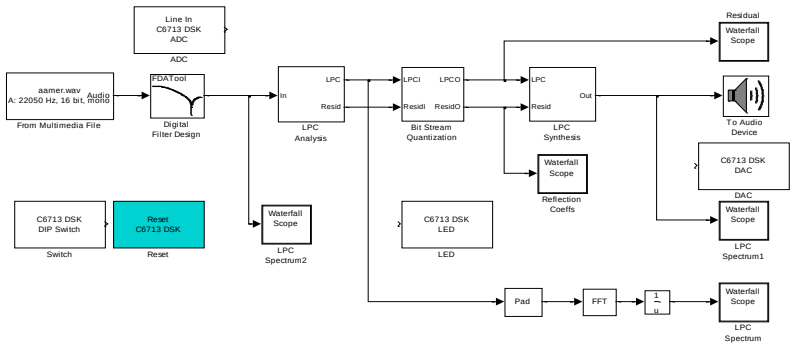
$$Km = \frac{-E[fm-1(n)g^{n}m-1(n-1)]}{E[lgm-1(n-1)l^{2}} \quad (1)$$

We observe that the optimum choice of the reflection coefficients in the lattice predictor is the negative of the (normalized) cross correlation coefficients between the forward and backward errors in the lattice.' Since it is apparent from (1 1.2.28) that K, ((1. it follows that the minimum mean-square value of the prediction error, which can be expressed recursively asis a monotonically decreasing sequence. [10] Here we initialize some of the variables like the frame size and also instantiate the System objects used in our processing. These objects also pre-compute any necessary variables or tables resulting in efficient processing calls later inside a loop. We create a buffer System object and set its properties such that we get an

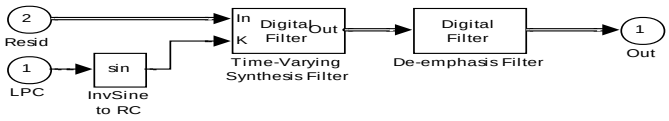
output of twice the length of the frame Size with an overlap length of frame Size. We also create a window System object. Here we will use the default window which is Hamming. By creating an autocorrelator System object and set its properties to compute the lags in the range [0:12] scaled by the length of input. We create a System object which computes the reflection coefficients from auto-correlation function using the Levinson-Durbin recursion. We configure it to output both polynomial coefficients and reflection coefficients. The polynomial coefficients are used to compute and plot the LPC spectrum. By creating an FIR digital filter System object used for analysis. Also create two all-pole digital filter System objects used for synthesis and de-emphasis.

2.2. Stream Processing Loop

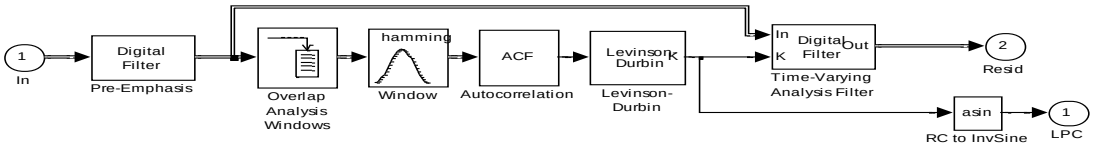
Here we call our processing loop where we do the LPC analysis and synthesis of the input audio signal using the System objects we have instantiated. The loop is stopped when we reach the end of the input file, which is detected by the AudioFileReader System object. Following fig shows the signal & LPC Spectrum.



"Figure 2: Block representation of system implementation using Simulink"



"Figure 3: LPC Analysis"

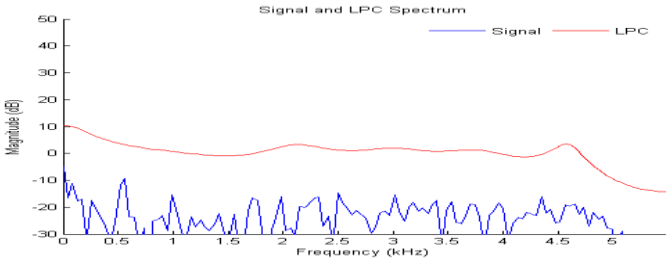


"Figure 4:LPC Synthesis"

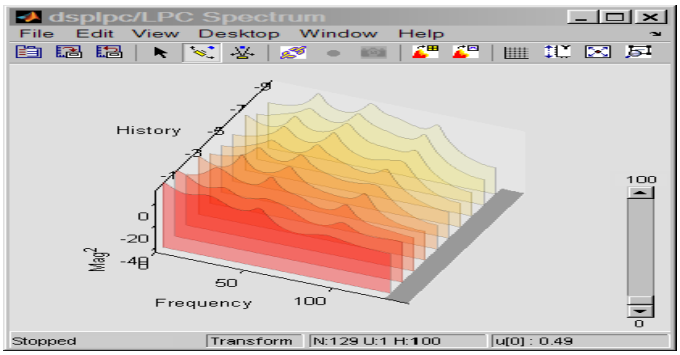
3. Performance Analysis

LPC determines the coefficients of a forward linear predictor by minimizing the prediction error in the least squares sense. It has applications in filter design and speech coding. [a,g] = lpc(x,p) finds the coefficients of a pth-order linear predictor (FIR filter) that predicts the current value of the real-valued time series x based on past samples.p is the order of the prediction filter polynomial, a = [1 a(2) ... a(p+1)]. If p is unspecified, lpc uses as a default p = length(x)-1. If x is a matrix containing a separate signal in each column, lpc returns a model estimate for each column in the rows of matrix a and a column vector of prediction error variances g. The length of p must be less than or equal to the length of x.Algorithms for lpc uses the autocorrelation method of autoregressive (AR) modeling to find the filter coefficients. The generated filter might not model the process exactly even if the data sequence is truly an AR process of the correct order. This is because the autocorrelation method implicitly windows the data, that is, it assumes that signal samples beyond the length of x are 0. [11]

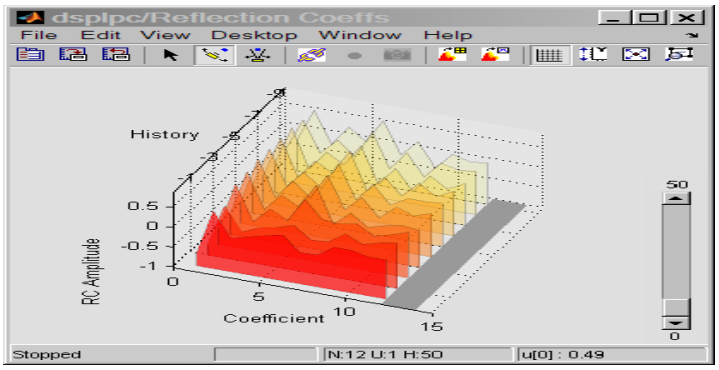
4. Figures and Graphs



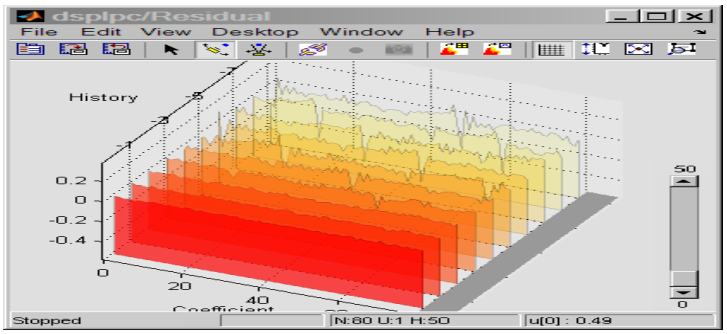
"Figure 5: Graph shows the the signal & LPC Signal"



"Figure 6: LPC Spectrum of signal"

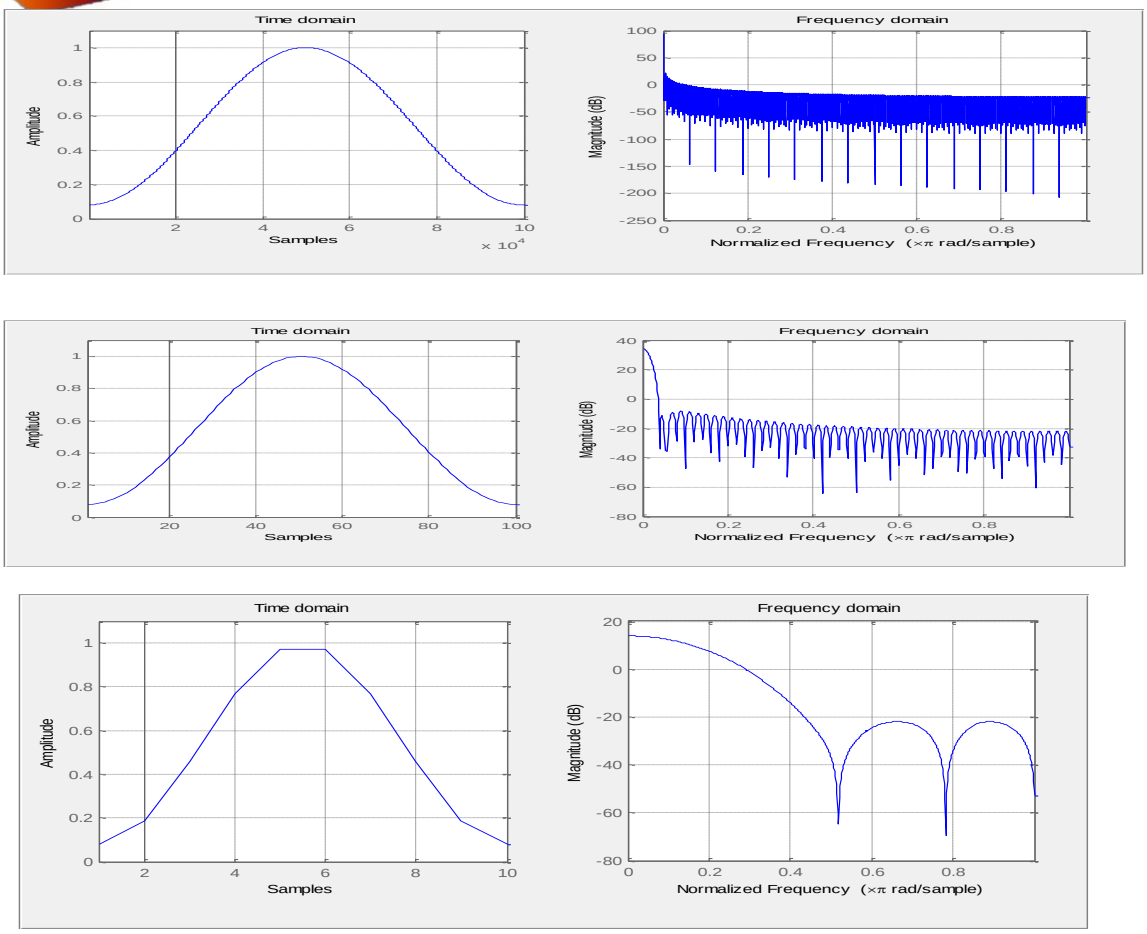


"Figure 7: Reflection coefficients of signal"



"Figure 8: Residual of Signal"

Following figures shows the analysis of the system for various orders of filters & windows



"Figure 9: Time & Frequency representation"

5. Conclusion

This analysis gives the performance of the system implemented using the DSK TMS3206711/13 for various values of the order of filters & various windows. We have seen here the implementation of speech compression technique using Linear Prediction Coding. The implementation used the DSP System Toolbox functionality available at the MATLAB command line. The code involves only calling of the successive System objects with appropriate input arguments. This involves no error prone manual state tracking which may be the case for instance for a MATLAB implementation of Buffer. From the performance it is observed that the optimized speech recognition can be archive

References

[1] Ahmed, M.S. Dept. of Syst. Eng., King Fahd Univ. of Pet. & Miner., Dhahran ,Comparison of noisy speech nhancement
algorithms in terms of LPC perturbation, Acoustics, Speech and Signal Processing, IEEE Transactions on Date of Publication: Jan 1989,Volume: 37 , Issue: 1 Page(s): 121 - 125

[2] Ying Cui; Takaya ,Recognition of Phonemes In a Continuous Speech Stream By Means of PARCOR Parameter In LPCVocoder, K.Electrical and Computer Engineering, 2007. CCECE 2007. Canadian Conference on Digital Object Identifier:10.1109/CCECE.2007.402 Publication Year: 2007, Page(s): 1606 – 1609

[3] Speech 'McLaughlin, M.; Linder, D. Carney. S, Design and Test of a Spectrally Efficient Land Mobile Communications

System Using LPC, Selected Areas in Communications, and IEEE Journal on Volume: 2, Issue: 4 Digital Object Identifier: 10.1109/JSAC.1984.1146086 Publication Year: 1984, Page(s): 611 – 620

[4] Keshavarz, A.; Mosayyebpour, S.; Biguesh, M.; Gulliver, T.A.; Esmaeili M,Speech-Model Based Accurate Blind Reverberation Time Estimation Using an LPC Filter, Audio, Speech, and Language Processing, IEEE Transactions on Volume: 20 , Issue: 6 Digital Object Identifier: 10.1109/TASL.2012.2191283 Publication Year: 2012 , Page(s): 1884 – 1893

[5] Bhattacharya, S.; Singh, S.K.; Abhinav, T, Performance evaluation of LPC and cepstral speechcoder in simulation and in real-time Recent Advances in Information Technology (RAIT), 2012 1st International Conference on Digital Object Identifier:10.1109/RAIT.2012.6194531 Publication Year: 2012, Page(s): 826 - 831

[6] Fliege, N.J., Mulitrate Digital Signal Processing (John Wiley and Sons, 1994).

[7] Mitra, S.K., Digital Signal Processing (McGraw-Hill, 1998).

[8] Orfanidis, S.J., Introduction to Signal Processing (Prentice-Hall, Inc., 1996).

[9] Vaidyanathan, P.P., Multirate Systems and Filter Banks (Prentice-Hall, Inc., 1993).

[10] Proakis, Digital Signal Processing (third edition pp. 863-64).

[11] Jackson, L.B., Digital *Filters and Signal Processing* (Second Edition, Kluwer Academic Publishers, 1989. pp.255-257).

Kadam V.K[1]
Associate Professor & Research Student,
Department of Electronics,
P.E.S College of Engineering
Nagsenvan, Aurangabad-431002 (M.S)
Dr.Babasaheb Ambedkar Marathwada University,
Aurangabad-431002 (MS)

Dr.R.C Thool[2]
Department of Information Technology
SGGS Institute of Engineering & Technology
Vishnupuri, Nanded 431606 (M.S)
(An autonomous institute set up and 100% funded by Government of Maharashtra)

Convexity of Minimal Total Dominating Functions Of Quadratic Residue Cayley Graphs

[1]S.Jeelani Begum [2] B.Maheswari
[1]Dept. of Mathematics, Madanapalle Institute of Technology & Science, Madanapalle, Andhra Pradesh, India
[2]Dept. of Applied Mathematics, S.P.Women's University, Tirupati, Andhra Pradesh, India

Abstract

Nathanson [17] paved the way for the emergence of a new class of graphs, namely,Arithmetic Graphs by introducing the concepts of Number Theory, particularly, the Theory of congruences in Graph Theory. Cayley graphs are another class of graphs associated with the elements of a group. If this group is associated with some arithmetic function then the Cayley graph becomes an arithmetic graph. Domination theory is an important branch of Graph Theory and has many applications in Engineering, Communication Networks and many others. In this paper we study the minimal total dominating functions of Quadratic Residue Cayley graphs and discuss the convexity of these functions in different cases.

Keywords: Arithmetic graph, Cayley graph, Total dominating set, Neighbourhood set, Quadratic Residue Cayley Graph, Total Dominating Functions, Minimal Total Dominating Functions, Convexity.

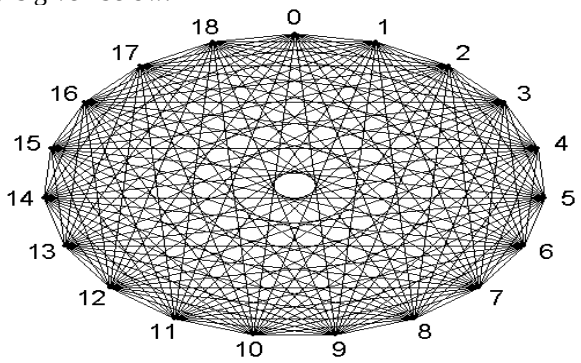
1. Introduction

There is a class of graphs, namely, Cayley graphs, whose vertex set V is the set of elements of a group (G, .) and two vertices x and y of G are adjacent if and only if xy^{-1} is in some symmetric subset S of G. A subset S of a group (G, .) is called a symmetric subset of G if s^{-1} is in S for all s in S. If the group (G, .) is the additive group $(Z_n, \oplus)$ of integers 0,1,2,……..,n-1 modulo n, and the symmetric set S is associated with some arithmetic function, then the Cayley Graph may be treated as an arithmetic graph. In this paper we consider Quadratic Residue Cayley graphs. A detailed study of convexity and minimality of dominating functions and total dominating functions are given in Cockayne et al. [2,3-12] Chesten et al. [1], Yu [18] and Domke et al. [13,14]. Algorithmic complexity results for these parameters are given in Laskar et al. [15] and Cockayne et al.[3].We start with the definition of a Quadratic Residue Cayley graph.

Quadratic Residue Cayley Graph

Let p be an odd prime and n, a positive integer such that $n \not\equiv 0 \pmod p$. If the quadratic congruence, $x^2 \equiv n \pmod p$ has a solution then, n is called a quadratic residue mod p.

The Quadratic Residue Cayley graph $G(Z_p, Q)$, is the Cayley graph associated with the set of quadratic residues modulo an odd prime p, which is defined as follows. Let p be an odd prime, S, the set of quadratic residues modulo p and let S* = {s, n - s / $s \in S, s \neq n$ }. The quadratic residue Cayley graph $G(Z_p, Q)$ is defined as the graph whose vertex set is Z_p = { 0 , 1 , 2 , …… p – 1} and the edge set is E = { (x , y) / x – y or y - x $\in$ S*}.
For example the graph of $G(Z_{19}, Q)$ is given below.



2. Total Dominating Functions

Total Dominating Set : Let G(V, E) be a graph without isolated vertices. A subset T of V is called a total dominating set (TDS) if every vertex in V is adjacent to at least one vertex in T.

Minimal Total Dominating Set : If no proper subset of T is a total dominating set, then T is called a minimal total dominating set (MTDS) of G.

Neighbourhood Set : The open neighbourhood of a vertex u is the set of vertices adjacent to u and is denoted by N(u).

Total Dominating Function : Let G(V, E) be a graph without isolated vertices. A function $f:V\to[0,1]$ is called a total dominating function (TDF) if $f(N(v))=\sum_{u\in N(v)} f(u)\geq 1$ for all v ∈ V.

Minimal Total Dominating Function : Let *f* and *g* be functions from $V\to$ [0,1]. We define f < g if f(u) ≤ g(u), ∀ u ∈ V, with strict inequality for at least one vertex u. A TDF f is called a minimal total dominating function (MTDF) if for all g < f, g is not a TDF.

We require the following results whose proofs are presented in [16].

Lemma 1: The Quadratic Residue Cayley graph $G(Z_p, Q)$ is $|S^*|$- regular, and the number of edges in $G(Z_p, Q)$ is $\frac{|Z_p||S^*|}{2}$.

Theorem 1: The Quadratic Residue Cayley graph $G(Z_p, Q)$ is complete if p is of the form 4m+3.

Suppose p = 4m + 3. Then $G(Z_p, Q)$ is complete. Then each vertex is of degree p – 1. That is the graph $G(Z_p, Q)$ is (p – 1) – regular.

∴ $|S^*| = p - 1$.

Hence each N(v) consists of p-1 vertices , ∀v∈V.

We consider the case p = 4m+3 of $G(Z_p, Q)$ and prove the following results.

3. MAIN RESULTS

Theorem 3.1: Let T be a MTDS of $G(Z_p, Q)$. Let $f:V\to[0,1]$ be a function defined by

$$f(v)=\begin{cases}1, & if\ v\in T,\\ 0, & if\ v\in V-T.\end{cases}$$

Then f becomes a MTDF of $G(Z_p, Q)$.

Proof: Consider $G(Z_p, Q)$. Let T be a MTDS of $G(Z_p, Q)$.

Since $G(Z_p, Q)$ is complete, $|T|=2$.

Also every neighbourhood N(v) of v ∈ V consists of (p-1) –vertices.

Let f be a function defined on V as in the hypothesis.

Then the summation values taken over the neighbourhood N(v) of v ∈ V is

$$\sum_{u\in N(v)} f(u)=\begin{cases}2, & if\ u\in V-T,\\ 1, & if\ u\in T.\end{cases}$$

Therefore

$$\sum_{u\in N(v)} f(u)\geq 1, \forall\ v\in V.$$

This implies that f is a TDF.

We now check for the minimality of f.

Define a function $g:V\to[0,1]$ by

$$g(v)=\begin{cases} r, & if\ v\in T,\ v=v_k,\\ 1, & if\ v\in T-\{v_k\},\\ 0, & if\ v\in V-T. \end{cases}$$

where $0 < r < 1$ and $v_k \in V$.
Since strict inequality holds at the vertex $v = v_k \in T$ of V, it follows that $g < f$.
Then

$$\sum_{u\in N(v)} g(u)=\begin{cases} 1+r, & if\ v\in V-T,\\ 1, & if\ v\in T,\ v=v_k,\\ r, & if\ v\in T,\ v\neq v_k. \end{cases}$$

Thus $\sum_{u\in N(v)} g(u) \not\geq 1, \ \forall\ v\in V.$
This implies that g is not a TDF. Since $r < 1$ is arbitrary it follows that there exists no $g < f$ such that g is a TDF.
Thus f is a MTDF.

Theorem 3.2: Let T_1 and T_2 be two MTDSs of $G(Z_p, Q)$ and $f_1,\ f_2$ be two functions of $G(Z_p, Q)$ defined by

$$f_1(v)=\begin{cases} 1, & if\ v\in T_1,\\ 0, & otherwise. \end{cases}$$

and $$f_2(v)=\begin{cases} 1, & if\ v\in T_2,\\ 0, & otherwise. \end{cases}$$

Then the convex combination of f_1 and f_2 becomes a TDF of $G(Z_p, Q)$ but not minimal.

Proof: Let T_1 and T_2 be two MTDSs of $G(Z_p, Q)$ and f_1, f_2 be the functions defined as in the hypothesis. Then by Theorem 3.1, the above functions are MTDFs of $G(Z_p, Q)$.
Let $h(v)=\alpha f_1(v)+\beta f_2(v)$, where $\alpha+\beta=1$ *and* $0<\alpha<1,\ 0<\beta<1.$

Case 1: Suppose T_1 and T_2 are such that $T_1 \cap T_2 \neq \phi$.
Then the possible values of h(v) are

$$h(v)=\begin{cases} \alpha, & if\ v\in T_1\ and\ v\notin T_2,\\ \beta, & if\ v\in T_2\ and\ v\notin T_1,\\ \alpha+\beta, & if\ v\in \{T_1\cap T_2\},\\ 0, & otherwise. \end{cases}$$

Since each neighbourhood N(v) of v in V consists of (p-1) vertices of $G(Z_p, Q)$, the summation value of h(v) taken over N(v) is

$$\sum_{u\in N(v)} h(u)=\begin{cases} \alpha+\beta+\beta, & if\ v\in T_1\ and\ v\notin T_2,\\ \alpha+\beta+\alpha, & if\ v\in T_2\ and\ v\notin T_1,\\ \alpha+\beta, & if\ v\in \{T_1\cap T_2\},\\ 2(\alpha+\beta), & otherwise. \end{cases}$$

This implies that $\sum_{u\in N(v)} h(u)\geq 1, \forall$ v ∈ V.
Therefore h is a TDF. We now check for the minimality of h.
Define a function $g: V\to[0,1]$ by

$$g(v)=\begin{cases}\alpha, & \text{if } v\in T_1 \text{ and } v\notin T_2,\\ \beta, & \text{if } v\in T_2 \text{ and } v\notin T_1,\\ r, & \text{if } v\in \{T_1\cap T_2\},\\ 0, & \text{otherwise.}\end{cases}$$

where 0 < r < 1.

Since strict inequality holds at $v\in \{T_1\cap T_2\}$, it follows that g < h.

Now $\sum_{u\in N(v)} g(u)=\begin{cases}r+\beta, & \text{if } v\in T_1 \text{ and } v\notin T_2,\\ \alpha+r, & \text{if } v\in T_2 \text{ and } v\notin T_1,\\ \alpha+\beta, & \text{if } v\in \{T_1\cap T_2\},\\ r+\alpha+\beta, & \text{otherwise.}\end{cases}$

where $r+\beta<1+\beta$ and $\alpha+r<\alpha+1$.

Thus $\sum_{u\in N(v)} g(u)\geq 1, \quad \forall v\in V$.

Therefore g is a TDF. Hence it follows that h is a TDF but not minimal.

Case 2: Suppose T_1 and T_2 are disjoint.

Then the possible values of h(v) are

$$h(v)=\begin{cases}\alpha, & \text{if } v\in T_1,\\ \beta, & \text{if } v\in T_2,\\ 0, & \text{otherwise.}\end{cases}$$

Since each neighbourhood N(v) of v in V consists of (p-1) vertices of $G(Z_p, Q)$, the summation value of h(v) taken over N(v) is

$$\sum_{u\in N(v)} h(u)=\begin{cases}\alpha+2\beta, & \text{if } v\in T_1,\\ \beta+2\alpha, & \text{if } v\in T_2,\\ 2(\alpha+\beta), & \text{otherwise.}\end{cases}$$

This implies $\sum_{u\in N(v)} h(u)\geq 1$, $\forall$ v ∈ V, since $\alpha+\beta=1$.

Therefore h is a TDF. We now check for the minimality of h.

Define a function $g:V\rightarrow[0,1]$ by

$$g(v)=\begin{cases}r, & \text{if } v\in T_1, v=v_i,\\ \alpha, & \text{if } v\in T_1, v\neq v_i,\\ \beta, & \text{if } v\in T_2,\\ 0, & \text{otherwise.}\end{cases}$$

where 0 < r < α.

Since strict inequality holds at $v=v_i\in T_1$, it follows that g < h.

Then $\sum_{u\in N(v)} g(u) = \begin{cases} \alpha+2\beta, & if\ v\in T_1, v=v_i, \\ r+2\beta, & if\ v\in T_1, v\neq v_i\ , \\ r+\alpha+\beta, & if\ v\in T_2, \\ r+\alpha+2\beta, & otherwise. \end{cases}$

where $r+2\beta \le \alpha+2(1-\alpha) = 2-\alpha > 1$.

i.e., $r+2\beta > 1$.

Thus $\sum_{u\in N(v)} g(u) > 1, \quad \forall v\in V$.

Therefore g is a TDF. Hence it follows that h is not a MTDF.

References

[1]. G.A. Chesten , G. Fricke , S.T. Hedetniemi and D.P. Jacobs. On the Computational complexity of upper fractional domination, Discrete. Appl. Math., 27: 195-207, 1990.

[2]. E.J. Cockayne, R.M.. Dawes and S.T. Hedetniemi. Total domination in graphs, Networks, 10: 211- 219, 1980.

[3]. E. J. Cockayne , G., Macgillivray and C.M. Mynhardt . A linear algorithm for universal minimal dominating functions in trees, J. Combin. Math. Comput., 10: 23-31, 1991.

[4]. E.J. Cockayne , G. Macgillivray, and C.M. Mynhardt . Convexity of minimal dominating functions and universals in graphs, Bull. Inst. Combin. Appl., 5: 37- 48, 1992.

[5]. E.J. Cockayne and C.M. Mynhardt . Convexity of minimal dominating functions of trees: A survey, Quaestiones Math., 16:301 – 317, 1993.

[6]. E.J. Cockayne, C.M. Mynhardt and B. Yu. Universal minimal total dominating functions in graphs, Networks, 24: 83- 90, 1994.

[7]. E.J. Cockayne, G. Macgillivray and C.M. Mynhardt. Convexity of minimal dominating functions of trees-II, Discrete Math., 125: 137-146, 1994.

[8]. E.J Cockayne and C.M. Mynhardt. A characterization of universal minimal total dominating functions in trees, Discrete Math., 141 :75-84, 1995.

[9]. E.J. Cockayne, G. Macgillivray and C.M. Mynhardt. Convexity of minimal dominating functions of trees, Utilitas Mathematica, 48: 129-144, 1995.

[10]. E.J. Cockayne, G. Fricke, S.T. Hedetniemi and C.M. Mynhardt. Properties of minimal dominating functions of graphs, ARS. Combinatoria, 41: 107-115, 1995.

[11]. .E.J. Cockayne, C.M. Mynhardt and B. Yu.Total dominating functions in trees: Minimality and convexity, Journal of Graph Theory, 19: 83 – 92,1995.

[12]. E.J. Cockayne and C.M. Mynhardt. Minimality and convexity of dominating and related functions in graphs: A unifying theory, Utilitas Mathematica, 51:145-163, 1997.

[13]. G.S. Domke, S.T. Hedetniemi and R.C. Laskar. Fractional packing covering and irredundance in graphs, Congr. numer., 66: 227-238, 1988.

[14]. G. Fricke, E.O. Hare, D.P. Jacobs and A. Majumdar. On integral and fractional total domination, Congr.Numer., 77: 87-95, 1990.

[15]. R. Laskar, J. Pfaff , S.M. Hedetniemi and S.T. Hedetniemi. On the algorithmic complexity of total domination, SIAM. J. Alg. Disc. Math., 5: 420-425, 1984.

[16]. L. Madhavi. Studies on domination parameters and enumeration of cycles in some Arithmetic Graphs, Ph.D. thesis, S.V.University, 2002.

[17]. Nathanson and B. Melvyn. Connected components of arithmetic graphs, Monat. fur. Math, 29,1980.

[18]. B.YU. Convexity of minimal total dominating functions in graphs, Journal of Graph Theory, 24 : 313-321, 1997.

Classification of Lung Tumor Using SVM

[1]Ms. Swati P. Tidke, [2]Prof. Vrishali A. [3]Chakkarwar
Department of computer science and engineering, Government College of engineering, Aurangabad, Maharashtra

Abstract—Lung cancer is the most important cause of cancer death for both men and women. Early detection is very important to enhance a patient's chance for survival of lung cancer. This paper provides a Computer Aided Diagnosis System (CAD) for early detection of lung cancer nodules from the Chest Computer Tomography (CT) images. There are five main phases involved in the proposed CAD system. They are image pre-processing, extraction of lung region from chest computer tomography images, segmentation of lung region, feature extraction from the segmented region, classification of lung cancer as benign or malignant. Initially total variation based denoising is used for image denoising, and then segmentation is performed using optimal thresholding and morphological operations. Textural features extracted from the lung nodules using gray level co-occurrence matrix (GLCM). For classification, SVM classifier is used. The main aim of the method is to develop a CAD (Computer Aided Diagnosis) system for finding the lung tumor using the lung CT images and classify the tumor as Benign or Malignant.

Keywords—Computer Aided Diagnosis System, optimal thresholding, gray level co-occurrence matrix (GLCM), Support vector machine (SVM).

I. Introduction

Lung cancer is the leading cause of tumor-related deaths in the world [1]. At the same time, it appears that the rate has been steadily increasing. Lung cancer is caused by the uncontrolled growth of tissues in the lung. The American cancer society estimates that 213, 380 new cases of lung cancer in the U.S will be diagnosed and 160, 390 deaths due to lung cancer will occur in 2007.The majority of all cases are caused by tobacco smoking. Exposure to asbestos, radon, uranium and arsenic are other factors for lung cancer. Lung cancer is a deadly disease and has chances to spread to other parts of the body, e.g. the brain, liver, bone and bone marrow. The early detection and diagnosis of nodules in CT image are among the most challenging clinical tasks performed by radiologists. Radiologists can miss up to 25% of lung nodules in chest radiographs due to the background anatomy of the lungs which can hide the nodules. Computer aided diagnosis system helps the radiologists by doing preprocessing of the images and recommending the most possible regions for nodules. The complexity for finding the lung nodules in radiographs are as follows:

1. A nodule diameter may be differed from a few millimeters
2. Nodules vary widely in density.
3. As nodules can be found anywhere in the lung region, they can be hidden by ribs and structures below the diaphragm, resulting in a large variation of contrast to the background.
4. To overcome these difficulties, the author proposed a Computer Aided Diagnosis (CAD) [2] system for detection of lung nodules [3]. The lung tumor prediction system is shown in Figure 1:

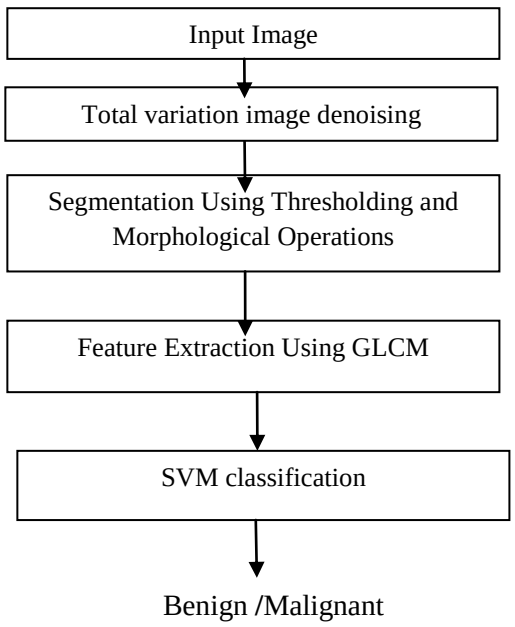
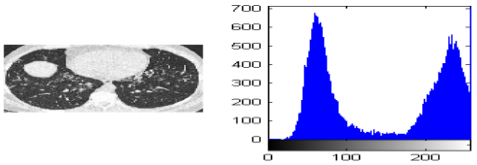


Fig. 1 lung tumor prediction system

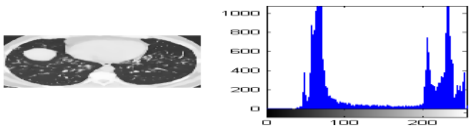
This paper initially apply the different image processing techniques such as image denoising for removing noise in the image, optimal thresholding for converting gray level image into binary image, morphological operations for segmentation, feature extraction using GLCM and Support vector machine is used for classification.

II. IMAGE PREPROCESSING

CT imaging process may contain various noises. These images are not clean so as to use directlyforprocessing,thuswe need to denoise these images. The segmenting scheme introduced in this paper performs an image preprocessing task to remove noise in a lung CT image at first. Total variation denoising is a preprocessing step to reduce the effect of undesired perturbations. Total variation denoising is very effective at simultaneously preserving edges while smoothing away noise in flat regions, even at low signal-to-noise ratios. total variation denoising results are shown in figure 2.



(a) Noisy image (b) Histogram noisy image



(c) Denoised image (d) Histogram denoised image

Fig. 2 Histogram of lung CT before and after denoising

III. Segmentation Of Lung Region

In chest CT images, there are two types of pixels with different density distribution:

1) Pixels within the very dense body and chest wall structures (the body pixels) and 2) low-density pixels. Optimal thresholding proposed by Shiying et al. [3], is applied on the pre-processed lung image to select a segmentation threshold to separate the body and non-body pixels through an iterative procedure. Let Ti be the segmentation threshold after the ith step. To choose a new segmentation threshold, we apply Ti to the image to separate the pixels into body and nobody pixels. Let μ_b and μ_n be the mean grey level of body and non-body pixels. Then new threshold will be:

$$T^{i+1} = \frac{1}{2}(\mu_b + \mu_n)$$

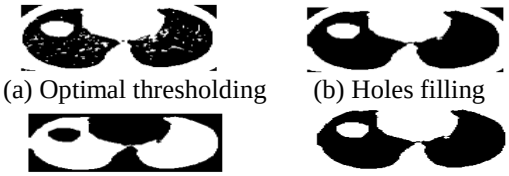
The pixels with a density lower than the threshold value are recognized assigned a value 1 and appear white, whereas other pixels are assigned the value of 0 and appear black. The lung image after segmentation with optimal threshold value contains non-body pixels such as the air surrounding the lungs, body and other low-density regions within the image and is removed through morphological operations such as erosion, dilation and labeling.

IV. Roi Extraction

Lung nodules are the small masses of tissue in the lung. They appear approximately round and white in CT scan or X-ray images. In the proposed method our region of interest is lung nodule and labeling algorithm is applied for region extraction. Connected component labeling is method of addressing different group of pixels based on their characteristics, in this case intensity values of pixels and their neighborhood. There are number of locations where pixels have same intensity values they are gathered as one group and uniquely labelled.Labeling is usually used to detect connected regions in binary images. Color images and data with higher-dimensionality can also be processed. In proposed method 4-connected labeling algorithm is used. Overview of 4-connected labeling algorithm is as follows:

- Given a binary image.
- Negate the image.
- For every pixel check the north and west pixel.
- If the pixel to the west or north has the same intensity value, the pixel belongs to same region. Assign the same label to the current pixel.
- If the pixel to the west has a different value and the pixel to the north has the same value, assign the north pixel's label to current pixel.
- If the north and west neighbors of pixel have different pixel values, create a new label and assign that label to the current pixel.
- Do this recursively for all pixels that have a 4-connectivity.

All pixels above the threshold that have a 4-connectivity will get the same number and thereby each region a unique label. Find the centroid of each label, if centroid of label is not present at significant height and width considering our region of interest eliminates that label. In this way we will get desired lung region. The extracted ROIs are then subject to feature extraction for analysis.



(a) Optimal thresholding (b) Holes filling

(c) Labeling (d) Lung region

Fig. 3 Optimal thresholding and morphological operations

V. Feature Extraction

The extracted ROIs can be distinguished as either cancerous or not using their texture properties. Gray Level Co-occurrence Matrix (GLCM) is one of the most popular ways to describe the texture of an image. A GLCM denote the second order conditional joint probability densities of each of the pixels, which is the probability of occurrence of grey

level i and grey level j within a given distance 'd' and along the direction 'θ'.7 features are considered for proposed method.

1. Area: It gives the actual number of pixels in the ROI.

2. Convex *Area*: It gives the number of pixels in convex image of the ROI.

3. Equivalent Diameter: It is the diameter of a circle with the same area as the ROI.

$$\text{Equivalent Diameter} = \frac{\sqrt{4 \cdot \text{Area}}}{\sqrt{\pi}} \qquad (3)$$

4. Solidity: It is the proportion of the pixels in the convex hull that are also in the ROI.

$$\text{Solidity} = \frac{\text{Area}}{\text{ConvexArea}} \qquad (4)$$

5. Energy: It is the summation of squared elements in the GLCM and its value ranges between 0 and 1.

$$\text{Energy} = \sum_{k=0}^{n} p^2(i,j) \qquad (5)$$

6. Contrast: It is the measure of contrast between an intensity of pixel and its neighboring pixels over the whole ROI.where N is the number of different gray levels.

$$\text{Contrast} = \sum_i^N \sum_j^N (i-j)^2 \, p(i,j) \qquad (6)$$

7. Homogeneity: It is the measure of closeness of the distribution of elements in the GLCM to the GLCM of each ROI and its Value ranges between 0 and 1.

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \qquad (7)$$

8. Correlation: It is the measure correlation of pixel to its neighbor over the ROI.

$$\text{Correlation} = \sum_i^N \sum_j^N \frac{p(i,j) - \mu r \, \mu c}{\sigma r \, \sigma c} \qquad (8)$$

9. Eccentricity: The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length.

II. Svm Classification

Support vector machines are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification The basic SVM takes a set of input data and for each given input, predicts, which of two classes forms the input, making it a non-probabilistic binary linear classifier. From given set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. In the proposed method we are using linear classifier. Best hyper plane is the one that represents the largest separation or margin between the two classes. So we choose the hyperplane so that the distance from it to the nearest data point on each side is maximized. If such a hyper plane exists, it is known as the maximum margin hyperplane and the linear classifier it defines is known as a maximum classifier, which is shown in fig.4

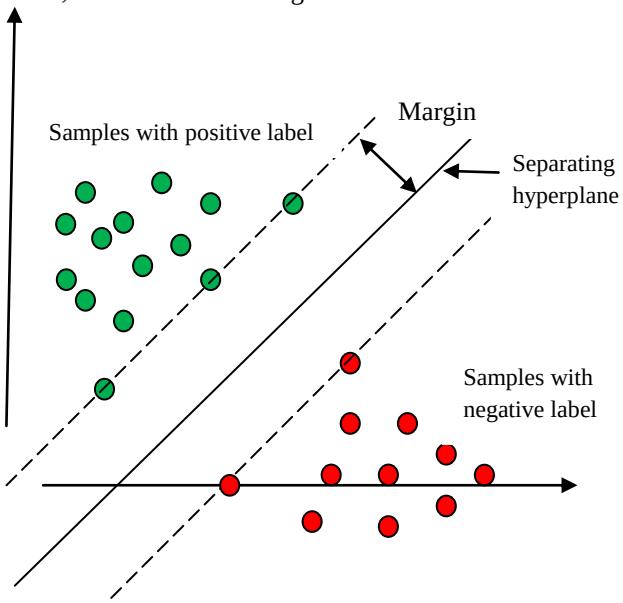
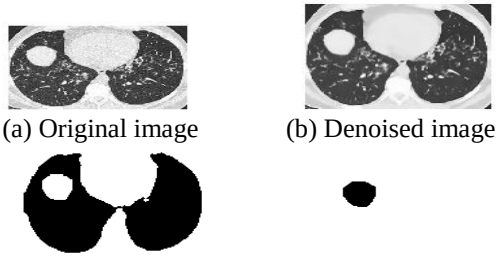


Fig. 4 Maximum margin classifier

III. Experimental Results And Analysis

The data set used for study of the proposed system consists of 25 diseased lung CT JPEG images of size196x257. A total of 40 ROIs were extracted. The system was tested with 15 diseased lung images. The results obtained for a diseased lung image is shown in Fig. 5



(a) Original image (b) Denoised image

(c) segmented image (d) tumor region

Fig. 5 Results obtained for a diseased lung image In case of SVM classifier out of 9 features at a time only twofeatures are selected for classification, which produces result as either benign or malignant. In the proposed method first two features provides the best results. A few classification results using Support vector machine is listed in table I

TABLE I
TUMOR CLASSIFICATION USING SVM

Image	Output Value	Classification
Image_1	0	Benign
Image_2	1	Malignant
Image_3	0	Benign
Image_4	1	Malignant
Image_5	1	Malignant
Image_6	1	Malignant
Image_7	1	Malignant
Image_8	1	Malignant
Image_9	0	Benign
Image_10	1	Malignant

Conclusion

Proposed system helps physician to extract he tumor region and evaluate whether the tumor is benign or malignant. The computer tomography image is used in this paper. SVM provides the accuracy of 92.5%.The accuracy of system can be improved if training is performed by using a very large image database. The different basic image processing techniques are used for prediction purpose. In the first phase the image is denoised. In second phase lung region is separated from surrounding anatomy. In next phase ROI is extracted, after ROI extraction features extraction is performed by GLCM. Finally with the obtained texture features classification is performed to detect the occurrence of cancer nodules.

REFERENCES

[1] Weir HK et. al.,"Annual report to the nation on the status of cancer, 1975-2007"*Journal National Cancer Institute*, vol. 95 , No. 17.

[2] Ayman El-Baz, Aly A. Farag, Robert Falk, Renato La Rocca, "Detection, Visualization and identification of Lung Abnormalities in Chest Spiral CT Scan: Phase-I", *International Conference on Biomedical Engineering*, Cairo, Egypt, 12-01-2002.

[3] Shiying Hu, Eric A. Huffman, and Joseph M. Reinhardt," Automatic lung segmentation for accurate quantification of volumetric X-Ray CT images", *IEEE Transactions on Medical Imaging*, vol. 20, no. 6 ,pp. 490-498, June 2001

[4] Gurcan, M.N, Sahiner, B., Patrick, N., Chan, H., Kazerooni, E.A., Cascade, P.N. and Hadjiiski,L., "Lung Nodule Detection on Thoracic Computed Tomography Images: Preliminary Evaluation of a Computer-Aided Diagnosis System", *Medical Physics*, Vol. 29, No. 11, Pp. 2552- 2558, 2002.

[5] J.Padmavathi, "A comparative study on breast cancer prediction using RBF and MLP", *International Journal of Scientific and Engineering Research*", Vol.2, Issue.1,2011.

[6] R.M. Haralick, K. Shanmugam and I.H. Dinstein," Textural features for image classification," *IEEE Transactions on Systems, Man and Cybenatics*, vol.3, no. 6, pp. 610-621, Nov. 1973.

[7] R.C. Gonzales and R.E. Woods, *Digital Image Processing*: Pearson Education, 2003.

[8] Kanazawa, K., Kawata, Y., Niki, N., Satoh, H., Ohmatsu, H., Kakinuma, R., Kaneko, M., Moriyama, N. and Eguchi, K., "Computer-Aided Diagnosis for Pulmonary Nodules Based on Helical CT Images", *Compute Med. Image Graph*, Vol. 22, No. 2, Pp. 157-167, 1998.

[9] Ayman El-Baz, Aly A. Farag, Robert Falk, Renato La Rocca, "A Unified Approach for Detection, Visualization and Identification of Lung Abnormalities in Chest Spiral CT Scan", *Proceedings of Computer Assisted Radiology and Surgery*, London 2003.

[10] Samuel G. Armato III, Maryellen L. Giger and Catherine J. Moran, "Computerized Detection of Pulmonary Nodules on CT Scans", *Radio Graphics*, vol. 19, pp. 1303-1311, and 1999.

[11] S.K.Vijai Anand "Segmentation coupled Textural Feature Classification for Lung Tumor Prediction", *International Conference on Communication, Control and Computing Technologies 2010.*

[12] M. Gomathi, P.Thangaraj "A Computer Aided Diagnosis System for Lung Cancer Detection using Machine Learning Technique", *European Journal of scientific research*, Vol.51 No.2, pp.260-275, 2011.

[13] B.Magesh, P.Vijayalakshmi, M. Abirami "computer aided diagnosis system for identification and classification of lesions in lungs", *International Journal of Computer Trends and Technology*- May to June Issue 2011.

[14] Nisar Ahmed Memon, Anwar Majid Mirza, S.A.M. Gilani "Deficiencies Of Lung Segmentation Techniques using CT Scan Images for CAD", *world Academy of Science, Engineering and Technology 2006.*

[15] Laura Auria1 and Rouslan A. Moro, "Support Vector Machines (SVM) as a Technique for Solvency Analysis", Berlin August 2008

[16] H. Samet and M. Tamminen (1988). "Efficient Component Labeling of Images of Arbitrary Dimension Represented by Linear Bintrees". *IEEE Transactions on Pattern Analysis and Machine Intelligence.*

[17] Michael B. Dillencourt and Hannan Samet and Markku Tamminen (1992)."A general approach to connected-component labeling for arbitrary image representations".

[18] Kenji Suzuki and Isao Horiba and Noboru Sugie. "Linear-time connected-component labeling based on sequential local operations". *Computer Vision and Image Understanding, 2003.*

[19] Boser, Bernhard E.,Guyon, Isabelle M., and Vapnik, Vladimir N."A training algorithm for optimal margin classifiers".

SD-miner System to Retrieve Probabilistic Neighborhood Points in Spatial Data Mining

[1]Asst. Prof. S. G. Kulkarni, [2]Asst. Prof. Padma D, [3]Mr. Manjunath R. H
[1,2,3,]Department of Computer Science, Gogte Institute of Technology, Belgaum.

Abstract:
In GIS or Geographic Information system technology, a vast volume of spatial data has been accumulated, thereby incurring the necessity of spatial data mining techniques. Displaying and visualizing such data items are important aspects. But no RDBMS software is loaded with displaying the spatial result over a MAP overlay or answer spatial queries like "all the points within" certain Neighborhood. In this project, we propose a new spatial data mining system named SD-Miner. SD-Miner consists of three parts: A Graphical User Interface for inputs and outputs, a Data Mining Module that processes spatial data mining functionalities, a Data Storage Model that stores and manages spatial as well as non-spatial data by using a DBMS. In particular, the data mining module provides major spatial data mining functionalities such as spatial clustering, spatial classification, spatial characterization, and spatio-temporal association rule mining. SD-Miner has its own characteristics:
(1) It supports users to perform non-spatial data mining functionalities as well as spatial data mining functionalities intuitively and effectively.
(2) It provides users with spatial data mining functions as a form of libraries, thereby making applications conveniently use those functions.
(3) It inputs parameters for mining as a form of database tables to increase flexibility. Result shows that significantly reduced and precise data items are displayed through the result of this technique.

I. Introduction

Due to the development of information technology, a vast volume of data is accumulated on many fields. Since automated methods for filtering/analyzing the data and also explaining the results are required, a variety of data mining techniques finding new knowledge by discovering hidden rules from vast amount of data are developed. In the field of geography, due to the development of technology for remote sensing, monitoring, geographical information systems, and global positioning systems, a vast volume of spatial data is accumulated. Also, there have been many studies of discovering meaningful knowledge from the spatial data. Since the spatial data has its own characteristics different from the non-spatial data, direct using of general data mining techniques incurs many difficulties. So there have been many studies of spatial data mining techniques considering the characteristics of the spatial data. However, commercial tools for spatial data mining have not been provided. Currently, many commercial data mining tools are available, but these tools not support the spatial data mining functionalities. Also, while some academic spatial data mining tools such as Geo-Miner are available, there are almost no commercial spatial data mining tools. So, for easy using of spatial data mining for real spatial data applications, developments of spatial data mining tools are needed. In this paper, author proposes a new spatial data mining system named SD-Miner. SD-Miner supports four important spatial data mining functionalities: spatial clustering, spatial classification, spatial characterization, and spatio-temporal association rule mining. We first analyze characteristics of previous spatial data mining techniques and suggest techniques to improve their efficiency in developing SD-Miner [1].

II. Motivation

Spatial data mining that is, discovery of interesting, implicit knowledge in spatial databases is a highly demanding field because very large amount of spatial data have been collected in various applications, ranging from remote sensing, to Geographical Information Systems (**GIS**), computer cartography, environmental assessment and planning etc. Spatial data has its own characteristic different from non spatial data. Direct use of general data mining techniques incur many difficulties, so we need a spatial data mining tool for easy using of spatial data, for spatial data application, which can be easily done by SD-miner. The SD-miner supports four important spatial data mining functionalities as spatial clustering, spatial classification, spatial characterization and spatial temporal association rule mining [1].

III. Proposed Approach

This work attempts to develop a spatial data mining system, the SD-miner which consists of three parts: A graphical user interface for inputs and outputs, a data mining module that processes spatial data mining functionalities, a data storage model that stores and manages spatial as well as non spatial data by using a DBMS.

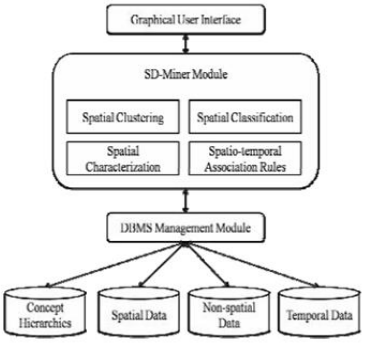
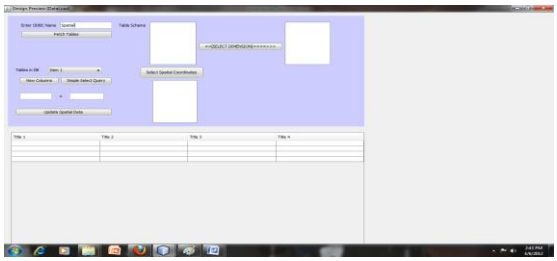


Fig 1: Architecture of SD-Miner

SD-Miner is composed of three main parts (Figure 1): Graphical User Interface (GUI), SD-Miner Module, and Data Storage Module (also, called DBMS management module) [1].

Graphical User Interface

A GUI is a type of computer human interface on a computer. It solves the blank screen problem that confronted early computer users. These early users sat down in front of a computer and faced a blank screen, with only a prompt. The computer gave the user no indication what the user was to do next. GUI is an attempt to solve this blank screen problem. Graphical user interface (GUI), a computer program that enables a person to communicate with a computer through the use of symbols, visual metaphors, and pointing devices. Best known for its implementation in Apple Inc.'s Macintosh and Microsoft Corporation's Windows operating system, the GUI has replaced the arcane and difficult textual interfaces of earlier computing with a relatively intuitive system that has made computer operation not only easier to learn but more pleasant and natural. The GUI is now the standard computer interface, and its components have themselves become unmistakable cultural artifacts.



Snap1- GUI

At a conceptual level, a computer human interface is a means by which "people and computers communicate with each other". In computer science terms, the GUI is a visual operating display that the monitor presents on the monitor to the computer operator. More specifically, a GUI is a specification for the look and feel of the computer system. GUI usually have common characteristic such as windows, icons, menus, and push-buttons (WIMP). Collectively, WIMP (Windows, Icons, Mouse, Pointer) are pictures that bring forth a certain action or an action space. The user issues commands via the GUI to computer applications. GUI usually has three major components. These three components are: a windowing system, an imaging model, and an application program interface (API). The windowing system builds the windows, menus, and dialog boxes that appear on the screen. The imaging model defines the fonts and graphics that appear on the screen. WIMP is products of both the windowing system and imaging model. Finally, the API is the means in which the user specifies how and what windows and graphics appear on the screen.The GUI contains some of the buttons and combo boxes which performs actions like:

1. Fetch tables- Which fetches the tables from the database.
2 .View column-Which displays columns according to the query written.
3. Simple select query-which specifies the condition in term of query.
4. Select spatial co-ordinates- this button helps us to select the required longuted and latitude.
5. Select dimensions.

SD-miner Module

The SD-Miner module processes respective data mining functions and transfers the results to the data storage module. This module provides four data mining functions as shown in fig-1.

Four important characteristics of SD-miner system:

Spatial Clustering

Spatial clustering classifies spatial objects as multiple groups according to its positional or geographical characteristics. This proposed work uses K means clustering.K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters). The main idea is to define k centroids, one for each cluster [2].

Spatial Classification

Usually, in the spatial classification, the objects are classified with considering of spatial and non-spatial attributes. The spatial classification also uses the decision tree. A big difference between the spatial classification and the classification is that the aggregation value of the spatial objects in a near region is used in the spatial classification. For making of a decision tree, the technique additionally uses predicates on relationship among spatial objects as decision criteria. For this, as the first step, the spatial attributes are represented as spatial predicates and then the possible useful predicates are extracted. For the second step, the decision tree is constructed with the predicates. The benefits of this technique are as follows. Since the decision tree is constructed after discarding ineffective predicates, the tree construction cost is greatly reduced. Speedy and correct classification can be made [3].

Spatial Characterization

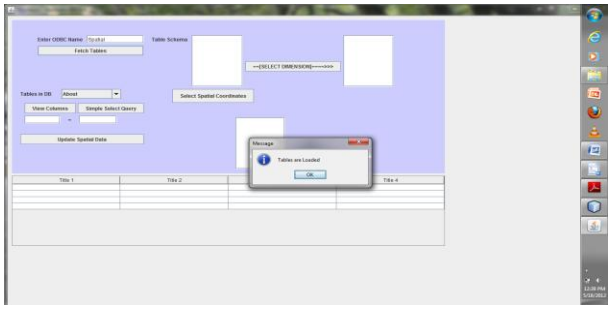
Spatial characterization extracts a global outline of data classes for a given spatial region by using the spatial objects of the region. It gives simple and clear abstract information of the region. Spatial characterization evaluates whether the characteristics of given spatial objects are expanded to near region. To do this, the objects are defined as a neighbor of each other with considering of their distance or direction The neighbor information is managed by using the neighbor table. The region handled by spatial characterization can be expanded with a spatial expansion algorithm using the neighbor table [4].

Spatio-Temporal Association Rule Mining

By using spatial association rule mining, proposed system can represent topological relationship and distance relationship of spatial objects via analyzing the relationship among spatial data and between spatial and non-spatial data. Also, by adding of temporal data analysis, proposed system can use spatio-temporal association rule mining.SD-Miner uses spatio-temporal association rule mining. In order to use spatial association rule mining, the spatial relationships between spatial objects must be defined. The relationships are represented as spatial predicates. The predicates defined by a user are stored as concept hierarchy data in the database of SD-Miner. If the spatial relationship is defined as the predicates, the predicates can be regarded as non-spatial attributes. So spatio-temporal association rules can be extracted by using the well-known Apriori algorithm. Spatial association rule mining can be applied whether the predicates are spatial or non-spatial. So association rules can be extracted among spatial predicates or among spatial and non-spatial predicates. This technique can be applied hierarchically by using level-based predicates if exist. So, proposed system can extract detailed as well as abstracted association rules in this case [5].

RESULTS

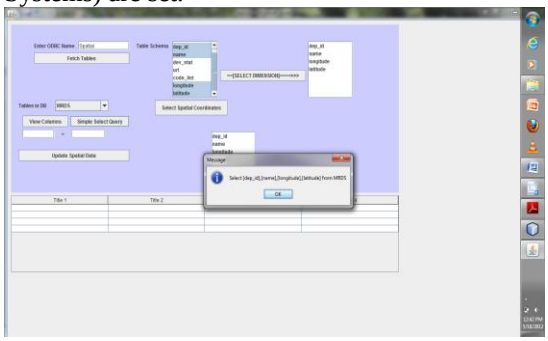
STEP 1: Load Data Base



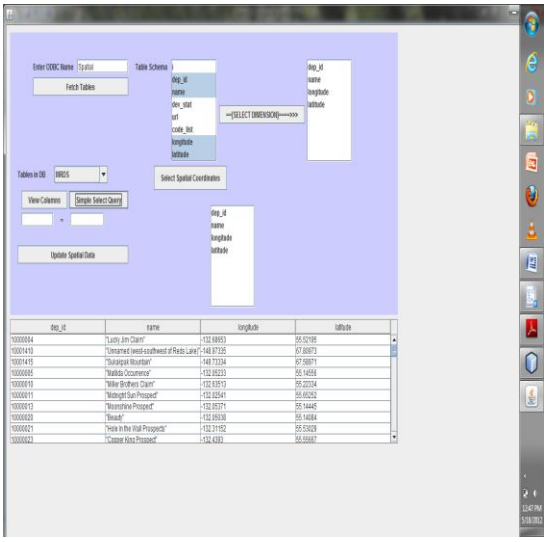
Snap 2: Illustrates Loading of Tables.

STEP 2: In Spatial Characterization phase, target objects will be set with respect to the database containing them as a description of the spatial properties.

In this proposed approach spatial properties like longitude, latitude and MRDS (Mineral Resources Data Systems) are set.



Snap 3: Illustrates Selecting of Required Characteristics.



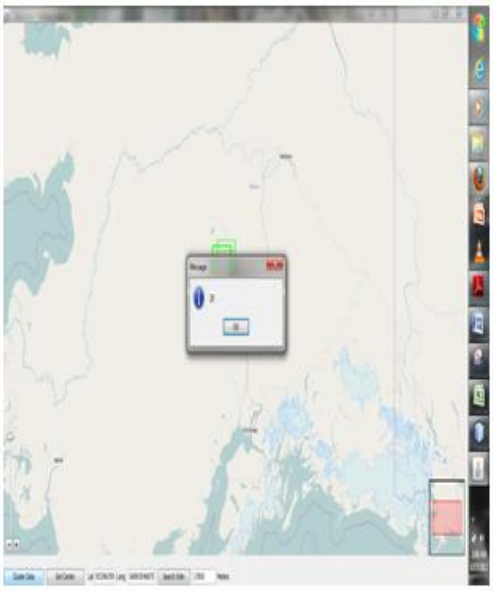
Snap 4: Illustrates list of selected characteristics in the form of columns and rows

Number of clusters	Range in meters
1	1000
4	2000
8	3000
25	4000
30	5000
38	6000
45	7000
52	8000
65	9000
71	10000

STEP 3: In Classification phase, the objects (MRDS) are assigned to a respective classes from a given set of classes based on the attribute values like longitude and latitude of the object.

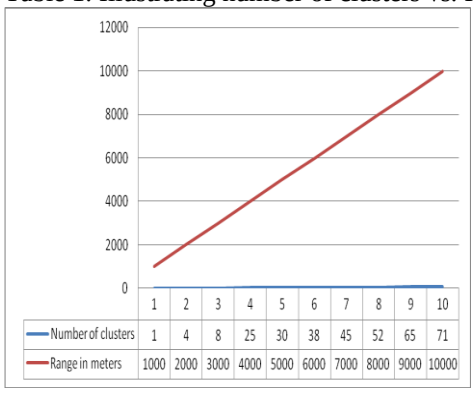
Snap 5: Illustrates The Classification Processes.

STEP 4: clustering classifies spatial objects (MRDS) as multiple groups according to specified longitude and latitude.



Snap 6: Illustrates clustering process.

Table 1: Illustrating number of clusters vs. Range



Graph: Number clusters vs. Range

The above table and graph shows us the variations of clusters according to the given range. As and when we change the parameters longitude and latitude the number of clusters will change, because it searches for the specified characteristic in the given range.

CONCLUSION

Spatial data has positional and topological data that do not exist in general data, and its structure is different according to the kinds of spatial data. Also, the objects on space affect each other and the relationship of objects is also different according to the kinds of objects. There have been many researches on spatial data mining considering these characteristics of spatial data.In this work, author explains the concept of spatial clustering, spatial classification, spatial characterization, and spatio-temporal association rule mining. We present our experiences in developing a spatial data mining system called SD-Miner that provides proper spatial data mining techniques with improved effectiveness and efficiency for real applications. SD-Miner adopts following techniques. For spatial clustering, proposed system adapt NN Superseeding algorithm. For spatial characterization, proposed system uses a binary decision tree and the RELIEF algorithm for efficiency. For spatial characterization, we use the neighbor table for spatial extension of the current characterization method. For spatio-temporal association rule mining, we use temporal data with spatial association rule mining using the spatial concept layer. SD-Miner uses spatial data mining functions extended from general mining functions. So, it can be applied to both of spatial and non-spatial data. Without special intervention of a user, it automatically recognizes which type of data is used. All functions are developed in library style, and the functions can be reused in another system.

IV. REFERENCES

[1] SD-Miner: A SPATIAL DATA MINING SYSTEMDuck-Ho Bae, Ji-Haeng Baek, Hyun-Kyo Oh, Ju-Won Song, Sang-Wook Kim

[2] J. Sander et al., "Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications," *Data Mining and Knowledge Discovery*, Vol. 2, No. 2, pp. 169-194, 1998

[3] M. Ester et al., "Algorithms for Characterization and Trend Detection in Spatial Databases," In *Proc. Int'l. Conf. on Knowledge Discovery and Data Mining*, KDD, pp. 44-50, 1998.

[4] K. Koperski, J. Han, and N. Stefanovic, "An Efficient Two-Step Method for Classification of Spatial Data," In *Proc. Int'l. Symp. On Spatial Data Handling*, SDH, pp. 45-54, 1998.

[5] Spatio-Temporal Data Mining On Moving Objects In DBMS Yahaya Bin Abd Rahim GIS Section OTB Research Institute for Housing, UrbanandMobility Studies Delft University of Technology November 15 2007.

[6] M. Ester et al., "Spatial Data Mining: Database Primitives, Algorithms and Efficient DBMS Support," *Data Mining and Knowledge Discovery*, Vol. 4, pp. 193-216, 2000.

[7] M. Ester, H. Kriegel, and J. Sander, "Algorithms and Applications for Spatial Data Mining," Geographic *Data Mining and Knowledge discovery*, 2001.

[8] J. Han, K. Koperski, and N. Stefanovic, "GeoMiner: A System Prototype for Spatial Data Mining," In *Proc. ACM Int'l. Conf. on Management of Data*, ACM SIGMOD, pp. 553-556, 1997.

[9] W. Lu, J. Han, and B. Ooi, "Discovery of General Knowledge in Large Spatial Databases," In *Proc. Far East Workshop on Geographic Information Systems*, pp. 275-289, 1993.

[10] E. Knorr and R. Ng, "Finding Aggregate Proximity Relationships and Commonalities in Spatial Data Mining," *IEEE Trans. On Knowledge and Data Engineering*, IEEE TKDE, Vol. 8, pp. 884-897, 1996.

Implementation of Elliptic Curve Digital Signature Algorithm Using Variable Text Based Message Encryption

[1]Jayabhaskar Muthukuru, [2]Prof. Bachala [3]Sathyanarayana
[1,2] Research Scholar, Department of Computer Science & Technology,
[1,2],Professor, Sri Krishnadevaraya University, INDIA

Abstract:
Digital Signatures are considered as digital counterparts to handwritten signatures, and they are the basis for validating the authenticity of a connection. It is well known that with the help of digital signature, forgery of digital information can be identified and it is widely used in e-commerce and banking applications. Elliptic curve digital signatures (ECDSA) are stronger and ideal for constrained environments like smart cards due to smaller bit size, thereby reducing processing overhead. We have implemented ECDSA over Elliptic Curve (EC) P-192 and P-256 using various Text Message encryptions which are Variable Size Text Message(VTM), Fixed Size Text Message(FTM) and Text Based Message(TBM) encryption methods and compared their performance.

Keywords: Digital Signature, Elliptic Curve Digital Signature Algorithm, Elliptic Curve Cryptography, ECDLP.

1. Introduction

Cryptography is the branch of cryptology dealing with the design of algorithms for encryption and decryption, intended to ensure the secrecy and/or authenticity of message. The Digital Signature Algorithm (DSA) was proposed in August 1991 by the U.S. National Institute of Standards and Technology (NIST). Digital signature authentication schemes provide secure communication with minimum computational cost for real time applications, such as electronic commerce, electronic voting, etc. The sender generates the signature of a given message using his secret key; the receiver then verifies the signature by using sender's public key. The ECDSA have a smaller key size, which leads to faster computation time and reduction in processing power, storage space and bandwidth. This makes the ECDSA ideal for constrained devices such as pagers, cellular phones and smart cards. The Elliptic-Curve Digital Signature Algorithm (ECDSA) is a Digital Signature Scheme based on ECC. ECDSA was first proposed in 1992 by Scott Vanstone in response of NIST (Nation Institute of Standards and Technology) request for public comments on their proposal for Digital Signature Schemes[1].
Digital Signature authenticated schemes, have the following properties.

1. **Confidentiality.** Secret information shared between sender and receiver; any outsider cannot read the information.
2. **Authentication.** The sender imprints his identity by means of the digital signature, which only the designated receiver can unravel and verify. An anonymous adversary cannot send a malicious message impersonating the genuine sender, because he does not have the necessary tools to generate the signature.
3. **Non-repudiation.** The signature firmly establishes the identity of the sender. The sender cannot deny having sent the message and the signature.

In this paper we discuss ECC in detail and ECDSA Implementation with different Text Message encryption methods and compared the results.

2. Elliptic Curve Discrete Logarithm Problem

An elliptic curve E, [2] defined over a field K of characteristic $\neq 2$ or 3 is the set of solutions $(x, y) \in K'$ to the equation

$$y^2 = x^3 + ax + b \quad (1)$$

$a, b \in K$ (where the cubic on the right has no multiple roots).
Two nonnegative integers, a and b, less than p that satisfy:

$$4a^3 + 27b^2 \pmod p = 0 \quad (2)$$

Then $E_p(a, b)$ denotes the elliptic group mod p whose elements (x, y) are pairs of nonnegative integers less than p satisfying:

$$y^2 = x^3 + ax + b \pmod p \quad (3)$$

together with the point at infinity O.
The elliptic curve discrete logarithm problem(ECDLP) can be stated as follows. Fix a prime p and an elliptic curve.

$$Q = xP \quad (4)$$

where xP represents the point P on elliptic curve added to itself x times. Then the elliptic curve discrete logarithm problem is to determine x given P and Q. It is relatively easy to calculate Q given x and P, but it is very hard to determine x given Q and P.

ECC is based on ECDLP. ECDH and ECDSA are cryptographic schemes based on ECC. The best known algorithm for solving ECDLP is Pollard-Rho algorithm which is fully exponential having a running time of $\sqrt{(\Pi * n / 2)}$.

3. Elliptic Curve Cryptography

The Elliptic curve cryptosystems (ECC) were invented by Neal Koblitz [2] and Victor Miller[3] in 1985. They can be viewed as elliptic curve analogues of the older discrete logarithm (DL) cryptosystems in which the subgroup of Z_p^* is replaced by the group of points on an elliptic curve over a finite field. The mathematical basis for the security of elliptic curve cryptosystems is the computational intractability of the elliptic curve discrete logarithm problem (ECDLP) [4].

ECC is a relative of discrete logarithm cryptography. An elliptic curve E over Z_p as in Figure 1 is defined in the Cartesian coordinate system by an equation of the form:

$$y^2 = x^3 + ax + b \qquad (5)$$

where a, b $\in Z_p$, and $4a^3 + 27b^2 \pmod p \neq 0 \pmod p$, together with a special point O, called the point at infinity. The set $E(Z_p)$ consists of all points (x, y), $x \in Z_p$, $y \in Z_p$, which satisfy the defining equation, together with O.

Each value of a and b gives a different elliptic curve. The public key is a point on the curve and the private key is a random number. The public key is obtained by multiplying the private key with a generator point G in the curve.

The definition of groups and finite fields, which are fundamental for the construction of elliptic curve cryptosystem are discussed in next subsections.

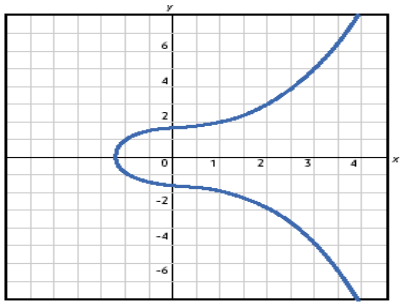


Figure 1. An Elliptic Curve

3.1. Groups

A group with an operation * is defined on pairs of elements of G. The operations satisfy the following properties:

- Closure: a * b $\in$G for all a, b $\in$G
- Associativity: a * (b * c) = (a * b) * c for all a, b, c $\in$G
- Existence of Identity: There exists an element e $\in$G, called the identity, such that e * a = a * e = a for all a $\in$G.
- Existence of Inverse: For each a $\in$G there is an element b$\in$G such that a * b = b * a = e. The element b is called the inverse of a.

Moreover, a group G is said to be abelian if a * b = b * a for all a, b $\in$G. The order of a group G is the number of elements in G.

3.2. Finite Field

A finite field consists of a finite set of elements together with two binary operations called addition and multiplication, which satisfy certain arithmetic properties. The order of a finite field is the number of elements in the field. There exists a finite field of order q if and only if q is a prime power. If q is a prime power, then there is essentially only one finite field of order q; this field is denoted by F_q. There are, however, many ways of representing the elements of F_q. Some representations may lead to more efficient implementations of the field arithmetic in hardware or in software. If $q = p^m$ where p is a prime and m is a positive integer, then p is called the characteristic of F_q and m is called the extension degree of F_q.

3.2.1. Prime Field F_p

Let p be a prime number. The finite field F_p called a prime field, is comprised of the set of integers {0,1,2,….,p-1} with the following arithmetic operations:

- Addition: If a, b$\in F_p$ then a + b = r, where r is the remainder when a + b is divided by p and $0 \leq r \leq p-1$ known as addition modulo p.
- Multiplication: If a, b $\in F_p$ then a.b = s, where s is the remainder when a.b is divided by p and $0 \leq s \leq p-1$ known as multiplication modulo p.

- Inversion: If a is non-zero element in F_p, the inverse of modulo a modulo p, denoted by a^{-1}, is the unique integer $c \in F_p$ for which a.c = 1.

3.2.2. Binary Field F_2^m

The field F_2^m, called a characteristic two finite field or a binary finite field, can be viewed as a vector space of dimension m over the field F_2 which consists of the two elements 0 and1. That is, there exist m elements $\alpha_0, \alpha_1, \ldots, \alpha_{m-1}$ in F_2^m such that each element α can be uniquely written in the form:

$$\alpha = a_0 \alpha_0 + a_1 \alpha_1 + \ldots\ldots\ldots + a_{m-1} \alpha_{m-1}, \text{ where } a_i \in \{0,1\}$$

Such a set $\{\alpha_0, \alpha_1, \ldots, \alpha_{m-1}\}$ is called a basis of F_2^m over F_2. Given such a basis, a field element α can be represented as the bit string $(a_0 + a_1 \ldots\ldots\ldots + a_{m-1})$ Addition of field elements is performed by bitwise XOR-ing the vector representations. The multiplication rule depends on the basis selected. ANSI X9.62 permits two kinds of bases: polynomial bases and normal bases.

3.2.3. Domain Parameters

The domain parameters for ECDSA consist of a suitably chosen elliptic curve E defined over a finite field F_q of characteristic p, and a base point $G \in E(Fq)$. Domain parameters may either be shared by a group of entities, or specific to a single user. To summarize, domain parameters are comprised of:

1. A field size q, where either q = p, an odd prime, or $q = 2^m$
2. An indication FR (field representation) of the representation used for the elements of F_q
3. (optional) a bit string seed E of length at least 160 bits
4. Two field elements a and b in F_q which define the equation of the elliptic curve E over F_q' (i.e., $y^2 = x^3 + ax + b$ in the case p > 3, and $y^2 + xy = x^3 + ax + b$ in the case p = 2)
5. Two field elements x_G and y_G in F_q which define a finite point $G = (x_G, y_G)$ of prime order in $E(F_q)$
6. The order of the point G, with $n > 2^{160}$ and $n > 4\sqrt{q}$ and
7. The cofactor $h = \#E(F_q)/n$

3.3. Elliptic Curve Operations over Finite Fields[8]

The main operation is Point multiplication is achieved by two basic elliptic curve operations.

i. Point addition, adding two points P and Q to obtain another point R i.e. R= P + Q.
ii. Point doubling, adding a point P to itself to obtain another point R i.e. R = 2P.

3.3.1. Point Addition

Point addition is the addition of two points P and Q on an elliptic curve to obtain another point R on the same elliptic curve. Consider two points P and Q on an elliptic curve as shown in Figure 2. If $P \neq -Q$ then a line drawn through the points P and Q will intersect the elliptic curve at exactly one more point –R. The reflection of the point –R with respect to x-axis gives the point R, which is the result of addition of points P and Q. Thus on an elliptic curve R = P + Q. If Q = -P the line through this point intersect at a point at infinity O. Hence P + (-P) = O. A negative of a point is the reflection of that point with respect to x-axis.

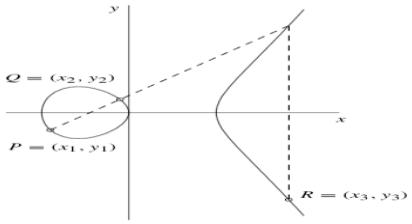


Figure 2: Point Addition

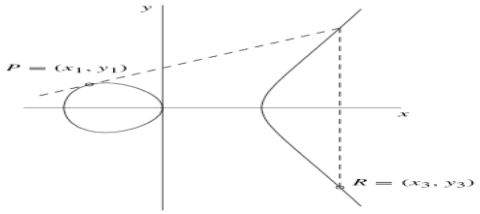


Figure 3: Point Doubling

3.3.2. Point Doubling

Point doubling is the addition of a point P on the elliptic curve to itself to obtain another point R on the same elliptic curve To double a point J to get L, i.e. to find R = 2P, consider a point P on an elliptic curve as shown in Figure 3. If y coordinate of the point P is not zero then the tangent line at P will intersect the elliptic curve at exactly one more point –R. The reflection of the point –R with respect to x-axis gives the point R, which is the result of doubling the point P, i.e., R = 2P. If y coordinate of the point P is zero then the tangent at this point intersects at a point at infinity O. Hence 2P = O when $y_j = 0$. Figure 3 shows point doubling.

3.3.3. Algebraic Formulae over F_p

Let p be a prime in F_P and a, b$\in$ FP such that $4a^3 + 27b^2 \neq 0$ mod p in F_P, then an elliptic curve E (F_P) is defined as

$$E(F_P) := \{ p(x, y), x, y \in F_P \}$$

Such that $y^2 = x^3 + ax + b$ mod p together with a point O, called the point at infinity. Below is the definition of addition of points P and Q on the elliptic curve E (F_P). Let $P(x_1, y_1)$ and $Q(x_2, y_2)$ then

$$R = P+Q = \begin{cases} \text{If } x_1 = x_2 \text{ and } y_2 = -y_1 \\ Q = Q+P \quad \text{If } P = O \\ (x_3, y_3) \quad \text{otherwise} \end{cases}$$

$$\text{Where } x_3 = \begin{cases} \lambda^2 - x_1 - x_2 & \text{If } P \neq \pm Q \text{ (Point Addition)} \\ \lambda^2 - 2x_1 & \text{If } P = Q \text{ (Point Doubling)} \end{cases}$$

$$y_3 = \lambda(x_1 - x_3) - y_1, \text{ and}$$

$$\lambda = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1} & \text{If } P \neq \pm Q \text{ (Point Addition)} \\ \dfrac{3x_1^2 + a}{2y_1} & \text{If } P = Q \text{ (Point Doubling)} \end{cases}$$

The point p(x, -y) is said to be the negation of p(x, y).

3.3.4. Algebraic Formulae over F_2^m

Denote the (non-super singular) elliptic curve over F_2^m by E (F_2^m). If a, b $\in F_2^m$ such that b $\neq 0$ then

$$E(F_2^m) = \{p(x, y), x, y \in F_2^m\}$$

such that $y^2 + xy = x^3 + ax^2 + b \in F_P^m$ together with a point O, called the point at infinity.
The addition of points on E (F_2^m) is given as follows: Let $P(x_1, y_1)$ and $Q(x_2, y_2)$ be points on the elliptic curve E(F_2^m), then

$$R = P+Q = \begin{cases} O & \text{If } x_1 = x_2 \text{ and } y_2 = -y_1 \\ Q = Q+P & \text{If } P = O \\ (x_3, y_3) & \text{otherwise} \end{cases}$$

$$\text{Where } x_3 = \begin{cases} \lambda^2+\lambda + x_2 + x_1 + a & \text{If } P \neq \pm Q \text{ (Point Addition)} \\ \lambda^2 + \lambda + a & \text{If } P = Q \text{ (Point Doubling)} \end{cases}$$

$y_3 = \lambda (x_1 + x_3) + x_3 + y_1$ and

$$\lambda = \begin{cases} \frac{y_2 + y_1}{x_2 + x_1} & \text{If } P \neq \pm Q \text{ (Point Addition)} \\ x_1 + \frac{x_1}{y_1} & \text{If } P = Q \text{ (Point Doubling)} \end{cases}$$

4. Implementation

This paper presents VTM Encryption, VTM decryption [5], ECDSA key generation, signature generation and signature verification algorithms [8] and ECDSA was implemented over Elliptic Curve (EC) P-192 and P-256 using Text Message Encryption methods which are VTM [5], FTM[5] and TBM [6] encryption methods and compared their performance.

Algorithm-1

VTM Encryption Algorithm[5]

NOTATION: TM - Text message

M - Message units
VS - variable size
IV - Initial Vector
k - Auxiliary base parameter
XRM - XORed message
Block – a word with followed space

INPUT: sextuple T = (p, a, b, G, n, h), Text Message

OUTPUT: Encrypted Message

Begin
n = wordCount(TM)
for i = 1 to n **do**
XRM = IV ⊕ Block[i]
M = ASCII(XRM)
for j = 0 to k-1 **do**
let $x_j = M * K + j \bmod p$
if $z_j = x_j^3 + x_j + b$ has a square root mod p **then**
break
end if
end for
if j < k **then**
compute y_j a square root of $z_j \bmod p$
map M to (x_j, y_j)
else
output "unsuccessful in attempt to map M to an EC point"
end if
Cm[i] = { kG ,Pm + kP_B}
IV = XRM
end for
End

Algorithm-2

VTM Decryption Algorithm[5]

INPUT: sextuple T = (p, a, b, G, n, h), Encrypted Message

OUTPUT: Decrypted/Plain text Message

Begin
for i = 1 to n **do** //where n is number of cipher texts
Pm(x, y) = Pm + K(n_BG) - n_B(kG) // n_B receivers private key
M = x/k
Dm = Text(M) // M is decimal value of base 256 format
TM[i] = Dm ⊕ IV
IV = Dm
TM = TM || TM[i]
end for
End

Algorithm-3

ECDSA Key pair generation Algorithm[8]

INPUT: Domain parameters D= (q, FR, a, b, G, n, h).

OUTPUT: Public key Q, private key d.

Select d∈ [1, … , n-1]
Compute Q = dG
Return (Q, d)

Algorithm-4

ECDSA Signature Generation Algorithm[8]

INPUT: Domain parameters D= (q, FR, a, b, G, n, h) , private key d, Encrypted message m'.

OUTPUT: Signature (r,s)

begin
repeat
k = Random[1, … , n-1] // select random value
r = x-coord([k]G) mod n
e = H(m')
s = k^{-1}(e+dr) mod n
until r ≠ 0 and s ≠ 0
return (r,s).
end

Algorithm-5

ECDSA Signature Verification Algorithm[8]

INPUT: Domain parameters D= (q, FR, a, b, G, n, h) , public key Q, Encrypted Message m', Signature (r, s).
OUTPUT: Acceptance or rejection of the signature.

begin
if r, s ∉ [1,…, n] **then**
Return (“Reject the signature”)
end if
e = H(m')
w = s^{-1} mod n
u_1 = ew mod n
u_2 = rw mod n

$x = u_1G + u_2Q$
if $x = \infty$ **then**
Return ("Reject the signature")
end if
v = x-coord(X) mod n
if v = r **then**
Return ("Accept the signature")
else
Return ("Reject the signature")
end if
end.

Elliptic Curve based Signature Generation & Signature Verification processes are described below and the same is represented in graphical format in figure 4 and figure 5.

Signature Generation steps:

1. Encrypt the message using EC Encryption algorithm which is VTM/FTM/TBM
2. Compute signature for Encrypted message using Algarithm-4
3. Send the digitally signed message

Signature Verification Steps:

1. Verify Signature using Algarithm-5.
2. If verification fails then reject the signature
3. If verification success, then decrypt the message using respective EC Decryption Algorithm.

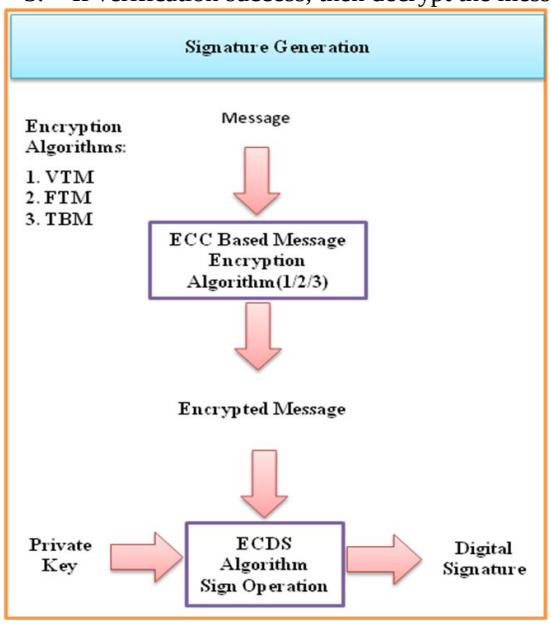


Figure 4: Signature Generation Process

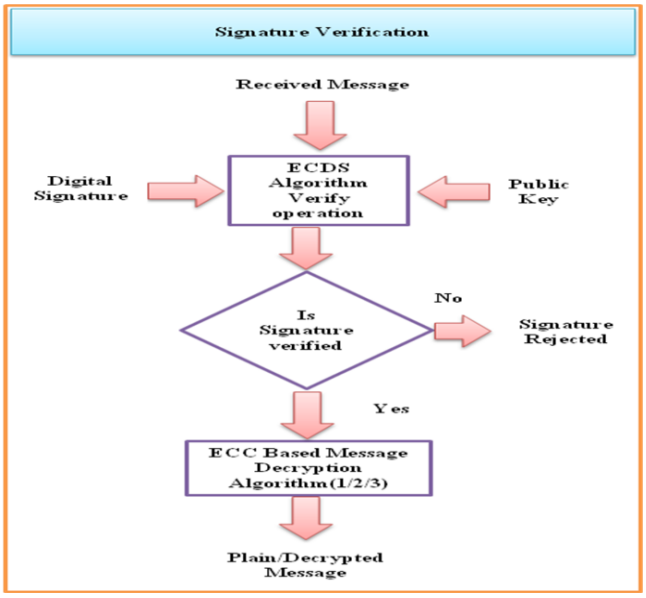


Figure 5: Signature Verification Process

5. Results and Discussion

In this section represents implementation results of ECDSA using VTM encryption over EC P-192 and P-256.

5.1. Results over Elliptic Curve P-192

Message m = "Test Run"
Private key = 2055107281
Public Key = (584194271639147920155034229735108596327098351992499437760 2,
5584890377300947026793868981513336619407548239394095574193)
This message encrypted and follows Signature Generation and Verification as mentioned below.
Encrypted message hash value H(E(m)) = -26821089969772781569684086062354389451610645 54

- ECDSA SIGNATURE as follows:

Select k= 1583021364
Compute kG = (37921946278159604401180029145945511663128641788889626308 82,

2891190659620656059990718022662146728564853605540168001982)
r = 37921946278159604401180029145945511663128641788889626308 82
Compute s = k^{-1} (e + dr) mod n = 341118468161025230839050235906555456270860509373907548348 3
Signature for the message m is (r, s).

- ECDSA VERIFICATION as follows:

Compute w = 577748014580366974157342368892617697941708250527103236026 8
Compute u_1 = 466642252724903410004202294633709000851059727718411130369 6
u_2 = 445590792742988647327720447499023685312487717133566127164 9
u_1G = (392970898996946769719748671667212244694231563209483104336 7,
453700345657110338028450481372179209611919804754395949167 1)
u_2Q = (127766171580020534806742076601680647595413362692969638337 0,
438080846038756764910705428973258588684808820612544874244 7)
v = 37921946278159604401180029145945511663128641788889626308 82
We obtain v = r, that is accept the signature.

5.2. Results over Elliptic Curve P-256

Message m= "How are you?"
Private Key = 978425864
Public Key = (118910487909274429022743485742135581553673510998540082125096949934594470938 22,
136698797209684711142721957596171372481001364004993589753744001635050991639 86)
This message encrypted and follows Signature Generation and Verification as mentioned below.
Encrypted message hash value H(E(m)) = 53770309037964977040219539705106232306909249184 6

- ECDSA SIGNATURE as follows:

Select k= 11579208921035624876269744694940757352999695522413576034242225906106838350224 3
Compute
KG = (86500881224166483227925267313354237293018428812409245047778807509807358555053,
39579053610346434470532506438011786967057506613223689314593851851982117599776)
r = 86500881224166483227925267313354237293018428812409245047778807509807358555053
Compute s = k^{-1} (e + dr) mod n
= 10438970071550173279661477973785546374937584448654061862201805470297056109170 8
Signature for the message m is (r, s).

- ECDSA VERIFICATION as follows:

Compute w = 10650639697755614553541805405233944739307883299318145000266847025131237147427 6
Compute u_1 =43824495211803284954034352427133274304161118431427286644319226927046995292 09
u_2=57692616982311160984176366728847647733800539362706147029132815066162592219439
u_1G = (10147462789339256415094921370320020372887311198 48 92002825714765996844262058436,
60937423109159230990348336949980 80 45643619656906462116717265149991515547954 08)
u_2Q = (10932210314568305562895697128244517730737835573471227859803024987190651216376 6,
42753639382524136274231334284305572212602843186842236043136827079395299552547)
v = 86500881224166483227925267313354237293018428812409245047778807509807358555053
We obtain v = r, that is accept the signature.

In the same way we have used FTM and TBM encrypted message for Signature generation and signature verification. ECDSA using Variable Size Text Message Encryption is better in performance aspect when compare with the other two methods and the results comparison is presented graphically in the next section.

6. Comparison Of ECDSA Using Various Text Based Cryptosystems

We compare the results of ECDSA using Variable Size Text Message(VTM) Encryption[5] with ECDSA using Fixed Size Text Message(FTM) Encryption[5] and Text Based Message(TBM) Encryption[6]. Figure 6 and Figure 7 presents total time taken for Signature Generation and Signature Verification when we use different text based encryption methods in ECDSA implementation. From Figure 6 and Figure 7, performance of ECDSA using Variable Size Text Message Encryption is better when compare with ECDSA using FTM Encryption and TBM Encryption. The reason is VTM based ECDSA used less number of point additions and multiplications compare with other two methods. Performance of ECDSA is inversely proportional to key size, and security of the system depends on key size.

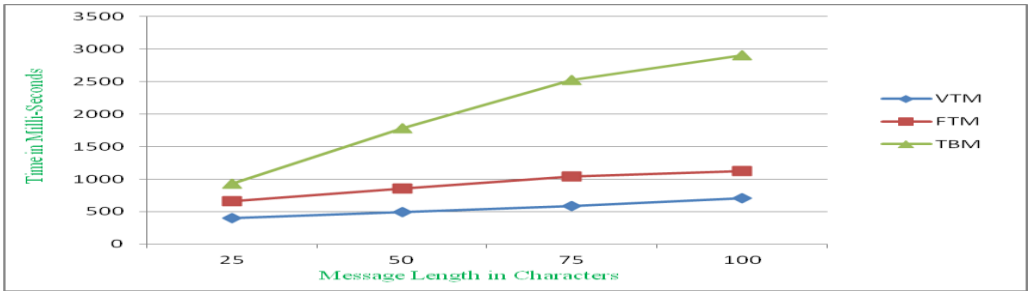


Figure 6: Performance comparison of various ECDSA methods for over EC P-192

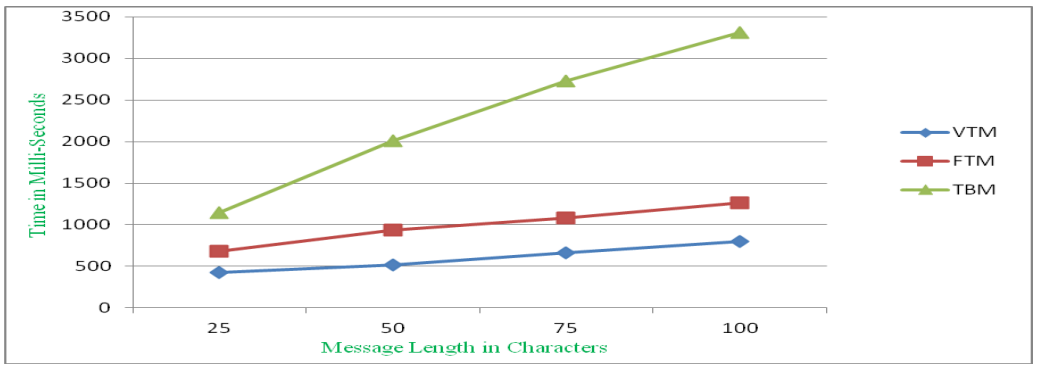


Figure 7: Performance comparision of various ECDSA methods for over EC P-256

7. Conclusion

In this paper we have implemented ECDSA for various domain parameters, after observing the results when the key size increases then complexity increases and performance decreased. After comparing VTM, FTM and TBM based ECDSA methods, ECDSA using Variable Text Message Encryption is better when comparing with Fixed Length Text Message and Text Based Encryption used ECDSA. The main reason is, the speed of scalar multiplication which plays an important role in the efficiency of whole system [7]. In VTM based ECDSA method, number of scalar multiplications are reduced, so this method is efficient when compared with FTM and TBM based methods.

References

[1] Navneet Randhawa, Lolita Singh, A Systematic Way to Provide Security for Digital Signature Using Elliptic Curve Cryptography, IJCST Vol.2, Issue 3, Sep-2011, 185-188
[2] Koblitz, N., 1987. Elliptic curve cryptosystems. Mathematics of Computation 48, 203-209.
[3] Miller, V., 1985. Use of elliptic curves in cryptography. CRYPTO 85.
[4] Certicom ECC Challenge. 2009. Certicom Research
[5] Jayabhaskar Muthukuru, Bachala Sathyanarayana, Fixed and Variable Size Text Based Message Mapping Techniques Using ECC, GJCST Vol.12, Issue 3, Feb-2012, 25-30.
[6] S. Maria Celestin Vigila , K. Muneeswaran "Implementation of Text based Cryptosystem using Elliptic Curve Cryptography", IEEE Sep-2009, pp. 82-85.
[7] Harsandeep Brar , Rajpreet Kaur, "Design and Implementation of Block Method for Computing NAF" IJCA, Volume 20– No.1, April 2011, pp. 37-41.
[8] Hankerson, D., Menezes, A., Vanstone, S., Guide to Elliptic Curve Cryptography (Springer, 2004).

Colorization Of Gray Image In *Lαβ* Color Space Using Texture Mapping And Luminance Mapping

[1]Mrs. Smriti Kumar, [2]Mr.Deepak Singh

[1,2],Department of Electronics and Telecommunication Engineering, CSIT Durg, India.

Abstract:
Gray Scale image colorization is an appealing area in the world of image processing. This work presents a simple process for colorizing gray image, using a colored image which is similar to this gray scale image but not the colorized version of the gray image. Here we convert both these images to a decorrelated color space *lαβ* and then divide these images into small windows of equal size, then we calculate mean and standard deviation for data points based on luminance values in each of these windows to compute a luminance measure and we also extract texture features like energy, entropy, contrast, correlation and homogeneity based on correlation matrix for each of these windows. We apply proper weights to these texture features to calculate a texture similarity measure and then calculate a similarity measure based on the texture similarity measure and luminance measure for each of these windows in both the images. Then we compare each window in gray scale image with every other window in colored image using this similarity measure and find the best matching window. Then the chromatic properties of some randomly selected data points of this colored window are transferred to the gray window pixel by pixel, based on the luminance mapping and thus colorization of gray scale image is achieved, which produces believable results.

Keywords: Colorization, Contrast, Correlation, Energy, Entropy, Gray Scale Images, Homogeneity, lαβ Color Space, Luminance Measure, Mean And Standard Deviation, Similarity Measure, Texture Feature, Texture Similarity Measure.

1. Introduction

Color can increase the visual appeal of grayscale images such as the black and white photos, old movies or scientific data. In medicine, image modalities which only acquire grayscale images such as Computerized Tomography (CT), X-ray images and Magnetic Resonance Imaging (MRI) can be enhanced when colorized for demonstrations and presentations. Even for some scientific images information content can be enhanced with color when variations in chromaticity as well as luminance are exploited. The task of adding colors to a grayscale image involves assigning three-dimensional (RGB) pixel values to a gray scale image but it is also a fact that gray scale images vary along only one dimension (luminance or intensity).This fact is also well known that a variety of colors may have the same luminance value although hue or saturation variation is possible, the problem of colorizing grayscale images has no inherently "correct" solution. Due to these uncertainties, in the colorization process there is a large role of human interaction. Colorization is a term introduced in 1970 and was later patented by Wilson Markle. Colorization trends have been classified into three categories - Hand coloring is the first one, Semi automatic coloring stands next, and automatic coloring is the third one. Semary N. A. [1]. Hand coloring has been a way of showing talent in art. Usually image editing software's like Adobe Photoshop or Paint shop Pro are used similarly to convert gray images to color images. Semiautomatic coloring refers to mapping luminance values to color values, which is done by converting a gray image in to an indexed image. Pseudo coloring is a common example [2][3]. Automatic coloring can be classified into three categories [1], which are transformational coloring, Image Matching and User Selection, this classification is made according to the source of colors to be transferred to the gray image pixels. Technique of our interest here is Image Matching/ Coloring by Reference. This technique is based on the fact that for each gray pixel $I_g(x, y)$ in target gray scale image there exists a colored pixel $I_s(x_2, y_2)$ in reference or source image such that the distance E (which is some similarity measure) between them exhibits the minimum value. The chromatic properties in "I_s" are transferred to "I_g" and the achromatic value of "I_g" are retained. When one represents any typical three channel image in any of the well-known color spaces, the different channels' values will be correlated. For instance, in RGB color space, most pixels for the red and green channel will have large values if the blue channel is large. This implies that if we want to alter the chromatic properties of a pixel in a coherent way, we must modify all color channels in tandem. Thus any type of color modification process gets complicated. So we require an orthogonal color space without correlations between the axes. Ruderman et al [4] developed a color space, called *lαβ*, which minimizes correlation between channels for many natural scenes, since it is based on data-driven human perception research that assumes that the human visual system is ideal for processing natural scenes. There's little correlation between the axes in *lαβ* space, which allows the application of different operations in different color channels without the troubles, due to undesirable cross-channel artifacts. Apart from this, *lαβ* color space is logarithmic, which means that to a first approximation the uniform changes in channel intensity are equally detectable [5]. This makes *lαβ* space, a very popular choice where colorization is concerned. The complete coding of this work has been done in MATLAB 10.

2. Methodology

For colorization purpose we convert our source (i.e. Color Image) and target (i.e. Gray Scale Image) images into decorrelated color space *lαβ* and then divide these images into small windows, whose size may vary according to the type of image (We tried various images and found that, a window size of '5 X 5' suffices for most of the images but some challenging images require larger window size). After this we extract texture features like energy, entropy, homogeneity, contrast and correlation based on correlation matrix for the purpose of texture matching, for each window. We also calculate mean and standard deviation for each pixel in this window and find out an average value of these. Then we compare each target (gray scale) image's window with all the windows of source (colored) image and find out the best matching window in source image based on a similarity measure which gives equal weight to both texture feature matching as well as mean and standard deviation matching (luminance matching). Then the chromatic values of some randomly selected data points from the source image window are transferred to the target image window, based on luminance mapping. The resulting image of this process is colorized version of target gray scale image. The basic method here matches the three-dimensional distribution of texture features and luminosity values between the images and then transforms the color distribution of the source (colored) image to match the distribution of the target (grayscale) image. A grayscale image is represented by a one dimensional distribution, hence between the two images one can match only the luminance channels. Apart from this we also match texture features like energy, entropy, homogeneity, contrast and correlation based on correlation matrix. Once a target image window matches a source image window, the color information is transferred between pixels but the original luminance value is retained. The *l* channel in *lαβ* color space determines the luminance value. For colorization purpose, mean and standard deviations along each of the three axes are required. Hence, we compute these measures for the windows in both the source and target images, for each pixel in each of the windows. We calculate the mean and standard deviation [9][10] for each axis separately in *lαβ* space. For this first, we subtract the mean from the data points:

$$l^* = l - \langle l \rangle, \quad (1.1)$$
$$\alpha^* = \alpha - \langle \alpha \rangle, \quad (1.2)$$
$$\beta^* = \beta - \langle \beta \rangle, \quad (1.3)$$

Then, we scale the data points by factors determined by the respective standard deviations:

$$l' = (\sigma_\tau^{l} / \sigma_\sigma^{l})\, l^*, \quad (2.1)$$
$$\alpha' = (\sigma_\tau^{\alpha} / \sigma_\sigma^{\alpha})\, \alpha^*, \quad (2.2)$$
$$\beta' = (\sigma_\tau^{\beta} / \sigma_\sigma^{\beta})\, \beta^*, \quad (2.3)$$

Using above formulae is luminosity match measure 'LM' is calculated for each window in both the images. 'LM' is average of luminosity scale calculated by (1) and (2) in a window, for all the pixels.
Then we calculate texture features like energy, entropy, homogeneity, contrast and correlation based on correlation matrix for each of the windows in both images based on following formulae [6]:

$$S = -c\sum_j P[i,j] \log P[i,j], \quad (3.1)$$
$$J = -\sum_i \sum_j P^2[i,j], \quad (3.2)$$
$$Con = -\sum_i \sum_j (i-j)^2 P[i,j], \quad (3.3)$$
$$H = -\sum_i \sum_j P[i,j]/(1+|i-j|), \quad (3.4)$$
$$Cor = -[\sum_i \sum_j ijP[i,j] - \mu_\iota \mu_\varphi] / \sigma_\iota \sigma_j, \quad (3.5)$$

Where S is entropy, J is energy, Con is contrast, H is homogeneity, Cor is correlation. P[i,j] is the value of element in co-occurrence matrix and [i,j] is the coordinate, μ_i and μ_j are mean of p_i and p_j, respectively, where [6] :

$$p_i = -\sum_j P[i,j], \quad (4.1)$$
$$p_j = -\sum_i P[i,j], \quad (4.2)$$

Now, these texture features are weighted appropriately, weights are calculated for each of these features based on the formula [7]:

$$W(i, j) = e^{-d(x_i - x_j)/2} \quad (5.1)$$

Where ω is the parameter under consideration and

$$d(x_i - x_j) = \| x_i - x_j \|^2, \quad (5.2)$$

Then a texture similarity measure 'TM' based on Euclidean distance is calculated for each window in both source and target images. Now, a similarity measure 'E' is computed for each of these windows which give equal weight to 'TM' and 'LM'. Based on 'E' each window in target gray scale image is compared to every window of source colored image and the best match is found out.

Now we select some data points randomly from the source image window using jitter sampling and the chromatic values of these pixels in source image window are transferred to the pixels of target image window, based on the mean and standard deviation calculations for luminance mapping, but their luminosity values are retained.

We intend here to transfer one image's appearance to another; hence it's possible to select source and target images that may not work well together, but the result's quality depends on the images' similarity in composition.We have also used the conventional method for colorization in *lαβ* space and compared the resulting images obtained with those obtained in Lαβ method based on time required for processing, Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Normalized Cross correlation (NK), Structural Content (SC), Maximum Difference (MD), Normalized Absolute Error (NAE), Visual Quality with respect to the target gray scale image. This calculation is based on following formulae [8]:

$$MSE = [\sum_{j=1}^{M}\sum_{k=1}^{N}(x_{j,k} - x'_{j,k})^2] / MN, \quad (6.1)$$

$$PSNR = 10\log[(2^n - 1)^2 / MSE], \quad (6.2)$$

$$NK = \sum_{j=1}^{M}\sum_{k=1}^{N}(x_{j,k} .x'_{j,k}) / \sum_{j=1}^{M}\sum_{k=1}^{N}(x_{j,k})^2, \quad (6.3)$$

$$SC = \sum_{j=1}^{M}\sum_{k=1}^{N}(x_{j,k})^2 / \sum_{j=1}^{M}\sum_{k=1}^{N}(x'_{j,k})^2, \quad (6.4)$$

$$MD = Max(|x_{j,k} - x'_{j,k}|), \quad (6.5)$$

$$NAE = \sum_{j=1}^{M}\sum_{k=1}^{N}|x_{j,k} - x'_{j,k}| / \sum_{j=1}^{M}\sum_{k=1}^{N}|x_{j,k}|, \quad (6.6)$$

The results are tabulated in Table 1and2. Apart from this to show the significance of choosing the right color space, we compared four different color spaces. These color spaces are RGB as used by most graphics algorithms, CIECAM97s, and *lαβ* which is very popular decorrelated color space and which we have used in our method. The reason behind choosing the CIECAM97s color space is that it closely relates to lαβ color space. It has three channels A, C1 and C2 [9][10]. Its chromatic channels C1 and C2 resemble the chromatic channels in *lαβ* space; accept a scaling of the axes. The achromatic channel A is different). It is also important to note here that the CIECAM97s operates in linear space, and *lαβ* are defined in log space. Using image in Figure 1, we have produced scattered plots of 2,000 data points, which were chosen randomly [9][10] in *lαβ*, RGB, and CIECAM97s. Figure 2 depicts these scatter plots that show the pairs of axes (any two axes out of three axes) plotted against each other. The data points in sample image are decorrelated if the data are aligned to the axis, as can be seen for all three pairs of axes in *lαβ* and CIECAM97s spaces where as RGB color space exhibits correlation to a large extent between all pairs of axes as the data cluster around a 45-degree slope line, between the two axes. This validates for *lαβ* color space because different input data (randomly selected), has been used.

3. Result

Our colorization method produces believable colorized output images, as is obvious from the results in Figure (3 –5). We have presented a comparison between results obtained by colorization in *lαβ* color space using our method and by conventional colorization technique in *lαβ* color space in Table (1-2). The results are compared with the original target image based on various parameters like Time Taken, Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Normalized Correlation (NK), Maximum Difference (MD), Normalized Absolute Error (NAE) and Visual Quality Obtained. As is apparent from this comparison our proposed colorization method in *lαβ* color space produces much better results. Images of various dimensions ranging from 256X256 to 1024x1024 were tested and the method proved successful, the time taken for colorization ranges from 25 seconds to 275 seconds.

Figure. 1 Image Analyzed

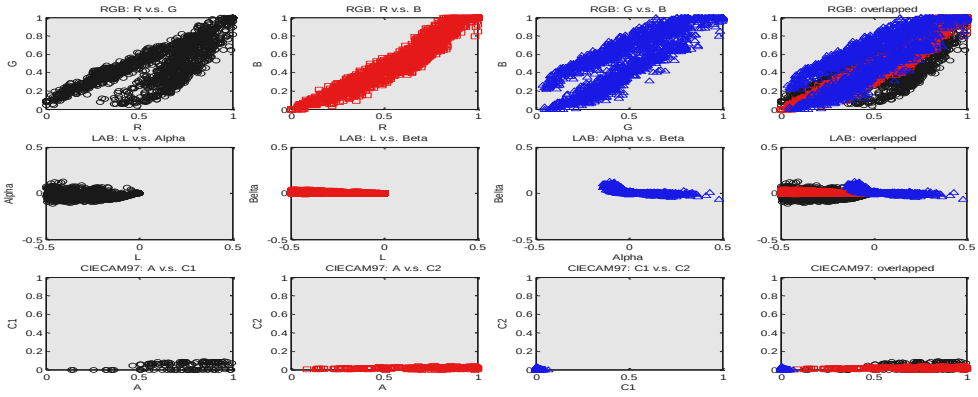


Figure. 2 Comparision between various color spaces

Figure. 3(a) Gray Scale Target Image

Figure. 3(b) Colored Source Image

Figure. 3(c) Target Image Colored by Conventional Method

Figure. 3(d) Target Image Colored by Proposed Method

Figure. 4(a) Gray Scale Target Image

Figure. 4(b) Colored Source Image

Figure. 4(c) Target Image Colored by Conventional Method

Figure. 4(d) Target Image Colored by Proposed Method

Table 1. Comparison between results of Figure 3.

DIFFERENT METHODS ► / PARAMETERS ▼	PROPOSED TECHNIQUE	CONVENTIONAL TECHNIQUE
Time taken (sec)	11.5790	68.3276
Mean square error	309.7544	1315.7
Peak signal to noise ratio	23.2206	15.2509
Normalized cross correlation	1.2155	0.8782
Structural content	1.2927	0.6498
Maximum difference	19	101
Normalized absolute error	0.1466	0.3362
Visual quality	Good	Average

Table 2. Comparisons between results of Figure 4.

DIFFERENT METHODS ► / PARAMETERS ▼	PROPOSED TECHNIQUE	CONVENTIONAL TECHNIQUE
Time taken (sec)	12.25	94.49
Mean square error	348.8706	3079.8
Peak signal to noise ratio	22.7042	13.2455
Normalized cross correlation	1.4651	0.8380
Structural content	1.4215	0.4603
Maximum difference	19	68
Normalized absolute error	0.1693	0.4966
Visual quality	Good	Average

Figure. 5(a) Gray Scale Target Image

Figure. 5(b) Colored Source Image

Figure. 5(c) Target Image Colored by Proposed Method

Conclusion

In this work we have successfully colorized gray scale images of various sizes in *lαβ* color space. Here we take a reference colored image or source image which is similar in texture to the target gray scale image, it is not necessary that this image is colorized version of the target gray scale image, it can be any similar colored image. The quality of output depends on the choice of this reference or source image. We convert both the target and source images in *lαβ* decorrelated color space, to avoid cross channel artifacts, and divide these images in windows of equal size. Then for each of these windows a luminosity measure based on mean and standard deviation is calculated. Again texture features like energy, entropy, contrast, homogeneity and correlation based on correlation matrix are extracted for each of these features are weighted properly and then for each window a texture similarity measure is calculated. After this we have calculated a similarity measure that gives equal weight to both luminosity measure as well as texture similarity measure for each window. After this each window in target image is compare with every window in source image and the best matching window is found out. Then we select some data points randomly from the source image window using jitter sampling and the chromatic values of these pixels in source image window are transferred to the pixels of target image window, based on the mean and standard deviation calculations for luminance mapping, but their luminosity values are retained and colorized version of target image is obtained. We compare our results with the output obtained by colorization in *lαβ* color space and come to the conclusion that the results obtained by colorization in *lαβ* color space using our proposed technique are of better quality visually as well as on the basis of image quality parameters like Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Normalized Correlation (NK), Maximum Difference (MD), Normalized Absolute Error (NAE) etc.

References

[1] Semary N. A., "Texture recognition techniques for natural gray images", Radio Science Conférence, 2007. NRSC April, 2007.

[2] Gonzalez R. C., Woods R. E., and Eddins, S. L., "Digital Image Processing Using Matlab." Pearson Education New Delhi (India), Edition No. 1 (Reprint), 2004.

[3] Solomon C., Breckon T., "Fundamentals of Digital Image Processing.", Wiley – Blackwel , UK, Edition No. 1, 2011.

[4] Daniel L. Ruderman, Thomas W. Cronin and Chuan-Chin Chiao, "Statistics of cone responses to natural images: implications for visual coding", J. Opt. Soc. Am. A/ Vol. 15, pp. 8, August, 1998.

[5] D.R.J. Laming, Sensory Analysis, Academic Press, London, 1986.

[6] Shaoyuan Sun, Jing Z., Liu G., Li Z., "Transfer Color to Night Vision Images.", Chinese Optics Letters, Vol.3.(8), August, 2005.

[7] Weiss, Y., "Segmentation using eigenvectors: A unifying view", In Proc. IEEE Intl. Conf. Computer Vision, Vol. 2, Corfu, Greece, pp. 975–982, 1999

[8] http://www.mathworks.fr/matlabcentral/Image Quality

[9] Smriti Kumar, Ayush Swarnkar, "Colorization of Gray Scale Images in *lαβ* Color Space using Mean and Standard Deviation" IEEE SCEECS 2012, MANIT, Bhopal (India), IEEE Conference Publications, March 2012.

[10] Welsh, T., Ashikhmin, M. and Mueller, K., "Transferring color to grey scale images", In: Proceedings of 29th Annual Conference on Computer Graphics and Interactive Techniques, ACM-SIGGRAPH2002, pp. 277- 280, ACM Press, New York, USA, 2002.

Weighted Analysis on Evaluation Criteria of the Most Advantageous Bid

Han-Chen Huang
Department of Leisure Management, Yu Da University, Taiwan.

Abstract:
In procurement operations, if the lowest bid proposed by a vendor is accepted as the awarding criterion that is less than the base price, it is not uncommon that the bidder may first offer an unreasonably low price to gain the right to supply, and then provide low-quality products in the future. Regardless of whether this fraudulent behavior can be identified, the company or agency inviting the bid is consequently impeded from receiving products that meet their needs. To address this issue, the most advantageous bid (MAB) method can be adopted as an alternative awarding criterion. However, when practicing MAB, the weighting of evaluation criteria and sub-criteria frequently presents a challenge for companies or government agencies offering an invitation for bid. Based on extant literature on supplier evaluation theories, this study conducts interviews with experts to determine evaluation criteria and sub-criteria for MAB, and analyzes the weights of evaluation criteria and sub-criteria. A fuzzy analytic hierarchy process (FAHP) is performed to analyze the obtained MAB evaluation criteria and sub-criteria. The results of the study provide a reference for any company or government agency seeking to evaluate MAB.

Keywords: Lowest Bid, Most Advantageous Bid, Fuzzy Analytic Hierarchy Process.

1. Introduction

Numerous companies or government agencies that engage in procurement activities adopt the lowest bid qualifying for the minimum specifications required as the awarding criterion. This approach frequently causes overly low-priced bidding and impedes tenderees from receiving products that meet their requirements. A solution to this issue is to adopt the most advantageous bid (MAB) system [1,2], which enables tenderees to select the optimal bidder by comprehensively evaluating its technology, quality, functionality, commercial terms, and prices based on pre-specified evaluation criteria [3].

The biggest difference between MAB and the lowest price bid approach is that MAB is awarded through "establishing an evaluation committee to rate each of the evaluation criteria and sub-criteria, so as to select suppliers that are most suitable for the purchaser's needs or most favorable to the purchaser." Therefore, the establishment of an impartial evaluation committee and an appropriate weighting of the selection criteria and sub-criteria are critical preliminary tasks [4]. When adopting the MAB method, the absence of appropriate evaluation criteria and sub-criteria weights may result in incorrect evaluation results. Therefore, this study reviews extant research on supplier evaluation theories [4-9] and conducts interviews with experts to determine suitable evaluation criteria and sub-criteria for MAB, and analyzes the weights of the determined evaluation criteria and sub-criteria using a fuzzy analytic hierarchy process (FAHP) [10-13]. The results of the study provide a reference for any company or government agency seeking to evaluate MAB.

2. Research Results

The purpose of this study is to construct an evaluation model for MAB. After reviewing extant literature [4-9], the preliminarily evaluation criteria and sub-criteria are determined. The Delphi method is applied to establish the hierarchy for the evaluation model that comprises three levels. The first level is the target level; the second level contains the evaluation criteria, comprising seven items (technological capabilities, management system, prices and costs, cooperative abilities, delivery and warranty, performance of prior contract fulfillment, and quality control abilities); and the third level is sub-criteria comprising the 29 indicators shown in Fig. 1. The questionnaire and telephone interviews were conducted with people who have been responsible for MAB operations in private Taiwanese companies and government agencies. Fifty-one responses were collected, among which 41 are considered valid and can be used to calculate the weights of the determined evaluation criteria and sub-criteria.

The weights and sorting of the evaluation criteria and sub-criteria are shown in Table 1. The evaluation criteria are sorted in descending order of significance (i.e., weight) as follows: "Quality Control Capabilities", "Performance of Prior Contract Fulfillment", "Price and Cost", "Technological Capabilities", "Delivery and Warranty", "Management System", and "Cooperative Abilities". The first five evaluation criteria account for 90.1% of the total weight, indicating that product quality, delivery and warranty, price, production technology, and performance of prior contract fulfillment receive the most attention. The weights of the "management system of the contractor and the cooperative abilities combined account for only 9.9%, indicating that these two categories receive only minor attention.

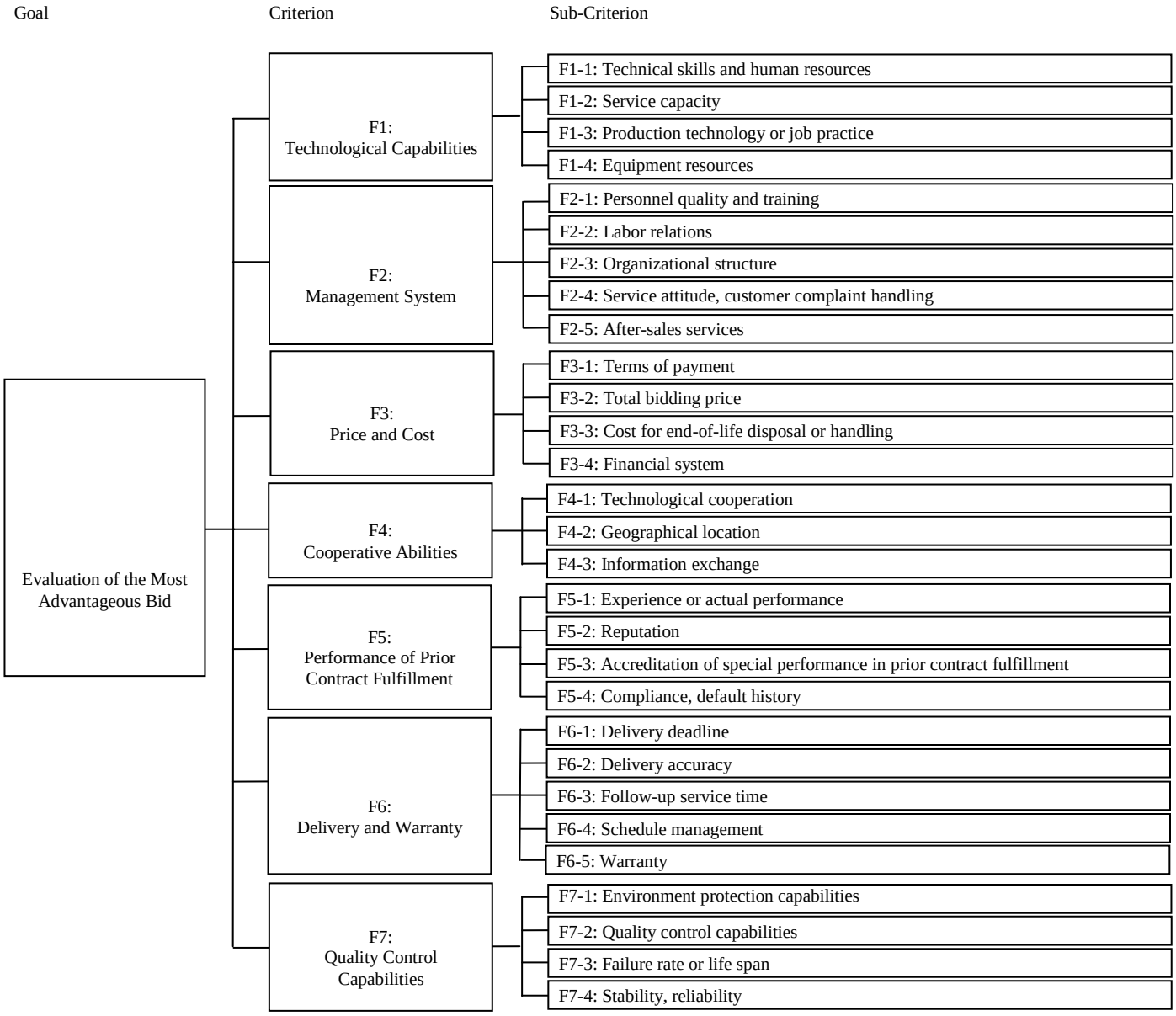


Figure 1. Hierarchy of the proposed MAB evaluation model

Among all evaluation criteria, the weight of "Quality Control Capabilities" accounts for 26.57% of this evaluation, and that of the "Price and Cost" accounts for 18.65%, indicating that the MAB procurement method can shift the award of contracts from the lowest price oriented approach to a quality oriented approach, in which the price proposal is considered as secondary. This enables purchasers to buy products with the best possible quality within the allocated budget, thereby making optimal use of available funds. To sort the sub-criteria in descending order of their significance (i.e., weight), the first five items are: "Stability, reliability", "Failure rate or life span", "Total bidding price", "Delivery accuracy", and "Terms of payment", and the final five items (in ascending weight order) are: "Organizational structure", "Personnel quality and training", "Labor relations", "Information exchange", and "Geographical location".

Table 1. Local weight and global weight for each criterion

Criterion [A]	Local Weights[B]	Ranking	Sub-Criterion[A]	Local Weights[B]	Ranking	Global Weights[C]	Ranking
F1	0.1387	4	F1-1	0.2393	3	0.03319	14
			F1-2	0.2931	2	0.04065	8
			F1-3	0.3120	1	0.04327	7
			F1-4	0.1556	4	0.02158	18
F2	0.0552	6	F2-1	0.1116	4	0.00616	28
			F2-2	0.1184	3	0.00654	27
			F2-3	0.1084	5	0.00598	29
			F2-4	0.3261	2	0.01800	22
			F2-5	0.3355	1	0.01852	21
F3	0.1865	3	F3-1	0.2654	2	0.04950	5
			F3-2	0.4461	1	0.08320	3
			F3-3	0.1735	3	0.03236	15
			F3-4	0.1150	4	0.02145	19
F4	0.0438	7	F4-1	0.3915	1	0.01714	23
			F4-2	0.3298	2	0.01444	25
			F4-3	0.2787	3	0.01220	26
F5	0.0987	5	F5-1	0.4008	1	0.03957	10
			F5-2	0.1949	3	0.01924	20
			F5-3	0.2376	2	0.02346	17
			F5-4	0.1667	4	0.01646	24
F6	0.2114	2	F6-1	0.3279	1	0.06932	4
			F6-2	0.1472	5	0.03112	16
			F6-3	0.1715	3	0.03623	12
			F6-4	0.1881	2	0.03976	9
			F6-5	0.1723	3	0.04578	6
F7	0.2657	1	F7-1	0.1477	4	0.03924	11
			F7-2	0.3254	2	0.08646	2
			F7-3	0.3546	1	0.09422	1
			F7-4	0.2393	3	0.03319	14

A. For An Explanation of the Codes, Please Refer to Fig. 1.
B. Local Weight is Determined based on Judgments of a Single Criterion.
C. Global Weight is Determined by Multiplying the Weight of the Criteria.

3. Conclusions

Based on supplier selection theory, this study conducted interviews with experts to formulate evaluation criteria and sub-criteria for the MAB approach. The obtained seven evaluation criteria and 29 sub-criteria can provide a reference for companies or government agencies engaged in processing the MAB. This study applied FAHP to determine the weights of evaluation criteria and sub-criteria of the MAB approach. The results show that the weights of the evaluation criteria are as follows: "Quality Control Capabilities"(0.2657), "Delivery and Warranty"(0.2114), "Price and Cost"(0.1865), "Technological Capabilities"(0.1387), "Performance in Prior Contract Fulfillment"(0.0987), "Management System"(0.0552), and "Cooperative Abilities"(0.0438). The first five items account for 90.1% of the total weight, whereas the combined weight for "Management System" and "Cooperative Abilities" only accounts for 9.9%. Enterprises or government agencies that need to formulate their evaluation criteria and sub-criteria for MAB may refer to the results of this study. Alternatively, they may apply the method proposed by this study to construct their own MAB evaluation model to improve the accuracy of supplier selection.

References

[1] J. Y. Wang and L. K. Chen. Most Advantageous Tender of the Law of Government Procurement is Applying to School Procurement. School Administration, 65(1): 155-164, 2010.

[2] X. H. Chen, C. L. Pan, and Y. F. Huang. The Integration of AHP and ELECTRE in the Selection Process for the Most Advantageous Bid. Journal of Commercial Modernization, 4(3): 99-119, 2008.

[3] H. P. Tserng, W. K. Teng, and S. J. Chen. The Study of Evaluation Criteria for Selecting Design-Build Contractor of Infrastructure Project. Journal of the Chinese Institute of Civil and Hydraulic Engineering, 20(3): 415-426, 2008.

[4] H. Y. Tsai and G. Lee. Role Analysis of Procurement Evaluation Committee Member. Journal of Building and Construction Technology, 7(1): 91-100, 2010.

[5] F. S. Liou, R. J. Dzeng, and S. S. Wang. A Study on Contractor's Attributes and Construction Performance for Public Building Procurement Using the Most Advantageous Bid System. Chung Hua Journal of Architecture, 1(1): 3-14, 2005.

[6] G. W. Dickson. An Analysis of Vendor Selection System and Decisions. Journal of Purchasing, 2(1): 5-17, 1966.

[7] T. Al-Faraj, A. Alidi, and J. Al-Zayer. Vendors Selection via a Spreadsheet Analytical Hierarchy Process. Computers and Industrial Engineering, 25(1): 65- 68, 1993.

[8] Jitendra Kumar and Nirjhar Roy. Analytic Hierarchy Process for a Power Transmission Industry to Vendor Selection Decisions. International Journal of Computer Applications, 12(11): 26-30, 2011.

[9] Tanmoy Chakraborty, Tamal Ghosh, and Pranab K Dan. Application of Analytic Hierarchy Process and Heuristic Algorithm in Solving Vendor Selection Problem. Business Intelligence Journal, 4(1): 167-178, 2011.

[10] T. L. Saaty. The Analytic Hierarchy Process: Planning, Priority Setting and Resource Allocation. McGraw-Hill, New York, 1980.

[11] L. A. Zadeh. Fuzzy Set. Information and Control, 8(1): 338-353, 1965.

[12] T. S. Li and B. Y. Huang. The Study of Applying Fuzzy Analytic Hierarchy Process in Supplier Selection. Journal of Quantitative Management, 5(2): 39-56, 2008.

[13] H. C. Huang. Decision-Making Model for Convention Site Selection. Advanced Materials Research, 538-541: 895-900, 2012.

Delaying Transmissions in Data Communication Networks to Improve Transport-Layer Performance

[1]Pangambam Uttam Kumar, [2]Jyoti Gupta

[1, 2] Department of Electronics and Communication Engineering, M.M. University, Mullana, Ambala, India

Abstract

The performance improvement of the network with small buffers needs to be focused. The main factor that reduces the performance of the network is the packet loss in the network layer in the event of data transmission. Packet loss occurred due to bit errors in physical layer and congestion in network layer degrades the performance in the transport layer, in terms of delay and bandwidth. Existing approaches improving the performance of the transport layer requires either much delay or bandwidth. So, Queue Length Based Pacing (QLBP) is presented, which is a novel pacing algorithm that decreases the burstiness of network traffic by delaying packets based on the length of the local packet buffer. Stability control mechanism is applied as contribution which further reduces the packet loss. Simulation is performed over the network of dumbbell topology with QLBP technique using Network simulator 2 (NS 2). Performance metric such as throughput, delay, packets drops over buffer size is analyzed.

Keywords: Dumbbell Topology, Packet Drop, Propagation Delay, Small Buffer Network, TCP, Throughput, Traffic Pacing and Transport Layer.

"1.Introduction"

TCP mechanisms

TCP is a sliding window-based protocol. The effective window used by a TCP sender is the minimum of the congestion window (cwnd – set to match the bandwidth of the network) and the receiver's buffer size. Since the window is the number of bytes the sender is allowed to send without an acknowledgment, the average rate at which traffic enters the network is governed by the window size divided by the round trip time (RTT). The sender uses the incoming acknowledgments to determine when to send new data, a mechanism referred to as ACK-clocking. The congestion window adjustment algorithm has two phases. In the slow start phase, the sender increases the congestion window rapidly in order to quickly identify the bottleneck rate while at the same time theoretically establishing a stream of well-spaced acknowledgments. The sender typically starts with a window of one packet; each acknowledgment increases the window by 1, effectively doubling the window every round trip time. Assuming the sender is not limited by the receiver's buffer space, the sender increases its congestion window until it detects that a packet loss has occurred; the loss is taken as a signal that the sender is transmitting packets faster than the network can handle. At this point the window is cut in half, and the sender enters the congestion avoidance phase. The sender then increases the window by 1/cwnd on every acknowledgment, effectively increasing it by 1 packet every round trip time. Again, the sender increases the window until it detects a packet loss; on a loss, the window is cut by half, and the sender resumes increasing the window by 1 packet per round trip. As a result, the sender's congestion window (controlling the rate at which packets are sent) at first increases along an exponential curve during slow start, then over time repeats a saw-tooth pattern of a factor of two decrease and a subsequent slow, linear increase. Since the receipt of acknowledgments governs the rate of increase in the congestion window, connections with longer round trip times have windows that grow at a slower rate (both slow start and the saw-tooth have a proportionately lower slope).

Each TCP acknowledgment reports the highest-numbered packet (more precisely, byte) that had been received along with all earlier packets. Assuming there are no losses, the information in each acknowledgment therefore makes all earlier acknowledgments redundant. As an optimization, TCP receivers try to reduce the number of acknowledgments by delaying returning an acknowledgment for a single received packet (assuming all prior packets had been correctly received) until either another packet arrives or a timer fires. Because half as many acknowledgments are received, and since acknowledgments trigger increases in the congestion window, the principal effect of delayed acknowledgments is to slow the rate at the congestion window is increased: during slow start from a factor of 2 to a factor of 1.5 per round trip and during congestion avoidance from 1 packet to 0.5 packets per round trip [21]. During slow start, every successfully acknowledged packet increases the window size by one packet. Thus, the sender transmits two packets for every new acknowledgment. Queuing theorists are used to thinking of sizing buffers so as to prevent them from overflowing and losing packets. But TCP's "saw

tooth" congestion control algorithm is designed to fill any buffer, and deliberately causes occasional loss to provide feedback to the sender. No matter how big we make the buffers at a bottleneck link, TCP will cause the buffer to overflow.

Goal of pacing

The goal of pacing is to evenly spread the transmission of a window of packets across the entire duration of the round trip time [21]. This can be implemented either by the sender or the receiver. At the sender, instead of transmitting packets immediately upon receipt of an acknowledgment, the sender can delay transmitting packets to spread them out at the rate defined by the congestion control algorithm – the window size divided by the estimated round-trip time. Alternatively, a receiver can delay acknowledgments to spread them across the round trip time, so that when they arrive at the sender, they will trigger spaced data packets. Of course, receiver pacing is less effective, since as we discussed earlier, acknowledgments arriving at the sender can trigger multiple data sends; with receiver pacing, these packets will be sent in a burst. Further, receiver pacing is susceptible to ACK compression. Therefore, only sender-based pacing is simulated. As a traditional window based protocol, TCP uses a window to determine the number of packets that can be sent and uses the receipt of acknowledgments to trigger the sending of packets [10]. Pure rate based schemes, on the other hand, use rates to determine both how much and when to send. Pacing is a hybrid between these two approaches – it uses the TCP window to determine how much to send but uses rates instead of acknowledgments to determine when to send.

Uses of pacing

One way to understand the impact of pacing is to consider the router from a queuing theory perspective. With bursty traffic, packets arrive all at once. As a result, queuing delay grows linearly with load, even when the load is below capacity. TCP, due to its burstiness, can be thought of as being close to this worst case curve. With pacing, traffic is evenly spaced out; so there is minimal queuing until the load matches the bottleneck capacity [29]. The queuing delay increases linearly once the bottleneck is saturated. This represents the best possible time curve for a queuing system and is theoretically even better than the curve for a random traffic source. TCP uses feedback from the network to detect congestion and adjust to it. With pacing, this feedback is delayed until the network is saturated, making it difficult for senders to "avoid" overwhelming the network.

Queue length based pacing algorithm

The general ideal of Queue Length Based Pacing (QLBP) is to dynamically adjust the sending rate of a queue according to the queue length, rather than to send packets at a constant rate [8]. Traffic pacing is based on TCP, but uses traffic conditioning techniques in the network to reduce traffic bursts. By delaying some packet transmissions, less packet losses occur and thus less retransmission are needed. Traffic pacing incurs a small additional delay, but uses less bandwidth than TCP since fewer retransmissions are necessary. Pacing is deployed on several (but not necessarily all) nodes in the network. This pacing process can be implemented on the outgoing interfaces of routers. These routers have sufficiently large buffers that allow moderate traffic bursts to be absorbed and paced without packet loss. At the network edge, routers with pacing capabilities reduce the burstiness of traffic before it enters the small-buffer network core. The inter-packet pacing delay is calculated based on the packet arrival curve and the packet deadline curve within the same pacing interval. QLBP, determine this delay based on some very simple rules: (a) if the pacing queue increases due to a higher input traffic rate, QLBP intentionally lowers the introduced pacing delay. This rule ensures that the link can be fully utilized under heavy load. (b) For packets that arrive at a rate lower than μmin, they do not get delayed. This rule ensures that pacing is only activated when packets arrive at a certain high rate.

Problems

One of the most problematic events for data transmissions in the network layer is a packet loss. The two main causes for packet loss in networks are:

• Bit errors in physical layer: Bit errors in the physical layer most commonly occur in wireless transmissions due to interference, but can also occur in wired links. These bit errors causes checksums in the data link layer to fail, triggering a packet drop.

• Congestion in network layer: Statistical multiplexing of network traffic implies that there are no guarantees about the available bandwidth on any given link. Thus, network traffic can congest the outgoing port of a router and cause transmission buffers to fill up. If a packet arrives at such a transmission queue when no more buffer space is available, then it is dropped. While these causes of packet loss are fundamentally different, their effects result in the same performance degradation in the transport layer.

"2. Literature Survey"

A. Vishwanath et al. [1] proposed that internet traffic is expected to grow phenomenally over the next five to ten years, and to cope with such large traffic volumes, core networks are expected to scale to capacities of terabits-per-second and beyond. Increasing the role of optics for switching and transmission inside the core network seems to be the most promising way forward to accomplish this capacity scaling. Unfortunately, unlike electronic memory, it remains a formidable challenge to build even a few packets of integrated all-optical buffers. In the context of envisioning a bufferless (or near-zero buffer) core network, their contributions are three fold: First, a novel edge-to-edge based packet-level forward error correction (FEC) framework as a means of combating high core losses was proposed and investigated via analysis and simulation the appropriate FEC strength for a single core link. Secondly, considering a realistic multi-hop network and develop an optimization framework that adjusts the FEC strength on a per-flow basis to ensure fairness between single and multi-hop flows. Third, studying the efficacy of FEC for various system parameters such as relative mixes of short-lived and long-lived TCP flows, and average offered link loads. Their study is the first to show that packet-level FEC, when tuned properly, can be very effective in mitigating high core losses, thus opening the doors to a buffer less core network in the future.

M. Shifrin et al. [2] proposed that because of TCP dynamics, Internet backbone routers hold large packet buffers, which significantly increase their power consumption and design time. Recent models of large-buffer networks have suggested that these large buffers could be replaced with much smaller ones. Unfortunately, it turns out that these large-buffer network models are not valid anymore in small-buffer networks, and therefore cannot predict how these small-buffer networks will behave. In this paper, a new model that provides a complete statistical description of small-buffer Internet networks was introduced. Novel models of the distributions of several network components, such as the line occupancies of each flow, the instantaneous arrival rates to the bottleneck queues, and the bottleneck queue sizes was presented. Later, all these models are combined in a single fixed-point algorithm that forms the key to a global statistical small-buffer network model. In particular, given some QoS requirements, also showed how this new model can be used to precisely size small buffers in backbone router designs.

M. Aron et al. [3] proposed that for the optical packet-switching routers to be widely deployed in the Internet, the size of packet buffers on routers has to be significantly small. Such small-buffer networks rely on traffic with low levels of burstiness to avoid buffer overflows and packet losses. The proposed work presented a pacing system that proactively shapes traffic in the edge network to reduce burstiness. The queue length based pacing uses an adaptive pacing on a single queue and paces traffic indiscriminately where deployed. In this work, they showed through analysis and simulation that this pacing approach introduces a bounded delay and that it effectively reduces traffic burstiness. The work also showed that it can achieve higher throughput than end-system based pacing.

A. Lakshmikantha et al. [4] suggested that traditionally, it had been assumed that the efficiency requirements of TCP dictate that the buffer size at the router must be of the order of the bandwidth-delay ($C \times$ RTT) product. Recently, this assumption was questioned in a number of papers, and the rule was shown to be conservative for certain traffic models. In particular, by appealing to statistical multiplexing, it was shown that on a router with N long-lived connections, buffers of size $O(C \times \text{RTT}/\sqrt{N})$ or even O (1) are sufficient. In this paper, the buffer-size requirements of core routers when flows arrive and depart are examined. The conclusion is as follows: If the core-to-access-speed ratio is large, then O (1) buffers are sufficient at the core routers; otherwise, larger buffer sizes do improve the flow-level performance of the users. From a modeling point of view, our analysis offers two new insights. First, it may not be appropriate to derive buffer-sizing rules by studying a network with a fixed number of users. In fact, depending upon the core-to-access-speed ratio, the buffer size itself may affect the number of flows in the system, so these two parameters (buffer size and number of flows in the system) should not be treated as independent quantities. Second, in the regime where the core-to-access-speed ratio is large, we note that the O (1) buffer sizes are sufficient for good performance and that no loss of utilization results, as previously believed.

Y. Huang et al. [6] proposed that Queuing analysis is important in providing guiding principles for packet network analysis. Stochastic fluid queuing models have been widely used as burst scale models for high speed communication networks. In this paper, the authors proposed a novel two-level Markov on-off source model to model the burstiness of a packet stream at different time scales. Analytical results are obtained to reveal the impact of traffic burstiness at two levels on the queue lengths in a tandem queue system. The method combines the modeling power of the Poisson processes with that of stochastic differential equations to handle the complex interactions between the packet arrivals and the queue content. Results for the tandem queuing network could be used to further justify the packet spacing scheme in helping deploying small buffer routers.

Y. Gu et al.[7] proposed that there is growing interest in designing high speed routers with small buffers that store only tens of packets. Recent studies suggest that TCP New Reno, with the addition of a pacing mechanism, can interact with such routers without sacrificing link utilization. The work showed in this paper, as workload requirements grow and connection bandwidths increase, the interaction between the congestion control protocol and small buffer routers produce link utilizations that tend to zero. This is a simple consequence of the inverse square root dependence of TCP throughput on loss probability. In this paper they presented a new congestion controller that avoids this problem by allowing a TCP connection to achieve arbitrarily large bandwidths without demanding the loss probability go to zero. Authors showed that this controller produces stable behaviour and, through simulation, they showed its performance to be superior to TCP New Reno in a variety of environments. Lastly, because of its advantages in high bandwidth environments, they compared their controller's performance to some of the recently proposed high performance versions of TCP including HSTCP, STCP, and FAST. Simulations illustrate the superior performance of the proposed controller in a small buffer environment.

M. C. Weigle et al. [8] proposed that when doing simulations, authors were confronted with the problem of choosing a good workload model. Often, realistic workload models are difficult to come up with. So, this paper proposed a tool, called Tmix, that allows to automatically extract communication workloads from packet traces and then to replay those workloads in the ns simulator or in test beds.

V. Sivaraman et al.[9] proposed that in the absence of a cost-effective technology for storing optical signals, emerging optical packet switched (OPS) networks are expected to have severely limited buffering capability. To mitigate the performance degradation resulting from small buffers, this paper proposes that optical edge nodes "pace" the injection of traffic into the OPS core. The contributions relating to pacing in OPS networks are three-fold: first, real-time pacing algorithms of poly-logarithmic complexity that are feasible for practical implementation in emerging high-speed OPS networks was developed. Second, an analytical quantification of the benefits of pacing in reducing traffic burstiness and traffic loss at a link with very small buffers was proposed. Third, authors showed via simulations of realistic network topologies that pacing can significantly reduce network losses at the expense of a small and bounded increase in end-to-end delay for real-time traffic flows. It was argued that the loss-delay trade off mechanism provided by pacing can be instrumental in overcoming the performance hurdle arising from the scarcity of buffers in OPS networks.

M. Enachescu et al. [10] proposed that internet routers require buffers to hold packets during times of congestion. The buffers need to be fast, and so ideally they should be small enough to use fast memory technologies such as SRAM or all-optical buffering. Unfortunately, a widely used rule-of-thumb says we need a bandwidth-delay product of buffering at each router so as not to lose link utilization. In this paper, they explored how buffers in the backbone can be significantly reduced even more, to as little as a few dozen packets, if they are willing to sacrifice a small amount of link capacity. Authors argued that if the TCP sources are not overly bursty, then fewer than twenty packet buffers are sufficient for high throughput. Specifically, authors argued that O (logW) buffers are sufficient, where W is the window size of each flow. The proposed worked supported the claim with analysis and a variety of simulations. The change they need to make to TCP is minimal-each sender just needs to pace packet injections from its window. Moreover, there is some evidence that such small buffers are sufficient even if we don't modify the TCP sources so long as the access network is much slower than the backbone, which is true today and likely to remain true in the future. The proposed work concluded that buffers can be made small enough for all-optical routers with small integrated optical buffers.

J. Naor et al. [11] considered the problem of scheduling a sequence of packets over a linear network, where every packet has a source and a target, as well as a release time and a deadline by which it must arrive at its target. The model considered is bufferless, where packets are not allowed to be buffered in nodes along their paths other than at their source. This model applies to optical networks where opto-electronic conversion is costly, and packets mostly travel through bufferless hops. The offline version of this problem was previously studied in M. Adler et al. (2002). Authors studied the online version of the problem, where they are required to schedule the packets without knowledge of future packet arrivals. Competitive analysis to evaluate the performance of our algorithms was used. The first deterministic online algorithms for several versions of the problem were presented. For the problem of throughput maximization, where all packets have uniform weights, they gave an algorithm with a logarithmic competitive ratio, and present some lower bounds. For other weight functions, algorithms that achieve optimal competitive ratios were showed.

G. Raina et al. [12] described how control theory has been used to address the question of how to size the buffers in core Internet routers. Control theory aims to predict whether the network is stable, i.e. whether TCP flows are desynchronized. If flows are desynchronized then small buffers are sufficient. The theory here shows that small buffers actually promoted resynchronization—a virtuous circle.

J. DeHart et al. [13] proposed that the Open Network Laboratory (ONL) is a remotely accessible network test bed designed to enable networking faculty, students and researchers to conduct experiments using high performance routers and applications. The system is built around a set of extensible, high-performance routers and has a graphical interface that enables users to easily configure and run experiments remotely. ONL's Remote Laboratory Interface (RLI) allows users to easily configure a network topology, configure routes and packet filters in the routers, assign flows or flow aggregates to separate queues with configurable QoS and attach hardware monitoring points to real-time charts. The remote visualization features of the RLI make it easy to directly view the effects of traffic as it moves through a router, allowing the user to gain better insight into system behaviour and create compelling demonstrations. Each port of the router is equipped with an embedded processor that provides a simple environment for software plugins allowing users to extend the system's functionality. This paper described the general facilities and some networking experiments that can be carried out.
M. Adler et al. [15] proposed that the time-constrained packet routing problem is to schedule a set of packets to be transmitted through a multi-node network, where every packet has a source and a destination (as in traditional packet routing problems) as well as a release time and deadline. The objective is to schedule the maximum number of packets subject to deadline constraints. In this paper authors extended the results in two directions. First, the more general network topologies of trees and 2-dimensional meshes were considered. Secondly, authors associated with each packet a measure of utility, called a weight, and study the problem of maximizing the total weight of the packets that are scheduled subject to their timing constraints. For the bufferless case, constant factor approximation algorithms for the time-constrained scheduling problem with weighted packets on trees and mashes were provided. Provisions of logarithmic approximations for the same problems in the buffered case were made. These results are complemented by new lower bounds, which demonstrate that they cannot hope to achieve the same results for general network topologies.

C.V.Hollot et al. [16] proposed that routers handle data packets from sources unresponsive to TCP's congestion avoidance feedback. Interest was in the impact these sources have on active queue management (AQM) control of long-lived TCP traffic. In these paper models of TCP/AQM dynamics was combined with models of unresponsive traffic to analyze the effects on AQM performance.

Y. Wu et al. [17] studied a queue where the rapidly varying component of input traffic is treated as a noise. The queue length statistics are compared between the cases with and without noise smoothing. This relates to abstract simulation and traffic modeling. The system load and traffic burstiness at large timescales have critical impacts on the evaluation errors caused by such smoothing.

M. Adler et al. [18] studied the problem of centrally scheduling multiple messages in a linear network, when each message has both a release time and a deadline. The proposed work showed that the problem of transmitting optimally many messages is NP-hard, both when messages may be buffered in transit and when they may not be; for either case, efficient algorithms that produce approximately optimal schedules were presented. In particular, the bufferless scheduling algorithm achieves throughput that is within a factor of two of optimal. It was showed that buffering can improve throughput in general by a logarithmic factor (but no more), but that in several significant special cases, such as when all messages can be released immediately, buffering can help by only a small constant factor. Finally, the proposed work showed how to convert their centralized, offline bufferless schedules to equally productive fully distributed online buffered ones. Most of the results extend readily to ring-structured networks.
R. Ahlswede et al. [20] introduced a new class of problems called network information flow which is inspired by computer network applications. Consider a point-to-point communication network on which a number of information sources are to be multicast to certain sets of destinations. Authors assumed that the information sources are mutually independent. The problem is to characterize the admissible coding rate region. This model subsumes all previously studied models along the same line. In this paper, authors studied the problem with one information source, and obtained a simple characterization of the admissible coding rate region. The result can be regarded as the Max-flow Min-cut Theorem for network information flow. Contrary to one's intuition, their work reveals that it is in general not optimal to regard the information to be multicast as a "fluid" which can simply be routed or replicated. Rather, by employing coding at the nodes, which was referred to as network coding, bandwidth can in general be saved. This finding may have significant impact on future design of switching systems.
R. W. Brockett et al. [22] suggested that recent study for congestion control in high speed networks indicates that the derivative information for the congestion at the common buffer for multiple sources could be useful in achieving efficient and fair allocation of the bandwidth. In this paper they presented an algorithm that estimates such derivatives for multiple on-off sources. The algorithm has its root in the infinitesimal perturbation analysis (IPA) for the classical queuing systems. Although the traditional IPA algorithm does not give unbiased derivative estimates for multi-class arrivals, the proposed work was able

to prove the unbiasedness in the case of multi-class ON-OFF sources. The results in this paper may motivate a new look at the end-to-end congestion control issue.

J. D. Salehi et al. [23] proposed that VBR compressed video is known to exhibit significant, multiple-time-scale bit rate variability. In this paper, the transmission of stored video from a server to a client across a high speed network were considered, and explore how the client buffer space can be used most effectively toward reducing the variability of the transmitted bit rate. Two basic results were presented. First, an optimal smoothing algorithm for achieving the greatest possible reduction in rate variability when transmitting stored video to a client with given buffer size was presented. A formal proof of optimality was presented, and demonstrated the performance of the algorithm on a set of long MPEG-1 encoded video traces. Second, the impact of optimal smoothing on the network resources needed for video transport, under two network service models: Deterministic Guaranteed service and Renegotiated CBR (RCBR) service was evaluated. Under both models, the impact of optimal smoothing was found to be dramatic.

M. Aron et al. [24] studied the start up dynamics of TCP on both high as well as low bandwidth delay network paths and proposed a set of enhancements that improve both the latency as well as throughput of relatively short TCP transfers. Authors also suggested that numerous studies have shown that the timer and congestion control mechanisms in TCP can have a limiting effect on performance in the start up phase. Based on the results of the study, mechanisms for adapting TCP in order to yield increased performance were proposed. First, a framework for the management of timing in TCP was proposed. Second, the work showed how TCP can utilize the proposed timer framework to reduce the overly conservative delay associated with a retransmission timeout. Third, the use of packet pacing in the initial slow-start to improve the performance of relatively short transfers that characterize the web traffic was proposed. Finally, the importance of estimating the initial slow-start threshold in TCP, mainly on high bandwidth-delay paths was quantified.

V. Visweswaraiah et al. [25] proposed that TCP congestion avoidance mechanisms are based on adjustments to the congestion-window size, triggered by the ACK clock. These mechanisms are not well matched to large but intermittent bursts of traffic, such as responses from a HTTP/1.1-based web server. Idle periods between bursts (web page replies) stop the ACK clock and hence disrupt even data flow. When restarting data flow after an idle period, current implementations either enforce slow start (SSR) or use the prior congestion window (NSSR). The former approach, while conservative, leads to low effective throughput in cases like P-HTTP. The latter case optimistically sends a large burst of back-to-back packets, risking router buffer overflow and subsequent packet loss. This paper proposed a third alternative: pacing some packets at a certain rate until the ACK clock can be restarted. The motivation and implementation of this third alternative was described and present simulation results which show that it achieves the elapsed-time performance comparable to NSSR and loss behaviour of SSR.

W. Willinger et al. [26] proposed that a number of empirical studies of traffic measurements from a variety of working packet networks have demonstrated that actual network traffic is self-similar or long-range dependent in nature-in sharp contrast to commonly made traffic modeling assumptions. A plausible physical explanation for the occurrence of self-similarity in local-area network (LAN) traffic was provided. The explanation is based on convergence results for processes that exhibit high variability and is supported by detailed statistical analyzes of real-time traffic measurements from Ethernet LANs at the level of individual sources. Here the mathematical results concerning the superposition of strictly alternating ON/OFF sources was developed. The key mathematical result states that the superposition of many ON/OFF sources (also known as packet-trains) with strictly alternating ON- and OFF-periods and whose ON-periods or OFF-periods exhibit the Noah effect produces aggregate network traffic that exhibits the Joseph effect. There is, moreover, a simple relation between the parameters describing the intensities of the Noah effect (high variability) and the Joseph effect (self-similarity). An extensive statistical analysis of high time-resolution Ethernet LAN traffic traces confirms that the data at the level of individual sources or source-destination pairs are consistent with the Noah effect. Implications of this simple physical explanation for the presence of self-similar traffic patterns in modern high-speed network traffic were discussed.

M. Mathis et al. [27] proposed that TCP may experience poor performance when multiple packets are lost from one window of data. With the limited information available from cumulative acknowledgments, a TCP sender can only learn about single lost packet per round trip time. An aggressive sender could choose to retransmit packets early, but such retransmitted segments may have already been successfully received. A Selective Acknowledgment (SACK) mechanism, combined with a selective repeat retransmission policy, can help to overcome these limitations. The receiving TCP sends back SACK packets to the sender informing the sender of data that has been received. The sender can then retransmit only the missing data segments. This memo proposes an implementation of SACK and discusses its performance and related issues.

I.Nikolaidis et al. [28] evaluated the statistical multiplexing gain in ATM networks for bursty as well as variable bit rate (VBR) traffic using a fluid-flow approximate model. The required bandwidth per source in a finite buffer multiplexer in order to achieve a given Grade of Service (GOS), expressed by the cell loss probability was obtained. For both bursty and VBR traffic sources, they performed a sensitivity analysis of significant parameters in the homogeneous case (all traffic sources are of the same type). The required bandwidth for bursty sources is shown to depend on burst and buffer length through their ratio. Finally, the mixing of bursty traffic and variable bit rate is considered. The results obtained through simulation with approximations proposed in the literature were compared.

L. Zhang et al. [29] used simulation to study the dynamics of the congestion control algorithm embedded in the BSD 4.3-Tahoe TCP implementation. The simple case of a few TCP connections, originating and terminating at the same pair of hosts, using a single bottleneck link were investigated. In this paper the dynamics that results from two-way traffic (in which there are data sources on both hosts) was also investigated. The one-way traffic clustering and loss-synchronization a phenomenon was discussed. In addition, there are two new phenomena: (1) ACK-compression, which is due to the interaction of data and ACK packets and gives rise to rapid fluctuations in queue length, and (2) an out-of-phase queue-synchronization mode which keeps link utilization less than optimal even in the limit of very large buffers. These phenomena are helpful in understanding results from an earlier study of network oscillations.

D. D. Clark et al. [31] proposed that Bulk data transmission is now finding more and more application in various fields. The major performance concern of a bulk data transfer protocol is high throughput. Theoretically, a packet switched network allows any single user an unlimited share of the network resources. In the absence of other traffic, therefore, a user should be able to transmit data at the raw bandwidth of a network channel. In reality, achievable end-to-end throughputs over high bandwidth channels are often an order of magnitude lower than the provided bandwidth. Experience shows that the throughput is often limited by the transport protocol and its flow control mechanism. It is especially difficult to achieve high throughput, reliable data transmissions across long delay, unreliable network paths. In this paper they introduced a new transport protocol, NETBLT, which was designed for high throughput, bulk data transmission applications. They first analyze the impact of network unreliability and delay on the end-to-end transport protocol; they then summarize previous experience; next they show the design and implementation of NETBLT, followed by the initial experience. Generally speaking, errors and variable delays are two barriers to high performance for all transport protocols. The NETBLT design and experience explores general principles for overcoming these barriers.

H. Heffes et al. [32] studied the performance of a statistical multiplexer whose inputs consist of a superposition of packetized voice sources and data. The performance analysis predicts voice packet delay distributions, which usually have a stringent requirement, as well as data packet delay distributions. The superposition is approximated by a correlated Markov modulated Poisson process (MMPP), which is chosen such that several of its statistical characteristics identically match those of the superposition. Matrix analytic methods are then used to evaluate system performance measures. In particular, moments of voice and data delay distributions and queue length distributions were obtained. Laplace-Satieties transform of the voice and data packet delay distributions, which are numerically inverted to evaluate tails of delay distributions, was also obtained. It is shown how the matrix analytic methodology can incorporate practical system considerations such as finite buffers and a class of overload control mechanisms discussed in the literature. Comparisons with simulation show the methods to be accurate. The numerical results for the tails of the voice packet delay distribution show the dramatic effect of traffic variability and correlations on performance.

"3.RELATED WORK"

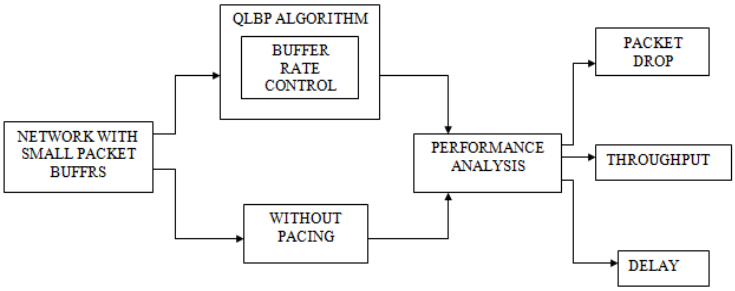
There are several possible approaches to addressing the problem of reducing the impact of packet loss on the delay in transport layer communication. The main techniques noted in this figure are:

• Lossy transmission: Using lossy transmission protocols (e.g., User Datagram Protocol (UDP)) places the bandwidth needs and delay close to the ideal lower bounds. Marginal amounts of additional bandwidth are necessary for packet headers and additional delay is incurred due to the packetized transmission of data.

• Reliable transmission: The baseline protocol for reliable transmission is the Transmission Control Protocol (TCP). Compared to UDP, TCP requires more bandwidth since some packets need to be retransmitted. It also incurs additional delay due to these retransmissions.

• Network coding: There are several coding techniques to reduce packet loss in networks. To reduce bit errors, error correction coding can be used. To avoid packet losses, transmission information can be spread across multiple paths in the network using network coding. These techniques require additional bandwidth since they rely on redundant transmission of information. These also exhibit increased delay over a lossy transmission due to the need for data reconstruction at the receiver. However, these techniques incur less delay than TCP.

In this work, the performance of the network is to be checked and a comparison is made with the proposed technique and the technique of non paced one. Firstly the data packets are allowed to send using the non paced technique and its performance in terms of throughput, packet drop and delays are analyzed. To make a comparison, the data packets are again sent using the proposed technique and various performances in terms of throughput, packet drop and delays are considered. Finally, it is found that the technique of Queue Length Based Pacing yields a better performance compared to the non paced technique, because eventhough some delays are intentionally made, all the data packets are successfully sent and received at the same time.So, there is no point of losing any data while transmitting. The general ideal of Queue Length Based Pacing (QLBP) is to dynamically adjust the sending rate of a queue according to the queue length, rather than to send packets at a constant rate.



"Figure 1. Block diagram of proposed work"

An extensive simulation model having scenario of 14 (user defined) wired nodes connected with duplex link is used to study inter-layer interactions and their performance implications. The other parameters used in this model are as under:

"Table 1. Simulation model specification"

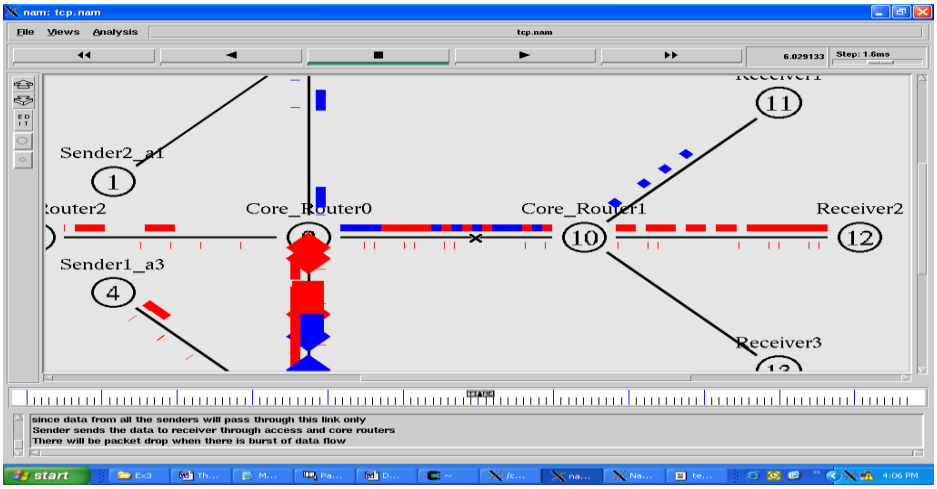
Software for Simulation	Network Simulator 2
Channel	Wired
Simulation run time	30 seconds
Area in which node move	600×600
Packet size	522 bytes
Application type	FTP/CBR
Transport agent	TCP
Link Type	Duplex Link
Queue Type	Drop Tail, FQ
Q- length	50-100Packets (variable)
Routing protocol	rlm
Propagation delay	variable

The type of link used for the proposed model is the Duplex link. The routing protocol used is the rlm (reactive layer multicast) protocol, the protocol used for wired routing. The queue type used in the proposed method is the Drop tail and the Fair Queue (FQ). After selecting the requirements for data transmission, the results are taken after simulation using the Network Simulator.

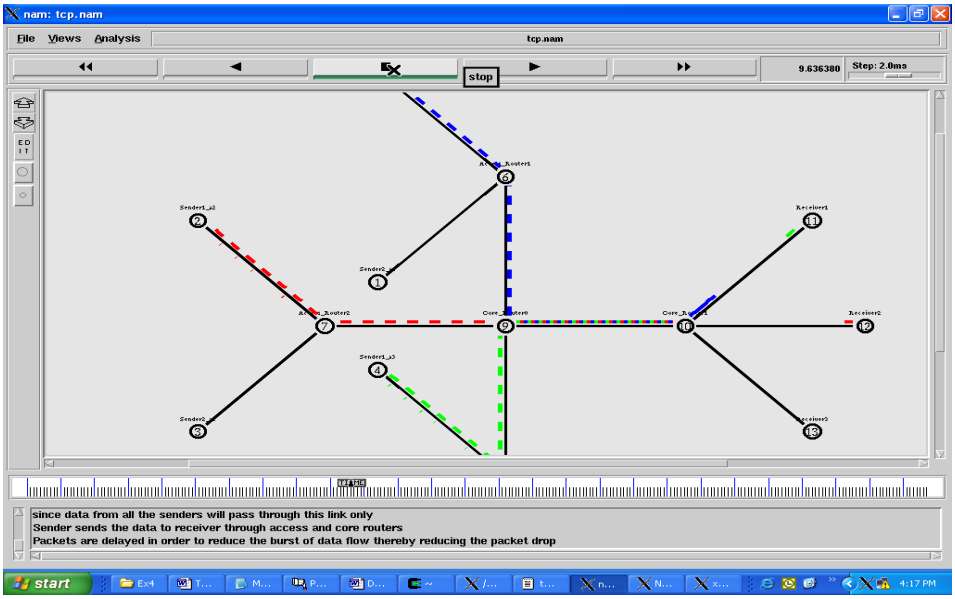
"4. Results and discussions"

The results are based on the simulations done on the non paced and paced techniques. Firstly, the simulations are performed on non paced techniques and finally on the paced techniques. The results are compared to verify the performances. The performance differentials are analyzed using throughput, delay, packet drops and buffer size.

As shown in the figure 2, due to the increased number of data packets being sent by the sending nodes packet drop occurs at the bottle link between core_Router0 and Core_Router1 due to the buffer over flow. The data packets are not yet received at the receivers but the sending process is not terminated but is continuing and due to this the receiver is not in a position to accumulate them all at their buffers. This result the packets to get dropped instead of being stored at the buffer. Thus, losing packets during transmission is like sending incomplete and meaningless data.



"Figure 2. Packets drop due to bursty traffic when applying non paced technique"



"Figure 3. Reducing bursty traffic by applying pacing technique"

As seen in the figure 3, all the data are received at the destinations, even though there is a burst of traffic in the network. So, there are no dropping of packets and due to this, a meaningful transmission is achieved just by delaying the packets during transmission.

Output comparison between existing and proposed technique

Propagation delay of the packet over the link is set as 50ms in the network in which no pacing technique is applied. Propagation delay of the packet over the link set as 100ms in the network in which pacing technique is applied since we intentionally delay the packet to improve the network performance.

Throughput of 4.35156e+09 bits/s is achieved in the network of dumbbell topology without pacing technique in 11.9 seconds with packet drop of 145 packets. But the same throughput is achieved over the same network with pacing technique in 26.6 seconds with no packet drop. So some delay is incurred using the proposed technique but the bursty traffic can be reduced thereby reducing the packet drop.

Propagation delay with pacing = 100ms

Propagation delay without pacing = 50ms

Buffer size = 100

"Table 2. Delay versus throughput and packet drop"

Delay(s)	Without pacing Throughput(bits/sec)	Without pacing Packet drop	With pacing Throughput (bits/sec)	With pacing Packet drop
11.998752	4.35156e+09	145	2090004480	No
26.67936	-	145	4.54045e+09	No

Propagation delay with pacing = 100ms

Propagation delay without pacing = 50ms

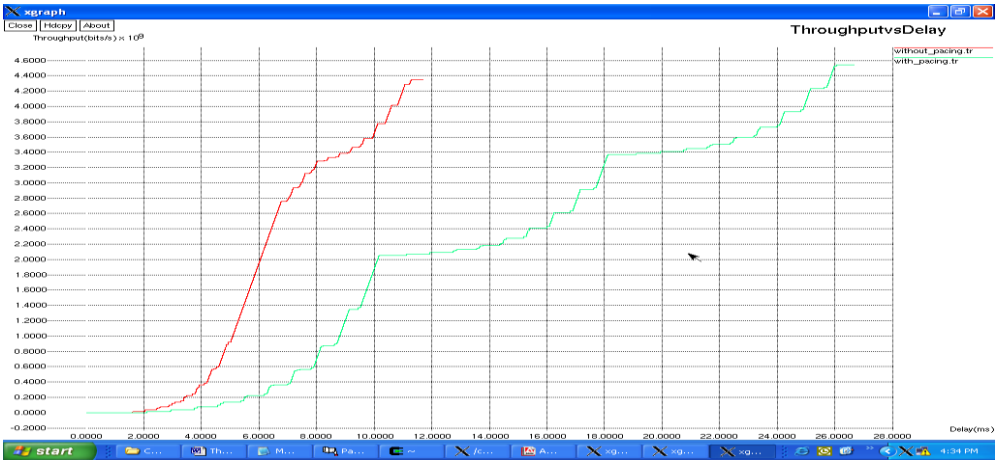
Buffer size = 100 and 50

"Table 3. Buffer size versus throughput and packet drop."

Buffer size (Packets)	Without pacing Throughput(bits/sec)	Without pacing Packet drop	With pacing Throughput (bits/sec)	With pacing Packet drop
100	4.35156e+09	50	4.54045e+09	No
50	3.24154e+09	145	4.54045e+09	No

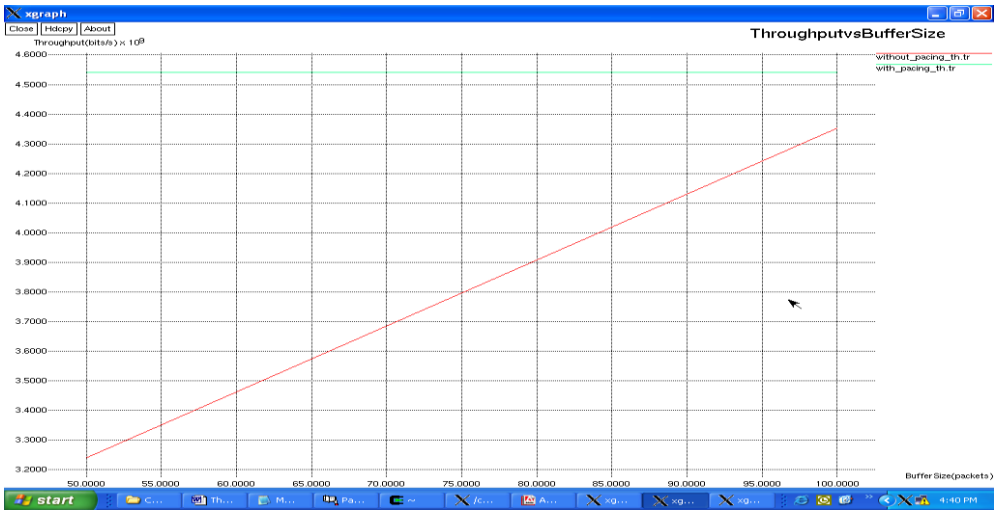
The network without pacing technique produces the lesser throughput when buffer size is decreased. The network with pacing technique produces the stable throughput even when the buffer size is decreased.

Comparative graph of transport layer performance with and without pacing technique



"Figure 4. Delay versus Throughput of non paced and pacing technique"

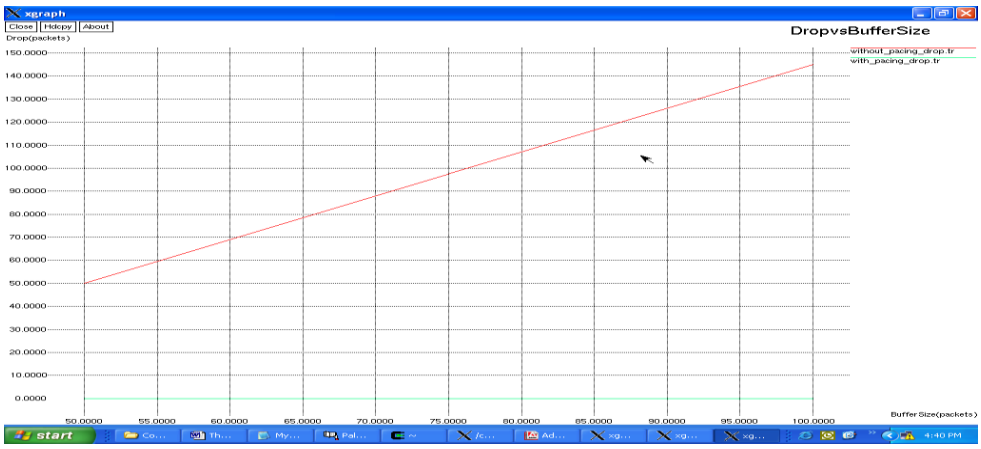
Network without pacing technique produces the higher throughput in 11 seconds. But the network with pacing technique produces the same throughput with some intentional delay but with no packet drop. Thus, achieving the same throughput and with no packet drop means a reliable and secure transmission since all the required data are received to the destination. So, this proposed technique of pacing is useful compared to the non paced method, because data packets are lost in the non paced one.



"Figure 4. Throughput versus buffer size"

Network without pacing technique produces the higher throughput when buffer size is increased. But the network with pacing technique produces the stable throughput even when buffer size is increased. Thus, in case of paced technique there is no effect of buffer size to the throughput. In the case of non paced technique, the throughput and buffer size increases linearly. Once the buffer size is increased the throughput also increases but not up to the paced technique, since in the paced technique the buffer size increases but the throughput remains at the same level.

As shown in figure 5, network without pacing technique produces the higher packet drop when buffer size is increased. But the network with pacing technique produces no packet drop when buffer size is increased. Thus, the pacing techniques is quite useful compared to the non paced one, since there is no packet lost at all.



"Figure 5. Packet drop versus buffer size"

Conclusion

This work presents a novel view on the tradeoff between link bandwidth and packet delay. Instead of using an error correction or network coding approach where more bandwidth is used to avoid packet losses, and proposed to delay packet transmissions to reduce the burstiness of traffic and thus reduce packet losses in small-buffer networks. This present Queue Length Based Pacing, which is a pacing technique that uses a single pacing queue on router ports and adapts its sending rate based on the amount of traffic that is buffered at that port. The analysis shows that pacing delay due to QLBP is bounded and that the variance of the instantaneous traffic rate is reduced. This shows that the effectiveness of QLBP through a prototype implementation and simulation. Specifically, the TCP connections in a small-buffer network with QLBP pacing achieve higher link utilization than in non-paced networks. Therefore, QLBP is an effective approach in improving the operation of networks and improving the effective bandwidth of connections at the cost of only small amounts of additional delay. Stability control mechanism is applied along with proposed technique using Fair Queue that helped to improve the network performance further.

References

[1] A. Vishwanath, V. Sivaraman, M. Thottan, and C. Dovrolis. Enabling a buffer less core network using edge-to-edge packet-level fec. In Proc. IEEE INFOCOM 10, San Diego, CA, March 2010.

[2] M. Shifrin and I. Keslassy. Small-buffer networks. Computer Networks, Volume 53, No. 14, pages 2552–2565, September 2009.

[3] Y. CAI, S. Hanay, and T. Wolf. Practical packet pacing in small-buffer networks. In ICC '09, Dresden, Germany, June 2009.

[4] A. Lakshmikantha, R. Srikant, and C. Beck. Impact of File Arrivals and Departures on Buffer Sizing in Core Router. In Proc. IEEE INFOCOM, 2008.

[5] O. Alparslan, S. Arakawa, and M. Murata. Node pacing for optical packet switching. In Proc. Photonics in Switching, Sapporo, August 2008.

[6] Y. Huang, Y. Liu, W. Gong, and D. Towsley. Two-level Stochastic Fluid Tandem Queuing Model for Burst Impact Analysis. In IEEE CDC, December 2007.

[7] Y. Gu, D. Towsely, C. V. Hollot, and H. Zhang. Congestion control for small buffer high speed networks. In Proc. IEEE INFOCOM 07, Anchorage, Alaska, pages 1037–1045, May 2007,.

[8] M. C. Weigle, P. Adurthi, F. H.-C. K. Jeffay, and F. D. Smith. Tmix: A tool for generating realistic TCP application workloads in ns-2. SIGCOMM Computer Communication Review, Volume 36, No. 3, pages 67–76, 2006.

[9] V. Sivaraman, H. Elgindy, D. Moreland, and D. Ostry. Packet pacing in short buffer optical packet switched networks. In Proc. IEEE INFOCOM 06, Spain, April 2006.

[10] M. Enachescu, Y. Ganjali, A. Goel, N. McKeown, and T. Rough garden. Routers with very small buffers. In Proc. Twenty-fifth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2006), Barcelona, Spain, April 2006.

[11] J. Naor, A. Rosen, and G. Scalosub. Online time-constrained scheduling in linear networks. In Proc. IEEE INFOCOM 05, Miami, FL, March 2005

[12] G. Raina, D. Towsely, and D. Wischik. Part II: Control theory for buffer sizing. ACM SIG COMM Computer Communication Rev, pages 79–82, July 2005.
[13] J. DeHart, F. Kuhns, J. Parwatikar, J. Turner, C. Wiseman, and K. Wong. The open network laboratory: a resource for networking research and education. ACM SIGCOMM Computer Communication Review, Volume 35, No. 5, pages 75–78, October 2005.
[14] D. Wischik and N. McKeon. Part I: Buffer sizes for core routers. ACM SIGCOMM Computer Communication Rev, pages 75–78, July 2005.
[15] M. Adler, S. Khanna, R. Rajaraman, and A. Rosen. Time-constrained scheduling of weighted packets on trees and meshes. Algorithmic, Volume 36, No. 2, pages 123–152, 2003
[16] C. V. Hollot, Y. Liu, V. Misra, and D. Towsley. Unresponsive Flows and AQM Performance. In Proc. IEEE INFOCOM, April 2003.
[17] Y. Wu, W. Gong, and D. Towsley. Analysis of abstract simulation via stochastic differential equation models. In IEEE CDC '03, December 2003.
[18] M. Adler, A. L. Rosenberg, R. K. Sitaram, and W. Unger. Scheduling time-constrained communication in linear networks. Theoretical Comp. Sc, Volume. 35, No. 6, pages 559–623, 2002.
[19] A. Razdan, A. Nandan, R. Wang, M. Y. Sanadidi and M. Gerla. Enhancing TCP Performance in Networks with Small Buffers. Network Research Laboratory Computer Science Department University of California, Los Angeles, CA 90095-1596, 0-7803-7533-X/02/$17.00 (C) IEEE 2002.
[20] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network information flow. IEEE Trans. Inf. Theory, Vol. 46, No. 4, pages 1204–1216, July 2000.
[21] A. Aggarwal, S. Savage, and T. Anderson. Understanding the performance of TCP pacing. In Proc. IEEE INFOCOM 2000, Tel Aviv, Israel, pages 157–1165, Mach 2000.
[22] R. W. Brockett, W. Gong, and Y. Guo. Stochastic analysis for fluid queueing systems. In IEEE CDC, December 1999.
[23] J. D. Salehi, Z. Zhang, J. Kurose, and D. Towsley. Supporting stored video: Reducing rate variability and end-to-end resource requirements through optimal smoothing. IEEE/ACM Trans. Netw. Volume 6, No. 4, pages 397–410, 1998
[24] M. Aron and P. Druschel. TCP: Improving start up dynamics by adaptive timers and congestion control. Rice University, Technical Report TR98-318, 1998.
[25] V. Visweswaraiah and J. Heidermann. Improving restart of idle TCP connections. University of Southern California, Technical Report TR97-661, 1997.
[26] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson. Self similarity through high-variability: Statistical analysis of Ethernet LAN traffic at the source level. IEEE/ACM Trans. Netw., Volume. 5, pages 71–86, 1997.
[27] M. Mathis, J. Mahdavi, S. Floyd, and A.] J. Hoe. Start-up dynamics of TCP's congestion control and avoidance schemes. Masterthesis, MIT, June 1995.
[28] I. Nikolaidis and I. Akyildiz. Source characterization and statistical multiplexing in ATM networks. Georgia Tech., Technical Report GITCC 92-24, 1992.
[29] L. Zhang, S. Shenker, and D. D. Clark. Observations on the dynamics of a congestion control algorithm: the effects of two way traffic. In Proc. ACM SIGCOMM 91, Zurich, Switzerland, pages 133–147 September 1991.
[30] Van Jacobson. Congestion avoidance and Control. In Proceedings of the ACM SIGCOMM '88 Conference on Communications Architectures and Protocols, pages 314–329, August 1988.
[31] D. D. Clark, M. M. Lambert, and L. Zhang. NETBLT: A high throughput transport protocol. ACM SIGCOMM Comp. Comm. Rev., Volume 17, pages 353–359, August 1987.
[32] H. Heffes and D.Lucantoni. A markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance. In IEEE J. Sel. Areas Communication. Pages 856–868 September 1986.
[33] J. Postel. User Datagram Protocol. Information Sciences Institute, RFC 768, August 1980.
[34] Vilas Bagad, Iresh Dhotre. Computer Communication Networks. Technical Publications Pune, First edition: October 2005, Chapter 6, pages 1-24, reprint: August 2006.

Performance analysis of Energy detection, Matched filter detection & Cyclostationary feature detection Spectrum Sensing Techniques

Mr. Pradeep Kumar Verma[1], Mr. Sachin Taluja[2], Prof. Rajeshwar Lal Dua[3]
[1]M.Tech Scholar, Department of Electronics & Communication Engineering, Jaipur National University, Jaipur, India
[2]M.Tech Scholar, Department of Electronics & Communication Engineering, Jaipur National University, Jaipur, India
[3]HOD, Electronics & Communication Engineering, Jaipur National University, Jaipur, India.

Abstract- The growing demand of wireless applications has put a lot of constraints on the usage of available radio spectrum which is limited and precious resource. However, a fixed spectrum assignment has lead to under utilization of spectrum as a great portion of licensed spectrum is not effectively utilized. Cognitive radio is a promising technology which provides a novel way to improve utilization efficiency of available electromagnetic spectrum. Spectrum sensing helps to detect the spectrum holes (underutilized bands of the spectrum) providing high spectral resolution capability. This is a review paper that compares the performance of three main spectrum sensing techniques.

Keywords- Cognitive Radio (CR), Energy Detection (ED), Matched Filter Detection (MFD), Cyclostationary feature Detection.

I. Introduction

The available radio spectrum is limited and it is getting crowded day by day as there is increase in the number of wireless devices and applications. In the studies it has been found that the allocated radio spectrum is underutilized because it has been statistically allocated not dynamically (allocated when needed). Also the approach of radio spectrum management is not flexible, since, each wireless operator is assigned a license to operate in a certain frequency band. In the present scenario, it has been found out that these allocated radio spectrums are free 15% to 85% most of the time i.e. they are inefficiently used depending upon the geographical area. Since most of the useful radio spectrum already allocated, it is difficult to find vacant frequency bands to either deploy new services or to enhance the existing ones. In order to overcome this situation, we need to come up with a means for improved utilization of the spectrum creating opportunities for dynamic spectrum access. [1]-[3].

The issue of spectrum underutilization in wireless communication can be solved in a better way using Cognitive Radio. Cognitive Radio is characterized by the fact that it can adapt according to the environment by changing its transmitting parameters, such as modulation, frequency, frame format, etc. The main challenges with cognitive radios are that it should not interfere with the licensed users and should vacate the band when required. For this it should sense the signals faster. This work focuses on the spectrum sensing techniques that are based on primary transmitter detection. In this category, three major spectrum sensing techniques "energy detection", "matched filter detection", and "cyclostationary feature detection" are addressed. This paper involves the comparative analysis of these spectrum sensing techniques for efficient working of cognitive radios.

II. Cognitive Radio

The concept behind the Cognitive users is that they have the ability to continuously sense the licensed spectrum to search the unused locations, once the hollow locations are identified; Cognitive radio users utilize those locations for transmission without interrupting the primary users. [4].The primary aim of a cognitive radio system is to get hold of a best available channel by using the cognitive capability and re-configurability. Cognitive capability is defined as the capacity of the radio system to gather information from the surroundings [5]. It requires very complex or sophisticated techniques in order to observe the sudden variations and changes in the radio environment without interfering with the existing users. Cognitive capability plays a major role to identify the unused or white spaces in the frequency spectrum at a particular time so as to select a suitable spectrum along with the suitable operating parameters. These unused channels are called spectrum holes or white spaces [5]. The cognitive radio enables the use of white spaces in the spectrum that become available temporarily. As soon as the primary user returns to its band, the cognitive user switches to a different spectrum hole or may stay in the same band but alters the power level and modulation method for avoiding interference to the existing licensed users in that band.

III. Spectrum Sensing

A major challenge in cognitive radio is that the secondary users need to detect the presence of primary users in a licensed spectrum and quit the frequency band as quickly as possible if the corresponding primary radio emerges in order to avoid interference to primary users. This technique is called spectrum sensing. Spectrum sensing and estimation is the first step to implement Cognitive Radio system [6].

Energy detection Spectrum Sensing

It is a non coherent detection method that detects the primary signal based on the sensed energy [1]. Due to its simplicity and no requirement of a priori knowledge of primary user signal, energy detection (ED) is the most popular sensing technique in cooperative sensing [7][8][9].

The block diagram for the energy detection technique is shown in the Figure 1. In this method, signal is passed through band pass filter of the bandwidth W and is integrated over time interval. The output from the integrator block is then compared to a predefined threshold. This comparison is used to discover the existence of absence of the primary user. The threshold value can set to be fixed or variable based on the channel conditions.

The ED is said to be the Blind signal detector because it ignores the structure of the signal. It estimates the presence of the signal by comparing the energy received with a known threshold ν derived from the statistics of the noise. Analytically, signal detection can be reduced to a simple identification problem, formalized as a hypothesis test,

$$y(k) = n(k) \quad \ldots\ldots\ldots\ldots\ldots\ldots H_0$$

$$y(k) = h * s(k) + n(k) \quad \ldots\ldots\ldots H_1 \qquad (1)$$

Where y(k) is the sample to be analyzed at each instant k and n (k) is the noise of variance σ^2. Let y(k) be a sequence of received samples k Є {1, 2….N} at the signal detector, then a decision rule can be stated as,

$$H_0 \ldots if\ \varepsilon < v$$

$$H_1 \ldots if\ \varepsilon > v \qquad (2)$$

Where $\varepsilon = |E\ y(k)|^2$

The estimated energy of the received signal and *v* is chosen to be the noise variance σ^2 [10].

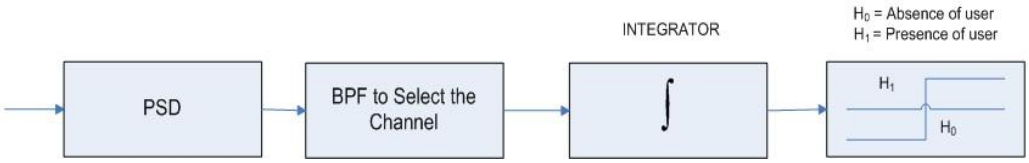


Fig. 1: Block diagram of Energy detection [1].

The "probability of primary user detection" and the "probability of false detection" for the energy detection method can be calculated by the given equations [10]:

$$P_d = P\ [Y > \lambda / H1] = Q_m\ (\sqrt{2\gamma}, \sqrt{\lambda}),$$

$$P_f = P\ [Y > \lambda / H0] = \Gamma\ (m, \lambda/2) / \Gamma\ (m) \qquad (3)$$

Where λ= SNR,

n= TW (Time bandwidth product)

Γ(.)= complete gamma function,

Γ(.,.)= incomplete gamma function,

Qm= Generalized Marcum function [11].

Matched Filter Spectrum Detection

A matched filter (MF) is a linear filter designed to maximize the output signal to noise ratio for a given input signal. When secondary user has a priori knowledge of primary user signal, matched filter detection is applied. Matched filter operation is equivalent to correlation in which the unknown signal is convolved with the filter whose impulse response is the mirror and time shifted version of a reference signal. The operation of matched filter detection is expressed as:

$$Y[n] = \sum_{k=-\infty}^{\infty} h[n-k]x[k] \qquad (3)$$

Where 'x' is the unknown signal (vector) and is convolved with the 'h', the impulse response of matched filter that is matched to the reference signal for maximizing the SNR. Detection by using matched filter is useful only in cases where the information from the primary users is known to the cognitive users [10].

Fig. 2: Block diagram of matched Filter detection [1].

Cyclostationary feature Spectrum Detection

It exploits the periodicity in the received primary signal to identify the presence of primary users (PU). The periodicity is commonly embedded in sinusoidal carriers, pulse trains, spreading code, hopping sequences or cyclic prefixes of the primary signals. Due to the periodicity, these cyclostationary signals exhibit the features of periodic statistics and spectral correlation, which is not found in stationary noise and interference [11].

Thus, cyclostationary feature detection is robust to noise uncertainties and performs better than energy detection in low SNR regions. Although it requires a priori knowledge of the signal characteristics, cyclostationary feature detection is capable of distinguishing the CR transmissions from various types of PU signals. This eliminates the synchronization requirement of energy detection in cooperative sensing. Moreover, CR users may not be required to keep silent during cooperative sensing and thus improving the overall CR throughput. This method has its own shortcomings owing to its high computational complexity and long sensing time. Due to these issues, this detection method is less common than energy detection in cooperative sensing [12].

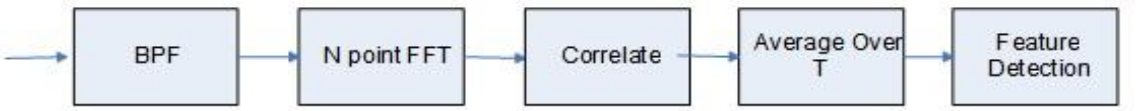


Fig. 3: Block diagram of Cyclostationary feature detection [1].

Process flow diagrams

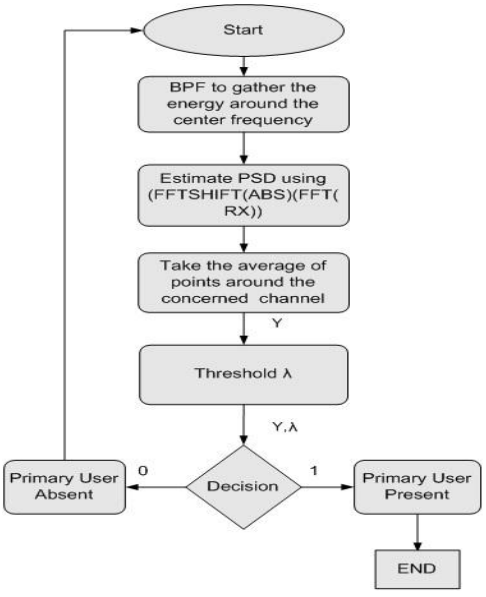


Fig. 4: Process flow diagram of Energy Detection.

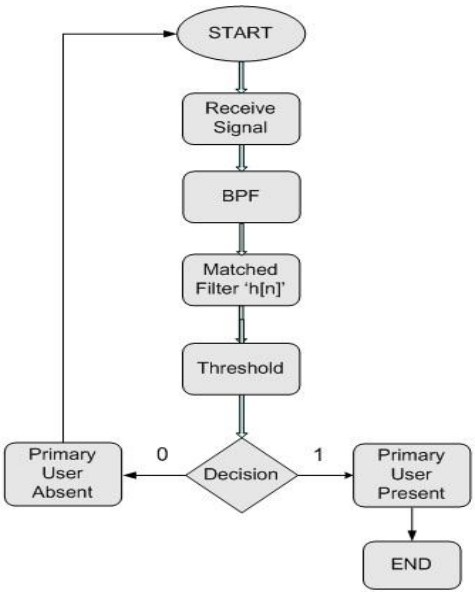


Fig. 5: Process flow diagram of Matched filter Detection.

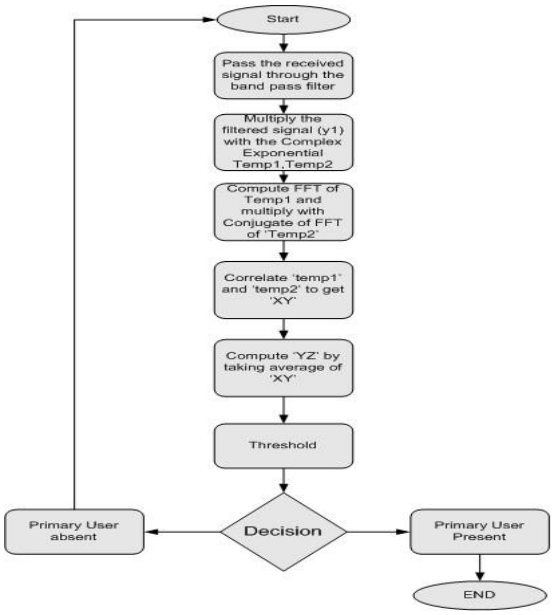


Fig. 6: Flow diagram of Cyclostationary feature Detection.

IV. Results And Analysis

An extensive set of simulations have been conducted using the system model as described in the previous section. The emphasis is to analyze the comparative performance of three spectrum sensing techniques. The performance metrics used for comparison include the "probability of primary user detection" and "probability of false detection". The number of channels and the number primary users considered in this analysis is twenty five and respectively. The SNR of the channels is considered to be precisely same and the channel model is AWGN with zero mean. The results are shown in Figure-7 and Figure-8.

Probability of Primary Detection

Figure-7 depicts the "probability of primary user detection" as a function of SNR for the three cases: (i) energy detection, (ii) matched filter detection and (iii) cyclo-stationary feature detection.

It is observed that for energy detection and matched filter detection, much higher SNR is required to obtain a performance comparable to cyclostationary feature detection. For energy detection, about 16 dBs higher SNR is needed to achieve 100% probability of detection whereas for matched filter detection, about 24 dB s higher SNR is required. For cyclostationary feature detection, 100% probability of detection is attained at -8 dB s. Cyclostationary feature detection performs well for very low SNR, however the major disadvantage is that it requires large observation time for occupancy detection. Matched filter detection performs well as compared to energy detection but restriction lies in prior knowledge of user signaling. Further, cyclostationary feature detection algorithm is complex as compared to other detection techniques.

Probability of False Detection

Figure-8 illustrates the "probability of false detection" for three transmitter detection based spectrum sensing techniques versus SNR.

It is observed that "probability of false detection" of cyclostationary feature detection is much smaller as compared to other two techniques. In fact, it is zero for the range of SNR considered in this study i.e., -30 dB to +30 dBs. It is further seen that the "probability of false detection" for energy detection technique is inversely proportional to the SNR. At low SNR we have higher probability of false detection and at high SNR we have lower probability of false detection, because energy detection cannot isolate between signal and noise. The probability of false detection for energy detection and matched filter detection approaches zero at about +14 dBs and +8 dBs respectively.

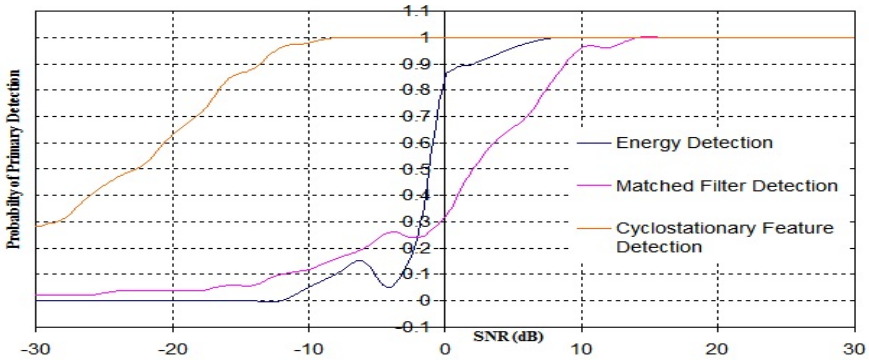


Fig. 7: Probability of Primary Detection.

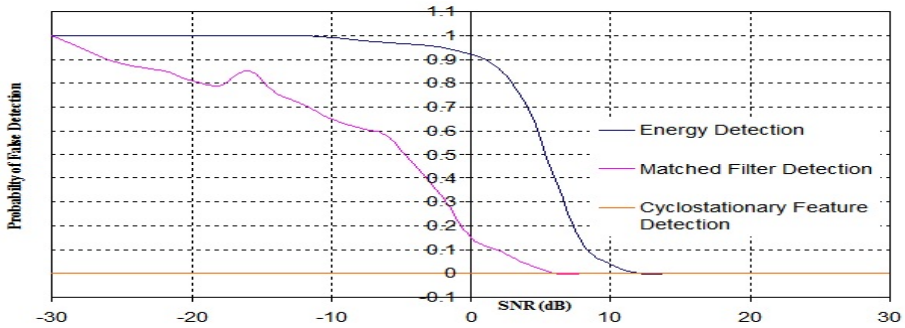


Figure 8: Probability of False Detection.

I. Conclusion

To efficiently utilize the wireless spectrum cognitive radios were introduced which opportunistically utilize the holes present in the spectrum. The most essential aspect of a cognitive radio system is spectrum sensing and various sensing techniques which it uses to sense the spectrum. In this paper the main focus was on Energy Detection, Matched Filter Detection and Cyclostationary feature Detection spectrum sensing techniques. The advantage of Energy detection is that, it does not require any prior knowledge about primary users. It does not perform well at low SNR values, it requires a minimum SNR for its working. The result in the paper shows that Energy detection starts working at -7 dB s of SNR. Matched filter detection is better than energy detection as it starts working at low SNR of -30 dB s. Cyclostationary feature detection is better than both the previous detection techniques since it produces better results at lowest SNR, i.e. for values below -30 dB s. the results shows that the performance of energy detection gets better with increasing SNR as the "probability of primary detection" increases from zero at -14 dB s to 100% at +8 dB s and correspondingly the "probability of false detection" improves from 100% to zero. Similar type of performance is achieved using matched filter detection as "probability of primary detection" and the "probability of false detection" shows improvement in SNR as it varies from -30 dB s to +8 dB s. the cyclostationary feature detection outclasses the other two sensing techniques as 100% "probability of primary detection" and zero "probability of false detection" is achieved at -8 dB s, but the processing time of cyclostationary feature detection is greater than the energy detection and matched filter detection techniques.

References

[1] Shahzad A. et. al. (2010), "Comparative Analysis of Primary Transmitter Detection Based Spectrum Sensing Techniques in Cognitive Radio Systems,'' Australian Journal of Basic and Applied Sciences,4(9), pp: 4522- 531, INSInet Publication.

[2] Weifang Wang (2009), "Spectrum Sensing for Cognitive Radio'', Third International Symposium on Intelligent Information Technology Application Workshops, pp: 410- 12.

[3] V. Stoianovici, V. Popescu, M. Murroni (2008), "A Survey on spectrum sensing techniques in cognitive radio" Bulletin of the Transilvania University of Braşov, Vol. 15 (50).

[4] Mitola, J. and G.Q. Maguire, 1999 "Cognitive Radio: Making Software R adios More Personal," IEEE Personal Communication Magazine, 6(4): 13-18.

[5] Haykin, S., 2005. "Cognitive Radio: Brain Empowered Wireless Communications", IEEE Journal on Selected Areas in Communications, pp: 201-220.

[6] D. B. Rawat, G. Yan, C. Bajracharya (2010), "Signal Processing Techniques for Spectrum Sensing in Cognitive Radio Networks'', International Journal of Ultra Wideband Communications and Systems, Vol. x, No. x/x, pp:1-10.

[7] Ekram Hossain, Vijay Bhargava (2007), "Cognitive Wireless Communication Networks", Springer.
[8] Linda Doyle (2009), "Essentials of Cognitive Radio", Cambridge University Press. [9] D. Cabric, A. Tkachenko, and R. Brodersen, (2006) "Spectrum sensing measurements of pilot, energy and collaborative detection," in Proc. IEEE Military Commun. Conf., Washington, D.C., USA, pp: 1-7.
[10] Ian F. Akyildiz, Brandon F. Lo, Ravikumar (2011), "Cooperative spectrum sensing in cognitive radio networks: A survey, Physical Communication", pp: 40-62.
[11] A. Tkachenko, D. Cabric, and R. W. Brodersen, (2007), "Cyclostationary feature detector experiments using reconfigurable BEE2," in Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks, Dublin, Ireland, Apr, pp: 216-219.
[12] R. Tandra and A. Sahai (2007), "SNR walls for feature detectors", in Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks, Dublin, Ireland, Apr, pp: 559–570.
[13] Parikshit Karnik and Sagar Dumbre (2004), "Transmitter Detection Techniques for Spectrum Sensing in CR Networks", Department of Electrical and Computer Engineering Georgia Institute of Technology.

Authors

Pradeep Kumar Verma- Student of M. Tech (Communication and signal processing) final semester at Jaipur National University, Jaipur. Completed B. Tech from Northern India Engineering College, Lucknow from Uttar Pradesh Technical University in Electronics and Communications Engineering in 2009. Worked as Site Engineer for 9 months in Telecom Industry. I have keen interest in subjects like signal and systems, digital communications, information theory and coding and wireless communications.

Sachin Taluja M.Tech Scholar at Jaipur National University, Jaipur. He received B.E. from M.D.University Rohtak, Haryana in Electronics and Communication . He has over 5 years of Industrial experience in the Field of Computers. His Area of interest includes Network Security, Artificial intelligence, Communication system, Computer architecture, Wireless Communications, Digital Communications, fiber optics, Nano Technology. He has attended various workshops on different domains of computers.

Professor Rajeshwar Lal Dua – He is a fellow life member of IETE and also a life member of I.V.S & I.P.A, former-scientist F of the Central Electronics Engineering Research Institute (CEERI), Pilani. Has been one of the most well known scientists in India in the field of Vaccum Electronics Devices for over the and half decades. His professional achievements span a wide area of vaccum microwave devices ranging from crossed-field and linear-beam devices to present-day gyrotrons. He was awarded a degree of M. Sc (Physics) and M. Sc Tech (Electronics) from BITS Pilani. He started his professional carrier in 1966 at Central Electronics Engineering Research Institute (CEERI), Pilani. During this period he designed and developed a specific high power Magnetron for defense and batch produced about 100 tubes for their use. Trained the Engineers of Industries with know how transfer for further production of the same.
In 1979 he visited department of Electrical and Electronics Engineering at the University of Sheffield (UK) in the capacity of independent research worker and Engineering Department of Cambridge University Cambridge (UK) as a visiting scientist. He has an experience of about 38 years in area of research and development in Microwave field with several papers and a patent to his credit. In 2003 retired as scientist from CEERI, PILANI & shifted to Jaipur and joined the profession of teaching. From last eight years he is working as professor and head of electronics department in various engineering colleges. At present he is working as head and Professor in the department of Electronics and communication engineering at JNU, Jaipur. He has guided several thesis of M.tech .of many Universities.

Landuse/Landcover Mapping of Achanakmar Amarkantak Biosphere Reserve, India Using Unsupervised Classification Technique

Sateesh Karwariya[1], Shashikant Tripathi[2]

1 (Research Scholar, Mahatma Gandhi Chitrakoot Gramodaya Vishwavidyalaya, Chitrakoot-485331, Satna (M.P.) INDIA,

2 (Associate Professor & Course coordinator Dept. of Science & Environment M.G.C.G.V. Chitrakoot-Satna (M.P.) INDIA,

Abstract

Achanakmar Amarkantak Biosphere Reserve located at the junction of hill ranges of Madhya Pradesh and Chhattisgarh state, India occupying total area of 3835.51sq.km. with topography ranging from high mountains, shallow valleys and plains. The core region of Achanakmar Amarkantak Biosphere Reserve falls in Chhattisgarh state lies between 22015' to 220 58' N and 810 25' to 820 50' E ,falls under the survey of India toposheet No. 64 F5,6,7,9,10,11,1314,15,64J1,J3.The Biosphere is bounded by Anuppur, Dindori and Bilaspur district.

Landuse and Landcover is an important parameter for developmental planning. In the present study an attempt has been made to generate the landuse landcover map from IRS satellite image using unsupervised classification. The study is based on secondary data, and using ERDAS IMAGINE software for all processing of the study area. The satellite imagery has used to prepare the land use and land cover map using unsupervised classification. The land use and land cover map clearly shows that area of Forest land is higher than others.

The main objective of this research paper is to analyzing the disturbance gradient in the Biosphere and development of wildlife information base including inventory data on habitats and species.

Keywords: Landuse, Landcover, LISS-3 Data, Biosphere reserve, Wildlife,Satellite imagery, Secondary Data Unsupervised classification technique.

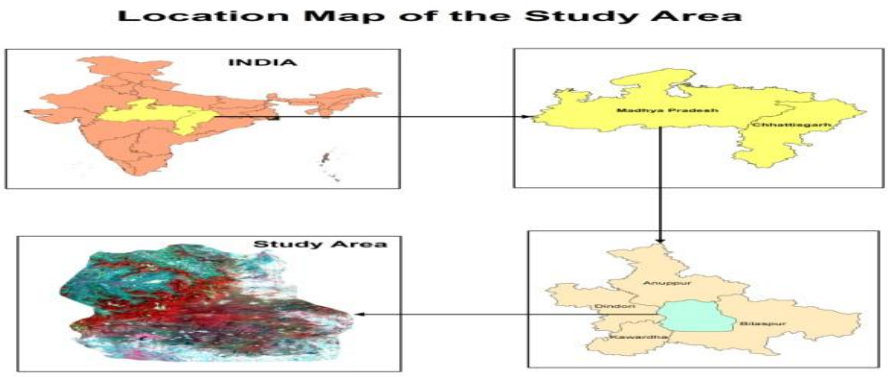
I. Introduction

Satellite Remote Sensing has become an important tool for monitoring and management of natural resources and the environment. Remotely sensed data are widely used in landuse/landcover classification. Landcover relates to the discernible Earth surface expressions, such as vegetation, soil, water or anthropogenic features, and thus describes the Earth's physical state in terms of the natural environment and the man-made structures (Xavier Baulies and Gerard Szejwach, 1998). Essentially, land cover can have only one class or category at a given time and location, and can be mapped using suitable remote sensing data with spectral signatures. Land use is an expression of human uses of the landscape, e.g. for residential, commercial, or agricultural purposes, and has no spectral basis for its unique identification. Thus it cannot be explicitly derived from image data, but only inferred by visual interpretation or assessed in the framework of object-based contextual analysis Landuse is obviously constrained by environmental factors such as soil characteristics, climate, topography and vegetation. But it also reflects the importance of land as a key and finite resource for most human activities including agriculture, industry, forestry, energy production, settlement, recreation, and water catchment and storage. Land is a fundamental factor of production, and through much of the course of human history, it has been tightly coupled with economic growth. Often improper Landuse is causing various forms of environmental degradation for sustainable utilization of the land ecosystems, it is essential to know the natural characteristics, extent and location, its quality, productivity, suitability and limitations of various landuses. Landuse is a product of interactions between a society's cultural background, state, and its physical needs on the one hand, and the natural potential of land on the other (Balak Ram and Kolarkar 1993). In order to improve the economic condition of the area without further deteriorating the bio environment, every bit of the available land has to be used in the most rational way.

As a result of technological advancements, changes of the earth's surface have become visible by satellite imagery as a result remote sensing has become the most effective tool for assessing and monitoring all these transition (Deer,1995) Therefore satellite remote sensing has become a major data source for different change detection applications, because of the repetitive data acquisition capabilities, digital format suitability for computer processing and lower cost than those associated with traditional methods (Coppin et al. 2002; Deer 1995; Lu et al. 2004)

II. Study Area

The Achanakmar-Amarkantak Biosphere Reserve Located in the states of Madhya Pradesh and Chhattisgarh, the Achanakmar-Amarkantak Biosphere Reserve is one of the premium biosphere reserves in India. The reserve covers a huge area of 3835.51 sq. km. and it falls in almost northern part of Bio-geographic zone.About 68.10% out of the total area of this reserve lies in the Bilaspur district in Chhattisgarh.The other major portions of the reserve fall in the Anuppur (16.20 %) and Dindori (15.70 %) districts of Madhya Pradesh. The protected area, Achanakmar Sanctuary is located in Bilaspur district, within the area of the Biosphere Reserve. The sanctuary has a total geographical area of 551.15 sq. km. The Achanakmar-Amarkantak Biosphere Reserve has been divided into core and buffer zone. The entire area of the Achanakmar Sanctuary is designated as the core zone of the reserve and the rest of the 3284.36 sq. km. are serving as the buffer zone, of this reserve. Out of the total area of the buffer zone, an area of 1224.98 sq. km. falls in the state of Madhya Pradesh and the remaining area of 2059.38 sq. km. falls in the Chhattisgarh state. The topography of the Biosphere reserve varies from the rice fields below in Bilaspur and Anuppur and the wheat fields in Dindori, to the hills of Maikal range of Satpuras. The topography of the soil in the Amarkantak plateau is of bauxite rocks. Several streams and Nallas are flowing through the reserve and many of them are perennial. The area of the Achanakmar-Amarkantak Biosphere Reserve is considered as one of the major watershed of peninsular India. It separates the rivers that drain into the Arabian Sea and Bay of Bengal. The reserve is also unique or being source of three major river systems like Narmada, Johilla and Sone of the Ganga basin, and also the Ama Nalla stream. The junction of the hill ranges like Vindhya and Satpura, the Maikal hill ranges lie within the Achanakmar-Amarkantak Biosphere Reserve.



Data Used

1. SOI Toposheet
2. Satellite Data

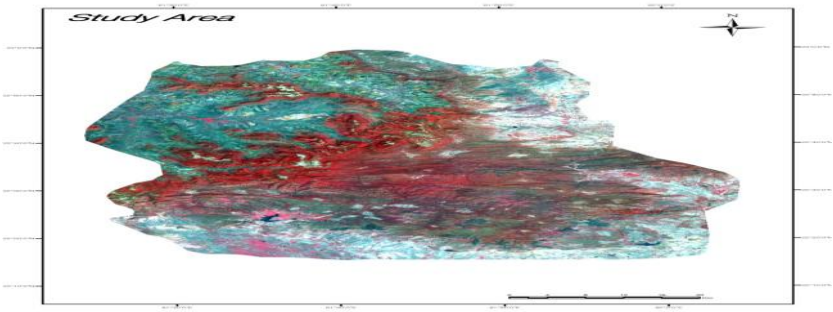


Figure 1: IRS LISS-III satellite image of the study area

III. Aim and Objective

- Preparation of various thematic data such land use and Land cover using LISS-3 Data.
- Create a land use land cover map from satellite imagery using unsupervised classification.
- To analyzing the disturbance gradient in the Biosphere and development of wildlife information base including inventory data on habitats and species.

IV. Methodology

- Geometric corrections of IRS-1C LISS-III data using survey of India (SOI) Toposheet at 1:50,000 scales.
- Selection of study area
- Land use/Land cover classification using unsupervised classification
- The extraction of thematic layers.
- Comparison of overall accuracies of each method with respect to performance Evaluation /Accuracy assessment.
- Output generation

Data Processing

- Analysis and interpretation of satellite data will be done by digital image processing as depicted the process generally includes 3 steps:

1. IMAGE PRE-PROCESSING
2. IMAGE ENHANCEMENT
3. IMAGE CLASSIFICATION

Field Surveys

Field surveys will be conducted within the study areas to determine the major types of landuse and landcover. Such data would be used in two aspects of the mapping of land use land cover. Firstly it will aid in land use and land cover classification, by associating the ground features of a specific type of land use and land cover with the relevant imaging and spectral characteristics. Secondly, ground data will be used for accuracy assessment of the developed land use and land cover maps.

V. Result

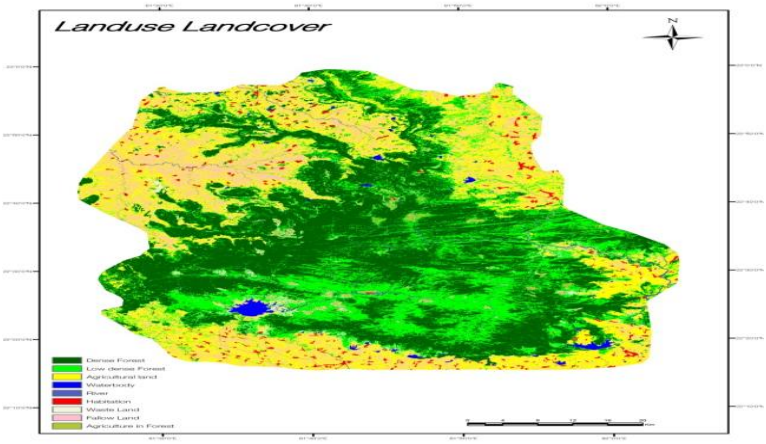


Figure2: Land use and Land cover details of the study area

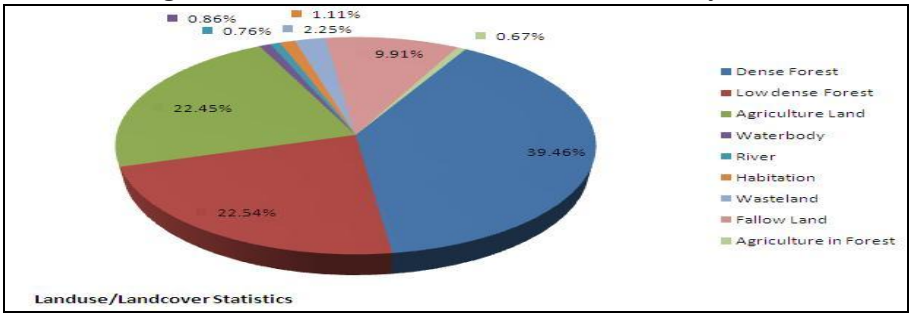


Figure 3: landuse landcover Statistics of Area

The land use and land cover map clearly shows that area of Forest Land is higher than others. Agriculture Land occupies second while fallow land hold the third place in given map.

Class	Area (ha)
Dense Forest	151555.14
Low dense Forest	86574.24
Agriculture Land	86223.87
Water body	3291.39
River	2900.61
Habitation	4270.14
Wasteland	8656.92
Fallow Land	37500.72
Agri. in Forest	2581.2
Total Area	383554.23

VI. Accuracy Assessment

The classification accuracy is most important aspect to assess the reliability of maps, especially when comparing different classification techniques. During this study the accuracy assessment method were used. Accuracy assessment, automatic random point to be selected by software. This method shows above 90% accuracy of map.

VII. Conclusion

This study shows how to classify land use and land cover map from Multispectral satellite imagery using unsupervised classification technique. In this method we calculated land use and land cover classes and their area using image interpretation keys and their unique spectral signature, the land use and land cover map clearly shows that area of Forest is higher than others. The present study supports their results by achieving highest accuracy even in case of Land use land cover mapping.

The classification accuracies were interpreted in terms of

(a) Effect of spatial resolution with same bandwidth.
(b) Comparison of three band set with MIR as either an additional Band or a replacement,
(c) Effect of date of acquisition

VIII. References

[1] Karwariya Sateesh, Goyal Sandip (2011). Land use and Land Cover mapping using digital classification technique in Tikamgarh district, Madhya Pradesh, India using Remote Sensing. INTERNATIONAL JOURNAL OF GEOMATICS AND GEOSCIENCES Volume 2, No 2, 2011.
[2] Ram Chandra, T.V. S Uttam Kumar (2005): from Abstract of their paper presented at map India; geomedia 2005, Image fusion in GRDSS for land use mapping.
[3] Xavier Baulies and Gerard Szejwach, 1998
[4] S.Sudhakar et, al. (1999). Techniques of Classification for Land use/Land cover with special reference for Forest type mapping in Jaldapara Wild life Sanctuary. Journal of the Indian society of Remote Sensing, Vol. 27.No.4, 1999.
[5] Arvik, J.H. 1997. A review of remote imaging source provider. Modern Agriculture, journal for site-specific crop management.
[6] Grenzdorffer, G. 1997. Remote sensing and geographic information system for site-specific farm management system. In J.V. Stafford (ed) Precision Agriculture

Books

[1] Lillesand, J.M. and Kiefer, R.W., "Remote sensing and image interpretation
[2] Jenson, J.R., 1986 "Digital image processing.
[3] Jenson, J.R, 2002"Digital image processing.

A Data Throughput Prediction Using Scheduling And Assignment Technique

M.Rajarajeswari [1], P.R.Kandasamy[2,] T.Ravichandran[3]
1. Research Scholar, Dept. of Mathematics Karpagam University,Coimbatore.
2.Professor and Head, Dept. of M.CA Hindusthan Institute of Technology,coimbatore.
3.The Principal, Hindusthan Institute of Technology, coimbatore.

Abstract:
Task computing is a computation to fill the gap between tasks (what user wants to be done), and services (functionalities that are available to the user). Task computing seeks to redefine how users interact with and use computing environments. Wide distributed many-task computing (MTC) environment aims to bridge the gap between two computing paradigms, high throughput computing (HTC) and high-performance computing (HPC). In a wide distributed many-task computing environment, data sharing between participating clusters may become a major performance constriction. In this project, we present the design and implementation of an application-layer data throughput prediction and optimization service for many-task computing in widely distributed environments using Operation research. This service uses multiple parallel TCP streams which are used to find maximum data distribution stream through assignment model which is to improve the end-to-end throughput of data transfers in the network. A novel mathematical model (optimization model) is developed to determine the number of parallel streams, required to achieve the best network performance. This model can predict the optimal number of parallel streams with as few as three prediction points (i.e. three Switching points). We implement this new service in the Stork Data Scheduler model, where the prediction points can be obtained using Iperf and GridFTP samplings technique. Our results show that the prediction cost plus the optimized transfer time is much less than the non optimized transfer time in most cases. As a result, Stork data model evaluates and transfer jobs with optimization service based sampling rate and no. of task is given as input, so our system can be completed much earlier, compared to non optimized data transfer jobs.

Key words: Optimization, Assignment Technique, Stork scheduling Data throughput.

Modules:
1) Construction of Grid Computing Architecture.
2) Applying Optimization Service.
3) Integration with Stork Data cheduler.
4) Applying Quantity Control of Sampling Data.
5) Performance Comparison.

Existing System:
TCP is the most widely adopted transport protocol but it has major bottleneck. So we have gone for other different implementation techniques, in both at the transport and application layers, to overcome the inefficient network utilization of the TCP protocol. At the transport layer, different variations of TCP have been implemented to more efficiently utilize high-speed networks. At the application layer, other improvements are proposed on top of the regular TCP, such as **opening multiple parallel streams** or **tuning the buffer size**. **Parallel TCP streams** are able to achieve high network throughput by behaving like a single giant stream, which is the combination of n streams, and getting an unfair share of the available bandwidth.

Disadvantage Of System:
- In a widely distributed many-task computing ernvionment, data communication between participating clusters may become a major performance bottleneck.
- TCP to fully utilize the available network bandwidth. This becomes a major bottleneck, especially in wide-area high speed networks, where both bandwidth and delay properties are too large, which, in turn, results in a large delay before the bandwidth is fully saturated.
- Inefficient network utilization.

Proposed System:

We present the design and implementation of a service that will provide the user with the optimal number of parallel TCP streams as well as a provision of the estimated time and throughput for a specific data transfer. A novel mathematical model (optimization model) is developed to determine the number of parallel streams, required to achieve the best network performance. This model can predict the optimal number of parallel streams with as few as three prediction points (i.e. three Switching points). We implement this new service in the Stork Data Scheduler model, where the prediction points can be obtained using Iperf and GridFTP samplings technique.

Advantage of System:

- The prediction models, the quantity control of sampling and the algorithms applied using the mathematical models.
- We have improved an existing prediction model by using three prediction points and adapting a full second order equation or an equation where the order is determined dynamically. We have designed an exponentially increasing sampling strategy to get the data pairs for prediction
- The algorithm to instantiate the throughput function with respect to the number of parallel streams can avoid the ineffectiveness of the prediction models due to some unexpected sampling data pairs.
- We propose to find a solution to satisfy both the time limitation and the accuracy requirements. Our approach doubles the number of parallel streams for every iteration of sampling, and observes the corresponding throughput.
- We implement this new service in the Stork Data Scheduler, where the prediction points can be obtained using Iperf and GridFTP samplings

Implementation module:

In this project we have implemented the optimization service, based on both Iperf and GridFTP. The structure of our design and presents two scenarios based on both, GridFTP and Iperf versions of the service. For the GridFTP version, these hosts would have GridFTP servers. For the Iperf version, these hosts would have Iperf servers running as well as a small remote module (TranServer) that will make a request to Iperf. Optimization server is the orchestrator host, designated to perform the optimization of TCP parameters and store the resultant data. It also has to be recognized by the sites, since the third-party sampling of throughput data will be performed by it. User/Client represents the host that sends out the request of optimization to the server. All of these hosts are connected via LAN. When a user wants to transfer data between two site , the user will first send a request that consists of source and destination addresses, file size, and an optional buffer size parameter to the optimization server, which process the request and respond to the user with the optimal parallel stream number to do the transfer. The buffer size parameter is an optional parameter, which is given to the GridFTP protocol to set the buffer size to a different value than the system set buffer size. At the same time, the optimization server will estimate the optimal throughput that can be achieved, and the time needed to finish the specified transfer between two site. This information is then returned back to the User/Client making the request. Stork is a batch scheduler, specialized in data placement and movement. In this implementation, Stork is extended to support both estimation and optimization tasks. A task is categorized as an estimation task, if only estimated information regarding to the specific data movement is reported without the actual transfer. On the other hand, a task is categorized as optimization, if the specific data movement is performed, according to the optimized estimation results.

Mathematical Model :

A novel mathematical model (optimization model) is developed to determine the number of parallel streams, required to achieve the best network performance. This model can predict the optimal number of parallel streams with as few as three prediction points (i.e. three Switching points).We propose to find a solution to satisfy both the time limitation and the accuracy requirements. Our approach doubles the number of parallel streams for every iteration of sampling, and observes the corresponding throughput. While the throughput increases, if the slope of the curve is below a threshold between successive iterations, the sampling stops. Another stopping condition is: if the throughput decreases compared to the previous iteration before reaching that threshold.

Assignment Problem:

Consider an matrix with n rows and n columns, rows will be consider as grid let and columns will as jobs.
Like,

	Job 1	Job 2	…….	Job n
Grid 1	Task	Task2	…	Task
Grid 2	Tas	Task	…	Task
…….				
Grid n	Tas	Task	…	Task

There will be more than one job for each grid so assign problem occur. to solve this problem we are going for new mathematical model to solve this problem. Three condition occur

1) Find out the minim value of each row and subtract least value with same row task value. Make least value as zero. If process matrix diagonally comes as "0" then process stop and job will assign successfully and job assign successfully.

	Job 1	Job2		Job n
Grid	10	3	…	0
Grid	6	0	…	7
…….				
Grid	0	6	…	7

2) Find out the minim value of each row and Subtract least value with same row task value. Make least value as zero. if column wise matrix diagonally comes as "0" then , then Find out minim value of each column and subtract least value with same row column value. Make least value as zero. Then if process matrix diagonally comes as "0" then process stop and job will assign successfully and job assign successfully.

	Job 1	Job 2	…….	Job n
Grid 1	10	0	…	3
Grid 2	6	0	…	7
…….				
Grid n	6	0	…	7

3) Find out the minim value of each row and subtract least value with same row task value. Make least value as zero. If column wise matrix diagonally comes as "0" then, then Find out minim value of each column and subtract least value with same row column value. Make least value as zero. Then if process matrix diagonally will comes as "0" then process stop and that job will be discard .

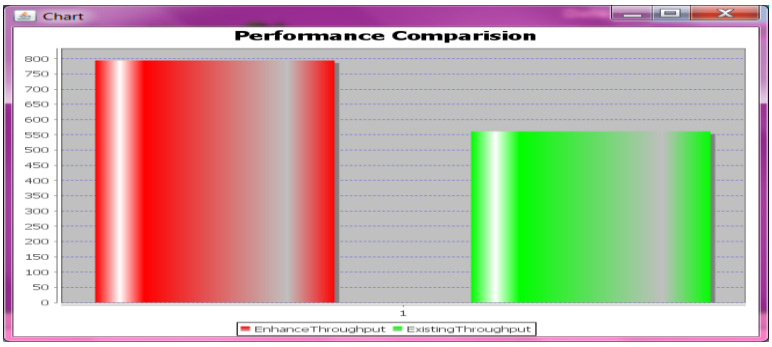
	Job 1	Job 2	…….	Job n
Grid 1	10	0	…	3
Grid 2	6	0	…	7
…….				
Grid n	6	0	…	7

	Job 1	Job 2	…….	Job n
Grid 1	10	0	…	4
Grid 2	6	0	…	7
…….				
Grid n	0	6	…	7

Experimental results:

Test Scenario	Pre-Condition	Test Case	Expected Output	Actual Output	Result
Applying Optimization Service	Client send request to EOS server then that request will be transfer to Gridftp server	Check the client request will be Valid request or not(only txt format file can be download/Estimate from Gridftp server)	Client Requested will be response to client with some information (number of stream taken to transfer the file and throughput achieve at the time of transfer file)	File received successfully	Pass
Integration with Stork Data Scheduler	If More than one Client make request to EOS server	Stork data scheduler make schedule for the incoming request exceed than one request/ checking the is Estimation or optimization Of service	Priority will be assigned to each user request(first come first process) / if estimation request will be give most first priority	According to priority the request process in Gridftp server and response given to client	Pass

Performance Comparison



Conclusion:

This study describes the design and implementation of a network throughput prediction and optimization service for many-task computing in widely distributed environments. This involves the selection of prediction models, the mathematical models. We have improved an existing prediction model by using three prediction points and adapting a full second order equation, or an equation where the order is determined dynamically. We have designed an exponentially increasing sampling strategy to get the data pairs prediction. We implement this new service in the Stork Data Scheduler, where the prediction points can be obtained using Iperf and GridFTP samplings. The experimental results justify our improved models as well as the algorithms applied to the implementation. When used within the Stork Data Scheduler, the optimization service decreases the total transfer time for a large number of data transfer jobs submitted to the scheduler significantly compared to the non optimized Stork transfers.

Reference:

[1] I. Raicu, I. Foster, and Y. Zhao, "Many-Task Computing for Grids and Supercomputers," Proc. IEEE Workshop Many-Task Computing on Grids and Supercomputers (MTAGS), 2008.
[2] Louisiana Optical Network Initiative (LONI), http://www.loni.org/, 2010.
[3] Energy Sciences Network (ESNet), http://www.es.net/, 2010. [4] TeraGrid, http://www.teragrid.org/, 2010.
[5] S. Floyd, "Rfc 3649: Highspeed TCP for Large Congestion Windows," 2003.
[6] R. Kelly, "Scalable TCP: Improving Performance in Highspeed Wide Area Networks," Computer Comm. Rev., vol. 32, Experiments," IEEE Network, vol. 19, no. 1, pp. 4-11, Feb. 2005. no. 2, pp. 83- 91, 2003.
[7] C. Jin, D.X. Wei, S.H. Low, G. Buhrmaster, J. Bunn, D.H. Choe, R.L.A. Cottrell, J.C. Doyle, W. Feng, O. Martin, H. Newman, F. Paganini, S. Ravot, and S. Singh, "Fast TCP: From Theory to Networks," Computer Comm. Rev., vol. 32, Experiments," IEEE Network, vol. 19, no. 1, pp. 4-11, Feb. 2005.
[8] H. Sivakumar, S. Bailey, and R.L. Grossman, "Psockets: The Case for Application-Level Network Striping for Data Intensive Applications Using High Speed Wide Area Networks," Proc. IEEE Super Computing Conf. (SC '00), p. 63, Nov. 2000.
[9] J. Lee, D. Gunter, B. Tierney, B. Allcock, J. Bester, J. Bresnahan, and S. Tuecke, "Applied Techniques for High Bandwidth Data Transfers across Wide Area Networks," Proc. Int'l Conf. Computing in High Energy and Nuclear Physics (CHEP '01), Sept. 2001.
[10] H. Balakrishman, V.N. Padmanabhan, S. Seshan, and R.H.K.M. Stemm, "TCP Behavior of a Busy Internet Server: Analysis and Improvements," Proc. IEEE INFOCOM '98, pp. 252-262, Mar. 1998.
[11] T.J. Hacker, B.D. Noble, and B.D. Atley, "Adaptive Data Block Scheduling for Parallel Streams," Proc. IEEE Int'l Symp. High Performance Distributed Computing (HPDC '05), pp. 265-275, July 2005.
[12] L. Eggert, J. Heideman, and J. Touch, "Effects of Ensemble TCP," ACM Computer Comm. Rev., vol. 30, no. 1, pp. 15-29, 2000.
[13] R.P. Karrer, J. Park, and J. Kim, "TCP-ROME: Performance and Fairness in Parallel Downloads for Web and Real Time Multimedia Streaming Applications," technical report, Deutsche Telekom Labs, 2006.
[14] D. Lu, Y. Qiao, and P.A. Dinda, "Characterizing and Predicting TCP Throughput on the Wide Area Network," Proc. IEEE Int'l Conf. Distributed Computing Systems (ICDCS '05), pp. 414-424, 2005.
[15] E. Yildirim, D. Yin, and T. Kosar, "Prediction of Optimal Parallelism Level in Wide Area Data Transfers," to be published in IEEE Trans. Parallel and Distributed Systems, 2010.

Effect of Traffic Type on the Performance of Table Driven and On Demand Routing Protocols of MANET

Patil V.P.
Smt. Indira Gandhi College of Engineering, New Mumbai, India.

Abstract:
An ad hoc network is a collection of mobile nodes that dynamically form a temporary network. Routing in a MANET is challenging because of the dynamic topology and the lack of an existing fixed infrastructure. In such a scenario a mobile host can act as both a host and a router forwarding packets for other mobile nodes in the network. Routing protocols used in mobile ad hoc networks (MANET) must adapt to frequent or continual changes of topology, while simultaneously limiting the impact of tracking these changes on wireless resources. In his paper investigation has been done on the effect of change in number of nodes on MANET routing protocols. Here, it has been analyzed and compared the performance of MANET routing protocols AODV and DSDV based on both CBR and TCP based traffic patterns. The NS-2 simulator is used for performing various simulations. Simulation results show that Reactive protocols better in terms of packet delivery ratio and average end-to-end delay.

Keywords: AODV, CBR, DSDV, MANET, NS2, Proactive Routing, Reactive Routing, TCP.

I. Introduction

A MANET[1,2] consists of a number of mobile devices that come together to form a network as needed, without any support from any existing Internet infrastructure or any other kind of fixed stations. Formally, a MANET can be defined as an autonomous system of nodes or MSs also serving as routers connected by wireless links, the union of which forms a communication network modeled in the form of an arbitrary communication graph. In such environment, Neighbor nodes communicate directly with each other's while communication between non-neighbor nodes performed via the intermediate nodes which act as routers. As the network topology changes frequently because of node mobility and power limitations, efficient routing protocols are necessary to organize and maintain communication between the nodes.
MANETs have several salient characteristics: i) Dynamic topologies ii) Bandwidth constrained, variable capacity links, iii) Energy-constrained operation and limited physical security etc. Therefore the routing protocols used in ordinary wired networks are not well suited for this kind of dynamic environment. In this paper an effort has been done to evaluate the routing performance of AODV and DSDV using Network Simulator NS2 and results have been analyzed.
The rest of the paper is organized as follows: Section II presents the mobile ad hoc routing protocols. Section III provides an overview and general comparison of the routing protocols used in the study. The related work is described in section IV. The simulation environment and performance metrics are described in Section V and then the results are presented in Section VI. Section VII concludes the paper.

II. Routing Protocols

Routing protocols for Mobile ad hoc networks can be broadly classified into two main categories:
1) Proactive or table-driven routing protocols
2) Reactive or on-demand routing protocols.
3) Hybrid routing protocols

2.1. On-Demand Routing Protocols (Reactive)

Reactive routing protocols [3], [4] try to utilize network bandwidth by creating routes only when desired by the source node. Once a route has been established, it is maintained by some route maintenance mechanism as long as it is needed by the source node. When a source node needs to send data packets to some destination, it checks its route table to determine whether it has a valid route. If no route exists, it performs a route discovery procedure to find a path to the destination. Hence, route discovery becomes on-demand. These routing approaches are well known as Reactive routing. Examples of reactive (also called on-demand) ad hoc network routing protocols include ad hoc on-demand distance vector (AODV), temporally ordered routing algorithm (TORA), dynamic source routing (DSR)[5].

2.2. Table Driven Routing Protocols (Proactive)

In proactive or table-driven routing protocols, each node continuously maintains up-to-date routes to every other node in the network. Routing information is periodically transmitted throughout the network in order to maintain routing table consistency. Thus, if a route has already existed before traffic arrives, transmission occurs without delay. Otherwise, traffic packets should wait in queue until the node receives routing information corresponding to its destination. However, for highly dynamic network topology, the proactive schemes require a significant amount of resources to keep routing information up-to-date and reliable. Certain proactive routing protocols are Destination-Sequenced Distance Vector (DSDV) [9], Wireless Routing Protocol (WRP) [10, 11], Global State Routing (GSR) [11] and Cluster head Gateway Switch Routing (CGSR) [11].

2.3 Hybrid Protocols

Hybrid protocols are the combinations of reactive and proactive protocols and takes advantages of these two protocols and as a result, routes are found quickly in the routing zone. Example Protocol**:** ZRP (Zone Routing Protocol), GPSR (Greedy perimeter stateless routing).

Iii. Overview of Aodv, Dsdv Routing Protocols and Traffic Pattern Types

Every routing protocol has its own merits and demerits, none of them can be claimed as absolutely better than others. In this paper the two reactive routing protocols – AODV, DSDV has been selected for evaluation.

3.1. Ad hoc On-demand Distance Vector Routing (AODV)

Ad-hoc On-demand distance vector (AODV) [12, 13] is another variant of classical distance vector routing algorithm, a confluence of both DSDV [9] and DSR [14]. It shares DSR's on-demand characteristics hence discovers routes whenever it is needed via a similar route discovery process. However, AODV adopts traditional routing tables; one entry per destination which is in contrast to DSR that maintains multiple route cache entries for each destination. The initial design of AODV is undertaken after the experience with DSDV routing algorithm. Like DSDV, AODV provides loop free routes while repairing link breakages but unlike DSDV, it doesn't require global periodic routing advertisements. AODV also has other significant features. Whenever a route is available from source to destination, it does not add any overhead to the packets. However, route discovery process is only initiated when routes are not used and/or they expired and consequently discarded. This strategy reduces the effects of stale routes as well as the need for route maintenance for unused routes. Another distinguishing feature of AODV is the ability to provide unicast, multicast and broadcast communication. AODV uses a broadcast route discovery algorithm and then the unicast route reply massage.

3.2 Destination-Sequenced Distance Vector (DSDV)

Destination-Sequenced Distance Vector (DSDV) [6] is a hop-by-hop distance vector protocol based on the classical Bellman-Ford mechanism. In DSDV, each node maintains a routing table which contains an entry for destination node in the network. The routing table contains entries such as the next hop address, metric or the number of hop counts, and the sequence number. Sequence numbers are assigned by destination node for identification of the routes. DSDV tags each route with a sequence number and considers a route X more favorable than Y if X has a greater sequence number, or if the two routes have equal sequence numbers but X has a lower metric. This was done so that the routing tables have the latest updated path. The sequence number for a route is updated every time a new route discovery is initiated. When a broken link is encountered, the sequence number is set to infinity and it is broadcasted to all the nodes so that the routing tables of the node containing the broken link can be updated to infinity and the link is discarded. The sequence number of every route is assigned by the destination and it is incremented for every route discovery operation. Thus in case of mobile ad-hoc networks, the sequence numbers enable DSDV to maintain up to date routing information at the nodes ensuring the consistency of routing data across all routing tables. Both periodic and triggered route updates are initiated by DSDV to maintain consistency of routing information. In case of periodic updates, fresh route discovery operations are initiated after the elapse of fixed interval of time. Triggered route updates are initiated whenever a node encounters a broken link which can be a result of sudden network topology change or communication link failure.

3.3 Traffic type:

There are two types of traffic patterns used in this paper a) TCP and b) UDP (CBR)

A) Transmission Control Protocol (TCP)

It is often referred to as TCP/IP due to the importance of this protocol in the Internet Protocol Suite. TCP operates at a higher level, concerned only with the two end systems, (e.g. between web browser and a web server). TCP provides reliable, sequential delivery of a stream of data from one program on one computer to another program on another computer. Common uses of **TCP** are e-mailing support, file transfer, Web applications. Among its features, TCP controls message size,

the rate at which messages are exchanged, and network traffic congestion. As for **IP**, it handles lower-level transmissions from computer to computer as a message transferred across the Internet.

B) User Datagram Protocol (UDP)

It is part of the base protocols of the Internet Protocol Suite. Programs on networked computers can send short messages sometimes called datagrams. UDP does not guarantee any reliability (it happens that datagram may be out of order, duplicated, or goes missing without any notice). The fact that no checking whether all packets are actually delivered is made, UDP proves to be faster and more efficient, for applications that do not need guaranteed delivery. UDP find its uses in situations like Time-sensitive applications where the problems due to delayed packets are avoided and It is also useful for servers that answer small queries from huge numbers of clients. UDP supports packet broadcast (conveys to all on local network) and multicasting (conveys to all subscribers).

IV. Related Work

Several researchers have done the qualitative and quantitative analysis of Ad-hoc Routing Protocols by means of different performance metrics. They have used different simulators for this purpose. Broch et al. [19], conducted experiments with DSDV, TORA, DSR and AODV. They used a constant network size of 50 nodes, 10 to 30 traffic sources, seven different pause times and various movement patterns. Packet delivery fraction (PDF), number of routing packets and distribution of path lengths were used as performance metrics. They extended *ns-2* discrete event simulator [20], developed by the University of California at Berkeley and the VINT project [21], to correctly model the MAC and physical-layer behavior of the IEEE 802.11 wireless LAN standard. Ehsan and Uzmi [22], presented the performance comparison of DSDV, AODV, DSR and TORA based on simulations performed using network simulator-2. Three metrics: normalized routing overhead, packet delivery fraction and average end to end delay, were used to measure performance. Karthikeyan et al. [23] studied the Reactive protocols, DSR and AODV as well as a Proactive Protocol, DSDV and their characteristics with respect to different mobility were analyzed based on packet delivery fraction, routing load, end to-end delay, number of packets dropped, throughput and jitter using Network Simulator (ns-2). However, in all the cases, only CBR sources were used for generating traffic.

Performance comparison of AODV and DSR routing protocols in a constrained situation is done in [24]. The authors claim that the AODV outperforms DSR in normal situation but in the constrained situation DSR out performs AODV, where the degradation is as severe as 30% in AODV whereas DSR degrades marginally as 10%.Ahmed and Alam [25] compare three routing protocols (DSR, AODV, and TORA) through simulations conducted with a discrete-event simulator (OPNET Modeler 10.5 version).Simulation results indicate that under specific simulation parameters TORA presents a higher performance than AODV and DSR.

Perkins ET all [26] show the performance of two on demand routing protocols namely DSR and AODV. Though both AODV and DSR use on demand route discovery, they have different routing mechanics. The authors observe that for application oriented metrics such as delay, throughput DSR outperforms AODV when the numbers of nodes are smaller. AODV outperforms DSR when the number of nodes is very large. The authors do show that DSR consistently generate less routing load than AODV. In Kumar et al. [27], a comparison of the performance of two prominent on-demand reactive routing protocols for MANET (DSR and AODV) is presented, along with the traditional proactive DSDV protocol. In Rahman and Zukarnain [28] the performance comparison between three routing protocols, namely AODV, DSDV and an improvement of DSDV, is presented. The authors use three network metrics, namely packet delivery ration, end-to-end delay, and routing overhead.

V. Proposed Methodology and Performance Metrics

5.1 Problem formulation:

The research in this MANET area has continued with prominent studies on routing protocols such as Ad hoc On-demand Distance Vector (AODV), Destination-Sequenced Distance-Vector Routing protocol, (DSDV) and Dynamic Source Routing Protocol (DSR). Several studies have been done on the performance evaluation of routing protocols based on CBR traffic pattern using different evaluation methods. Different methods and simulation environments give different results and consequently, there is need to broaden the spectrum to account for effects not taken into consideration in a particular environment. It is observed that most of the research work is based on CBR traffic pattern whereas most of the traffic approximately 95% on the Internet carries TCP. It is desirable to study and investigate the Performance of different MANET routing protocols under both CBR and TCP traffic patterns. In this paper, we will evaluate the performance of Reactive protocols (AODV) and Proactive protocols (DSDV) of mobile ad-hoc network routing protocols for both CBR and TCP traffic patterns. The performance of these routing protocols is evaluated with respect to various parameters such as average end-to-end delay, throughput and packet delivery ratio.

5.2 Performance metrics:

Design and performance analysis of routing protocols used for mobile ad hoc network (MANET) is currently an active area of research. To judge the merit of a routing protocol, one needs metrics both- qualitative and quantitative- with which to measure its suitability and performance. Specifically, this paper evaluates the performance comparison of AODV and DSDV routing protocols. The following performance metrics are used to compare the performance of these routing protocols in the simulation:

a) Packet Delivery Ratio

Packet delivery ratio is calculated by dividing the number of packets received by the destination through the number of packets originated by the application layer of the source. It specifies the
packet loss rate, which limits the maximum throughput of the network. The better the delivery ratio, the more complete and correct is the routing protocol.

b) Average End-To-End Delay

Average End-to-End delay (seconds) is the average time it takes a data packet to reach the destination. This metric is calculated by subtracting time at which first packet was transmitted by source from time at which first data packet arrived to destination. This includes all possible delays caused by buffering during route discovery latency, queuing at the interface queue, retransmission delays at the MAC, propagation and transfer times. This metric is significant in understanding the delay introduced by path discovery.

c) Throughput

The throughput of the protocols can be defined as percentage of the packets received by the destination among the packets sent by the source. It is the amount of data per time unit that is delivered from one node to another via a communication link. The throughput is measured in bits per second (bit/s or bps).

VI. Simulation, Results and Performance Analysis

6.1 Simulations

In this paper, two different scenarios have been taken. In the first scenario, traffic pattern is taken as CBR and no. of nodes have been varied and performance comparison has been made between AODV and DSDV protocols. In the second scenario, traffic pattern is taken as TCP and no. of nodes have been varied and a performance comparison has been made between AODV and DSDV protocols. Identical mobility pattern are used across protocols to gather fair results.

6.1.1 Test Scenario 1

In the first scenario, the simulation based on TCP traffic pattern is selected. Parameters of this scenario are summarized in table 1. Here, TCP sources are used which use flow control and retransmission feature.

6.1.2 Test case Scenario 2

In the second scenario, the simulation based on CBR traffic pattern has been chosen. Parameters of this scenario are summarized in table 2. CBR sources are used that started at different times to get a general view of how routing protocol behaves.

6.2 Results and performance comparison

Performance of AODV and DSDV protocols is evaluated under both CBR and TCP traffic pattern. Extensive Simulation is done by using NS-2.

Table 1: Simulation Parameters for Test Case of TCP

PARAMETERS	VALUES
No. Of Nodes	25, 50,75, 100
Simulation Time	50 Seconds
Environment Size	1000x1000
Simulation Speed	20 m/sec
Pause Time	5 Seconds
Packet Size	512 bytes
Packet Rate	4 pkts/sec
Traffic Type	TCP
Simulator type	NS2 -2.34

Table 2: Simulation Parameters for Test Case of UDP (CBR)

PARAMETERS	VALUES
No. Of Nodes	25, 50,75, 100
Simulation Time	50 Seconds
Environment Size	1000x1000
Simulation Speed	20 m/sec
Pause Time	5 Seconds
Packet Size	512 bytes
Packet Rate	4 pkts/sec
Traffic Type	UDP (CBR)
Simulator type	NS2 -2.34

6.2.1 Throughput

As it can be seen in case of CBR traffic pattern that throughput of AODV protocol is better than as compared to DSDV protocols. In case of CBR traffic, throughput remains almost constant for these two protocols irrespective of number of nodes. In case of TCP traffic, throughput changes rapidly with respect to change in the number of nodes. This is shown in fig 1 and 2.

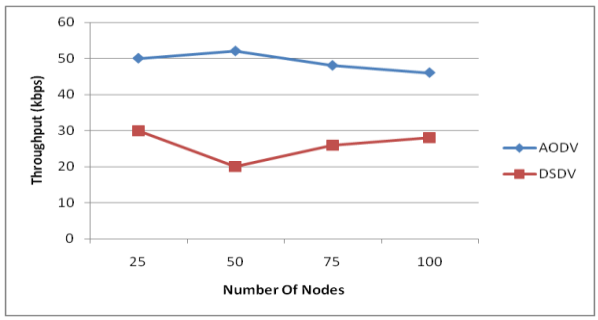


Figure 1: Throughput for CBR Traffic Pattern

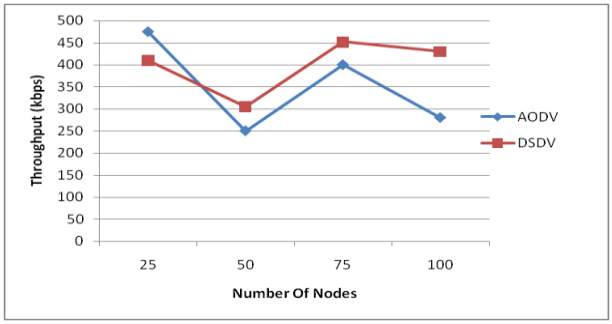


Figure 2: Throughput for TCP Traffic Pattern

6.2.2 Average end-to-end Delay

It is seen that Average end-to-end Delay of proactive routing protocols (DSDV) is less as compared to reactive routing protocol (AODV) in any kind of traffic pattern i.e. either CBR (Figure 3) or TCP (Figure 4). Average end-to-end Delay is also remains almost constant in DSDV whereas it varies in the case of AODV protocol with respect to change in number of nodes.

6.2.3 Packet Delivery Ratio

It is observed that in case of CBR traffic Reactive protocols deliver almost all the originated data packets converging to 100% delivery whereas Proactive protocols (DSDV) Packet Delivery Ratio is approx 50% (Figure 5). Reactive protocols perform better than the proactive protocols in case of CBR traffic pattern. In the case of TCP traffic pattern (Figure 6), Packet delivery ratio of AODV protocols remains almost constant whereas it changes rapidly for DSDV protocol irrespective of the network load.

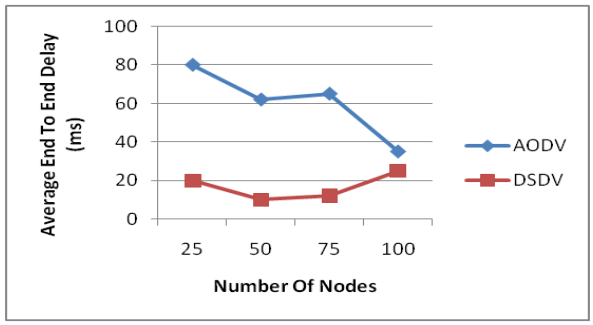


Figure 3: Average End to End Delay for CBR Traffic.

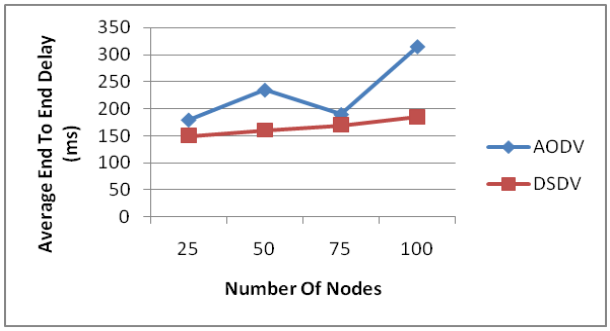


Figure 4: Average End to End Delay TCP Traffic

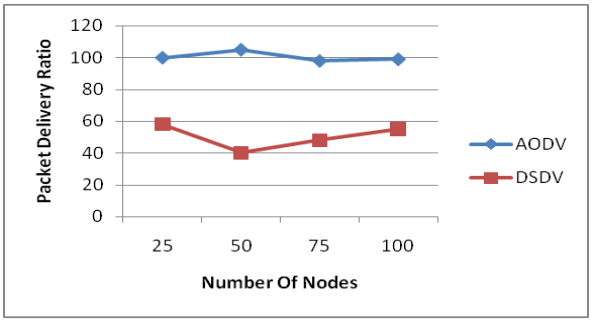


Figure 5: Packet Delivery Ratio for CBR Traffic Pattern.

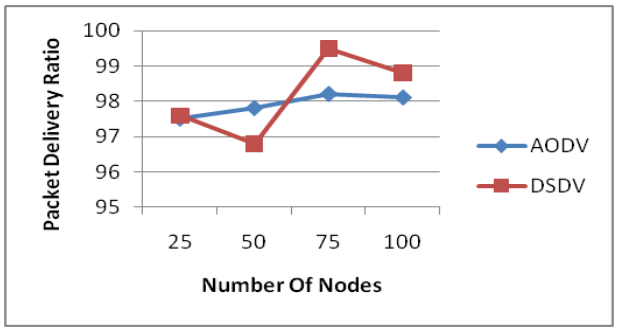


Figure 6: Packet Delivery Ratio for TCP Traffic Pattern

VII. Conclusion

This study was conducted to evaluate the performance of two MANET protocols i.e. AODV, and DSDV based on CMU's generated TCP & CBR traffic. These routing protocols were compared in terms of Packet delivery ratio, Average end-to-end delay, and throughput when subjected to change in no. of nodes & traffic type. In this paper, evaluation of the routing protocols, unlike previous works of similar nature, brings out the effect of routing protocols on the TCP performance. Through experiments, first the essential differences between TCP and UDP traffic and therefore need to consider TCP traffic for routing protocol performance evaluation. Based on these experiments three performance metrics i.e. Packet delivery ratio, Average end-to-end delay, and throughput were measured for quantifying the performance of the protocols. Simulation results show that Reactive protocols better in terms of packet delivery ratio and average end-to-end delay. Performance can also be analyzed for other parameters like Jitter, Routing Overhead. By evaluating the performance of these protocols new protocols can be implemented or changes can be suggested in the earlier protocols to improve the performance.

References

[1] C. E. Perkins and E. M. Royer. "The ad hoc on-demand distance vector protocol," Ad Hoc Networking, Addison-Wesley, pp. 173-219, 2001.

[2] D. B. Johnson, D. A. Maltz, and J. Broch, "DSR: the dynamic source routing protocol for multi-hop wireless ad hoc networks," in Ad Hoc Networking, Addison-Wesley, 2001, pp. 139-172.

[3] Johnson, D. B., Maltz, D. A., Hu, Y. C. and Jetcheva, J. G. 2002. "The dynamic source routing protocol for mobile ad hoc networks (DSR)". Draft-ietf-manet-dsr-07.txt.

[4] erkins, C. E., Royer, E. M. and Das, S. R. 1999. "Ad Hoc On-Demand Distance Vector Routing". IETF Internet Draft. http://www.ietf.org/internet-drafts/draft-ietf-manetaodv-03.txt.

[5] Anuj K. Gupta, Dr. Harsh Sadawarti, Dr. Anil K. Verma, "Performance analysis of AODV, DSR and TORA Routing Protocols", International Journal of Engineering & Technology (IJET), ISSN: 1793-8236, Article No. 125, Vol.2 No.2, April 2010, pp. – 226-231

[6] Narendran Sivakumar, Satish Kumar Jaiswal, "Comparison of DYMO protocol with respect to various quantitative performance metrics", ircse 2009

[7] NS-2, The ns Manual (formally known as NS Documentation) available at http://www.isi.edu/nsnam/ns/doc.NAM

[8] ttp://www.isi.edu/nsnam/nam/

[9] C. E. Perkins and P. Bhagwat, "Highly Dynamic Destination- Sequenced Distance-Vector Routing (DSDV) for Mobile Computers," SIGCOMM, London, UK, August 1994, pp. 234-244.

[10] Vincent D. Park and M.Scott Corson. "A highly adaptive distributed routing algorithm for mobile wireless networks".In Proceedings of INFOCOM 1997, 1997.

[11] Saleh Ali K.AI-Omari & Putra Sumari "An overview of mobile adhoc networks for the existing protocols and applications"Internattional Journal on Applications of Graph theory in wireless adhoc networks and sensor networks (Graph-Hoc) , Vol2 No1 March 2010.

[12] C. E. Perkins and E. M. Royer, "Ad Hoc On-demand Distance Vector Routing," In Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, New Orleans, LA, February 1999, pp. 90-100.

[13] C. Perkins, "Ad hoc On demand Distance Vector (AODV) routing", IETF Internet draft (1997),http://www.ietf.org/internet-drafts/draftietf- manet-aodv-00.txt.

[14] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad-Hoc Ad hoc Networks," Mobile Computing, ed. T. Imielinski and H. Korth, Kluwer Academic Publishers, 1996, pp. 153-181.

[15] Dyer T. D., Boppana R. V., " A Comparison of TCP performance over three routing protocols for mobile adhoc networks ", ACM Symposium on Mobile Adhoc Networking & Computing (Mobihoc) , October 2001.
[16] Suresh Kumar, R K Rathy and Diwakar Pandey, "Traffic pattern based performance comparison of two reactive routing protocols for ad hoc networks using NS2", © 2009 IEEE.
[17] Vikas Singla, Rakesh Singla and Ajay Kumar,"Performance Evaluation and Simulation of Mobile Ad-hoc Network Routing Protocols", International Journal of Engineering and Information Technology, Volume 1 No. 1 October 2009.
[18] Yogesh Chaba, Yudhvir Singh, Manish Joon, "Simulation Based Performance Analysis of On-Demand Routing Protocols in MANETs," Second International Conference on Computer Modeling and Simulation, 2010.
[19] J. Broch, D. A. Maltz, D. B. Johnson, Y. C. Hu and J. Jetcheva, "A Performance Comparison of Multi-Hop Wireless Ad-hoc Network Routing Protocols," in Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM'98), October 1998, pp. 85–97.
[20] Information Sciences Institute, "ns-2 network simulator," Software Package, 2003. [Online]. Available: http://www.isi.edu/nsnam/ns/
[21] "The VINT Project," USC/ISI, Xerox PARC, LBNL and UC Berkeley,1997. [Online].Available: http://www.isi.edu/nsnam/vint/
[22] Humaira Ehsan and Zartash Afzal Uzmi, "Performance Comparison of Ad-hoc Wireless Network Routing Protocols", IEEE Transactions, 2004.
[23] N. Karthikeyan, V. Palanisamy and K. Duraiswamy, "A Performance Evaluation Of Proactive And Reactive Protocols Using ns-2 Simulation", International J. of Engg. Research & Indu. Appls. (IJERIA).ISSN 0974-1518, Vol.2, No.II (2009), pp 309-326.
[24] R. Misra, C. R. Manda (2005)l, "Performance Comparison of AODV/DSR On-Demand Routing Protocols for Ad Hoc Networks in Constrained Situation", *IEEE ICPWC* 2005.
[25]. Ahmed S, Alam MS. "Performance evaluation of important ad hoc network protocols." EURASIP Journal on Wireless Communications and Networking 2006; (2): 42–42.
[26] S. R. Das, C. E. Perkins and E. M. Royer (2000), "Performance comparison of Two On-Demand Routing protocols for Ad hoc Networks", *In Proc. of INFOC OM* 2000, Tel Aviv, Israel, March 2000.
[27] Kumar BRA, Reddy LC, Hiremath PS." Performance comparison of wireless mobile ad-hoc network routing protocols." International Journal of Computer Science and Network Security 2008; 8(6): 337–343.
[28]. Rahman AHA, Zukarnain ZA." Performance comparison of AODV, DSDV and I-DSDV routing protocols in mobile ad hoc networks.", European Journal of Scientific Research 2009; 31(4): 566–576.

AUTHOR

Er. PATIL V.P. is currently working as a faculty member in Electronics and Telecommunication Engineering department in Smt. Indira Gandhi college of Engineering, New Mumbai. He is graduate in B.E. and post graduate in M.TECH (ELECTRONICS DESIGN AND TECHNOLOGY) .He is having 25 years of experience in teaching in engineering colleges. His area of research is in computer communication networking and microwave engineering.

Case study of CAD Based Mammographic Lesions using Wavelet Decomposition

Elayabharathi.T[1], Dr.Nagappan.A[2]

[1] is a research Scholar in the Department of computer science &engineering at the Vinayaka mission research foundation deemed university, Salem, Tamilnadu, India.

Abstract – This paper describes the efforts by study of the characteristics of true masses compared to the falsely detected masses is carried out using wavelet decomposition transform. According to the cancer statistics, the breast cancer incidence rate is increased almost every year since 1980, the death rate is shown a substantial decrease. Both trends may be attributed, in large part to mammography which is widely recognized as the most sensitive technique for breast cancer detection

Index terms –Lesions, Wavelet, contourlet, Mammogram, CAD

1. Introduction

Conventional mammography is a film based x-ray Technique referred as a screen-film mammography. Full field digital mammography is a new technology in which a solid state detector is used instead of film for the generation of the breast image. Modern applications, including computer aided detection and computer aided diagnosis, computer display and interpretation, digital image and transmission and storage, require a digital format of the mammogram

1.1 Objective

The purpose of our study was to retrospectively evaluate the impact on recall rates and cancer detection when converting from film-screen to digital mammography in a small community-based radiology practice.

Digital mammography offers considerable advantages over film-screen mammography [1-3]. Despite advantages, it has been slow to be adopted. This reluctance is due to many factors, including the high initial capital expenditure and the question of whether the added expense results in a better mammography "product" [4-7]. The reluctance to upgrade to digital mammography is especially true of small community-based imaging centers, where capital is less prevalent and patient volumes are lower than in larger metropolitan locations.

1.2 statistics and discussions

Breast cancer ranks first in the causes of cancer deaths among women and is second only to cervical cancer in developing countries[8]. The best way to reduce death rates due to this disease is to treat it at an early stage. Early diagnosis of breast cancer requires an effective procedure to allow physicians to differentiate between benign tumors from malignant ones. Developing computer-aided diagnosis (CAD) systems to help with this task is a non-trivial problem, and current methods employed in pursuit of this goal illustrate the difficulty in maximizing both sensitivity to tumoral growths and specificity in identifying their nature.

X-ray mammography is the best current method for early detection of breast cancer, with an accuracy of between 85% and 95%[3]. Identifying abnormalities such as calcifications and masses often requires the eye of a trained radiologist. As a result, some anomalies may be missed due to human error as a result of fatigue, etc. The development of CAD systems that assist the radiologist has thus become of prime interest, the aim being not to replace the radiologist but to offer a second opinion. Eventually, the state-of-the-art could advance to the point where such systems effectively substitute for trained radiologists, an eventuality that is desirable for small outfits that cannot afford to have an expert radiologist at their continuous disposal. For example, a CAD system could scan a mammogram and draw red circles around suspicious areas. Later, a radiologist can examine these areas and determine whether they are true lesions or whether they are artifacts of the scanning process, such as shadows.

To our knowledge, no prior study has compared cancer detection and recall rates at a single center before and after the installation of a digital mammography system, keeping the interpreting radiologists constant. Such a study would limit the number of uncontrolled variables, allowing potential outcomes to be mediated only by the introduction of the technology and the variability in the women undergoing screening.

2. Background

Considerable effort has been expended to develop CAD systems to aid the trained radiologist identify areas with possible pathology on an image. Most of these efforts have concentrated on X-ray mammography and chest radiography. A number of CAD schemes

have been investigated in literature. These include:

- Subtraction techniques that identify anomalies by comparison with normal tissue
- Topographic techniques that perform feature extraction and analysis to identify anomalies
- Filtering techniques that use digital signal processing filters, often developed especially to augment anomalies for easy detection
- staged expert systems that perform rule-based analysis of image data in an attempt to provide a correct diagnosis

The majority of CAD systems attempt to identify anomalies by either looking for image differences based on comparison with known normal tissue (subtraction techniques)[4] or by image feature identification and extraction of features that correlate with pathological anomalies, such as in texture analysis (topographic techniques)[11, 4, 6, 36]. Most systems proceed in stages, first examining the image data and extracting pre-determined features, then localizing regions of interest or ROIs which can be examined further for potential anomalies. High degrees of sensitivity have been achieved using several of these techniques, but many have been hampered by high false-positive rates and hence low specificity. The problem of false positives is compounded further by the fact that false positive rates are reported per image, not per case. Since many radiological examinations include more than one image, the actual number of false positives may be a multiple of those reported.

A number of different approaches have been employed in an effort to reduce false positive rates, many of them focusing on the use of artificial neural networks (ANNs). A common metric used for evaluating the performance of CAD systems, the receiver operating curve or ROC (see Appendix A), is commonly used to evaluate a CAD scheme's degree of tradeoff between sensitivity and specificity. The area under this curve, A_Z, is a measure of overall performance, with a value of A_Z closer to 1 indicating better performance. Since sensitivity in most techniques is quite high, specificity often becomes the limiting factor, with techniques displaying higher specificity performing at higher A_Z values.

This study decomposes several techniques and identifies their salient features and characteristics with respect to performance. The extent of the array of techniques examined herein is by no means all-inclusive; rather, a number of techniques are described and their performance evaluated.

3. Methods

By careful consideration of the design of various CAD schemes, it is possible to categorize the techniques employed under three broad headings:

- Data reduction - the image is examined in order to identify the ROIs.
- Imageenhancement - the ROIs are subjected to processes that enhance or augment the visibility of pathological anomalies, such as microcalcifications and lesions.
- Diagnosis - the ROIs are subjected to one or more of the broad categories of procedures mentioned in Section 2 in order to arrive at a diagnosis, most commonly in the form of "benign" or "malignant"

These categories are extremely broad, and there may exist CAD systems that subject images to techniques that do not fall under one of them. However, most of the CAD systems employ methods that can be classified under one or more of them.

3.1 Data Reduction

Data reduction is the process by which an image is decomposed into a collection of regions that appear to contain anomalies that differ from the surrounding tissue. These regions are usually a strict subset of the original image and are subregions of the original image that may contain ROIs. By doing this, the CAD system need only process those subregions identified by the data reduction step, rather than the entire input image. Data reduction accomplishes two objectives simultaneously[34]:

- An increase in throughput via a reduction in input data
- A reduction in false positives by limiting the scope of the detection algorithms in the rest of the CAD system to the ROIs only. With less of the original image to worry about, the CAD system gains specificity since less image means less false-positives in general, assuming that the detection algorithms work as intended.

It is clear that the most obvious way to perform data reduction is to have a trained radiologist identify the ROIs for the CAD system. This can be accomplished through a graphical interface to the CAD system that allows the radiologist to specify suspicious regions. It should be noted that some CAD systems do not require this step at all due to the nature of their diagnostic process, such as that those that employ subtraction techniques.

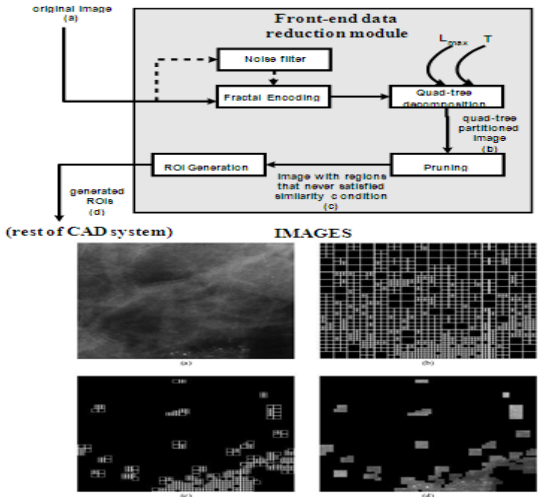


Figure 1: overview of front-end data reduction module with images

3.2 Image Enhancement

Mammographic image enhancement methods are typically aimed at either improvement of the overall visibility of features or enhancement of a specific sign of malignancy. Various schemes for doing this exist, with most of them based in signal processing techniques used either in their original form (such as simple histogram equalization) or adapted for specific use in mammography.

A number of generic image enhancement methods exist. Histogram equalization and fuzzy image enhancement[35] are just two examples. Though a whole slew of image enhancement techniques exist in the general domain, very few are specifically targeted at the enhancement of mammographic images. Section 4 describes one of them.

3.3 Statistical Techniques

Anguh et al[2] propose a multiscale method for segmenting and enhancing lesions of various sizes in mammograms. The first stage applies a multiscale automatic threshold esti- mator based on histogram moments to segment the mammogram at multilevels. The second stage then converts the segmented image using pseudo-color mapping to produce a color image[2]. The final result is analogous to a breast map which provides an adequate basis for radiological breast tissue differentiation and analysis in digital mammography. Their paper provides a treatment on the mathematical theory of moments before present an algorithm for the multiscale thresholding of the mammogram. The result of this thresholding technique is a mammographic map or breast map based on various thresholds with varying object sizes. This map can then be used by a radiologist; however, the CAD scheme proposed in

[2] uses pseudo-color mapping[10] to convert the grayscale to a color image. This is done since human vision can only discern a limited number of grayscale levels. The end results is a pseudo-color breast map in which the lesions have been highlighted in different colors and confirmed by visual inspection by a trained radiologist. Anguh et al[2] claim that this multiscale segmentation and enhancement method detects virtually all lesions identified by an expert radiologist in the process of visual inspection in initial tests on 25 mammograms.

4. Wavelet-Based Enhancement

Koren et al[18] developed a contrast enhancement method based on the adaptation of specific enhancement schemes for distinct mammographic features, which were then used to combine the set of processed images into an enhanced image. In their scheme, the mammo- graphic image is first processed for enhancement of microcalcifications, masses and stellate lesions. From the resulting enhanced image, the final enhanced image is synthesized by means of image fusion[20]. Specifically, their algorithm consisted of two major steps:

1. the image is first subjected to a redundant B-spline wavelet transform decomposition[18] from which a set of wavelet coefficients is obtained
2. the wavelet coefficients are modified distinctly for each type of malignancy (microcal- cifications, stellate lesions or circumscribed masses).
3. the multiple sets of coefficients thus obtained are fused into a single set from which the reconstruction is computed The algorithm is illustrated in Figure 2, as applied to a digitized mammogram that they obtained from the University of Florida database. The theoretical treatment for the mathematics involved in this scheme is beyond the scope of this study. However, it is interesting to note that the enhance image produced by this scheme is "more easily interpreted by a radi- ologist compared to images produced via global enhancement techniques"[18]. It is yet to be seen what improvement this enhancement scheme can contribute to existing CAD schemes.

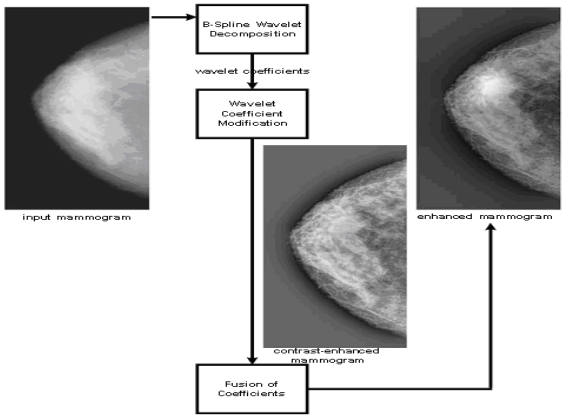


Figure 2. Overview of the image fusion algorithm based on B-Spline Wavelet Transform

5. Classification

Apply Bézier splines[17] to both lesion detection and characterization, where lesion detection is achieved by segmentation using a threshold computed from the Bézier smoothed histogram and lesion characterization is achieved by means of fitness between Gaussian and Bézier histograms of data projected on principal components of the segmented lesions. The most interesting component of their systems in the use of the Bézier splines as a basis of thresholding of the mammographic image - the overall performance of their classification scheme is significantly worse than that seen from, for example, the ANN-based scheme used by Chen et al[2],

Bézier splines are a spline approximation method, developed by the French engineer Pierre Bézier for use in the design of Renault automobile bodies. Since a Bézier curve lies within the convex hull of the control points on which it is fitted, applying it to the histogram of the original image produces a smoothed histogram from which a threshold can be easily chosen by simply finding the largest minimum or the rightmost inflection point, which is where the highest brightness level is located. As a rule, a Bézier curve is a polynomial of degree one less than the number of control points used. Since a typicalgrayscale image consists of 256 brightness levels, the histogram values of these levels can be used as the control points for a Bézier curve polynomial of degree 255. If the histogram levels are denoted by $p_k = (x_k, y_k)$, where both k and x_k vary from 0 to 255, then these coordinate points can be blended to produce a position vector P (u) which describes the path of an approximating Bézier polynomial between p_0 and p_{255}:

$$P(u) = \sum_{k=0}^{255} p_k BEZ_{k,255}(u)$$

where $0 \leq u \leq 1$.

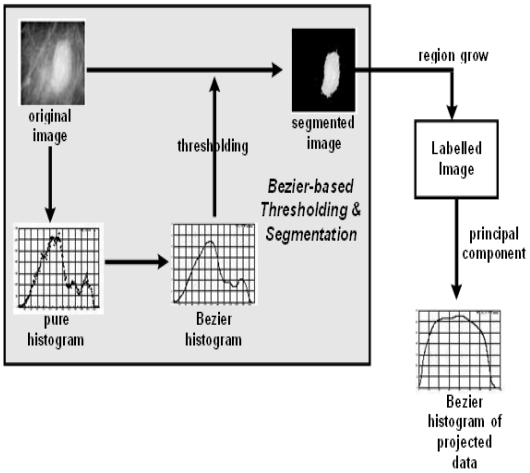


Figure 7. System overview of the Bézier spline-based thresholding and segmentation algorithm.

6. Results

The database used for this work comes from vinayaka missions hospital and is a set of 43 digitalized mammographic images, 25 of them corresponding to benign pathologies and 18 to breast cancer. Due to the good performance of the detection stage, only few microcalcifications are not detected. Detection of the maximum possible number of microcalcificactions is very important for the success of the system, being very critical the correct adjustment of noise thresholds in the contourlet pre-processing stage.

7. Conclusions

The proposed system combines several state of the art image processing techniques, namely contourlet transforms for the noise removal of the mammographic images and border detection with the wavelet transform modulus maxima lines. The tested wavelet based compression method proved to be an accurate approach for digitized mammography.

References

[1] Predrag Baki'c And. Application of neural networks in computer aided diagnosis of breast cancer. http://citeseer.nj.nec.com/404677.html.

[2] M.M. Anguh and A.C. Silva. Multiscale segmentation and enhancement in mammo- grams.

http://www.visgraf.impa.br/sibgrapi97/anais/ART03/ART03.ps.gz.

[3] L. Baker. Breast cancer demonstration project: five year summary report. Cancer, pages 194–225, 1982.

[4] Keir Bovis and Sameer Singh. Detection of masses in mammograms using texture features. http://www.dcs.ex.ac.uk/research/pann/master/cameo/paper80.ps.

[5] Garra B.S., Krasner B.H., Horii S.C., Archer S., Muk S.K., and Zerman R.K. Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis. Ultrasound Imaging, 15:267–285, 1993.

[6] I. Christoyianni, E. Dermatas, and G. Kokkinakis.

[7] Shane Dickson. Investigation of the use of Neural Networks for Computerised Medical Image Analysis. PhD thesis, Department of Computer Science, University of Bristol, February 1998.

[8] Chen D.R., Chang R.F., and Huang Y.L. Computer-aided diagnosis applied to us of solid breast nodules by using neural networks. Radiology, 213(2):407–412, November 1999. http://www.cs.ccu.edu.tw/rfchang/radio9911.pdf.

[9] V. Goldberg and A. Manduca. Improvement in specificity of ultrasonography for diag- nosis of breast tumors by means of artificial intelligence. Medical Physics 1992, 19:1475–1481, 1992.

[10] R.C. Gonzalez and R.C. Woods. Digital Image Processing. Addison Wesley., 1992.

[11] R. Gupta and P.E. Undrill. The use of texture analysis to identify suspicious masses in mammography. http://www.biomed.abdn.ac.uk/Abstracts/A00337/

[12] R. Hanka, T. Harte, A. Dixon, D. Lomas, and P. Britton. Neural networks in the inter- pretation of contrast-enhanced magnetic resonance images of the breast. In Proceedings Current Perspectives in Healthcare Computing Conference, Harrogate, UK, pages 275–283, March 1996.

[13] Trevor Hastie, Debra Ikeda, and Robert Tibshirani. Computer-aided diagnosis of mam- mographic masses. Technical report, Departments of Preventive Medicine & Biostatistics, University of Toronto, June 1996. ftp://utstat.toronto.edu/pub/tibs/mammo.ps.Z.

[14] Trevor Hastie, Debra Ikeda, and Robert Tibshirani. Statistical measures for the computer-aided diagnosis of mammographic masses. Journal of Computational and Graphical Statistics, 8(3), September 1999.

[15] D. Hearn and M.P. Baker. Computer Graphics. Prentice Hall, Inc., 2nd edition, 1994.

[16] Y. Hirose, K. Yamashita, and S. Hijiva. Backpropagation algorithm which varies the number of hidden units. Neural Networks, 4:61–66, 1991.

[17] International Commission on Radiological Units and Measurements. Medical Imaging - The Assessment of Image Quality, 1996.

[18] Andrew Laine Iztok Koren and Fred Taylor. Enhancement via fusion of mammographic features. http://citeseer.nj.nec.com/288745.htm.

[19] Barry Kalman, Stan C. Kwasny, and William R. Reinus. Neural classification of ab- normal tissue in digital

[20] I. Koren, A. Laine, and F. Taylor. Image fusion using steerable dyadic wavelet transform. Proceedings of the IEEE International Conference on Image Processing, 3:232–235, 1995.

[21] A. Laine, J. Fan, and W. Yang. Wavelets for contrast enhancement of digital mam- mography. IEEE Engineering in Medicine and Biology Magazine, 14(5):536–550, 1995. http://citeseer.nj.nec.com/349213.html.

[22] A. Laine, S. Schuler, J. Fan, and W. Huda. Mammographic feature enhancement by multiscale analysis. IEEE Transactions in Medical Imaging, MI-13:725–740, 1994. http://citeseer.nj.nec.com/laine94mammographic.html.

[23] H. Li, K. Liu, and S. Lo. Fractal modeling and segmentation for the enhancement of microcalcifications in digital mammograms. IEEE Trans. Med. Imaging, 16:785–798,1997. http://citeseer.nj.nec.com/li97fractal.html.

[24] S. Liu and E. Delp. Multiresolution detection of stellate lesions in mammograms. In Proceedings of the IEEE International Conference on Image Processing, pages 109–112, October 1997. http://citeseer.nj.nec.com/article/liu97multiresolution.html.

[25] Sheng Liu. Multiresolution detection of spiculated lesions in digital mammograms. http://citeseer.nj.nec.com/179226.html.

[26] Laura N. Mascio, Jose M. Hernandez, and Clinton M. Logan.

[27] A. Mendez, P. Tahoces, M. Lado, M. Souto, and J. Vidal. Computer aided diagnosis: Automatic detection of malignant masses in digitized mammograms. Medical Physics, 25(6):957–964, 1998.

[28] A. Ossen, T. Zamzow, H. Oswald, and E. Fleck.

Segmentation of medical images using neural-network classifiers. In Proceedings of the International Conference on Neural Networks and Expert Systems in Medicine and Healthcare (NNESMED'94), pages 427– 432,1994. http://citeseer.nj.nec.com/ossen94segmentation.html

[29] Riccardo Poli and Guido Valli. Optimum segmentation of medical images with hopfield neural networks. Technical Report CSRP-95-12, University of Birmingham School of Computer Science, October 1995. http://citeseer.nj.nec.com/156628.html.

[30] Qi and Snyder. Lesion detection and characterization i n digital mammography by bézier histograms. http://citeseer.nj.nec.com/348097.html.

[31] Gonzalez R.C. and Woods R.E. Image Compression, pages 312–315. Reading, Mass.: Wesley, 1991.

[32] Guido Valli Riccardo. Neural networks and prior knowledge help the segmentation of medical images. http://citeseer.nj.nec.com/336457.html.

[33] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. Nature, 323:533–536, 1986.

[34] H. Sari-Sarraf, S. S. Gleason, and R. M. Nishikawa. Front-end data reduction in computer-aided diagnosis of mammograms: A pilot study. In SPIE's Medical Imag- ing Conference, February 1999. http://www-ismv.ic.ornl.gov/publications/spie99.pdf.

[35] Sameer Singh and Reem Al-Mansoori. Identification of regions of interest in digital mammograms. Journal of Intelligent Systems, 10(2):183–217, 2000. http://www.dcs.ex.ac.uk/research/pann/pdf/pann SS 005.pdf.

[36] P. Undrill and R. Gupta. Texture analysis and boundary refinement to outline mam- mography masses. IEEE Colloquium (Digest), 072:5/1–5/6, 1996.

[37] Kevin S. Woods. Automated image analysis techniques for digital mammography. http://citeseer.nj.nec.com/woods94automated.html.

Digital Video Watermarking Using Discrete Cosine Transform And Perceptual Analysis

[1]B.R.Darshan, [2]R.Triveni
Department of Ece, Madanapalle Institute of Technology & Sciences
Madanapalle

Abstract

Due to the extensive use of digital media applications, multimedia security and copyright protection has gained tremendous importance. Digital Watermarking is a technology used for the copyright protection of digital applications. In this paper, a comprehensive approach for watermarking digital video is introduced. We propose a hybrid digital video watermarking scheme based on Discrete Cosine Transform (DCT) The video frames are first decomposed using DCT and the binary watermark is embedded In this paper, we describe a visible watermarking scheme that is applied into the host video in the DCT domain. A mathematical model has been developed for that purpose. We have also proposed a modification of the algorithm to make the watermark more robust.

Introduction

The popularity of digital video based applications [1] is accompanied by the need for copyright protection to prevent illicit copying and distribution of digital video. Copyright protection inserts authentication data such as ownership information and logo in the digital media without affecting its perceptual quality. In case of any dispute, authentication data is extracted from the media and can be used as an authoritative proof to prove the ownership. As a method of copyright protection, digital video watermarking [2, 3] has recently emerged as a significant field of interest and a very active area of research. Watermarking is the process that embeds data called a watermark or digital signature into a multimedia object such that watermark can be detected or extracted later to make an assertion about the object. The object may be an image or audio or video. For the purpose of copyright protection digital watermarking techniques must meet the criteria of imperceptibility as well as robustness against all attacks [4-6] for removal of the watermark.

Many digital watermarking schemes have been proposed for still images and videos [7]. Most of them operate on Uncompressed videos, while others embeded watermarks directly into compressed videos. The work on video specific watermarking can be further found in [08-10].

Video watermarking introduces a number of issues not present in image watermarking. Due to inherent redundancy between video frames, video signals are highly susceptible to attacks such as frame averaging, frame dropping, frame swapping and statistical analysis. Video watermarking approaches can be classified into two main categories based on the method of hiding watermark bits in the host video. The two categories are: Spatial domain watermarking where embedding and detection of watermark are performed by directly manipulating the pixel intensity values of the video frame.Transform domain techniques, on the other hand, alter spatial pixel values of the host video according to a pre-determined transform and are more robust than spatial domain techniques since they disperse the watermark in the spatial domain of the video frame making it difficult to remove the watermark through malicious attacks like cropping, scaling, rotations and geometrical attacks. The commonly used transform domain techniques are Discrete Fourier Transform (DFT), the Discrete Cosine Transform (DCT), and the Discrete Wavelet Transform (DWT).

Digital watermarking is defined as a process of embedding data (watermark) into a multimedia object to help to protect the owner's right to that object. The embedded data (watermark) may be either visible or invisible. In visible watermarking of Video, a secondary image (the watermark) is embedded in a primary (host) video such that watermark is intentionally perceptible to a human observer whereas in the case of invisible watermarking the embedded data is not perceptible, but may be extracted/detected by a computer program.

Some of the desired characteristics of visible watermarks are listed below .

- A visible watermark should be obvious in both color and monochrome images.
- The watermark should spread in a large and important area of the image in order to prevent its deletion by clipping.
- The watermark should be visible yet must not significantly obscure the image details beneath it.
- The watermark must be difficult to remove, rather removing a watermark should be more costly and labor intensive than purchasing the image from the owner.
- The watermark should be applied automatically with little human interventionand labor.

There are very few visible watermarking techniques available in current literature.
The IBM Digital Library Organization has used a visible watermarking technique to mark digitized pages of manuscript form the Vatican archive. Rajmohan proposes a visible watermarking technique in DCT domain. He divides the image into different blocks, classifies these blocks by perceptual classification methods as proposed in [5] and modifies the DCT coefficients of host image as follows.

$$X'n = \alpha_n Xn + \beta_n Wn \qquad (1.1)$$

The α_n and β_n coefficients are different for different classes of blocks. Xn are the DCT coefficient of the host image blocks and Wn are the DCT co-efficients of the watermark image block. Here, we propose a visible watermarking scheme that modifies gray values of each pixel of the host image. The modification is based on the local as well as global statistics of the host image. The characteristics of the Human Visual System (HVS) are taken into consideration so that the perceptual quality of the frame is not very much affected.

1. Proposed Watermarking Technique

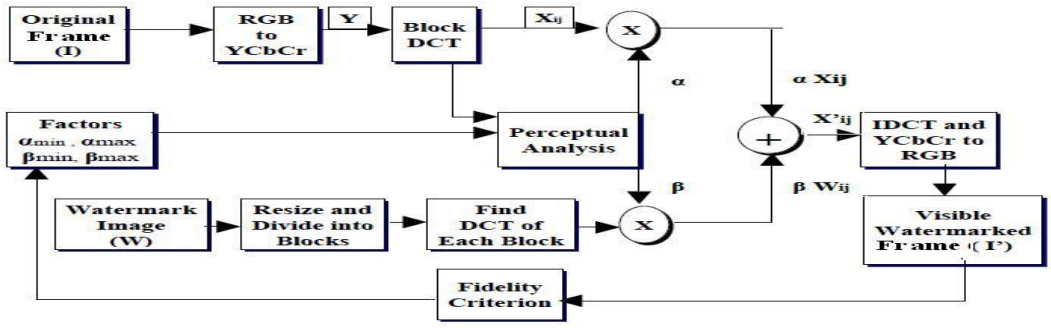


Figure 1.0 Schematic representation of visible watermarking

The steps for watermark insertion are discussed below.

- The original Frame I (one to be watermarked) and the watermark image W are dividedinto blocks (both the images may not be of equal size, but blocks should be of equal size).
- μ and σ , the global mean and variance of the image I are computed.
- For each block the local statistics mean μn and variance σn are computed.
- Let **i**n denote the nth block of original image I ,and **wn** denote the nth block of watermark image W. Denoting the nth block of watermarked image by **in',**

we have,

$$i_{n'} = \alpha_n . i_n + \beta_n W_n \qquad n = 1,2, \ldots\ldots \qquad (1.2)$$

where α_n and β_n are scaling and embedding factors respectively for each bloc Computed as described below.

Fig.1.0 gives the schematic representation of the insertion process.

The choice of α_n's and β_n's are governed by certain characteristics of Human Visual System (HVS) which for watermark images can be translated into following requirements .

- The edge blocks should be least altered to avoid significant distortion of the image. So one can add only small amount of watermark gray value in the edge block of host image.
- This means that scaling factor α_n should be close to α_{max} , (the maximum value of the scaling factor) and embedding factor β_n should be close to β_{min} (the minimum value of the embedding factor).
- It is also pointed out that blocks with uniform intensity (having low variance) are more sensitive to noise than the blocks with non-uniform intensity (having high variance). So one can add less to the blocks with low variance and add more to the blocks with high variance. We assume the scaling factor α_n is inversely proportional to variance whereas β_n directly proportional to variance.
- Yet another characteristic of HVS is that the blocks with mid-intensity are more sensitive to noise than that of low intensity blocks as well as high intensity blocks. This means that the α_n should increase with local mean gray value up to mid gray value and again decrease with local mean gray value. The variation of α_n with mean block gray value is assumed to be gaussian in nature. The variation β_n with mean gray value is reverse to that of α_n .

Basing on the above discussion we propose the following mathematical model.

$$\alpha_n = \begin{cases} \alpha_{max} , & \text{for edge blocks} \\ \alpha_{min} + (\sigma_{min}(\alpha_{max} - \alpha_{min})/\sigma_n) \exp(-((\mu_n - \mu)/\sigma)^2 / 2), & \text{for other blocks} \end{cases}$$

$$\beta_n = \begin{cases} \beta_{min} , & \text{for edge blocks} \\ \beta_{min} + (\sigma_n (\beta_{max} - \beta_{min}) / \sigma_{max}) [1 - \exp(-((\mu_n - \mu)/\sigma)^2 / 2)], & \text{for other blocks} \end{cases}$$

For other blocks
Where,
α_{min} and α_{max} are respectively minimum and maximum values of scaling factor,
β_{min} and β_{max} are respectively minimum and maximum values of embedding factor,
σ_{min} and σ_{max} are respectively minimum and maximum values of block variances,
μ_n and σ_n are respectively normalized mean and variance of each block, and
μ and σ are respectively normalized mean & variances of the frame.

2. Results

Figure 2.1 original frame

Figure 2.2 watermarked image

Figure 2..3 watermarked at different positions

A video clip of size 176X144 is used in our experiments and the watermarked logo can be the variable size. The experiments are done on a desktop computer with Pentium 4 CPU 2.00GHz and 1GB RAM.

The proposed algorithm is applied to a sample video sequence "sea.avi" using a variable watermark logo. Fig 2.1 represents the original video frame is embedded with watermark logo of Fig 2.2 and the resultant watermarked frame obtained at Fig 2.3 respectively. In the Fig 2.3 represents the watermarked logo is embedded at different positions of the frame.

3. Conclusions

A visible watermarking technique has been proposed here in DCT domain. A mathematical model is developed for this purpose exploiting the texture sensitivity of the HVS. The typical values of α_{min} , α_{max} , β_{min}, and β_{max} are 0.95, 0.98, 0.05 and 0.17 respectively. . For more robustness watermark should not be publicly available, the watermark should be used in different sizes and should be put in different portions for different frames. The watermark may find application in digital TV, digital libraries and e-commerce.

References

[1] Yeo and M.M. Yeung, "Analysis and synthesis for new digital video applications," icip, International Conference on Image Processing (ICIP'97),vol. 1, pp.1,1997.

[2] M. Natarajan , G. Makhdumi1,"Safeguarding the Digital Contents: Digital Watermarking," DESIDOC Journal of Library & Information Technology, Vol. 29, May 2009, pp. 29-35.

[3] C.I. Podilchuk, E.J. Delp "Digital watermarking: algorithms and applications," Signal Processing Magazine, Vol 18,pp. 33-46, IEEE, July 2001.

[4] G. Doërr, J.L. Dugelay, "Security Pitfalls of Frame-by-Frame Approaches toVideo Watermarking," Signal Processing, IEEE Transactions, vol. 52, pp. 2955 – 2964, 2004.

[5] M. K. Thakur, V. Saxena, J. P.Gupta, "A Performance Analysis of Objective Video Quality Metrics for Digital Video Watermarking," Computer Science and Information Technology (ICCSIT), 2010, 3rd IEEE International Conference ,Vol. 4, pp. 12- 17,2010.

[6] S. Voloshynovskiy, S. Pereira, T. Pun, "Watermark attacks," Erlangen Watermarking Workshop 99, October 1999.

[7] G. Langelaar, I. Setyawan, and R. Lagendijk, "Watermarking Digital Image and Video Data: A State of - Art Overview," IEEE Signal Processing Magazine, vol. , pp. 20-46, Sep. 2000.

[8] F. Hartung and B. Girod, "Watermarking of uncompressed and compressed video," Signal Processing, 1998,vol. 66, no. 3,pp. 283-301.

[9] T. Khatib, A. Haj, L. Rajab, H. Mohammed, " A Robust Video Watermarking Algorithm", Journal of Computer Science, vol. 4, pp. 910-915, 2008.

[10] T. Tokar, T. Kanocz, D. Levicky, "Digital watermarking of uncompressed video in spatial domain," 9th International Conference on Radioelectronica, IEEE, pp. 319-322, 2009.

Web Personalization using Efficient Ontology Relations

Mohd. Sadik Ahamad [1], **S. Naga Raju** [2]
[1] Kakatiya University, Kakatiya Institute of Technology and Science,
Warangal, Andhra Pradesh, India.
[2]Kakatiya University, Kakatiya Institute of Technology and science,
Assoc.Prof. Department of computer science engineering, Warangal, Andhra Pradesh, India.

Abstract— on the last decades, the amount of web-based information available has increased dramatically. How to gather useful information from the web has become a challenging issue for users. Current web information gathering systems attempt to satisfy user requirements by capturing their information needs. For this purpose, user profiles are created for user background knowledge description. As a model for knowledge description and formalization, ontologies are widely used to represent user profiles in personalized web information gathering. However, when representing user profiles, many models have utilized only knowledge from either a global knowledge base or user local information. In this project, a personalized ontology model is proposed for knowledge representation and reasoning over user profiles.

Keywords— Ontology, Semantic Relations, Web Mining

1. INTRODUCTION

Today, Global analysis uses existing global knowledge bases for user background knowledge representation. Commonly used knowledge bases include generic ontologies e.g., WordNet, thesauruses (e.g., digital libraries), and online knowledge bases (e.g., online categorizations and Wikipedia). The global analysis techniques produce effective performance for user background knowledge extraction. However, global analysis is limited by the quality of the used knowledge base. For example, WordNet was reported as helpful in capturing user interest in some areas but useless for others.

Local analysis investigates user local information or observes user behavior in user profiles. For example, taxonomical patterns from the users' local text documents to learn ontologies for user profiles. Some groups learned personalized ontologies adaptively from user's browsing history. Alternatively, analyzed query logs to discover user background knowledge. In some works, such as, users were provided with a set of documents and asked for relevance feedback. User background knowledge was then discovered from this feedback for user profiles. However, because local analysis techniques rely on data mining or classification techniques for knowledge discovery, occasionally the discovered results contain noisy and uncertain information. As a result, local analysis suffers from ineffectiveness at capturing formal user knowledge. From this, we can hypothesize that user background knowledge can be better discovered and represented if we can integrate global and local analysis within a hybrid model. The knowledge formalized in a global knowledge base will constrain the background knowledge discovery from the user local information. Such a personalized ontology model should produce a superior representation of user profiles for web information gathering.

In this paper, an ontology model to evaluate this hypothesis is proposed. This model simulates users' concept models by using personalized ontologies, and attempts to improve web information gathering performance by using ontological user profiles. The world knowledge and a user's local instance repository (LIR) are used in the proposed model. World knowledge is commonsense knowledge acquired by people from experience and education; an LIR is a user's personal collection of information items. From a world knowledge base, we construct personalized ontologies by adopting user feedback on interesting knowledge. A multidimensional ontology mining method, Specificity and Exhaustivity, is also introduced in the proposed model for analyzing concepts specified in ontologies. The users' LIRs are then used to discover background knowledge and to populate the personalized ontologies. The proposed ontology model is evaluated by comparison against some benchmark models through experiments using a large standard data set.

2. PREVIOUS WORK

Electronic learning (e-Learning) refers to the application of information and communication technologies (e.g., Internet, multimedia, etc.) to enhance ordinary classroom teaching and learning. With the maturity of the technologies such as the Internet and the decreasing cost of the hardware platforms, more institutions are adopting e-Learning as a supplement to traditional instructional methods. In fact, one of the main advantages of e-Learning technology is that it can facilitate *adaptive learning* such that instructors can dynamically revise and deliver instructional materials in accordance with learners' current progress. In general, adaptive teaching and learning refers to the use of what is known about learners, a priori or through interactions, to alter how a learning experience unfolds, with the aim of improving learners' success and satisfaction. The current state-of the- art of e-Learning technology supports automatic collection of learners' performance data (e.g., via online quiz). [1]

However, few of the existing e-Learning technologies can support automatic analysis of learners' progress in terms of the knowledge structures they have acquired. In this paper, we illustrate a methodology of automatically constructing concept maps to characterize learners' understanding for a particular topic; thereby instructors can conduct adaptive

teaching and learning based on the learners' knowledge structures as reflected on the concept maps. In particular, our concept map generation mechanism is underpinned by a context-sensitive text mining method and a fuzzy domain ontology extraction algorithm.

The notion of ontology is becoming very useful in various fields such as intelligent information extraction and retrieval, semantic Web, electronic commerce, and knowledge management. Although there is not a universal consensus on the precise definition of ontology, it is generally accepted that ontology is a formal specification of conceptualization.

Ontology can take the simple form of a taxonomy of concepts (i.e., light weight ontology), or the more comprehensive representation of comprising a taxonomy, as well as the axioms and constraints which characterize some prominent features of the real-world (i.e., heavy weight ontology). Domain ontology is one kind of ontology which is used to represent the knowledge for a particular type of application domain. On the other hand, concept maps are used to elicit and represent the knowledge structure such as concepts and propositions as perceived by individuals. Concept maps are similar to ontology in the sense that both of these tools are used to represent concepts and the semantic relationships among concepts. [1]

However, ontology is a formal knowledge representation method to facilitate human and computer interactions and it can be expressed by using formal semantic markup languages such as RDF and OWL, whereas concept map is an informal tool for humans to specify semantic knowledge structure. Figure shows an example of the owl statements describing one of the fuzzy domain ontologies automatically generated from our system. It should be noted that we use the (rel) attribute of the <rdfs:comment> tag to describe the membership of a fuzzy relation (e.g., the super-class/sub-class relationship). We only focus on the automatic extraction of lightweight domain ontology in this paper. More specificially, the lightweight fuzzy domain ontology is used to generate concept maps to represent learners' knowledge structures.

With the rapid growth of the applications of e-Learning to enhance traditional instructional methods, it is not surprising to find that there are new issues or challenges arising when educational practitioners try to bring information technologies down to their classrooms. The situation is similar to the phenomenon of the rapid growth of the Internet and the World Wide Web (Web). The explosive growth of the Web makes information seekers become increasingly more difficult to find relevant information they really need [1].

This is the so-called problem of information overload. With respect to e-learning, the increasing number of educational resources deployed online and the huge number of messages generated from online interactive learning (e.g., Blogs, emails, chat rooms) also lead to the excessive information load on both the learners and the instructors. For example, to promote reflexive and interactive learning, instructors often encourage their students to use online discussion boards, blogs, or chat rooms to reflect what they have learnt and to share their knowledge with other fellow students during or after normal class time. With the current practice, instructors need to read through all the messages in order to identify the actual progress of their students.

3. PROPOSED SYSTEM

A. Ontology Construction

The subjects of user interest are extracted from the WKB via user interaction. A tool called Ontology Learning Environment (OLE) is developed to assist users with such interaction. Regarding a topic, the interesting subjects consist of two sets: positive subjects are the concepts relevant to the information need, and negative subjects are the concepts resolving paradoxical or ambiguous interpretation of the information need. Thus, for a given topic, the OLE provides users with a set of candidates to identify positive and negative subjects. For each subject, its ancestors are retrieved if the label of contains any one of the query terms in the given topic. From these candidates, the user selects positive subjects for the topic. The user-selected positive subjects are presented in hierarchical form. The candidate negative subjects are the descendants of the user-selected positive subjects. From these negative candidates, the user selects the negative subjects. These positive subjects will not be included in the negative set. The remaining candidates, who are not fed back as either positive or negative from the user, become the neutral subjects to the given topic. Ontology is then constructed for the given topic using these users fed back subjects. The structure of the ontology is based on the semantic relations linking these subjects. The ontology contains three types of knowledge: positive subjects, negative subjects, and neutral subjects.

B. Semantic Specificity

The semantic specificity is computed based on the structure inherited from the world knowledge base. The strength of such a focus is influenced by the subject's locality in the taxonomic structure. The subjects are graph linked by semantic relations. The upper level subjects have more descendants, and thus refer to more concepts, compared with the lower bound level subjects. Thus, in terms of a concept being referred to by both an upper and lower subjects, the lower subject has a stronger focus because it has fewer concepts in its space. Hence, the semantic specificity of a lower subject is greater than that of an upper subject. The semantic specificity is measured based on the hierarchical semantic relations (is-a and part-of) held by a subject and its neighbors. The semantic specificity of a subject is measured, based on the investigation of subject locality in the taxonomic structure. In particular, the influence of locality comes from the subject's taxonomic semantic (is-a and part-of) relationships with other subjects.

C. Topic Specificity

The topic specificity of a subject is performed, based on the user background knowledge discovered from user local information. User background knowledge can be discovered from user local information collections, such as a user's stored documents, browsed web pages, and composed/received emails. The ontology constructed has only subject labels and semantic relations specified. we populate the ontology with the instances generated from user local information collections. We call such a collection the user's local instance repository. The documents may be semistructured (e.g., the browsed HTML and XML web documents). In some semistructured web documents, content-related descriptors are specified in the metadata sections. These descriptors have direct reference to the concepts specified in a global knowledge base. These documents are ideal to generate the instances for ontology population. When different global knowledge bases are used, ontology mapping techniques is used to match the concepts in different representations. The clustering techniques group the documents into unsupervised clusters based on the document features. These features, usually represented by terms, can be extracted from the clusters. The documents can then be classified into the subjects based on their similarity. Ontology mapping techniques can also be used to map the features discovered by using clustering and classification to the subjects, if they are in different representations.

D. Analysis of Subjects

The exhaustivity of a subject refers to the extent of its concept space dealing with a given topic. This space extends if a subject has more positive descendants regarding the topic. In contrast, if a subject has more negative descendants, its exhaustivity decreases. Based on this, we evaluate a subject's exhaustivity by aggregating the semantic specificity of its descendants where Subjects are considered interesting to the user only if their specificity and exhaustivity are positive. A subject may be highly specific but may deal with only a limited semantic extent.

4. RESULTS

The concept of this paper is implemented and different results are shown below, The proposed paper is implemented in Java technology on a Pentium-IV PC with 20 GB hard-disk and 256 MB RAM with apache web server. The propose paper's concepts shows efficient results and has been efficiently tested on different Datasets. The Fig 1, Fig 2, Fig 3 and Fig 4 shows the real time results compared.

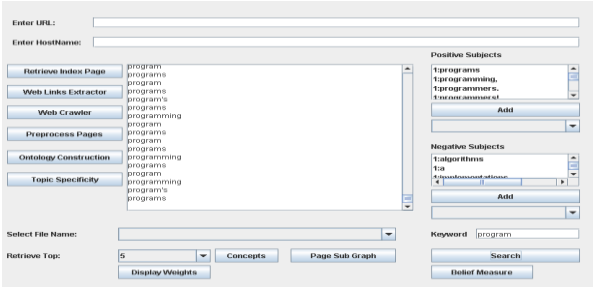


Fig. 1 Computation of Positive and Negative Subjects.

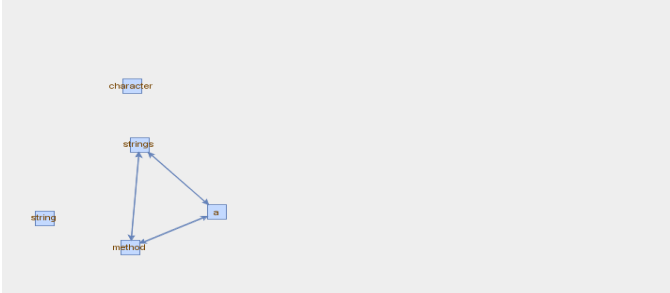


Fig. 1 Displaying Constructed Ontology

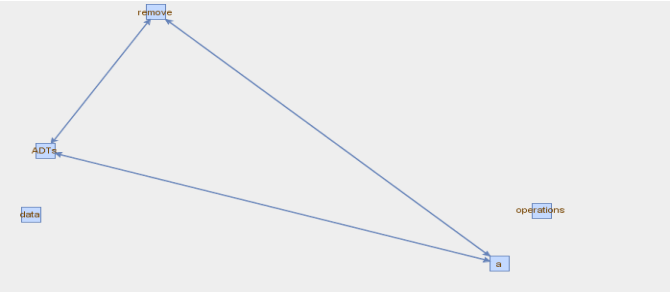


Fig. 3 Displaying Constructed Ontology

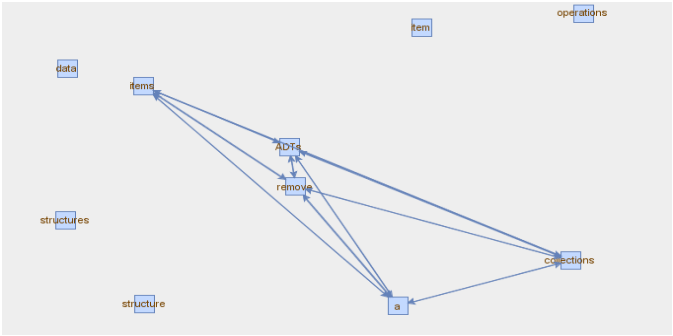


Fig. 4 Displaying Constructed Ontology

5. SYNONYM HANDLING

When we handle the retrieved ontology keywords we would drag the semantic relationships between the instance sets of subject headings. Although we get the results related to user knowledge but there may be a chance of losing data because of the synonyms. Sometimes synonyms of keywords may give us the results better that user expected.

For this reason synonyms of keywords retrieved and maintained later by using all these words we form the instance sets and retrieve more subject headings from LCSH and add to LIR.

We can derive the probability of the results before synonym handling case and after synonym handling case. For example if we got **M** words without synonym case, and the probability is **P1.** For the synonym case if we got **N** words, and calculated probability is **P2**. If we compare these two probabilities, definitely

P1 < P2 (M<N)

Finding the right direction for searching ontology related words is difficult. Ontology is vast and there could be many directions the user may not able to find the relevant results of interest according to him. The problem becomes bigger if we consider the synonyms of the words. To find the more related suitable synonyms and words we find the probability of result set for each synonym and compare them with the existing results. We consider only the synonyms that gives more results according to user and the user interest. With this approach we can refine the search.

If we could drag the relationship between the resultset and the no of synonym words added, through probability then we can predict the results for other cases. This helps to formulate the analysis. By taking some small cases we can do that and it helps to solve and predict complex cases.

6. CONCLUSION

In this project, an ontology model is proposed for representing user background knowledge for personalized web information gathering. The model constructs user personalized ontologies by extracting world knowledge from the LCSH system and discovering user background knowledge from user local instance repositories.

A multidimensional ontology mining method, exhaustivity and specificity, is also introduced for user background knowledge discovery. In evaluation, the standard topics and a large testbed were used for experiments. The model was compared against benchmark models by applying it to a common system for information gathering. The experiment results demonstrate that our proposed model is promising. Sensitivity analysis was also conducted for the ontology model. In this investigation, we found that the combination of global and local knowledge works better than using any one of them. In addition, the ontology model using knowledge with both is-a and part-of semantic relations works better than using only one of them. When using only global knowledge, these two kinds of relations have the same contributions to the performance of the ontology model. While using both global and local knowledge, the knowledge with part-of relations is more important than that with is-a. The proposed ontology model in this project provides a solution to emphasizing global and local knowledge in a single computational model. The findings in this project can be applied to the design of web information gathering systems. The model also has extensive contributions to the fields of Information Retrieval, web Intelligence, Recommendation Systems, and Information Systems. Synonyms will give us more directions to choose user interests. It refines the search.

REFERENCES

[1] R.Y.K. Lau, D. Song, Y. Li, C.H. Cheung, and J.X. Hao, "Towards a Fuzzy Domain Ontology Extraction Method for Adaptive e- Learning," IEEE Trans. Knowledge and Data Eng., vol. 21, no. 6, pp. 800-813, June 2009.

[2] T. Tran, P. Cimiano, S. Rudolph, and R. Studer, "Ontology-Based Interpretation of Keywords for Semantic Search," Proc. Sixth Int'l Semantic Web and Second Asian Semantic Web Conf. (ISWC '07/ ASWC '07), pp. 523-536, 2007.

[3] C. Makris, Y. Panagis, E. Sakkopoulos, and A. Tsakalidis, "Category Ranking for Personalized Search," Data and Knowledge Eng., vol. 60, no. 1, pp. 109-125, 2007.

[4] S. Shehata, F. Karray, and M. Kamel, "Enhancing Search Engine Quality Using Concept-Based Text Retrieval," Proc. IEEE/WIC/ ACM Int'l Conf. Web Intelligence (WI '07), pp. 26-32, 2007.

[5] A. Sieg, B. Mobasher, and R. Burke, "Web Search Personalization with Ontological User Profiles," Proc. 16th ACM Conf. Information and Knowledge Management (CIKM '07), pp. 525-534, 2007.

[6] D.N. Milne, I.H. Witten, and D.M. Nichols, "A Knowledge-Based Search Engine Powered by Wikipedia," Proc. 16th ACM Conf. Information and Knowledge Management (CIKM '07), pp. 445-454, 2007.

[7] D. Downey, S. Dumais, D. Liebling, and E. Horvitz, "Understanding the Relationship between Searchers' Queries and Information Goals," Proc. 17th ACM Conf. Information and Knowledge Management (CIKM '08), pp. 449-458, 2008.

[8] M.D. Smucker, J. Allan, and B. Carterette, "A Comparison of Statistical Significance Tests for Information Retrieval Evaluation," Proc. 16th ACM Conf. Information and Knowledge Management (CIKM '07), pp. 623-632, 2007.

[9] R. Gligorov, W. ten Kate, Z. Aleksovski, and F. van Harmelen, "Using Google Distance to Weight Approximate Ontology Matches," Proc. 16th Int'l Conf. World Wide Web (WWW '07), pp. 767-776, 2007.

[10] W. Jin, R.K. Srihari, H.H. Ho, and X. Wu, "Improving Knowledge Discovery in Document Collections through Combining Text Retrieval and Link Analysis Techniques," Proc. Seventh IEEE Int'l Conf. Data Mining (ICDM '07), pp. 193-202, 2007.

AUTHORS

Mohd. Sadik Ahamad received the B.E. degree in computer science Engineering from Muffakham jha College of Engineering and Technology, Osmania University in 2008 and M.Tech. Software engineering from kakatiya institute of technology and science, affiliated to Kakatiya University in 2012. During 2009-2010, he worked as an Asst. Prof in Balaji institute of technology and science.

S. Naga Raju Assoc. Prof. of Department of computer science engineering, kakatiya institute of technology and science, Warangal, researcher in web mining, member of ISTE, B.Tech. from Amaravati University, M.Tech from JNTUH.

Heat and Mass Transfer with Variable Temperature and Exponential Mass Diffusion

[1]I. J. Uwanta and [2]M. N. Sarki

[1]Department of mathematics, Usmanu Danfodiyo University Sokoto. Nigeria.
[2]Department of mathematics, Kebbi State University of Science and Technology, Aliero. Nigeria

Abstract
In this paper an analytical study is performed to study heat and mass transfer with variable temperature and exponential mass diffusion, the results were obtained for velocity, temperature and concentration, the dimensionless governing equations are tackled by the Laplace transform method, and computed for parameters namely thermal Grashof number Gr, mass Grashof number Gc, Schmidt number Sc, Prandtl number Pr, time t, and acceleration a . It is observed that the velocity increases with increasing values of Gr, Gc, a and t, It was also observed that velocity, temperature and concentration decreases with increasing Pr and Sc respectively.

Key Word: exponential, mass transfer, variable temperature, mass diffusion.

1 Introduction

Heat and mass transfer plays an important role in drying, filtration processes, saturation of porous materials by chemicals, solar energy collectors, nuclear reactors, in manufacturing industries for the design fins, steel, rolling, nuclear power plants, gas turbines and various propulsion devices for aircraft, missiles, space craft design, satellites, combustion and furnace design, material processing, energy utilization, temperature measurement, remote sensing for astronomy and space exploration, food processing and cryogenic Engineering, as well as numerous agricultural, health and military application. The study of convection flow with mass transfer along a vertical plate is receiving considerable attention to many researchers because of its vast application in the field of cosmically and geophysical science. England and Emery (1969) have studied the thermal radiation effects of an optically thin gray gas bounded by a stationary vertical plate, Gupta et al. (1979) have studied free convective effects flow past accelerated vertical plate in incompressible dissipative fluid, Mass transfer and free convection effects on the flow past an accelerated vertical plate with variable suction or injection, was studied by Kafousia and Raptis (1981), Jha et al. (1991)analyzed mass transfer effects on exponentially accelerated infinite vertical plate with constant heat flux and uniform mass diffusion. Raptis and Perdikis (1999) analyzed the Radiation and free convection flow past a moving plate, Chamkha and Soundalgekar (2001) have analyzed radiation effects on free convection flow Past a semi-infinite vertical plate with mass transfer, Chaudhary and Jain (2006) analyzed Influence of fluctuating surface temperature and velocity on medium with heat absorption, Toki (2006) studied unsteady free convective flow on a vertical oscillating porous plate with heat , Alam et al. (2008) have analyzed the effects of variable suction and thermophoresis on steady MHD combined free – forced convective heat and mass transfer flow over a semi-infinite permeable inclined plate in the presence of thermal radiation, Muthucumaraswamy et al. (2009) have studied the exact solution of flow past an accelerated infinite vertical plate with heat and mass flux. It is proposed to study heat and mass transfer with variable temperature and exponential mass diffusion. The dimensionless governing equations are solved using the Laplace transform technique.
The solutions are in terms of exponential and complementary error function.

2 Problem Formulation:

Governing equation for Heat and mass transfer with variable temperature and exponential mass diffusion. Then under usual Boussinesq's approximation the unsteady flow equations are presented as momentum equation, energy equation, and mass equation respectively.

$$\frac{\partial u}{\partial t'} = g\beta(T - T_\infty) + g\beta^*(C' - C'_\infty) + \nu\frac{\partial^2 u}{\partial y^2} \tag{1}$$

$$\rho C_\rho \frac{\partial T}{\partial t'} = K\frac{\partial^2 T}{\partial y^2} + q_o(T - T_\infty) \tag{2}$$

$$\frac{\partial C'}{\partial t'} = D\frac{\partial^2 C'}{\partial y^2} - KC' \tag{3}$$

The initial and boundary conditions are:

$$\left.\begin{array}{llll} U=0 & T=T_\infty & C'=C'_\infty & for\ all\ \ y,t'\le 0 \\ t'>0:\ \ U=u_0t' & T=T_\infty+(T_w-T_\infty)At' & C'=C'_\infty+(C'_w-C'_\infty)e^{a't'} & at\quad y=0 \\ U\to 0 & T\to T_\infty & C'\to C'_\infty & as\quad y\to\infty \end{array}\right\} \quad (4)$$

$$where \qquad A=\frac{u_0^{\,2}}{\nu}$$

Where u is the velocity of the fluid, T is the fluid temperature, C' is the concentration, g is gravitational constant, β *and* β^* are the thermal expansion of fluid, t' is the time, ρ is the fluid density, C_ρ is the specific heat capacity, V is the velocity of the fluid, k is the thermal conductivity. The non-dimensional quantities are:

$$\left.\begin{array}{l} U=\frac{u}{u_0},\quad t=\frac{t'u_0^{\,2}}{\nu},\quad Y=\frac{yu_0}{\nu},\quad \theta=\frac{T-T_\infty}{T_w-T_\infty} \\ \Pr=\frac{\mu C_p}{k},\ \ a=\frac{a'\nu}{u_0^{\,2}},\ Sc=\frac{\nu}{D},\ F=\frac{q_0u_0^{\,2}}{k},\ R=\frac{k\nu}{u_0^{\,2}}, \\ Gr=\frac{g\beta\nu(T_w-T_\infty)}{u_0^{\,3}},\quad C=\frac{C'-C'_\infty}{C'_w-C'_\infty},\quad Gc=\frac{\nu g\beta^*(C'_w-C'_\infty)}{u_0^{\,3}}, \end{array}\right\} \quad (5)$$

Substituting the non-dimensional quantities of (5) in to (1) to (4) leads to dimensionless equations as:

$$\frac{\partial u}{\partial t}=Gr\theta+GcC+\frac{\partial^2u}{\partial y^2} \quad (6)$$

$$\frac{\partial\theta}{\partial t}=\frac{1}{\Pr}\frac{\partial^2\theta}{\partial y^2}-\frac{F\theta}{\Pr} \quad (7)$$

$$\frac{\partial C}{\partial t}=\frac{1}{Sc}\frac{\partial^2C}{\partial y^2}-RC \quad (8)$$

Where Sc is the Schmidt number, Pr is Prandtl number, and Gr, Gc are the Grashof numbers, F is the heat source, R is the Concentration parameter.

The initial and boundary conditionsare reduces to:

$$\left.\begin{array}{llll} U=0, & \theta=0, & C=0, & for\ all\quad y,t\le 0 \\ t>0:\ \ U=t, & \theta=t, & C=e^{at}, & at\quad y=0 \\ U\to 0, & \theta\to 0, & C\to 0, & as\quad y\to\infty \end{array}\right\} \quad (9)$$

3 Method Of Solution

The dimensionless governing equations (6) to (8) with initial boundary conditions are solved using Laplace transform techniques and the results for temperature, concentration and velocity in terms of exponential and complementary error function:

$$L(\theta)=\frac{e^{-y\sqrt{(S\Pr+F)}}}{s^2} \quad (10)$$

$$L(C)=\frac{e^{-y\sqrt{Sc(s+R)}}}{s-a} \quad (11)$$

$$L(U)=\frac{e^{-y\sqrt{s}}}{s^2}+\frac{Gr(e^{y\sqrt{s}}-e^{-y\sqrt{(S\Pr+F)}})}{d^2(1-\Pr)(s-d)}-\frac{Gr(e^{-y\sqrt{s}}-e^{-y\sqrt{(S\Pr+F)}})}{d(1-\Pr)s^2}$$
$$-\frac{Gc(e^{-y\sqrt{s}}-e^{-y\sqrt{Sc(s+R)}})}{(1-Sc)(a-b)(s-a)}-\frac{Gc(e^{-y\sqrt{s}}-e^{-y\sqrt{Sc(s+R)}})}{(1-Sc)(b-a)(s-b)} \quad (12)$$

The Laplace inversion gives,

$$\theta=\frac{t}{2}\left[(2\eta\sqrt{Ft})erfc(\eta\sqrt{\Pr}+\sqrt{ct})+\exp(-2\eta\sqrt{Ft})\ erfc(\eta\sqrt{\Pr}-\sqrt{ct})\right] \quad (13)$$
$$-\frac{\eta\Pr\sqrt{t}}{2\sqrt{Ft}}\left[\exp(-2\eta\sqrt{Ft})\ erfc(\eta\sqrt{\Pr}-\sqrt{ct})+\exp(2\eta\sqrt{Ft})\ erfc(\eta\sqrt{\Pr}+\sqrt{ct})\right]$$

$$C=\frac{\exp(at)}{2}\left[\exp\left(2\eta\sqrt{Sc(a+R)t}\right)erfc\left(\eta\sqrt{Sc}+\sqrt{(a+R)t}\right)+\exp\left(-2\eta\sqrt{Sc(a+R)t}\right)erfc\left(\eta\sqrt{Sc}-\sqrt{(a+R)t}\right)\right] \quad (14)$$

$$U = t\left[(1+2\eta^2)ercf\,\eta - \frac{2\eta\exp(-\eta^2)}{\pi}\right] - \frac{Gr\exp(dt)}{2d^2(1-\Pr)}\left[\exp\left(2\eta\sqrt{dt}\right)erfc\left(\eta+\sqrt{dt}\right)+\exp\left(-2\eta\sqrt{dt}\right)erfc\left(\eta-\sqrt{dt}\right)\right]$$

$$+\frac{Gr\exp(dt)}{2d^2(1-\Pr)}\left[\exp\left(2\eta\sqrt{\Pr(d+F)t}\right)erfc\left(\eta\sqrt{\Pr}+\sqrt{(d+F)t}\right)+\exp\left(-2\eta\sqrt{\Pr(d+F)t}\right)erfc\left(\eta\sqrt{\Pr}-\sqrt{(d+F)t}\right)\right]$$

$$-\frac{Gr\,t}{d(1-\Pr)}\left[(1+2\eta^2)ercf\,\eta-\frac{2\eta\exp(-\eta^2)}{\pi}\right]$$

$$+\frac{Gr\,t}{2d(1-\Pr)}\left\{\begin{array}{l}\left[\exp\left(2\eta\sqrt{Ft}\right)\,erfc\left(\eta\sqrt{\Pr}+\sqrt{ct}\right)+\exp\left(-2\eta\sqrt{Ft}\right)\,erfc\left(\eta\sqrt{\Pr}-\sqrt{ct}\right)\right]\\ -\frac{\eta\Pr\sqrt{t}}{2\sqrt{Ft}}\left[\exp\left(-2\eta\sqrt{Ft}\right)\,erfc\left(\eta\sqrt{\Pr}-\sqrt{ct}\right)+\exp\left(2\eta\sqrt{Ft}\right)\,erfc\left(\eta\sqrt{\Pr}+\sqrt{ct}\right)\right]\end{array}\right\}$$

$$-\frac{Gc\exp(at)}{2(1-Sc)(a-b)}\left[\exp\left(2\eta\sqrt{at}\right)erfc\left(\eta+\sqrt{at}\right)+\exp\left(-2\eta\sqrt{at}\right)erfc\left(\eta-\sqrt{at}\right)\right]$$

$$+\frac{Gc\exp(at)}{2(1-Sc)(a-b)}\left[\exp\left(2\eta\sqrt{Sc(a+R)t}\right)erfc\left(\eta\sqrt{Sc}+\sqrt{(a+R)t}\right)+\exp\left(-2\eta\sqrt{Sc(a+R)t}\right)erfc\left(\eta\sqrt{Sc}-\sqrt{(a+R)t}\right)\right]$$

$$-\frac{Gc\exp(bt)}{2(1-Sc)(b-a)}\left[\exp\left(2\eta\sqrt{bt}\right)erfc\left(\eta+\sqrt{bt}\right)+\exp\left(-2\eta\sqrt{bt}\right)erfc\left(\eta-\sqrt{bt}\right)\right]$$

$$+\frac{Gc\exp(bt)}{2(1-Sc)(b-a)}\left[\exp\left(2\eta\sqrt{Sc(b+R)t}\right)erfc\left(\eta\sqrt{Sc}+\sqrt{(b+R)t}\right)+\exp\left(-2\eta\sqrt{Sc(b+R)t}\right)erfc\left(\eta\sqrt{Sc}-\sqrt{(b+R)t}\right)\right] \quad (15)$$

$$where \qquad d=\frac{F}{(1-\Pr)}, \quad b=\frac{ScR}{(1-Sc)}, \quad c=\frac{F}{\Pr}, \quad \eta=\frac{y}{2\sqrt{t}}.$$

4. Results and Discussion

The problem of heat and mass transfer has been formulated, analyzed and solved analytically, for physical understanding to the problems numerical computations were carried out for different physical parameters such asthermal Grashof number Gr, mass Grashof number Gc, Schmidt number Sc, Prandtl number Pr, time t, and acceleration a, upon the nature of flow and transport, the value of the Schmidt number Sc is taken to be 0.6 which corresponds to water-vapor, also the value of Prandtl number Pr are chosen such that they represent air (Pr=0.71). It is observed that the velocity increases with increasing values of Gr, Gc, and a.To access the effects of the various parameters in the flow fields, graphs are presented as follows:

4.1 Velocity profiles

Figures 1 to 6 represent velocity profile for the flow

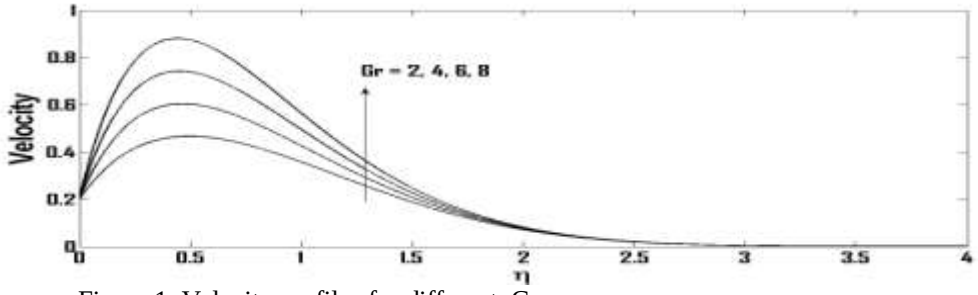


Figure 1: Velocity profiles for different Gr

The velocity profiles for different values of thermal Grashof number (Gr=2, 4, 6, 8) is presented in figure 1. It observed that velocity increases with increasing Gr.

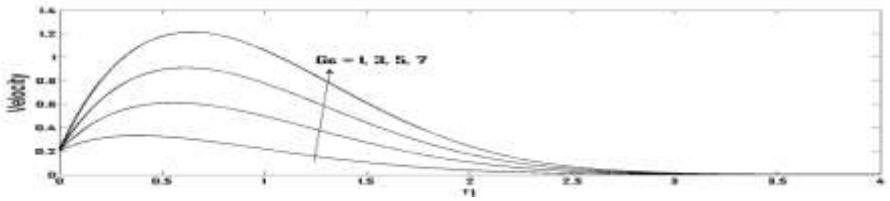


Figure 2 : Velocity profiles for different Gc

The velocity profiles for different values of mass Grashof number (Gc=1, 3, 5, 7) is presented in figure 2. It observed that velocity increases with increasing Gc.

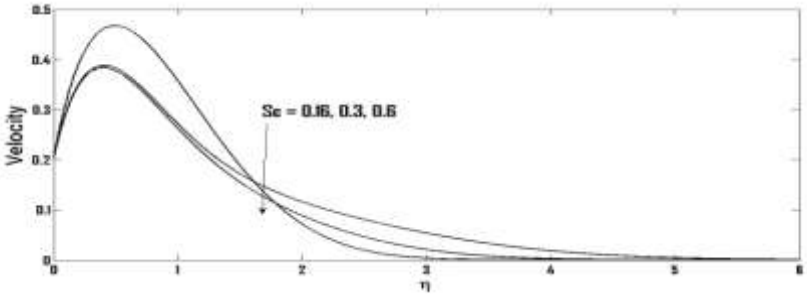


Figure 3 : Velocity profiles for different Sc

The velocity profiles for different values of Schmidt number (Sc= 0.16, 0.3, 0.6) is presented in figure 3. It observed that velocity decreases with increasing Sc.

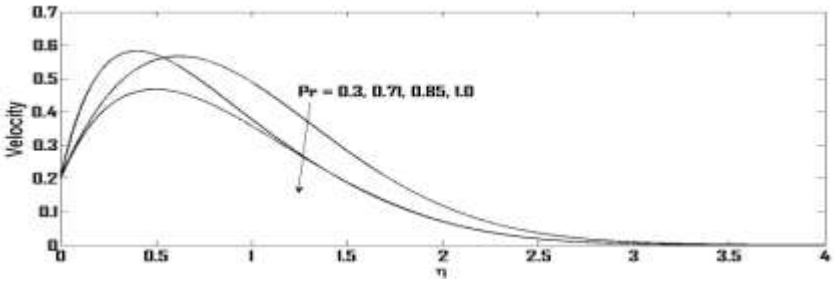


Figure 4 : Velocity profiles for different Pr

The velocity profiles for different values of Prandtl number (Pr= 0.3, 0.71, 0.85, 1.0) is presented in figure 4. It observed that velocity decreases with increasing Pr.

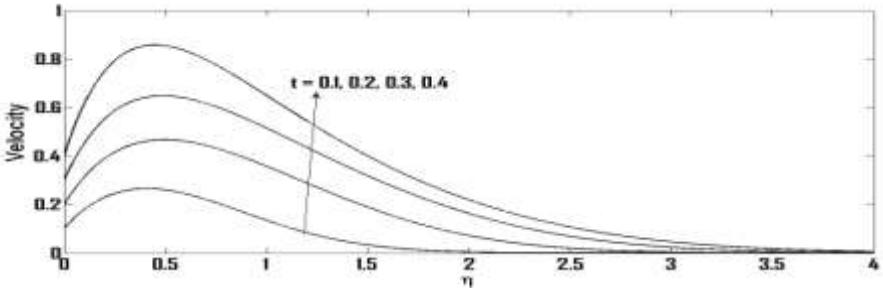


Figure 5 : Velocity profiles for different t

The velocity profiles for different values of time (t= 0.1, 0.2, 0.3, 0.4) is presented in figure 5. It observed that velocity increases with increasing t.

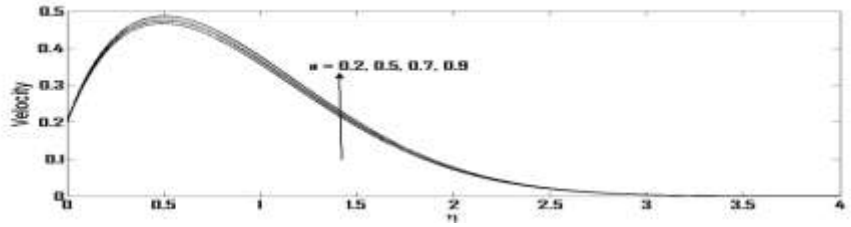


Figure 6 : Velocity profiles for different a .

The velocity profiles for different values of a (a= 0.2, 0.5, 0.7, 0.9) is presented in figure 6. It observed that velocity increases with increasing a.

4.2 Temperature profiles

Figures 7 and 8 represent temperature profiles for the flow

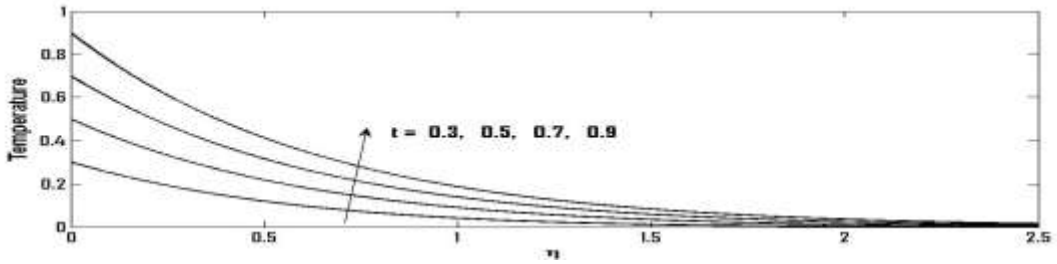


Figure 7: Temperature profiles for different t

The temperature profiles for different values of time (t=0.3, 0.5, 0.7, 0.9) is presented in figure 8. It is observed that temperature increases with increasing t.

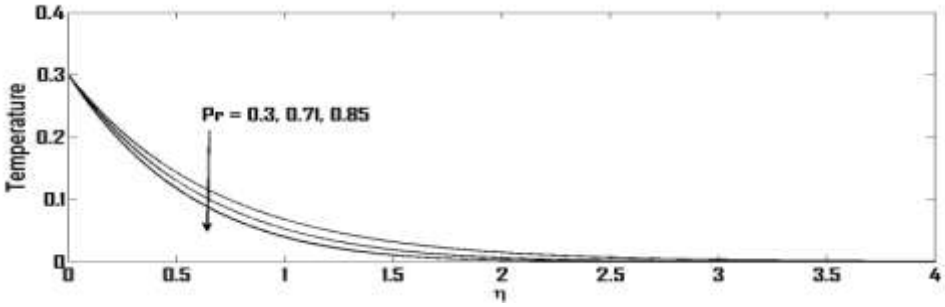


Figure 8 Temperature profiles for different Pr

The temperature profiles for different values of prandtl number (Pr=0.3, 0.71, 0.85) is presented in figure 8. It is observed that temperature decreases with increasing Pr.

4.3 Concentration profiles

Figures 9 and 10 represent concentration profiles for the flow

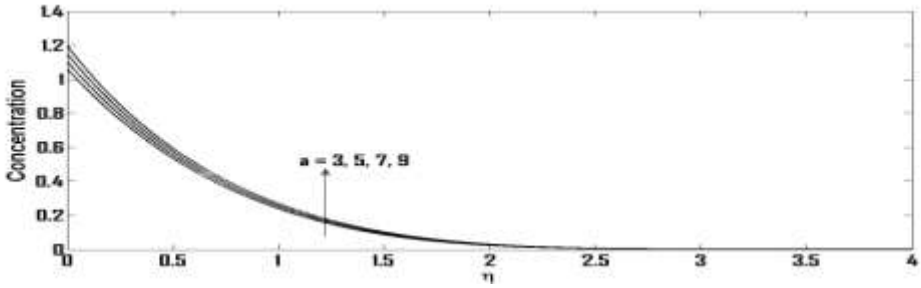


Figure 9: Concentration profiles for different a

The concentration profiles for different values of a (a=3, 5, 7, 9) is presented in figure 9. It is observed that concentration increases with increasing a.

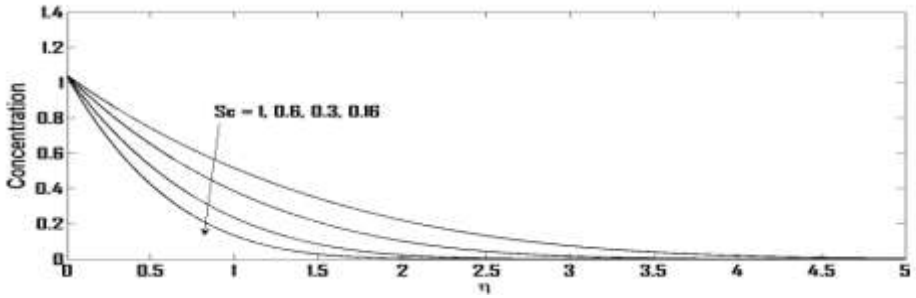


Figure 10: Concentration profiles for different Sc

The concentration profiles for different values of Schmidt number (Sc=1, 0.6, 0.3, 0.16) is presented in figure 10. It is observed that concentration decreases with increasing Sc.

Conclusion:

Analytical solutions of heat and mass transfer with variable temperature and exponential mass diffusion have been studied. The dimensional governing equations are solved by Laplace transform technique. The effect of different parameters such as Schmidt number, Prandtl number, mass Grashof number, thermal Grashof number, and time are presented graphically. It is observed that velocity profile increases with increasing parameter a, t, Gc, and Gr and also decreases with increasing Sc and Pr respectively, it is also observed that temperature and concentration profile increases with increasing t, and inversely, decreases as Sc and Pr increases respectively.

References

Alam, M. S., Rahaman, M. M., and Satar, M. A. (2008). ''Effects of Variable Suction and Thermophoresis on Steady MHD Combined Free – Forced convective Heat and mass Transfer Flow over a Semi Infinite Permeable Inclined Plate in the Presence of Thermal Radiation''.

1. *Internati onal Journal of Thermal Sciences.47(6).758-765.*
2. Abramo witz M. and Stegun I. A. (1972) ''Handbook of Mathematical functions'' Dover Publication, inc, New York.
3. Basanth Kumar Jha., Ravindra Prasad.andSurendraRai,(1991)''Mass
4. Transfer Effects on the Flow Past an Exponentially Accelerated Vertical Plate With Constant Heat Flux''. Astrophysics and
5. Space Science, 181.125-134.
6. Basanth Kumar Jha., Ravindra Prasad. (1990). ''Free Convective
7. Mass Transfer on the Flow Past an Accelerated Plate with Heat
8. Source''.Mechanics Research Communications. 17. 143-148.
9. Hossain M.A. and Shayo L.K.(1986).''The Skin Friction in the Unsteady
10. Free Convection Flow Past an Accelerated Plate''.
11. Astrophysics and Space Science.125: 315-324.
12. Chamkha, A. J., Takhar, H. S., and Soundalgekar, V. M.(2001)
13. ''Radiation Effects on Free Convection Flow Past a Semi
14. Infinite Vertical Plate With Mass Transfer''.
15. Chemical Engineering Science 84.335-342.
16. Chaudhary, R. C., and Jain, P.(2006) ''Influence of Fluctuating Surface
17. Temperature and Velocity on Medium With Heat Absorption''.
18. *Journal of Technical Physics, 47(4).239-254.*
19. Das, U. N., Deka, R. K., and Soundalgekar, V. M.(1999). ''Transient free
20. Convective Flow Past an Infinite Vertical Plate With
21. Temperature Variation''.*ASME Journal of Heat Transfer.*
22. *121. 1021-1097.*

23. England W. G., and Emery A. F.(1969). ''Thermal Radiation Effects on the
24. Laminar Free Convection Boundary Layer of an Absorbing Gas''
25. *Journal of Heat Transfer. 91. 37-44.*
26. Gupta, A. S., Pop, I., and soundalgekar, V. M.(1979).''Free Convective
27. Effects on Flow Past Accelerated Vertical Plate in
28. Incompressible Dissipative Fluid''.
29. Rev. Roum. Science Tech-Mec.Apl. 24. 561-568.
30. Jha B. K., Prasad R., and Rai S.(1991).''Mass Transfer Effects on
31. Exponentially Accelerated Infinite Vertical Plate With Constant
32. Heat Flux''.Astrophysics and Space Science. 181. 125-134.
33. Kai-Long Hsiao.(2010). ''Heat Mass for Viscous Flow With Radiation
33.1.1. Effects Past a Non-Linearly Stretching Sheet''. World Academy
34. Of Science Engineering and Technology.62(2).12-21.
35. Kafousias, N. G., and Williams,, E. W.(1995). ''Thermal Diffusion and
36. Diffusion Thermo Effects on Mixed Free-Forced Convective and
37. Mass Transfer Boundary Layer Flow With Temperature
38. Dependent Viscosity''.*International Journal of Engineering*
39. *Science.33(9).1369-1384.*
40. Kafousias, N. G., and Raptis A. A.(1981). ''Mass Transfer and Free
41. Convection Effects on the Flow Past an Accelerated Vertical
42. Plate With Variable Suction or Injection''. 26. 11-22.
43. Muthucumaraswamy R., Sundar Raj M. and Subramanian V.S.A. (2009) ''
44. Exact Solution of Flow Past an Accelerated Infinite Vertical
45. Plate With Heat and Mass Flux'',*International Journal of*
46. *Applied Mechanics and Engineering 14.585–592.*
47. Raptis, A., and Massalas, C. V. (1998). ''Magneto hydrodynamic Flow
48. Past by the Presence of Radiation''.
49. Heat andMass Transfer. 34. 107-109
50. Soundalgekar, V. M.(1982) ''Effects of Mass Transfer on Flow Past a
51. Uniformly Accelerated Vertical plate''. Letters in Heat and
52. Mass Transfer 9. 65-72.
53. Soundalgekar, V. M.(1965) ''Hydro magnetic Flow Near an Accelerated
54. Plate in the Presence of a Magnetic Field''. Applied Scientific
55. Research 12(1).152-156.
56. Soundalgekar, V. M.(1971). Unsteady Free Convective Flow of an Elastic
57. -Visco Fluid Past an infinite Plate With Suction''.
58. Chemical Engineering Science 26.2043-2053.
59. Stokes, G. G.(1851). ''On the Effects of Internal Frictions of Fluid on the
60. Motion of Pendulum''.Transactions Cambridge
61. Philosophical society. 9. 8-106.
62. Toki, C. J.(2006) ''Unsteady Free Convective Flow on a Vertical
63. Oscillating Porous Plate With Heat''. Journal of Applied
64. Mechanics, 76.4-8.
65. Verma, A. V. (1983). ''An Asymptotic Analysis of Free Convection
66. Boundary layer on a Horizontal Plate due to Small
67. Fluctuation in Surface Temperature''. International
68. *Journal of Engineering Science, 21(1). 35-43*

69. 6 Abbreviations

70. C' Species concentration in the fluid $kg: m^{-3}$
71. C dimensionless concentration
72. C_p Specific heat at constant pressure $J:k\,g^{-1}:K$
73. D mass diffusion coefficient m^2, s^{-1}

74. Gc mass Grashof number
75. Gr thermal Grashof number
76. g acceleration due to gravity $m\,s^{-2}$
77. k thermal conductivity $W: m^{-1}s^{-1}$
78. Pr Prandtl number
79. Sc Schmidt number
80. T temperature of the fluid near the plate K
81. t' times
82. t dimensionless time
83. u veocity of the fluid in the x'-direction $m\,s^{-1}$
84. u_0 velocity of the plate $m\,s^{-1}$
85. u dimensionless velocity
86. y coordinate axis normal to the plate m
87. Y dimensionless coordinate axis normal to the plate
88. α thermal diffusivity $m^{-2}s^{-1}$
89. β volumetric coefficient of thermal expansion k^{-1}
90. β^* volumetric coefficient of expansion with concentration k^{-1}
91. μ coefficient of viscosity $Ra.s$
91.1. kinematic viscosity $m^{-2}s^{-1}$
92. ρ density of the fluid $kg\,m^{-3}$
93. T dimensionless skin-friction $kg, m^{-1}s^{-2}$
94. θ dimensionless temperature
95. η similarity parameter
96. $ercf$ complementary error function

Development of Efficient Decoding Technique for Convolutionaly Encoded Telemetry Data

Namratha M[1], Pradeep[2]

[1,2] (Dept. of information science, PESIT/visvesvaraya technological university, India)

Abstract

Telemetry system of spacecraft is meant for radioing information from a spacecraft to the ground. Telemetry is typically a mixture of payload or housekeeping health data. Engineering or health data is composed of a wide range of measurements, from switch positions and subsystem states to voltages, temperatures and pressures. Telemetry may be transmitted in real time, or it may be written to a data storage device until transmission is feasible. As the spacecraft health monitoring is very important to maintain the spacecraft in condition, Telemetry system becomes mission critical system in the spacecraft. Packet Telemetry and Telemetry Channel Coding services provide to the user reliable and transparent delivery of telemetry information. With the error detecting and correcting capability of the channel code chosen, errors which occur as a result of the physical transmission process may be detected and corrected by the receiving entity. The implementation of convolution encoder is simple, this means these encoders are small and consumes less power on-board spacecraft and thus it has a good attribute as a spacecraft hardware. The proposed paper aims at decoding the convolution encoded system. There are various approaches to decode and broadly they are classified as

- Sequential decoding
- Maximum likelihood decoding based on Viterbi decoding

The development of an algorithm, implementation of the same and testing the same with simulated convolution encoded data stream is planned as part of this paper.

Keywords: Branch metrics, Convolution encoder, Forward error correction, Path metrics, Survivor paths, Trellis diagram, Viberti algorithm

I. INTRODUCTION

The probability of bit error P_e[1] in a digital transmission can be improved to attain acceptable reliability by increasing E_b/N_o. As E_b/N_o is directly proportional to $C/N_{0(}$(the ratio of the average carrier power to the noise power density) and the desired E_b/N_0 can be achieved by increasing the transmitted power and/or reducing the system noise temperature (to reduce N_0). Both these measures in a spacecraft are limited by cost, size, power and thermal constraints.. Channel coding is an acceptable alternative. In fact for a fixed E_b/N_0, coding is the method of lowering the BER. Channel Coding performs two functions, error detection and error correction. What is termed as forward error correction (FEC) allows errors to be corrected without the need for retransmission.In order to overcome this limitation convolutional coding and decoding method can be used. Convolutional coding is particularly suited to space and telemetry systems that require simple and small encoders and that obtain excellent coding gain by using sophisticated decoders at the ground station.
The functions of this software product are as follows:

- Configure the software based on the user requirement (like constraint length).
- Encodes data either represented in binary or hexadecimal format
- Decodes the data to recover the original data.

The system at present is configured for encoding a block of data obtained . It cannot be used to encode a continuous stream of data. In order to do it the system must be configured accordingly. It is assumed that the decoder is synchronized with the stream of incoming data .If the decoder doesnot know which of the n symbols in block initiates a branch of the trellis , then the data will be decoded with a very large number of errors .It also assumed that encoded data is not modified such that some part of it is missing. It contains all the software requirements to level of detail sufficient to enable designers to design a system to satisfy those requirements , and testers to test that the system satisfies those requirements.

II. DESIGN CONSIDERATIONS

- Design should allow for users to specify the two generator polynomials based on which the data is encoded.
- Should also for request constraint length
- Other options such as inversion of the output if required should be provided
- The data to be encoded in binary or hexadecimal format is generated
- The encoded data in binary format
- If any of the input like generator polynomial is missing error message should be generated.

2.1 System Design:

The system that implements forward error correction mainly consists of a channel encoder which adds redundant information to the stream of input symbols in a way that allows errors which are in the channel to be corrected .This redundancy is provided in a structured way to provide error control capability .Because of the redundancy introduced by the channel coder, there must be more symbols at the output of the coder than the input .The input to the channel coder is referred to as the message symbols .The input may also be referred to as information symbols .The stream that consists of the coded message and injected noise is input to the Viterbi decoder. Within the decoder, a Metric Update kernel is performed, which produces two streams – a state metric stream, which contains the accumulated state metrics for all delay states, and a transition stream, which contains the optimal path chosen for each delay state. These two streams are passed to a trace back stream, which traverses the state metric stream and employs the transition stream to find the optimal path through the Trellis.

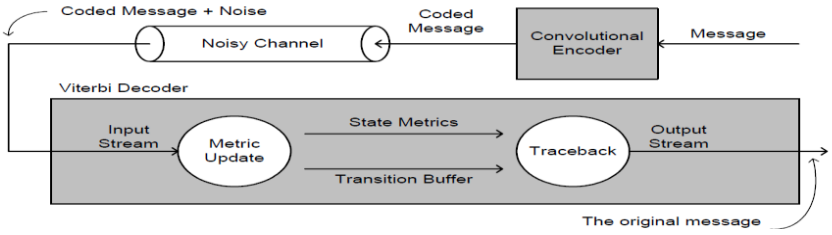


Fig 1: Simplified block of forward error correction system.

Like any error-correcting code, a convolutional code works by adding some structured redundant information to the user's data and then correcting errors using this information. A convolutional encoder is a linear system .A binary convolutional encoder can be represented as a shift register. The outputs of the encoder are modulo 2 sums of the values in the certain register's cells. The input to the encoder is either the un encoded sequence (for non-recursive codes) or the un encoded sequence added with the values of some register's cells (for recursive codes). Convolutional codes are frequently used to correct errors in noisy channels. They have rather good correcting capability and perform well even on very bad channels (with error probabilities of about 10^{-3}).

2.2 Convolutional Encoder

A binary convolutional encoder can be represented as a shift register. The encoder for a binary (2,1,4) code is shown in figure. Note that encoder consists of an m=3 stage shift register together with an n=3 modulo 2 adders and a multiplexer for serializing the encode outputs. The mod-2 adders can be implemented as EXCLUSIVE-OR gates. Since The mod-2 addition is a linear operation. The encoder[2] is a linear feed forward shift register. All the convolutional encoders can be implemented using a linear feed forward shift register of this type.

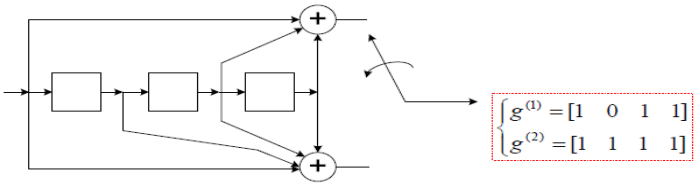


Fig 2: Convolution Encoder using shift registers

The information sequence $u=(u_0,u_1,u_2\ldots.)$ enters the encoder one bit at a time. Since the encoder is a linear system ,the two encoder output sequences $v^{(1)} =(v_0^{(1)} , v_1^{(1)} , v_2^{(1)} \ldots.)$ and $v^{(2)} =(v_0^{(2)} , v_1^{(2)} , v_2^{(2)} \ldots.)$ can be obtained as the convolution of the

input sequence with the two encoder "impulse" responses .The impulse responses are obtained by letting u=(1 0 0…) and observing output sequences. Since the encoder has m –time unit memory, the impulse responses can last at most m+1 time units, and are written $g^{(1)} =(g_0^{(1)}, g_1^{(1)} \ldots\ldots g_m^{(1)})$ and written $g^{(2)} =(g_0^{(2)}, g_1^{(2)} \ldots\ldots g_m^{(2)})$ .The impulse responses for the fig are

G(1)=1011 G(2)=1111

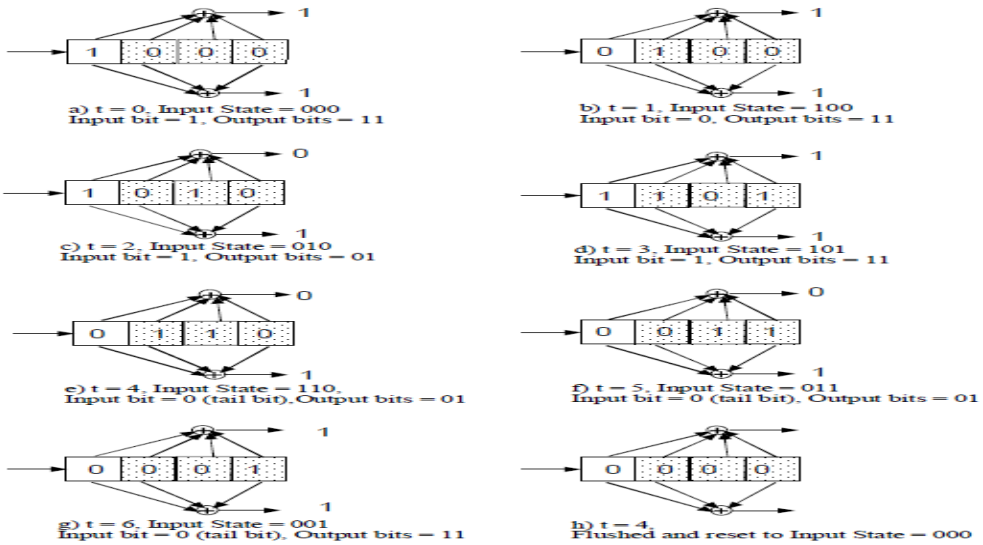
The impulse responses g(1) , g(2) are called generator sequences of the code. The encoding equations can be written as $v^{(1)} =u*g^{(1)}$ $v^{(2)} =u*g^{(2)}$ where * denotes discrete convolution and all operation are modulo-2.the convolution operation implies that for all l>=0,

$$v_l^{(j)} = \sum_{i=0}^{m} u_{l-i}\, g_i^j = u_l g_0^{(j)} + u_{l-1} g_1^{(j)} + \ldots\ldots\ldots\ldots\ldots + u_{l-m} g_m^{(j)} \quad \text{for } j=1,2..$$

After encoding,the two output sequences are multiplexed into a single sequence,called the code word,for transmission over channel

$V=(v_0^{(1)}\ v_0^{(2)}, v_1^{(1)}\ v_1^{(2)}, v_2^{(1)}\ v_2^{(2)}, \ldots)$

The encoding procedure is described here for sequence of 10 with the (2,1,4) code for the input sequence 1011 a) At time t = 0, we see that the initial state of the encoder is all zeros (the bits in the right most register positions). The input bit 1 causes two bits 11 to be output by a mod2 sum of all bits in the registers for the first bit and a mod2 sum of three bits for second output bit per the polynomial coefficients.b) At t = 1, the input bit 0 moves forward one register. The encoder is now in state 100. The output bits are now again 11 by the same procedure.c)At time t=2,the input bit 1 moves forward one register .the encoder is now in state 010.The output bits are produced similarly.At t=3,4 the same process continues the output produced is11 11 01 11.The entire process is summarized in the Fig 3 below.



A convolutional encoder[3] is often seen as a finite state machine. Each state corresponds to some value of the encoder's register. Given the input bit value, from a certain state the encoder can move to two other states. These state transitions constitute a diagram which is called a trellis diagram. A trellis diagram for the code on the Figure 2 is depicted on the Figure 3. A solid line corresponds to input 0, a dotted line – to input 1 (note that encoder states are designated in such a way that the rightmost bit is the newest one).

2.3 Trellis Diagram

Each path on the trellis diagram[4] corresponds to a valid sequence from the encoder's output. Conversely, any valid sequence from the encoder's output can be represented as a path on the trellis diagram. One of the possible paths is denoted as red (as an example). Each state transition on the diagram corresponds to a pair of output bits. There are only two allowed transitions for

every state, so there are two allowed pairs of output bits, and the two other pairs are forbidden. If an error occurs, it is very likely that the receiver will get a set of forbidden pairs, which don't constitute a path on the trellis diagram. So, the task of the decoder is to find a path on the trellis diagram which is the closest match to the received sequence.

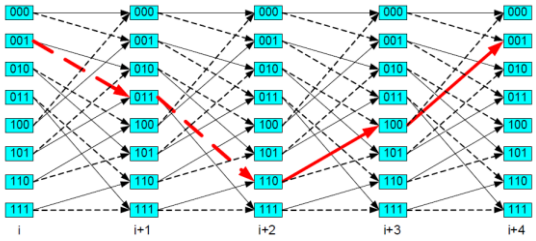


Fig 4: Trellis diagram

2.4 Viterbi algorithm

Viterbi algorithm[5] reconstructs the maximum-likelihood path given the input sequence. A soft decision decoder – a decoder receiving bits from the channel with some kind of reliability estimate. Three bits are usually sufficient for this task. Further increasing soft decision width will increase performance only slightly while considerably increasing computational difficulty. For example, if we use a 3-bit soft decision, then "000" is the strongest zero, "011" is a weakest zero, "100" is a weakest one and "111" is a strongest one.A hard decision decoder – a decoder which receives only bits from the channel (without any reliability estimate). A branch metric – a distance between the received pair of bits and one of the "ideal" pairs ("00","01", "10", "11").A path metric is a sum of metrics of all branches in the path.A meaning of distance in this context depends on the type of the decoder:

• for a hard decision decoder it is a Hamming distance, i.e. a number of differing bits;

• for a soft decision decoder it is an Euclidean distance.

In these terms, the maximum-likelihood path is a path with the minimal path metric. Thus the problem of decoding is equivalent to the problem of finding such a path.Let's suppose that for every possible encoder state we know a path with minimum metric[6] ending in this state. For any given encoder state there is two (and only two) states from which the encoder can move to that state, and for both of these transitions we know branch metrics. So, there are only two paths ending in any given state on the next step. One of them has higher metric, it is a survivor path.A Viterbi algorithm[7] consists of the following three major parts:

1. Branch metric calculation – calculation of a distance between the input pair of bits and the four possible "ideal" pairs ("00", "01", "10", "11").
2. Path metric calculation – for every encoder state, calculate a metric for the survivor path ending in this state (a survivor path is a path with the minimum metric).
3. Traceback[8] – this step is necessary for hardware implementations that don't store full information about the survivor paths, but store only one bit decision every time when one survivor path is selected from the two.

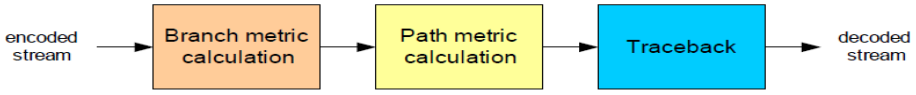


Fig. 5: Viterbi decoder data flow

Branch Metric Calculation:

Methods of branch metric calculation are different for hard decision and soft decision decoders.

For a *hard decision* decoder, a branch metric is a Hamming distance between the received pair of bits and the "ideal" pair. Therefore, a branch metric can take values of 0, 1 and 2. Thus for every input pair we have 4 branch metrics (one for each pair of "ideal" values).

For a *soft decision* decoder, a branch metric is measured using the Euclidean distance. Let x be the first received bit in the pair, y – the second, $x0$ and $y0$ – the "ideal" values. Then branch metric is $Mb=(x-x_0^2)+(y-y_0^2)$.

Furthermore, when we calculate 4 branch metric for a soft decision decoder, we don't actually need to know absolute metric values – only the difference between them makes sense. So, nothing will change if we subtract one value from the all four branch metrics:

$Mb=(x2-2\,x\,x_0+x_0^{\,2})+(y^2-2\,y\,y_0+y_0^{\,2})$;

$Mb^* = Mb - x^2 - y^2 = (x_0^2 - 2\,x\,x_0) + (y_0^2 - 2\,y\,y_0)$.

Path Metric Calculation:

Path metrics are calculated using a procedure called *ACS* (Add-Compare-Select). This procedure is repeated for every encoder state.

1. Add – for a given state, we know two states on the previous step which can move to this State , and the output bit pairs tha t correspond to these transitions. To calculate new path Metrics , we add the previous path metrics with the corresponding branch metrics.
2. Compare, select – we now have two paths, ending in a given state. One of them (with greater metric) is dropped.

As there are 2^{K-1} encoder states, we have 2^{K-1} survivor paths at any given time.

It is important that the difference between two survivor path metrics cannot exceed $\delta \log(K-1)$, where δ is a difference between maximum and minimum possible branch metrics.

The problem with path metrics is that they tend to grow constantly and will eventually overflow.

But, since the absolute values of path metric don't actually matter, and the difference between them is limited, a data type with a certain number of bits will be sufficient.

There are two ways of dealing with this problem:

1. Since the absolute values of path metric don't actually matter, we can at any time subtract an identical value from the metric of every path. It is usually done when *all* path metrics exceed a chosen threshold (in this case the threshold value is subtracted from every path metric). This method is simple, but not very efficient when implemented in hardware.
2. The second approach allows overflow, but uses a sufficient number of bits to be able to detect whether the overflow took place or not. The *compare* procedure must be modified in this case. The whole range of the data type's capacity is divided into 4 equal parts. If one path metric is in the 3-rd quarter, and the other – in the 0-th, then the overflow took place and the path inthe 3-rd quarter should be selected. In other cases an ordinary compare procedure is applied. This works, because a difference between path metrics can't exceed a threshold value, and the range of path variable is selected such that it is at least two times greater than the threshold.

Traceback:

It has been proven that all survivor paths merge after decoding a sufficiently large block of data (D in fig) i.e. they differ only in their endings and have the common beginning.

If we decode a continuous stream of data, we want our decoder to have finite latency. It is obvious that when some part of path at the beginning of the graph belongs to every survivor path, the decoded bits corresponding to this part can be sent to the output. Given the above statement, we can perform the decoding as follows:

1. Find the survivor paths for $N+D$ input pairs of bits.
2. *Trace back* from the end of any survivor paths to the beginning.
3. Send N bits to the output.
4. Find the survivor paths for another N pairs of input bits.
5. Go to step 2.

In these procedure[9] D is an important parameter called decoding depth. A decoding depth should be considerably large for quality decoding, no less then $5K$. Increasing D decreases the probability of a decoding error, but also increases latency.

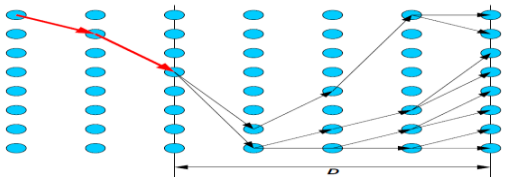


Fig. 6: Survivor paths graph example.

Blue circles denote encoder states. It can be seen that all survivor paths have a common beginning (red) and differ only in their endings.

Iii. Testing

A few of the sample test cases are as follows:

Serial No. of test case:	UTC-1
Module under test:	ConfigureCODEC
Description:	Configures the user input
Sample input:	User input
Expected Output:	it should configure the user input
Actual Output:	It is configuring the user input
Remarks:	Test is successful

Serial No. of test case:	UTC-2
Module under test:	CCSDS_Encoder
Description:	Encodes the input data
Sample input:	Data given by the user
Expected Output:	It should encode the data
Actual Output:	It is encoding the data
Remarks:	Test is successful

Serial No. of test case:	UTC-3
Module under test:	CCSDS_Decoder
Description:	Decodes the received data
Sample input:	Input from the encoder
Expected Output:	Decode the received data
Actual Output:	Decoding the received data
Remarks:	Test is successful

IV. Conclusion And Future Work

This paper attempted implementation of encoding and decoding of telemetry data. The major portion of the project is concentrated on C implementation on Windows platform which has been done successfully, experimentation with different input. In this application, the data is not a continuous stream and the noise is not considered . Further work on how to makethedata a continuous stream can be done. Noise should be simulated. Graphical representation of output from encoder and decoder.

REFERENCES

[1] http://www.radio-electronics.com/info/rf-technology-design/ber/bit-error-rate-tutorial-definition.php
[2] http://apogeelabs.com/pdffiles/convolutional_encoding.pdf
[3] http://www.comlab.hut.fi/opetus/333/2004_2005_slides/Convolutional_Coding_Viterbi_Algorithm.pdf
[4]http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=650373
[5] http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=5443417&contentType=Conference+Publications
[6] http://www.1-core.com/library/comm/viterbi/
[7] http://www.ece.umd.edu/~tretter/enee722/ungerboeck_paper.pdf
[8]http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1450960&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D1450960
[9]http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=410439&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D410439

Design of Classifier for Detecting Image Tampering Using Gradient Based Image Reconstruction Technique

Sonal Sharma, Preeti Tuli

Abstract
Image tampering detection is a significant multimedia forensics topic which involves, assessing the authenticity or not of a digital image. Information integrity is fundamental in a trial but it is clear that the advent of digital pictures and relative ease of digital image processing makes today this authenticity uncertain. In this paper this issue is investigated and a framework for digital image forensics is presented, individuating if the tampering has taken place. Based on the assumptions that some processing must be done on the image before it is tampered, and an expected distortion after processing an image, we design a classifier that discriminates between original and tampered images. We propose a novel methodology based on gradient based image reconstruction to classify images as original or tampered. This methodology has its application in a context where the source image is available (e.g. the forensic analyst has to check a suspect dataset which contains both the source and the destination image).

Index Terms — Gradient, Poisson equation, Region of interest (ROI), Digital image forensics, Authenticity verification, Image reconstruction from gradients.

1 INTRODUCTION

IN today's digital age, the creation and manipulation of digital images is made simple by digital processing tools that are easily and widely available. As a consequence, we can no longer take the authenticity of images for granted especially when it comes to legal photographic evidence. *Image forensics*, in this context, is concerned with determining the source and potential authenticity of an image. Although digital watermarks have been proposed as a tool to provide authenticity to images, it is a fact that the overwhelming majority of images that are captured today do not contain a digital watermark. And this situation is likely to continue for the foreseeable future. Hence in the absence of widespread adoption of digital watermarks, there is a strong need for developing techniques that can help us make statements about the origin, veracity and authenticity of digital images.

In this paper we focus on the problem of reliably discriminating between "tampered" images (images which are altered in order to deceive people) from untampered original ones. The basic idea behind our approach is that a tammpered image (or the least parts of it) would have undergone some image processing. Hence, we design a classifier that can distinguish between images that have and have not been processed. We apply it to a suspicious image of a target image and classify the suspicious image as tampered or untampered. The rest of this paper is organized as follows: In Section 2we present a method to verify the authenticity of images that is used in the classifier we design for image forensics, i.e. we formulate the problem and present solution methodology. Statistical performance results are given in Section 3, with conclusions drawn in section 4.

2 PROBLEM FORMULATION AND SOLUTION METHODOLOGY

The problem of fraud detection has been faced by proposing different approaches each of these based on the same concept: a forgery introduces a correlation between the original image and the tampered one. Several methods search for this dependence by analyzing the image and then applying a feature extraction process. In [1] the direct approach proposed by Fridrich et al. comprises of performing an exhaustive search by comparing the image to every cyclic – shifted versions of it, which requires (MN^2) steps for an image sized M by N. This computationally expensive search does not work where the copied region has undergone modifications. In [2] A.N. Myma et al. presented an approach of first applying wavelet transform to the input image to yield a reduced dimension representation, then exhaustive search is performed to identify similar blocks in the image by mapping them to log polar co-ordinates and using phase correlation as the similarity criterion. But the performance relies on the location of copy- move regions. In [3] Weihai Li et al. utilized the mismatch of information of block artifact grid as clue of copy paste forgery. A DCT grid is the horizontal lines and the vertical lines that partition an image into blocks and a block artifact grid (BAG) is the grid embedded in an image where block artifact appears. The DCT grid and BAG match together in untampered images. But if the copied area is from the other different image it cannot be detected by the method, also the complexity of algorithm is high. In [4] Bayram et al. proposed Fourier – Mellin transform (FMT). But the algorithm works for the case of only slight rotation. In [5] Xu Bo et al. proposed a method in which Speed up Robust features (SURF) key points are extracted and their descriptors are matched within each other with a threshold value. This method fails to automatically locate the tampered region and its boundary.

None of these approaches [1, 2, 3, 4, and 5] conducts authentication verification using gradient maps in the image reconstruction.

The approach presented in this paper verifies the authentication in two phases *modeling phase* and *simulation phase.* In *modeling phase* the image is reconstructed from the image gradients by solving a poisson equation and in the *simulation phase* absolute difference method and histogram matching criterion between the original and test image is used. The solution methodology is discussed in the subsequent paragraphs.

2.1 Image Reconstruction

In the year 1993 Luc Vincent [6] carried out the work in morphological grayscale reconstruction. In 2004 Di Zang and G. Sommer [7] carried out phase based image reconstruction in the monogenic scale space. In 2005 S. Leng et al. [8] presented fan-beam image reconstruction algorithm to reconstruct an image via filtering a back projection image of differentiated projection data. In 2008 A. L. Kesidis and N. Papamarkos [9] presented a new method for the exact image reconstruction from projections. The original image is projected into several view angles and the projection samples are stored in an accumulator array. In 2011 P. Weinzaepfel et al [10] proposed another novel approach which consists using an off-the-shelf image database to find patches visually similar to each region of interest of the unknown input image.

The approach presented in this paper is gradient based image reconstruction by solving poisson equation. The details of the method are described as under:

2.1.1 Gradient Based Image Reconstruction:

As already stated in our previous work [11] image reconstruction from gradient fields is a very active research area. The gradient-based image processing techniques and the poisson equation solving techniques have been addressed in several related areas such as high dynamic range compression [12], Poisson image editing [13], image fusion for context enhancement [14], interactive photomontage [15], Poisson image matting [16] and photography artifacts removal [17]. A new criterion is developed, where the image is reconstructed from its gradients by solving a poisson equation and hence used for authenticity verification [11].

In 2D, a modified gradient vector field,

$$G' = [G'_x, G'_y] \quad (1)$$

may not be integrable.

Let I' denote the image reconstructed from G', we employ one of the direct methods recently proposed in [12] to minimize,

$$\| \nabla I' - G \| \quad (2)$$

so that,

$$G \approx \nabla I' \quad (3)$$

By introducing a Laplacian and a divergence operator, I' can be obtained by solving the Poisson differential equation [18, 19]

$$\nabla^2 I' = div([G'_x, G'_y]) \quad (4)$$

Since both the Laplacian and *div* are linear operators, approximating those using standard finite differences yields a large system of linear equations. The full multigrid method [20] is used to solve the Laplacian equation with Gaussian-Seidel smoothing iterations. For solving the poisson equation more efficiently, an alternative is to use a rapid poisson solver, which uses a sine transform based on the method [18] to invert the laplacian operator. Therefore, the rapid poisson solver is employed in our implementation. The image is zero-padded on all sides to reconstruct the image.

2.1.2 Poisson Solvers:

A Poisson solver produces the image whose gradients are closest to the input manipulated gradient domain image in a least squares sense, thereby doing a kind of inverse gradient transform. Note that if the input were a gradient domain image whose gradients had not been manipulated, the inverse gradient transformation would have an exact solution, and the poisson equation would give a perfect reconstruction of the image. Both the FFT-based solver and the poisson solver using zero Dirichlet boundary condition work successfully in obtaining an inverse gradient transformation in the sense that they give a perfect reconstruction of the image when the input gradient domain image is not manipulated. This section details the poisson solver which has been used in the present research work.

In this section, we describe the standard gradient integration problem and its poisson solution and then expand this result to include a data function term. The problem of computing a function f (x,y) whose gradient ∇f (x,y) is as close as possible to a given gradient field g (x,y) is commonly solved by minimizing the following objective:

$$\int\int \| \nabla f - g \|^2 \, dx \, dy. \quad (5)$$

Note that g is a vector-valued function that is generally not a gradient derived from another function. (If g were derived from another function, then the optimal f would be that other function, up to an unknown constant offset.)

It is well-known that, by applying the Euler-Lagrange equation, the optimal f satisfies the following Poisson equation:

$$\nabla^2 f = \nabla \cdot g, \quad (6)$$

which can be expanded as $f_{xx} + f_{yy} = g_x^x + g_y^y$, where $g = (g^x, g^y)$. Subscripts in x and y correspond to partial derivatives with

respect to those variables. We have superscripted g^x and g^y to denote the elements of g rather than subscript them, which would incorrectly suggest they are partial derivatives of the same function.We now expand the objective beyond the standard formulation. In particular, we additionally require f(x, y) to be as close as possible to some data function u(x, y). The objective function to minimize now becomes:

$$\int\int \lambda_d (f-u)^2 + \|\nabla f - g\|^2 \, dx \, dy, \quad (7)$$

Where, λ_d is a constant that controls the trade-off between the fidelity of f to the data function versus the input gradient field.
To solve for the function f that minimizes this integral, we first isolate the integrand:

$$L = \lambda_d (f-u)^2 + \|\nabla f - g\|^2 = \lambda_d (f-u)^2 + (f_x - g^x)^2 + (f_y - g^y)^2 \quad (8)$$

The function f that minimizes this integral satisfies the Euler-Lagrange equation:

$$(dL / df) - (d/dx) \, . \, (dL/dfx) - (d/dy).(dL/dfy) = 0 \quad (9)$$

Substituting and differentiating, we then have:

$$2 \lambda_d (f-u) - 2(f_{xx} - g^x) - 2(f_{yy} - g^y) = 0 \quad (10)$$

Rearranging gives us:

$$\lambda_d f - (f_{xx} + f_{yy}) = \lambda_d u - (g_{xx} + g_{yy}) \quad (11)$$

or equivalently:

$$\lambda_d f - \lambda^2 f = \lambda_d u - \lambda \cdot g \quad (12)$$

The left-hand side of this equation is poisson equation, typically studied in three dimensions in physics. Our analysis will be in 2D. As expected, setting $\lambda_d = 0$ nullifies the data term and gives us the Poisson equation.

2.1.3 Discrete Sine Transform

In this section we analyze the 2D Poisson equation in the sine domain. As with fast Poisson solvers, we can solve the screened poisson equation (Equation 8) by taking its sine transform.
The **discrete sine transform** (DST) is a Fourier-related transform similar to the discrete Fourier transform (DFT), but using a purely real matrix. It is equivalent to the imaginary parts of a DFT of roughly twice the length, operating on real data with odd symmetry (since the Fourier transform of a real and odd function is imaginary and odd), where in some variants the input and/or output data are shifted by half a sample.
Formally, the discrete sine transform is a linear, invertible function F : RN → RN (where R denotes the set of real numbers), or equivalently an N × N square matrix. There are several variants of the DST with slightly modified definitions. The N real numbersx0,...., x_{N-1} are transformed into the N real numbers X_0, ..., X_{N-1} according to the formula:

$$X_k = \sum_{n=0}^{N-1} x_n \sin\left[\frac{\pi}{N+1}(n+1)(k+1)\right] \qquad k = 0, \ldots, N-1 \quad (13)$$

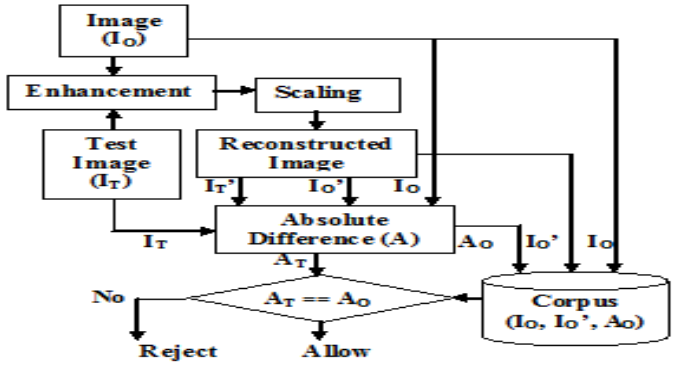


Figure 1: Schematic diagram for Modeling and Simulation Phase

2.1.4 Inverse Discrete Sine Transform

The inverse of DST is DST multiplied by 2/ (N+1). Like for the DFT, the normalization factor in front of these transform definitions is merely a convention and differs between treatments.

2.2 Absolute Difference

In the present work our approach is to find the absolute difference between the original and the reconstructed image. Subtraction gives the difference between the two images, but the result may have a negative sign and can be lost. The function that finds how different the two images are- regardless of the arithmetic sign- is the absolute difference:

$$N(x, y) = |O_1(x, y) - O_2(x, y)|, \quad (14)$$

where, $O_1(x, y)$ and $O_2(x, y)$ are pixels in the original images, |x| is the absolute difference operator, and N(x, y) is the resultant new pixel. The absolute difference operator returns +x whether the argument is –x or +x.

2.3 Histogram Normalization

Histogram is a graphical representation of the intensity distribution of an image. It quantifies the number of pixels for each intensity value considered.Histogram normalization is a method that improves the contrast in an image, in order to stretch out the intensity range. Equalization implies mapping one distribution (the given histogram) to another distribution (a wider and more uniform distribution of intensity values) so that the intensity values are spread over the whole range.

To accomplish the equalization effect, the remapping should be the cumulative distribution function (CDF)

For the histogram H(i), its cumulative distribution H'(i) is:

$$H'(i) = \Sigma H(j), \text{ where } 0 \leq j < i \quad (15)$$

To use this as a remapping function, we have to normalize H'(i) such that the maximum value is 255 (or the maximum value for the intensity of the image). Finally, we use a simple remapping procedure to obtain the intensity values of the equalized image:

$$equalized(x, y) = H'(src(x,y)) \quad (16)$$

In our work, first we perform the histogram normalization and then the histogram equalization criteria is used where the normalized histogram values of the original and test image are utilized for matching the two images.

2.4 Algorithm Used:

Algorithm 1: Modeling and Simulation of original and reconstructed image

Modeling phase

Step 1: Read an image (I_O).

Step 2: Convert into grayscale image, say R.

(Enhancement stage)

Step 3: Perform Scaling on the image.

Step 4: Enhance the image using median filtering and convolution theorem (I_O).

Step 5: Reconstruct the image using proposed methodology (I_O').

Step 6: Find the absolute difference between original and reconstructed image (A_O).

Step 7: Store the original image, reconstructed image and absolute difference (I_O, I_O', A_O)

Simulation phase

Step 8: Input a test image (I_T)

Step 9: Reconstruct I_T to obtain I_T' and find the absolute difference (A_T) between I_T and I_T'

Step 10: Compare A_T and A_O to find a match and hence allow or reject the subject accordingly.

2.5 Modeling and Simulating

As shown in Fig. 1, in the modeling phase, let I_O be the original image of a subject which has to be modeled for the formation of knowledge based corpus. After enhancing and proper scaling of the original image I_O, the image is poisson reconstructed from its gradients as:

$$I_O' = \text{Poisson\_reconstruction } (I_O) \quad (17)$$

Now the absolute difference between the original and reconstructed image is calculated as:

$$A_O = \text{Absolute\_difference } (I_O, I_O') \quad (18)$$

Now store the triplet (I_O, I_O', A_O) in the corpus so as to form the knowledge based model (corpus). The equations (17) and (18) can be repeatedly used to register n number of subjects, and store their details for authentication verification. In the simulation phase, when the tampered or forged image will be presented to the security system for authentication, the system will reconstruct the test image (I_T) as:

$$I_T' = \text{Poisson\_reconstruction } (I_T) \quad (19)$$

And, then the absolute difference between the original test image (I_T) and reconstructed test image (I_T') is calculated as:

$$A_T = \text{Absolute\_difference } (I_T, I_T') \quad (20)$$

Now, the resultant A_T is compared with A_O (the absolute difference stored in corpus of the original and reconstructed original image in modeling phase)

If (A_T == A_O)

"Authenticity Verified as TRUE!"

Else

"Authenticity Verified as FALSE!"

Hence, the result will reject the subject due to a mismatch and the images obtained by forgery or tampering for authenticity verification will be classifiedd as fake or invalid and any hidden data (for destroying the security system or secret communication) will be clearly identified.

3 RESULTS AND DISCUSSION

The solution methodology for the above stated problem is implemented using proposed algorithm and the experimental outcomes are shown below:

3.1 Results for modeling phase (Original Image).

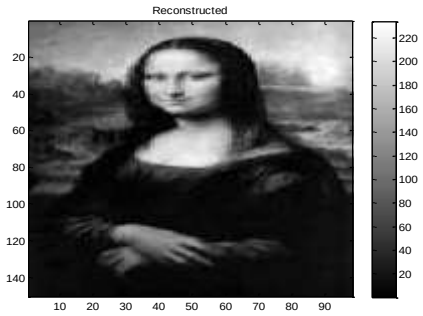


Figure 2.1: Original Image (I_O) **Figure 2.2: Reconstructed Image (I_O')**

Solving Poisson Equation Using DST
Time for Poisson Reconstruction = 0.703055 secs
(Image Size: 98×150)

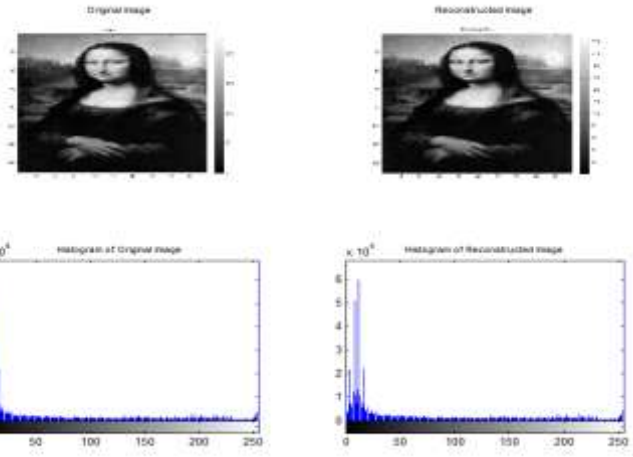


Figure 2.3: Histogram of original image (I_O) and Reconstructed Image (I_O')

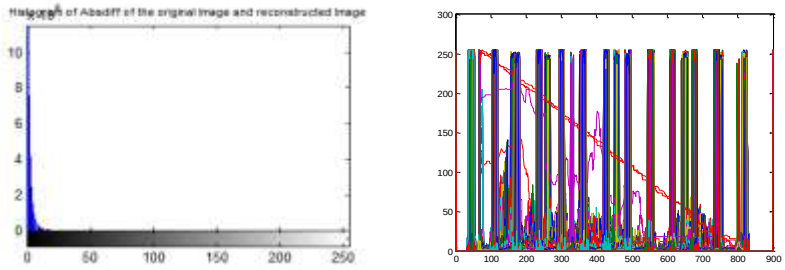


Figure 2.4: Histogram of absolute difference of original image (I_O) and reconstructed original image (I_O') (left), Plot of absolute difference of original image (I_O) and reconstructed original image (I_O') (right)

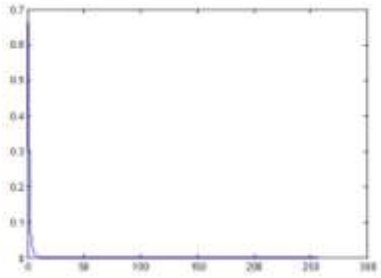


Figure 2.5: Plot of the Normalized Histogram of absolute difference of original image (I_O) and reconstructed original image (I_O')

As shown in Fig. 2.1 to Fig. 2.5, the registration of the images for authenticity verification has been done by using the steps of the modeling phase mentioned in Algorithm 1. The original image I_O (Fig. 2.1) is converted to grayscale image, then the image is scaled and enhanced and then the scaled image is poisson reconstructed and the resultant image (I_O') is shown in Fig. 2.2. The histogram of the original and reconstructed image is shown in subplots in Fig. 2.3. For authenticity verification the absolute difference (A_O) of the original image (I_O) and reconstructed image (I_O') is calculated. Now the histogram of A_O is obtained and the result is shown in Fig. 2.4 (left). The plot of the histogram so obtained is shown in Fig. 2.4 (right), and the plot of normalized histogram is shown in the Fig. 2.5. This plot of the normalized histogram will be compared with that of the test image during simulation. Now the corpus contains the triplet (I_O, I_O', A_O) for the registered subject's original image.

The above can be repeated for the registration of n number of subject's.

3.2 Results for simulation phase [Test (tampered) Image]

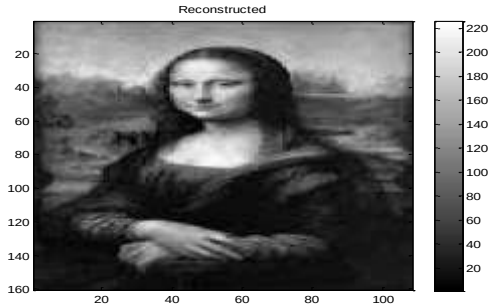


Figure 3.1: Test Image (I_T) **Figure 3.2: Reconstructed Test Image (I_T')**

Solving Poisson Equation Using DST
Time for Poisson Reconstruction = 0.060161 secs
Image Size: (184×274)

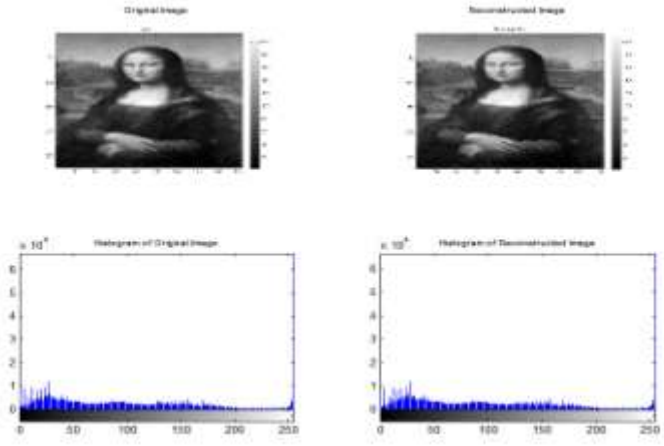


Figure 3.3: Histogram of original Test image (I_T) and Reconstructed Test Image (I_T')

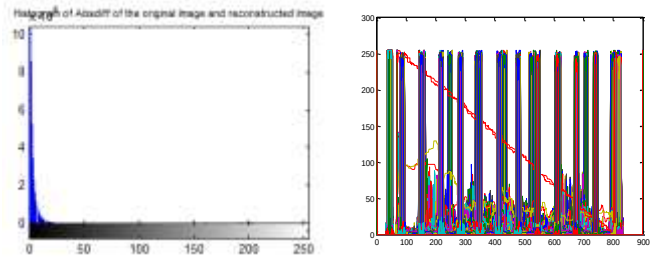


Figure 3.4 : Histogram of absolute difference of test image (I_T) and reconstructed test image (I_T') (left), Plot of absolute difference of test image (I_T) and reconstructed test image (I_T') (right)

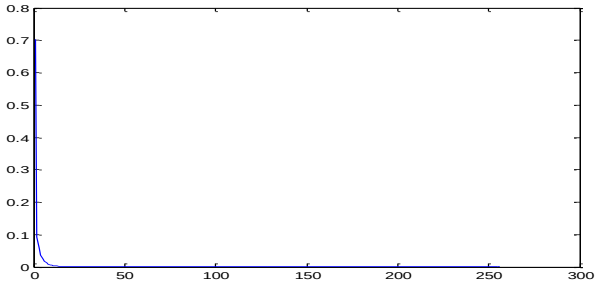


Figure 3.5: Plot of the Normalized Histogram of the absolute difference of test image (I_T) and reconstructed test image (I_T')

The test image (tampered) is passed through the steps of simulation phase mentioned in Algorithm 1 and the results are shown in Fig. 3.1 to Fig. 3.5. The authenticity verification of the test image (I_T) has been done by using the steps of the simulation phase mentioned in Algorithm 1. The test image I_T (Fig. 3.1), is converted to grayscale image, then the image is scaled and enhanced. Now the scaled image is poisson reconstructed and the resultant image (I_T') is shown in Fig. 3.2. The histogram of the test and reconstructed test image is shown in subplots in Fig. 3.3. For authenticity verification the absolute difference (A_T) of the test image (I_T) and reconstructed test image (I_T') is calculated and then the histogram of A_T is obtained and the result is shown in Fig. 3.4 (left) and the plot of the histogram so obtained is shown in Fig. 3.4 (right). Finally, the plot of normalized histogram is shown in the Fig. 3.5 which will be compared with that stored in corpus.

4 CONCLUSION

The normalized histogram of absolute difference of the test image and reconstructed test image (Fig. 3.5) is compared with the normalized histogram of absolute difference of the original image and reconstructed original image (Fig. 2.5), and the so obtained result is inequality, since, the value of the difference is not zero and comes to be-
0.0049units.
And, hence the image is classified as tampered and finally rejected. If the image was not tampered then the so obtained difference (between the the normalized histogram of absolute difference of the test image and reconstructed test image (Fig. 3.5) and the normalized histogram of absolute difference of the original image and reconstructed original image (Fig. 2.5) would be-
0.0000units.
In this manner the authenticity of the individual's can be verified and the test images can be classified as tampered (or forged) or original, and hence the tampering can be detected. We also observed that the time required to reconstruct the original image is-
0.703055 secs,
But, the time required to reconstruct the tampered image is-
0.060161 secs.
Hence, we also conclude that the tampering can be detected by the time our poisson solver takes to reconstruct the original image and test image.

REFERENCES

[1] J. Fridrich, D. Soukal, and J. Lukas, "Detection of copy-move forgery in digital images," Proc. Digital Forensic Research Workshop, Cleveland, OH, August 2003.

[2] A.N.Myrna, M.G.Venkateshmurthy, "Detection of Region Duplication Forgery in Digital Images Using Wavelets and Log-Polar Mapping," Conference on Computational Intelligence and Multimedia Applications, Dec. 2007, Vol.3, pp. 371-377.

[3] Weihai Li, Nenghai Yu and Yuan Yuan, "Doctored jpeg image detection," IEEE ICMS 2008, pp 253-256.

[4] S. Bayram, H. T. Sencar, and N. Memon, "An Efficient and Robust Method for Detecting Copy-move Forgery," Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, pp. 1053-1056, 2009.

[5] Xu Bo, Wang Junwen, Liu Guangjie and Dai Yuewei, "Image Copy-move Forgery Detection Based on SURF," International Conference on Multimedia Information Networking and Security, IEEE pp 889-892.

[6] Luc Vincent, "Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms," IEEE Transactions on Image Processing, vol. 2, no. 2, 1993.

[7] Di Zang and G. Sommer, "Phase Based Image Reconstruction in the Monogenic Scale Space," DAGM-Symposium, 2004.

[8] S. Leng, T. Zhuang, B. Nett and Guang-Hong Chen, "Exact fan-beam image reconstruction algorithm for truncated projection data acquired from an asymmetric half-size detector," Phys. Med. Biol. 50 (2005) 1805–1820.

[9] A. L. Kesidis, N. Papamarkos, "Exact image reconstruction from a limited number of projections," J. Vis. Commun. Image R. 19 (2008) 285–298.

[10] P. Weinzaepfel, H. Jegou, P. Perez, "Reconstructing an image from its local descriptors," Computer Vision and Pattern Recognition (2011).

[11] S. Sharma, P. Tuli, " Fraud and Tamper Detection in Authenticity Verification through Gradient Based Image Reconstruction Technique for Security Systems," IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-0661 Volume 3, Issue 4 (July-Aug. 2012), PP 01-06

[12] R. Fatta, D. Lischinski, M. Werman, "Gradient domain high dynamic range compression" ACM Transactions on Graphics 2002;21(3):249-256.

[13] P. P´erez ,M. Gangnet , A. Blake, " Poisson image editing" ACM Transactions on Graphics 2003;22(3):313-318.

[14] R. Raskar, A. Ilie , J.Yu, " Image fusion for context enhancement and video surrealism", In: Proceedings of Non-Photorealistic Animation and Rendering '04, France, 2004. p. 85-95.

[15] A. Agarwala , M. Dontcheva, M. Agrawala , S. Drucker, A.Colburn, B. Curless, D Salesin , M. Cohen M, " Interactive digital photomontage. ACM Transactions on Graphics" 2004;23(3):294-302.

[16] J. Sun, J. Jia, CK. Tang , HY Shum , "Poisson matting. ACM Transactions on Graphics" 2004;23(3):315-321.

[17] A. Agrawal , R. Raskar, SK. Nayar , Y. Li, "Removing flash artifacts using gradient analysis" ACM Transactions on Graphics 2005;24(3):828-835.

[18] R. Raskar, K. Tan, R. Feris , J. Yu, M. Turk "Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging" ACM Transactions on Graphics 2004;23(3):679-688.

[19] J. Shen, X. Jin, C. Zhou, Charlie C. L. Wang, "Gradient based image completion by solving the Poisson equation," PCM'05 Proceedings of the 6th Pacific-Rim conference on Advances in Multimedia Information Processing – Volume Part I 257-268

[20] W. Press, S. Teukolsky, W. Vetterling, B. Flannery "Numerical Recipes in C: The Art of Scientific Computing" Cambridge University Press; 1992.

Control Parameters Optimization of Laser Beam Machining Using Genetic Algorithm

Ruben Phipon[1], B.B.Pradhan[2]
[1] Mechanical Engineering Department, SMIT, Sikkim, India.
[2] Mechanical Engineering Departments, SMIT, Sikkim, India.

Abstract
To improve and optimize the responses of a Laser Beam machining process, the various input machining control parameters are to be set at an optimal value. As such one has to adopt experimental methods, which are cumbersome, time consuming, costly and at times not feasible. During such situations, optimization techniques like Genetic Algorithm (GA) can be used as it provides a cost effective method for solution of such complex problems. Unlike traditional optimization techniques, GA is a robust and performs well in multimodal optimization problems. Considering these advantages of GA, optimization of Nd:Yag Laser beam machining process is done using this technique. In this research work, the desired responses are minimum kerf taper and surface roughness. The process control parameters considered are Oxygen pressure, pulse width, pulse frequency and cutting speed. Experiments are designed and the mathematical models correlating the desired responses and the control parameters are established using Response Surface Methodology (RSM). Finally, GA is applied to search the optimal parametric values for the optimal responses. Using Genetic Algorithm, minimum Kerf taper obtained is 0.14695° which is 0.313° less in magnitude than experimentally measured value. Also, minimum surface roughness predicted using GA is 1.2625µm which is 0.3375µm better in value compared to the experimental measured value. The average percentage prediction error of GA is found to be 3.35% for kerf taper and 4.02% for surface roughness. Thus, the results prove GA to be a novel optimization technique which can be used to optimize Laser beam machining processes.

Keywords: Laser Beam Machining (LBM), Response surface methodology (RSM), Genetic Algorithm (GA), Optimization.

1. Introduction

Laser beam machining (LBM) is a novel thermal energy based advanced machining process which can be used for machining a wide range of materials. In this process, a laser beam is focused for melting and vaporizing to remove material from the workpiece as per the desired shape [1]. Hence, the characteristic of non-contact between the tool and the workpiece makes this machining process desirable as it removes chances of workpiece deterioration due to cutting tool force. It is suitable for cutting complex geometric profiles, for drilling miniature holes in sheet metal and precision machining of micro-parts. However, improvement in LBM process performance can be made by studying the different factors that affect the quality characteristics. Thus, process performance can be improved by proper selection of process control parameters.
Kuar et al.[2] performed experiments to investigate into CNC pulsed Nd:YAG laser micro-drilling of zirconium oxide (ZrO2). The optimal setting of process parameters such as pulse frequency and pulse width, lamp current, assist air pressure for achieving minimum HAZ thickness and taper of the micro-hole was determined. Dubey and Yadav [3] while cutting thin sheet of aluminium alloy using pulsed laser performed multi-objective optimization of kerf quality such as kerf deviation and kerf width. They observed assist gas pressure and pulse frequency make significant affect on the kerf quality in the operating range of process parameters. Sharma et al. [4] conducted experiments based on the Taguchi quality design concept for parameter optimization of the kerf quality characteristics during pulsed Nd:YAG laser cutting of nickel based superalloy thin sheet. The results indicate that the optimum input parameter levels suggested for curved cut profiles are entirely different from straight cut profiles.

Optimization of the machining process first requires a mathematical model to be established to correlate the desired response and the process control parameters. Thereafter an optimization technique is applied to find optimal setting of the control parameters to derive the desired responses. Mukherjee and Ray [5] presented a generic framework for parameter optimization in metal cutting processes for selection of an appropriate approach. Response Surface Methodology (RSM) is generally employed to design experiments with a reduced number of experimental runs to achieve optimum responses. Lalwani et al. [6] applied RSM to investigate the effect of cutting parameters on surface roughness in finish hard turning of MDN250 steel using coated ceramic tool. Soveja et al. [7] studied the influence of the operating factors on the laser texturing process using two experimental approaches: Taguchi methodology and RSM. Dubey and Yadav [8] present a hybrid Taguchi method and response surface method (TMRSM) for the multi-response optimization of a laser beam cutting process.

Yildiz [9] demonstrated the superiority of the hybrid optimization approach over the other techniques in terms of convergence speed and efficiency. Yusup et al. [10] discussed evolutionary techniques in optimizing machining process parameters for both traditional and modern machining. They observed evolutionary techniques while optimizing machining process parameters positively gives good results. Samanta and Chakraborty [11] proved the applicability and suitability of evolutionary algorithm in enhancing the performance measures of non traditional machining processes. Jain et al. [12] used GA for optimization of process parameters of mechanical type advanced machining processes Traditional optimization methods are not suitable to solve problems where the formulated objective functions and constraints are very complicated and implicit functions of the decision variables. Unlike conventional optimization techniques, GA is a robust and can be effectively applied for multi modal problems. Hence, considering these advantages of GA, an attempt has been made to optimize the LBM process in this research paper using this technique.

2. Mathematical Modeling Using Rsm

RSM is a statistical technique employed to design experiments with a reduced number of experimental runs to achieve optimum responses. It is used to establish mathematical models which correlate the responses and the independent control parameters. Sharma and Yadav [13] performed experiments in a 200W pulsed Nd: YAG laser beam machine system with CNC work table. Surface quality i.e. kerf taper and surface roughness are the required measures of response. Process parameters considered that affect these responses are assist gas pressure (x_1), pulse width (x_2), pulse frequency (x_3), and cutting speed (x_4).The relationship of process parameters and output responses is represented mathematically in the form of a polynomial. The 1st order polynomial does not provide higher order interaction and the 3rd order are not desirable as they are difficult to solve. As such 2nd order polynomial is suitable for higher order interaction and gives better result. CCRD technique is applied to provide good predictions throughout the permissible region of interest. CCRD requires a minimum of five levels of all control parameters for the calculation of regression coefficients. The process control parameters and their values at different levels are shown in Table 1. A total of 31 experimental runs are designed which consist of 2p factorial runs, 2p axial runs and 7 centre point runs where p is the number of control parameters. The designed experiments and the recorded response values are listed in Table 2.

Thus, the 2nd order polynomial correlating independent process control parameters and responses are given in equations 1 and 2.

Table 1 Process control parameter values at different levels.

Input parameters	Symbol	Units	Coded Levels				
			-2	-1	0	1	2
Oxygen pressure	x_1	Kg/cm^2	4	5	6	7	8
Pulse width	x_2	ms	1.6	1.7	1.8	1.9	2
Pulse frequency	x_3	Hz	8	9	10	11	12
Cutting speed	x_4	mm/min	6	7	8	9	10

Table 2 Experimental design and measured responses

Expt. run	Parameter levels				Ta(deg)	Ra(μm)
	x_1	x_2	x_3	x_4		
1	0	0	0	0	0.16370	2.54
2	-1	-1	-1	-1	0.22100	2.00
3	1	-1	-1	-1	0.19372	1.60
4	-1	1	-1	-1	0.19646	2.42
5	1	1	-1	-1	0.17740	1.90
6	0	0	0	0	0.16650	2.42
7	-1	-1	1	-1	0.22650	2.78
8	1	-1	1	-1	0.16920	2.53
9	-1	1	1	-1	0.19100	2.96
10	1	1	1	-1	0.18282	2.90
11	0	0	0	0	0.15558	2.58
12	-1	-1	-1	1	0.26740	3.03
13	1	-1	-1	1	0.22100	2.46
14	-1	1	-1	1	0.33830	2.96
15	1	1	-1	1	0.15553	2.44
16	0	0	0	0	0.17189	2.64
17	-1	-1	1	1	0.30834	2.94

18	1	-1	1	1	0.18010	2.68
19	-1	1	1	1	0.31921	2.54
20	1	1	1	1	0.19920	2.05
21	0	0	0	0	0.18554	2.39
22	-2	0	0	0	0.29195	3.01
23	2	0	0	0	0.16095	2.67
24	0	-2	0	0	0.20463	2.60
25	0	2	0	0	0.16100	2.42
26	0	0	0	0	0.17740	2.48
27	0	0	-2	0	0.19650	2.81
28	0	0	2	0	0.18010	3.06
29	0	0	0	-2	0.16659	2.51
30	0	0	0	2	0.22922	3.12
31	0	0	0	0	0.15280	2.60

The models are developed using Minitab software. Mathematical model developed for minimum Kerf Taper is as follows:

$$Ta = 0.167621-0.035356\,x_1-0.004663\,x_2-0.001023\,x_3+0.023064\,x_4+0.018484\,x_1^2+0.007575x_2^2+0.008947x_3^2+0.011348\,x_4^2-0.004594\,x_1x_2-0.002558\,x_1x_3-0.022681x_1x_4+0.002551\,x_2x_3+0.006302\,x_2x_4+0.002899\,x_3x_4 \quad (1)$$

Similarly, the mathematical models developed for surface roughness is,

$$Ra = 2.52143-0.15625\,x_1-0.00875\,x_2+0.12792\,x_3+0.13458\,x_4+0.03579\,x_1^2-0.04671\,x_2^2+0.05954\,x_3^2+0.02954\,x_4^2-0.00688\,x_1x_2+0.05937\,x_1x_3-.03812\,x_1x_4-0.06937\,x_2x_3-0.14938\,x_2x_4-0.24563x_3x_4 \quad (2)$$

where, Ta and Ra are the desired responses for kerf taper and surface roughness respectively.
x_1, x_2, x_3, x_4 are the process control parameters of oxygen pressure, pulse width, pulse frequency and cutting speed respectively.

3. Optimization Using Ga

3.1 Genetic Algorrrithm

Genetic algorithm replicates the idea of survival of the fittest using an interbreeding population to create a robust search strategy. A population of solutions to a specified problem is maintained. It then iteratively creates new populations from the old by ranking the solutions according to their fitness values through the process of selection . Selection in GA is based on biological evolution where only the fittest survive and their gene pool contributes to the creation of the next generation. Hence, the likelihood of a chromosome (solution point) being selected as a good one is proportional to its fitness value. This is followed by interbreeding the fittest to create new offsprings which are optimistically closer to the optimum solution to the problem at hand. This process of crossover may be regarded as artificial mating of two fit chromosomes to create the chromosome for the next generation. The idea is some genes with good characteristics from one chromosome may combine with some good genes in the other chromosome to create a better solution represented by the new chromosome. Lastly, mutation makes random adjustment in the genetic composition. The mutation operator changes the current value of a gene to a different one. It is useful for introducing new traits in the solution pool

3.2 Optimization of LBM process using GA

The present research work optimizes the desired responses and control parameters by writing .M-files and then solved by GA using the MATLAB software. Figure 1 shows the GA output of best measured response of minimum kerf taper as 0.14695°. GA was run for 50 generations as the result remained stagnant even after increasing the number of generations further.
Three different initial population sizes were considered while running the GA. A test of 20 runs were conducted for each population size and the best five results have been shown. Table 3 lists the values of control parameters and the response predicted using GA for minimum kerf taper.

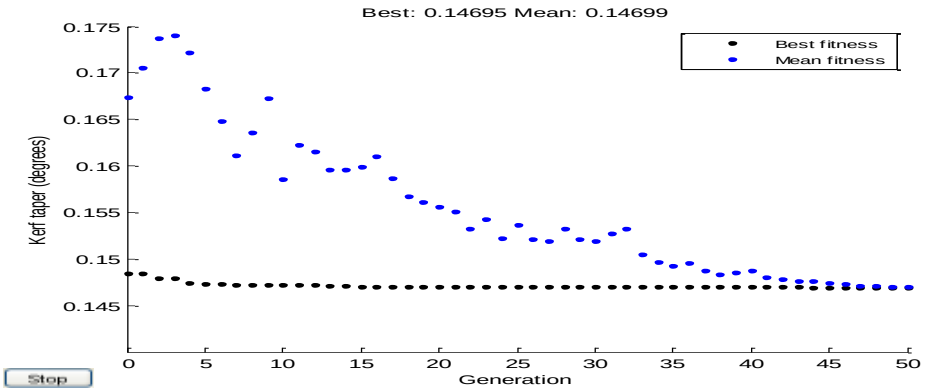


Fig.1 Plot of GA predicted result for minimum Kerf taper

Table 3 GA predicted test results for Kerf taper

Expt. No.	Process variables				Response
	Oxygen pressure. (Kg/cm^2)	Pulse width. (ms)	Pulse frequency. (Hz)	Cutting speed. (mm/min)	Kerf taper, T_a
Population size 30					
1	5	1.9	9	9	0.34724
2	5	1.7	11	9	0.33891
3	5	1.9	11	9	0.33971
4	4	1.8	10	8	0.32975
5	7	1.7	9	9	0.31837
Population size 60					
6	5	1.7	11	7	0.20013
7	5	1.7	9	7	0.20413
8	7	1.9	11	9	0.19853
9	6	1.8	10	8	0.19413
10	7	1.9	9	9	0.18765
Population size 90					
11	7	1.9	11	7	0.17452
12	6	1.8	10	8	0.16352
13	5	1.9	9	7	0.15445
14	7	1.7	9	7	0.14895
15	7	1.8	10	9	0.14695

Similarly, Figure 2 shows the GA output of best measured response of minimum surface roughness as 1.2625 μm. Also, Table 4 lists the values of control parameters and the response predicted using GA for minimum surface roughness.

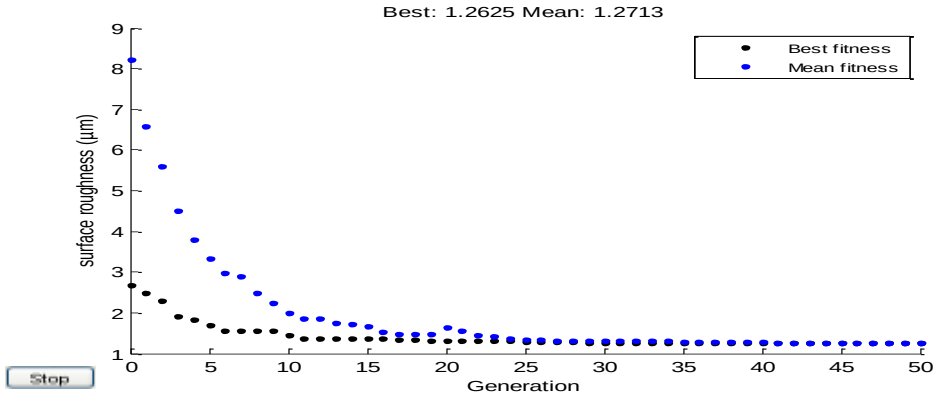


Fig. 2 Plot of GA predicted result for minimum Surface roughness

Table 4 GA predicted test results for surface roughness

Expt. No.	Process variables				Response
	Oxygen pressure. (Kg/cm^2)	Pulse width. (ms)	Pulse frequency. (Hz)	Cutting speed. (mm/min)	Surface roughness, Ra
Population size 30					
1	5	1.7	9	9	3.9567
2	4	1.8	10	8	3.7532
3	6	1.8	10	8	3.7256
4	6	1.8	10	9	3.5245
5	6	1.9	10	8	3.4245
Population size 60					
6	6	1.8	10	8	2.9785
7	5	1.7	11	7	2.8684
8	5	1.9	11	7	2.8622
9	5	1.9	9	9	2.6231
10	6	1.8	10	8	2.4146
Population size 90					
11	6	1.8	10	8	2.3856
12	5	1.7	9	7	2.2345
13	8	1.8	10	7	1.7562
14	7	1.9	9	7	1.8553
15	7	1.7	9	7	1.2625

3.3 Validation

Validation of the GA predicted results with the experimental results is done in order to conform the GA predicted results to be acceptable for practical use. Percentage of prediction error shows the amount of variation with the actual experimental results. The percentage of prediction error is calculated as

Prediction error%

$$= \frac{\text{Experimental result - GA predicted result}}{\text{Experimental result}} \times 100$$

In order to validate the test results predicted by GA, six random experimental results are compared with the GA predicted results as shown in Table 5.

Table 5 Comparison of experimental and GA predicted results.

Sl.no.	Experimental result		GA predicted result		Prediction error %	
	Kerf taper	Surface roughness	Kerf taper	Surface roughness	Kerf taper	Surface roughness
1	0.15552	1.9	0.14695	1.83	5.51	3.68
2	0.221	2.53	0.23	2.6	3.91	3.84
3	0.19372	3.01	0.189	2.95	2.43	1.99
4	0.30834	3.12	0.3112	3.25	0.919	4
5	0.29195	2.78	0.2845	2.65	2.543	4.67
6	0.3383	2.05	0.3554	2.18	4.811	5.96
Average percentage of error					3.35	4.02

Figures 3 and 4 show the plot for comparison of experimental results with the GA predicted results for minimum kerf taper and surface roughness respectively.

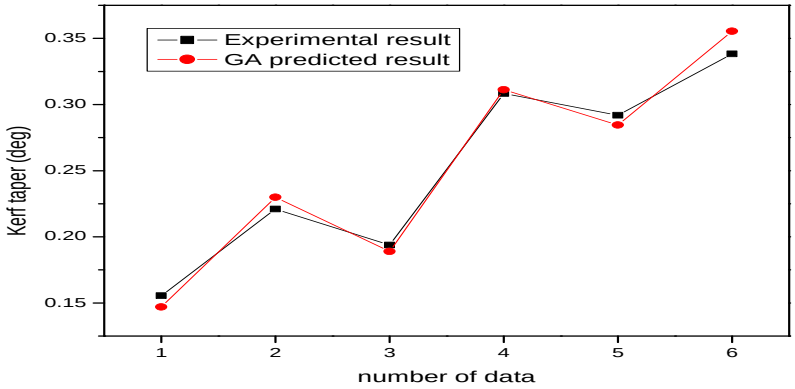


Fig. 3 Plot for comparison of experimental and GA predicted results for minimum kerf taper

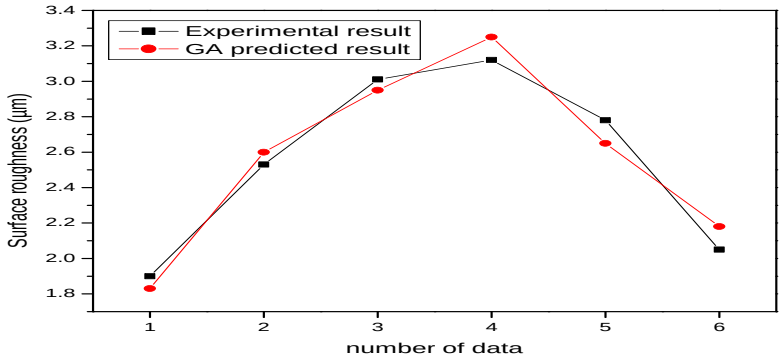


Fig. 3 Plot for comparison of experimental and GA predicted results for minimum surface roughness.

4. Result and Analysis

Heuristic analyses using GA for optimizing the cut quality namely kerf taper and surface roughness during pulsed Nd:Yag laser cutting of thin Al-alloy sheet for straight profile is performed. Tables 3 and 4 list the values of process control parameters and the GA predicted responses of kerf taper and surface roughness respectively. It is observed from these tables that as the population size (possible solutions) in GA increases, the responses decrease showing improvement of the desired response quality. This can be attributed to the fact that more number of possible solutions provide opportunities of reproducing better offsprings or solutions.

However, the result remains stagnant when population size of more than 50 is used. Thus, global optima is achieved at population size 50 and no further improvement in the response values are attained by further increase in the population size.

This research shows the comparison of results between GA and past researches. Tables 5 highlight the comparison of outcome for optimality analyses. On comparison of the test results for the desired responses, GA based optimality analysis achieve better fitness function values as compared to those derived by the past researchers. Through GA, minimum kerf taper obtained is 0.14695° which is 0.313° less in magnitude than experimentally measured value. The result suggests that average to low values of oxygen pressure and pulse width combined with average to high values of pulse frequency and cutting speed gives optimal results for minimum kerf taper. Similarly, using GA the minimum surface roughness predicted at optimal parameter setting is 1.2625µm which is 0.3375µm better in value compared to the experimental measured value. The result suggests that a medium high value of oxygen pressure and medium low values of pulse width, pulse frequency and cutting speed are to be set for obtaining better or minimum surface roughness.

The tabulated values while comparing experimental and GA predicted results at RSM predicted combination of optimal parametric setting are listed in Table 6.

Table 6 Comparison of experimental and GA predicted results at RSM predicted combination of Optimal parametric setting

Results	Expt.	GA
Kerf taper	0.15553	0.14695
Surface roughness	1.6	1.53159
Percentage error (%) for kerf taper		5.5
Percentage error (%) for surface roughness taper		4.27

From this table, it is observed that GA based predicted results at optimal parametric setting is closer to the values as measured by actual experiments. This is shown by percentage prediction error which is 5.5 % and 4.27 % for kerf taper and surface roughness respectively. Thus, it is observed for GA to be a cost effective, robust yet simple and fast method for optimizing the process control parameters of a Nd:Yag LBM for desired responses of minimum kerf taper and minimum surface roughness.

5. Conclusion

The following conclusions are drawn from this research work.

(i) GA provides a cost effective soft computing technique for optimizing machining operations.
(ii) Based on the test results predicted using GA, this technique can be accommodated within an intelligent manufacturing system for automated process planning.
(iii) future work can be done taking into consideration more independent control parameters.
(iv) Multiobjective optimization can be taken as further work.

References

[1] Avanish Kumar Dubey and Vinod Yadava, "Laser beam machining—A review,"International Journal of Machine Tools & Manufacture, vol. 48, 2008, pp.609– 628.

[2] A.S. Kuar, B. Doloi, B. Bhattacharyya, " Modelling and analysis of pulsed Nd:YAG laser machining characteristics during micro-drilling of zirconia (ZrO2)," International Journal of Machine Tools & Manufacture, vol. 46,2006, pp. 1301–1310.

[3] Avanish Kumar Dubey and Vinod Yadava, "Optimization of kerf quality during pulsed laser cutting of aluminium alloy sheet," Journal of materials processing technology, vol. 204, 2008, pp.412–418.

[4] Amit Sharma,VinodYadava, Raghavendra Rao, "Optimization of kerf quality characteristics during Nd: YAG laser cutting of nickel based superalloy sheet for straight and curved cut profiles," Optics and Lasers in Engineering, vol. 48,2010, pp. 915–925.

[5] Mukherjee, I and Ray, P.K, "A review of optimization techniques in metal cutting processes", Computers& Industrial Engineering, Vol. 50, 2006, pp. 15-34.

[6] D.I. Lalwani, N.K. Mehta, P.K. Jain, "Experimental investigations of cutting parameters influence on cutting forces and surface roughness in finish hard turning of MDN250 steel," Journal of materials processing technology, vol.206, 2008, pp.167–179.

[7] A. Soveja, E. Cicala, D. Grevey, J.M. Jouvard, " Optimization of TA6V alloy surface laser texturing using an experimental design approach," Optics and Lasers in Engineering, vol. 46, 2008, pp. 671– 678.

[8] Avanish Kumar Dubey and Vinod Yadava, " Multi-objective optimization of laser beam cutting process," Optics & Laser Technology, vol. 40, 2008, pp. 562–570.

[9] Ali R. Yildiz, "A comparative study of population-based optimization algorithms for turning Operations," Information Sciences, vol. 210, 2012, pp. 81–88.

[10] Norfadzlan Yusup, Azlan Mohd Zain, Siti Zaiton Mohd Hashim, "Evolutionary techniques in optimizing machining parameters: Review and recent applications (2007–2011)," Expert Systems with Applications, vol. 39, 2012, pp. 9909–9927.

[11] Suman Samanta and Shankar Chakraborty, "Parametric optimization of some non- traditional machining processes using artificial bee colony algorithm," Engineering Applications of Artificial Intelligence, vol. 24, 2011, pp. 946–957.

[12] Neelesh K. Jain, V.K. Jain, Kalyanmoy Deb, " Optimization of process parameters of mechanical type advanced machining processes using genetic algorithms," International Journal of Machine Tools & Manufacture, vol. 47, 2007, pp. 900–919.

[13] Sharma, A. and Yadava,V. (2012) "Modelling and optimization of cut quality during pulsed Nd:YAG laser cutting of thin Al-alloy sheet for straight profile", *Optics & Laser Technology*, Vol. 44, 2012, pp. 159–168.

Waste Source Separation Management for Urban Mining: A Change Strategy to Improve Quality

[1]Helen Morabi Heravi[a], [2]Mohammad Reza Sabour[b]

[A] (Environmental Management, Environmental Study/ Universiti Putra Malaysia, Malaysia 43300)
[B] (Environmental Engineering, Civil Engineering/ K.N.Toosi University, Tehran, Iran 1346)

Abstract : The aim of Urban Mines is to provide information, expert advice, support, direction and financial solutions to waste, energy, and climate change problems [1]. There are several effective ways of separating recyclable materials; one possible way is to separate it as mixed waste in suitable plants i.e. curbside, MRF (Material Recycling Facility) or landfill. It identifies, evaluates, analyzes past experiences and their compliance with new methods of solid waste management based on social, economic, technical, environmental and health considerations associated with the management concerns that have a positive and valuable effect in solving problems and optimization of source separation schemes in large cities. The aim of this study is to evaluate past experiences in Tehran, study positive and negative points of view of different methods and provide a model for source separation. On the basis of an experimental work and the merits of a new system, results suggested an "integrated waste source separation system" should be adopted. In preparing a new plan on the resolution of the problem previous experience in Tehran was examined, as well as the other metropolitan cities in Iran and similar projects that have been implemented in other countries.

Keywords - Waste management, Waste source separation, Urban mining

INTRODUCTION

The concept of Urban Mining challenges the status quo. Urban mining is the process of reclaiming compounds and elements from products, building and waste. This definition is based on the concept of Sustainable Development, the fact that our ecosystem is finite, non-growing and materially closed and uses industrial ecology, which examines materials and energy flows in products, processes, industrial sectors and economies [2]. Urban mining includes actions that can recover resources from waste as secondary materials or energy by separate collection or recovery of resources etc. [3, 4].Most countries that have prepared reports on recycling household waste undertook them to assist those involved in the waste management services and recommended collection techniques required for the recovery of recyclable materials such as urban mining. The effectiveness of each regulatory schema studied and a brief review of some of many countries have shown that both source separation and resource recovery schemes need to have an active long-term role and a national strategy [5]. The results of the research for information on source separation could not answer the question as to whether source separation of mixed waste should be before or after collection, or whether both should be pursued in parallel. The results were based on very limited existing information and evidence [6].

The aim of the waste source separation of paper was gathering qualitative information on the waste collection experiences and planning new methods of source separation. Source separation schemes are a step in developing new environmental policy, the use of new technology and the development of social services. Now several local contractors pay for the cost of the selected material (bread and ferrous-metal) in Tehran. Markets for the sale of all recyclables should be secured - this is the first recovery strategy that has priority over other factors. Also it should be ensured that there are incentive laws for the products collected for use. Skills and equipment are effective factors for reduction or increase of the amount to be collected. If the income from this process does not cover the cost it will not find acceptance and the only choice will be collection as mix-waste. In the present study, the objective is to investigate points of strength and weakness in strategies adopted between the years 1997 to 2006 and from 2006 to 2010 in Tehran to help the local and regional managers to establish the best collection system for household waste. The study covers source separation in the residential area and not market issues.

I. MATERIAL AND METHODS

In Tehran, an average of 7,000 metric tons of waste are generated per day by each of the municipal regions. Currently, more than two and a half million tons of waste annually is generated in Tehran. According to the waste analysis that was carried out, 32% of dry material is recyclable. The promotion of source separation of recyclable materials will be of help in the recycling and prevention of environmental pollution and loss of national capital. Several waste source

separation trials have been conducted in Tehran in recent years and were divided into two groups which have been well documented.

a. Traditional strategies (1997 to 2006)

Resource recovery from mixed waste is the traditional method of separation that has been able to attract the most participation of people. The major obstacles to effective dry waste collection operations in each of these systems are examined. The implementation of better plans is considered separately in addition to traditional systems. Different ways of motivating and encouraging people to cooperate with the municipality is a necessary requirement of any strategy adopted. Waste pickers in the city collected waste door-to-door and purchased waste (bread and ferrous-metal) separated by citizens. Waste pickers in this system, especially those collecting paper and plastic recyclables from contaminated sources such as household garbage and hospital waste negatively affected the quality of recycled products collected. Equipment used in this method includes a wide range of handheld devices such as carts, as well as pickup trucks. In this program, citizens recycling material such as bread ,glass and ferrous-metal from home and then selling them for cash or exchanging them for other materials and separated recyclable materials in the following ways:

At source using specific recycling centers and stores; returning waste to the producers and institutions and advertising it for sale or free collection in newspapers; Taking it to storage sites and making use of collection and transportation services provided by itinerant workers and by pickup from the door ;Disposing of it through municipal refuse collection services at landfill sites.

b. Waste source separation strategies (2006 until 2010)

Tehran's waste management organization in most areas monitored performance of contractors in order to try to make the process more efficient. The separation process was originally done with the evaluation contractor. Activities carried out in this process included separation, collection and transportation. Scavengers often purchased goods from the citizen, and then transferred them to a collection center where they could sell them at a higher price. Contractors also offered these materials to other facilities at higher prices.

Between the years 2006 to 2010, several schemes were implemented for collecting recyclable materials by the Tehran waste management organization as a pilot in different areas of Tehran. Trialing a system before full-scale implementation is generally recommended. Short term trials were performed in Tehran as part of this scheme and were divided into four groups:

Improvement of traditional separating by existing contractors; Use of public service vehicles for collecting dry waste each day; Use of vehicles with towable containers for collecting wet and dry waste ; Use of public service vehicles for collecting dry waste at night .

The buyback center system, door-to-door services and curbside sorting are the source separation initiative undertaken in Tehran City. The area covered is 22 regions of the city. Buyback centers are the most common source separation system currently in operation in Tehran, and are the preferred approach for high-density communities. Buyback centers are privately operated. A standard one-ton recyclable materials collection pick-up truck with one driver and one container is used, from which the collected material is sorted at the buyback center. The citizens keep recyclable materials separated from other refuse, and can also deliver these materials to the buyback center during its hour's operation (5:00-7:00 p.m. in the afternoon). With the cooperation of the buyback center system, recyclable material is collected as co-mingled materials of paper, glass, metal, and plastic. Citizens who keep their recyclable materials separate and sell it to the center receive either promotional materials or a ticket for earn money, as described by the scheme in Figure 1.

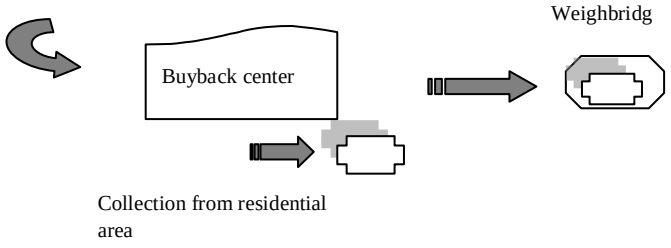


Figure 1: Source separation by the buyback centers

The curbside sorting collection is a reuse system undertaken in Tehran. Curbside collection is made daily. Curbside collection sorting of recyclable materials was started as a youth employment project over 40 years ago. Materials are left at the curbside in a co-mingled state. At the MRF plant, sorted material is separated into specific categories which can be sent directly for recycling. A motorized three-wheeled vehicle with one container is used for collecting recyclable material at the curbside, passing four times daily. In co-mingled collection, sorting will result in a quality product to sell to industry. In curbside co-mingled collection, all recyclable materials are collected except for bread, and

non-ferrous metal in the existing wheeled bins. Glass, because of the risk of breakage during due to both collection and storage and mixing with other refuse is often not collected in collection systems. The 22 areas of the city are divided into several posts per week, and the vehicles are assigned to these areas and their specific posts, as described by the scheme in Figure 2.

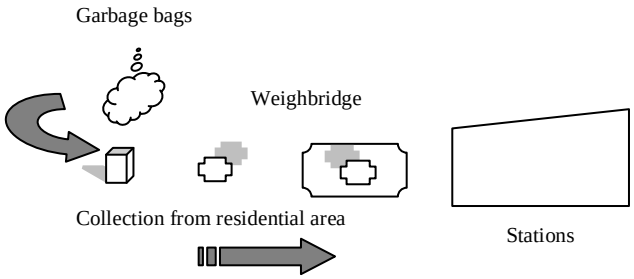


Figure 2: Curbside sorting

The door-to-door services collection system is a reduction system undertaken in Tehran; the coverage area is 22 regions of Tehran City. The collection vehicle is a 500 kg pickup truck; this vehicle has curbside loading and is unloaded by hand. The sorting position is available to the collector outside the body of the vehicle. Weekly loading is utilized. Citizens deliver recyclables using blue polyethylene bags, with recyclable materials separated as a co-mingled for the scavenger. For a building complex, a wheeled 60-, 120-, or 240-liter bin is preferred; these are used for the collection of co-mingled clean recyclables. The number of collections at each post is a two times daily. By undertaking this system, we can achieve separation of glass. The door-to-door system is supported by a private sector contractor chosen by regional municipalities, as described by the scheme in Figure 3.

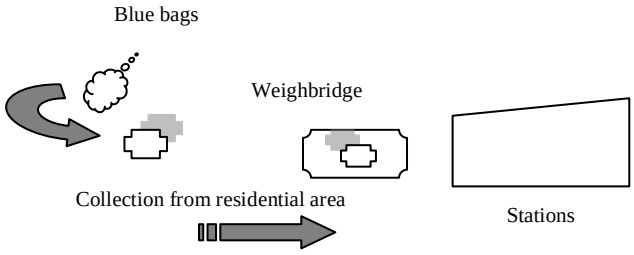


Figure 3: Source separation by the door-to-door services

At the MRF (Material Recovery Facility) plant, recyclable material is separated into specific categories which can be marketed directly. An MRF will be required to deal with the unsorted products of the collection system, and this has not yet been detailed or financially programmed. The ability to achieve realistic standards of products from a co-mingled collection is as yet unproven. Service stations located in urban areas and all mechanical devices belong to the waste management organization. Silos are located at an intermediate station for storage, and separation of collected recyclable materials has weighing and administrative locations divided by types of materials and machinery. More recyclable materials are collected by pickup track or three-wheeled vehicles that belong to the contractors than are collected and carried by truck to recycling facilities. More recyclable materials are discharged onto the platforms; then the mixed refuse is unloaded onto conveyors, where workers hand-pick some easy-to-remove items.

All waste separation is done by the workers immediately after they receive the recyclable materials at the station. Garbage bags are broken, and the refuse is separated into components using various combinations of magnetic separation and hand-sorting. Segregated materials include: PET, various plastics, ferrous and nonferrous metals, glass, paper, and cardboard. Funeral waste is discharged into semi-trailers and sent to landfills. Recyclable materials such as ferrous metal, cardboard, and aluminum cans are pressed and baled, then loaded onto trucks and transferred to recycling facilities. The MRF is operated by the solid waste management organization. The tonnages handled emanate from the curbside, door-to-door service, buyback center, and office paper collection rounds as well as the product, as described by the scheme in Figure 4.

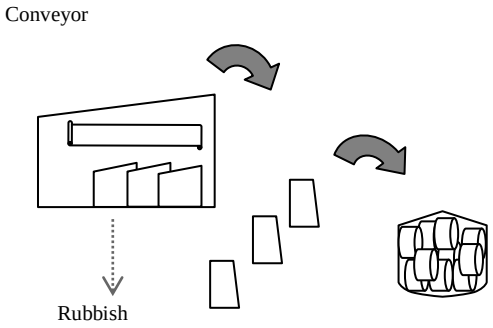


Figure 4: Separation at the MRF plant

II. RESULTS

Results in this paper show that carefully implemented waste collection systems should be successful, if they had high participation rates and were carefully managed. This paper identifies strengths and weaknesses of traditional and new schemes of waste collection in Tehran, and identifies matters required to implement a successful collection system. Identifying points of strength and weakness of the traditional approach help to ensure that the best way forward is found. Some reasons for failure in resolving the problems of traditional schemes implemented in Tehran before the summer of 2006 are:

a) Lack of coordination between municipalities (22region) and the multiplicity of authorities involved in decision making
b) Lack of clearly defined goals (short term and transient)
c) Lack of a systematic approach to schemes and applying non-expert opinion in decision-making
d) Lack of the necessary support to provide facilities, equipment and financial resources for the work involved

Local conditions in each country will affect the choice of the most viable waste collection system. Consideration of market issues is also an important factor in the choice of waste-collection systems. In the past there were not many differences between the buying and selling of recyclable materials. Now contractors offer their materials to recycling facilities and manufacturers at higher prices. The reasons for the success of traditional systems may be cited as follows:

a) Lower costs and therefore a real and higher purchase price
b) The direct purchase system of the people
c) It was known as the native form, a very strong relationship with people and was usually punctual (true to one's promise)

III. DISCUSSION

In preparing a new plan on the resolution problem previous experience in Tehran has been considered and also plans from other metropolitan cities in Iran have been examined as well as the implementation of similar projects in other countries. Positive and negative points from all these examples were studied from social, economic, technical, environmental and health points of view. The performance of suitable methods of source separation were considered from different parameters such as the type of urban location, type of roads, as well as the social, cultural and economic levels of households. The recommended actions to be taken to solve the problem were divided in two ways: firstly, the effect on solid waste management system elements and secondly an emphasis on public awareness.

Eliminating the traditional system was possible, by attracting people to alternative clearing methods. This could include buying waste from them or incentivizing donations through the award of prizes or points that can be redeemed for goods and services; restrictions on waste disposal by the enforcement of local regulations using a more coercive approach that could involve financial penalties prescribed by a court of law; establishing area offices for recycling ; creating tax incentives for recycling and correct waste disposal ; creating cultural incentives that encourage the protection and preservation of the local environment and human health; creating awareness of ethical considerations; encouraging awareness of religious and moral requirements for cleanliness and correct disposal of waste; making it a matter of national pride. By using all the integrated and overlapping motivational techniques suggested here and promoting them through slogans, advertising and public service messages that can be combined to create the most powerful campaign that can win hearts and minds and create a social transformation. A solid waste management system was tried providing an "interim plan for the collection of dry and wet waste (through source separation)"; Recycling Offices would be established in 22 regions of the Tehran municipality; a "technical and performance manual for collecting dry waste" would be prepared; reform of "the technical and performing system of collecting dry waste" would take place; reform of "technical and performing system with a new approach to the integrated collection of dry waste" would also be designed.

IV. CONCLUSION

To achieve their aims the Tehran waste management organization should put more emphasis on the technical performance of each contractor by ranking them to create more motivation for effective collecting and separating of waste (based on tonnage). It also should create more incentives for private sector investment and participation. The municipality should provide systematic surveillance, and maximum participation by authorities in the process of creating more effective, more desirable and more efficient services of contractors. Contractors should perform their services based on time table and mechanized dry waste collections also performing an integrated approach to ensure no overlapping of services. There should be better use of economic incentives that have an effect on performance. There further needs to be better education and information provided through effective supervision by the councils. Additionally, there must be improved coordination and integration of dry waste management. Finally, the systems must be evaluated on their levels of planning and performance.

Tehran waste management organization's best approach is known as "integrated waste source separation" using buyback centers, a door-to-door service, and storage containers. These systems use a standard method of source separation under consumer responsibility.

V. Acknowledgements

The authors gratefully acknowledge valuable comments on an earlier version of the manuscript by experts from the recycling office of Tehran waste management organization.

REFERENCES

[1] Urban Mines team,2012 , http://www.urbanmines.org.uk/

[2] Thomas E. Graedel, (2011), What is Urban Mining?, Yale School of Forestry and Environmental Science, takes it further and includes energy as part of the concept of urban mining.

[3] Cossu R. (2012) The environmentally sustainable geological repository: The modern role of landfilling. Waste Management, vol.32, 243–244

[4] Baccini P., P.H. Brunner (2012) Metabolism of the Anthroposphere - Analysis, Evaluation, Design. The MIT Press, Massachusetts Institute of Technology, Cambridge ISBN 978-0-262-01665-0

[5] Coleman. R.W,C.G.Hill,A.Marsden,(1991),*Recycling Household Waste: the way ahead ,Association of Municipal Engineers of the Institution of Civil Engineers* (Thomas Telford ,London E144JD,printed by Faygate printing services , British Library)

[6] MIESZKIS. K. W. and F. E. THOMAS, (1980), *Conservation & Recycling, Pergamon* (The Welsh School of Architecture, UWIST, Cardiff, U.K. Press Ltd., 1980. Printed in Great Britain. Vol. 3, pp. 413 -425)

Load Flow Analysis of Distribution System Including Wind Turbine Generating System Models

P.Srihari[1], G.Srinivasa Rao[2]

[1,2],Post Graduate Student , Department of EEE, V R Siddhartha Engg. College (Autonomous), Vijayawada-7.

Abstract— The power world is sauntering towards eco-friendly distributed generation (DG), their integration into the distribution network postures challenges to existing load flow techniques. This paper is concerned with developing models of various types of wind turbine generating systems (WTGSs) used as distributed generation (DG) sources and demonstrating their application for steady state analysis. As wind energy is used as a DG power source, its impact on the distribution system needs to be analyzed. This requires a load flow algorithm capable of simulating both radial and weakly mesh systems along with Wind Turbine Generating Systems (WTGSs). The compensation based distribution load flow algorithm is extended including WTGS modeling in the present work. The application of the proposed models for the load flow analysis of radial systems having WTGS has been demonstrated. Based on these studies we evaluate the impact of wind based DG on the voltage profile and losses of radial distribution networks. Simulation studies have been conceded out on radial distribution systems having WTGS as DG sources to illustrate the application of the proposed models.

Keywords— Distributed Generation, Distribution load flow, Wind Turbine Generating Systems, Forward and Backward sweep, Voltage Profile, Losses.

I.INTRODUCTION

The supply of electric power to homes, offices, schools, factories, stores, and nearly every other place in the modern world is now taken for granted. Electric power has become a fundamental part of the infrastructure of contemporary society, with most of today's daily activity based on the assumption that the desired electric power is readily available. The power systems which provide this electricity are some of the largest and most complex systems in the world. They consist of three primary components: the generation system, the transmission system, and the distribution system. Each component is essential to the process of delivering power from the site where it is produced to the customer who uses it. Wind energy continues to be one of the fastest growing energy technologies and it looks set to become a major generator of electricity throughout the world.

The role of wind energy in electrical power generation is increasing day by day and new advances in power electronic Equipment's are underpinning this strategy. Wind energy is converted into mechanical power by wind turbines. This mechanical power is fed to the generator directly or through gear system. This cumulative arrangement of turbine and Generator is called a wind turbine generating unit (WTGU). A large number of wind turbine generating systems (WTGSs) are already in operation and many new systems are being planned. The integration of WTGS with the power systems is taking place at both the transmission and the distribution voltage levels. This growth in wind generation has spurred investigations, to understand the behavior of the wind turbine generating Systems (WTGSs) as well as their impact on the power grid.

As distributed generation penetrates into the distribution network its impact on the voltage profile and total power losses needs to be analyzed by appropriate load flow methods. Due to high R/X ratio of distribution lines, the performance of load flow methods used for the transmission network are inadequate for distribution systems. These methods do not make use of the structural advantage (radial / weakly mesh) of distribution system. This led to the development of separate load flow methods for distribution system. These methods are broadly classified into three categories: direct methods, backward / forward sweep methods and Newton-Raphson based methods [1]. The backward/forward sweep methods have an edge over the others as they are simple, flexible and the computational time requirement is low [1].

The basic backward/forward sweep method described in [2] is applicable for radial networks only. The practical distribution systems are weakly meshed. The compensation based distribution load flow algorithm proposed in [3] address this issue. It is improved to adapt the PV bus modeling in [4]. The method was further extended to the distribution networks containing DG's, where these are modeled as constant PQ or PV nodes [5].

Ii. Wtgs Modeling

The mechanical power converted from wind energy by a wind turbine is described as

$$P_m = \frac{1}{2}\rho A C_p(\lambda,\upsilon)V^3 \quad (1a)$$

$$\lambda = \frac{\omega \eta R}{v} \quad (1b)$$

Where, Pm is the mechanical power developed in watts, ρ is density of air (kg/m^3), V is wind speed(m/s), C_p is power coefficient, υ is pitch angle, λ is tip speed ratio, η is gear ratio and A is the area swept by the rotor. ω is the angular velocity(rpm),R is the turbine rotor radius(m) and $c_1 - c_6$,x are constants.

The variation of power coefficient C_p with a variation of υ, λ is non linear in nature. It can be estimated by the relation.

$$C_p(\lambda,\upsilon) = c_1\left(c_2\frac{1}{\wedge} - c_3\upsilon - c_4\upsilon^x - c_5\right)e^{\left(\frac{-c_6}{\wedge}\right)} \quad (1c)$$

Where, $\frac{1}{\wedge} = \frac{1}{\lambda+0.08\upsilon} - \frac{0.0035}{1+\upsilon^3}$ (2)

The induction generators are generally used in wind turbine for electricity generation and modeled with standard equivalent circuit from using the concept of a rotating transformer [6–9]. In the study of [6], Feijoo and Cidras proposed two models based on the steady- state equivalent circuit of the induction generator. In the study, wind farms with asynchronous generators are modeled as PQ or RX buses in the load flow analysis. When the WTGS node is modeled as a PQ-bus, produced active and consumed reactive power. The other model, in that study, is called RX bus model in which active and reactive powers are calculated by using equivalent circuit parameters of the induction machine. Both are solved iteratively using a Newton type algorithm. Implementation of RX model into distribution system load flow analysis is found to be problematic, therefore, excluded from this study and the PQ model is used for the analyses.

Model-I for WTGS:

Let us consider an induction motor equivalent circuit referred to the stator side as given in Fig. 1. In the figure, V_S and V_r stand for the magnitude of the stator voltage and referred rotor voltage, respectively. R_{SC} and X_{SC} show short-circuit equivalent resistance and reactance, respectively. R_m resistance represents core losses and X_m shows magnetizing reactance of the motor. P_S and Q_S stand for the active and reactive power demand of the motor, respectively. From figure, recognizing the fact that $Q_r = 0$, the referred rotor voltage (V_r) can be obtained by solving well-known bi-quadratic equation given in [10] by neglecting the short-circuit losses as follows.

$$V_r = \sqrt{\frac{V_s^4 - 2P_m R_{sc} V_s^2 + P_m^2\left(R_{sc}^2 + X_{sc}^2\right)}{V_s^2}} \quad (3)$$

From the equivalent circuit, given, in Fig.1,the reactive power balance can be written as follows

$$Q_s = X_{sc}\frac{P_m^2}{V_r^2} + \frac{V_s^2}{X_m} \quad (4)$$

Substituting the voltage equation, Eq.(3) into Eq(4), we get

$$Q_s = \frac{V_s^2}{X_m} + \frac{X_{sc}P_m^2 V_s^2}{V_s^4 - 2P_m R_{sc} V_s^2 + P_m^2\left(R_{sc}^2 + X_{sc}^2\right)} \quad (5)$$

Similarly, from the equivalent circuit, the active power balance can be written as

$$P_s = P_m + R_{sc}\frac{P_m^2}{V_r^2} + \frac{V_s^2}{R_m} \quad (6)$$

And, substituting voltage equation, Eq.(3), into Eq.(6), we get the active power demand of the machine as follows

$$P_s = P_m + \frac{V_s^2}{X_m} + \frac{R_{sc}P_m^2 V_s^2}{V_s^4 - 2P_m R_{sc} V_s^2 + P_m^2\left(R_{sc}^2 + X_{sc}^2\right)} \quad (7)$$

Equ.s (5) &(7) have a straight solution and depend on the terminal voltage magnitude, mechanical power input and the equivalent circuit [6-8].

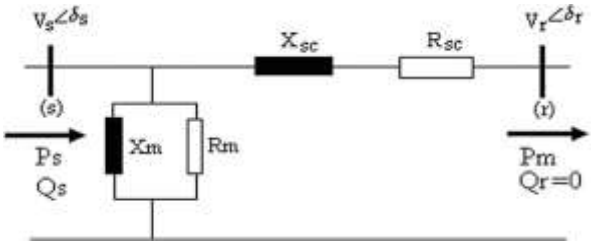


Fig.1. Induction motor equivalent circuit.

Model-II for WTGS:

Similarly, the referred rotor voltage (V_r) of the induction motor, given in Fig. 1, can also be obtained by using modified bi-quadratic equation given in [12] as follows

$$\Phi_r = tan^{-1}\left(\frac{Q_r}{P_r}\right) \quad (8a)$$

$$K = V_s^2 - 2P_r(R + Xtan\Phi_r) \quad (8b)$$

$$V_r = \sqrt{\frac{K \pm \sqrt{K^2 - 4(R^2+X^2)P_r^2 sec^2\Phi_r}}{2}} \quad (8c)$$

From Fig.1,recognizing the fact that Q_r=0,the referred rotor voltage can be written as follows

$$\Phi_r = 0 \quad (9a)$$

$$K = V_s^2 - 2P_r R \quad (9b)$$

$$V_r = \sqrt{\frac{K \pm \sqrt{K^2 - 4(R^2 + X^2)P_r^2}}{2}} \quad (9c)$$

Substituting the voltage equation, Eq.(9c) into Eq.(4), we get

$$Q_s = \frac{V_s^2}{X_m} + \frac{2X_{sc}P_m^2}{V_s^2 - 2P_m R_{sc} - \sqrt{V_s^4 - 4P_m R_{sc} V_s^2 - 4X_{sc}^2 P_m^2}} \quad (10a)$$

$$Q_s = \frac{V_s^2}{X_m} + \frac{2X_{sc}P_m^2}{V_s^2 - 2P_m R_{sc} + \sqrt{V_s^4 - 4P_m R_{sc} V_s^2 - 4X_{sc}^2 P_m^2}} \quad (10b)$$

Similarly, Substituting the voltage equation, Eq.(9c) into active power balance, Eq.(6),we get

$$P_s = P_m + \frac{V_s^2}{R_m} + \frac{2R_{sc}P_m^2}{V_s^2 - 2P_m R_{sc} - \sqrt{V_s^4 - 4P_m R_{sc} V_s^2 - 4X_{sc}^2 P_m^2}} \quad (11a)$$

$$P_s = P_m + \frac{V_s^2}{R_m} + \frac{2R_{sc}P_m^2}{V_s^2 - 2P_m R_{sc} + \sqrt{V_s^4 - 4P_m R_{sc} V_s^2 - 4X_{sc}^2 P_m^2}} \quad (11b)$$

Likewise, Equations (10) and (11) have also a straightforward solution and depend on terminal voltage magnitude, mechanical power input and the equivalent circuit parameters of the induction machine except the rotor slip. It is noted that for the induction generators the mechanical power is negative due to the change on the flow of power which is transferred from ''rotor'' (bus-r) to the ''stator'' (bus-s). Hence, the mechanical power input, P_m, must be taken in negative sign for both developed models.

Since there are two solutions of real and reactive power outputs of the developed Model-II[11], the required root of the equation, given in Equations (10) and (11), must be determined. For this reason, variation of the active and reactive power output of the generator with different mechanical power input (P_m) are calculated using both expressions, for a hypothetical machine. On the other hand, the solution of Equations (10b) and (11b) remain on the feasible region. Therefore, the solution of Equations (10b) and (11b) is unique for the reactive and active power outputs of the induction generator, respectively. They can facilitate the computation of real and reactive powers of the induction generator for a specified mechanical power input and terminal voltage, in a simple way.

Presently various types of WTGU have been installed and they can be broadly classified into three categories, namely fixed, semi-variable and variable speed types[7]. The models developed here for the WTGU are intended to obtain the power output of the WTGU for a given terminal voltage and wind speed.

Pitch regulated fixed speed WTGU[7]

In this class of WTGU, the pitch angle controller regulates the wind turbine blade angle (ν) according to the wind speed variations. This designed power output Pe of the WTGU with wind speed is provided by the manufacturer in the form of a power curve. Hence, for a given wind speed Pe can be obtained from the power curve of the WTGU, but Qe needs to be computed. With the slip known, the reactive power output Qe is calculated from the induction generator equivalent circuit. Pitch regulated fixed speed WTGU power output (Pe and Qe)

$$P_e = \frac{[R_1(R_2^2 + s^2(X_m + X_{12})^2) + sR_2X_m^2]|V|^2}{\left[R_2R_1 + s\left(X_m^2 - (X_m + X_{12})(X_m + X_{12})\right)\right]^2 + [R_2(X_m + X_{12}) + sR_1(X_m + X_{12})]^2} \quad (12)$$

$$Q_e = \frac{[X_mX_{12}s^2(X_m + X_{12}) + X_{12}s^2(X_m + X_{12})^2 + R_2^2(X_m + X_{12})]|V|^2}{\left[R_2R_1 + s\left(X_m^2 - (X_m + X_{12})(X_m + X_{12})\right)\right]^2 + [R_2(X_m + X_{12}) + sR_1(X_m + X_{12})]^2} \quad (13)$$

Semi Variable Speed WTGU [7]

This class of WTGU consists of a pitch controlled wind turbine and a wound rotor induction generator. The rotor circuit of the generator is connected to an external variable resistance. There are two controllers, a pitch controller and rotor resistance controller. These two controllers are designed to operate in a coordinated manner. This design guarantees that the active power output is equal to the maximum power at wind speeds below nominal and equal to rated power above nominal wind speeds. For this class of WTGU also, the manufacturer provides the designed real power out-put versus wind speed characteristics [7]. We note that the expression for Pe and Qe can be recast as a quadratic function of a new variableR2/s. Hence, even when R2 and s are unknown the quantityR2 /s (say Req) can be computed by solving the quadratic equation in Req involving Pe. To compute Qe, this value of Req is used in the modified expression for Qe. The quadratic equation for Req is as follows:

$$aR_{eq}^2 + bR_{eq} + c = 0 \quad (14)$$

Where

$$a = P_e(R_1^2 + (X_{11} + X_m)^2 - |V|^2R_1^2) \quad (15)$$

$$b = 2R_1P_eX_m^2 - X_m^2|V|^2 \quad (16)$$

$$c = P_e\left(R_1^2 + (Z_{11} + Z_m)^2 + P_e(Z_e^2 - (Z_m + Z_{22})(Z_m + Z_{11}))^2 - R_1(Z_m + Z_{22})^2|V|^2\right) \quad (17)$$

Given wind speed u_w and the terminal voltage V

1. For the given u_w obtain Pe from the power curve of the WTGU (provided by the manufacturer).

2. Compute $R_{eq} = \min \left| \frac{-b \pm \sqrt{b^2-4ac}}{2a} \right|$ (18)

3. Knowing Req , compute Reactive power Qe as

$$Q_e = \frac{[R_{eq}^2(X_m+X_{l1})-(X_m+X_{l2})\cdot(X_m^2-(X_m+X_{l2})(X_m+X_{l1}))]|V|^2}{[R_{eq}R_2+(X_m^2-(X_m+X_{l2})(X_m+X_{l1}))]^2+[R_{eq}(X_m+X_{l1})+R_2(X_m+X_{l2})]^2} \quad (19)$$

The Pitch regulated fixed speed WTGS model and Semi variable speed WTGS model considered here are the Vestas unit of 1MW rating .The Induction generator circuit parameters are given in Table I.

TABLE I

Induction generator circuit parameters

Type of WTGS Model	Pitch Regulated Fixed Speed	Semi Variable Speed
R_1(pu)	0.005986	0.005671
X_{l1}(pu)	0.08212	0.15250
R_2(pu)	0.01690	0.00462
X_{l2}(pu)	0.107225	0.096618
X_m (pu)	2.5561	2.8985
X_c (pu)	2.5561	2.8985

TABLE II

Induction generator circuit parameters (Model I & II)

1.	Generator short circuit resistance(ohm)	7.81
2.	Short circuit reactance(ohm)	18.76
3.	Open circuit resistance (ohm)	1230
4.	Open circuit reactance(ohm)	500.6
5.	System rms base voltage(kV)	12.66

III. LOAD FLOW ANALYSIS USING Wtgss

As the wind generator is of induction type the bus connected to wind generator can be modeled as PQ bus in load flow algorithm. The load flow algorithm including wind generator is explained through flow chart.

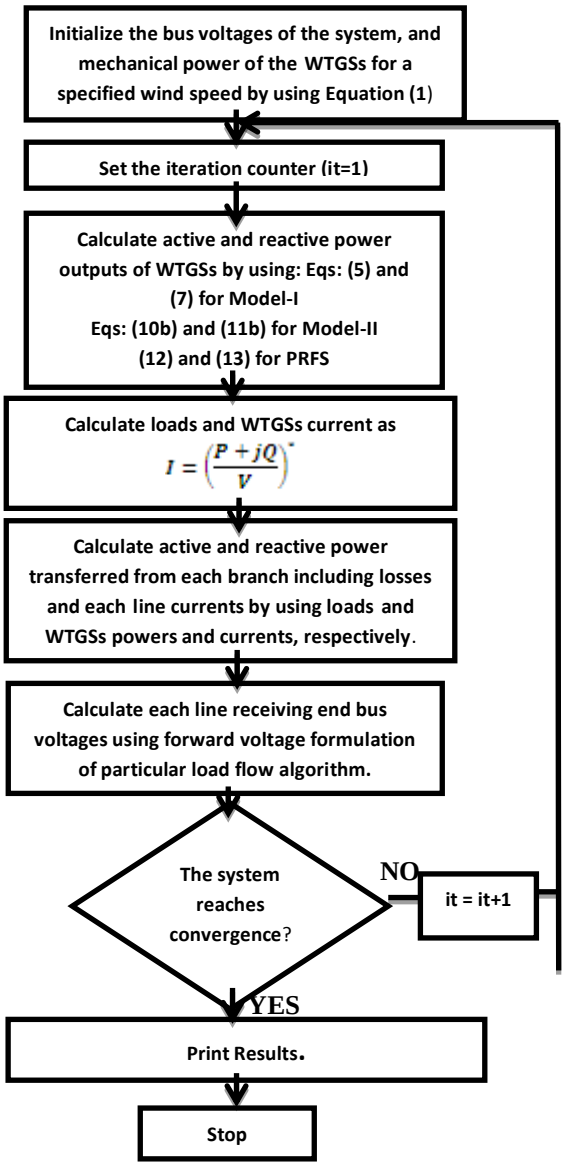


Fig. 2. Flow chart for distribution load flow with WTGS models

IV. Results And Discussion

A 33-bus distribution system is used to evaluate the effect of WTGS on the power losses, voltage profile of radial distribution system, and on the performance of load flow algorithms. Sweep-based algorithms are used and they all have been coded in

Matlab. Transformer parameters are taken from [7]. Tables III & IV represents the load flow results of the systems with and without WTGS modeled by using both developed models, pitch regulated and semi variable speed WTGS models are also depicted. Similarly, in load flow analyses, initial voltage magnitude at all buses is considered to be the same as the source bus (1.0 pu) for test system and a tolerance of 10^{-4} on convergence criteria is used. Mechanical power input is selected as P_m =1.0 MW. From the tables, it is clearly seen that the inclusion of a WTGS significantly, alters the power intake from the main substation, and total power losses of the test systems are also decreased. Moreover, the voltage magnitude of the bus at which the WTGS is connected is increased due to injected active power. The Induction machine circuit parameters and the data is taken from [11] for Model I & Model II, whereas for the remaining models i.e., Pitch regulated fixed speed and Semi variable speed is taken from [7].It is clearly observed that from the Tables III & IV developed models are more advantageous than the Pitch regulated fixed speed WTGS and Semi variable speed WTGS models.

$|V|_{pu}$ is the WTGS bus voltage.

P_m is the mechanical power input(1MW).

$P_s(MW)$ is the WTGS active power generation.

$Q_s(MVAr)$ is the WTGS reactive power demand.

$P_l(MW)$ is the active power loss.

$Q_l(MVAr)$ is the reactive power loss.

Semi Variable Speed(SVS) WTGS.

Pitch Regulated Fixed Speed(PRFS) WTGS.

ABLE III

Load flow result of the 33-bus system for Pitch Regulated and Semi Variable Speed WTGS models

Bus No.	WTGS Model	P_s (MW)	Q_s (MVAr)	P_l (MW)	Q_l (MVAr)	$\|V\|_{pu}$
6	–	–	–	0.369	0.247	0.9331
18	–	–	–			0.8959
25	–	–	–			0.9648
32	–	–	–			0.8788
6	SVS	1	0.624	0.316	0.219	0.9614
18	SVS	1	0.640			0.9243
25	SVS	1	0.616			0.9737
32	SVS	1	0.650			0.9246
6	PRFS	1	0.5665	0.285	0.197	0.9593
18	PRFS	1	0.5640			0.9218
25	PRFS	1	0.572			0.9731
32	PRFS	1	0.564			0.9202

TABLE IV

Load flow result of the 33-bus system with and without WTGS models.(Model I & Model II)

Bus No.	WTGS Model	P_s (MW)	Q_s (MVAr)	P_l (MW)	Q_l (MVAr)	$\|V\|_{pu}$
6	–	–	–	0.369	0.247	0.9331
18	–	–	–			0.8959
25	–	–	–			0.9648
32	–	–	–			0.8788
6	II	0.919	0.923	0.181	0.129	0.9680
18	II	0.919	0.923			0.9295
25	I	0.992	0.017			0.9750
32	I	0.992	0.017			0.9449

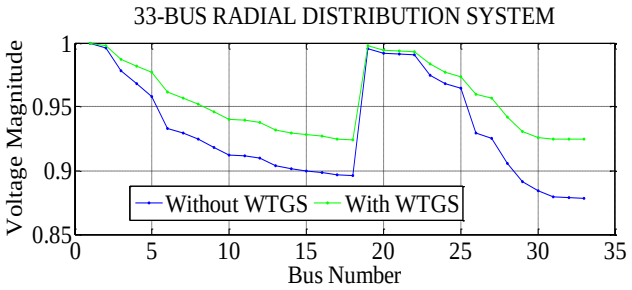


Fig. 3 Voltage profile variation of the 33-bus system for PRFS WTGS model

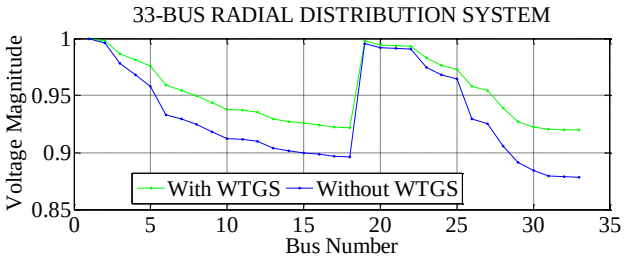


Fig. 4 Voltage profile variation of the 33-bus system for SVS WTGS model

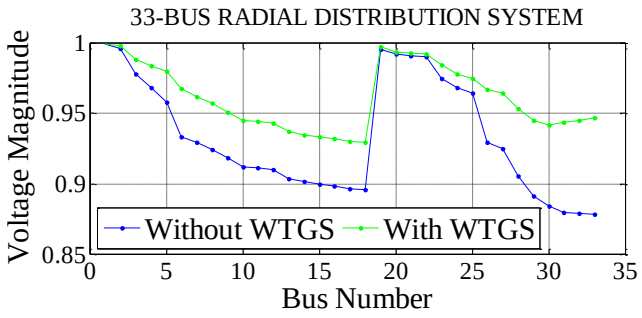


Fig. 5 Voltage profile variation of the 33-bus system (with and without WTGS models)

V. CONCLUSION

In this paper, two new models[11] for wind turbine generating systems are provided. The main advantage of the developed analytical models for WTGS is that they facilitate the computation of real and reactive power outputs for a specified mechanical power input and terminal voltages, in a simple way such that it does not require computation of any system parameters (i.e., rotor slip).Whereas the other proposed models in [7] rotor slip and power curves plays a crucial role.

These developed models [11] can overcome the computational complexity and the computational time problem. The WTGS impact on the system losses and the voltage profile is analyzed. The effectiveness of linking two buses can be studied in terms of decrement in losses and improvement in voltage profile. And the same load flow analysis can be done for other models i.e., Variable speed WTGS and Stall regulated fixed speed WTGS are also useful for the future planning of the system.

References

[1] A. Losi, M. Russo, "Object Oriented Load Flow for Radial and Weakly Meshed Distribution Networks," IEEE Trans. Power Syst., vol. 18, no.4, pp. 1265-1274, Nov. 2003.

[2] W.H. Kersting, "Distribution System Modeling and Analysis," CRC Press ,New York,2002.

[3] D.Shirmohammadi, H.W.Hong, A.Semlyen, and G.X.Luo,"A compensation based power flow for weakly meshed distribution and transmission networks," IEEE Trans. Power Syst., vol.3, pp. 753-762, may 1998.

[4] C. S. Cheng, D. Shirmohammadi, "A three-phase power flow method for real-time distribution system analysis," *IEEE Trans. Power Syst.*, vol10, pp. 671–679, May 1995.

[5] Y. Zhu, K. Tomsovie, "Adaptive Power Flow Method for Distribution Systems with Dispersed Generation," IEEE Trans. Power Del, vol. 17, pp. 822-827, July 2002.

[6] Feijoo A, Cidras J. "Modeling of wind farms in the load flow analysis," IEEE Transaction on Power Systems 2000;15(1):110–6.

[7] Divya KC, Rao PSN. "Models for wind turbine generating systems and their application in load flow studies," Electric Power Systems Researc2006;76:844-56.

[8] Cutsem TV, Vournas C. "Voltage stability of electric power systems," Boston: Kluwer Academic Publishers Springer; 1998.

[9] Hatziargyriou ND, Karakatsanis TS, Papadopoulos M. "Probabilistic load flow in distribution systems containing dispersed wind power generation," IEEE Transaction on Power Systems 1993;8(1):159–65.

[10] Haque MH." Load flow solution of distribution systems with voltage dependent load models," Electric Power Systems Research 1996;36(3):151–6.

[11] Ulas Eminoglu "Modeling and application of wind turbine generating system (WTGS) to distribution systems," Renewable Energy, vol.34,no.11, pp.2474-2483, Nov.2009.

[12] Satyanarayana S, Ramana T, Sivanagaraju S, Rao GK. An efficient load flow solution for radial distribution network including voltage dependent load models. Electric Power Components and Systems 2007;35(5):539-51.

Mobile Networking and Ad hoc networking technologies

[1]Simanta Sarma, [2]Dr. Sarbananda Das
[1](HOD & Asstt. Professor, Department of Computer Science, S.B.M.S College, Sualkuchi, Assam, India)
[2](Rtd. Principal, North Gauhati College, North Guwahati, Assam, India)

Abstract: In this paper we describe mobile network & mobile ad hoc networking technology. Moreover we discuss generation of mobile network and a lean and efficient routing protocol for wireless ad hoc networks. We report on its implementation, on performance comparisons and on a formal validation result. Moreover we discuss Formal Protocol Verification and operating over infrared or Bluetooth. This paper evaluates two model checking tools, SPIN and UPPAAL, using the verification of the Ad hoc Routing protocol as a case study. Insights are reported in terms of identifying important modeling considerations and the types of ad hoc protocol properties that can realistically be verified.

Keywords. Cellular Phone network, mobile ad hoc networks, routing protocols, Wireless networks, ad hoc routing, routing protocol Implementation, formal validation, model checking, Infrared or Bluetooth.

1. INTRODUCTION:

Cellular communications has experienced explosive growth in the past two decades. Today millions of people around the world use cellular phones. In modern area Cellular phones are most important factor in human life. Cellular phones allow a person to make or receive a call from almost anywhere. Likewise, a person is allowed to continue the phone conversation while on the move. Cellular communications is supported by an infrastructure called a cellular network, which integrates cellular phones into the public switched telephone network. Cellular service has seen tremendous acceptance, especially in the last few years, with millions of new subscribers each year and the new subscriber rate growing. Some estimates predict that half a billion cellular phones will be in service by the end of the next decade. AD-HOC networks are typically described as a group of mobile nodes connected by wireless links where every node is both a leaf node and a router.

1.1 EVOLUTION OF MOBILE NETWORKS:

The first systems offering mobile telephone service (car phone) were introduced in the late 1940s in the United States and in the early 1950s in Europe. Those early single cell systems were severely constrained by restricted mobility, low capacity, limited service, and poor speech quality. The equipment was heavy, bulky, expensive, and susceptible to interference. Because of those limitations, less than one million subscribers were registered worldwide by the early 1980s.

1.1.1 GENERATIONS OF MOBILE COMMUNICATION:

1G (First Generation):

The introduction of cellular systems in the late 1970s and early 1980s represented a big leap in mobile communication especially in capacity and mobility. Semiconductor technology and microprocessors made smaller, lighter weight and more sophisticated mobile systems a practical reality for many more users. These 1G cellular systems still transmit only analog voice information. The most prominent 1G system is Advanced Mobile Phone System (AMPS), Nordic Mobile Telephone (NMT), and Total Access Communication System (TACS).

2G (Second Generation):

The development of 2G cellular systems was driven by the need to improve transmission quality, system capacity, and coverage. 2G cellular systems include GSM, Digital AMPS (D-AMPS), code division multiple access (CDMA), IS-95 and Personal Digital Communication (PDC). The most important distinction in 1G and 2G is that 1G networks use analog signals, while 2G networks use digital. Today, multiple 1G and 2G standards are used in worldwide mobile communications. Many standards are used only in one country or region, and most are incompatible. GSM is the most successful family of cellular standards (GSM900, GSM–railway [GSM–R], GSM1800, GSM1900, and GSM400).

3G (Third Generation):

3G systems promise faster communications services, including voice, fax and Internet, anytime and anywhere with seamless global roaming. The first 3G network was deployed in Japan in 2001. 3G technology supports 144 Kbps bandwidth, with high speed movement (e.g. vehicles), 384 Kbps (e.g. on campus) & 2 Mbps for stationary. **"True" 3G**
Third-generation specifications call for even higher speeds 144 kbps in vehicles, 384 Kbps for pedestrians outdoors, and 2.48 Mbps in indoor offices. Some carriers are calling their current deployments 3G. This is contested by others as being the lowest rung of the 3G specification, and hence prefers to use the term 2.5G. As expected, each of the 2.5G technologies has a forward path to the 3rd generation.

- EDGE (Enhanced Data Rates for Global [or GSM] Evolution) is the true 3G offering along the GSM path. It provides data rates three times greater than GSM/GPRS, with speeds in the range 100 - 130 kbps (up to 200 kbps in bursts). EDGE was rolled out across Canada in 2004. Being an extension of GSM/GPRS, EDGE will be widely available internationally, and supported by network operators in many countries, and over 60 network operators in over 40 countries have committed to EDGE for their next generation services.
- There are a couple of forward paths from CDMA2000 offering substantially higher data rates. Neither of the CDMA based carriers (Telus Mobility, Bell Mobility) had announced offerings or pilots at the time or writing.

Next Generation

There are specifications for higher speed services. We will update you when these become closer to reality in Canada.

1.2. BRIEF HISTORY

Many channel allocation algorithms have been proposed during the last thirty years for cellular networks to avoid channel interference and efficiently utilize the limited frequencies. These algorithms can be classified into three types: fixed, flexible, and dynamic. Among them, dynamic channel allocation (DCA) strategies have been the focus of recent research. With DCA strategies, a cell may use any channel that will not cause channel interference. Typically, each channel is associated with a priority; when a cell needs a channel, it picks the available channel which has the highest priority. All the previous algorithms, which are referred to as *centralized* channel allocation algorithms, rely on a *mobile switching centre* (MSC) to accomplish channel allocation.

2. MOBILE AD-HOC NETOWRK

Theoretical mobile ad hoc networking research [CCL03] started some decades ago. But commercial digital radio echnologies appeared in the mid-nineties. Since then, few proposal for enabling ad hoc communications were made. The first technology (IEEE802.11, also referred to as Wi-Fi [ANS99]) is still strongly leading the market, although there is great room for improvement. This section provides an overview and a technical description of the technologies that have been proposed hitherto. A common feature of most wireless networking technologies is that they operate in the unlicensed Industrial Scientific and Medical (ISM) 2.4GHz band. Because of this choice of frequency band, the network can suffer interferences from microwave ovens, cordless telephones, and other appliances using this same band plus, of cours, other networks. In particular, Farrell and Abukharis studied the impact on Bluetooth on IEEE802.11g [ST04]

2.1 Packet radio

Packet radio [GFS78] was used for the earliest versions of mobile ad hoc networks. It was sponsored by DARPA in the 1970s. It allows the transmission of digital data over amateur radio channels. Using special radio equipment, packet radio networks allowing transmissions at 19.2 kbit/s, 56 kbit/s, and even 1.2 Mbit/s have been developed. Since the modems employed vary in the modulation techniques they use, there is no standard for the physical layer of packet radio networks. Packet radio networks use the AX.25 data link layer protocol, derived from the X.25 protocol suite and designed for amateur radio use. AX.25 has most frequently been used to establish direct, point-topoint links between packet radio stations, without any additional network layers. However, in order to provide routing services, several network layer protocols have been developed for use with AX.25. Most prominent among these are NET/ROM, ROSE, and TexNet. In principle, any network layer protocol may be used, including the Internet protocol (IP), which was implemented in the framework of the AMPRNet project.

2.2 IEEE802.11

Wi-Fi is a wireless networking technology based on the IEEE802.11 specifications. The first—and still most used—Wi-Fi standard is referred to as IEEE802.11b in the scientific literature. It was then declined into IEEE802.11a, IEEE802.11g

and IEEE802.11n. IEEE802.11i and IEEE802.11h, which respectively focus on Quality of Service (QoS) and security, are out of the scope of this document. All Wi-Fi technologies operate on the 2.4GHz band, except from IEEE802.11a which operates within the 5GHz band. These technologies use significantly different PHY layers which, from the user point of view, make them differ in term of the bandwidth (i.e. the data rate) that they provide. Typically, Wi-Fi enabled devices have coverage distances ranging from 50 to more than 100 meters. In practice, this coverage distance depends greatly on the nature of the antenna and on the environment in which the devices evolve.

2.2.1 IEEE802.11a

IEEE802.11a uses Orthogonal Frequency Division Multiplexing (OFDM). It is the only wireless radio technology that works in the 5GHz band. The main idea behind OFDM is that since low-rate modulations (i.e modulations with relatively long symbols compared to the channel time characteristics) are less sensitive to multipath, it should be better to send a number of low rate streams in parallel than sending one high rate waveform. OFDM then works by dividing one high-speed signal carrier into several lower-speed subcarriers, which are transmitted in parallel. High-speed carriers, which are 20MHz wide, are divided into 52 sub channels, each approximately 300KHz wide. OFDM uses 48 of these sub channels for transporting data, while the four others are used for error correction. OFDM delivers higher data rates and a high degree of multipath reflection reconstruction, thanks to its encoding scheme and error correction.

2.2.2 IEEE802.11b

IEEE 802.11b uses Direct Sequence Spread Spectrum (DSSS) as the physical layer technique for the standard. DSSS uses a complex technique which consists in multiplying the data being transmitted by a *noise* signal. This noise signal is a pseudo-random sequence of 1 and −1 values, at a frequency much higher than the original signal. The resulting signal wave looks much like white noise. This white noise can be filtered at the receiving end in order to recover the original data. This filtering happens by again multiplying the same pseudo-random sequence by the received signal (because $1 \times 1 = 1$, and $-1 \times -1 = 1$). This process, known as "de-spreading", mathematically constitutes a correlation of the transmitted pseudo-random sequence with the receiver's assumed sequence. For allowing de-spreading to work correctly, the transmit and received sequences must *synchronized*. So far, IEEE 802.11b is the implementation of the IEEE 802.11 standard that has been most heavily studied in the framework of mobile ad hoc networks.

2.2.3 IEEE802.11g

IEEE802.11g, just like IEEE802.11a, uses orthogonal frequency-division multiplexing (OFDM), it then boasts similar bandwidths. OFDM is described in Section 2.2.1. But unlike IEEE802.11a, IEEE802.11g works in the 2.4 GHz band. Since the draft 802.11g standard combines fundamental features from both 802.11a and 802.11b, it leads to the development of devices that can inter-operate with technologies based on both of the previous versions of the specification.

2.3 Bluetooth

Bluetooth is essentially the same kind of microwave radio technology that has given us wireless door chimes and automatic garage door openers. It was initially restricted to an operating distance of just 10 meters and a speed of approximately 1 Mbit/s. When Bluetooth devices come within range of each other, they establish contact and form a temporary network called a Personal Area Network (PAN). In the Bluetooth terminology, this is also known as a Pico net. A multi-hop ad hoc network formed by the interaction of Bluetooth devices is called a Scatter net. When using Bluetooth, the devices must establish a network session before being able to transmit any data. Bluetooth uses the Frequency-Hopping Spread Spectrum (FHSS) technique. Unlike IEEE802.11 which establishes a communication link on a certain frequency (a channel), FHSS breaks the data down into small packets and transfers it on a wide range of frequencies across the available frequency band. Bluetooth transceivers jump among 79 hop frequencies in the 2.4 GHz band at the rate of 1,600 frequency hops per second. 10 different types of hopping sequences are defined, 5 of the 79 MHz range/79 hop system and 5 for the 23 MHz range/23 hop system.

This technique trades off bandwidth, in order to be robust and secure. More precisely, Spread Spectrum communication techniques have been used for many years by the military because of their security capabilities.

2.4 Hiperlan

The HiperLAN2 standard is very close to 802.11a/g in terms of the physical layers it uses—both use OFDM technology—but is very different at the MAC level and in the way the data packets are formed and devices are addressed. On a technical level, whereas 802.11a/g can be viewed as true wireless Ethernet, HiperLAN2 is more similar to wireless

Asynchronous Transfer Mode (ATM). It operates by sharing the 20MHz channels in the 5GHz spectrum in time, using Time Division Multiple Access (TDMA) to provide QoS through ATM-like mechanisms. It supports two basic modes of operation: centralized mode and direct mode. The centralized mode is used in the cellular networking topology where each radio cell is controlled by an access point covering a certain geographical area.

2.5 ZigBee

ZigBee-enabled devices conform to the IEEE 802.15.4-2003 standard. This standard specifies its lower protocol layers, the physical layer (PHY), and the medium access control (MAC). It targets Low-Rate Wireless Personal Area Network (WPAN). ZigBee-style networks research began in 1998. Zigbee was intended to operate in contexts in which both Wi-Fi and Bluetooth are not suitable. Zigbee operates in the unlicensed 2.4 GHz, 915 MHz and 868 MHz ISM bands. It uses direct-sequence spread spectrum (DSSS) coding. This makes the data rate to reach 250 kbit/s per channel in the 2.4 GHz band, 40 kbit/s per channel in the 915 MHz band, and 20 kbit/s in the 868 MHz band. The maximum output power of ZigBee antennas being generally 1 mW, the transmission range of ZigBee nodes is between 10 and 75 meters. Observations have shown that the transmission range is heavily dependent on the environment.

2.6 Broadband wireless networking

WiMAX (IEEE 802.16) stands for Worldwide Interoperability for Microwave Access. IEEE 802.16 boasts data rates up to 70 Mbit/s over a distance of 50 km. However practical limits from real world tests seem to be between 500 kbit/s and 2 Mbit/s at a distance of around 5-8kms. WiBro is a wireless broadband internet technology being developed by the Korean telecoms industry. It has been announced that WiBro base stations will offer an aggregate data throughput of 30 to 50 Mbit/s and cover a radius of up to 5 km. The technology will also offer Quality of Service.

HIPERMAN [HPF03, HPF04], which stands for High Performance Radio Metropolitan Area Network, is a European alternative to WiMAX. The standards were created by the European Telecommunications Standards Institute (ETSI). It provides a wireless network communication in the 2-11 GHz bands. The adequation of these technologies to ad hoc networking is discussable, since they would permit to establish ad hoc networking at a level at which technologies for infrastructure networks (like GSM or UMTS) are available.

3. Protocol Validation

3.1 Survey of Methods

Computer networking protocol validation is commonly done using a combination of simulation and testing. These are both valid approaches that to some extent complement each other. Simulation offers the possibility to run a large batch of tests under identical circumstances whereby some parameter can be varied and the effect studied. A very common assisting tool, or framework, is the network simulator - ns-2 [26]. Live testing is often applied to some extent during protocol development. An important application for the method is when checking interoperability between different implementations. Live testing poses the difficulty of conducting several comparable tests, but if done in a structured way it may very well expose errors or problems not visible in simulations. The gray zone problem, reported by Lundgren et al. [34] is one example of such a discrepancy. In Paper C initial results from a structured live testing study are presented. The tool we use is called the APE testbed [38]. A third alternative is to use formal verification in order to be sure to cover all situations possible in a system model. Testing and simulation are not exhaustive methods and cannot guarantee that there are no undiscovered subtle errors or design flaws in a protocol. The objective of formal verification is to improve on reliability by reasoning about systems using mathematical logic. A formal system model can thereby be checked to fully comply with a given set of requirements. There have been comparatively few efforts at formal verification of ad hoc routing protocols. The reason for this is twofold. First, there is the presumption that the methods are difficult to use which is to some extent true since there really is a threshold to cross before becoming proficient. The deductive methods usually require more experience before it is possible to carry out a proof for a non trivial system. Even then, it is often a very time consuming process. In the case of deductive methods they have a potential to be very powerful and can be used to construct proofs for large or even infinite state systems. However, the proof may be notoriously difficult to find or it may not even exist because the problem is not well formulated. Algorithmic verification methods, commonly known as model checking [9], have been more successful in terms of industrial deployment because of their easier usage. These methods have another problem that surfaces for systems composed from a set of different components that can interact in a non deterministic manner. Many possible interleaving of execution are thereby possible, leading to exponential growth of the searched state space; the state explosion problem [47]. These new techniques thus have the potential for verifying infinite state systems automatically by abstract interpretation [15] followed by, for example, symbolic model checking [36]. There is ongoing work on many fronts in order to lower the threshold of use as well as on coping with the state explosion

problem. Here, we concentrate on some of the more user friendly tools, namely automatic model checkers. Our hope is to advocate the use of formal verification by the average protocol designer.

3.2 Formal Protocol Verification

3.2.1 System Description Languages

In order to verify a particular protocol it first has to be described in a structured and unambiguous way. For this, there are two main choices. Either, one can write an implementation in a regular programming language such as C and thereafter verify the code directly. This approach has been used by Engler and Musuvathi [18] to find errors in different AODV implementations. It is most often not used for exhaustive verification but rather as a method of finding program bugs, even though Engler and Musuvathi were also able to identify a routing loop error in the AODV specification. The second approach to describing the protocol is to use a formal description language. This can either be a subset of first order predicate logic or some more high level formalism such as PROMELA (PROcess Meta LAnguage) used in the SPIN [23] tool. In reality, these languages are just representations of transition systems. It is essential that the formal description matches that of the real system implementation, but normally some parts have to be abstracted away from in order to make the problem feasible for verification. In the case of deductive verification the proof may otherwise be too difficult to construct and in the case of model checking the state space can easily blow up. When abstracting, one has to make sure that the system model retains the same behavior as the implementation for the properties of interest.

Table 1. SPIN verification results

Scenario	States generated	Transitions	All states searched	Memory used [Mb]	Time used
(a)	5715	12105	Yes	4.242 (6.188)	0.20 (0.20) s
(b)	269886	731118	Yes	33.05 (124.7)	12.33 (10.48) s
(c)	53614	128831	Yes	8.836 (30.12)	2.19 (1.92) s
(d)	4.58e+07 (8.15e+06)	1.33e+08 (2.21e+07)	No	4083 (4083)	5 h:57 min (8 min:56 s)
(e)	1.41e+06	4.59e+06	Yes	170.4 (806.6)	1:36 (1:26) min:s
(f)	3.40e+07 (7.27e+06)	1.22e+08 (2.50e+07)	No	4083 (4083)	4 h:2 min (9 min:43 s)

3.2.2 Requirement Properties and Specification Languages

Requirements on the system are commonly expressed in a temporal logic such as LTL (Linear Temporal Logic) [43] or CTL (Computation Tree Logic) [10]. Requirement properties can be categorized as either liveness or safety properties [29]. Characteristic for a safety property is that a violation is detectable using a finite system run. It can informally be described using the sentence "something bad will never happen" provided that the property holds in all reachable system states. In contrast, a liveness property corresponds to the sentence "something good will eventually happens". In order to produce a counter example for a liveness property it is sometimes necessary to study infinite system runs. An example of a liveness property is the one we used in Paper B and Paper D, expressed somewhat simplified: under the appropriate premises, a given routing protocol will eventually find a route to a particular destination.

3.2.3 Applying the Method

Model Checking

There are two main advantages of model checking in comparison to deductive methods. The first one is that once the system model has been constructed and the verification properties devised, the process is completely automatic and outputs a "yes" or "no" answer. The other advantage is the possibility to generate error traces in case a property is not fulfilled by the system. This makes it possible for the user to modify the model accordingly. The main disadvantage of model checking is its limitation to finite state systems. It can, however, be used in hybrid infinite state verification approaches where model checking is, for example, a component in a CEGAR (Counter-Example Guided Abstraction Refinement) loop [12]. Furthermore, model checking of symbolically represented systems can be regarded as infinite state since the original system may contain an unlimited element (such as continuous time). Using model checking, one can check safety as well as liveness properties. Model checking algorithms work by exploring the state space whereby the search stops at the first violation or when the complete execution tree has been examined. Methods can be divided into

explicit state and symbolic model checking depending on if the individual states or groups (sets) of states are used to represent the state space.

Deductive Verification

In deductive verification the goal is to prove that a conclusion, the property to be verified, can be drawn from a given set of premises, the system description. This was previously a tedious manual process which has been speeded up with the emergence of semi-automatic tools, so called theorem provers. One advantage of this method is that it can be used to prove properties of infinite state systems, for example a protocol running in a network with an Unbounded number of nodes. An invariant is an assertion that is true in all states of a system. A safety property, expressed as an invariant, can be proven using mathematical induction. First it needs to be proven that the initial system configuration implies the assertion. In the inductive step it is then checked whether all state transitions preserve the property, that is, if the assertion holds before the transition it will also hold after it. Hence, the verification does not require an explicit state space search. This avoids the state explosion problem at the cost of a more cumbersome proof process. The manual method was used by Ogier [40] to make a proof of correctness for the TBRPF [39] protocol. For the discovery module he further presents a proof that the neighbor information exchanged is sufficient for the functionality of the protocol.

3.2.4 The State Explosion Problem and Remedies

The state explosion problem in model checking refers to the situation in which the state space storage overhead grows exponentially with the size of the model. This problem occurs because of the large number of possible interleaving between processes in a reactive concurrent system. Verification may thereby fail simply because the available amount of computer memory is limited. There have been a number of suggestions for coping with the state explosion, that is, to make verification feasible for realistically sized systems. We list the major remedies below following the description by Clarke et al. [9].Symbolic representation. Symbolic representation refers to the use of compact data structures for representing state space. For example, by encoding the transition relations of a Kripke structure as a Binary Decision Diagram (BDD) it is possible to save storage by exploiting the often inherent regularity of a hardware or software system. Constraint system representation of continuous parameters such as clock ranges, which is done in UPPAAL, is another example of a symbolic representation. In that case it would not even be possible to store all time points explicitly regardless of the amount of available memory.Partial order reduction. Partial order reduction [24] is an optimization, for example implemented in the SPIN tool. If a group of concurrently executing processes do not exchange any data throughout their lifetime, then it does not make a difference for the end result if they are run one after the other or in parallel. This makes verification simpler since the processes can be verified in isolation. However, once processes cooperate, for example by message passing, which is certainly the case for protocol implementations, then the possible interleaving of operation have to be taken into account when verifying the system. Partial order reduction is a way of disregarding process interleavings that produce the same global state as some other interleaving. Note that the verification property also needs to be taken into account since it might introduce additional data dependencies between processes. Keeping as much as possible local to each modeled process can thus promote partial order reduction.

Compositional reasoning. This technique [2] involves decomposing the system into components which are verified separately and in isolation from the rest. Global properties can then be inferred from the composition of the components. If there are mutual dependencies between components one can still verify each component separately under the assumption that the other components work as expected; assume-guarantee reasoning. There are both manual and automatic approaches available for compositional reasoning.

4. Related Work

Routing below the IP layer for ad hoc networks was independently adapted by [1] using label switching which is equivalent to the selectors. A similar project is [2] where the authors also aim at putting L2.5 routing logic inside the (wireless) network interface card. For AODV, formal validations have been carried out by the Verinet group [19]. Using a theorem prove and a SPIN model of AODV in a 2 node setup (with an AODV router environment), it is in fact a loop free routing protocol. The Verinet group [23] have carried out formal validation of AODV [13] and identified a flaw that could lead to loop formation. This was done using the HOL [24] theorem prover and a SPIN model of AODV in a two node setup (with an AODV router environment). They have also suggested a modification and verified that, after this, the protocol was loop free. Their approach verified the general case, but the methodology involves substantial user interaction.

5. Conclusions

This work is to our knowledge the first to study a range of topologies in order to determine where the limit actually is when performing model checking on an ad hoc routing protocol. We have introduced the ad hoc routing protocol which targets the common-case of network clouds with 10-15 nodes and a diameter of up to three hops. We believe that such settings will be the most popular ones where ad hoc networks can and will be put into operation. More specially, in larger settings and for IEEE 802.11 there are such severe degradations occurring under any ad hoc routing scheme that we do not consider this to be a relevant use case that a routing protocol should try to address. When verifying both the data and control aspects of the protocol using SPIN and when verifying the timing properties using UPPAAL the size of network, i.e. the number of nodes involved, as well as the nature of the topological scenarios is limited due to state space storage overhead. Even if parallel model checking approaches were used, our conclusion is that it is at this point not feasible to provide a proof for topologies of any significant size by modeling the protocol directly. On the other hand, our study enables us not only to analyze the modeling considerations that have to be imposed, but also provides us with a solid starting point for the further work we intend to pursue in the direction of infinite-state verification of ad hoc routing protocols.

6. REFERENCES

1. Somnath Sinha Maha Patra, "Improved Genetic algorithm for channel allocation scheme with channel borrowing in Mobile Computing", IEEE Transactions on Mobile Computing, vol-5, No.7, July 2006.
2. Arup Acharya, Archan Misra and Sorav Bansal. .A Label-switching Packet Forwarding architecture for Multi-hop Wireless LANs.
3. Proc WoWMoM'02, Sep 2002, Atlanta, USA Lundgren, H.: Implementation and Real-world Evaluation of Routing Protocols for Wireless Ad hoc Networks. Licentiate thesis, Uppsala University (2002)
4. IETF MANET Working Group: MANET charter. http://www.ietf.org/html. charters/manet-charter.html (2004
5. Perkins, C., Belding-Royer, E., Das, S.: Request for Comments: Ad hoc on-demand distance vector (AODV) routing. http://www.ietf.org/rfc/rfc3561.txt (2003).
6. Johnson, D.B., Maltz, D.A., Hu, Y.C.: Internet draft: The dynamic source routing protocol for mobile ad hoc networks (DSR). http://www.ietf.org/internet-drafts/draft-ietf-manet-dsr-09.txt (2003).
7. Xiong, C., Murata, T., Tsai, J.: Modelling and simulation of routing protocol for mobile ad hoc networks using coloured Petri nets. In: Proc. Workshop on Formal Methods Applied to Defence Systems in Formal Methods in Software Engineering and Defence Systems. (2002)
8. Tian-Tsair SutS, Po-Chiun Huangt, and Chung-Ju Changt,"A channel borrowing protection scheme for handoffs in a cellular based pcs system" , 0-7803-2955-4/95/ IEEE
9. Holzmann, G.: The Spin Model Checker, Primer and Reference Manual. Addison-Wesley, Reading, Massachusetts (2003)
10. Barnat, J., Brim, L., Stribrna, J.: Distributed LTL model-checking in SPIN. Technical Report, Masaryk University (2000)

GSM Based Anti-theft Security System Using AT&T Command.

Visa M. Ibrahim[1]. Asogwa A. Victor[2]. S. Y. Musa[3].
[1,2,3]Department of Electrical and Electronics Engineering, School Of Engineering and Engineering Technology, Modibbo Adamma University of Technology Yola, Adamawa State. Nigeria

Abstract
Antitheft security system security system using AT&T COMMAND utilizes an embedded system design with GSM to monitor and safeguard a car. It secures the car against theft. Upon activation, it automatically demobilizes the car by disconnecting the ignition key supply from the car battery. This now makes it impossible for anybody so starts the car, let alone moving with it. In an attempt of theft through the car doors or boot, the system sends text message to the car owner and at the same time starts up an alarm. This design popped out due to the increasing rate at which packed cars are stolen especially in our country, but with this design this packed car is being monitored irrespective of where it is packed, provided there is GSM network coverage. From the research conducted, it was found out that majority of the existing car security system uses only alarm, and doesn't send text message to the car owner let alone of demobilizing the car. But with the use of GSM network, the owner is guaranteed that the car will send text message to his phone, and at the same time, have people around alerted of what is happening. Added to this is that the car will not be move because it has been demobilized. The use of AT&T COMMAND made the design use very few electronics component, look very small and compact that it can be hardly seen when mounted in the car [2].

Keywords- Communication, AT&T Command, Microcontroller, GSM, Networks, Text message, Feedback.

1. Introduction

In a situation where there is high level of t heft, there is need for better security system. It is much safer to have a system that monitors and communicates to the device owner without putting human life to risk in the name of "Watchman". This tends to utilize the availability of GSM network, mobile phone and electronics circuit to achieve an automated system which is programmed to work as a thinking device to accomplish this purpose. By simply dialing the phone number of the mobile phone attached to the circuit and sending a code after the phone has automatically been answered, puts the system to either "active or inactive" state, and on any attempt of theft the system sends a text message to the device owner, demobilizes the system (car) and then starts up an alarm. With this, the car is always protected. The total absence of sufficient security personnel in a packed car is a great discomfort to car owners. This insecurity has paved way to increasing rate of stealing packed cars – even with security.In order to enhance an improved and life risk free security system, the purpose of this study is to aid a better security system of cars with the use of GSM. This system monitors one's car against theft, and has a text message sent to the car owner, telling him that his car is being tampered, and at which part of the car (either doors or boot) is being tampered. The system will also demobilize the car (that is stopping the car from moving), set up an alarm for the people around to notice what is happening.

1.1 System Description

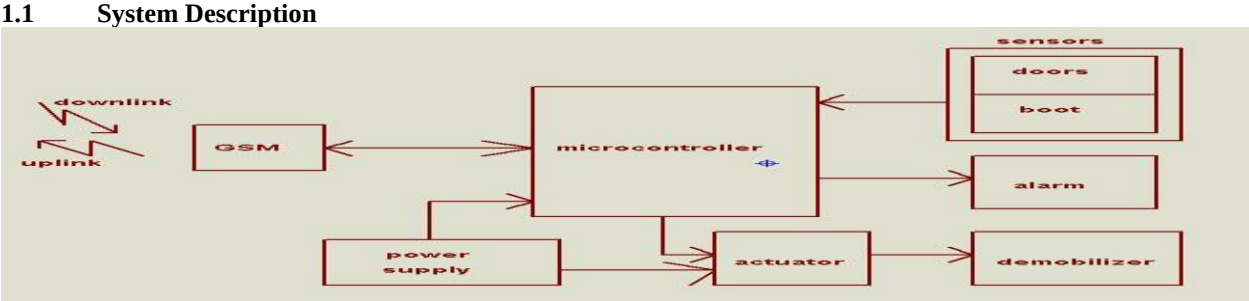


Figure 1: Functional Block Diagram of the System.

The diagram in figure 1. Describe the overall system. The system comprises of a GSM which serves as an intermediate between the outside world (car owner) and the entire system. It is responsible for both receiving of calls for system activation and deactivation, and sending of text message to the car owner. Its output its output fed into the microcontroller. The micro controller which is the brain and control circuit behind the entire design is controlled with a written program stored into its Read Only Memory (ROM). The controller with the direction of the written program co-ordinate the system's operation as follows:

i. If the system is active, it monitors both the car doors and boot to check if anyone is opened.
ii. Demobilizes the car by disconnecting the ignition line from the battery supply via the actuator
iii. Sends text message to the car owner signifying where the attempt is made from, and finally
iv. Starts up an alarm for alert purpose.
v. But, if the system is inactive, the microcontroller disables all outputs and sensor inputs.

In this paper, the design utilized locally assembled microcontroller and few other basic electronics components to achieve both the control and the controlled. Evidently, microcontrollers have been used to perform such simple security task in which human performance might degrade over time. The reliability of this design makes it durable and capable for offering guarantee security at all time. The power supply provides the voltage and current required for effective performance of the system. This supply is tapped from the car battery and then regulated before being fed to the system.

2. Design Frame Work

2.1 System Design

The design process is divided into two: Hardware and Software designs.

2.1.1 Hardware Design

2.1.1.1 The microcontroller.

he microcontroller is actually a computer on chips. It is a member of the 8051 family of microcontrollers and is 8 bit wide. It has 256 bytes of RAM, 8K of on chip ROM, three timers, 32 inputs and output pins making a total of four ports, each 8 bit wide. 1 serial pin and 8 interrupt sources all on a single chip. In all is a forty pin IC. The AT89C51 microcontroller is an 8 bit processor meaning that the central processing unit CPU can work only on 8 bit. Data larger than 8 bit must be broken down into 8 bit pieces to be processed by the CPU, it is a programmable chip: the availability of software development tools such as compilers, assemblers and debuggers so that it behave exactly as the designer program it. This makes it more effective and reliable in carrying out task especially in a design that involves logic decision and control. The alarm was implemented using a 12V/15W siren.The BC337 bipolar junction transistor is a switching device which switches on the alarming unit [1]. And FET (IRF3205) is a power mosfet [1]. It functions as an actuator to the demobilization unit. Relay is the demobilizer in the system. Voltage regulation is achieved by the use of LM7805 [8]. The figure 2 below gives clearer view of these hardwares.

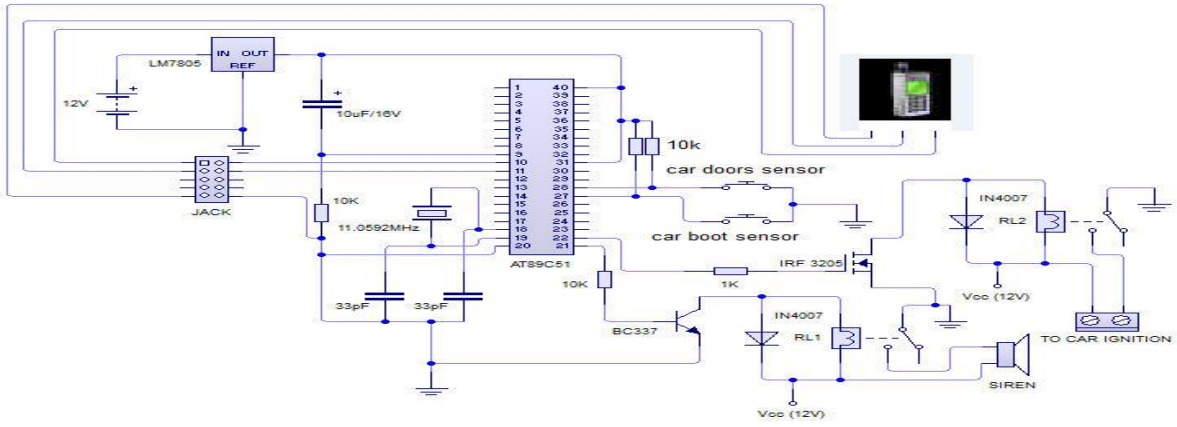


Figure 2: Complete circuit diagram.

2.1.1.2 Interfacing the microcontroller to coordinate the entire system.

Port two bit six (P2.6) and Port two bit seven (P2.7) of the AT89C51 are respectively used as input for doors sensor and boot sensor. Port two bit zero (P2.0) and Port two bit one (P2.1) are used for controlling the alarming unit and the actuator for demobilizing the car.

2.2 Software Design

The program was written in assembly language using the 8051 instruction set. It utilizes AT&T Command codes for its operation and performance. The AT&T command codes are special codes made for easy control and accessibility of compatible phone through programs written to microcontroller (AT89C51) in serial communication mode [4]. The simplicity of the hardware is as a result of the serial communication compatibility of the AT89C51 P3.0 (RXD) and P3.1 (TXD). The AT&T Command is indeed the enhancement of the microcontroller based anti theft security system with GSM (text message) feedback. The flowchart for the program is shown in Figure 3.

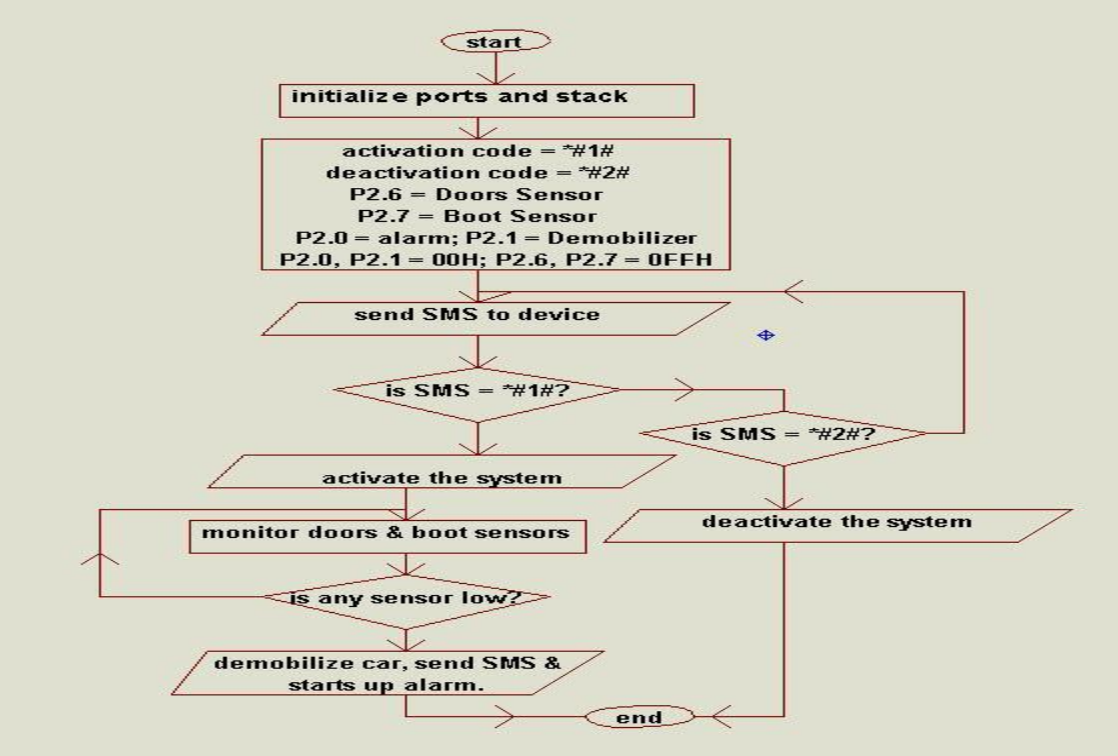


Figure 3: Controlled Program Flowchart.

3.0 PERFORMANCE EVALUATION AND TESTING

After the construction of the entire system, the program was written and burned into the microcontroller chip. On interfacing it to the car, it was tested and found ok as it performed the objective of the design. Satisfaction was gotten from the moment the system was activated and the car opened, the system automatically disconnected the ignition, sent a text message to a programmed recipient and then started up an alarm. The mission of the design was accomplished.

The entire system has serial inputs through P3.0 (RXD) and P3.1 (TXD). The outputs are the text message controller, alarming system and demobilization unit. The system is battery powered and was tapped from car battery and for this reason it made it easier for dc use. With these a system that sends text message to car owner is designed and implemented.

4.0

Conclusion

A cost effective, simple, miniature and robust GSM based anti-theft security system using AT&T command has been successfully designed and constructed. From the input to the output of each stage was a success. The major input is like an interrupt and upon reception by microcontroller it goes through the iteration of demobilizing the car, sending text message and starting up an alarm. The system exhibits a satisfactory performance.

BIBLIOGRAPHY

[1]. Paul Horowtz and Winfield Hill (2002) The Art of Electronics, 2nd Edition, Cambridge University press.
[2]. TIA/EIA-592 (1998) A Facsimile Digital Interfaces-Asynchronous Facsimile DCE Control Standard, Service Class 2.
[3]. Ronald J. (Monroe Community College) and Neal S.Widmer (Purdue University) (2005) Digital Systems Principles and Application, 8th Edition, Prentice- Hall India.
[4]. AT Command Set (2005) For Nokia GSM/SAGEM and WCDMA Product Version 1.2.
[5]. Robert R. (1998) Modern Control System, 8th Edition, New York.
[6]. Muhammad Ali Mazidi and Janice Gillispie Mazadi (www.asadali.tk)" the 8051 microcontroller and embedded systems" ; Department of Computer Science and Information Engineering, National Cheng Kung University of Taiwan.
[7]. Zarlink Semiconductor (1993) Application Note on MT8870 Integrated DTMF Receiver.
[8]. Forrest M. MIMS III Getting Started in Electronics.
[9]. Muhammed Ali Mazidi (2nd edition) The 8051 Microcontroller and Embedded System Using Assembly and C
[10]. Rashid R. (1998) Power Electronics, Circuit Device and Application, 2th Edition, Prentice hall.
[11]. Donal Heffernan (2002) 8051 Tutorial, University of Limerick.
[12]. Theraja B. L. and Theraja A. K. (2003) A Text of Electrical Technology, 23th Edition, Schand and Company Ltd. New Delhi.
[13]. Mazidi M. (1987) 8051 Assembly Language Programming and Interfacing, Prentice Hall, USA.
[14]. http//www.gogle/Ebedtroniccs.
[15]. Hall F. and Lister P.F (1980), Microprocessor Fundamentals, Pitman, London.
[16]. Horowitz P. and Hill W. (1989), The Art of Electronics, 2nd Edition, Cambridge University press.

SURVEY OF FORMAT PRESERVING ENCRYPTION

S.Vidhya[1] , K.Chitra[2]

[1] Ph.D Scholar, Department of Computer Science, SCSVMV University, Kanchipuram, India.
[2]Assistant Professor, Govt.Arts College, Melur, Madurai, India

Abstract

Cryptography is a technique used to transmit data in a secured way through the internet. Encryption is the process of converting information from its original form (called plaintext) into an encoded, unreadable form (called cipher text). Format preserving encryption (FPE) refers to a set of techniques for encrypting data such that the cipher text has the same format as the plaintext. A format-preserving encryption scheme is applicable for many real-life applications. FPE is a good encryption scheme that allows for encryption with minimal modifications to the original plain text. I examine the FPE model constructed by Black and Rogaway.

Keywords – Analysis of FPE, Data type preserving encryption, Format preserving encryption, FPE, Survey of FPE

I.Introduction

"Security, like correctness, is not an add-on feature."--Andrew S. TanenbaumThe above quote (taken from Cryptography Quotation page) is trying to say that security is not something extra, but it is something essential. Encryption and decryption are the methods used to transmit messages and other sensitive documents and information. During the last few years, format-preserving encryption (FPE) has developed as a useful tool in applied cryptography. The goal is this: under the control of a symmetric key K, deterministically encrypt a plaintext X into a cipher text Y that has the same format as X .

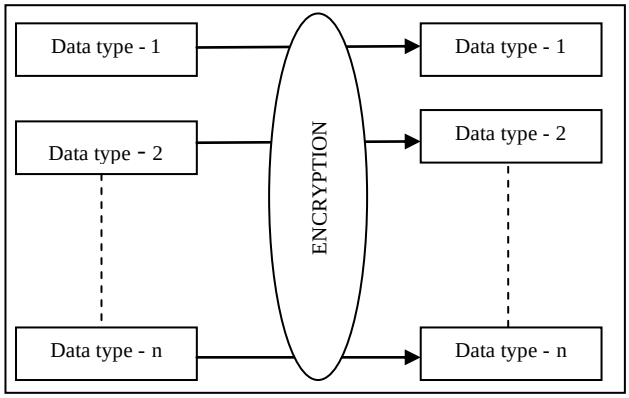


Fig.1 Format Preserving Encryption

The another names for FPE are Data type Preserving Encryption (DPE) and Feistel Finite Set Encryption Mode (FFSEM). There are many names to indicate theFPEtechnique. The main aim of all the techniques is to get back the same size, and data type as the original plain text is being encrypted. Transmitting sensitive data securely over the multi-system environments with minimum changes [1].

II.Fpe Importance

During Encryption and Decryption there is a need for changing the database to store the encrypted text. The main disadvantage in normal encryption method is the cost of modifying the existing databases and applications to process an encrypted information. These costs are related with two important criteria. First, sensitive information like credit card numbers is frequently used as primary key in databases, so changes of this field by encrypting data may require significant schema changes. Second, applications are related to specific data format; encryption will require a format change. In format preserving encryption there is no need to alter the database. A database field which contains a sixteen digits credit cannot store the DES generated cipher text. A Front end program cannot read it [2]. A Graphical User Interface would not display it. The normal encryption should provides lot of facilities for changes in data format throughout an application program and physical database schema. The main aim of FPE is to encrypt the data without the need to modify all of the systems that use that data; such as database field, queries, and all the application program.

A. Fpe Uses

1. Encrypt all types of data including numeric and Alpha numeric
2. Reduce changes to database or application schemas . The data is suitable for the existing data base field.
3. It supports referential integrity
4. Enables encryption of primary and foreign keys
5. It also supports reversible and non-reversible data masking

Iii. Fpe Mechanism

FPE security mechanism needs to be strong with the following limitations:

1. The attackers familiar with format and type of data in the database.

2. Data cannot be extended. If the FPE algorithm encrypts an N-digit number, the output also an N-digit number.[3]

The FPE algorithm should be satisfied the above mentioned conditions.

Iv. Existing Fpe Techniques

Cryptographers John Black and Phillip Rogaway proposed three techniques for FPE [4]. All the techniques are based on secured block ciphers (AES). This section provides analysis of three FPE techniques.

1.Prefix cipher
2.Cycle Walking
3. Feistel Network

A. Prefix Cipher

In prefix cipher FPE algorithm each integer in a plaintext is assigned by pseudo random weights. The weights are calculated by using AES or 3DES block cipher to each integer.

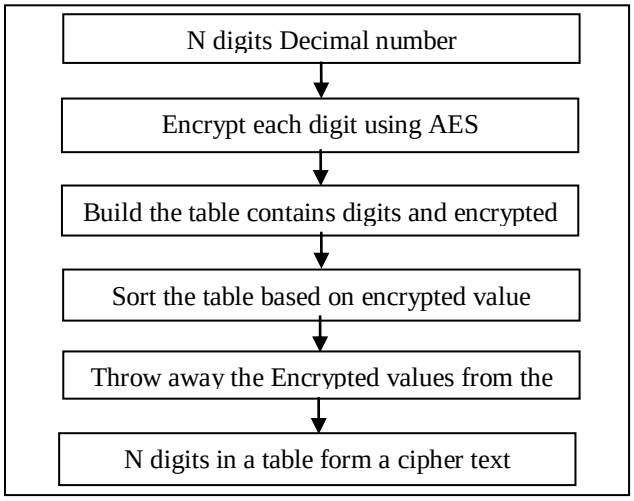


Fig.2 Prefix Cipher

This method is applicable only for small length of plaintexts. For larger plaintext, the entries in the the lookup table and the required number of encryptions to create the table is too big. To build the table, AES or 3DES algorithm is used to encrypt the digits in the plaintext. The table contains input digitand the encrypted value, then sort by the encrypted value. The tiny domain X = {0, 1, 2, 3, 4} having just five possible plaintexts. = {0, 1, 2, 3, 4} having just five possible plaintexts. Under the control of a key K , say having 128 bits, compute Y(0)=AESK (0), Y(1)=AESK (1),Y(2)=AESK (2), Y(3)=AESK (3), Y(4)=AESK (4), and Y(5)=AESK (5). Use the relative ordering of Y(0),Y(1),Y(2),Y(3),Y(4) and Y(5) to determine the desired permutation [5].Suppose we want to encrypt set of 8 digits. Applying AES algorithm for each digit to build the table which contains digits and encrypted value. Sort the table based on encrypted value. The 8 digit number is 34567812.

TABLE I. ENCRYPTED VALUE BEFORE SORTING

Digit	AES encryption of digit
3	49d68753999ba68ce3897a686081b09d
	19ad2b2e346ac238505d365e9cb7fc56
5	9b82998964728141405e23dd9f1dd01b
6	d45efc5268a9afeac1d229e7a1421662
7	b9322f19c62b38e9bed82bd3e67b1319
8	a524c76df94fdd98f7d6550dd0b94a93
1	7346139595c0b41e497bbde365f42d0a
2	3063b6df0a2cdbb0851251d2c669d1bf

TABLE 2. ENCRYPTED VALUE AFTER SORTING

Digit	AES encryption of digit
4	19ad2b2e346ac238505d365e9cb7fc56
2	3063b6df0a2cdbb0851251d2c669d1bf
3	49d68753999ba68ce3897a686081b09d
1	7346139595c0b41e497bbde365f42d0a
5	9b82998964728141405e23dd9f1dd01b
8	a524c76df94fdd98f7d6550dd0b94a93
7	b9322f19c62b38e9bed82bd3e67b1319
6	d45efc5268a9afeac1d229e7a1421662

In this example encryption of 3 is 4, encryption of 4 is 2 and encryption of 2 is 6.

Prefix Cipher Performance

It needs N AES calls to encrypt N digits number. For example to encrypt 16 digits credit card number it needs 16 AES calls. It is not a efficient scheme. The this method is interesting for small values of |M| but is completely unpractical otherwise since$2^{|M|}$time and memory are required in order to start using the cipher[5]. The Prefix method having a very slow key generation and a very fast encryption . The needs more amount of time to build the table, and the memory to hold the table.

Prefix Cipher Optimization

Encrypt the elements using AES and storing only the first 32 bits or 64 bits in the table instead of storing entire 128 bits as a cipher text. The table is sorted using 32 bit elements, and if two entries in the table are same, re-encrypt and compare the entire encrypted value. This Optimization makes the intermediate table smaller, and lowers the amount of copying required during the sort.

B. Cycle Walking

The Cycle-walking construction works by encrypting the plaintext with an existing block cipher (AES or 3DES) repeatedly until the cipher becomes in acceptable range If we have a plaintext X, create an FPE algorithm from the block cipher by repeatedly applying the AES OR 3DES until the result is satisfying required FPE range.

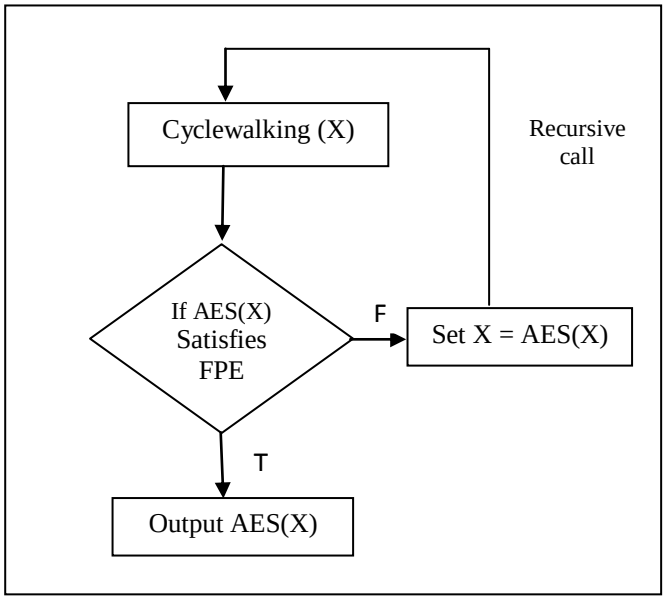


Fig. 3 Cycle Walking

```
CycleWalking FPE(x)
 {
  if AES(x) is an element of M
    return AES(x)
  else
    return CycleWalking FPE(AES(x))
 }
```

The recursion is guaranteed to terminate.

1) Cycle Walking Performance

The larger the difference is between the size of the cipher text and the size of the required FPE output range. [6]. It needs too many iterations. For example if we want to encrypt 64 bits input. The standard block cipher such as AES produces 128 bits cipher text as output. To maintain FPE, AES is repeatedly applied until the higher order 64 bits becomes zero. when plaintext is much smaller than AES domain, that large number of iterations are required for each operation.

C**. Feistel Network**

The Feistel + Cycle construction is the combination of two main techniques. First, the Feistel network that is generated for the size of the given plaintext. This network used to encrypt the data. The cycle-walking technique is applied to the cipher text to provide the cipher text in appropriate range. In Feistel network the sub keys are calculated at each round, The pseudo random values generated by AES algorithm are used as a sub key. Like cycle walking repeatedly executing the Feistel network until the required FPE range is reached.

1) Feistel Mechanism

The standard version of a Feistel network works like this. Assume that X has an even number of bits. Partition it in to a left-hand side L and a right-hand side R. Take the right hand side R and apply to it some key-dependent round function. Next xor together the already-mentioned left-hand side L and the processed right-hand side R∗ to get the new right hand side R'.. The old right-hand side becomes the new left-hand side L'. . This is round-1 of the classical Feistel network, mapping a left and right side, (L, R), to a new left and right side, (L',R'). Each round the round function will differ [5]. We can use any number of rounds. The round function performs the following operation to calculate the new L and R values:

R' = L XOR f(R)

L' = R

Feistel structure to encipher a 5-digit number, say the number 23456. To encode a 5-digit number needs 16 bits are required. Converting 23456 as a 16-bit binary number, getting 0101101110100000. F is a round function. Superscript specifies round number and subscript specifies key value. Here I never specify the round function.

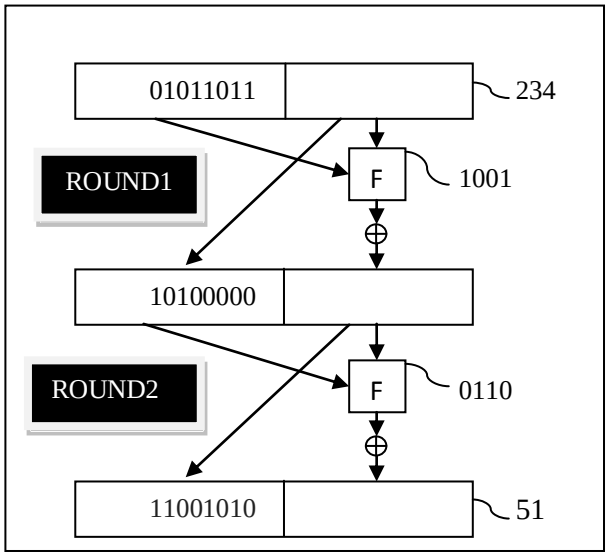


Fig . 4 Feistel Network

The initial left-hand side L; the next 10 bits are the initial right-hand side R . The 10-bit numbers that result from each round function have been randomly generated in the figure

4; I specify only 2 rounds of Feistel network. We can use any number of rounds. Encrypting 5 digit number we get 5 digit number as a output.

1) **Feistel + Cycle Method**

In an above example there is a possibility for getting 6 digit number as output. At that time keep rechipering until get the 5 digit output. This method is called cycle walking. Repeatedly applying the Feistel network until we get the desired output.

2) Feistel + Cycle Method Performance

The generalized Feistel has been the most widely used method to build FPE scheme. The main advantage in Feistel network is the size of the input can be changed. The Feistel + Cycle construction's performance is dependent upon the number of rounds used and the specific PRF (Permutation Round Function) that is used in the round function. For any plaintext that is smaller than the block size of the PRF, the performance is essentially i*r*cost(PRF), where r is the round count and i is the average number of times the cipher cycles to get an acceptable output[7].

V. Comparative Study

The following table shows the performance of the FPE techniques on a 2.34 Ghz Pentium IV with 1 GB memory running Microsoft Windows XP Professional.[6]

TABLE III. COMPARATIVE TABLE

ECHNIQUE NAME	NUMBER OF BITS ENCRYPTED	TIME REQUIRED IN MILLI SECONDS
Prefix cipher (using AES -256)	20	760
Cycle Walking (using 3DES)	64	15000
Feistel + Cycle (using AES-256)	56 (32 Rounds)	10500

The prefix method works on only small data set. The Cycle-walking construction, like the Prefix method, is quite simple, but works on a limited class of sets. The performance of the Feistel + Cyclic method is based on number of rounds constructed and round function PRF used in the network.

VI. Conclusion

Most of our real life applications such as credit card number and social security number require format preserving encryption. Using Format preserving encryption the data base schema and applications will never changed. The cost and time for modifying the data base is reduced. An individual technique alone is not secured For better security we use combination of more than one techniques and also increase the number of permutations at the time of encryption. In future FPE will be applied to all the data types.

References

[1] M. Bellare, T. Ristenpart, P. Rogaway, and T. Stegers. *Format-preserving encryption*. Full version of this paper. 2009.

[2] H. E. Smith and M. Brightwell. *Using Datatype-Preserving Encryption to Enhance Data Warehouse Security.* NIST 20th National Information Systems Security Conference, pp.141, 1997.

[3] *BPS: A Format-Preserving Encryption Proposal* Eric Brier, Thomas Peyrin and Jacques SternIngenico, France

[4] J. Black and P. Rogaway. Ciphers with Arbitrary Finite J. Black and P. Rogaway. *Ciphers with Arbitrary Finite Domains.* RSA Data Security Conference, Cryptographer's Track (RSA CT '02), Lecture Notes in Computer Science, vol. 2271, pp. 114-130, Springer, 2002.

[5] *A Synopsis of Format-Preserving* Encryption Phillip Rogaway March 27, 2010

[6] *Format Preserving Encryption* Terence SpiesVoltage Security, Inc.

[7] M. Bellare, P. Rogaway, and T. Spies. *The FFXmode of operation for format-preserving encryption*(Draft 1.1). February, 2010. Manuscript (standards proposal) submitted to NIST.

Implementation of Register Files in the Processor of Hard Real Time Systems for Context Switching

[1] Prof. B Abdul Rahim, [2] Mr. S.Narayana Raju, [3] Mr. M M Venkateswara Rao
[1] Head, Dept. Of ECE AITS, Rajampet, kadapa (Dist.), A.P, India
[2] M.Tech (Embedded Systems), NECT, Hyderabad, RR (dist.), A.P, India
[3] M.Tech (Embedded Systems), AITS, Rajampet, kadapa (Dist.), A.P, India

Abstract
Embedded Real Time applications use multi threading to share of real time application. The advantage of multi-threading include greater throughput, more efficient CPU use, Better system reliability improved performance on multiprocessor computer.
Real time systems like Flight control systems require very precise timing; multi threading itself becomes an overhead cost mainly due to context switching of the real-time operating system (RTOS). In this paper we propose a new approach to improve the overall performance of embedded systems by implementing register files into processor register bank itself. So that the system use multithreading by moving the context switching component of the real-time operating system (RTOS) to the processor hardware. This technique leads to savings of processor clock cycles used by context switching, By this approach the hard real time embedded systems performance can be improved

Keywords: Context Switching, Operating System, Memory, Simulation, Programming, Hard Real Time, Log Terminal

I. INTRODUCTION

In general, an operating system (OS) is responsible for managing the hardware resources of a computer and hosting applications that execute on the computer. A RTOS Is a Specialized type of operating system designed to execute applications with very precise timing and a high degree of reliability. They are intended for use with real time applications. Such applications include embedded systems (such as programmable thermostats and household appliance controllers), industrial robots, spacecrafts, industrial controllers, scientific research equipments, etc. RTOS can be divided into two categories, hard real-time and soft real-time systems. In a hard real-time or immediate real-time system, the completion of an operating after its deadline is considered useless, and this may cause a critical failure of the complete system and can lead to an accident (e.g. Engine Control Unit of a car, Computer Numeric Control Machines). Usually the kernel divides the application into logical pieces Commonly called threads and a kernel that coordinates their execution. A thread is an executing Instance of an Application and its context is the contents of the processor registers and program counter at any point of time. A scheduler, a part of the Real Time operating System's kernel, schedules threads execution based upon their priority. Context switching function can be described in slightly more detail as the Kernel performs different activities with regard to threads on the CPU as follows.

1. Suspend the progress of current running thread and store the processor's state for that thread in the memory.
2. Retrieve the context of the next thread, in the scheduler's ready list, from memory and restore it in the processor's registers.

Context Switching occur as a result of threads voluntarily relinquishing their allocated execution time or as a result of the scheduler making the context switch when a process has used ip its allocated time slice. A context switch can also occur as a result of a hardware interrupt, which is a signal from a hardware device to the kernel indicating that an event has occurred. Storing and restoring processor's registers to/from external memory (RAM) is a time consuming activity and may take 50 to 80 processors clock cycles depending upon context size and RTOS design. If the system needs to respond to an event in less than this time, the event response has to be implemented as an Interrupt Routine (ISR). On the other hand, if several events happen continuously, then the overall performance of the system may not be acceptable as most treads may not get a chance to execute. To improve responsiveness, the context switch time needs to be reduced. In general, there are two factors that effect the context switching cost. Direct cost due to moving the processor's registers to and from external memory or cache and indirect cost because of perturbation of cache, CPU, pipeline, etc. This presents difficulty in estimating the total cost of context switching cost[2], several algorithms have been developed and implemented to reduce the direct cost of context switching[3][4][5]. As discussed earlier, context registers need to be saved externamemory one at a time. A MIPS processor with 12 registers (9 temporary registers, stack pointer, global pointer and

program counter), need to be saved, it requires at least 2 X 2 X 12=48 clock cycles to switch the context. Using our suggested approach, this context switch time would be reduced drastically to 4 processor's clock cycles independent of the number of context registers.

The approach presented in this paper is to save the context in newly created context register files. These register files are implemented in the processor hardware itself as part of the processor's register bank module. To test and achieve the required performance, software also needs to be modified to exploit the suggested hardware design.

II. HARDWARE IMPLEMENTATION

To prove the concept and measure the performance of our suggested approach, MIPS processor architecture was selected [6] and the suggested approach was implemented on top of it. The register bank module of the processor was modified by adding register files in the register bank to save context. The size of each register file is equal to context size. We have implemented 6 register files. The block diagram of the modified MIPS architecture as shown in figure 1.

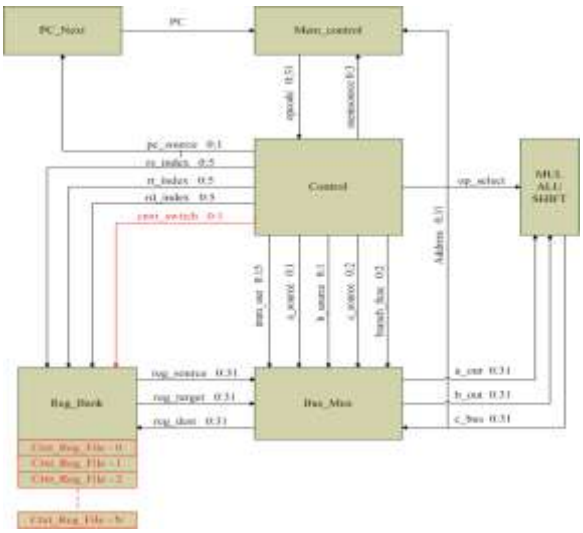


Figure 1. Modified MIPS processor architecture

To access these register files, two special context switch CPU instructions have been implemented: 'sext' and 'rext' to save and restore the processor's context respectively. The index will be stored in temporary register that to which register file the context will be stored and from which register file the context will be restore. Example register $4 contains the index of the register file.

III. SOFTWARE IMPLEMENTATION

To implement the software we have installed VMWARE for Linux operating system environment. VMWARE is an utility to enter from one operating system to another operating system.

The software implementation is divided into two parts. The first part deals with modification of GNU MIPS tool-chain by adding the 'scxt' and 'rcxt' instructions to the GNU-MIPS assembler [8].

The second part mainly deals with the implementation of a small co-operative operating system that executes the threads in round robin fashion using the newly implemented context switch instructions. The'sext' and 'rext' was developed using mips assembly language and the cooperating operating system was developed using C- language. Those two files executed in VMWARE linux environment. These object files added to the GNU MIPS assembler so that the GNU MIPS tool-chain gets modified. This operating system supports context switching using external RAM locations as well as internal register files.

Figure 2. Shows the co-operating operating system's tread structure. To achieve the fast context switch using internal register files, application has to set the 'FastCtxtSwitch' member of the task structure to 1 for that particular thread at the time of thread creation.

```
#include "plasma's"

#define CONTXT_SIZE 15
typedef void (*TaskFunc)(void)
typedef  struct task
{
Void (*Taskptr)();
……..
……..
……..
}

Void createTask (….)
  {
If (cnxt_type= =0)
{Threads[TaskID].State=0;
Threads[TaskID].Fastctxtswitch=1;
}else
{
Threads[TaskID].FastCtxtSwitch=0;
}
```

Figure 2. Task Structure

IV. SIMULATION RESULTS

Xilinx Spartan 3E 1600E the advanced development board was used for our test and experimentation [9]. Xilinx ISE 10.1 was used to design the hardware and the simulations were executed using the advanced simulation and debugging toolset, ModelSim [10]. Figure 3 shows the simulation wave form. Context registers are saved in the context switch register file-2 in 2 clock cycles.

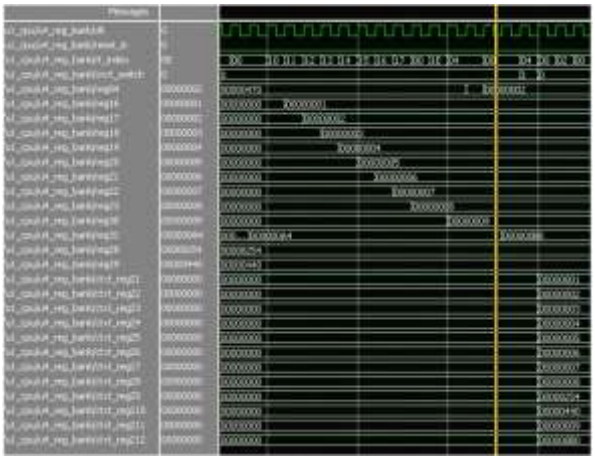


Figure 3. Simulation results for 'scxt' instruction

Similarly figure 4. Shows the simulation waveform of the 'rext' instruction. The context registers are also being restored from the context switch register file-2 in 2 clock cycles.

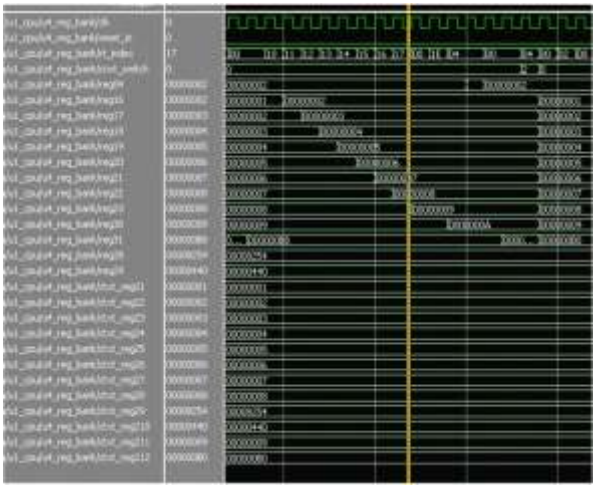


Figure 4. Simulation results for 'rext' instruction

V. PROGRAMMING VIA IMPACT

After successfully compiling an FPGA design using the Xilinx development software, the design can be downloaded using the iMPACT programming software and the USB cable. To begin programming, connect the USB cable to the starter kit board and apply power to the board. Then double-click configure Device (iMPACT) from within Project Navigator. To start programming the FPGA, right click the FPGA and select program. When the FPGA successfully programs, the iMPACT software indicates success, as shown in Figure 5.

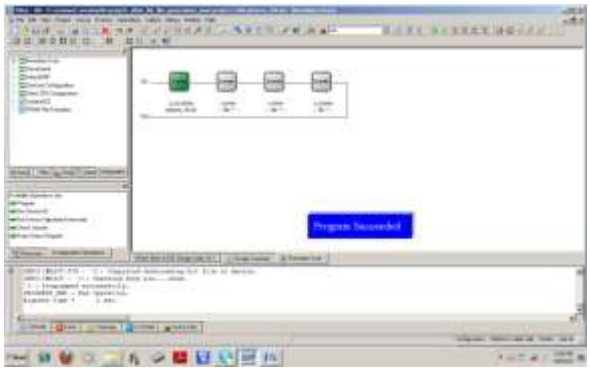


Figure 5. iMPACT Programming Succeeded

VI. TEST APPLICATION 1

This Application tests the successful operation of the proposed approach by switching four threads using internal register files. This test is used to ensure that data between threads is not corrupted and thread's context switching is correct. The flow chart of this application is shown in figure 6.

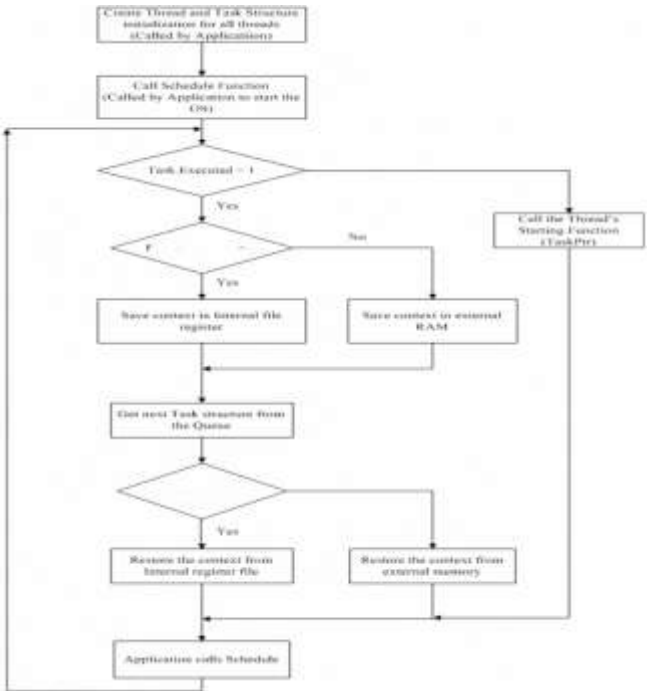


Figure 6:Flowchart for Test Application -1

There are four Tasks Executes in round robin fashion. TaskID=0, TaskID=1, TaskID=2, TaskID=3 executes one by one, finally the TaskID3 calculates and prints the Number of Clock cycles consumed to process one data sample. These Tasks will send the messages to the debug terminal port, the output log received on the debug terminal as shown in figure 7.

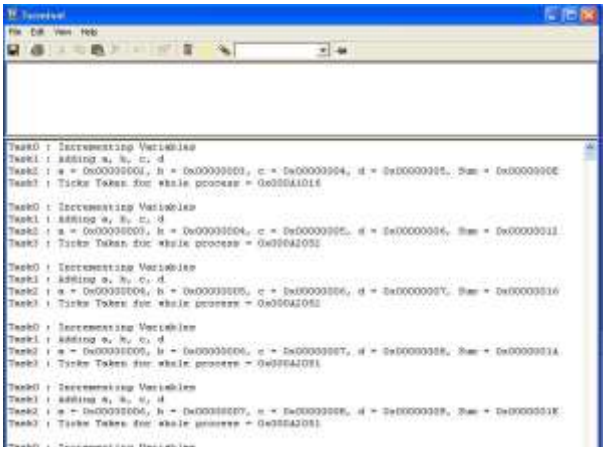


Figure 7: Serial Debug Log from Test Application -1

VII. TEST APPLICATION-2

The Second application is designed to measure the performance improvement, in clock cycles. It creates four threads that executes in never-ending loops. First two threads does context switching using internal register files. Next two threads does context switching using external memory. Message send to the debug serial port include: Number of Clock cycles consumed by first two threads, Number of clock cycles consumed by the next two threads. Finally, the difference between these two. Figure 8 shows the output log for this test.

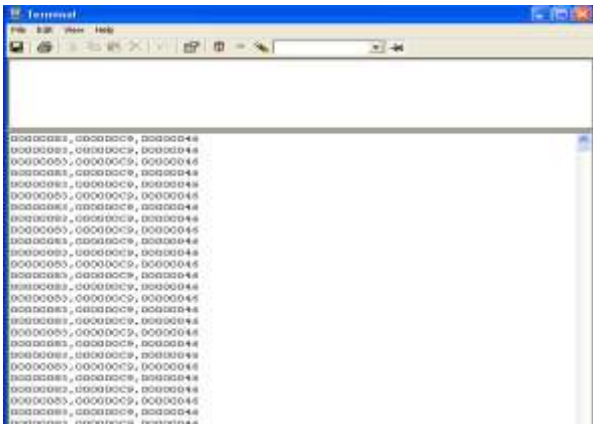


Figure 8: Serial Debug Log from Test Appl-2

VIII. CONCLUSION

This paper Presents a new approach as an attempt to improve the hard RTOS system performance which helps in meeting the deadlines in an efficient Way. It boosts the system performance for such systems where lot of context switching is required in a small period of time. This approach can be easily extended to soft RTOS and regular operating systems to improve the overall system performance. In these systems, threads are created at run time and it is difficult to know the number of threads at the design time. Therefore, threads that are part of frequent context switching using internal register files using the specially designed scheduler algorithm. This paper also takes another step forward in moving the real-time operating system kernel to hardware and programming iMPACT using Xilinx software.

IX. REFERENCES

[I] http://www.rtos.com/PDFs/ AnalyzingReal-TimeSystemBehavior.pdf

[2] Francis M. David, Jeffery C. Carlyle, Roy H. Campbell "Context Switch Overheads for Linux on ARM Platforms" San Diego, California Article No. : 3 Year of Publication: 2007 ISBN:978-I-59593-751-3

[3] Zhaohui Wu, Hong Li, Zhigang Gao, Jie Sun, Jiang Li An Improved Method of Task Context Switching in OSEK Operating System" Advanced Information Networking and Applications, 2006. AlNA 2006. Publication date: 18-20 April 2006 ISSN : 1550-445X

[4] Jeffrey S. Snyder, David B. Whalley, Theodore P. Baker "Fast Context Switches: Compiler and rchitectural support for Preemptive Scheduling" Microprocessors and Microsystems, pp.35-42, 1995. Vailable:citesser.ist.psu.edul33707.html

[5] Xiangrong Zhou, Peter Petrov "Rapid and low-cost context-switch through embedded processor customization for real-time and control applications" DAC 2006, July 24- 28 San Francisco, CA [6] http://www.opencores.org/projecLplasma

[7] MIPS Assembly Language Programmer's Guide, ASM - 0 I-DOC, PartNumber 02-0036-005 October, 1992

[8] http://ftp. gnu.org/gnuibinutilsl

[9] MicroBlaze Development Kit Spartan-3E 1600E Edition User Guidewww.xilinx.com

[10] http://model.com/contentlmodelsim-pe-simulation-and-debug

A Benefits Estimation Model for Software Reuse Based Program

[1]Shobha Rani Malik, [2] Dr. Saba Hilal

[1] (Lecturer,Technical Edu. Deptt. Haryana and Research Scholar at Lingayas' University Faridabad, India)
[2](Research Guide, Lingayas' University Faridabad, India)

Abstract:
In software engineering, reuse of artifacts is considered as silver-bullet for project development specially in component based development technic.The benefits of reuse are neither linear nor exact they are estimated with the help of metrics and models. To measure benefits of software reuse is a challenging task since they are multi dimensional mainly earlier time-to-market ,better quality, increased productivity etc. and of two dimensional characteristics- quantitative and qualitative . Reuse is highly potential in correlated domains due to similar requirements.

Keywords: Software Reuse Benefits estimation, Producer and User Benefits, Software Reuse Benefits increasing factors

1.Introduction

"Actually it is hard to see how much time or budget u have save during the project developing. it is also quite impossible to set up the goal and timetable about the reuse work in the beginning of the project" [14].Cost savings is the most promoted benefit for reuse, but benefits also exist in risk, schedule, and performance [9]. Estimation of benefits is not easy since they depend up on management, technology and organizational factors[38] . Generally the largest payoffs of software process assessment and improvement are in human terms - pride in work, increased job satisfaction, improved ability to attract and retain software experts and not in dollars [3][4]. Benefits include total life cycle costs saved as well as additional profits resulting from earlier completion of the product[7].

1.1 Software Reuse Benefits

Benefits in software engineering [48][49][50][51][26][28][8] due to reuse of artifacts are mainly as shown in figure1.

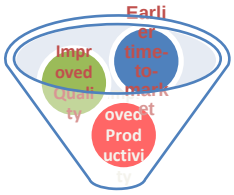


Fig 1. Reuse Benefits

Both the producer and the user enjoy benefits of reuse artifacts in a reuse program.

- User Benefits

User enjoys benefits of reusing well tested, more reliable, higher standards artifacts by reducing development and maintenance costs[1][49]with better quality , improved productivity that results additional revenue, improved sales and increased market share due to earlier market delivery of product.

- Producer Benefits

Producer takes benefits of selling product with high price, fees and royalties, [1] that results increased market share.
When producer does not explicitly charge for its components or service, reuse may be economically feasible for the user but not for the producer [7].
A producer assumes a higher cost to design and implement a reusable asset than a custom developed asset, but the consumer saves time and money by not having to develop the component [35]. If a software developer have to play both user and producer roles then for organization's benefits, creation cost should be less than consuming cost.

2. Related Work

Various studies [20][6][21][10][29][18] [19][26][24] have estimated only direct(quantitative) benefits of reuse generally measured costs avoidance that is a major factor of reuse benefits but a very few measured[28][15][13] some limited indirect(qualitative) benefits of reuse. in a In this research, a model is proposed for a reuse oriented organization which can

estimate direct as well as indirect reuse benefits by considering almost all factors that contribute in benefits of reuse by extending and filling the gap left by other studies. We reviewed some measurement tools-metrics and models to estimate benefits of software reuse and also have suggested some benefits increasing factors.

3. Proposed Work

3.1 Reuse Benefits Estimation Metrics and Models

To estimate benefits of reuse suitable measurement technique should be applied. Various metrics and models are proposed by researcher to measure reuse benefits , some of them are reviewed here. In software engineering reuse has been claimed as an important source of saving costs [2].

- Cost Saving:

$Cost_{saved} = Cost_{scratch} – cost_{with\ reuse}$ So, as the number of element increases the savings increases also[46].
$Cost_{saving} = Cost_{scratch} - Cost_{reuse} - Cost_{delivered}$ [42]
Where $Cost_{scratch}$ = cost of developing software from scratch, $Cost_{reuse}$ = slide cost associated with reuse and C_d = cost of software delivered
According[13], Saving =∑((cost of artifacts * reuse rate of artifacts) – cost (avg. integration effort of artifacts* reuses of artifacts))
According [20][21] , Total savings due to reuse= [(Savings due to avoided cost -Relative cost of reuse)*No of uses]-[Cost to add the compnent to the library+Cost to maintain the component in the library]
% Savings = [0.92 - 1.5*j/i]*100 , where j = number of reusable software components that have been built, i = number of attempted instances of reuse[34].

Benefits = activity cost without reuse- activity cost with reuse [5].
Benefits = ∑(Development without reusable products - Cost of adptation)- Total cost of all resources[6].
Benefits = graphs of the reuse cost ratio (cost with reuse/cost without reuse) versus the breakeven point. [29].
In the GTE-Contel model[11]
Net benefits =∑ (net cost no reuse - net cost with reuse reuse) –total reuse investment
Benefits (system S) = [Cost of developing S without reuse - Cost of developing S with reuse]/ Cost of S with out reuse. [18]
These benefits are sensitive to implementations, reuse strategy and reused components cost.
As reuse leverage increases, benefits also increases [7].
Reuse Cost Avoidance (RCA)= Development Cost Avoidance (DCA) +Service Cost Avoidance(SCA) [19].
Benefit investment = (cost of project without reuse - actual cost of project with reuse)/ cost of producing reusable components[26]

- Quality :

It is very hard to quantify quality benefit due to its intangible nature. It is generally measured either in savings in maintenance or in finding and removing errors.
According[17] , Defect removal efficiency $=DRE_i = E_i / (E_i + E_{i+1})$, where E_i = the number of errors found in the i^{th} component and E_{i+1} =the number of errors found after integrating $(i+1)^{th}$ component with the i^{h} component.
Quality is savings in maintenance[28].
Quality is finding and removing errors[15].
High level of reuse correlates with a low defect [52].
Reuse rate, development time and decreases in number of errors are highly correlated [53].
Adequacy (library contents) = number of artifacts reused + number of not available but required artifacts + number of artifacts in library and Goodness (artifacts repository)= Defect density of reused artifacts + Avg. reusability score [13] .
Quality of an instrument (Q) = B/R Where Q <B(producer activities) and R(consumer activities).[31]

- Productivity:

According [18] , Productivity = size of the system /cost spent to develop system i,e
Productivity = a(1+ Reuse benefit)b , a, b are Coefficients estimated with standard least squares regression
Productivity= (Reuse KNCSS +Modified KNCSS)/ Product Total KNCSS]*100 [10].
Productivity =1/cost [16].

3.2 Benefits estimation for a reuse program

We are estimating benefits of a reuse program in a hypothetical scenario of a corporation that starts its reuse initiative with domain engineering in 2007 developing reusable components as shown in Table1, that are used in applications

internallyasshown in Table2, and are also sold externally to corporation for a period of 4 years.

Table1: Components Details

Year	Component	Size
2007	X1	5k
2008	X2	10k
2009	X3	15k
2010	X4	20k

The following assumptions are taken with respect to this program :

- ✓ Cost of the reused components is double as compared to similar components made for single use.
- ✓ Salary of employees is fixed, not depending upon size and quantity of components .Overhead cost to make reusable components is 10 % of salary of employees and all other details are as shown in Table3.
- ✓ Cost of purchased component(yi) that are used in applications is 25% extra of cost of internally developed component(yi).
- ✓ Sell price of internally developed components and applications is 25% extra of their respective costs.
- ✓ Set-up cost of corporation for reuse program is $3,000 and of application- engineering cycle is $1,000 .

Table 2: Applications Details

Year	Application	Component used internally developed	Component used externally developed	Additional-Code
2008	App(1)	X1	O_1	2K
2009	App(2)	X1,X2	O_1,O_2	4K
2010	App(3)	X1, X2,X3	O_1,O_2,O_3	6K

Table 3: Salary of Employees

Personnel	Initial Salary ($)	Increment/y (%)
component developer for reuse	1000	15
manager for reuse	800	10
librarian for reuse	700	10
Domain Analyst	800	10

3.2.1 Benefits Structure:

In above said organization , a component developer roles as creator (producer) in Domain Engg. cycle and component user as consumer in Application- Engg. cycle ,so all benefits are of Corporate. It is not possible to quantify all benefits of reuse specially qualitative benefits but it is tried to measure them. Benefits of reuse program for above said corporation are structured as shown in Fig.2. Time value of money is not considered here.

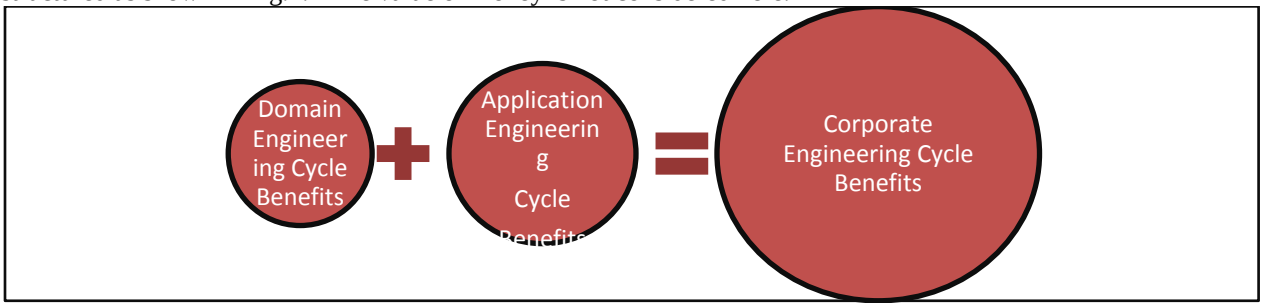


Fig 2. Benefits Structure

Corporate Engineering Cycle Benefits= Domain Engineering Cycle Benefits from selling components internal and external to the corporation + Application Engineering Cycle Benefits from selling applications external to the corporation.

3.2.1.1 Domain Engg. Cycle Benefits

$Benefits_{Domain\text{-}Engg} = Benefits_{comp\text{-}Engg} - Salary_{Domain\text{-}Analyst}$

$Benefits_{comp\text{-}Engg} = \sum Benefits_{comp}$

Where

$Benefits_{Domain\text{-}Engg}$ = Benefits of Domain Engg. Cycle

$Benefits_{comp\text{-}Engg}$ = Benefits of Component Engg. Cycle

$Benefits_{comp}$ = Benefits from selling a component internal and external to the corporation

$\sum Benefits_{comp} = \sum Direct\ Benefits_{comp} + \sum Indirect\ Benefits_{comp}$

It is assumed that all components (X1, X2, X3) made before 2010 are sold internal and external to the corporation and fees and royalties are 1% of sell price for externally sold components. Component X4 made in 2010 is not included in cost-benefit analysis since it will be used internally in next plan(App4,App5,App6) and sold externally in next coming year(2011) according proposed scheme.

- $\sum Direct\ Benefits_{comp}$ = [Saving due to sell - Additional cost due to make reusable]

[(Total sell price -Total cost to make reusable components) + fees and royalties] - [Total cost to make reusable components - Total cost to make for no reuse]

=[{(Sell price of components internal to the corporation+ Sell price of components external to the corporation)- Total cost to make reusable components } +fees and royalties]- [Total cost to make reusable components -Total cost to make for no reuse]

[{(9281.25+11601.56)- 9281.25}+116.0156] - [9281.25 -4640.625]

=11717.58-4640.625 = $7076.953

- $\sum Indirect\ Benefits_{comp}$ = increased market share +Improved quality Increased market share- due to revenue from selling components external to the corporation + Improved quality and productivity in terms of increased experience of staff (more trained staff) in same environment.

$Benefits_{Domain\text{-}Engg} = (7076.953\text{-}2648) + \sum Indirect\ Benefits_{comp}$

$Benefits_{Domain\text{-}Engg} = \$4428.953 + \sum Indirect\ Benefits_{comp}$

3.2.1.2 Application -Engg. Cycle Benefits

$Benefits_{App\text{-}Engg} = \sum Benefits_{App}$

Where

$Benefits_{App\text{-}Engg}$ = Benefits of Application Engg. Cycle

$Benefits_{App}$ = Benefits of a application

$\sum Benefits_{App} = \sum Direct\ Benefits_{App} + \sum Indirect\ Benefits_{App}$

$\sum$Direct $Benefits_{App}$ = [Saving due to sell- Saving due to reuse]

= [Total sell price-Total cost with reuse]

= [Total sell price of applications external to the corporation – (Total cost of applications with reuse +Set-up cost -Total cost of repeated components)]

= [(Total sell price of applications external to the corporation – Total cost of applications with reuse) - Set-up cost + repeated cost of components]

= (.25* Total cost of applications with reuse) -1000 + cost (2X1+X2)

= (.25* 45285.19 +2*2750+ 3080)-1000

= 11321.3+ 5500 + 3080-1000

=19901.3-1000 = $18901.3

- $\sum$Indirect Benefits$_{App}$ = additional revenue due to earlier market delivery of product + increased sales (increased productivity due to reusing components +Improved quality) +increased market share

3.2.1.3 Corporate Engg. Cycle Benefits

$Benefits_{Cor\text{-}Engg}$ = $Benefits_{Domain\text{-}Engg}$ + $Benefits_{App\text{-}Engg}$ –Set-up cost , Where

$Benefits_{Cor\text{-}Engg}$ = Corporate Engg. Cycle benefits

i, e $Benefits_{Cor\text{-}Engg}$ = [(Direct Benefits) $_{Domain\text{-}Engg}$ + $(Direct\ Benefits)_{App\text{-}Engg}$] + [$(Indirect\ Benefits)_{Domain\text{-}Engg}$ + $(Indirect\ Benefits)_{App\text{-}Engg}$] -3000
=[4428.953 +18901.3]-3000 + [$(Indirect\ Benefits)_{Domain\text{-}Engg}$ + $(Indirect\ Benefits)_{App\text{-}Engg}$]

= $20330.25 + [$(Indirect\ Benefits)_{Domain\text{-}Engg}$ + $(Indirect\ Benefits)_{App\text{-}Engg}$] = $20330.25 + Increased market share- due to additional revenue due to earlier market delivery of product and selling components external to the corporation+ Increased sales (Improved quality +increased productivity due to reusing components) +Improved quality and productivity in terms of increased experience of staff (more trained staff) in same environment .

3.4. Benefits increasing factors in software reuse scenario

Following efforts can help to increase benefits in context of above said corporation:

- Rich Component repository

Organization ‘s own repository should be rich of components since in house built components are lesser cost than components purchased from market. Library should be updated according needs and new components should be regularly introduced [39].

- Maximize black-box reuse

Black-Box Reuse(CBD form) is costly due to inefficiency in search techniques and components price [40] but in house built black-box components are designed according internal reuse needs and forward looking reuse as well as this reduce the understanding time since a reusable asset require more understanding time than any other part of reuse process [22].So search cost and component price are very reasonable in house built black-box component.

- Minimize Complexity

In our model , as Complexity of component increases ,cost increases also .So try to Minimize it.

- Effective Size

Cost increases according size , so it should be effective and can be calculated as

Effective size = (existing size × (0.4 × redesign% + 0.25 × reimplementation% + 0.35 × retest %)) [32].

- Avoid Reuse Failure

Failures of reuse increases programmer's time, decreases motivation, accountability and incentives so it should be avoided [33].

- Start with small scale and grow-up

At initial level organizations should start from low level of reuse with small size of components since they are easy to adapt[38] and make and then grow-up with increasing size of components.

- Include reuse as integral part

Reuse should be as an integral part of software life cycle process[28] by providing central support for reuse activities .

- Improve Quality

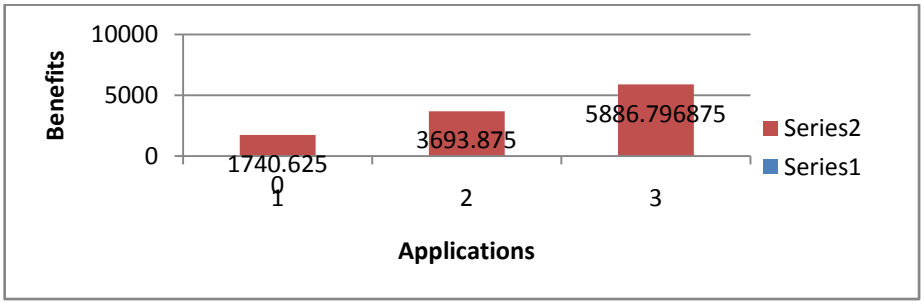
Quality benefits can be achieved by error reduction and standization . Reuse rate of component should be increased since defect are fixed each time that results –improved quality[10]. Try to remove defects in earlier phases as soon as possible since cost of prevention and debugging can be amortized number of uses [27] and the longer it takes to identify an error, the costlier it is to fix[35][36] .Initiatives should be provided to establish methods to avoid post release errors[37].

- Minimize product delivery time

Try to avoid missing a market window since in reused oriented environment ,sometimes earlier time-to-market can be more profitable and competitive advantage than direct cost reduction[41][44][45][47].

- Use common technical standards [12].
- Increase Scale of Reuse

In our model , Components are used across projects. For example , as component X1 is used in App1,App2 and App3 then benefits increase also according scale as shown in Graph1.



Graph1. Scale Vs. Benefits

- Share knowledge and experiences

Organization can introduce reuse groups and forums where people can discuss problems so that experiences should be shared by other organizations [23] .

- Make components according market needs

Design and functionality of components market should be according needs.

- Up gradation and incentives for staff

Organization should establish refresher courses, training for staff up gradation to compete market trends and recognition, royalty , lottery award and preferential funding policy for reuse achievements[25] . Set targets and use competitions to drive reuse [30] to increase and promote the reuse level.

- Improve Continuously

Perform walkthroughs, reviews , feedback from users and monitor reuse efforts throughout development life cycles.

4. Conclusion

Both producer as well as user enjoy benefits in a software reuse scenario specially user. Benefits may be operational or strategic ,direct or hidden, quantitative or qualitative. Some measures(metrics and models) to estimate the benefits of a reuse program are reviewed. In this research, a simplified and enhanced benefit estimation model is suggested for all cycles of software reuse development .The Proposed study also suggests that rich repository, black-box strategy, failure avoidance development according market requirements, better quality , standardization of components as well as continuously improvement with incremental business strategy by keeping staff up graded, happy and sharing knowledge ,experiences with others ,the corporate(organization) can increase its benefits.

5. Future Work

Economic worth of proposed scheme for above said organization will be estimated.

References

[1]. M.R.V. Chaudron," Software Engineering Economics with special attention to Component-& Reuse-based Software Engineering"www.win.tue.nl/~mchaudro,Dept. of Mathematics and Computing Science Eindhoven University of Technology, Based on IBM Systems Jnl, 1993, Vol 32, No, 4

[2]. Barry W. Boehm. Software Engineering Economics. Prentice Hall, Engle-wood Cli_s, NJ, 1981.

[3]. Krasner, H. ,The Payoff for Software Process Improvement: What it is and How to Get it. In K. Emam & N. Madhavji(Eds.), Elements of Software Process Assessment and Improvement (pp. 151-176) , Washington, DC: IEEE Computer Society Press,1999.

[4]. Ronald G. Wolak, DISS 725 – System Development: Research Paper 4,Software Process Assessment and Improvement Models,wolakron@nova.edu.

[5]. Barnes, B. and T. Bollinger. Making Reuse Cost Effective. IEEE Sqftware, Jan. 1991,8(l): 13-24.

[6] . Bollinger, T., and S. Pfleeger. Economics of Reuse: Issues and Alternatives. Information and Software Technology, December 1990, 32(10): 643-652.

[7]. Nasib S. Gill,Reusability Issues in Component-Based Development, nsgill_2k4@yahoo.com

[8]. Hui Zhou and Monan Yao, Software Reuse And Off shoring : A Study Of Benefits, Difficulties And Feasibility, Bachelor of Applied Information Technology Thesis, Report No. 2010:035,ISSN: 1651-4769

[9]. Jared Fortune1, Ricardo Valerdi2, Barry W. Boehm3 and F. Stan Settles4 . Estimating Systems Engineering Reuse, 7th Annual Conference on Systems Engineering Research 2009 (CSER 2009)

[10]. Lim, W. Effects of Reuse on Quality, Productivity and Economics. IEEE Software,September 1994, 11(5): 23-30.

[11]. Margano, J. and T. Rhoads. Software Reuse Economics: Cost-Benefit Analysis on a Large Scale Ada Project. In Proceedings, International Conference on Software Engineering, Melbourne, Australia, 11-15 May 1992: 338-348. 141-155.

[12]. McClure, C., "Software Reuse: A Standards-Based Guide", IEEE Computer Society, 2001, ISBN 076950874X.

[13]. Basili, In the context of the Goal Question Metric ,Lin98.

[14]. Henrik, a software development professor in Sweden.

[15]. Gaffney, J. and R. Cruickshank. SPC-TR-88-015 Version 1.0, Software Productivity Consortium, Reston, VA, April 1988.

[16]. Gaffney, J. and T. Durek. Software Reuse —Key to Enhanced Productivity: Some Quantitative Models. Information and Software Technology, June 1989, 31(5):258-267.

[17]. K. S. Jasmine, and R. Vasantha,' DRE - A Quality Metric for Component based Software Products, World Academy of Science, Engineering and Technology 34 2007.

[18]. Devanbu, P., S.Karstu, W.Melo, and W.Thomas. Analytical and Empirical Evaluation of Software Reuse Metrics. Proceedings, International Conference on Software Engineering, Berlin, 1996.

[19]. Poulin, J. The Economics of Software Product Lines. International Journal of Applied Software Technology , March 1997.

[20]. Poulin, J., and J. Caruso. A Reuse Metrics and Return on Investment Model. In Advances in Software Reuse; Proceedings of the Second international Workshop on SoftwareReusability, Lucca, Italy, 24-26 March 1993: 152-166.

[21]. Lim, W. Reuse Economics: A Comparison of Seventeen Models and Directions for Future Research, Proceedings, International Conference on Software Reuse, Orlando, FL, 23-26, April 1996: 41-50.

[22]. Jeffrey S. Poulin. Measuring software reuse: principles, practices, and economic models. Addison-Wesley Longman Publishing Co., Inc., Boston, MA,USA, 1996.

[23]. Ryan Gerard, Robert R. Downs, James J. Marshall, and Robert E. Wolfe.The software reuse working group: A case study in fostering reuse. In IRI, pages 24{29. IEEE Systems, Man, and Cybernetics Society, 2007.

[24]. Malan, R. and K. Wentzel. Economics of Reuse Revisited. Technical Report HPL-93-31, Hewlett Packard Laboratories, April 1993.
[25]. Reuse incentives, Planning Initiation Implementation Maturity, http://www.rise.com.br 43/46
[26]. Poulin, J. S., Caruso, J. M., and Hancock, D. R. The business case for software reuse. IBM Systems Journal, 32(4), 567-594.
[27]. M.D. Lubars,Affording Higher Reliability Through Software Reusability,Software Eng. Notes, Oct. 1986, p.39.
[28]. nta Fowler Chmiel , An Integrated Cost Model for Software Reuse, for the degree of Doctor of Philosophy in Computer Science, Dissertation submitted in partial fulfillment of the requirements to the College of Engineering and Mineral Resources at West Virginia University, Morgantown, West Virginia,2000
[29]. Schimsky, D. Software Reuse --Some Realities. Vitro Tech Journal, Summer 1992, 10(1):47-57.
[30]. Dr. Jeffrey S. Poulin Measuring Software Reuse Federal Chief Architects Forum11 May 2004, Jeffrey.Poulin@lmco.com,http://home.stny.rr.com/jeffreypoulin

[31]. Sarbjeet Singh, Manjit Thapa, Sukhvinder singh and Gurpreet Singh, Software Engineering - Survey of Reusability Based on Software Component," International Journal of Computer Applications (0975 – 8887) Volume 8– No.12, October 2010.

[32]. G.N.K.Suresh Babu and Dr.S.K.Srivatsa, Analysis and Measures of Software Reusability, International Journal of Reviews in Computing ,© 2009 IJRIC.

[33]. Robert G. Fichman ,Incentive Compatibility and Systematic Software Reuse, Appeared in: Journal of Systems and Software, New York; Apr 27, 2001; Vol. 57, Iss. 1; pg. 4

[34]. L. Amar and J. Cofey. Measuring the benefits of software reuse – examining three different approaches to software reuse. Dr Dobbs Journal, 30:73,76, 2005.
[35]. De Marco, T., Structured Analysis and System Specification, Prentice Hall, Englewood Cliffs, NJ, p. 26 (1979).
[36]. Grady, R. B., "Practical Results from Measuring Software Quality," Commun.ACM 36, 62–67 (1993).
[37]. William R. Bitman ,Balancing Software Composition and Inheritance to Improve Reusability, Cost, and Error Rate,Johns Hopkins APL Technical Digest, Volume 18, Number 4 (1997).
[38]. Mohsin Irshad ,Master Thesis Software Engineering Measuring Cost Avoidance Through Software Reuse, A model to measure costs avoided through software,reuse and guidelines to increase profits through,software reuse, School of Computing, Blekinge Institute of Technology,SE-371 79 ,Karlskrona, Sweden, 2010
[39]. Michael F. Dunn and John C. Knight. Software reuse in an industrial setting: a case study. In Proceedings of the 13th international conference on Software engineering, ICSE '91, pages 329{338, Los Alamitos, CA, USA, 1991. IEEE Computer Society Press.
[40]. By T. Ravichandran and Marcus A. Rothenberge ,Software Reuse Strategies and Component Markets , Communications of the ACM August 2003/Vol.46 No. 8 Pp. 109-114
[41]. William T. Ward. Calculating the real cost of software defects. Hewlett-Packard Journal, pages 55-58, October 1991.
[42]. A Comparison of Approaches to Reuse Investment Analysis John Favaro
[44]. Preston G. Smith and Donald G. Reinertsen. Developing Products in Half the Time. Van Nostrand Reinhold, New York, 1991.
[45]. Martin L. Griss,Software Reuse:From Library to Factory,Software Technology Laboratory,HPL-93-67,July, 1993
[46]. Jasmine K.S and Dr. R. Vasantha ,Cost Estimation Model For Reuse Based Software Products, Proceedings of the International Multi Conference of Engineers and Computer Scientists 2008 Vol I IMECS 2008, 19-21 March, 2008, Hong Kong
[47]. George Stalk Jr. Time - the next source of competitive advantage. Harvard Business Review, 66(4):41-51, Jul-Aug 1988.
[48]. Parastoo Mohagheghi , R. C. (2007). "Quality, productivity and economic benefits of software reuse: a review of industrial studies." Empirical Software Engineering Volume 12(Issue 5): pp: 471-516
[49]. Sommerville, I., "Software Reuse", Software Engineering, 7th Edition, Chapter 18, 2004 http://www.comp.lancs.ac.uk/computing/resources/IanS/SE7/Presentations/PDF/ch18.pdf
[50]. Mili, H., Mili, A., Yacoub, S., and Addy, E. (2002) Reuse-Based Software Engineering.
[51.] Anthony M. Wilson, Joseph G. San Miguel and Michael W. Boudreau, The Impact of Software Reuse on The Cost of Navy Sonar and Fire Control Systems, Naval Postgraduate School Monterey, California Thesis, Approved for public release; distribution is unlimited , June 2009
[52]. Agresti w. and Evanco W. , Projecting Software Defects in analyzing Ada designs,IEEE Trans.Softw. Engg.18,11,988-997,1992
[53]. Browne J., LeeT. and Werth J. , Experimental Evaluation of Reusability-Oriented Parallel Programming Environment, 16,2,111-120, 1990

Growth of Robotics Industry Early in 21st Century

[1], Manshi Shukla and [2], Amar Nath Shukla
[1,] Department of Computer Science, Maharana Pratap Engineering College, Uttar Pradesh
[2,] Department of Electronics, Maharana Institute of Professional Studies, Uttar Pradesh

Abstract:
Robotics Industry includes two types of robots: Industrial Robots and Service Robots. The automation and modernization of manufacturing processes depends on the industrial robots. Service robots are enriching innovation or products development in different applications, whether professional or personal. Industrial and service robotics is mature as a technology and is well established. This paper analyzes the growth in the Robotics Industry worldwide in the early age of 21st century both for industrial robots and service robots. The growth of industrial and service robots is explained on the basis of the average increase in the sales during 2000-2011. In this paper the market trends for both types of robots is compared which contributed in the expansion of robotics industry for the same period.

Keywords: Applications of robots, Automation, Industrial Robots, Market analysis, Robotics industry, Service Robots, Types of robots.

1. Introduction

Robots have a long history. The first digitally operated and programmable robot, the Unimate, was installed in 1961 at General Motors. Since 1961, the world around us is changing at unprecedented and unimaginable speed which gives new pave to Robotics Industry. The robotics industry includes vast range of industrial robots and service robots. As defined by ISO 8373, the industrial robot is 'an automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes, which may be either fixed in place or mobile for use in industrial automation applications.' According to the International Federation of Robotics, another professional organization, a service robot is 'a robot which operates semi or fully autonomously to perform services useful to the well being of humans and equipment, excluding manufacturing operations'. The early age of 21st century sees that both types of robots are capturing the market rapidly to fill the emptiness in our society world-wide. In this paper, the growth of robotics industry is analyzed between 2000 and 2011 for industrial and service robots. During the early age of this century, various applications of robots, ranging from manufacturing processes to non-manufacturing processes and from professional uses to personal or domestic uses, are changing our society. To do this robots are becoming more powerful, with more sensors, intelligence and cheaper components. In order to show the comparisons between service robots and industrial robots, the number of units sold worldwide is analyzed during 2000-2011.

2. Market analysis

Robotics is an exciting and multi-disciplinary area that dominates 21st century. Robotic industry is entering a period of rapid growth[1]. Industrial robots and service robots both contributed in the success of this industry. The statistical data used from Table 1. to Table 3. has been taken from the International Federation of Robotics (IFR) and the Economic Committee data from UN for Europe (UNECE) [2,3,4,5,6,7,8]. The graphical representation, has been showed from Figure 1a.&b. to Figure 8. is prepared after analyzing the Table 1., Table 2. and Table 3.

2.1. Analysis of market for industrial robots

Industrial robotics applications have evolved from simple tasks like 'pick & place' to complex functions like welding and assembling & dissembling. Post 2000, showed growing demand, competition and pressure to increase productivity due to increased globalization and drive of modernization. Since the industrial robots started to be introduced in industry at the end of the 1960s, total accumulated yearly sales amounted to more than 1,149,290 units worldwide at the end of 2000 which is the start of 21st century, where as 1,600,000 units sold worldwide at the end of 2005, that is amounted to more than 2,284,230 units by the end of 2011[2,6,8].

Following information is retrieved from Table 1. in order to explain the growth in industrial robotics worldwide in early 21st century:

1) In 2009, only 60,000 units were sold which is the least amount in early 21st century, because of recession.
2) Year 2011 is showing highest sales amounted to 1,66,028 units ever recorded .
3) During 2005-2011, average increase in sales of 13.65% has been estimated.
4) During 2000-2004, average increase in sales of 0.645% has been estimated.
5) Between year 2004-2005, 24.70% increase is calculated[Figure 2.].
6) In 2005, first time sales estimated above 100,000 units.

7) High intensification in sales is estimated during 2004-2011.
8) Industry showed average growth of 36.39% during 2000-2011.

Table 1. The number of units of industrial robots sold worldwide at and up to the end of the year during 2000-2011[2,3,4,5,6,7]

S. No.	At the End of Year	No. of Units Sold	Up to the end of year
1.	2000	99,000	1,149,290
2.	2001	78,210	1,227,500
3.	2002	68,600	1,296,100
4.	2003	81,800	1,377,900
5.	2004	95,400	1,473,300
6.	2005	126,700	1,600,000
7.	2006	112,200	1,712,200
8.	2007	114,365	1,826,565
9.	2008	113,300	1,939,865
10.	2009	60,000	1,999,865
11.	2010	118,337	2,118,202
12.	2011	166,028	2,284,230

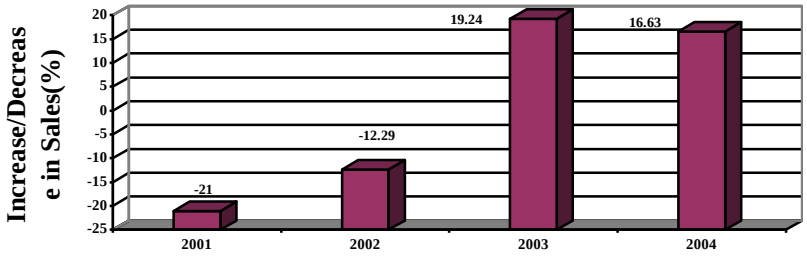


Figure 1a. Increase/decrease in sales (%) of industrial robots at the end of year during 2001-2004

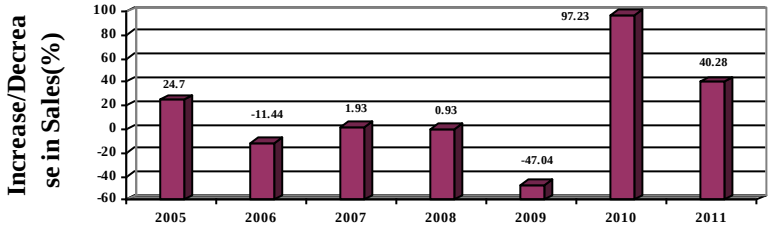


Figure 1b. Increase/decrease in sales (%) of industrial robots at the end of year during 2005-2011

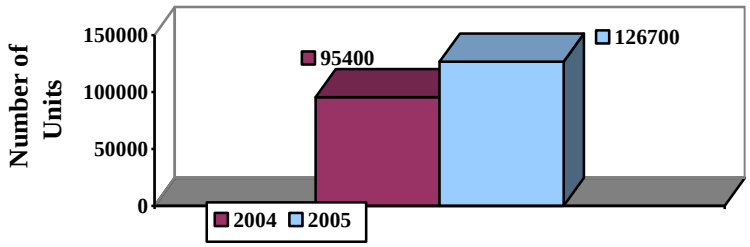


Figure 2. The number of units of industrial robots sold at the end of year 2004 and 2005

The industrial robots have captured market worldwide gradually, since 2005, with average growth of 13.65% [Figure 1b.]. 2009 is the exception, because the lowest sales are estimated at the end of year. Year 2011 is the 'golden year' for the industrial robotics.

2.2. Analysis of market for service robots

Service robots turn up in either of two key areas: the professional or home/domestic domain. Service robots for professional use are extremely diverse, since they are usually designed to perform a specific task. Professional service robots also can be found in public buildings for delivering goods, executing tasks in dangerous or hazardous environments, or even helping in cowsheds for automatic milking. The personal service robots have influenced the human life by doing domestic tasks, providing security & surveillance and facilitating through transportation, education & entertainment. The financial opportunities in these markets have already exceeded several billion dollars in the early decade of 21st century, and for components within them e.g., machine vision and mobile platforms.

2.2.1. Market trends of service robots for professional use.

The major application areas for professional robots are underwater robots, defence robots, medical robots, demolition robots, mobile robot platforms for multiple use, laboratory robots, and cleaning robots. Since 1998, the total no. of professional robots installed for these applications is amounted to 110,000 units worldwide in the current decade[2].

The Table 2. provides following information related to professional robots installed worldwide in early 21st century:

1) The average increase in the no. of units installed during 2000-2011 is about 872 units.
2) The no. of units installed is reached to 10,000 units first time in 2008.
3) The average growth of 28.97% in 2000-2011.

The decline in the sales of professional robots is seen in the year 2001, 2003 and 2009[Figure 3.]. Apart from this the demand of the professional robots increased worldwide.

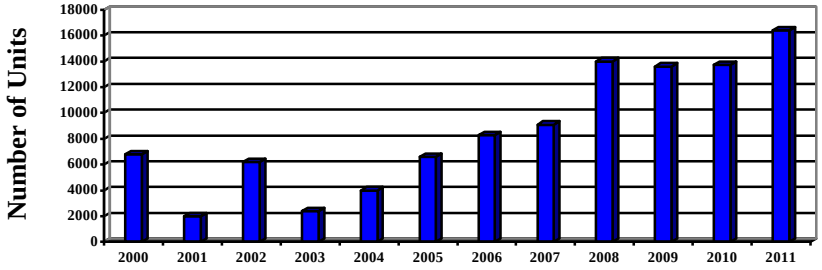


Figure 3. The number of units of professional service robots installed at the end of year during 2000-2011

Table 2. The number of units of professional service robots sold worldwide at and up to the end of the year during 2000-2011[2,3,4,5,6,7]

S. No.	At the End of Year	No. of Units Sold	Up to the end of year
1.	2000	6,815	10,415
2.	2001	1,985	12,400
3.	2002	6,200	18,600
4.	2003	2,400	21,000
5.	2004	4,000	25,000
6.	2005	6,600	31,600
7.	2006	8,300	39,900
8.	2007	9,100	49,000
9.	2008	14,000	63,000
10.	2009	13,600	76,600
11.	2010	13,741	90,341
12.	2011	16,408	106,749

2.2.2. Market trends of service robots for personal and private use. Service robots for personal and private use are mainly in the areas of domestic (household) robots, which include vacuum cleaning and lawn mowing robots, and entertainment robots, including toy and hobby robots. Vacuum cleaning robots were introduced on the market at end of 2001[7]. The total 13.4 million units were sold up to 2011 for personal and private usage of robots.

On the basis of the data given in the Table 3., the following information is complied:

1) The average no. of units sold during 2000-2011 estimated is about 217,318 units.
2) The no. of units sold is reached to 1 million first time in the year 2007.
3) The no. of units sold crosses the 2 million in 2010.
4) The average growth of 74.15% in 2000-2011.

The service robots for personal and private use are growing since 2000, except the years 2001, 2006, 2008 and 2009[Figure 4.].

Table 3. The number of units of personal & private service robots sold worldwide at and up to the end of the year during 2000-2011[2,3,4,5,6,7]

S. No.	At the End of Year	No. of Units Sold	Up to the end of year
1.	2000	109,500	112,500
2.	2001	64,000	176,500
3.	2002	427,500	604,000
4.	2003	706,000	1,310,000
5.	2004	790,000	2,100,000
6.	2005	800,000	2,900,000
7.	2006	640,000	3,540,000
8.	2007	1,860,000	5,400,000
9.	2008	1,800,000	7,200,000
10.	2009	1,500,000	8,700,000
11.	2010	2,200,000	10,900,000
12.	2011	2,500,000	13,400,000

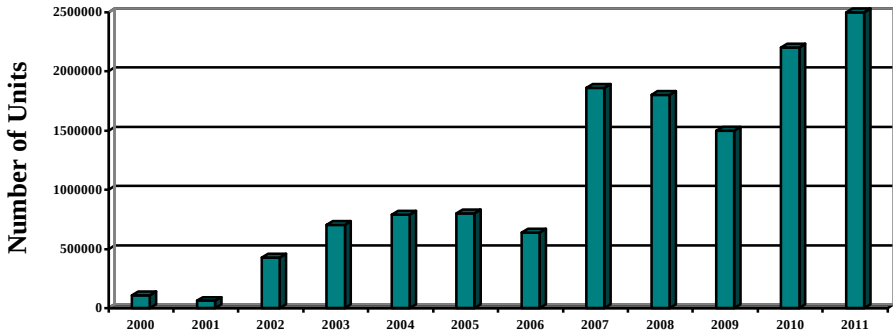


Figure 4. The number of units of personal & private robots sold at the end of year during 2000-2011

2.3. Growth in robotics industry

In early 21st century, the demand for both industrial and service robots is increasing, which in turn, boosted sales of both in order to enhance the growth of robotics industry. Since 1961, The total 15,790,979 units were sold worldwide up to the end of 2011, among these 14,734,089 units were sold in 2000-2011, for the applications area ranging from automotive to non automotive and from defense to domestic[Figure 5.]. This industry is intensified at the average of 46.50% during 2000-2011.

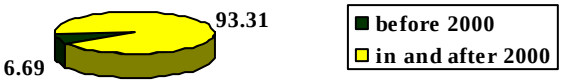


Figure 5. The number of units (%) of robots sold before 2000 and in & after 2000

3. Comparison of market trends of industrial and service robots

Both types of robots gave contribution in order to increase the market for robotics industrin 21st century. But the way they are contributing is different for these robots depending on the application area, they are being used. Following six comparisons have been identified for market trends of industrial and service robots worldwide in 2000-2011:

1) Since 1961, the total 2,284,230 units of industrial robots sold, while the total 13,506,749 units of service robots(including robots for professional use and personal & private use) sold, since 1998.
2) In early 21st century, the average increase in the sales of industrial robots is 36.39%, while it is 51.56% for service robots.
3) The sales of the industrial robots never reached to 1 million units up to the end of 2011, whereas the sales of service robots reached to same at the end of 2003[Figure 6.].
4) In 2009, the sales of industrial robots declined to 60,000 units by 47.04%, while for service robots it is declined to 1,513,600 units by 19.53%, as compared to 2008[Figure 6.]
5) The average increase in the no. of units of industrial robots is near to professional robots and very far from the personal & private robots[Figure 7.].

6) Since 1961, the contribution of industrial robots in the robotics industry is estimated to only 14.47% of the total units available worldwide [Figure 8.].

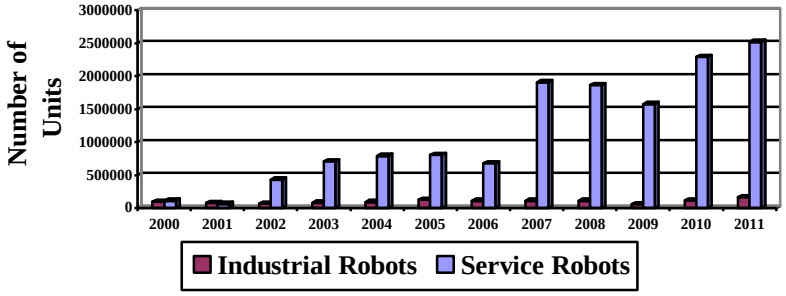


Figure 6. The number of units of industrial and service robots sold at the end of year during 2000-2011

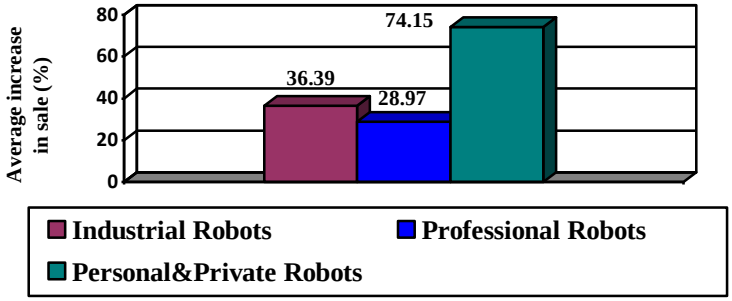


Figure 7. The average increase in sales(%) of industrial and service robots during 2000-2011

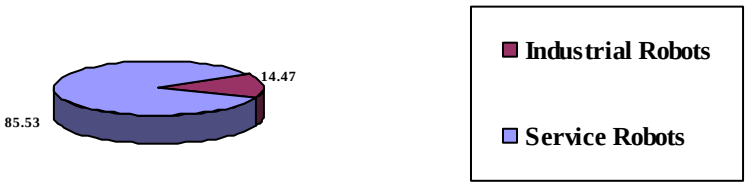


Figure 8. The contribution(% of total units) of industrial and service robots in robotics industry up to the end of 2011

4. Conclusion

The robotics industry has magnificent growth worldwide in early 21st century. Both industrial and service robots contributed in the growth. But the share of service robots is larger, among them personal robots added millions of units. So, it is expected that later in 21st century robots will be known much for household applications. Entry points into these markets worldwide are accelerating from a variety of corners and initiatives in industry, government, academia, large corporations and small startups. The technological changes and emptiness in our society increases the demand of robots in all types of industries. In order to continue this growth the robots require extensive improvement in technology.

References

[1] Dr. Prahlad Vadakkepat. Robotics, India and the Future. Digit Magazine, India, Jan 2010.
[2] World Robotics 2012. Taipei. 30 August 2012.
[3] World Robotics 2011. Frankfurt.1 September 2011.
[4] World Robotics 2010. United Nations, New York and Geneva. 2010.
[5] World Robotics 2008. United Nations, New York and Geneva. 2008.
[6] World Robotics 2006. United Nations, New York and Geneva. 2006.
[7] World Robotics 2002. United Nations, New York and Geneva. 2002.
[8] World Robotics 2001. United Nations, New York and Geneva. 2001.

FPGA Implementation and Functional Verification of a Pipelined MIPS Processor

[1] Balaji valli, [2] A. Uday Kumar, [3] B.Vijay Bhaskar
[1] PG Scholor, [2] Asst. Professors, [3]Assoc. Professor,
[1, 2, 3]Avanthi's St. Theressa Institute of Engineering & Technology, Garividi ,A.P.

Abstract:
This project targets the implementation design of a pipelined MIPS RISC Processor using VHDL (Very high speed integrated circuit Hardware Description Language). In this paper MIPS instruction format, instruction data path, decoder modules are analyzed. Furthermore, instruction fetch (IF) module of a CPU is designed based on RISC CPU instruction set. Function of IF module mainly includes fetch instruction and latch module address arithmetic module check validity of instruction module synchronous control module.

Keywords: MIPS, RISC, CPU, VHDL, FPGA, ID, IF, EX, MEM

1. Introduction

Pipeline is one of the basic techniques to improve the CPU's performance. This paper based on MIPS instruction set, designed a five-stage pipeline CPU. MIPS processor is basically a RISC micro processor. RISC CPU has extensive use in embedded system. Developing CPU with RISC structure is necessary choice. The MIPS Architecture defines thirty-two, 32-bit general purpose registers (GPRs). Instruction Set Architecture (ISA) of processor is composed of instruction set and corresponding registers. Program based on same ISA can run on the same instruction set. MIPS instruction has been developed from 32-bit MIPSI to 64-bit MIPSIII and MIPSIV since it was created. To assure downward compatibility, every generation production of MIPS instruction directly extends new instruction based on old instruction but not abnegates any old instruction, so MIPS processor of 64-bit instruction set can execute 32-bit instruction.

2. Implementation of Pipelined MIPS Processor

There are three formats of MIPS instructions:

(i) Register Format (R-type)

OPCODE (31 to 26)	RS (25to21)	RT (20to16)	RD (15to11)	Shift (10to6)	FUN (5to 0)

The first two 5-bit register specifications are the two read registers and the last 5-bit register specification is the destination register, that the instruction will use. The last 6-bits in the instruction are used as function bits.

(ii) Immediate Format (I-type)

OPCODE (31 to 26)	RS (25to21)	RT (20to16)	RD (15to11)	Shift (10to6)	FUN (5to 0)

The I-type is similar to the R-type except the second read register and the 5 function bits are replaced by a 16-bit immediate value. Each I-type opcode can have only one instruction because it has no function bits like the R-type.

(iii) Jump Type Format (J- type)

The J-type format consists of a 6-bit opcode and remaining bit indicates branching address.

OPCODE (31 to 26)	RS (25to21)	RT (20to16)	RD (15to11)	Shift (10to6)	FUN (5to 0)

3. Architecture

The MIPS pipelined processor involves five steps, the division of an instruction into five stages implies a five-stage pipeline:

1. Instruction Fetch (IF): fetching the instruction from the memory
2. Instruction Decode (ID): reading the registers and decoding the instruction

3. Execution (EX): executing an operation or calculating an address
4. Data Memory (MEM): accessing the data memory
5. Write Back (WB): writing the result into a register.

The key to pipelining the single-cycle implementation of the MIPS processor is the introduction of pipeline registers that are used to separate the data path into the five sections IF, ID, EX, MEM and WB. Pipeline registers are used to store the values used by an instruction as it proceeds through the subsequent stages. The MIPS pipelined registers are labeled according to the stages they separate. (e.g. IF/ID, ID/EX, EX/MEM, MEM/WB)

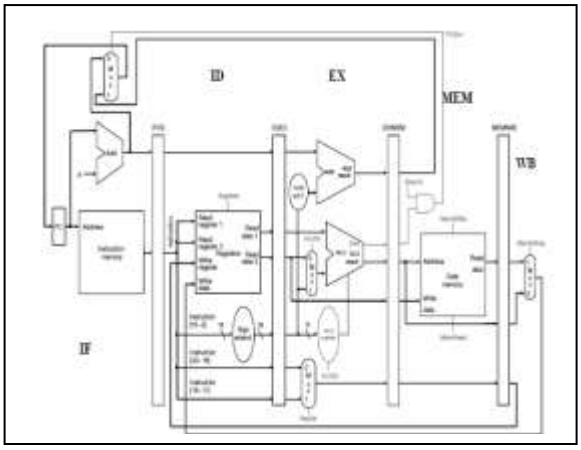


Fig.1. Pipelined MIPS Processor

To implement the MIPS pipelined processor, pipeline registers are placed into the corresponding VHDL modules that generate the input to the particular pipeline register. For example, the Instruction Fetch component will generate the 32-bit instruction and the PC+4 value and store them into the IF/ID pipeline register. When that instruction moves to the Instruction Decode stages it extracts those saved values from the IF/ID pipeline register.

4. RESULTS

Implementation is done using XILINX 10.1. RTL schematic and Floor plan view are shown in Fig.3 and Fig.4. Simulation snap is shown in Fig.2.

Fig.2. Functional simulation

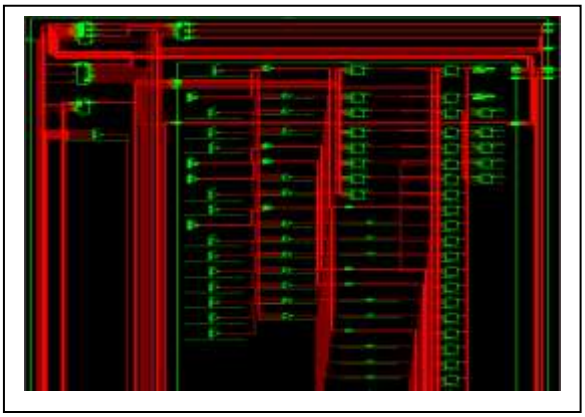


Fig.3. RTL Schematic

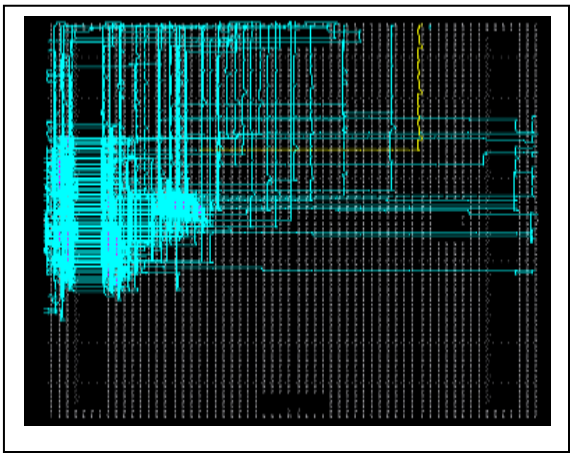


Fig.4. Floor Plan View

Hardware utilization summery is reported in Table. 1.

S.No.	Logic Utilization	Utilized Hardware
1.	Number of Slices	354
2.	Number of Slice Flip Flops	256
3.	Number of 4input LUTs	458
4.	Number of bonded IOBs	85

Table 1. Device Utilization Summery

5. Conclusion

In this paper, top-down design method adopted and VHDL is used for the implementation. Design is implemented on XILINX 11.1. Data Memory and Instruction Memory are generated using XILINX COREGEN.

REFERENCES

[1] Hui Ma, Dinglei Wang, The Design of Five-stage Pipeline CPU Based on MIPS, 978-1-4244-8165-1/11, 2011 IEEE.
[2] Mrs. Rupali S. Balpande, Mrs.Rashmi S. Keote, Design of FPGA based Instruction Fetch & Decode Module of 32-bit RISC (MIPS) Processor, 2011 International Conference on Communication Systems and Network Technologies, 978-0-7695-4437-3/11, 2011 IEEE,
[3] Wang-Yuan Zhen, IBM-PC Macro Asm Program, Huazhong University of Science and Technology Press, 1996.9.
[4] MIPS Technologies, Inc. MIPS32™ Architecture For Programmers Volume II: The MIPS32™ Instruction Set June 9, 2003.
[5] Zheng-WeiMin, Tang-ZhiZhong. Computer System Structure (The second edition), Tsinghua University Press, 2006.
[6] Pan-Song, Huang-JiYe, SOPC Technology Utility Tutorial, Tsinghua University Press, 2006.
[7] MIPS32 4KTMProcessor Core Family Software User's Manual, MIPS Technologies Inc. [M]. BeiJing, Engine Industry Press. 2003.
[8] Yi-Kui, Ding-YueHua, Application of AMCCS5933 Controller in PCI BUS, DCABES2007, 2007.7.759
[9] "Rapid Prototyping of digital Systems, a tutorial approach "By – James O. Hamblen and Michael D.Furman.
[10] Charles E. Gimarc, Veljko M. Mhtinovic, "RISC Principles, Architecture, and Design", Computer Science Press Inc., 1989.
[11] Whte paper, "Wide range of comprehensive tools speed, Development of high-pafomance embedded system", MIPS Technologies Inc.,hnp:/lwww.mips.com/whitepapers/030399Wl .hml (current Sep. 30,2002).
[12] S. Chen, B. Mulgrew, and P. M. Gran< "A clustering technique for digital communications channel equalization using radial basis function networks," *EEE Trans. Neuml Nerworkr*, vol. 4, pp. 570-578, July 1993.

Extraction of Edge Detection Using Digital Image Processing Techniques

[1]M. Kalpana, [2]G. Kishorebabu, [3]K.Sujatha
[1,2]Asst.professor, S.S.I.E.T ,
[3]Asst.professor, R.V.R.I.E.T

Abstract
Digital image Processing is one of the basic and important tool in the image processing and computer vision. In this paper we discuss about the extraction of a digital image edge using different digital image processing techniques. Edge detection is the most common technique for detecting discontinuities in intensity values. The input image or actual image have some noise that may cause the of quality of the digital image. Firstly, wavelet transform is used to remove noises from the image collected. Secondly, some edge detection operators such as Differential edge detection, Log edge detection, canny edge detection and Binary morphology are analyzed. And then according to the simulation results, the advantages and disadvantages of these edge detection operators are compared. It is shown that the Binary morphology operator can obtain better edge feature. Finally, in order to gain clear and integral image profile, the method of ordering closed is given. After experimentation, edge detection method proposed in this paper is feasible.

Index Terms: Image, Digital image, Edge, boundary, Edge detection, wavelet denoising, differential operators, and binary morphology.

I. Introduction
An edge in a digital image is a boundary or contour at which a significant change occurs in some physical aspect of an image, such as the surface reflectance, illumination or the distances of the visible surfaces from the viewer. Changes in physical aspects manifest themselves in a variety of ways, including changes in color, intensity and Texture. Edge always indwells in two neighboring areas having different grey level. It is the result of grey level being discontinuous. Edge detection is a kind of method of image segmentation based on range non-continuity. Image edge detection is one of the basal contents in the image processing and analysis, and also is a kind of issues which are unable to be resolved completely so far [1]. When image is acquired, the factors such as the projection, mix, aberrance and noise are produced. These factors bring on image feature is blur and distortion, consequently it is very difficult to extract image feature. Moreover, due to such factors it is also difficult to detect edge. The method of image edge and outline characteristic's detection and extraction has been research hot in the domain of image processing and analysis technique.Detecting edges is very useful in a number of contexts. For example in a typical image understanding task such as object identification, an essential step is to an image into different regions corresponded to different objects in the scene. Edge detection is the first step in the image segmentation. Edge feature extraction has been applied in many areas widely. This paper mainly discusses about advantages and disadvantages of several edge detection operators applied in the cable insulation parameter measurement. In order to gain more legible image outline, firstly the acquired image is filtered and denoised. In the process of denoising, wavelet transformation is used. And then different operators are applied to detect edge including Differential operator, Log operator, Canny operator and Binary morphology operator. Finally the edge pixels of image are connected using the method of bordering closed. Then a clear and complete image outline will be obtained

Ii. Image Denoising
As we all know, the actual gathered images contain noises in the process of formation, transmission, reception and processing. Noises deteriorate the quality of the image. They make image blur. And many important features are covered up. This brings lots of difficulties to the analysis. Therefore, the main purpose is to remove noises of the image in the stage of pretreatment. The traditional denoising method is the use of a low-pass or band-pass filter to denoise. Its shortcoming is that the signal is blurred when noises are removed. There is irreconcilable contradiction between removing noise and edge maintenance.Yet wavelet analysis has been proved to be a powerful tool for image processing [2]. Because Wavelet denoising uses a different frequency band-pass filters on the signal filtering. It removes the coefficients of some scales which mainly reflect the noise frequency. Then the coefficient of every remaining scale is integrated for inverse transform, so that noise can be suppressed well. So wavelet analysis widely used in many aspects such as image compression, image denoising [3][4], etc

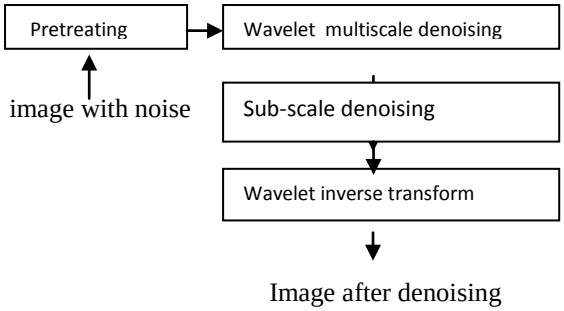


fig1: Sketch of removing image noises with wavelet transformation

The basic process of denoising making use of wavelet transform is shown in Fig1, its main steps are [3][4] as follows:

1) Image is preprocessed (such as the gray-scaleadjustment, etc.).
2) Wavelet multi-scale decomposition is adapted to process image.
3) In each scale, wavelet coefficients belonging to noises are removed and the wavelet coefficients are remained and enhanced.
4) The enhanced image after denoising is gained using wavelet inverse transform.

The common used operators are the Differential, Log, Canny operators and Binary morphology, etc. The simulation effect of wavelet denoising through Matlab is shown in Fig. 2.

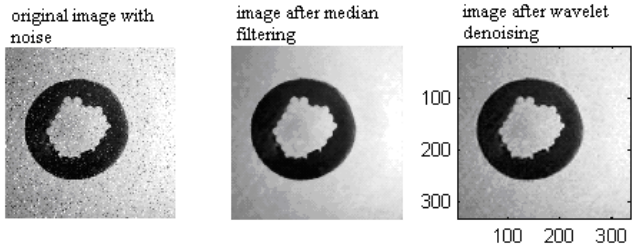


Fig.2.Comparison of denoising methods

Comparing with the traditional matched filter, the high-frequency components of image may not be destroyed using wavelet transform to denoise. In addition, there are Many advantages such as the strong adaptive ability, calculating quickly, completely reconstructed, etc. So the signal to noise ratio of image can be improved effectively making use of wavelet transform.

III. EDGE DETECTION

The edge detection of digital image is quite important foundation in the field of image analysis including image division, identification of objective region and pick-up of region shape and so on. Edge detection is very important in the digital image processing, because the edge is boundary of the target and the background. And only when obtaining the edge we can differentiate the target and the background.

The basic idea of image detection is to outstand partial edge of the image making use of edge enhancement operator firstly. Then we define the 'edge intensity' of pixels and extract the set of edge points through setting threshold. But the borderline detected may produce interruption as a result of existing noise and image dark. Thus edge detection contains the following two parts:

1) Using edge operators the edge points set are extracted.
2) Some edge points in the edge points set are removed and a number of edge points are filled in the edge points set. Then the obtained edge points are connected to be a line

A. Differential operator

Differential operator can outstand grey change. There are some points where grey change is bigger. And the value calculated in those points is higher applying derivative operator. So these differential values may be regarded as relevant 'edge intensity' and gather the points set of the edge through setting thresholds for these differential values. First derivative is the simplest differential coefficient. Suppose that the image is $f(x, y)$, and its derivative operator is the first order partial derivative $\partial f/\partial x$,$\partial f/\partial y$. They represent the rate-of-change that the gray f is in the direction of x and y . Yet the gray rate of change in the direction of α is shown in the equation (1):

$$\frac{\partial f}{\partial \alpha} = \frac{\partial f}{\partial x}\cos\alpha + \frac{\partial f}{\partial y}\sin\alpha \qquad (1)$$

Under consecutive circumstances, the differential of the function is:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy.$$

The direction derivative of function f (x, y) has maximum at a certain point. And the

direction of this point is arc tan $[\frac{\partial f}{\partial y}/\frac{\partial f}{\partial x}]$. The maximum of direction derivative is $\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$. The vector with this direction and modulus is called as the gradient of the function f , that is,$\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$.So the gradient modulus operator is designed in the equation (2).

$$G[f(x,y)] = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \qquad (2)$$

For the digital image, the gradient template operator is designed as:

$$\sqrt{(\Delta_x f(i,j)^2 + \Delta_y f(i,j)^2} \qquad (3)$$

$$\Delta_x f(i,j) = f(i,j) - f(i-1,j),$$
$$\Delta_y f(i,j) = f(i,j) - f(i,j-1),$$

Differential operator mostly includes oberts operator and Sobel operator.

(1) Roberts operator:

Roberts operator is a kind of the most simple operator which makes use of partial difference operator to look for edge. Its effect is the best for the image with steep low noise. But the borderline of the extracted image is quite thick using the Roberts operator, so the edge location is not very accurate.Roberts operator is defined as:

$$G[f(x,y)] = (\{[f(x+1,y+1) - f(x,y)^2] + [f(x+1,y) - f(x,y+1)^2\}^{1/2} \quad (4)$$

But absolute deviation algorithm is usually used to predigest the equation (4) in practice. The following equations (5) and (6)are the process of reduction.

$$G[f(x,y)] \approx |f(x+1,y) - f(x,y)| + |f(x,y+1) - f(x,y)| \quad (5)$$

$$G[f(x,y)] \approx |f(x+1,y+1) - f(x,y)| + |f(x,y+1) - f(x+1,y)| \quad (6)$$

The template of the Roberts operator is shown in Fig. 3.

1	0
0	-1

0	1
- 1	0

Fig.3.Roberts Operator

(2) Sobel and Prewitt operator

To reduce the influence of noise when detecting edge, the Prewitt operator enlarges edge detection operator template from two by two to three by three to compute difference operator. Using the Prewitt operator can not only detect edge points, but also restrain the noise. The Sobel operator counts difference using weighted for 4 neighborhoods on the basis of the Prewitt operator. The Sobel operator has the similar function as the Prewitt operator, but the edge detected by the Sobel operator is wider. Suppose that the pixel number in the 3×3 sub-domain of image is as follows:

A_0	A_1	A_2
A_7	$f(x, y)$	A_3
A_6	A_5	A_4

We order that

$$X = (A_0+2A_1+A_2) - (A_6+2A_5+A_4)$$

And $Y = (A_0+2A_7+A_6) - (A_2+2A_3+A_4)$.

Then Prewitt operator is as follows:

$$G[f(i,j)] = (\sqrt{X^2+Y^2}) \quad (7)$$

$$G[f(i,j)] = /X/+/Y/ \quad (8)$$

Prewitt operator is said in Fig.4 in the form of the template.

1	-1	1
1	-1	0
1	-1	-1

1	1	1
0	0	0
-1	-1	-1

Fig.4.prewitt operator

Sobel operator can process those images with lots of noises and gray gradient well. We order that

$X = (A_0+2A_1+A_2) - (A_6+2A_5+A_4)$
And $Y = (A_0+2A_7+A_6)-(A_2+2A_3+A_4)$.

Then Sobel operator is as follows:

$$G[f(i,j)] = (\sqrt{X^2+Y^2}) \quad (9)$$

$$G[f(i,j)] = /X/+/Y/ \quad (10)$$

The templates of sobel operator is shown in fig.5

1	2	1
0	0	0
-1	-2	- 1

1	0	-1
2	0	-2
1	0	-1

Fig.5.Sobel Operator

The original image of cable insulation layer and the edge detection drawing of Sobel operator gained using Mat Lab simulation are shown in Fig. 6 and Fig.7

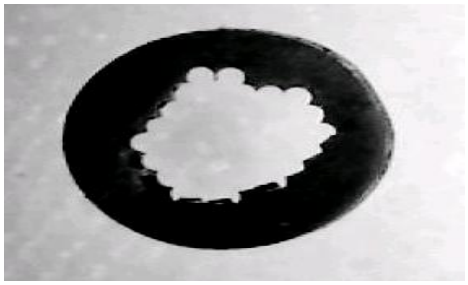


Fig.6. Original image

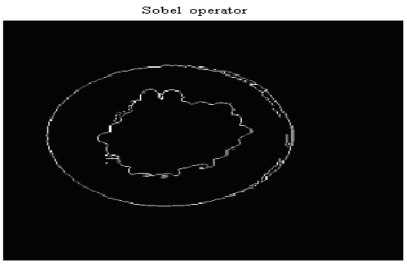


Fig. 7 the edge detection drawing of Sobel operator

From the simulation drawing Fig. we can know that the edge position is very accurate. And the effect of Sobel edge detection is very satisfying. In a word, the Sobel and Prewitt operators have a better effect for such images with grey level changing gradually and more noises.

B. Log operator

The Log operator is a linear and time-invariant operator. It detects edge points through searching for spots which two-order differential coefficient is zero in the image grey levels. For a continuous function $f(x, y)$, the Log operator is defined as at point (x, y) :

$$\Delta^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \qquad (11)$$

The Log operator is the process of filtering and counting differential coefficient for the image. It determines the zero overlapping position of filter output using convolution of revolving symmetrical Log template and the image. The Log operator's template is shown in Fig. 8

-1	- 1	-1
-1	8	-1
-1	-1	-1

0	- 1	0
-1	4	-1
0	-1	-1

Fig. 8 Log operator

In the detection process of the Log operator, we firstly pre-smooth the image with Gauss low-pass filter, and then find the steep edge in the image making use of the Log operator. Finally we carry on binarization with zero grey level to give birth to closed, connected outline and eliminate all internal spots. But double pixels boundary usually appears using the Log operator to detect edge, and the operator is very sensitive to noise. So the Log - operator is often employed to judge that edge pixels lie in either bright section or dark section of the image.

C. Canny operator

The Canny operator is a sort of new edge detection operator. It has good performance of detecting edge, which has a wide application. The Canny operator edge detection is to search for the partial maximum value of image gradient. The gradient is counted by the derivative of Gauss filter. The Canny operator uses two thresholds to detect strong edge and weak edge respectively. And only when strong edge is connected with weak edge, weak edge will be contained in the output value. The theory basis of canny operator is shown in equations (12)-(15)

Gauss:

$$G(x, y) = \exp[-(x^2 + y^2)/2\sigma^2] \qquad (12)$$

Edge normal

$$n1 = \nabla(g*p)/|\nabla(g*p)| \qquad (13)$$

Edge strengths

$$Gn\ P = \frac{\partial}{\partial n}[g * p] \qquad 14)$$

Maximal strengths:

$$0 = \frac{\partial}{\partial n} GnP = \frac{\partial}{\partial n^2}^2 [g * p] \qquad (15)$$

D. Binary morphology:

Mathematical morphology is a new method applied in image processing. The basic idea is to measure and extract the corresponding shape from image with structural elements having stated form. So that the image processing and analyzing can be completed [2].Using mathematical morphology to detect the edge is better than using differential treatment. Because it is not sensitive to noise, and the edge extracted is relatively smooth. Binary image is also known as black-and-white image. The object can be easily identified from the image background. So we adopt the combination of binary image and mathematical morphology to detect edge. It is called Binary morphology. Suppose that the region is shown in form of the set A. Its border is $\beta(A)$. B is an appropriate structure element, and it is symmetrical around the origin. Firstly we corrupt A with Recorded as

$A\Theta B = \{x|(B)_x \subseteq A\}$,

Where $(B)_x$ is a translation B along the vector. The interior of region is available with $A\Theta B$. And $A - (A\Theta B)$ is the borderline naturally. Then $\beta(A)$ is obtained. The equation of edge extraction can be said,

$\beta(A) = A - (A\Theta B)$.

Structuring element is larger, the edge gained will be wider.

E. Simulative results analysis

In order to know about the advantages and disadvantages of these edge detection operators, we detect edge using these different operators respectively. The simulation results are shown in Fig. 9 and Fig. 10.

Orirnal Image **Binary image** **Extracted edge**

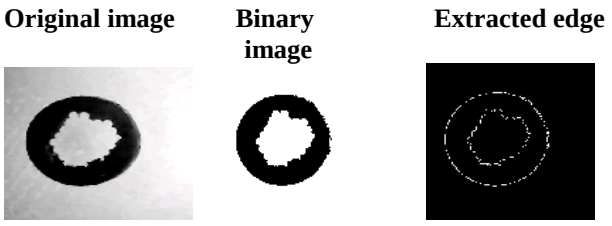
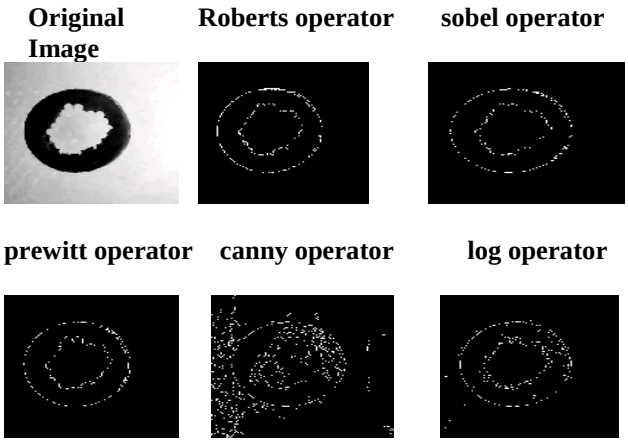


Fig.9.Detecting edge with binary morplology



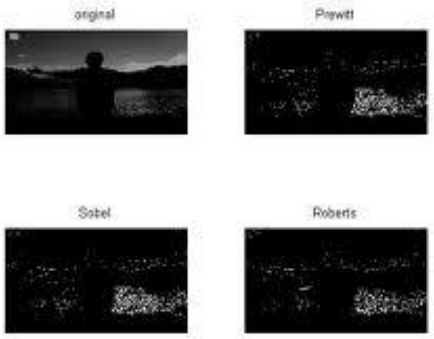


Fig. 10 several edge detection algorithm comparison

From the simulation results we can conclude that: the effect of detecting edge with sobel operator after wavelet de-noising and with Binary morphology directly is better. So these two methods can be used. But finally we choose Binary morphology method based on specific measurement errors.

IV. BORDERLINE CLOSED

Although image is denoised before detecting edge, yet noises are still introduced when detecting edge. When noise exists, the borderline, which is obtained using derivative algorithm to detect image, usually produces the phenomenon of break. Under this situation, we need to connect the edge pixels. Thus, we will introduce a kind of method closing the borderline with the magnitude and direction of pixels gradient. The basis of connecting edge pixels is that they have definite similarity. Two aspects' one using gradient algorithm to process image. One is the magnitude of gradient; the other is direction of gradient. According to edge pixels gradient's similarity on these two aspects, the edge pixels can be connected. Specific speaking, if Pixel(s,t) is in neighbor region of the pixel (x,y) and their gradient magnitudes and gradient directions must satisfy two conditions (16) and (17) respectively, then the pixels in (s,t) and the pixels in (x,y) can be connected. The closed boundary will be obtained if all pixels are judged and connected

$$|\nabla f(x,y) - \nabla f(s,t)| \leq T \qquad (16)$$

$$\varphi(x,y)-\varphi(x,t)|\leq A \qquad (17)$$

Where T is magnitude threshold, A is angle threshold.

V. CONCLUSION

These edge detection operators can have better edge effect under the circumstances of obvious edge and low noise. But the actual collected image has lots of noises. So many noises may be considered as edge to be detected. In order to solve the problem, wavelet transformation is used to denoising the paper. Yet its effect will be better if those simulation images processed above are again processed through edge thinning and tracking.

Although there are various edge detection methods in the domain of image edge detection, certain disadvantages always exist. For example, restraining noise and keeping detail can't achieve optimal effect simultaneously. Hence we will acquire satisfactory result if choosing suitable edge detection operator according to specific situation in practice.

References:

[1] Lei Lizhen, Discussion of digital image edge detection method, Mapping aviso, 2006, 3:40-42.

[2] Lai Zhiguo, etc, "Image processing and analysis based on MATLAB" , Beijing: Defense Industry Publication, 2007,4.

[3] Ma Yan, and Zhang Zhihui, Several edge detection operators comparation, Industry and mining automation, 2004, (1): 54-56.

[4] Gao Cheng, and Lai Zhiguoetc, Image analysis and application based on MATLAB, Beijing: Publishing House of National defence industry, 2007, 4: 133-175.

[5] Wang Zhengyao, Edge detection of digital image[Master paper], Xi'an: Xi'an Jiaotong University, 2003.

[6] Heung-Soo Kim and Jong-Hwan Kim. A two-step detection algorithm from the intersecting chords. Pattern Recognition Letters. 2001, 22:787-798.

[7] Canny J F. A computatioal approach tO edge detection[J]. IEEE Trans on PAMI, 1985, 8(6): 679-698.

Estimation and Mapping of Land Surface Temperature From AATSR Images And GIS: A Case Study In Kolondieba-Tiendaga Basin In Sudano-Sahelian Climate, Mali

Daou I[1]., Mariko A[2]., Rasmus F[3]., Menenti M[4]., Kourosh K[5]., Maïga H B[6]., Maïga S.M[1].,

[1]. University of Bamako, Faculty of Science and Techniques (FAST), Mali
[2]. National Engineering school-Aberhamane Baba Touré of Bamako, Mali
[3]. University of Geography and Geology of Copenhague, Denmark
[4]. Delft University of Technology, Department and flesh, Remote Sensing optical and laser Remote sensing, Netherlands.
[5]. University of Cheickh Anta Diop (UCAD), Laboratory of Education and Research in Geomatics

Abstract:
The knowledge of the various terms which intervene in the energy balance of surface is essential at the same time for the hydrologists, the amenagists, the agronomists, the meteorologists, and the climatologists. The land surface temperature of the making ground left the key parameters of this assessment, plays a fundamental part in the processes of interaction hydrosphere-biosphere, atmosphere. It represents, the temperature of canopy of the vegetation for the covered zones of vegetation, and the temperature of canopy of the vegetation plus that of the surface of the ground, for the zones of scattered vegetation. While it represents the temperature of surface of the ground for the zones bar soil. This land surface temperature can be collected through two approaches: conventional approaches, and them approaches of remote sensingThe conventional methods make it possible to collect the temperature of surface of the ground starting from measurements of weather stations. On the other hand, the approaches of remote sensing make it possible, as for them, to estimate the land surface through the use of model of energy balance of surface. In this study, model SEBS was used on satellite data AATSR to estimate and mapping the land surface temperature, on the catchment area of Kolondièba-Tiendaga, in southern Mali zone. These results show values of land surface temperature between 303 and 296 (°K) for standard deviations of 2, 66 and 0, 945. These results are similar to those already found elsewhere, in West Africa with the same types given of satellite images AASTR

Keywords: Land surface Temperature, model SEBS, Mapping, watershed of Kolondieba-Tiendaga, Mali

1. Introduction

The knowledge of the space-time variation of the land surface temperature intervening in the energy assessment of surface is key information in the energy exchanges and matter between the surface of the ground and the atmosphere [1, 2, 3, 4, 6, 12, 17, 36,]. Indeed, the land surface temperature is an important parameter in the energy assessment of surface and plays an essential part in the processes of interaction between the hydrosphere-Biosphere-Atmosphere and the cycle of water. In addition, the follow-up of the transfers of mass and energy to the level of a surface is dominating for the integrated management of the water resources and vegetable. It is also necessary for the good comprehension of the hydrological systems, the ecological and climatic ecosystems, like with the follow-up and the forecast of their evolutions [7, 8,9,10, 11, 13, 14].Information on the land surface temperature represents for the bar soil the temperature of surface of the ground, while it represents for the dense zones of vegetations, the temperature of surface of canopy of the vegetation[4]. However, for the zones of scattered vegetation, it determines the temperature of canopy of the vegetation, and ground surfaces it [4,17,18,19] . Traditionally, the data of land surface temperature are collected starting from the weather stations.Today, with the development of technologies of observation of the ground, the remote sensing and the GIS appears like a tool privileged for the collection and the follow-up of parameters biophysics, in particular the land surface temperature, the evapotranspiration, the albedo, because providing related information's to the mass transfers, and in particular to the processes of evapotranspiration. Several research related to the problems of measurement and collection of these parameters through the use of the technology of the thermal infra-red remote sensing [28, 29, 37, 20, 21, 22, 24]. Thus, the land surface temperature collected starting from the satellite data can be used for the evaluation of the evapotranspiration, the water stress of the vegetation and the requirement out of water for the annual cultures through the use of model of energy assessments of surface. With the international scales, these tools for remote sensing and GIS are requested more and more for the collection and the cartography of the parameters biophysics, particularly, the albedo, emissivity, the temperature of surface, the evapotranspiration, and the evaporative fraction, etcOur study proposes to measure and chart the temperature of surface of the ground starting from satellite images AATSR of platform ENVISAT of the European space agency (ESA) by using model SEBS (Surface Energy Balance System) developed

by [16, 19, 20, 29, 30, 31] on the catchment area of Kolondièba-Tiendaga, in soudano-sahelian medium, in southern the Mali zone.

2. Materials and Method

2.1 Study area

The study relates to the catchment area (Kolondièba-Tiendaga) located in southern Mali zone between Longitudes and Latitudes 34 °W and 6, 82°W and 10,15°N and 11,08°N. This basin of a surface of 3050 km ², is under basin of the transborder basin of the Outlaw (Fig.1). It is localized on the upstream reservoir of Niger, to approximately 262 km of Bamako (capital of Mali), and is limited to North by the town of Kolondieba, in the South by the commune of Sibirila to the border of the Ivory Coast, in the East by the commune of Fakola and Bougoula, in the West by the commune of Garalo (circle of Bougouni). Its discharge system is in Tiendaga on the road of Fakola. Its physical Characteristics are given in Tableau1. The catchment area of Kolondièba-Tiendaga has a tropical climate of transition or Sudanese Southern climate, characterized by the one hot season alternation and dries (November-April) and one rain season (May-October) where the annual pluviometric average is higher 1000mm/an. The monthly average temperatures vary between 38° in May and 23° in December. The values of the relative humidity oscillate between 82% in August and 38% in February. The radiation is strong during all the year, where the values of the potential evapotranspiration remain important, with a maximum in May (161) (DNH, Mali, 1990). The geological substratum is consisted metamorphic and granitic rocks covered with washed tropical ferruginous grounds, hydromorphic grounds ferralitic, and grounds. The relief is dominated by plates upstream of the basin, plains in the center and hollows downstream whose altitude can vary between 315 and 390m. the vegetation is characterized by savannas raised and shrubby of herbaceous, mixed with annual cultures.

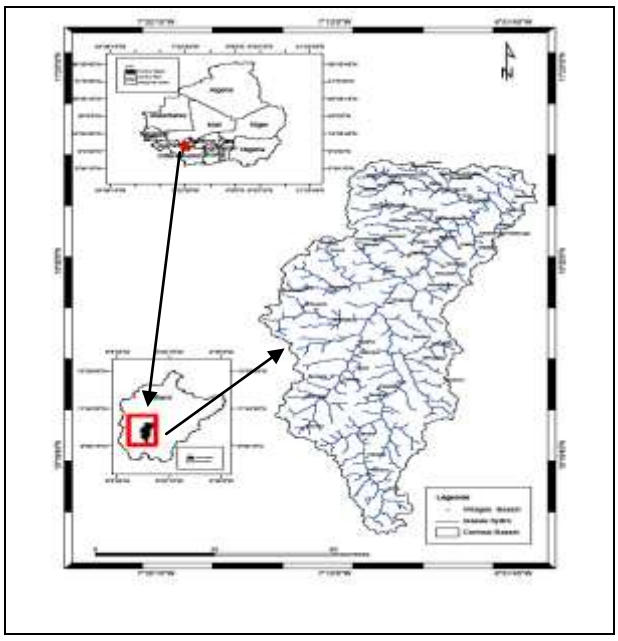


Figure 1: Presentation of the study area

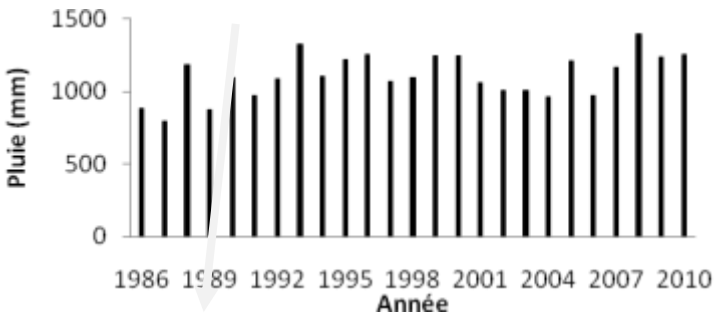


Figure2: Annual pluviometry recorded on the catchment area of Kolondièba-Tiendaga between 1986 and 2010

Tablel: Morphometric characteristics of the catchment area

Designation	Values
Discharge system	Longitudes : 11,066761 Latitudes : - 6,846652
Altitude discharge system	313 m
Surface basin	3050 Km^2
Length of the principal channel	158,30
The overall length of the hydrographic network	5854,26
Strahler order	7
Density of drainage	1,895km/Km^2
Perimeter	430 Km
Index of compactness	2,050 KG
Rectangle length equivalent	186 m
Largeur Rectangle equivalent	16m

2.2 Materials

The data of remote sensing used in this study are satellite images AATSR (Advanced Along Track Scanning Radiometer) of platform ENVISAT of the European space agency (ESA). Sensor AASTR is one of the six instruments of satellite ENVISAT launched by the European space agency, on March 1st, 2002 in polar orbit to provide data on the atmosphere, the ocean, the ground, and freezes it during the five next years. The data produced by ENVISAT will be exploited within the framework of the scientific research on the ground and the monitoring of the environmental and climatic changes. Moreover, they will facilitate the development of operational and commercial applications. Sensor AATSR is a radiometer with sweeping with a space resolution of 1Kmx1Km, recording 7 spectral bands in the visible one and the infra-red. The trace of sensor AATSR is selected in order to ensure a continuity with the data files of ATSR1 and ATSR2 of the series of ERS1 and ERS2 of ESA. The characteristics of the sensor are given in the tableau2. Products ATS-TOA level 1P of AATSR are used in this work. They were acquired near the European space agency (ESA) through the project of reinforcement of the capacities in Africa (TIGER II). These data were downloaded starting from the http://ats-merci-uk.eo.esa.int:8080/merci site, and relate to the period 2003 up to 2010. On the whole, we worked on more than 100 satellite images for the period quoted of in top (2003 up to 2010).

In addition to the satellite data, we used weather data obtained from National management of the Meteorology of Mali. They are relating to the temperature, the speed of the wind, the relative humidity. The data of solar radiation used in this work are collected in [35]

Table2: Characteristics of sensor AATSR of satellite ENVISAT (Esa, 2009)

Spectrales Bands spectrales (µm)	Central wavelength (µm)	Bandwidth (µm)	Spatial resolution (Km)	Application
0,0545-0,565	0,555	0,02	1	Chlorophyl
0549-0,669	0,659	0,02	1	Vegetation index
0,855-0,875	0,865	0,02	1	Vegetation index
1,580-1,640	1,61	0,03	1	Cleaning of cloud
3,50-3,89	3,70	0,30	1	Temperature of sea surface
10,40-11,30	10,85	1,00	1	Temperature of sea surface
11,50-12,50	12,00	1,00	1	Temperature of sea surface

As regards the geometrical resolution, sensor AATSR gathers these data according to two systems of curved sweepings:

A sweeping curved forwards (Forward curved scan)

A sweeping curved towards Nadir (Nadir curved scan)

2.3 Methods

The methodology used in this study is based on model SEBS (Surface Energy Balance System) developed the European space agency (ESA). This model was developed by [16,19,20, 29,30,31]. It is installed on software BEAM like module. Software BEAM was conceived by the European space agency for the treatment, the analysis, visualization, georeferencing satellite images, in particular those produced by ESA (MERIS, AATSR, ASAR, etc). Images AASTR were atmospheric corrected by model SMAC before being used in model SEBS. This model SMAC (Simplified Method for Atmospheric Corrections) is also a module of BEAM developed for the atmospheric correction of images MERIS and AATSR of satellite ENVISAT. The practical application of SMAC is to change the equation of radiative transfer, and to calculate the reflectance of surface of satellite measurements.

Model SEBS consists of a whole of tools for the determination of the physical properties of surface of the ground and the variables of state, such as emissivity, the albedo, the temperature and vegetable cover.

Based on contrast between the dry and wet zones, model SEBS was initiated by [19] which proposed a method to calculate the evapotranspiration starting from the evaporative fraction. This concept was improved by [30, 31] to establish model SEBS. To reach the temperature of surface, model SEBS bases emissivities derived and the temperature from brightness of the bands (in Brightness Temperaure) to calculate the land surface temperature through the technique of Split Window (SPT). The advantage of these techniques of Split Window is the elimination of the effects of the atmosphere, except the atmospheric water columns.
Several coefficients of calibration for the technique of Split Window are tested, model SEBS adopts for satellite images AATSR of satellite ENVISAT of the European space agency that of [26, 27, 28].
It should be noted that the study related to 21 pluviometric stations left again well on the catchment area of Kolondieba-Tiendaga

2.3.1 Estimation of the components of the energy assessment
The data of remote sensing and weather are used to calculate the heat flows of surface. Clear radiation is divided into heat flow latent, significant heat flow and heat flow of the ground. The latent heat flow is calculated like a residue of the energy assessment according to the following equation.

$$LE = R_n - G - H$$

Where R_n is clear radiation resulting respectively from the assessment of radiation entering and outgoing (W/m^2)
LE: latent heat flux (W/m^2)
G: heat flux of the ground (W/m^2)
H: Sensible heat (W/m^2)
R_n and G can be estimated locally by using weather data [1] and on a regional level by incorporation of emitted and reflected radiation distributed spatially[14,15].
Net radiation (R_n). The assessment of radiation of the surface of the ground is given by the following formula.

$$R_n = R_{s\downarrow} - R_{s\uparrow} + R_{L\downarrow} - R_{L\uparrow}$$

Where

R_n :Net radiation

$R_{s\downarrow}$: radiation short wavelength (0,14-4µm) coming from the sun

$R_{s\uparrow}$: radiation big wavelength) (>4µm) coming from the atmosphere

$R_{L\downarrow}$: Quantity of radiation emitted by terrestrial surface

$R_{L\uparrow}$; Quantity of energy reflected by terrestrial surface

$R_{s\downarrow}$ is calculated starting from the product of instantaneous radiation arriving on the ground,[24]. The solar radiation absorptive by surface of the ground is calculated according to the following relation:

$$R_{abs} = (1-\alpha) R_{s\downarrow}$$

Where α is the albedo of surface. It is given starting from measurement of narrow bands by the technique of average of weight suggested by [27]. The radiation arriving big length $R_{L\downarrow}$ is estimated starting from measurements on the ground of the temperature of the air and the steam pressure by using the following relation:

$$R_{L\downarrow} = \varepsilon_a \sigma R_a^4$$

Where is the emissivity of the atmosphere $\left[\varepsilon_a = 1{,}24 \left(e_d / T_a \right)^{1/7} \right]$, σ is the constant of Stefan-Boltzmann (5, $67x10^{-8} W/mk^{-4}$, your is the temperature of the air in (K) and e_d is the deficit of pressure in (mbar). The radiation big wavelength leaving ($R_{L\downarrow}$) is obtained starting from the temperature of surface by being unaware of the small contribution of the radiation of the sky reflected by using the following equation:

$$R_{L\uparrow} = \varepsilon_s \sigma T_s^4$$

Where ε_s is the emissivity of surface and is the temperature of surface (°K). According to [31], emissivity (εs) for the range of 8-14µm could be predicted starting from the NDVI with a strong correlation. Emissivity (ε_s), is thus calculated starting from the NDVI by using the relation following logarithmic curve:

$$\varepsilon_s = 1{,}0094 + 0{,}047 x \ln(NDVI)$$

The quantity of clear radiation received by surface is the entering and leaving sum of all radiations and is calculated according to the equation below.

$$R_n = (1-\alpha) R_{s\downarrow} + \varepsilon_a \sigma T_a^4 - \varepsilon_s T_s^4$$

2.3.1.1 Heat flux of the ground (G): The heat flux of the ground (G) is commonly considered like a fraction of clear radiation R_n depending on the index of the leaf area Index (LAI) and the NDVI (Normalize Vegetation Index) [11]. It can be higher than $0{,}3R_n$ for the grounds naked and lower than 0,1Rn for the covered zones of vegetation [14,15]. Several studies showed that the G/R_n ratio of the day is relative to other factors, such as the quantity of vegetation present [2, 3, 18, 10] presented the following relation to estimate the heat flow of the ground G.

$$G = R_n \left(\frac{T_s}{\alpha} \right) x \left(0{,}0038\alpha + 0{,}0074\alpha^2 \right) x \left(1 - 0{,}98\, NDVI^4 \right)$$

2.3.1.2 Sensible flux heat (H): The estimate of the reliable values of significant heat flow is the most difficult aspect of this methodology because it depends on aerodynamic resistance. H is commonly expressed like a function of T_s and T_a.

$$H = \rho C_p (T_s - T_a) / r_a$$

Where ρ is the density of the dry air (Kgm^{-3}), C_P is the capacity of specific heat of the air ($Jkg^{-1}C^{-1}$), and r_a is the aerodynamic resistance of the transport of heat (s/m^{-1}). r_a is estimated by the theory of similarity of Monin-Obukhov [5]. Many models semi-empirical were proposed to estimate r_a and H [6, 33, 34]. Among these models, that presented by [6], proposes simplest to estimate H starting from some points of measurement easily obtained, such as the NDVI, T_s and T_a. The principal problem of this method is the over-estimate of H in urban environment where the index of vegetation is weak from where an over-estimate of [25].

To avoid this problem a model of aerodynamic resistance semi-empirical was proposed by [33].

$$r_a = 4{,}72 \left\{ \ln(z / z_0) \right\}^2 / (1 + 0{,}54u)$$

Where z is the height the speed of the wind (m), z_0 is the length of roughness of surface (m) and u is the speed of wind (m/s^{-1}).

[13, 21] showed that z_0 can be estimated starting from index of vegetation, such as the NDVI and the ratio of NIR/R. [10] also connected z_0 (cm) with the NDVI to estimate flux of surface at the regional level.[21] used an exponential relation to estimate z_0 starting from ratio NIR/R

$$Z_0 = \exp(0{,}01021 + 0{,}1484(NIR / \mathrm{Re}\, d))$$

2.3.1.3 Estimation of land surface temperature

As announced higher, the temperature of surface was estimated according to the formula of [26], was integrated directly in module SEBS of BEAM. This formula is based on the temperature of brightness, the steam contents, the emissivity of surface and the difference in emissivity to calculate the temperature of surface of the ground

$$LST = 0{,}39T_1^2 + 2{,}34T_1 - 0{,}78T_1 x T_2 - 1{,}34T_2 + 0{,}39T_2^2 + 0{,}56$$

Where T1 and T2 represent respectively bands 6 and 7. The coefficients are empirically given on the basis of type of occupation of the ground, vegetable cover, the season and the duration of the day, the atmospheric steam pressure, the zenith angle and the emissivity of surface [23]

3. Results and Discussion

3.1 Space-time distribution of the values of temperature of surface of 2003 to 2010 analyzes

From the analysis of the **tableau3**, it appears that the median values of the temperatures of surface of the ground estimated starting from satellite images AATSR, using model SEBS (Surface Energy Balance System), on the catchment area of kolondièba-Tiendaga, oscillate between 300 and 292°K. At the same time, the maximum and minimal values vary respectively between 303 and 299°k (maximum values) and between 288 and 299°K for the minimal values.

As for the values of standard deviation, they vary between 1 and 3.

In addition, we found that the highest values of temperature of surface were recorded respectively into 2007,2003,2005,2006,2010,2009,2004 where they vary between 303 and 299 (°K). On the other hand, the low value estimated during this study was recorded in 2008 (296°K). These values are in the same fork as those found by [23], on the basin of Volta, in West Africa with the same types of satellite data (AASTR.)

The analysis of the space-time variation between 2003 and 2010 revealed that the year 2007 is hottest with the value of temperature of the highest surface (303°K) for a standard deviation of 2,67. Whereas year 2008 is shown the least hot, where 296,21°K was estimated with a standard deviation of 0,945.

It should however be recognized that these results were not validated, because of the limited means. Indeed, the project did not have average materials sufficiently to undertake studies of validation of ground.
The analysis of the various charts worked out starting from these values of temperature of surface of the ground shows that these values increase South in North, contrary to those of the evapotranspiration, over the same study period (2003 to 2010), under the same conditions.
Indeed, we noticed that the values of temperatures of surfaces of the ground are high on the parts of the catchment area where the percentage of naked ground is most important. While the low values of temperatures of surface of the ground are observed in the zones where vegetable cover is relatively much more important, and richer in hydrographic terms of networks and water resource.
We can conclude that the variation of the median values of temperature of surface on the catchment area of Kolondièba-Tiendaga, in southern Mali zone, follows the dynamics of occupation of the grounds.
Table 3: Comparaison des valeurs de température de surface mesurées (°K) à partir des images AATSR de 2003 - 2010 sur le Bassin versant de Kolondièba-Tiendaga

Stations	2003	2004	2005	2006	2007	2008	2009	2010	Moy	Max	Min	Std variat
1	296,92	293,17	295,10	298,94	295,39	295,67	291,48	290,96	294,70	298,94	290,96	2,70
2	299,06	294,13	294,58	296,14	297,29	295,22	291,65	294,81	295,36	299,06	291,65	2,21
3	298,57	296,31	296,23	295,59	296,69	295,66	292,68	292,51	295,53	298,57	292,51	2,03
4	297,57	297,77	297,72	296,02	298,86	296,18	292,63	291,16	295,99	298,86	291,16	2,71
5	298,19	298,25	297,82	296,69	296,61	295,67	293,83	291,59	296,33	298,25	291,59	2,16
6	299,84	297,61	298,42	296,94	295,35	292,90	296,71	292,66	296,30	299,84	292,66	2,54
7	302,50	297,45	297,24	295,40	295,23	296,04	298,56	291,57	296,75	302,50	291,57	3,13
8	301,92	297,18	296,03	293,49	293,61	295,17	296,80	291,49	295,71	301,92	291,49	3,15
9	301,80	298,62	298,16	297,15	295,89	296,14	291,88	298,10	297,22	301,80	291,88	2,83
10	300,69	296,12	298,82	298,82	297,31	294,84	290,88	295,70	296,65	300,69	290,88	3,02
11	298,52	296,31	299,52	299,13	297,95	295,58	292,61	298,47	297,26	299,52	292,61	2,31
12	298,23	294,65	297,64	297,83	300,01	293,58	293,69	297,69	296,66	300,01	293,58	2,37
13	299,73	296,44	298,67	299,73	300,48	295,92	294,05	297,59	297,83	300,48	294,05	2,23
14	302,80	295,44	298,51	295,30	300,59	295,72	290,76	298,66	297,22	302,80	290,76	3,72
16	301,28	296,39	298,02	299,48	300,00	294,44	290,12	298,65	297,30	301,28	290,12	3,60
15	302,83	295,89	296,83	297,24	302,20	294,22	290,46	296,75	297,05	302,83	290,46	4,02
18	300,63	291,70	295,36	299,37	300,76	294,36	288,54	297,36	296,01	300,76	288,54	4,38
17	301,56	293,89	298,74	293,33	298,94	294,44	290,94	297,84	296,21	301,56	290,94	3,58
19	298,19	291,50	298,25	295,17	301,29	293,76	288,88	298,76	295,73	301,29	288,88	4,17
20	299,87	293,94	298,65	295,49	303,49	295,60	291,63	299,88	297,42	303,49	291,63	3,83
21	299,30	293,48	297,47	294,78	301,05	296,09	289,97	300,26	296,55	301,05	289,97	3,75
Moy	300,00	295,54	297,51	296,76	298,52	295,10	292,42	295,83				
Max	302,83	298,62	299,52	299,73	303,49	296,18	298,56	299,88				
Min	296,92	291,50	294,58	293,33	293,61	292,90	288,54	290,96				
Std variat	1,80023	2,0569	1,35218	1,9491	2,666	0,9446	2,6806	3,23242				

3.2 Analyzes correlation between the temperature of surface and the indices of vegetation (NDVI)

Front, to carry out this analysis, we thought necessary to have the advantages and the limits of the NDVI.
Defines by [26], the NDVI is the index of vegetation most frequently used in the studies on the follow-up of the vegetation, and is admitted by the whole of the scientific community as being most reliable. This index gives an account of the chlorophyllian activity of canopy and makes it possible to quantify the produced vegetable biomass [26]. The applications of the NDVI are numerous: evaluation of deforestation, follow-up of forest fires, the turning into a desert and same of the devastations caused by insects [26].
Given that the vegetation absorbs an important part of the solar radiation in the red and that it reflects it in the infra-red close relation, its formula can be written

$$NDVI = \frac{\rho_{NIR} - \rho_{R}}{\rho_{NIR} + \rho_{R}}$$

Where

ρ_{NIR} is the spectral band corresponding to the infra-red close relation (ranging between 0,55 and 0,68 µm)

ρ_{R} that corresponding to the red (including between 0,73 and 1,1µm). The theoretical values of the NDVI vary [26] consider that this index is sufficiently stable between -1 and 1 (between 0,1 and 0,9 for the vegetation).
toallowcomparisons of the vegetable activity steps seasonal or interannual times. They add that the force of the NDVI even lies in its design. Indeed, this index being a ratio, calculated starting from two spectral bands, it makes it possible to eliminate part of certain noises like the atmospheric contaminations (aerosols…), differences in illumination of the ground or shade cloudy. In addition, the NDVI has an interesting resolving power: the clouds, water, and snow have a greater reflectance in the visible one in the infra-red, their NDVI is thus negative and one can insolate them easily. Contrary, the naked rocks and grounds will have a similar reflectance in the visible one and the infra-red close relation, but their very low value of NDVI, on both sides of zero, does not allow confusion with vegetation [26].This index has only advantages, [26, 37] recognize some limiting with the NDVI. In particular the recurring presence of atmospheric contaminations residual or of skews related on the camera angle and the relief. Its principal disadvantage is its tendency to saturation when the foliar density (LAI) is too strong on the level of the groundThe index saturates and accounts more for the phenologic variations so covered vegetable is too dense It arises from the analysis between the monthly values of temperature of surface (LST) and the index of vegetation (NDVI) which it exists a strong correlation between the two parameters. Indeed, our results show strong values

coefficients of correlation for the years (2005, 2004, 2007, 2006, 2003). We respectively obtained values of R2 (0,834; 0,716; 0,686; 0,582; 0,563). On the other hand, a low coefficient of correlation is observed for the years (2008, 2009, 2010), where respective values of R^2 (0,394; 0,376; 0,178) are obtained. The fall of R^2 for these years is explained by the lack of data for years. Indeed, for these three years. There were less data AATSR, in particular for the months from July-August. This insufficiency of data can be at the origin of this fall.

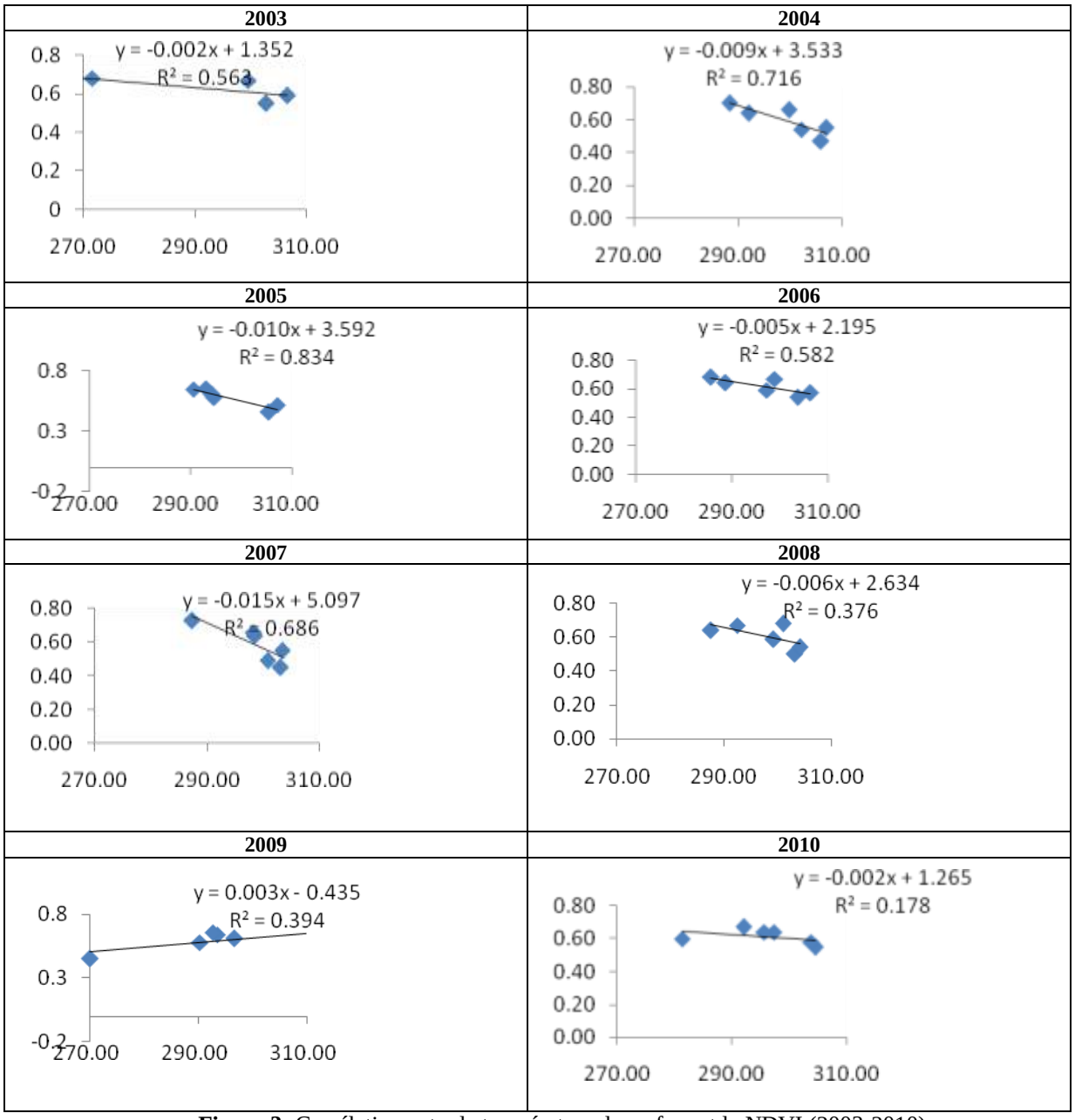


Figure 3: Corrélation entre la température de surface et le NDVI (2003-2010)

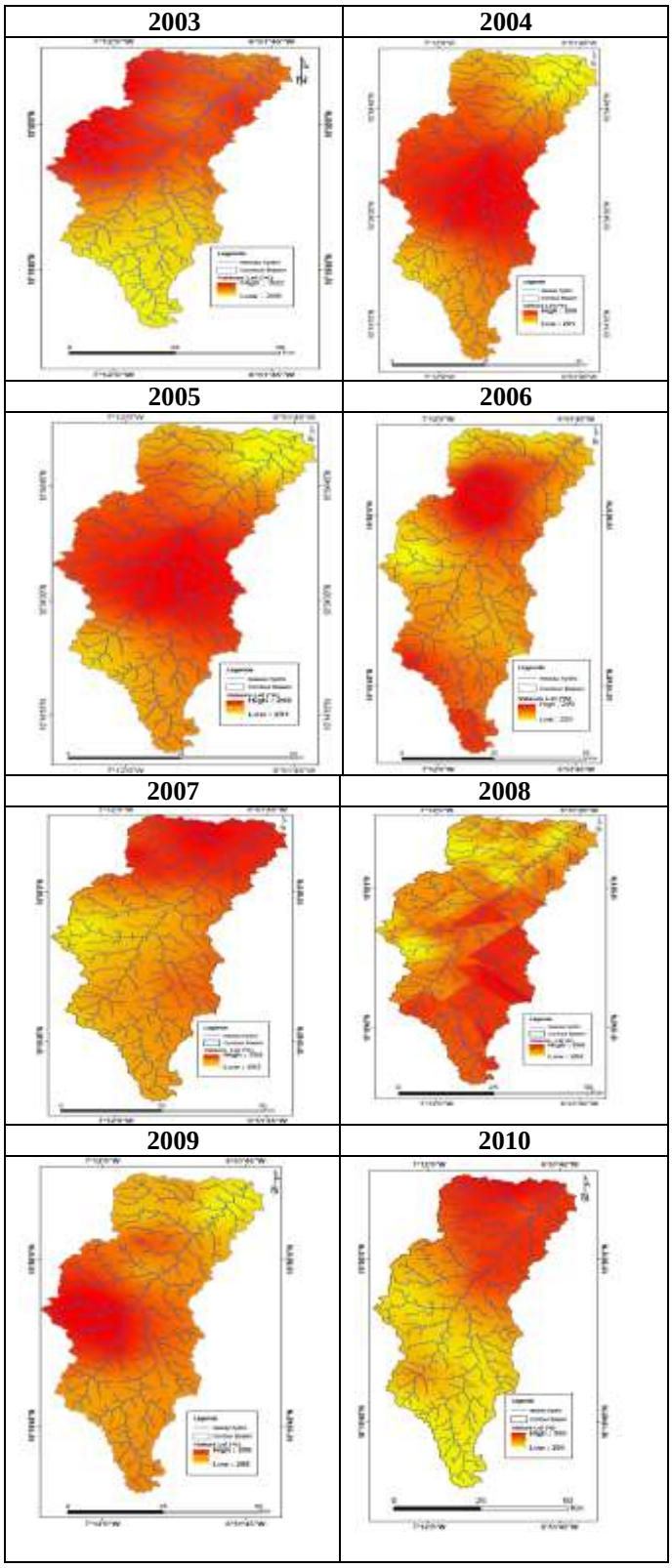


Figure 4: Charts of temperature of surface (°K) of the catchment area of Kolondièba-Tiendaga between 2003 and 2010 elaborate starting from images AATSR

4. Conclusion

The tools for remote sensing and GIS appear today like an essential element in the collection and the mapping of the parameters biophysics on broad scales, of which in particular the evapotranspiration, the land surface temperature, emissivity, albedo, etc

The use of model SEBS developed by [31] made it possible to determine and chart the temperature of surface starting from satellite images AATSR on the catchment area of Kolondièba-Tiendaga in climate soudano-saheliaen, in southern the Mali zone. The results obtained of the study show values of temperature which vary between 303 and 296 with standard deviations of (2,66 and 0,945). These same results show a very good tendency with the evapotranspiration calculated starting from the evaporative fraction.

For the later studies, it would be more adapted to make a comparison between the values of temperature of surface estimated starting from model SEBS and those collected starting from a station or a measuring device of land surface temperature.

Acknowledgement

Our thanks are addressed to DANIDA for the financial support of this study within the framework of project WANSEC (West African Network for Studies Environment Changes). Our thanks are also addressed to the European space agency (ESA) for its customer support within the framework of project TIGER II.

We thank the Eau unit and Environment for the DER of Geology of the ENI-ABT of Bamako (Mali) to have sheltered this work.

In end, we thank the IRD through project RIPIECSA for its support in the finalization for the drafting for this article.

References

[1] Allen,R., Pereira, L., Raes, D., Smith, M., (1998). Crop evapotranspiration guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper, 56. FAO, Rome.

[2] Bastiaanssen, W.G.M., (2000). SEBAL based sensible and latent heat flux in the irrigated Gediz Basin, Turkey. Journal of hydrology, 229, pp. 87-100

[3] Bastiaanssen, W.G.M., Menenti, M., Fedds, R.A. and Holtslag, A. A. M., (1998). A remote sensing surface energy balance algorithm for Land (SEBAL)-1. Formulation. Journal of hydrology, 212/213, pp. 198-212

[4] Bhattacharya, B. K., Dadhwal, V.K., (2003). Retrieval and Validation of land surface temperature (LST) from NOAA AVHRR thermal images of Gujarat , India. International Journal of Remote sensing, 24,06,pp. 1197-1206

[5] Businger, J.A. (1988). A note on the Businger Dyer profiles. Boundary- Layer Meteorological, 42: 145 151.

[6] Carlson, T. N., Capehart, W.J., and Gillies, R.R., (1995). A new look at the simplified method for remote sensing daily evapotranspiration. Remote sensing of Environment. 54, 2, pp. 161-167.

[7] Chen, D., Gao, G., Yu, C. Y., Guo, J., Ren, G., (2005). Comparison of the thornthwaite method and pan data with the standard Penman-Monteith estimates of references evapotranspiration in China. Climate Research, 28, pp. 123-132

[8] Courault, D., Seguin, B., Olioso, A., (2005). Review on estimation evapotranspiration from remote sensing data: from empirical to numerical modeling approaches. SPRINGER. Irrigation and Drainage system, 19, pp. 223-249

[9] French, A. N., Schmugge, T. J., kustas, W. P., Brubaker, K. L., et Prueger, J., (2003). Surface energy fluxes over EL Reno Oklahoma, using high resolution remotely sensed data. Water resources research 39(6), 1164

[10] Gao, W., Coulter, R. L., Lesht, B. M., Qui, J., Wesley, M. L., (1998). Estimating clear-sky regional surface fluxes in the southern great plains atmospheric radiation measurement site with ground measurement and satellites observations. Journal of Applied Meteorology, 37, pp. 5-22.

[11] Gowda, H.P., Chavez, L. J., Calaizzi, D. P., Evett, R. S., Howell, T. A., Tolk, A. J., (2007). ET mapping for agricultural water management: Present status and Challenges. Springer Verlag, Irrigation Science, 26, pp. 223-237.

[12] Gupta, R. K., Prasad, S., Sesh Sai, and Viswanadham, T. S., (1997). The estimation of land surface temperature over an agricultural area in the state of Haryana and Panjab, India, and its relationships with the Normalized Difference Vegetation Index (NDVI) using NOAA AVHRR data. International Journal of Remote sensing, 18, 18, pp. 3729-3741

[13] Hatfield, J. L., (1988). Large scale evapotranspiration from remotely sensed surface temperature in proceedings: planning now for irrigation and drainage, Lincoln NE, 18-21 July, American Society of Civil Engineers (ASCE). New York, pp. P205-P509.

[14] Jacob, F., Oliosio, A., Gu, X., Su, Z., & Seguin, B., (2002a). Mapping surface fluxes using airborne, visible, near infrared, thermal infrared remote sensing and a spatialized surface energy balance model. Agronomie, 22, 669-680

[15] Jacob, F., Schmugge, T., Ogawa, K. A. F., & Ritchie, J., (2002b). The potentialities of ASTER to retrieve radiative properties over semi arid regions. In J. Sobrino (Ed), First International Symposium on recent advances in quantitative remote sensing, pp. 913-920

[16] Jia, L., Su, Z., Van den Hurk, B., Menenti, M., Moene, A., De Bruin H. A. R ., Yrissary, J. J. B., Ibanez, M., and Cuesta, A., (2003). Estimation of sensible heat flux using the Surface Energy Balance System (SEBS) and ASTER measurements. Physics Chemistry of the Earth, 28(1-3), 75-88

[17] Kant, Y., Badarinath, K. V. S., (2000). Studies of land surface temperature over heterogeneous area using AVHRR data. International Journal of Remote sensing, 21, 8, pp, 1749-1756

[18] Kustas, W. P., and Norman, J. M., (1999). Reply to comments about the basic equations of dual-source vegetation-atmosphere transfer models. Agricultural and Forest Meteorology, 94, 275-278.

[19] Menenti, M., Choudhury, B. J., (1993). Parameterization of land surface evapotranspiration using a location dependent potential evapotranspiration and surface temperature range. In: Boll, A.J. et al., (eds) Proceedings of exchange of processes of land surface for the range of space and time scales. IAHS Publ 212, pp. 561-568.

[20 Menenti, M., Jial, L., Su, Z., (2003). On SEBI-SEBS validation in France, Italy, Spain, USA, and China. In: Co-chair Allen, R. G., Bastiaanssen, W., (eds) Proceedings of the workshop on use of remote sensing of crop evapotranspiration for large regions. International Commission on Irrigation and Drainage (ICID) Monptellier

[21] Moran, M. S., (1990). A satellite based approach for evaluation of the spatial distribution of evapotranspiration from agricultural lands (dissertation abstract). PhD dissertation, University of Arizona, Tuscan.

[22] Norman. J. M., Kustas, W. P., Humes, K. S., (1995). A two- sources approach for estimating soil and vegetation energy fluxes from observation of directional radiometric surface temperature. Agric for Meteorol 77, 263-293

[23] Opoku Duah, S., Donoghue, D. N. M., Burt, T .P., (2008). Intercomparison of evapotranspiration over the Savannah Volta Basin in West Africa using remote sensing data. SENSORS.ISSN, 1424-8220

[24] Parodi, G. N., (2000). AHVRR Hydrogical Analysis System, Algorithms and theory- Version 1.0. International Institute for Aerospace and Earth Science (ITC), Wageningen, the Netherlands.

[25] Richter, R., (2002). ATCOR for ERDAS imagine (Ver. 2.0). Manual User. Geosystems GmbH, Riesstr, 10, D-82110 Germering, Germany.

[26] Rousse, J.W., Haas, R.H., Schell, J.A., and Deering, D.W. (1973). Monitoring vegetation systems in the great plains with ERTS. In ERTS Symposium, NASA, SP-351, Washington DC, vol. 1, pp. 309-317.

[27] Sarwar, A., Bill, R., (2007). Mapping evapotranspiration in the Indus basin using ASTER data. International journal of remote sensing, 28, pp. 5037-5047

[28] Sobrino, J.A., Gomez, M., Jiménes-Munos, J. C., Olioso, A., (2007). Application of a simple algorithm to estimate daily evapotranspiration from NOAA-AVHRR images for the Liberian Peninsula. Remote sensing of Environment, 110, pp. 139-148

[29] Sobrino, J. A., Jiménez Munoz, L. C., (2005). Land surface temperature retrieval from thermal infrared data: An assessment in the context of the surface processes and ecosystem changes through response analysis (SPECTRA) mission, Journal of Geophysic al Research, 110, D16103, doi: 10. 1029/2004JD005588

[30] Su, Z., Pelgrum, H., and Menenti, M., (1999). Aggregation effects of surface heterogeneity in land surface processes. Hydrology and Eartth System Science, 3 (4), 549-563.

[31] Su, Z., Schmugger, T., Kustas, W. P., Massman, W. J., (2001). An evaluation two models for estimation of the roughness height for heat transfer between the land surface and atmosphere, J, Appl Meteorol, 40, 1933-1951

[32] Su, Z., (2002). The surface energy balance system (SEBS) for estimation of the turbulent heat fluxes. Hydrology and Earth Science, 6(1), 85-99

[33] Thom, A.S., and Oliver, H.R., (1977). On penman's equation for estimating regional evaporation. Quartely. Journal of the Royal Meteorological Society. 103, 436, pp. 345-357

[34] Viney, N.R., (1991). An empirical expression for aerodynamic resistance in the unstable boundary layer. Boundary Layer Meteorology. 56, 4, pp. 381-393.

[35] Wang, K., Wang, P., Li, Z., Cribb, M., & Sparrow, M., (2007). A simple method to estimate actual evapotranspiration from a combination of net radiation, vegetation index, and temperature. Journal of Geophysical research, 112, pp. D15107

[36] Wanger, W., (1998). Soil moisture retrieval from ERS scatterometer data. PhD dissertation to the technisch-naturwissenschaftliche fatkultät of the Vienna University of Technology, Karlsplatz, 13, A 1040 Wien Austria

[37] Zhang, J., Wang, Y., Wang, Z., (2007). Change analysis of land surface temperature based on robust statistics in the estuarine area of Peark River (China) from 1990 to 2000 by Landsat TM/EMT+ data. International Journal of remote sensing, 28, 10, 2383-2390

Routing and security for remote labs for teaching and research (SRS-E-LABO)

Alassane Diop
Research Associate, Center for Research LICEF, TELUQ / UQAM,
Montréal, Québec, Canada

Abstract:
In this paper, routing services and security are software components that can be integrated in an e-laboratory or remote laboratory as well as in mobile laboratories. We will show how to implement services to manage the routing and security using virtual routers. The virtual routers that we propose allow designers to online laboratories to adapt or supplement their learning environments or dedicated platforms using their services to route and secure access to physical devices and the information passed to the various network nodes to remote laboratories. The concept to be presented, based on WCF (Windows Communication Foundation), provides centralized network management at a single location while the remote labs are based on networks and distributed systems.

Keywords: Remote labs, online laboratory, computer network, WAP, virtual router, security, e-laboratory, e-learning, WCF

1. Objectives and routing services security

This article aims to contribute to the development of service concepts, routing and security using software components that can be integrated with laboratory environments online. These software components are presented in the form of services to manage 1) addressing the various nodes of connections, 2) the table router to forward packets from networks or sub networks connected directly or indirectly to the primary node, 3) transmission security with an encryption of information, 4) public and private keys for data privacy; 5) mobile devices and applications WAP (Wireless Application Protocol) and service use in web applications by exploiting the available services. Routing services and security for remote labs for teaching and research online (SRS) are reliable and secure transactional solutions that integrate across platforms and interoperate with existing infrastructure. The SRS will be based on the WCF architecture (Windows Communication Foundation) which is compatible with all other existing architectures because "The message parameters are defined by documents in XSD (XML Schema Definition), which allows any system comprising XML (eXtended Markup Language) to manage documents. The message contract defines specific message parts using SOAP (Simple Object Access Protocol), and allows finer control of them, when interoperability demands such precision. "[1]. Figure 1 illustrates the main layers of the WCF architecture used to implement the various services of SRS.
Service contracts illustrated in Figure 1 have spoken publicly in WSDL (Web Services Description Language) [2], a format that describes network services in terms of operating parameters on messages SOAP (Simple Object Access Protocol). This allows designers to use that information to public contracts to build service oriented applications that ensures interoperability. To transfer information on intranets and the Internet, the services use protocols such as HyperText Transfer Protocol (HTTP) and Transmission Control Protocol (TCP). Other types of transport mechanisms can be added via extension points of WCF integrated.

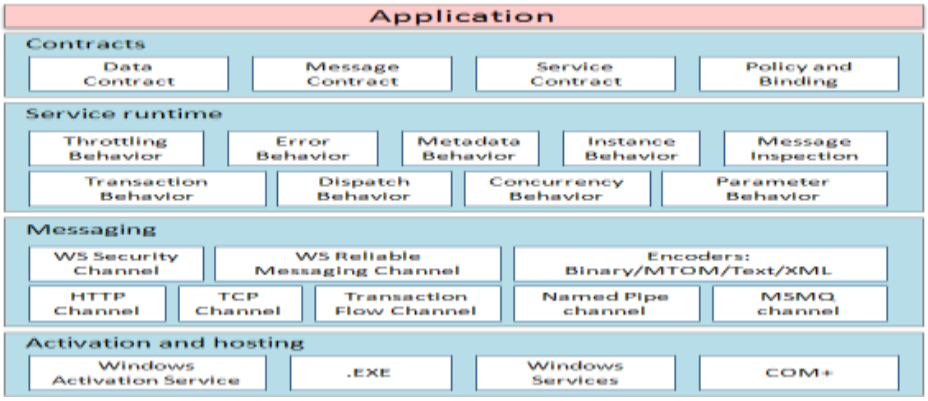


Figure 1. WCF Architecture [1]

2. Communication services routing and security

The SRS can be used either ad hoc or infrastructure. Figures 1 and 2 show the two modes of communications services, routing and security levels and applications to an established e-collaboration environment consisting of a platform for distance learning (online platform), a site laboratory Remote (e-laboratory) and another site that hosts communication tools useful for synchronous and asynchronous learning or collaborative research (Collab).

Ad hoc mode as the name suggests allows each item that uses routing services (SRS) to host and use its own services directly. These can be shared or restricted in Intranet. The communication mechanism is very simple. To deploy routing services must sequentially, initially, install a server that hosts all SRS available, then set the access "users", controls and security; mount the virtual router using local services or remote, and finally implement the global environment which exploits a remote laboratory using deletion of SRS. Information can be exchanged between sites represented by rectangles in Figure 2. Note that the communication is bidirectional.

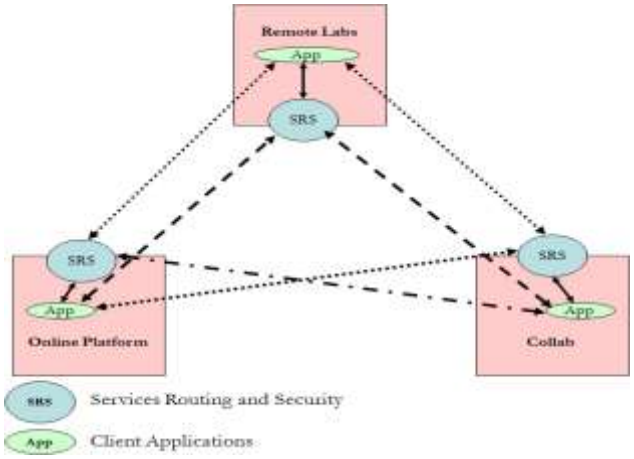


Figure 2. Communication services routing and security in ad hoc mode

The infrastructure mode shown in Figure 3 allows a centralized interconnection of SRS on a single server hosting services. This communication mode is easier to maintain when updating the system. The SRS is divided for the various sites where the possibility of overloading likely. In this case a policy of performance management and bandwidth can be introduced to give priority levels to the sites. The communication mechanism is more complex to manage, compared to ad hoc mode, but the steps to establish communication and use the services remain the same. Information can be exchanged between sites represented by rectangles in Figure 3. To do everything passes through the central node of the SRS.

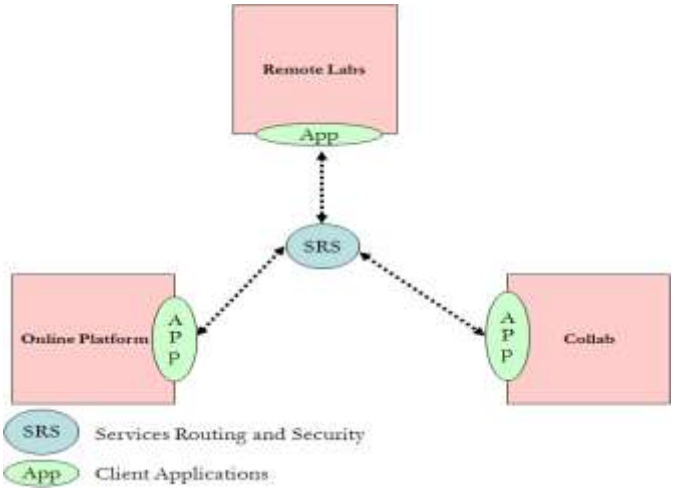


Figure 3. Communication Services Routing and Security Infrastructure mode

It is possible to combine the two basic modes for a mixed mode as shown in Figure 4.

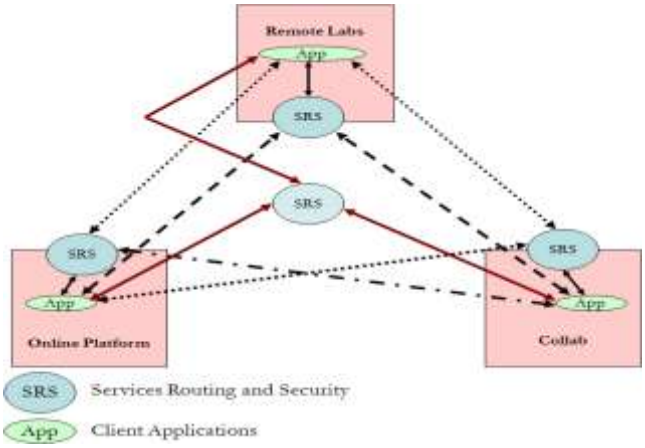


Figure 4. Communication Services Routing and Security in Mixed Mode

3. Description of services routing and security

Five management points will be treated as mentioned in Section I of this paper.

3.1. Address management of the various nodes of connections

These are services for managing the global addressing of nodes sites. As input parameters, we have the IP address of the equipment and the subnet mask that hosts SRS, the type of communication (Local, Remote, Internet-Remote Intranet VPN: Virtual Private Network, etc. ...) as well as the connection mode (ad-hoc, Infrastructure, Mixed). As output parameters, we have the IP address, physical address (MAC) and indicator connection (Connected or not connected). The information will be stored in a database for tracking, transactions and security. These services allow us to have as well an overview of the system based on information collected at nodes or sites. It is even possible to make the overall network diagram that uses SRS.

3.2. Management table router to forward packets from networks or sub networks connected directly or indirectly to node

These are services that manage the routing tables stored in a database routing. Virtual routes are defined to facilitate transport and routing information. As input parameters, we have the source address of the original node and the address (es) recipient (s) of the end node or end nodes. As output parameters, we have the packet flow of information from the nodes and the safety information required. It is also possible to use the broadcast address (broadcast) to transmit information at different nodes of the global network.

3.3. Security management system with a transmission of information encryption

These are services that manage the security of transmissions using an encryption system. All information transmitted from node to node will be protected by an encryption algorithm using the management services described public and private keys to the next depending on whether the information coming from the central server or a node of the central system. These services also use the information generated by the management services of the routing table. So we need the source address, destination addresses and information or applications that must pass at the nodes of the central system. Output, we have information packets encrypted and then decrypted at that node by a service recipient decryption of packets.

3.4. Management of public and private keys

These are services that manage the public and private keys. All nodes are equipped with SRS security certificate checked and authorized by a certificate authority or free. Services in place to manage effectively the keys generated by these authorities. These certificates require a subscription to these certification authorities such as, for example, Entrust and VeriSign. It is also possible, before moving to production, use some utilities to generate public and private keys as part of testing applications.

3.5. Management of mobile devices and applications and WAP services use in web applications by exploiting the available services

These services allow you to manage the connections between the central node and mobile devices. Routing services support mobile devices such as PDAs and cellular phones with WAP technology. The example of using WAP technology for online training will be used to validate the service [3]. Mobile applications can also use the services of routing and security.

4. WSDL Structure of services routing and security

The structure of the various services put in place meets the international standard WSDL documents namely:
The WSDL document will be divided into two groups of sections. The top group consists of abstract definitions, while the bottom group contains concrete descriptions. The abstract sections define SOAP messages (Simple Object Access Protocol) completely independently of the platform and language.
The definition of a set of services that can be implemented by different web sites or applications. Each WSDL document can have sections illustrated in the following diagram:

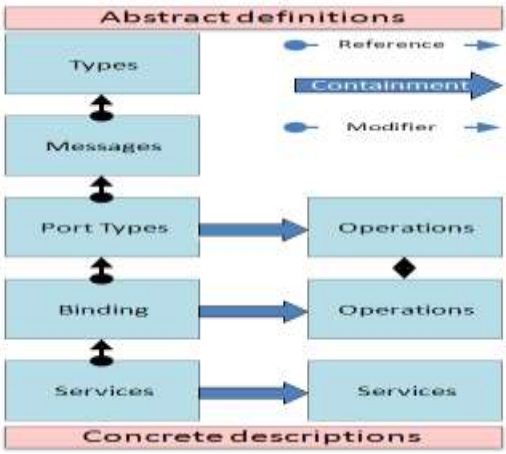


Figure 5. Structure of WSDL documents [1]

5. Results and conclusion

In this part of the article, the connection service owned management services addressing is presented. Figures 6 and 7 show the tests with the "svcutil.exe" tool from visual studio. This tool allows you to quickly test the functionality of the services developed.

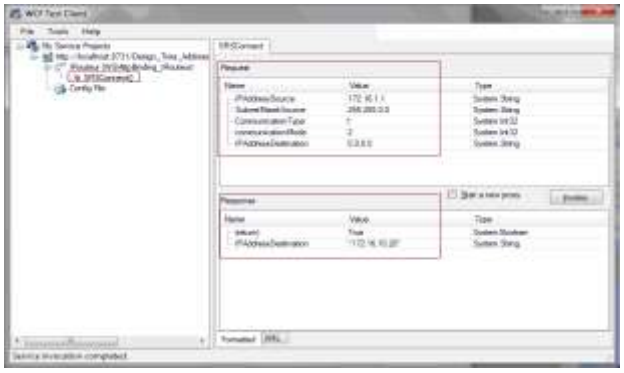


Figure 6. Testing Service SRSCONNECT

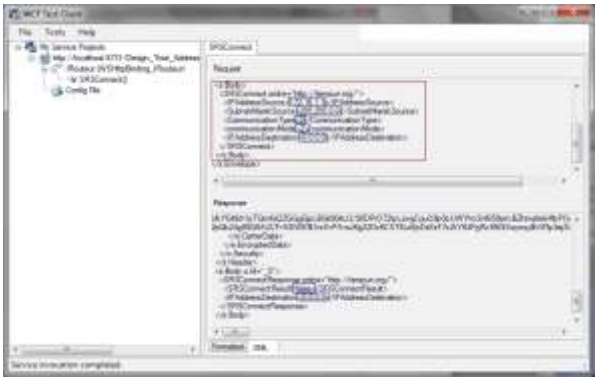


Figure 7. Representation of results in XML format

The service takes as input SRSCONNECT four parameters (information on the source node, the mode and type of communication) and establishes a connection between two nodes (virtual router). If the connection is established, the service outputs the logical and physical addresses of the remote site as well as the true indicator (connected) or false (not connected). Figures 8 and 9 show an example of using the SRS services in C # and Visual Basic. Figure 6 shows how to add references of routing services, while Figure 9 shows a web browser, the service description and a sample source code to call the SRS.

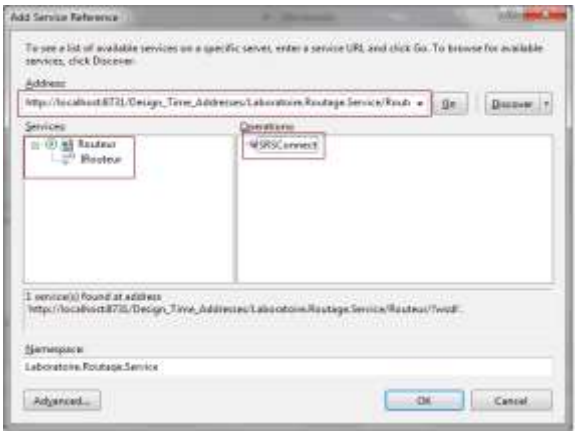


Figure 8. Reference Service SRSCONNECT on visual studio

Figure 9. Example of C # and Visual Basic provided by the online service

To test this service, the first thing to do is create a client and write the code to call the service. To do this, one possibility is to use the "svcutil.exe" tool from the command line using the following syntax:

```
svcutil.exe http://localhost:8731/Design_Time_Addresses/Laboratoire.Routage.Service/Routeur/?wsdl
```

This will create a configuration file and a code file containing the class of customer. Adding the two files to a client application and the generated client class to call up the service SRSCONNECT. The following table shows two examples, the first source code written in C # and Visual Basic in the second

C#	Visual Basic

class Test
{
static void Main() {
RouteurClient client = new RouteurClient();
//Use the variable "client" to call operations on the service.
Client.srsconnect(ipsource, marqueSource, Type,
Mode, IpDestination);
//Always close the client.
client.Close(); }
}
	Class Test
	Shared Sub Main()
	Dim client As RouteurClient = New RouteurClient()
	'Use the variable "client" to call operations on the service.
	Client.srsconnect(ipsource, marqueSource, Type,
	Mode, IpDestination);
	' Always close the client.
	client.Close()
	End Sub
	End Class

Table 1. Sample code to use the services of the routing software

Routing services and security contribute to online laboratories in the field of networks and security. With its various services in place, you can easily create an online laboratory that learners can use to do practical work and experimental without the need for physical facilities or to go to physical laboratories [4] [5]. This option allows to extend the use of laboratories using a platform accessible from a web browser or mobile device.

References

[1] Microsoft documentation, Architecture windows communication foundation, http://msdn.microsoft.com/fr-ca/library/ms733128.aspx, site consultée en Septembre 2012

[2] Microsoft documentation, Description du langage WSDL (Web Services Description Language), http://msdn.microsoft.com/fr-fr/library/bb469924.aspx, site consultée en Septembre 2012

[3] Diop Alassane, Distributed environment of learning suitable for telephony and mobile computing; International Journal of Theoretical and Applied Information Technology (JATIT); E- ISSN 1817-3195 / ISSN 1992-8645; pp 057 – 064, Vol 39. No. 1 – Mai 2012.

[4] Diop Alassane, Livre, Laboratoires à distance de télématique et réseaux, Analyse, conception et étude de cas pratiques pour l'enseignement en ligne. Edition : Éditions universitaires européennes; 2012; 144 pages; ISBN: 978-3-8417-8997-6; Parution: Février 2012.

[5] Diop Alassane, Architecture of a virtual and remote laboratory; International Journal of Reviews in Computing (IJRIC); ISSN : 2076-3328 | E-ISSN : 2076-3336; pp 017 – 021, Vol 10 – Juillet 2012.

Performance Evaluation of Aodv and Dsr Routing Protocols for Vbr Traffic for 150 Nodes in Manets

Gurpreet Singh, [1] Atinderpal Singh[2],
[1, 2]Department of CSE & IT, BBSBEC, Fatehgarh Sahib, Punjab, India
[3]Assistant Professor, Management Wing, SBSSTC, Ferozepur, Punjab, India

Abstract:
Mobile Ad hoc network (MANETs) are rapidly growing as an important area of wireless communication with the revolutionary inventions in the mobile devices. Efficient routing protocols make MANETs reliable. Despite the considerable simulation works, still more investigation is required in the performance evaluation of routing protocols for multimedia traffic especially Variable Bit Rate(VBR).In this paper, we will conduct a number of simulations for the performance evaluation of two popular routing protocols of MANET, namely AODV and DSR, for VBR multimedia traffic using Real Time Protocol(RTP).We will investigate the performance using four metrics-packet received, throughput, routing overhead and network load.

Keywords: MANETs; AODV; DSR; VBR; RTP.

I. INTRODUCTION

A Mobile Ad hoc network (MANET) [1] [2] is an autonomous collection of mobile routers or nodes communicating over radio links. MANET is a temporary network without infrastructure. The wireless routers or nodes moves randomly and organize themselves arbitrarily. The nodes directly communicate via wireless links within each other's radio range, while that are distant apart use other nodes as relay in a multihop routing function. As the nodes are mobile, the structure of the network changes dynamically and unpredictably over time.Adhoc networks are self-configuring and self-organizing, so to maintain communication between nodes in the network, each node behaves as a transmitter, a host and a router.

Due to growing usage of the portable devices and progress in the wireless communication, Mobile Ad hoc networks are gaining importance due to its wide range of applications [3].Mobile Ad hoc networks are used in military communication and operations, home appliances, emergency services, educational applications and entertainment.

MANET has few challenges and limitations that have to be addressed [4].Limited radio transmission range, routing overhead, battery constraints, security, etc. are some of inefficiencies in a MANET environment.

In [5], the performance of three popular routing protocols AODV, DSR and OLSR is evaluated for multimedia data in a multihop mobile network by taking different performance metrics. In this paper, we will evaluate the performance of two reactive protocols AODV and DSR by taking the VBR multimedia transmission in a multihop mobile network. The mobility scenario simulates the environment of a modern city, where mobile nodes are connected to each other and communicates.

This paper is organized as follows: The next section presents the main Ad hoc routing protocols AODV and DSR that are used in the performance evaluation process. Section III presents the simulation set up and the performance evaluation metrics. In section IV, we present the simulation results and performance comparison. And finally, we conclude the paper and present the plans for the future work in section V.

II. MOBILE ADHOC ROUTING PROTOCOLS

There are many different ways to classify the routing protocols of MANETs depending upon their network structure, communication model, routing strategy and state information [6]. Depending upon the routing strategy, there are two types of routing protocols: Table Driven and Source Initiated (On Demand).

The first type is a proactive approach i.e. table driven. Routing information in the network maintained up to date. Table driven protocols have lower latency but have high routing overhead. The routing overhead is high due to periodic updations of routing tables. Some of main proactive protocols of MANET are DSDV, FSR and OLSR.

The second type is source initiated. It is also known as On Demand or Reactive. Route id created only when demanded by the source node. When a node wants to communicate in the network, it initiates a route discovery process within the network. So, there is no need of periodic updates. Hence, the routing overhead is low but the latency is high. Some of main reactive protocols of MANET are AODV, DSR and TORA. The mobile Ad hoc routing protocols considered in this study are described below.

III. DYNAMIC SOURCE ROUTING (DSR)

DSR [7] [8] is an On Demand routing protocol. DSR is based on the theory of source based routing and is a very simple and efficient routing protocol. DSR is designed for use in multihop wireless Ad hoc networks of mobile nodes. The DSR protocol works into two main mechanisms called Route Discovery and Route Maintenance. Route Discovery is the mechanism in which a source node tending to send a packet to a destination obtains a source route to destination. It is initiated only when a source node wants to send packet to a destination and doesn't already know the route to it. And, then it maintains that route in the cache of the nodes. Route Maintenance is the mechanism in which source node is able to detect the link failures to the destination. It then repairs the route or finds an alternate route. It is used only when source is sending packets to a destination.

IV. B. ADHOC ON DEMAND DISTANT VECTOR PROTOCOL (AODV)

AODV [7] [9] is a reactive distant vector protocol. It mixes the property of DSR and DSDV. Routes discovered on demand are maintained as long as they are required. AODV routing table maintains routing information of any route that has been used recently within a time interval. The operation of AODV is loop free by use of sequence numbers which indicate the freshness of the route. When links break, AODV causes the affected set of nodes to be notified to invalidate the route. Route Request (RREQs), Route Replies (RREPs), and Route Errors (RRERs) are three message types defined by AODV for its working.

V. SIMULATION ENVIRONMENT

To evaluate and compare the performance of these routing protocols in Mobile Ad hoc network, we performed extensive simulations using NS-2 simulator [10]-[13]. Each simulator is carried out under the constant mobility.

TABLE 1
THE SIMULATION PARAMETERS

Experiment Parameter	Experiment Value	Description
Simulation Time	200 S	Simulation Duration
Terrain Dimension	1000*1000 m	X,Y Dimension of Motion
No. of Mobile Nodes	100	No. of nodes in a network
Node Placement	Random waypoint	Change Direction Randomly
Mobility Speed	0-50 mps	Mobility of Nodes
Mobility Model	Random	Mobility Direction
Routing Protocols	DSR, AODV	Path-finding
MAC protocol	802.11g	Wireless
Traffic	VBR	
Traffic rate	25 pkt/sec	
Packet Send rate	256kb	
Packet Size	1 kb	
Pause Time	100 sec	

Performance Metrics

Packet received [14]: It is the number of packets received by the application layer of destination nodes.

Throughput [14]: It is the average at which data packet is delivered successfully from one node to another over a communication network. It is usually measured in bits per second.
Throughput = (no of delivered packets * packet size) / total duration of simulation

Routing Overhead [14]: This is the total number of routing control packets generated by all nodes to the total data packets during the simulation time.

Network Load *[15]:* It is the total traffic received by the network layer from the higher MAC that is accepted and queued for transmission. It is measured as bits per second.

VI. Simulation Results And Performance Comparison

Performance of AODV and DSR routing protocols is evaluated under Variable Bit Rate multimedia traffic.

1) Packet Received

In the reactive protocols, AODV and DSR, AODV outperforms the DSR in terms of number of packets received by all destination nodes.

DSR	AODV
4569	45183

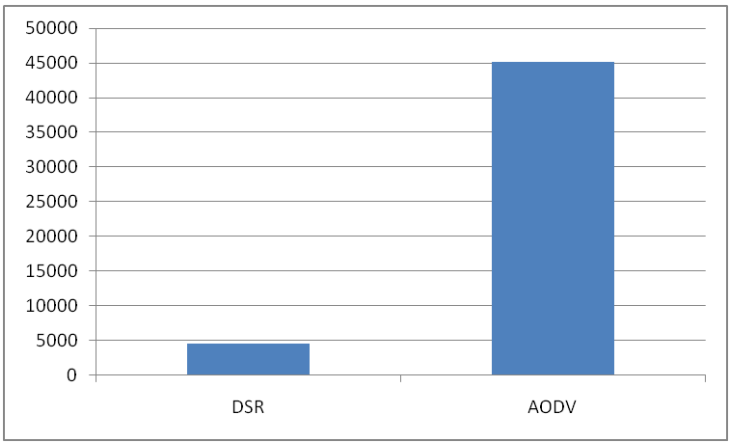


Figure 1: Number of Packet Received of AODV and DSR for 100 nodes

From the above figure, we have seen that the number of packets received in AODV protocol is very high than the number of packets received in DSR protocol for 100 nodes.

2) Throughput

Throughput is better for AODV protocol than DSR protocol.

DSR	AODV
103.85	262.59

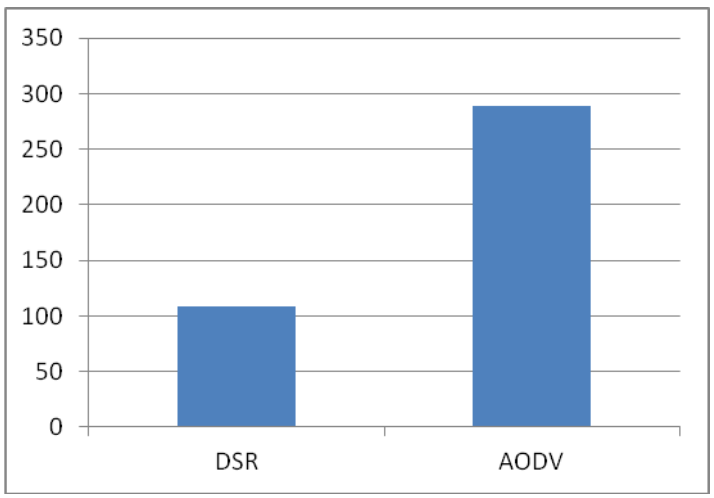


Figure 2: Throughput of AODV and DSR for 100 nodes

From above figure, it is observed that the AODV protocol outperforms the DSR protocol in terms of throughput when the number of nodes is 100.

3) Routing Overhead

Routing Overhead is higher in case of AODV than DSR.

DSR	AODV
37139	45285

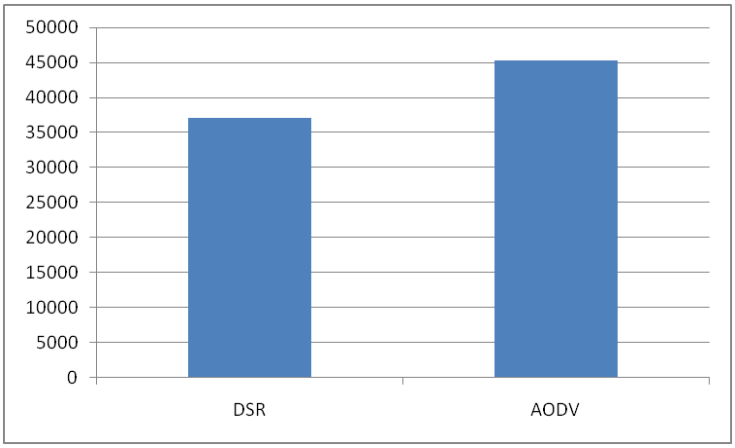


Figure 3: Routing Overhead for AODV and DSR for 100 nodes

It is observed from the figure above; in which AODV present the worse behavior in terms of routing overhead measurements than DSR for VBR traffic.

4) Network Load

Network Load is very high for DSR protocol than AODV protocol.

DSR	AODV
8.12	1.0

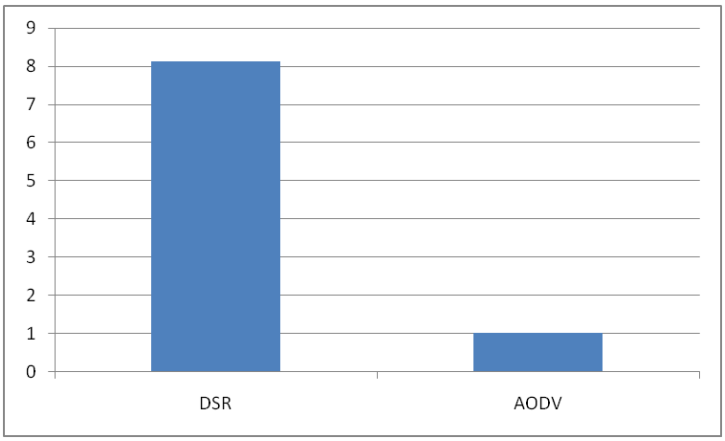


Figure 4: Network Load for AODV and DSR for 100 nodes

As shown in the above figure that DSR has higher Network Load than AODV. With the increase in hops the network overloads.

VII. CONCLUSION AND FUTURE WORK

Network simulator NS-2.34 is used to evaluate and compare the performance of AODV and DSR protocols under VBR traffic. The performance is compared in terms of number of packet received, throughput, routing overhead and network overload when number of nodes is constant. Simulation results show that in case of number of packets received AODV is ahead of DSR. And for throughput, AODV outperforms DSR. This shows that AODV has higher reliability. And high routing overhead in AODV for VBR data for 100 node is due to on demand route discovery for routing. The high network load of DSR worsens the performance of DSR as compared to AODV. This is due to source routing for route discovery and hop by hop packet forwarding in DSR. So, AODV is best suited for delay sensitive RTP applications which do not allow latency in the network. Thus from above investigations we conclude that AODV is a better option for VBR multimedia traffic.

In our future work, we will intend to do more simulations to evaluate more multimedia centric metrics for multimedia data transmission and to enhance the performance of DSR. In addition to it, we plans to investigate the performance of these protocols for congestion control mechanism by increasing the number of nodes for VBR traffic.

References

[1] Vasudha Arora and C. Rama Krishna, "Performance Evaluation of Routing protocols for MANETs under Different Traffic Conditions", 2nd IEEE International Conference on Computer Engineering and Information Technology, 2010.

[2] Vikas Singla and Parveen Kakkar, "Traffic Pattern based performance comparison of Reactive and Proactive Protocols of Mobile Ad-hoc Networks", International Journal of Computer Applications, Volume 5-No. 10, August 2010.

[3] Mobile Ad Hoc Networking: "An Essential Technology for Pervasive Computing", Jun-Zhao Sun Media Team, Machine Vision and Media Processing Unit.

[4] Krishna Moorthy Sivalingam, "Tutorial on Mobile Ad Hoc Networks", 2003.

[5] George Adam, Vaggelis Kapoulas, Chhristos Bouras, Georgios Kioumourtzis, Apostolos Gkamas and Nikos Tavoularis, "Performance Evaluation of Routing Protocols for multimedia transmission over mobile Ad hoc networks",IFIP WMNC'2011.

[6] Mehran Abolhasan, Tadeusz Wysocki, and Eryk Dutkiewicz, "A review of routing protocols for mobile ad hoc networks", Technical report, Telecommunication and Information Research Institute, University of Wollongong, Wollongong, NSW 2522; Motorola Australia Research Centre, 12 Lord St., Botany, NSW 2525, Australia,2003.

[7] Georgios Kiou Mourtzis, "Simulation and Evaluation of Routing Protocols for Mobile Ad hoc Networks", Master thesis in Computer Science and Engineering, Naval Postgraduate School, Monterey California, September, 2005.

[8] D.B Johnson, D.A Maltz, and Yih-Chun Hu., "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)", Internet draft (draft-ietf-manet-dsr- 10.txt), 19 July 2004.

[9] C. Perkins E. Belding-Royer, and S.Das, "Ad hoc On- Demand Distance Vector (AODV) Routing", RFC 3561, July2003.

[10] "The Network Simulator version 2", the source code of ns-allinone-2.34 can be downloaded from http://www.isi.edu/nsnam/ns/ns-build.html

[11] Kevin Fall, Kannan Varadhan, and the VINT project (May, 2010), available at http://www.isi.edu/nsnam/ns/ns-documentation.html

[12] Marc Gresis, "Tutorial for the network simulator (ns- 2)", available at http://www.isi.edu/nsnam/ns/tutorial/index.html

[13] NS by example available at http://nile.wpi.edu/NS

[14] Satveer Kaur, "Performance Comparison of DSR and AODV Routing Protocols with Efficient Mobility Model in Mobile Ad-Hoc Network", IJCST Vol. 2, Issue 2, June 2011.

[15] G.Rajkumar, R.Kasiram and D.Parthiban," Optimizing Throughput with Reduction in Power Consumption and Performance Comparison of DSR and AODV Routing Protocols", [ICCEET],2012.

Total Prime Graph

M.Ravi (a) Ramasubramanian[1], R.Kala[2]
1 Dept.of Mathematics, Sri Shakthi Institute of Engineering & Technology, Coimbatore – 641 062.
2 Dept. of Mathematics, Manonmaniam Sundaranar University, Tirunelveli – 627 012.

Abstract:
We introduce a new type of labeling known as "Total Prime Labeling". Graphs which admit a Total Prime labeling are called "Total Prime Graph". Properties of this labeling are studied and we have proved that Paths P_n, Star $K_{1,n}$, Bistar, Comb, Cycles C_n where n is even, Helm H_n, $K_{2,n}$, $C_3^{(t)}$ and Fan graph are Total Prime Graph. We also prove that any cycle C_n where n is odd is not a Total Prime Graph.

Keywords: Prime Labeling, Vertex prime labeling, Total Prime Labeling, Total Prime Graph

1. Introduction

By a graph G = (V,E) we mean a finite, simple and undirected graph. In a Graph G, V(G) denotes the vertex set and E(G) denotes the edge set. The order and size of G are denoted by 'p' and 'q' respectively. The terminology followed in this paper is according to [1]. A labeling of a graph is a map that carries graph elements to numbers. A complete survey of graph labeling is in [2]. Prime labeling and vertex prime labeling are introduced in [4] and [6]. Combining these two, we define a total prime labeling.

Two integers 'a' and 'b' are said to be relatively prime if their greatest common divisor is 1, (i.e.) $(a,b)=1$. If $(a_i,a_j)=1$, for all $i\neq j\ (1\le i,j\le n)$ then the numbers $a_1,a_2,a_3,\ldots,a_n$ are said to be relatively prime in pairs. Relatively prime numbers play a vital role in both analytic and algebraic number theory.

Definition 1.1 [4] Let G=(V,E) be a graph with 'p' vertices. A bijection $f:V(G)\to\{1,2,3,\ldots,p\}$ is said to be as "Prime Labeling" if for each edge e=xy the labels assigned to x and y are relatively prime. A graph which admits prime labeling is called "Prime Graph".

Definition 1.2 [6] Let G=(V,E) be a graph with 'p' vertices and 'q' edges. A bijection $f:E(G)\to\{1,2,3,\ldots,q\}$ is said to be a "Vertex Prime Labeling", if for each vertex of degree at least two, the greatest common divisor of the labels on its incident edges is 1.

2. Total Prime Graph:

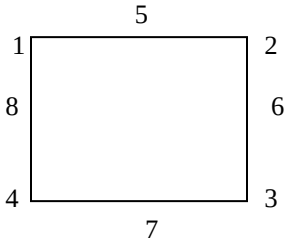
Definition 2.1 Let G=(V,E) be a graph with 'p' vertices and 'q' edges. A bijection $f:V\cup E\to\{1,2,3,\ldots,p+q\}$ is said to be a "Total Prime Labeling" if

(i) for each edge e=uv, the labels assigned to u and v are relatively prime.
(ii) for each vertex of degree at least 2, the greatest common divisor of the labels on the incident edges is 1.

A graph which admits "Total Prime Labeling" is called "Total Prime Graph".

Example 2.2

(1) C_4 is a Total Prime Graph.



(2) C_3 (or) K_3 is not a Total Prime Graph, because we can assign only one even label to an edge and one more even label to a vertex. But we have totally three even labels and the third even label can be assigned neither to any vertex not to any edge. Note that C_3 has Prime Labeling as well as Vertex Prime Labeling.

Notations 2.3

(1) Δ and δ denotes the maximum and minimum degree of a vertex respectively.

(2) $\lfloor n \rfloor$ denotes the greatest integer less than or equal to n.

(3) $\lceil n \rceil$ denotes the least integer greater than or equal to n.

(4) g.c.d denotes greatest common divisor.

Theorem 2.4 The path P_n is a Total Prime Graph.

Proof Let $P_n = v_1 v_2 v_3 \ldots v_n$. P_n has 'n' vertices and 'n-1' edges.

We define $f : V \cup E \to \{1, 2, 3, \ldots, (2n-1)\}$ as follows.

$$f(v_i) = i, 1 \le i \le n$$

$$f(e_j) = n + j, 1 \le j < n$$

Clearly f is a bijection.

According to this pattern, the vertices are labeled such that for any edge e=uv ∈ G, $\gcd [f(u), f(v)] = 1$.

Also the edges are labeled such that for any vertex v_i, the g.c.d of all the edges incident with v_i is 1.

Hence P_n is a Total Prime graph.

Definition 2.5 K_1 with 'n' pendent edges incident with $V(K_1)$ is called a Star Graph and is denoted by $K_{1,n}$.

Theorem 2.6 $K_{1,n}, (n > 1)$ is a Total Prime Graph.

Proof Let $V(K_1) = \{u\}$ and $v_i, 1 \le i \le n$ be the vertices adjacent to u.

Therefore $K_{1,n}$ has 'n+1' vertices and 'n' edges.

Now we define $f : V \cup E \to \{1, 2, 3, \ldots, (2n+1)\}$ as follows.

$$f(u) = 1$$

$$f(v_i) = 2i, 1 \le i \le n$$

$$f(e_j) = 2j + 1, 1 \le j \le n$$

Clearly f is a bijection.

According to this pattern, $K_{1,n}$ is a Total Prime Graph.

Definition 2.7 The graph obtained from $K_{1,n}$ and $K_{1,m}$ by joining their centers with an edge is called a Bistar and is denoted by B(n,m)

Theorem 2.8 Bistar B(n,m) is a Total Prime Graph.

Proof Let $V(K_2) = \{u, v\}$ and $u_i, 1 \le i \le n$; $v_i, 1 \le i \le m$ be the vertices adjacent to u and v respectively.

For $1 \le j \le n$, $e_j = uu_j$; $e_{n+1} = uv$; for $(n+2) \le j \le (n+m+1)$, $e_j = vv_j$.

Therefore B(n,m) has 'n+m+2' vertices and 'n+m+1' edges.

Now we define $f : V \cup E \to \{1, 2, 3, \ldots, (2n+2m+3)\}$ as follows.

$$f(u)=1$$
$$f(v)=2$$
$$f(u_i)=2(i+1),1\le i\le n$$
$$f(v_i)=2i+1,1\le i\le m$$
$$f(e_j)=(n+m+2)+j,1\le j\le(n+m+1)$$

Clearly f is a bijection.

According to this pattern, clearly B(n,m) is a Total Prime Graph.

Definition 2.9 A graph obtained by attaching a single pendent edge to each vertex of a path $P_n = v_1v_2v_3\ldots v_n$ is called a Comb.

Theorem 2.10 Comb is a Total Prime Graph.

Proof Let G be a Comb obtained from the path by joining a vertex u_i to v_i, $1\le i\le n$.

The edges are labeled as follows:

For $1\le i\le n$, $e_{2i-1}=v_iu_i$ and $e_{2i}=v_iv_{i+1}$

Therefore G has '2n' vertices and '2n-1' edges.

Now define $f:V\cup E\to\{1,2,3,\ldots,(4n-1)\}$ as follows.

$$f(v_i)=2i-1,1\le i\le n$$
$$f(u_i)=2i,1\le i\le n$$
$$f(e_j)=2n+j,1\le j\le(2n-1)$$

Clearly f is a bijection.

According to this pattern, Comb is a Total Prime Graph.

Theorem 2.11 Cycle C_n, n is even, is a Total Prime Graph.

Proof Let $C_n=(v_1e_1v_2e_2v_3\ldots v_ne_nv_1)$

Therefore C_n has 'n' vertices and 'n' edges.

Now we define $f:V\cup E\to\{1,2,3,\ldots,2n\}$ as follows.

$$f(v_i)=i,1\le i\le n$$
$$f(e_j)=n+j,1\le i\le n$$

Clearly f is a bijection.

According to this pattern, clearly Cycle C_n, n is even, is a Total Prime Graph.

Theorem 2.12 Cycle C_n, n is odd, is not a Total Prime Graph.

Proof Let $C_n=(v_1e_1v_2e_2v_3\ldots v_ne_nv_1)$

Therefore C_n has 'n' vertices and 'n' edges.

Define $f:V\cup E\to\{1,2,3,\ldots,2n\}$

Now, no. of even labels available is 'n'.

For any 3 consecutive vertices, we can assign at most one even label and so, number of vertices with even labels is at most $\left\lceil\frac{n}{3}\right\rceil$.

Also, out of 3 consecutive edges, we can assign at most one even label and so the number of edges with even labels is at most $\left\lceil \frac{n}{3} \right\rceil$.

Therefore the necessary condition for existence of total Prime Graph is $2\left\lceil \frac{n}{3} \right\rceil$=n.

Case 1: $n \equiv 0 \pmod 3$

(i.e.) n is a multiple of 3.

Therefore, in this case $2\left(\frac{n}{3}\right) = n$

(i.e.) $2n = 3n$

(i.e.) $2 = 3$

Which is a contradiction.

Case 2: $n \equiv 1 \pmod 3$

In this case $2\left(\frac{n+2}{3}\right) = n$

(i.e.) $2n + 4 = 3n$

(i.e.) $n = 4$

But n is odd, so it's not possible.

Case 3: $n \equiv 2 \pmod 3$

In this case $2\left(\frac{n+1}{3}\right) = n$

(i.e.) $2n + 2 = 3n$

(i.e.) $n = 2$

But n is odd, so it's not possible.

Therefore Cycle C_n, n is odd, is not a Total Prime Graph.

Definition 2.13 Helm H_n is a graph obtained from wheel by attaching a pendent edge at each vertex of n-cycle.

Theorem 2.14 Helm H_n is a Total Prime Graph.

Proof Here center vertex will be labeled as u and all the vertices on the cycle are labeled as $u_1, u_2, \ldots, u_n$. The corresponding pendent vertices are labeled as $v_1, v_2, \ldots, v_n$. The edges are labeled as $e_1, e_2, \ldots, e_{2n}$ starting from the pendent edge incident at vertex u_1 and with labeling the edge on the cycle alternatively in clockwise direction $e_1, e_2, \ldots, e_{2n}$ and the spokes of the wheels are labeled as $e_{2n+1}, e_{2n+2}, \ldots, e_{3n}$ starting from the edge uu_1 and proceeding in the clockwise direction.

Therefore Helm H_n has '2n+1' vertices and '3n' edges.

Now we define $f : V \cup E \to \{1, 2, 3, \ldots, (5n+1)\}$ as follows.

$$f(u) = 1$$

$$f(u_i) = 2i + 1, 1 \le i \le \left\lfloor \frac{2n+1}{2} \right\rfloor$$

$$f(v_i) = 2i, 1 \le i \le \left\lfloor \frac{2n+1}{2} \right\rfloor$$

$$f(e_j) = (2n+1) + j, 1 \le j \le 3n$$

Clearly f is a bijection.

According to this pattern, clearly Helm H_n is a Total Prime Graph.

Definition 2.15 $K_{m,n}$ is a complete bipartite graph with bipartition X and Y, in which any two vertices in X as well as any two vertices in Y are non-adjacent. Also every vertex of X is adjacent to every vertex of Y.

Theorem 2.16 $K_{2,n}$, is a Total Prime Graph.

Proof $K_{m,n}$ have 'm+n' vertices and 'mn' edges. Here m=2. Therefore $K_{2,n}$ has '2+n' vertices and '2n' edges.

Let $X=\{u_1,u_2\}$ and $Y=\{v_1,v_2,v_3,\ldots,v_n\}$. The edges are labeled in a continuous manner starting from $e_1=v_1u_1$ to $e_{2n-1}=v_nu_1$ and the last edge $e_{2n}=v_1u_2$.

Now we define $f:V\cup E\to\{1,2,3,\ldots,(3n+2)\}$ as follows:

$$f(u_i)=2i-1, 1\le i\le 2$$

The vertices $Y=\{v_1,v_2,v_3,\ldots,v_n\}$ are partitioned into $\left\lfloor\frac{n}{2}\right\rfloor$ sets as follows:

for $j=even$ and $0\le j\le\left\lceil\frac{n}{2}\right\rceil$, let $S_j=\{v_{j+1},v_{j+2}\}$

$$f(v_i)=2i+j, 1\le i<n \text{ and } v_i\in S_j$$

$$f(v_n)=\begin{cases}3n+1, & \text{n is odd}\\ 3n+2, & \text{n is even}\\ 3n+1, & \text{n is of the form 10r-2 and r=1,2,3,}\ldots\end{cases}$$

The edges are labeled as follows:-

Case 1: n is odd

(i) for $0\le k\le\left\lceil\frac{n}{2}\right\rceil-1$, $e_{4k+1}=u_1v_{\left\lceil\frac{4k+1}{2}\right\rceil}$

$$e_{4k+2}=u_1v_{2k+2}$$

$$e_{4k+3}=u_2v_{\left\lceil\frac{4k+3}{2}\right\rceil}$$

(ii) for $0\le k\le\left\lceil\frac{n}{2}\right\rceil-2$, $e_{4k+4}=u_2v_{2k+3}$

(iii) $e_{2n-2}=u_2v_n$, $e_{2n-1}=u_1v_n$, $e_{2n}=u_2v_1$

Case 2: n is even

(i) for $0\le k\le\frac{n}{2}-1$, $e_{4k+1}=u_1v_{\left\lceil\frac{4k+1}{2}\right\rceil}$

$$e_{4k+2}=u_1v_{2k+2}$$

$$e_{4k+3}=u_2v_{\left\lceil\frac{4k+3}{2}\right\rceil}$$

(ii) for $0\le k\le\frac{n}{2}-2$, $e_{4k+4}=u_2v_{2k+3}$

(iii) $e_{2n}=u_2v_1$

The unassigned labels given in their order to the edges in the order $\{e_1, e_2, e_3, \ldots, e_{2n}\}$.

Clearly f is a bijection.

According to this pattern, clearly $K_{2,n}$, is a Total Prime Graph.

Definition 2.17 $C_3^{(t)}$ denotes the one-point union of 't' cycles of length 3. $C_3^{(t)}$ is also called as Friendship Graph (or) Dutch t-windmill.

Theorem 2.18 $C_3^{(t)}$ is a Total Prime Graph.

Proof $C_3^{(t)}$ has '2t+1' vertices and '3t' edges.

Let the vertex set be $\{v_0, v_1, v_2, \ldots, v_{2t}\}$ with centre vertex v_0. Let the edge set be $\{e_1, e_2, e_3, \ldots, e_{3t}\}$ with $e_1 = v_0 v_1$ and label the edges in clockwise direction.

Now we define $f : V \cup E \to \{1, 2, 3, \ldots, (5t+1)\}$ as follows:

$$f(v_i) = i+1, 0 \le i \le 2t+1$$

$$f(e_j) = (2t+1) + j, 1 \le j \le 3t$$

Clearly f is a bijection.

According to this pattern, clearly $C_3^{(t)}$ is a Total Prime Graph.

Definition 2.19 The fan graph F_n is defined as $K_1 + P_n$, P_n is a path of n vertices.

Theorem 2.20 Fan graph $F_n, n \ge 3,$ is a Total Prime Graph.

Proof F_n has $n+1$ vertices and '2n-1' edges.

We define $f : V \cup E \to \{1, 2, 3, \ldots, 3n\}$ as follows.

$$f(v_i) = i, 1 \le i \le n$$

For $e_i = v_i v_{i+1}$, $1 \le i \le n$

$$f(e_i) = n+1+i$$

For $e_{n+j-2} = v_1 v_j$, $3 \le j \le n+1$

$$f(e_{n+j-2}) = 3n+3-j$$

Clearly f is a bijection.

According to this pattern, Fan Graph F_n is a Total Prime Graph.

References

[1]. F.Harary, *Graph Theory*, Addison Wesley, Reading, Mass., 1972.
[2]. J.A.Gallian , *A dynamic survey of graph labeling*, Elctronic J.Combinatorics, (January 2010).
[3]. T.M.Apostol, *Introduction to Analytic Number Theory*, Narosa Publishing House, 1998.
[4]. A.Tout, A.N. Dabbouey and K.Howalla, *Prime Labeling of graphs*, National Academy Science Letters, 11(1982), 365-368.
[5]. Fu,H.L and Huang,K.C (1994) on *Prime Labeling* Discrete mathematics, North Holland, 127, 181-186..
[6]. T.Deretsky, S.M.Lee and J.Mitchem, on *Vertex prime labeling of graphs in graph theory*, Combinatorics and Applications Vol.1, J.Alavi, G.Chartrand and O.Ollerman and A.Schwenk, eds…, Proceedings 6th international conference Theory and Application of Graphs (Wiley, Newyork,1991) 359-369.

A Study on Prosody Analysis

Padmalaya Pattnaik [1], Shreela Dash [2]
(Asst.Prof, C.V. Raman College of Engineering, Bhubaneswar, Odisha, India)

Abstract:
Speech can be described as an act of producing voice through the use of the vocal folds and vocal apparatus to create a linguistic act designed to convey information. Linguists classify the speech sounds used in a language into a number of abstract categories called phonemes. Phonemes are abstract categories, which allow us to group together subsets of speech sounds. Speech signals carry different features, which need detailed study across gender for making a standard database of different linguistic, & paralinguistic factors. Prosodic phenomena are specific to spoken language. They concern the way in which speech sounds are acoustically realized: how long they are, how high and how loud. Such acoustic modulations are used by human speakers to express a variety of linguistic or paralinguistic features, from stress and syntactic boundaries, to focus and emphasis or pragmatical and emotional attitudes. Linguistics and speech technology have approached prosody from a variety of points of view, so that a precise definition of the scope of prosodic research is not easy. A main distinction can be drawn between acoustic-phonetic analyses of prosody and more abstract, linguistic, phonological approaches When people interact with others they convey emotions. Emotions play a vital role in any kind of decision in affective, social or business area. The emotions are manifested in verbal, facial expressions but also in written texts. The objective of this study is to verify the impact of various emotional states on speech prosody analysis.

Keywords: *Duration, Emotion, Jitter, Prosody, Shimmer*

1. Introduction

No language is produced in a smooth, unvarying stream. Rather, the speech has perceptible breaks and clumps. For example, we can perceive an utterance as composed of words, and these words can be perceived as composed of syllables, which are composed of individual sounds. At a higher level, some words seem to be more closely grouped with adjacent words: we call these groups phrases. These phrases can be grouped together to form larger phrases, which may be grouped to form sentences, paragraphs, and complete discourses. These observations raise the questions of how many such constituents there are and how they are best defined. A fundamental characteristic of spoken language is the relation between the continuous flow of sounds on the one hand, and the existence of structured patterns within this continuum on the other hand. In this respect, spoken language is related to many other natural and man-made phenomena, which are characterized not only by their typically flowing nature but also by the fact that they are structured into distinct units such as waves and measures. Prosodic phonology is a theory of the way in which the flow of speech is organized into a finite set of phonological units. It is also, however, a theory of interactions between phonology and the components of the grammar. Although many speech interfaces are already available, the need is for speech interfaces in local Indian languages. Application specific Indian language speech recognition systems are required to make computer aided teaching, a reality in rural schools. This paper presents the preliminary work done to demonstrate the relevance of an Oriya Continuous Speech Recognition System in primary education. Automatic speech recognition has progressed tremendously in the last two decades. There are several commercial Automatic Speech Recognition (ASR) systems developed, the most popular among them are Dragon Naturally Speaking, IBM Via voice and Microsoft SAPI. Speech is a complex waveform containing verbal (e.g. phoneme, syllable, and word) and nonverbal (e.g. speaker identity, emotional state, and tone) information. Both the verbal and nonverbal aspects of speech are extremely important in interpersonal communication and human-machine interaction. Each spoken word is created using the phonetic combination of a set of vowel semivowel and consonant speech sound units. Different stress is applied by vocal cord of a person for particular emotion. The increased muscle tension of the vocal cords and vocal tract can directly or indirectly and adversely affect the quality of speech. We use emotions to express and communicate our feelings in everyday life. Our experience as speakers as well as listeners tells us that the interpretation of meaning orintention of a spoken utterance can be affected by the emotions that area expressed and felt.

2. Literature

According to the classic definition, prosody has to do with speech features whose domain is not a single phonetic segment, but larger units of more than one segment, possibly whole sentences or even longer utterances. Consequently, prosodic phenomena are often called supra-segmentals. They appear to be used to structure the speech flow and are perceived as stress or accentuation, or as other modifications of intonation, rhythm and loudness.

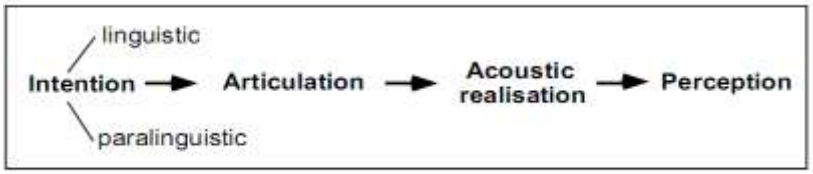


Fig: 1 Stages of oral communication

An emotion is a mental and physiological state associated with a wide variety of feelings, thoughts, and internal (physical) or external (social) behaviors. Love, hate, courage, fear, joy, sadness, pleasure and disgust can all be described in both psychological and physiological terms. An emotion is a psychological arousal with cognitive aspects that depends on the specific context. According to some researcher, the emotions are cognitive processes. Emotion is a process in which the perception of a certain set of stimuli, follows cognitive assessment which enables people to label and identify a particular emotional state. At this point there will be an emotional physiological, behavioral and expressive response. For example, the primordial fear, that alerts us as soon when we hear a sudden noise, allows to react to, dangerous situations and provides instantly resources to face them as escape or close the door. The emotional stimuli may be an event, a scene, a face, a poster, an advertising campaign. These events, as a first reaction, put on alert the organism with somatic changes as heart rate, increase of sweat, acceleration of respiratory rhythm, rise of muscle tensions.

Emotions give an immediate response that often don't use cognitive processes and conscious elaboration and sometimes they have an effect on cognitive aspects as concentration ability, confusion, loss, alert and so on. This is what is asserted in evaluation theory, in which cognitive appraisal is the true cause of emotions [2]. Two factors that emerge permanently are those related to signals of pleasure and pain and characterizing respectively the positive and negative emotions. It's clear that these two parameters alone are not sufficient to characterize the different emotions. Many authors debate on primary and secondary emotions other on pure and mixed emotions, leaving the implication that emotions can somehow be composed or added.

The systems based on the analysis of physiological response as blood pressure, heart rate, respiration change present an initial phase where the signals are collected in configurations to be correlated with different emotional states and a subsequently recognition basing on the measure of indicators. One of the interesting early works on the emotions was that one of Ortony [3]. From this work, through componential analysis, other authors constructed an exhaustive taxonomy on affective lexicon. According to Ortony, stimuli that cause emotional processes are of three basic types: events, agents and objects corresponding to three classes of emotions: satisfied/ unsatisfied (reactions to events), approve/disapprove(reaction to agents), appreciate/unappreciate (reaction to objects). According to Osgood [4] an emotion consists of a set of stages: stimulus (neural and chemical changes), appraisal and action readiness. Continuing the studies of Charles Darwin, the Canadian psychologist Paul Ekman [5] has confirmed that an important feature of basic emotions is that they are universally expressed, by everybody in any place, time and culture, through similar methods. Some facial expressions and the corresponding emotions are not culturally specific but universal and they have a biological origin. Ekman, analyzed how facial expressions respond to each emotion involving the same type of facial muscles and regardless of latitude, culture and ethnicity. This study was supported by experiments conducted with individuals of Papua New Guinea that still live in a primitive way.

3. Emotions

Human emotions are deeply joined with the cognition. Emotions are important in social behavior and to stimulate cognitive processes for strategies making. Emotions represent another form of language universally spoken and understood. Identification and classification of emotions has been a research area since Charles Darwin's age. In this section we consider facial, vocal and textual emotional expressions.

Emotion classifications of the researchers differ according to the goal of the research and the field .Also the scientist's opinion about the relevance of dividing different emotions is important. There is no standard list of basic emotions. However, it is possible to define list of emotions which have usually been chosen as basic, such as: erotic(love)(shringar), pathetic (sad) (karuNa), wrath (anger)(roudra), quietus (shAnta), normal(neutral).

The objective of this study is to analyze impact for different emotions on vowels in terms of certain parameters for stage prosody analysis.

4. Prosody

In linguistics, prosody is the rhythm, stress, and intonation of speech. Prosody may reflect various features of the speaker or the utterance: the emotional state of the speaker; the form of the utterance (statement, question, or command); the presence of irony or sarcasm; emphasis, contrast, and focus; or other elements of language that may not be encoded by grammar or choice of vocabulary. Prosody has long been studied as an important knowledge source for speech understanding and also considered as the most significant factor of emotional expressions in speech[16]. Prosody gives

naturalness and message intelligibility to speech Emotional prosody is the expression of feelings using prosodic elements of speech. Linguistically relevant prosodic events concur to express sentence structure: they highlight linguistic units by marking their boundaries and suggesting their function. Several types of prosodic units (differing mainly in their scope) have been proposed: paragraphs, sentences, intonation groups, intermediate groups, stress groups, syllables etc. Although prosody is by definition suprasegmental, prosodic analyses take often the phoneme as their minimal unit, where to measure rhythmical variations and locate intonation events. The family of prosodic phenomena includes the suprasegmental features of intonation, stress, rhythm and speech rate, whose variations are relevant to express the function of the different prosodic units: the prominent syllable in the word will be marked by stress, a falling intonation contour will mark the conclusion of a sentence, a faster speech rate and lower intonation characterize a parenthetical phrase.Such prosodic features are physically realized in the speech chain in terms of variations of a set of acoustic parameters. Acoustic-phonetic analyses identify the following 'phonetic correlates of prosody': fundamental frequency/pitch (f0), length changes in segmental duration, pauses, loudness, voice quality.

Prosody is the combination of voice's pitch, duration and energy variation during speech. It provides an additional sense to the words, which is extraordinarily important in natural speech. For example, interrogative and declarative sentences have very different prosody (especially intonation). Besides, the prosody of a sentence is one of the factors that make a speaker seem happy, sad, angry or frightened. We can even decide from prosody if the speaker is an energetic person or, on the contrary, a lazy one. When singing, intonation and timing evolution characterize melody. But prosody is not only important in natural speech but also in synthetic speech. Prosody is crucial in order to achieve an acceptable naturalness. If a TTS system does not have a good prosodic treatment, its output speech sounds completely monotonous and, moreover, it won't be able to distinguish between sentences of different kinds. The two main parameters of prosody are-

4.1 Intonation

In a first approximation, sounds can be classified invoiced and unvoiced sounds. Voiced sounds, unlike unvoiced ones, are produced making the vocal chordsvibrate. These vibrations provoke some periodicitiesin the speech signal and therefore a fundamental frequency (F0). This value is inversely proportional tothe distance between periodicities and it makesspeech sound with higher or lowers frequency. It iscommonly called pitch. On the contrary, as unvoiced sounds do not have any periodicities (vocal chords donot vibrate) and can be modeled as a filtered noise signal. So if we detect the pitch curve of a speech signal it will only exist in the voiced segments. Pitch is not constant but its value changes during asentence. That is called intonation. Thanks tointonation we can distinguish, for example, between adeclarative and an interrogative sentence or identifyfocused words inside a sentence.

4.2 Duration

The duration of speech segments is the other main parameter of prosody. The timing structure of a sentence is extremely important to give naturalness to speech. Phone duration depends on a great number of parameters, such as its phonetic identity, surroundingphones, level of stress or position in the sentence or in the word. What's more, duration of a word also depends on its importance in the sentence. For example, a focused word will generally have longer duration.

5. Prosody Analysis

TTS systems generate speech from a text. There is a need of prosodic assignment to phones to produce high quality speech. Once the phones to synthesize are determined, it is necessary to know the pitch and duration yet achieved the required quality needed for most applications. This is the main reason why the tool was created: the prosodic module requires models of prosodic patterns and these patterns have to be studied and tested before the application to TTS.The goal of the prosodic analysis is the generation of the pitch contour[17].

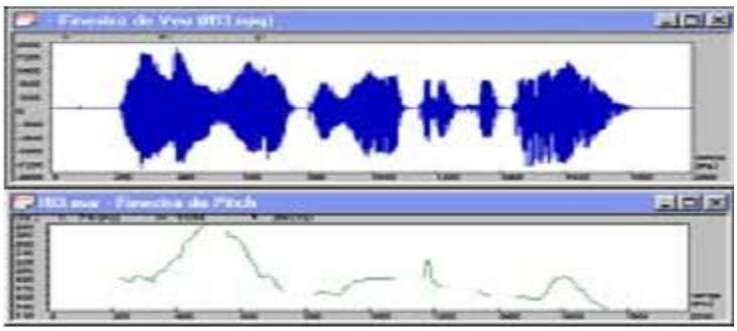


Fig: 2 Speech Segment & it's pitch contour

6. Parametric Measurements of Acoustics Signals for Prosody Analysis

Four acoustic signal features such as Duration of speech segment, Pitch, Jitter and Simmer were used to parameterize the speech.

6.1 Duration

Utterance durations, vowel durations were measured from the corresponding label files produced by a manual segmentation procedure. On an average, utterance durations become longer when speech is emotionally elaborated.

6.2 Fundamental Frequency (pitch)

We calculated the pitch contours of each utterance using speech processing software. Global level statistics related to F0 such as minimum, maximum, mean were calculated from smoothed F0 contours.

6.3 Jitter & Shimmer

Jitter and Shimmer are related to the micro-variations of the pitch and power curves. In other words, Shimmer and Jitter are the cycle-to-cycle variations of waveform amplitudes and fundamental periods respectively. The Jitter & Shimmer occur due to some undesirable effect in audio signal. Jitter is the period frequency displacement of the signal from the ideal location. Shimmer is the deviation of amplitude of the signal from the ideal location.

7. Methodology and Experimentation

There are many features available that may be useful for classifying speaker affect: pitch statistics, short-time energy, long-term power spectrum of an utterance, speaking rate, phoneme and silence durations, formant ratios, and even the shape of the glottal waveform [8, 11, 12, 9, 10]. Studies show, that prosody is the primary indicator of a speaker's emotional state [1, 13, 12]. We have chosen to analyze prosody as an indicator of affect since it has a well-defined and easily measureable acoustical correlate -- the pitch contour. In order to validate the use prosody as an indicator for affect and to experiment with real speech, we need to address two problems: First, and perhaps most difficult, is the task of obtaining a speech corpus containing utterances that are truly representative of an affect. Second, what exactly are the useful features of the pitch contour in classifying affect? Especially as many factors influence the prosodic structure of an utterance and only one of these is speaker's emotional state [6, 7, 9].

The data analyzed in this study were collected from semi-professional actors and actress and consists of 30 unique Odia language sentences that are suitable to be uttered with any of the five emotions i.e., roudra, shringar, shanta, karuna, and neutral. Some example sentences are "Jayanta jagi rakhhi kaatha barta kaara", "*Babu chuata mora tinni dina hela khaaini*". The recordings were made in a noise free room using microphone. For the study, samples have been taken from three male & three female speakers. The utterances are recorded at the bit rate of 22,050Hz. The Vowels are extracted from the words consisting of 3 parts i.e.CV, V, VC. CV stands for Consonant to Vowel transition, V for steady state vowel, VC for Vowel to Consonant transition.

8. Experimental Results

For experimental analysis data samples were created from recordings of male and female speakers in various emotions (mood). Vowels are then extracted and stored in a database for analysis. From this database after analysis the result of the utterance. Duration, fundamental frequency & variations of pitch of every vowel are measured and compared to give following results.

8.1 Duration

It is observed from the duration Table 1 that speech associated with the emotion "love(Shringar)"has higher duration with the vowel /i/ gets elongated both for male and female speakers where as the emotion "sad (karuna)" for male speakers vowel(/a/,/i/) gets elongated whereas for female (/i/,/u/) gets elongated.

Table (1): Average duration of vowels for speakers in different emotions

Emotions	Duration in Millie Seconds (male-average)				
	/a/	/i/	/u/	/e/	/o/
Neutral	67	70	60	39	53
Santa	58	75	60	58	43
Karuna	**106**	**106**	61	56	54
Raoudra	64	50	43	49	56
Shringar	**100**	**113**	**90**	54	56
Emotions	Duration in Millie Seconds (female-average)				
	/a/	/i/	/u/	/e/	/o/
Neutral	50	64	66	40	35
Santa	53	**99**	79	43	50
Karuna	**83**	**105**	**101**	**86**	63
Raoudra	50	43	58	38	41
Shringar	**80**	**161**	**118**	55	54

8.2 Fundamental Frequency

Figure3 shows the analysis result that the mean pitch for male speaker associated with emotion karuna for vowel /i/ has dominance where as in female the speech associated with emotion shringar for vowel /a/ plays dominant role.

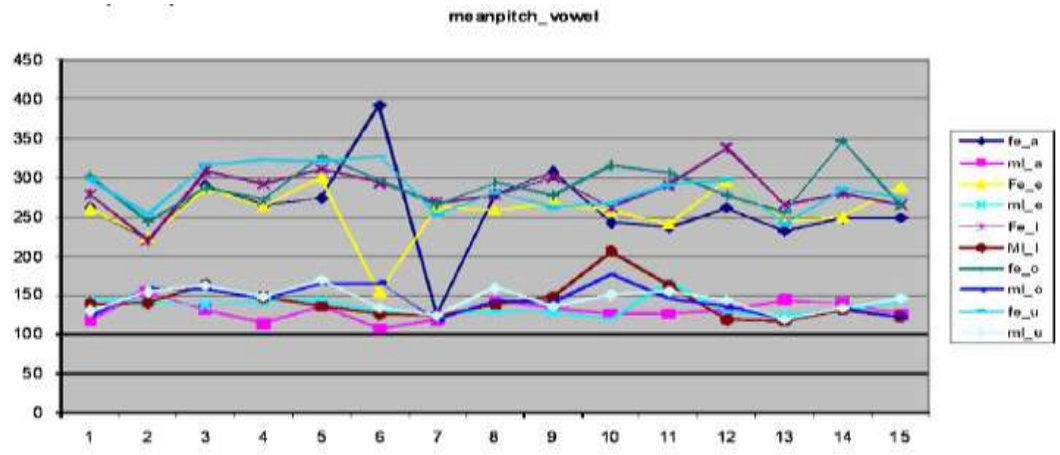


Fig: 3 mean pitch of vowels of speakers in different emotion

8.3 Jitter

Figure 4 show that the emotion "anger" of vowel /i/ is having a dominant role in jitter for male speaker. The jitter value for female speaker has dominant vowel /u/ for the emotion "love".

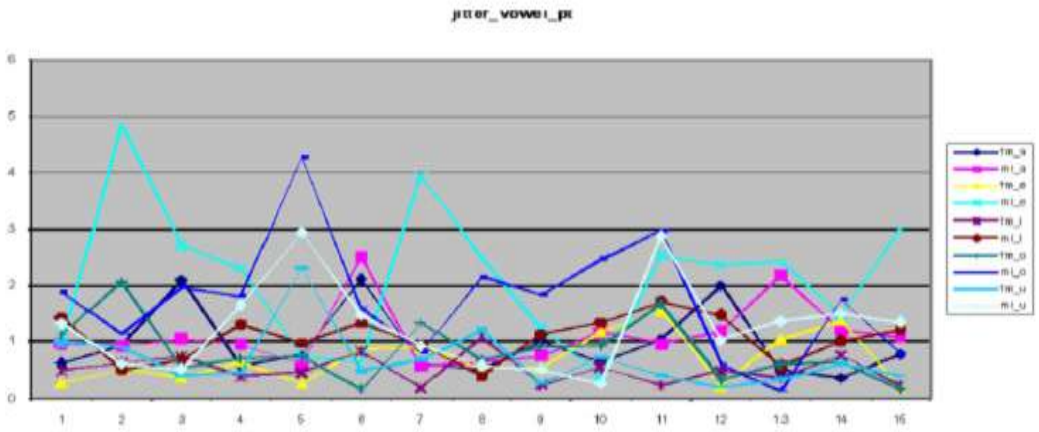


Fig:4 jitter of vowels of speakers in different emotion

8.4 Shimmer

It is observed from figure 5 that the shimmer in the emotion "anger" of vowel /o/ has dominant role for males & for female in emotion "anger" of vowel /i/ has dominant role.

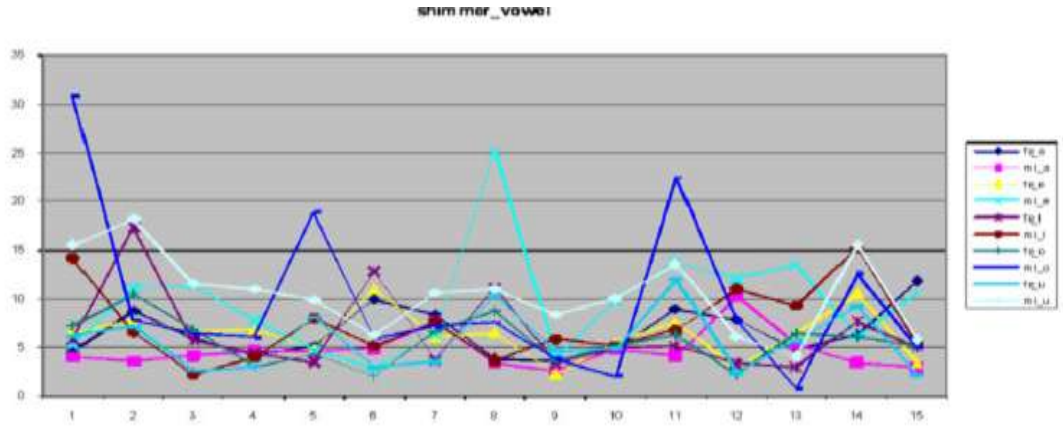


Fig:5 shimmer of vowels of speakers in different emotion

9. Conclusion

The great importance and complexity of prosody in speech makes this subject an important area of research in speech synthesis applications In this study, we investigate acoustic properties of speech prosody associated with five different emotions (love (shringar)), pathetic (sad) (karuNa), wrath (anger) (roudra), quietus (shAnta),normal (neutral) intentionally expressed in speech by male and female speakers. Results show speech associated with love (shringar) and sad (karuna) emotions are characterized by longer utterance duration, and higher pitch. However we observed that for jitter anger or love has dominance over others, whereas for shimmer the emotion anger plays a vital role. Future works of my research are the following. We have to collect synthetic speech and put emotion labels on them. We have to reconsider how to estimate emotion in speech using parallel programming.

References

[1] P. Olivier and J. Wallace, Digital technologies and the emotional family, International Journal of Human Computer Studies, 67 (2), 2009, 204-214.

[2] W. L. Jarrold, Towards a theory of affective mind: computationally modeling the generativity of goal appraisal, Ph.D. diss., University of Texas, Austin, 2004.

[3] C. G. Ortony and A. Collins, The cognitive structure of emotions, (Cambridge University Press: New York, 1990).

[4] M. M. C.E. Osgood and W.H. May, Cross-cultural Universals of Affective Meaning, (Urbana Chaimpaigne: University of Illinois Press, 1975).

[5] E. Paul, Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life, (NY: OWL Books, 2007).

[6] A. Ichikawa and S. Sato, Some prosodical characteristics in spontaneous spoken dialogue, International Conference on Spoken Language Processing, v. 1, 1994, 147-150.

[7] R. Collier, A comment of the prediction of prosody, in G. Bailly, C. Benoit, and T.R. Sawallis (Ed.), Talking Machines: Theories, Models, and Designs, (Amsterdam: Elsevier Science Publishers, 1992).

[8] H. Kuwabara and Y. Sagisaka, Acoustic characteristics of speaker individuality: Control and conversion, Speech Communication, 16(2), 1995, 165-173.

[9] K. Cummings and M. Clements, Analysis of the glottal excitation of emotionally styled and stressed speech, Journal of the Acoustical Society of America, 98(1), 1995, 88-98.

[10] D. Roy and A. Pentland, Automatic spoken affect classification and analysis, Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition, 1996, 363-367.

[11] A. Protopapas and P. Lieberman, Fundamental frequency of phonation and perceived emotional stress, Journal of Acoustical Society of America, 101(4), 1997, 2267-2277.

[12] X.Arputha Rathina and K.M.Mehata, Basic analysis on prosodic features in emotional speech, nternational Journal of Computer Science, Engineering and Applications (IJCSEA) , Vol.2, No.4, August 2012,99-107

[13] D. Hirst, Prediction of prosody: An overview, in G. Bailly, C. Benoit, and T.R. Sawallis (Ed.), Talking Machines: Theories, Models, and Designs, (Amsterdam: Elsevier Science Publishers, 1992).

[14] L. Rabiner and R. Shafer, Digital Processing of Speech Signals, (New York: Wiley and Sons, 1978)

[15] Wavesurfer, http://www.speech.kth.se/wavesurfer/

[16] Masaki Kuremastsu et eal, An extraction of emotion in human speech using speech synthesize and classifiers foe each emotion , International journal of circuits systems and signal processing, 2008.

[17] J .L. Navarro, I. Esquerra, A Time-FrequencyApproach to Epoch Detection, Proceedings ofEurospeech'95, 405-408, Madrid 1995.

Generalized PWM algorithm for Direct Torque Controlled Induction Motor Drives using the only Sampled Voltages

J.Bhavani[1], J.Amarnath[2], D.Subbarayudu[3]

1Associate professor, EEE Department, Malla Reddy Engineering college, Maisammaguda, Hyderabad, Andhra Pradesh, India.
2J.Amarnath, Professor, EEE Department, JNTUH College of Engineering, Hyderabad, Andhra Pradesh, India.
3D.subba Rayudu, Professor, EEE Department, Pulla Reddy Engineering College, Kurnool, A.P.

***Abstract*—**This paper presents a simple generalized pulse width modulation (GPWM) algorithm for direct torque controlled induction motor drives. Though the classical direct torque control (DTC) algorithm gives fast dynamic response, it gives large steady state ripples and variable switching frequency operation of the inverter. To overcome these problems, space vector pulse width modulation (SVPWM) algorithm is used for DTC drives. But, as the SVPWM algorithm requires angle and sector information the complexity involved is more. Also, it gives more switching losses of the inverter. To reduce the complexity involved the proposed algorithm uses the instantaneous phase voltages only. Moreover, the proposed approach gives the realization of various carrier based PWM algorithms that include both SVPWM and various discontinuous PWM (DPWM) algorithms by using a generalized control algorithm and is obtained via unequal sharing of zero states. In the proposed approach, by varying a constant (k) value from zero to one, various DPWM algorithms can be generated along with the SVPWM algorithm. To validate the proposed GPWM algorithm based DTC drive, several numerical simulation studies have been carried out and the results have been presented. The simulation results show the effectiveness of the proposed algorithm.

Index Terms—DPWM, DTC, GPWM, Induction motor drive, SVPWM.

I. INTRODUCTION

THE variable speed drives (VSDs) are becoming popular in many industrial applications. The invention of field oriented control (FOC) algorithm has been made a renaissance in the high-performance variable speed drive applications. The FOC algorithm gives the decoupling control of torque and flux of an induction motor drive and control the induction motor similar to a separately excited dc motor [1]. But, the complexity involved in the FOC algorithm is more due to reference frame transformations. To reduce the complexity Involved, a new control strategy called as direct torque control (DTC) has been proposed in [2]. A detailed comparison between FOC and DTC is presented in [3] and concluded that DTC gives fast torque response when compared with the FOC. Though, FOC and DTC give fast transient and decoupled control, these operate the inverter at variable switching frequency due to hysteresis controllers. Moreover, the steady state ripples in torque, flux and currents are high in DTC.

To reduce the harmonic distortion and to obtain the constant switching frequency operation of the inverter, nowadays many researchers have focused their interest on pulsewidth modulation (PWM) algorithms. A large variety of PWM algorithms have been discussed in [4]. But, the most popular PWM algorithms as sinusoidal PWM (SPWM) and space vector PWM (SVPWM) algorithms. The SVPWM algorithm offers more degrees of freedom when compared with the SPWM algorithms. Hence, it is attracting many researchers. The SVPWM algorithm is explained in detailed in [5]. Though the SVPWM and SPWM algorithms give good performance, these give more switching losses of the inverter due to the continuous modulating signals. Hence, to reduce the switching losses of the inverter, the discontinuous PWM (DPWM) algorithms are becoming popular. Also, the classical SVPWM algorithm requires angle and sector information to generate the actual gating times of the inverter. Hence, the complexity involved is more. To reduce the complexity involved in the algorithms and for easier implementation, nowadays, the carrier based PWM algorithms are attracting many researchers.

The magnitude tests based approach is presented to generate the carrier based SVPWM and various DPWM algorithms with reduced complexity in [6]-[7]. Also, by distributing the zero state time unequally, various PWM algorithms have been generated in [9]. The detailed explanation regarding to the correlation between carrier comparison and space vector approaches is given in [10]. However, the [6]-[10] gives the explanation under the linear modulation region only.

To reduce the complexity in the classical SVPWM algorithms and to extend the operation up to overmodulation region, various PWM algorithms have been generated in [11]-[12] by using the concept of offset time. By adding the

suitable offset time to the imaginary switching times, which are proportional to the instantaneous phase voltages, various PWM algorithms have been generated under both linear and overmodulation regions. Similarly, various approaches have been presented in [13]-[14] for the generation of various carrier based PWM algorithms.

This paper presents carrier based generalized PWM (GPWM) algorithm for direct torque controlled induction motor drives. In the proposed GPWM algorithm by changing a constant value between 0 and 1, various PWM algorithms have been generated. Moreover, the proposed algorithm uses instantaneous phase voltages only. Thus, the proposed GPWM algorithm will bring all modulators under a common roof with reduced complexity.

II. PROPOSED GPWM ALGORITHM

The proposed GPWM algorithm may be pursued by the definition of a duty cycle or modulating signal for phase *n* (with *n* = a, b and c), which is given as the ratio between pulsewidth and modulation period.

$$V_n^* = \frac{Pulsewidth}{Modualtionperiod} \quad (1)$$

Once the modulating signal V_n^* is calculated, the ON and OFF times of the inverter-leg devices can be via digital counters and comparators. For example, the duty cycle or modulating signal of SPWM algorithm can be obtained as follows [9]-[10]:

$$V_n^* = \frac{1}{2} + \frac{V_n}{V_{dc}}, \quad n = a, b \text{ and } c \quad (2)$$

where V_n is the instantaneous reference voltage of phase *n* and V_{dc} is the dc-link voltage. In the similar way, the modulating signals of the various DPWM algorithms and SVPWM algorithms can be obtained by adding a suitable zero sequence voltage (V_z) to the instantaneous phase voltages (V_n).

$$V_n^* = k_1 + \frac{V_n + V_z}{V_{dc}} \quad (3)$$

where $V_z = k_2[\min(V_n) - \max(V_n)] - \min(V_n)$ (4)

where k_2 is the parameter that takes into account the unequal null-state sharing, can be defined as follows:

$$k_2 = 0.5(1 + \text{sgn}(\cos(3\omega t + \delta)) \quad (5)$$

where sgn(X) is 1, 0 and –1 when X is positive, zero, and negative, respectively. As previously discussed, and k_1 is an additional parameter whose value may be equal to the value of k_2 or be fixed at 0.5. Thus, the proposed approach eliminates the calculation of both the hexagon sector, in which the reference-voltage space vector is located, and the related phase.

In all the other carrier-based techniques, it must be taken that $k_1 = k_2$. The standard SVPWM algorithm can be obtained by fixing the k_2 value at 0.5. Similarly, by fixing the k_2 value at 0 and 1, the DPWMMIN and DPWMMAX algorithms can be obtained. By varying the modulation angle δ in (5), various DPWM algorithms can be generated. The DPWM0, DPWM1, DPWM2 and DPWM3 can be obtained for $\delta = \pi/6, 0, -\pi/6$ and $-\pi/3$ respectively.

In conclusion, it is worth noticing that a mathematical expression of the modulating signal in SVPWM was, in effect, already known, but it was referred only to classical SVPWM operating in linear modulation range. Here, the use of the modulating signal in the synthesis of the switching pattern has been put in evidence, and as a novelty, it has been extended to the over modulation range and in generalized modulation by defining the new k_1 and k_2 parameters.

The overmodulation range is of remarkable importance, particularly in the adjustable speed drive applications, in order to well exploit the dc-link voltage, thus obtaining the rated voltage of the motor without using a large dc voltage. It is easy to realize that a good over modulation strategy should manage the transition beyond the linear range, going toward the six-step behavior and thus avoiding any abrupt change of the fundamental voltage components that is applied to the load. There are so many approaches to extend the operation in to the overmodulation region. However, to reduce the complexity burden, this paper presents a simplified approach to extend the operation in to the overmodulation region. In the overmodulation region the difference between the maximum and minimum voltages in a sampling time period, which is also known as effective voltage is greater than the V_{dc}. The effective voltage can be defined as follows:

$$V_{eff} = \max(V_n) - \min(V_n) \quad (6)$$

Hence, in the overmodulation region the modulating signal can be calculated as follows:

$$V_n^* = \frac{1}{2} + \frac{V_n + V_z}{V_{eff}} \quad (7)$$

In overmodulation region, the zero state time partition parameter value (k_2) is fixed at 0.5. However, k_2 is clearly no longer present in the duty-cycle expression because now the null states are not applied.

III. PROPOSED GPWM ALGORITHM BASED DTC

The reference voltage space vector can be constructed in many ways. But, to reduce the complexity of the algorithm, in this thesis, the required reference voltage vector, to control the torque and flux cycle-by-cycle basis is constructed by using the errors between the reference d-axis and q-axis

stator fluxes and d-axis and q-axis estimated stator fluxes sampled from the Previous cycle. The block diagram of the proposed GPWM based DTC is as shown in Fig. 1. From Fig. 1, it is seen that the proposed GPWM based DTC scheme retains all the advantages of the DTC, such as no co-ordinate transformation, robust to motor parameters, etc. However a space vector modulator is used to generate the pulses for the inverter, therefore the complexity is increased in comparison with the DTC method.

In the proposed method, the position of the reference stator flux vector $\overline{\psi}_s^*$ is derived by the addition of slip speed and actual rotor speed. The actual synchronous speed of the stator flux vector $\overline{\psi}_s$ is calculated from the adaptive motor model.

After each sampling interval, actual stator flux vector $\overline{\psi}_s$ is corrected by the error and it tries to attain the reference flux space vector $\overline{\psi}_s^*$. Thus the flux error is minimized in each sampling interval. The d-axis and q-axis components of the reference voltage vector can be obtained as follows:
Reference values of the d-axis and q- axis stator fluxes and actual values of the d-axis and q-axis stator fluxes are compared in the reference voltage vector calculator block and hence the

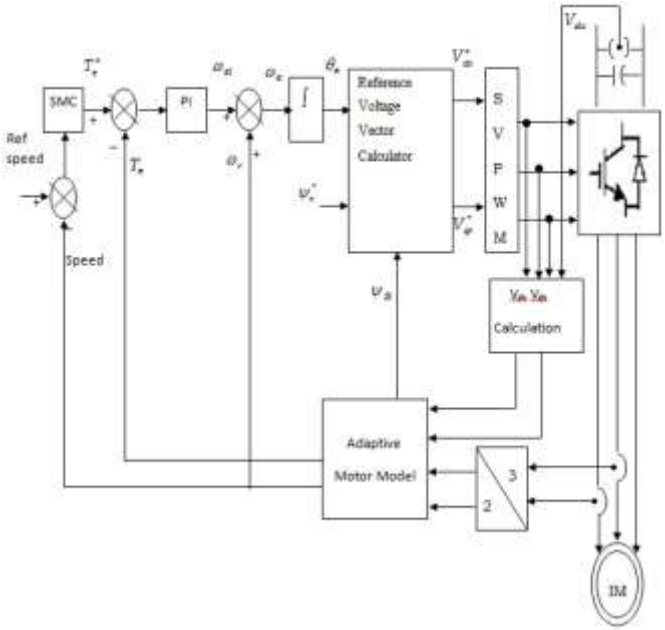


Fig.1 Block diagram of proposed SVPWM based DTC

errors in the d-axis and q-axis stator flux vectors are obtained as in (8)-(9).

$$\Delta\Psi_{ds} = \Psi_{ds}^* - \Psi_{ds} \quad (8)$$

$$\Delta\Psi_{qs} = \Psi_{qs}^* - \Psi_{qs} \quad (9)$$

The appropriate reference voltage space vectors due to flux error and stator ohmic drop are given as

$$v_{ds}^* = R_s i_{ds} + \frac{\Delta\psi_{ds}}{T_s} \quad (10)$$

$$v_{qs}^* = R_s i_{qs} + \frac{\Delta\psi_{qs}}{T_s} \quad (11)$$

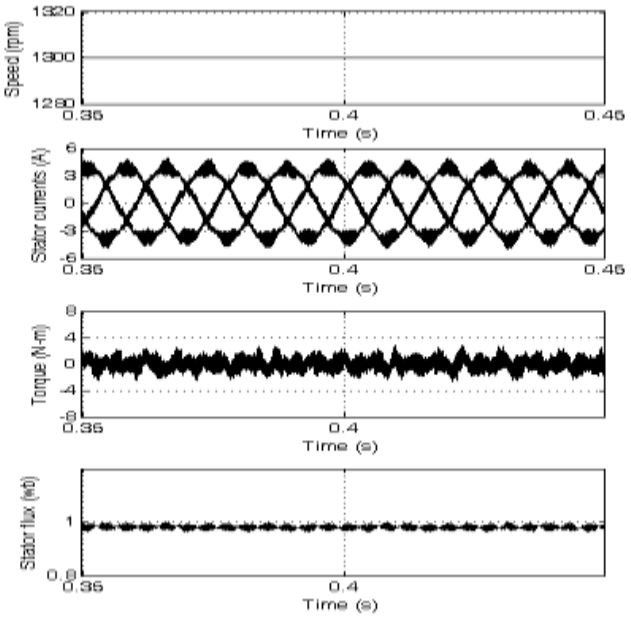
IV. SIMULATION RESULTS AND DISCUSSION

To validate the proposed generalized PWM algorithm based DTC, several numerical simulation studies have been carried out and results are presented. The details of the induction motor, which is used for simulation studies are as follows:

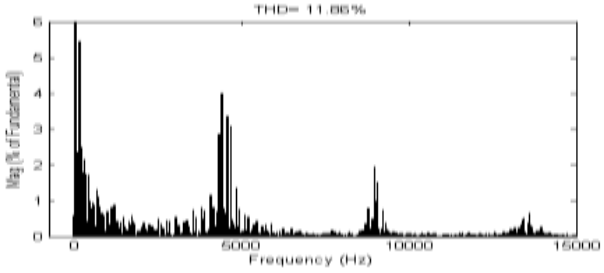
A 3-phase, 4 pole, 4kW, 1200rpm induction motor with parameters as follows:

$R_s = 1.57\Omega$, $R_r = 1.21\Omega$, $L_s = L_r = 0.17$H, $L_m = 0.165$H and $J = 0.089$Kg.m^2.

The steady state simulation results for various PWM algorithms based DTC are shown from Fig. 2 to Fig. 8. From the simulation results it can be observed that at higher modulation indices (motor is running at 1300 rpm, which pertains to the higher modulation region), the DPWM algorithms give reduced harmonic distortion when compared with the SVPWM algorithm. Moreover, it can be concluded that among the various DPWM algorithms, DPWM3 gives reduced harmonic distortion. From the simulation results, it can be observed that as the SVPWM algorithm is a continuous PWM algorithm, it gives continuous pulse pattern and more switching losses. Whereas, the DPWM algorithms clamp each phase to either positive or negative DC bus for 120 degrees over a fundamental cycle, these reduce the switching frequency and switching losses by 33.33% when compared with the SVPWM algorithm. Thus, the proposed GPWM algorithm generates a wide range of PWM algorithms at all modulation indices with reduced complexity by varying a parameter k_2 from 0 to 1.

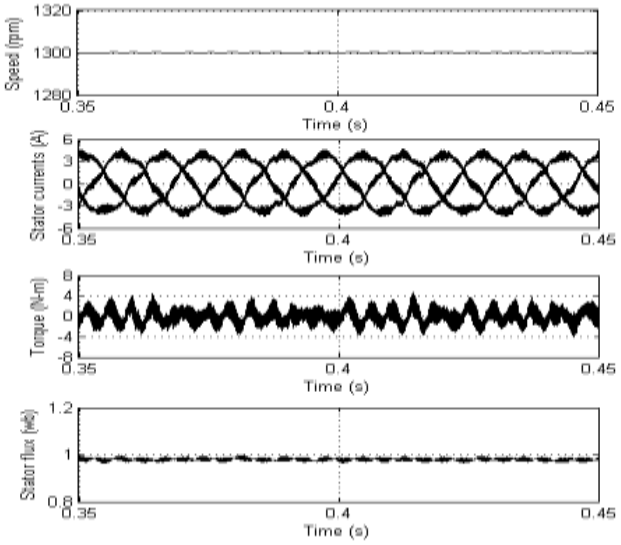


(a)

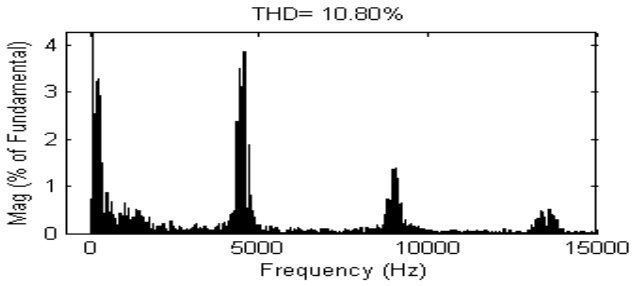


(b)

Fig. 2 Simulation results for SVPWM based DTC drive

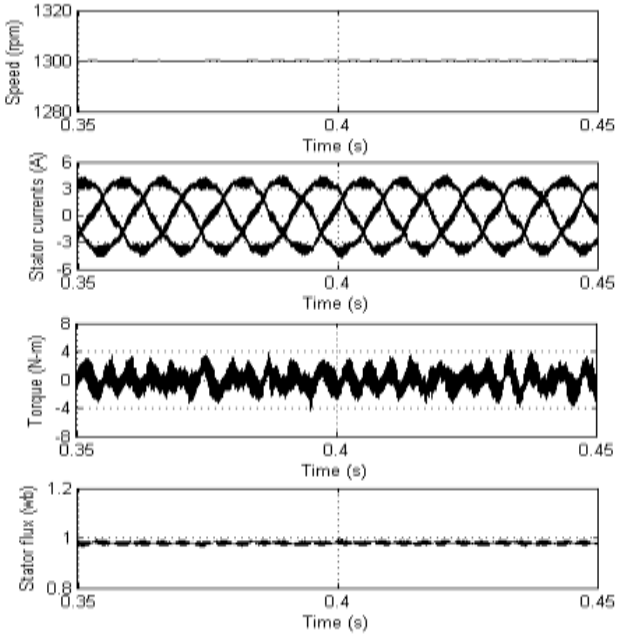


(a)

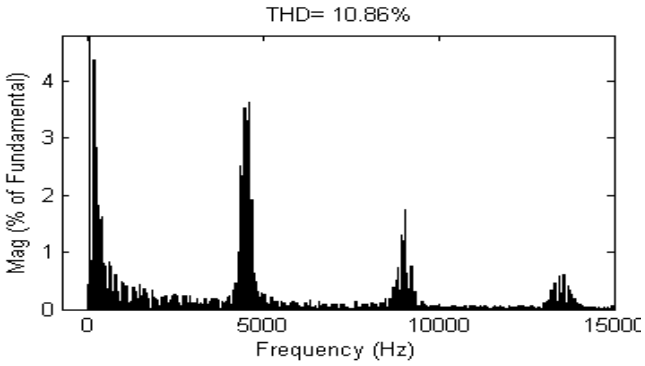


(b)

Fig. 3 Simulation results for DPWMMIN based DTC drive

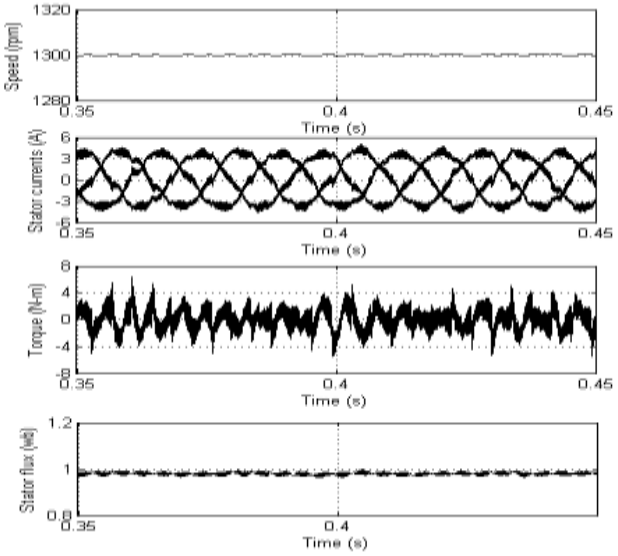


(a)

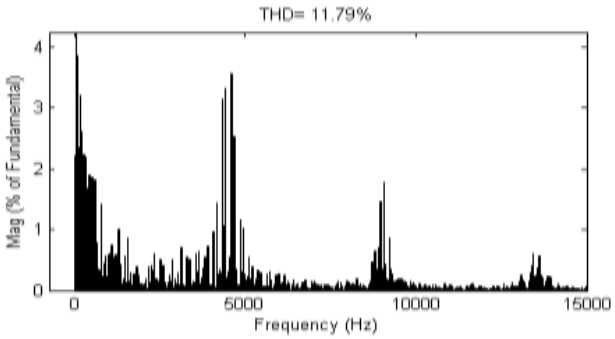


(b)

Fig. 4 Simulation results for DPWMMAX based DTC drive

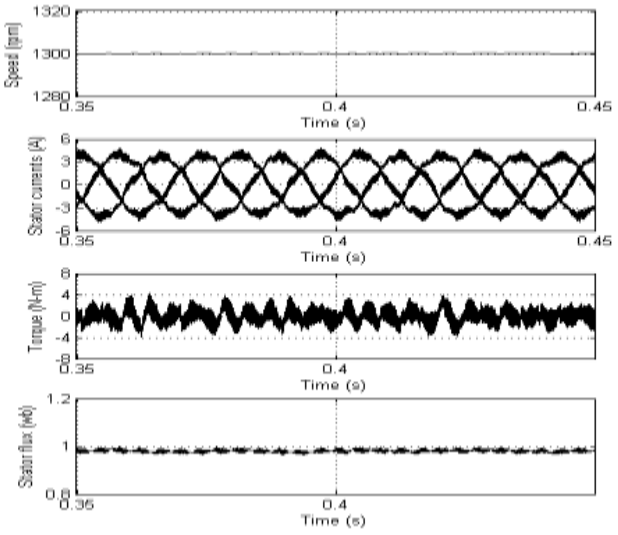


(a)

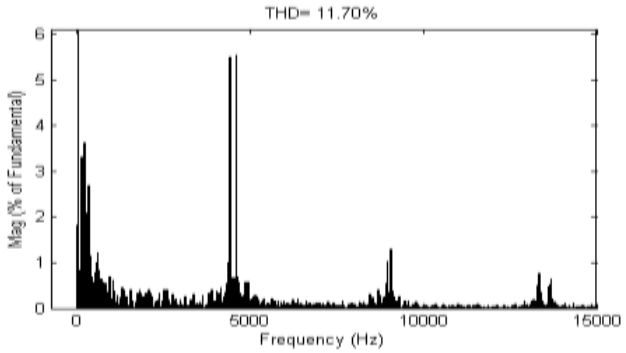


(b)

Fig. 5 Simulation results for DPWM0 based DTC drive

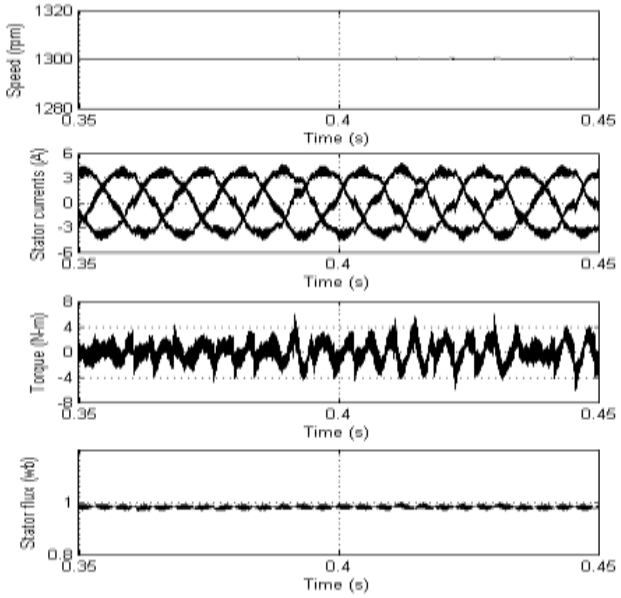


(a)

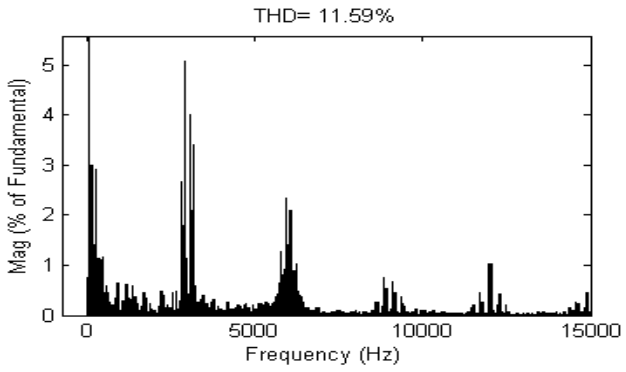


(b)

Fig. 6 Simulation results for DPWM1 based DTC drive

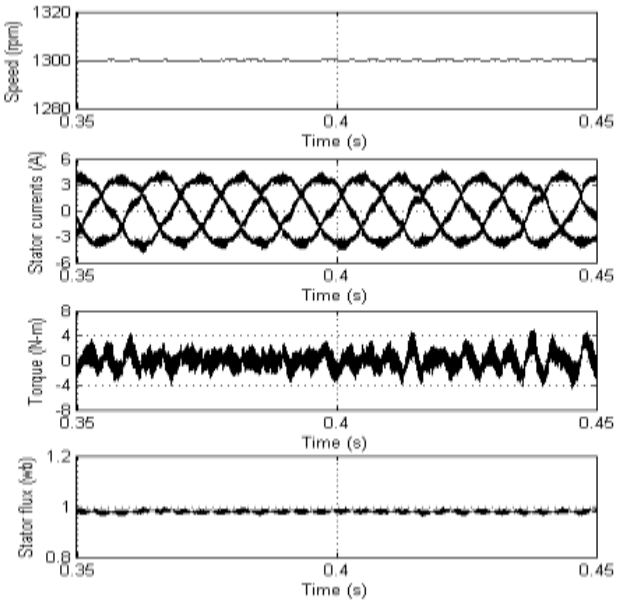


(a)

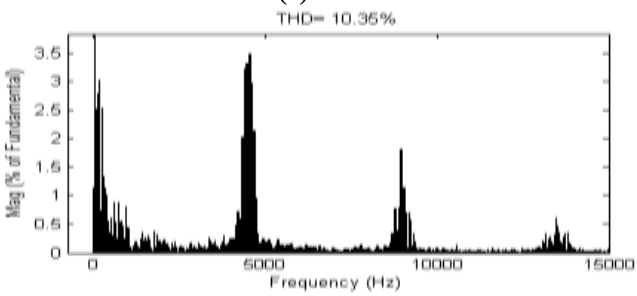


(b)

Fig. 7 Simulation results for DPWM2 based DTC drive



(a)



(b)

Fig. 8 Simulation results for DPWM3 based DTC drive

V. Conclusion

A simple and novel GPWM algorithm for direct torque controlled induction motor drives is presented in this paper. The proposed algorithm generates a wide range of DPWM algorithms along with SVPWM algorithm by using the instantaneous phase voltages only. From the simulation results it can be observed that the proposed GPWM algorithm gives all possible PWM modulators with reduced complexity. Moreover, from the simulation results it can be observed that at higher modulation indices (motor is running at 1300 rpm, which pertains to the higher modulation region), the DPWM algorithms give reduced harmonic distortion when compared with the SVPWM algorithm. Moreover, it can be concluded that among the various DPWM algorithms, DPWM3 gives reduced harmonic distortion.

References

[1] F. Blaschke "The principle of field orientation as applied to the new transvector closed loop control system for rotating-field machines," *Siemens Review*, 1972, pp 217-220.

[2] Isao Takahashi and Toshihiko Noguchi, "A new quick-response and high-efficiency control strategy of an induction motor," *IEEE Trans. Ind. Applicat.*, vol. IA-22, no.5, Sep/Oct 1986, pp. 820-827.

[3] Domenico Casadei, Francesco Profumo, Giovanni Serra, and Angelo Tani, "FOC and DTC: Two Viable Schemes for Induction Motors Torque Control" *IEEE Trans. Power Electron.*, vol. 17, no.5, Sep, 2002, pp. 779-787.

[4] Joachim Holtz, "Pulsewidth modulation – A survey" IEEE Trans. Ind. Electron., vol. 39, no. 5, Dec 1992, pp. 410-420.

[5] Heinz Willi Van Der Broeck, Hans-Christoph Skudelny and Georg Viktor Stanke, "Analysis and realization of a Pulsewidth Modulator based on Voltage Space Vectors", IEEE Trans. Ind. Applic., Vol. 24, No.1, Jan/Feb 1988, pp.142-150.

[6] Ahmet M. Hava, Russel J. Kerkman and Thomas A. Lipo, "A high-performance generalized discontinuous PWM algorithm" IEEE Trans. Ind. Applicat., vol. 34, no. 5, Sept/Oct, 1998, pp. 1059-1071.

[7] Ahmet M. Hava, Russel J. Krkman and Thomas A. Lipo, "Simple analytical and graphical methods for carrier-based PWM-VSI drives" IEEE Trans. Power Electron., vol. 14, no. 1, Jan 1999, pp. 49-61.

[8] V. Blasko, "Analysis of a hybrid PWM based on modified space-vector and triangle-comparison method," IEEE Trans. Ind. Applicat., vol. 33, pp. 756–764, May/Jun, 1997.

[9] Olorunfemi Ojo, "The generalized discontinuous PWM scheme for three-phase voltage source inverters" IEEE Trans. Ind. Electron., vol. 51, no. 6, Dec, 2004, pp. 1280-1289.

[10] Keliang Zhou and Danwei Wang, "Relationship Between Space-Vector Modulation and Three-Phase Carrier-Based PWM: A Comprehensive Analysis" *IEEE Trans. Ind. Electronics*, vol. 49, no.1, pp. 186-196, Feb, 2002.

[11] Joohn-Sheok Kim and Seung-Ki Sul, "A novel voltage modulation technique of the space vector PWM" in *Proc. IPEC*, Yokohama, Japan, 1995, pp. 742-747.

[12] Dae-Woong Chung, Joohn-Sheok Kim, Seung-Ki Sul, "Unified Voltage Modulation Technique for Real-Time Three-Phase Power Conversion" IEEE Trans. On Ind. Applications, vol. 34, no.2, pp 374-380, March/April, 1998, pp. 756-764.

[13] Antonio Cataliotti, Fabio Genduso, Angelo Raciti, and Giuseppe Ricco Galluzzo "Generalized PWM–VSI Control Algorithm Based on a Universal Duty-Cycle Expression: Theoretical Analysis, Simulation Results, and Experimental Validations" *IEEE transactions on Ind. Electron.*, vol. 54, NO. 3, June 2007, pp 1569-1580.

[14] Edison Roberto C. Da Silva, Euzeli CiprianoDos Santos, Jr., and Cursino Bradao Jacobina, "Pulsewidth modulation strategies" *IEEE Ind. Electron., Magazine*, no.2, pp.37-45, June, 2011.

Reducing Powersystem Oscillations Using Facts Controller(Tcsc)

N Mohan, M.Ramesh
(Student of EEE Department i, Madanapally Institute of Technology & Science)
(Associate Professor, EEE Department, Madanapally Institute of Technology & Science)

Abstract
The recently proposed phase imbalanced series capacitive compensation concept has been shown to be effective in enhancing power system dynamics as it has the potential of damping power swing as well as sub synchronous resonance oscillations. In this paper, the effectiveness of a "hybrid" series capacitive compensation scheme in damping power system oscillations is evaluated. A hybrid scheme is a series capacitive compensation scheme, where two phases are compensated by fixed series capacitor (C) and the third phase is compensated by a TCSC in series with a fixed capacitor (Cc). The effectiveness of the scheme in damping power system oscillations for various network conditions, namely different system faults and tie-line power flows is evaluated using the EMTP-RV time simulation program.

Index Terms—FACTS Controllers, phase imbalance, series compensation, thyristor controlled series capacitor.

Introduction:
FLEXIBLE AC Transmission Systems (FACTS) technology provides unprecedented way for controlling transmission grids and increasing transmission capacity.FACTS Controllers provide the flexibility of controlling both real and reactive power which could result in an excellent capability for improving power system dynamics. A problem of interest in the power industry at which FACTS Controllers could play a significant role in it is Increasing damping of low frequency power oscillations that often arise between areas in large interconnected power networks. These oscillations are termed inter-area oscillations, which are normally characterized by poor damping [2]. Inter-area oscillations can severely restrict system operations by requiring the curtailment of electric power transfers level as an operational measure. These oscillations can also lead to widespread system disturbances, e.g. cascading outages of transmission lines and, therefore, system wide voltage collapse Several

Studies have investigated the potential of using FACTS Controllers' capability in damping inter-area oscillations.
The use of Thyristor Controlled Series Capacitor (TCSC), and Static Synchronous Series Compensator (SSSC) have been the subjects of several studies evaluating their respective effectiveness in enhancing power system dynamics The recently proposed phase imbalanced series capacitive compensation concept has been shown to be effective in enhancing power system dynamics as it has the potential of damping power swing as well as subsynchronous resonance oscillations [7], [8]. Fig. 1 shows a scheme for a phase imbalanced capacitive compensation. It is a "hybrid" series compensation scheme, where the series capacitive compensation in one phase is created using a single-phase
TCSC in series with a fixed capacitor (Cc), and the other two phases are compensated by fixed series capacitors (C). The TCSC control is initially set such that its equivalent compensations at the power frequency combined with the fixed capacitor yield a resultant compensation equal to the other two phases. Thus, the phase balance is maintained at the power frequency while at any other frequency, a phase imbalance is created. To further enhance power oscillations damping, the TCSC is equipped with a supplementary controller.

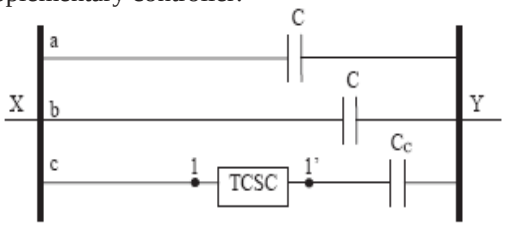


Fig. 1. A schematic diagram of the hybrid series compensation scheme

The phase imbalance of the proposed scheme can be explained mathematically as follows: 1) At the power frequency, the series reactances between buses X and Y, in Fig. 1, in phases a, b, and c are given by:

$$X_a = X_b = \frac{1}{j\omega_o C}$$

$$X_c = \frac{1}{j\omega_o C_c} - jX_{TCSCo}$$

Where − jX TCSCo is the effective capacitive reactance of the TCSC at the power frequency such that

$$X_a = X_b = X_c$$

2) During any other frequency, fe

$$X_c = \frac{1}{j\omega_e C_c} - jX_{TCSCo} - j\Delta X_{TCSC}$$

The first terms in (2) and (3) are different because of the difference in frequency. The third term in (3) represents the change in the effective capacitive reactance of the TCSC due to the action of the TCSC supplementary controller.

This scheme would, definitely, be economically attractive when compared with a full three-phase TCSC which has been used/proposed for power oscillations damping. Furthermore, reducing the number of thyristor valves to one third will also have a positive impact on system reliability. The effectiveness of the scheme in damping power swings and subsynchronous resonance oscillations is reported in [7], [8]. This paper evaluates the effectiveness of the scheme in damping power system oscillations. Time domain simulations were conducted on a benchmark network using the EMTP-RV.

Ii. Study Benchmark:

To demonstrate the effectiveness of the proposed scheme in power system oscillations damping, the system shown in Fig. 2 is adopted as a test benchmark. It consists of three large generating stations (G1, G2 and G3) supplying two load centers (S1 and S2) through five 500 kV transmission lines. The two double-circuit transmission lines L1 and L2 are series compensated with fixed capacitor banks located at the middle of the lines. The compensation degree of L1 and L2 is 50%. The compensation degree is defined as the ratio (XC/XL)* 100% for fixed capacitor compensated phases and (XCc+XTCSC)/XL* 100% for the hybrid compensated phase. The total installed capacity and peak load of the system are 4500 MVA and 3833 MVA respectively. Shunt capacitors are installed at buses 4 and 5 to maintain their voltages within 1±0.05 p.u. Two loading profiles designated as Load Profiles A and B are considered in the investigations of this paper. In Load Profile A, S1 = 1400 + j200 MVA and S2 = 2400 + j300 MVA while in Load Profile B, S1 = 2000 + j200 MVA and S2 = 1800 + j300 MVA. The power flow results for the bus voltages and the line real power flows of the system for these two loading profiles are shown in the Appendix. The EMTP- RV is used as the simulation study tool.

III. Modeling of The Single-Phase Tcsc:

The single-phase TCSC is modeled in the EMTP-RV as a single module using an ideal thyristor pair and an RC snubber circuit as shown in Fig. 3. A Phase Locked Loop (PLL) is used to extract phase information of the fundamental frequency line current, which will be used to synchronize TCSC operation. The thyristor gating control is based on the Synchronous Voltage Reversal (SVR) technique [9] - [11]. The TCSC impedance is measured in terms of a boost factor kB, which is the ratio of the apparent reactance of the TCSC seen from the line to the physical reactance of the TCSC capacitor bank. A positive value of kB is considered for capacitive operation. A low-pass filter based estimation algorithm is used to estimate the voltage and the current phasors. A boost measurement block performs complex impedance calculations for the boost factor of the TCSC as kBˆˆ= Imag VC / I C / X CTCSC , where, VC and I C are the estimated phase voltage and current and XCTCSC is the capacitive reactance of the TCSC capacitor branch at the fundamental frequency. A proportional-integral (PI) control based boost level controller is implemented to control the TCSC boost level to the desired value by adjusting the instant of the expected capacitor voltage zero crossing. The integral part of the controller helps in removing the steady state errors. The controller parameters were determined by performing repeated time domain simulations for the different operating conditions.

This algorithm uses the difference between the actual boost level and the reference boost level (err) shown in Fig. 3 as an objective function. The algorithm starts with arbitrary initial values for the control parameters and calculates the values of the objective function each time. The control parameters are incremented for the next iteration and the procedure is repeated until the objective function approaches a minimum value (below a threshold value). The procedure described above is widely used by industry for tuning of controller parameters. The multiple simulations run based tuning procedure similar to the above was reported in [12].

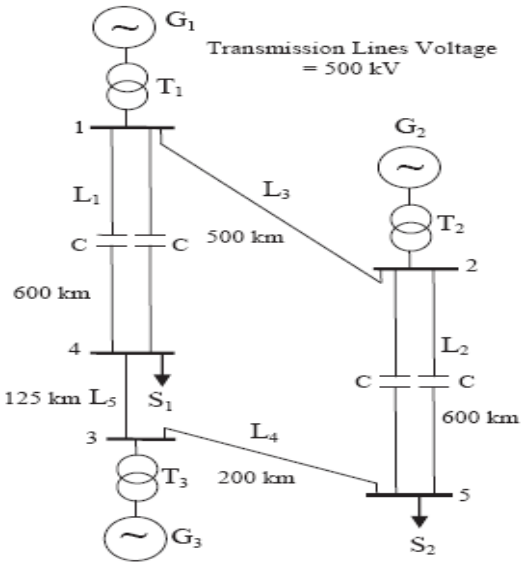


Fig. 2. Test benchmark

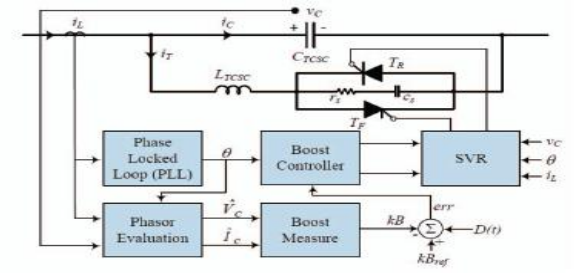


Fig. 3. Block diagram of a TCSC controller

In Fig. 3, D(t) is a supplemental signal generated from an m-stage lead-lag compensation based controller. As the real power flow in the transmission line is proportional to the inverse of the total line reactance, the power swing damping can be achieved by properly modulating the apparent TCSC reactance through this controller. The supplemental controller input (stabilizing) signals could be local (e.g. real power flows) or remote (e.g. load angles or speed deviations of remote generators). If a wide-area network of Synchronized Phasor Measurement (SPM) units is available, then the remote signals can be downloaded at the controller in real time without delay. Local signals are generally preferred over remote signals as they are more reliable since they do not depend on communications. In Fig. 3, kBref is the TCSC boost level set point. TheSynchronous Voltage Reversal block solves for angle □ □ fromthe non-linear relation,

$$u_{CZ} = X_o i_{LM}\left[\lambda\gamma - \tan(\lambda\gamma)\right],$$

wher e uCZ is the estimated capacitor voltage at the desired instant when the capacitor voltage zero crossing occurs, iLM is the measured value of the line current iL, X0 is the TCSC capacitor reactance at the TCSC resonance frequency, □ □ is theratio between the TCSC resonance frequency and the system fundamental frequency and □ □ is the angle difference betweenthe firing time and the voltage zero-crossing. The value of □ □ isused to calculate the exact firing instants of the individual thyristors. The non-linear relationship between the boost factor and the thyristor firing angle is shown in Fig. 4.

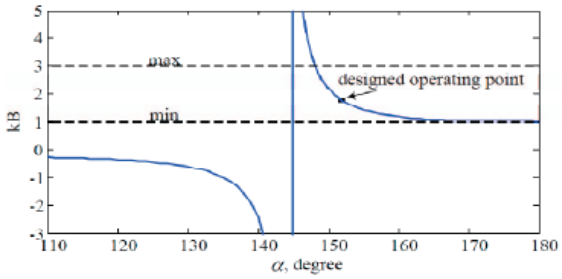


Fig. 4. TCSC boost factor as a function of the thyristor firing angle .

This section demonstrates the capability of the proposed hybrid series compensation scheme in power system oscillations damping. For this purpose, the scheme is assumed to be placed in the test benchmark replacing the fixed series capacitive compensation in L1 and L2. Moreover, it is assumed that each TCSC provides 50% of the total capacitive compensation and the disturbance is a three-cycle, three-phase fault at bus 4. Furthermore, the performance of the scheme is compared to the case with only fixed capacitor compensation which is labeled in the figures of the time responses as Fixed C.

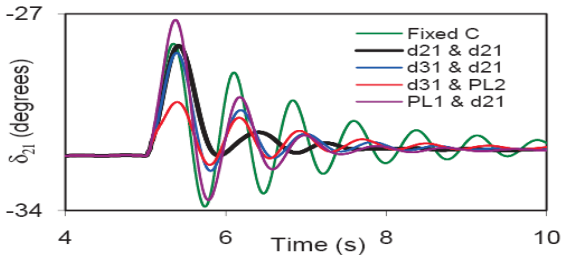
CASE STUDY I: Load Profile A

In this case, four different combinations of stabilizing signals (tabulated in Table I) are examined in order to determine the combination that would result in the best system transient time responses. The final results of the time-domain simulation studies (controllers tuning) are shown in Fig. 5 which illustrates the generator load angles, measured with respect to generator 1 load angle, during and after fault clearing. The transfer functions of the TCSC supplemental controllers for the four combinations are given in Table II. Comparing the responses of the fixed series capacitor compensation to the hybrid TCSC compensation scheme in

Fig. 5, the positive contribution of the proposed hybrid scheme to the damping of the system oscillations is very clear. As it can be seen from Fig. 5, the power swing damping controller effectively damps the system oscillations. It can also be seen from Fig. 5 that the best damping of the relative load angle responses are achieved with the 21- 21 combination. The second best damped responses are obtained with the 31- 21 combination. These results should be expected due to the direct relationship between the relative load angles and the generators that yield the problem. It can also be seen from Fig. 5 that the worst damped responses are obtained with PL1- 21 combination which results also in the increase of the first swings.

TABLE I: THE FOUR EXAMINED COMBINATIONS OF STABILIZING SIGNALS FOR CASE STUDY I

Combination	Each TCSC in L_1	Each TCSC in L_2
1	δ_{21}	δ_{21}
2	δ_{31}	δ_{21}
3	δ_{31}	P_{L2}
4	P_{L1}	δ_{21}



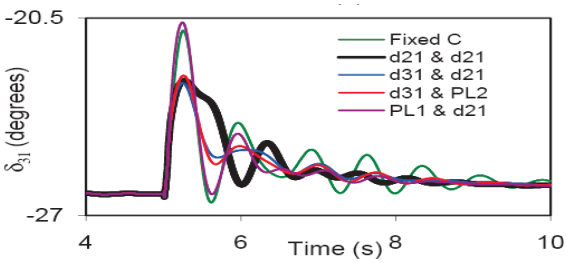


Fig. 5. Generator load angles, measured with respect to generator 1 load angle, during and after clearing a three-phase fault at bus 4 (Load Profile A).

CASE STUDY II: Load Profile B

In this case, 21 is used as the supplementary controllers stabilizing signal. The transfer functions of the TCSC supplemental controllers are given in Table III. Fig. 6 illustrates the generator load angles, measured with respect to generator 1 load angle, during and after fault clearing. It can be seen from Fig. 6 that, at this loading profile, the hybrid single-phase-TCSC scheme provides again a better damping performance to system oscillations compared to fixed capacitor compensation. It is observed, however, that there is a slight increase in the first swing of 21.0 TABLE II

Transfer Functions of the Tcsc Supplemental Controllers for Case Study I

Combi-nation	Each TCSC in L_1	Each TCSC in L_2
1	$G(s)=0.25\frac{10}{(s+10)}\frac{3s}{(3s+1)}$	$G(s)=-0.15\frac{10}{(s+10)}\frac{3s}{(3s+1)}$
2	$G(s)=0.05\frac{10}{(s+10)}\frac{3s}{(3s+1)}$	$G(s)=-0.15\frac{10}{(s+10)}\frac{3s}{(3s+1)}$
3	$G(s)=0.1\frac{10}{(s+10)}\frac{3s}{(3s+1)}$	$G(s)=-0.4\frac{10}{(s+10)}\frac{3s}{(3s+1)}$
4	$G(s)=-0.25\frac{10}{(s+10)}\frac{3s}{(3s+1)}$	$G(s)=-0.25\frac{10}{(s+10)}\frac{3s}{(3s+1)}$

TRANSFER FUNCTIONS OF THE TCSC SUPPLEMENTAL CONTROLLERS FOR CASE STUDY II

Each TCSC in L_1	$G(s)=0.3\frac{10}{(s+10)}\frac{3s}{(3s+1)}$
Each TCSC in L_2	$G(s)=-0.15\frac{10}{(s+10)}\frac{3s}{(3s+1)}$

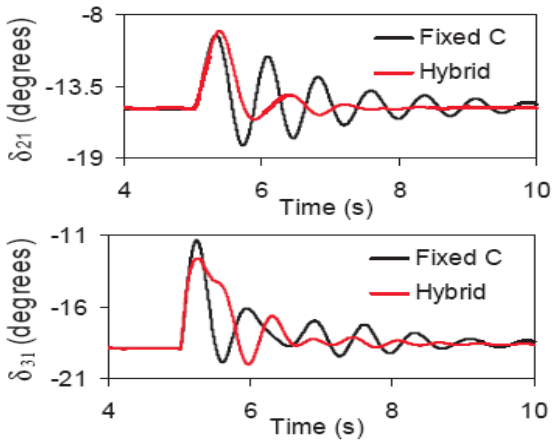
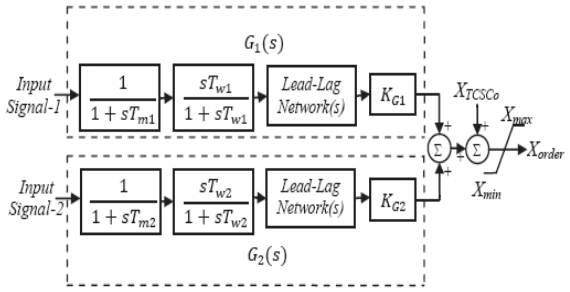


Fig. 6. Generator load angles, measured with respect to generator 1 load angle, during and after clearing a three-phase fault at bus 4 (Load Profile B).

CASE STUDY III:

A Dual-Channel Controller:

Any of the four signals, 21, 31, PL1 and PL2 contains the system's two natural modes of oscillations and can be used to add damping to these modes as it has been demonstrated in Case study I. The sum of two properly selected signals, however, should result in a more effective damping. The reason is that the two natural modes of oscillations are, in general, not in phase. A dual-channel controller would adjust separately the gain and phase of each mode of oscillations and, thus, provides a better damping.



The performance of the dual-channel TCSC supplemental controller shown in Fig. 7 in damping power system oscillations is examined using the six pairs of signals given in Table IV. Investigations are conducted on the test benchmark system at Load profile B.

TABLE IV

THE SIX EXAMINED COMBINATIONS OF STABILIZING SIGNALS FOR CASE STUDY III

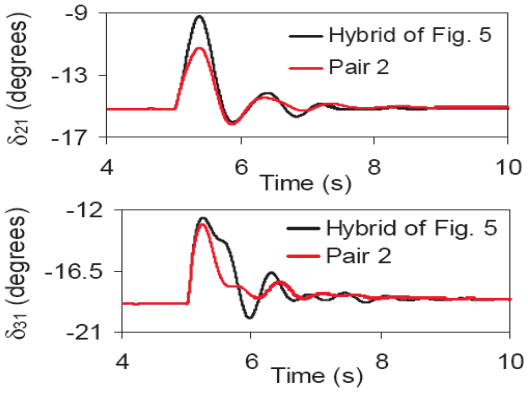
Pair number	Each TCSC (input signal-1, input signal-2)
1	δ_{21}, δ_{31}
2	δ_{21}, P_{L1}
3	δ_{21}, P_{L2}
4	δ_{31}, P_{L1}
5	δ_{31}, P_{L2}
6	P_{L1}, P_{L2}

The final results of the time-domain simulation studies (controllers tuning) show that the best and second best damped responses are obtained with pairs 2 and 5. The transfer functions of the TCSC supplemental controllers for the six pairs of signals are given in Table V. Fig. 8 illustrates the generator load angles, measured with respect to generator 1 load angle, during and after fault clearing. These results (in red color) are compared to the hybrid case of Fig. 6 (referred to as single channel).

TABLE V: TRANSFER FUNCTIONS OF THE TCSC SUPPLEMENTAL CONTROLLERS

Pair 2 Each TCSC in L_1	$G_1(s) = 0.25\frac{10}{(s+10)}\frac{0.5s}{(0.5s+1)}$ $G_2(s) = -0.5\frac{60}{(s+60)}\frac{0.01s}{(0.01s+1)}$
Pair 2 Each TCSC in L_2	$G_1(s) = 0.25\frac{10}{(s+10)}\frac{0.5s}{(0.5s+1)}$ $G_2(s) = -0.5\frac{60}{(s+60)}\frac{0.01s}{(0.01s+1)}$
Pair 5 Each TCSC in L_1	$G_1(s) = 0.28\frac{10}{(s+10)}\frac{3s}{(3s+1)}$ $G_2(s) = -2.5\frac{60}{(s+60)}\frac{0.01s}{(0.01s+1)}$ $*\frac{(s+0.1)}{(s+0.2)}\frac{(s+0.5)}{(s+3)}$
Pair 5 Each TCSC in L_2	$G_1(s) = 0.26\frac{10}{(s+10)}\frac{s}{(s+1)}$ $G_2(s) = 2\frac{60}{(s+60)}\frac{0.01s}{(0.01s+1)}$ $*\frac{(s+0.1)}{(s+0.2)}\frac{(s+0.5)}{(s+3)}$

Fig. 9 illustrates the three-phase voltages, VX-Y, across the hybrid single-phase-TCSC compensation scheme (installed in L1 and the controllers are Pair 2) during and after clearing the fault. The system phase imbalance during the disturbance is clearly noticeable especially in phase C.



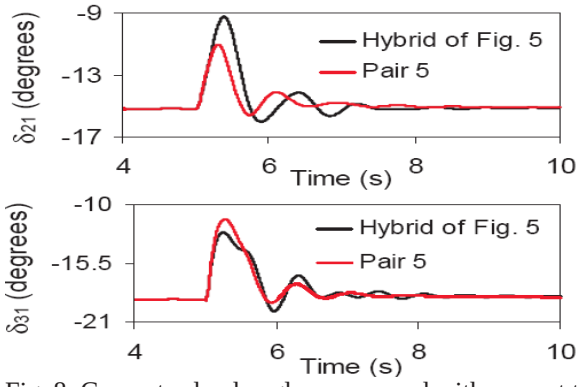


Fig. 8. Generator load angles, measured with respect to generator 1 load angle, during and after clearing a three-phase fault at bus 4 (Load Profile B, dual-channel controller).

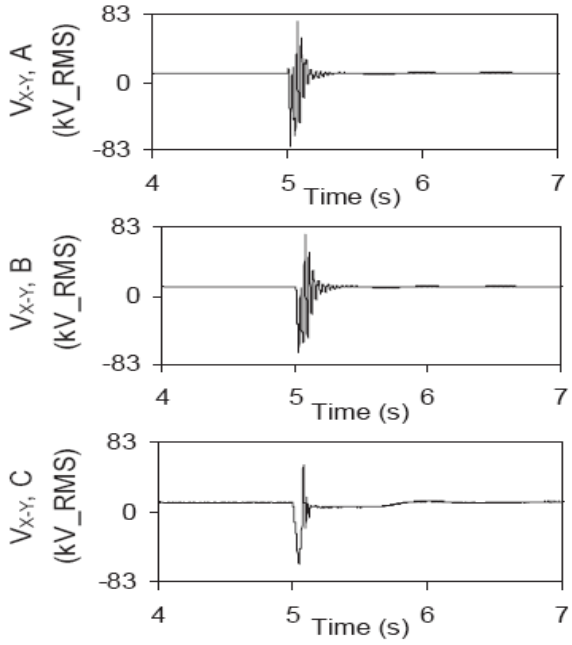


Fig. 9. Phase voltages, VX-Y across the hybrid single-phase-TCSC scheme on L1 during and after clearing a three-phase fault at bus 4 (Load Profile B, dual- channel supplemental controllers, Pair 2).

V. Conclusions:

The paper presents the application of a new hybrid series capacitive compensation scheme in damping power system oscillations. The effectiveness of the presented scheme in damping these oscillations is demonstrated through several digital computer simulations of case studies on a test benchmark. The presented hybrid series capacitive compensation scheme is feasible, technically sound, and has an industrial application potential.

References:

[1] Narain G. Hingorani and Laszlo Gyugyi, "Understanding FACTS, Concepts and Technology of Flexible AC Transmission Systems," IEEE Press, 2000.

[2] M. Klein, G.J. Rogers and P. Kundur, "A Fundamental Study of Inter- Area Oscillations in Power Systems," IEEE Transactions on Power Systems, Vol. 6, No. 3, 1991, pp. 914-921.

[3] E.V.Larsen J.J. Sanchez-Gasca and J.H. Chow, "Concepts for Design of FACTS Controllers to Damp Power Swings," IEEE Transactions on Power Systems, Vol. 10, No. 2, May 1995, pp. 948-956

[4] B. Chaudhuri, B. Pal, A. C. Zolotas, I. M. Jaimoukha, and T. C. Green, "Mixed-sensitivity Approach to H Control of Power System Oscillations Employing Multiple FACTS Devices," IEEE Transactions on Power System, Vol. 18, No. 3, August 2003, pp. 1149–1156.

[5] B. Chaudhuri and B. Pal, "Robust Damping of Multiple Swing ModEL Employing Global Stabilizing Signals with a TCSC," IEEE Transactions on Power System, Vol. 19, No. 1, February 2004, pp. 499–506.

[6] R. Majumder, B.C. Pal, C. Dufour and P. Korba, "Design and Real- Time Implementation of Robust FACTS Controller for Damping Inter- Area Oscillation," IEEE Transactions on Power Systems, Vol. 21, No. 2, May 2006, pp. 809-816.

[7] D. Rai, G. Ramakrishna, S.O. Faried and A. Edris," Enhancement of Power System Dynamics Using a Phase Imbalanced Series Compensation Scheme," IEEE Transactions on Power Systems, Vol. 25, No. 2, May 2010, pp. 966-974.

[8] D. Rai, S.O. Faried, G. Ramakrishna, and A. Edris, "Hybrid Series Compensation Scheme Capable of Damping Subsynchronous Resonance," IET Generation, Transmission and Distribution, Vol. 4, No. 3, March 2010, pp. 456-466.

[9] H. Xie and L. Ängquist, "Synchronous Voltage Reversal control of TCSC – impact on SSR conditions," Proceedings of the Nordic Workshop on Power and Industrial Electronics (NORPIE), 2004.

[10] Lennart Ängquist, "Synchronous Voltage Reversal Control of Thyristor Controlled Series Capacitor," Royal Institute of Technology, TRITA- ETS-2002-07, ISSN 1650-674X.

[11] L. Angquist and C. Gama, "Damping Algorithm based on Phasor Estimation," Proceedings of the IEEE Power Engineering Society Winter Meeting, Columbus, Ohio, January 28 – February 1, 2001, Vol. 3, pp. 1160-1165.

12] A.M. Gole, S. Filizadeh, R.W. Menzies, and P.L. Wilson "Optimization-enabled Electromagnetic Transien Simulation," IEEEnt Transactions on Power Delivery, Vol. 20, No. 1, January 2005, pp. 512–518.

Black Hole Attack and its Counter Measures in AODV Routing Protocol

Varsha Patidar[1], Rakesh Verma[2]
1Medicaps Institute of Technology and Management, Pigdamber, Rau (M.P)
2 Assistant Professor, CSE Department, Medicaps Institute of Technology and Management, Pigdamber, Rau (M.P)

Abstract:
Mobile Ad-Hoc network is a collection of mobile node. In this network, a wireless node usually moves randomly and dynamically forms a temporary network without a network infrastructure. Due to absence of infrastructure, MANET is used in various application like medical, battle field, business application and remote areas. MANET is vulnerable to attacks such as Black Hole Attack, Grey Hole Attack, and Worm Hole Attack. Wireless Ad- Hoc network may be unprotected against attacks by malicious nodes due to security vulnerabilities. Many mechanisms have been proposed to overcome the Black Hole Attack. A malicious node send Route Response (RREP) incorrectly of having route to destination with minimum hop count and when sender sends the data packet to this malicious node, it drops all the packet in the network. If there are multiple malicious nodes in the network, and are cooperating with each other for performing the attack than such attack is called Cooperative Black Hole Attack.

Keywords: *Black Hole Attack, Grey Hole Attack, malicious node, MANET, RREP, Worm Hole Attack.*

Introduction:

In Ad-Hoc Network, mobile nodes communicate with each other without any fixed infrastructure between them. It does not have pre-defined infrastructure to keep the network connected. Ad-Hoc Network create a network in such situation where creating the infrastructure is impossible. Destination Sequenced Distance Vector (DSDV) routing is a table driven routing protocol. In this, a routing table is maintained by each mobile node with entries of every possible destination nodes and number of hops to reach the destination node. DSDV update its routing table periodically for every change in network. Whereas, AODV is an On-Demand source initiated routing protocol when source wishes to route a packet to destination node. In all communication networks, security is the major concern, but due to dependence on other nodes for transmission, Ad-Hoc network face the greatest challenge. Many researchers have proposed solutions for mitigating and identifying the single black hole node. The packet delivery ratio will reduced if some malicious node is in the path of destination node. To overcome from this problem, identification of misbehaved nod is necessary. To improve the performance of network, trust value for node is introduced. With the help of trust value, the behavior of node can be judged. If a node has low trust value in a network, then we can identify the misbehaving node in the network. A single node or multiple nodes collectively can perform the black hole attack. When a Source node want to establish a route to Destination node, the source node S will broad cast the RREQ message to all the nodes in the network until it reaches to Destination node. This approach is followed when there is no black hole attack in the network.

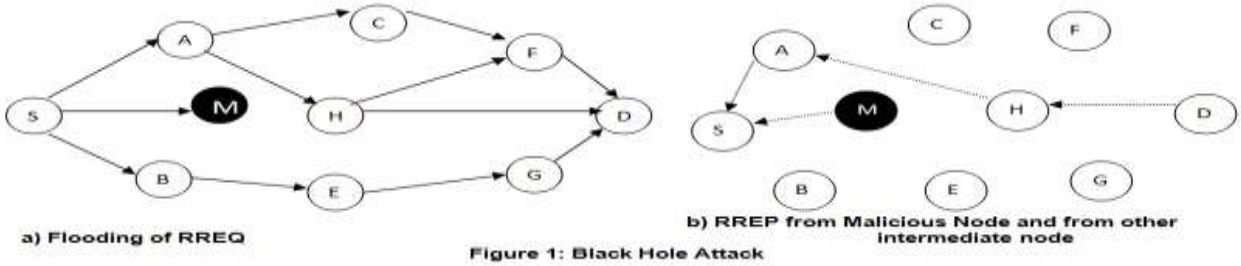


Figure 1: Black Hole Attack

In the above fig 1, when the source node broadcast the RREQ message, the black hole node will immediately reply RREQ through an RREP message. This RREP have an extremely large sequence number. Apart from this RREP, other normal nodes also receive the RREQ and destination node will select the route with minimal hop count and return the RREP. But as per AODV, largest sequence number and minimal hop count will be selected by source node. So, source node will select the black hole node for sending the data. Eavesdropping or direct dropping of received data packet is done by black hole node. Black hole node does not check its routing table and will respond to RREQ message before any other node check its routing table and respond to RREQ message.

2. AODV Routing Protocol

In proactive and reactive routing protocol, there must be good co-operation among the nodes so that the data can be routed successfully from source to destination. If nodes have good co-operation between them then there will be no packets

dropping or modification in the content. If there is no co-operation between the nodes, then there are high possibilities that an attacker take the advantage of situation and perform the malicious function. AODV routing protocol provides such situation when source node want to send the message to destination node which is not directly in contact with the source node then a route discovery process is initiated by broadcasting the RREQ message in the network. Now malicious node will take the advantage of this RREQ message and immediately send the RREP message to source node of having the route to the destination node without checking its routing table. This RREP message has the large sequence number and minimal hop count. When source node starts transmitting the data, the malicious nose will drop the packet rather than forwarding it to the destination node.

3. Co-operative Black Hole Attack

In Black Hole Attack, a node falsely advertises shortest and freshest path during route discovery process by the source node. When a group of malicious nodes works co-operatively than the condition become worse. According to Hongmei Dang et.al, [1] when source node get a RREP message of its RREQ message, from node B1 which is a malicious node, the source node sends a "further Request (FRq)" to the node B2 , asking it if it had route to destination node and also if it has route to node B1. Node B2 in response to these questions responds in YES to the source node. When source node sends a "Further Request (FRq)", it follows a different path other than node B1 and followed the path (S-3-4B2). Now source node is assured that the route S-B1-B2 is secure. Source node start sending data through node B1 and it consume the data rather than forwarding it further by compromising the security of the network.

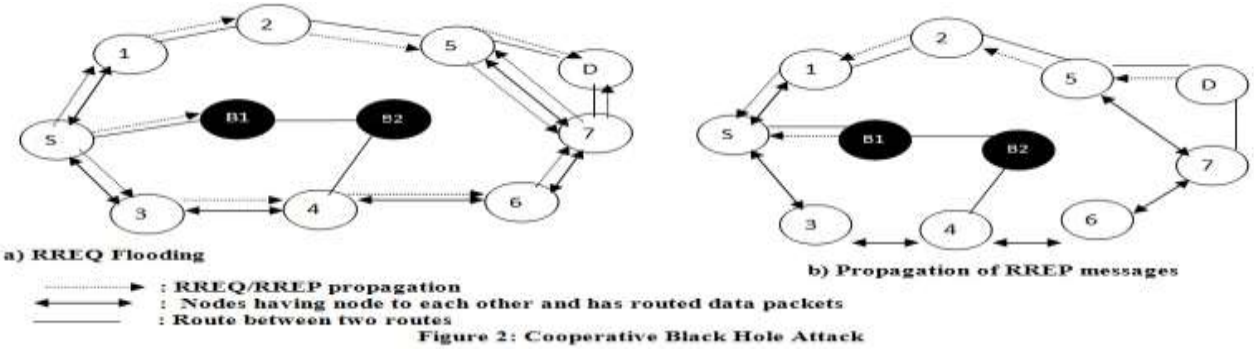


Figure 2: Cooperative Black Hole Attack

3. Related work on Black Hole Attack

There are many existing solution for detecting and mitigating the malicious node in the network. Marti et al. [2] proposed the use of watch dog and path rater. In watch dog, to detect the misbehavior in the path, it listens the transmission of the next node. In watch dog mechanism, the state information is maintained on maintained in monitored nodes, but the transmitted packets increase the memory overhead. Path rater works by keeping the rating of other nodes that ranges from 0 to 0.8 and the node that have the value 0.5 signifies as neutral. In Semih et al. [3], as the nodes are moved randomly due to this a different scenario is created by the network. Number of hopes for different path and conductivity of the network is tested. Zero hopes means there is no connection between a pair of nodes. In Buchegar [4], CONFIDANT (Cooperation of Nodes Fairness in Dynamic Ad hoc Network) was proposed. In existing watch dog and path rater scheme, this protocol adds trust manager and reputation system. The work of trust manager is based on watch dog. The trust manager works by reporting alarm to neighbor node if any malicious node is present. And this reporting is done by evaluating the events reported by watch dog. In this, with the help of trust manager, malicious node can be isolated from the network. In Ming – Yang Su et[5], they proposed that every Intrusion Detection System (IDS) node will execute a Anti Black Hole Mechanism (ABM) for estimating the suspicious value of a node by calculating the abnormal difference between RREQ and RREP message transmitted from the node.IDS will broadcast a block message if the suspicious value exceeds a thresh hold asking all nodes in the network to cooperatively isolate the malicious node with the help of this mechanism, cooperative black hole nodes can be detected in MANET. In Venkat Balakrishnan et al.[6], introduced a new mode known as secure MANET Routing with Trust Intrigue (SMRTI). This model works by capturing the evidence from neighboring node in order to identify their malicious behavior. This is also done through recommendation from other nodes in MANET. Evidence can also be captured by observing the interaction of neighbors. This model consists of two component, detection and reaction. In reaction component, whether to reject or accept a newly discovered route and also to predict the future behavior of node is done here by utilizing the evidences. In Alem Y.F et al. [7], proposed a solution based on Intrusion Detection using Anomaly Detection (IDAD) to prevent both single and multiple black hole nodes. In IDAD, it is assumed that user's activity can be monitored and this user's activity is compared with intruder's anomaly activities. A pre-collected set of anomaly activities known as audit data is provided to IDAD system and if any node activity is not listed in audit data than that node is isolated from the network. In Medadian. Met al.[8], an approach is proposed for mitigating the black hole attack. By using the opinion of neighbor node, honesty of nodes is judged. Node must show its honesty in order to transfer the data packets. The node which first receive the RREP packet, initiate the judgment process on replies and forward the packet to source. This judgment is based on opinion of network nodes about replier. After receiving the opinion of neighbor, the node

decides whether the replier is malicious node or not. The drawback of this solution is that there is no guarantee that the opinion of neighbor node is always correct. In Lalit Himral et al. [9], to find the black hole nodes, secure routes are discovered by checking the sequence number. If the difference between the sequence number of source node or intermediate (who sent first RREP) is large, then the probability of that node to be malicious is more. The first RREP by any intermediate node is usually comes from malicious node. It is recommended that such node should be immediately removed from routing table. In Michiradi et al. [10], Collaborative Reputation (CORE) protocol, each node have Reputation Table (RT) and watch dog mechanism. To implement this, there are three different levels, first one is subjective reputation, reputation is calculated from direct interaction between subject and its neighbor, second is indirect reputation, and is calculated by positive report by other nodes. Third is Functional reputation, based on behavior monitored during a specific task. These three reputations decide whether to include or exclude the node in the network.

Table1: Comparison of existing solutions to Black Hole Attack

Author's Name	Method	Disadvantage
Marti et al.	Watch Dog-detects the misbehavior in the path. Path rater- rates the other node ranges from 0 to 0.8.	Can't detect the selfish node, packet dropping and ambiguous collision in the network.
Venkat Balakrishnan et al.	SMRTI- captures evidence from neighbor by observing their behavior	Opinion of neighbor is not always correct. If malicious node increases the packet delivery ratio decreases.
Medadian et al.	Opinion from neighbor nodes and use honesty of nodes	No guarantee of neighbor's node opinion.
Buchegger et al.	CONFIDANT-Trust manager evaluates based on Watch Dog	Have complex reputation index
Michiardi et al.	CORE-Reputation based	Periodic exchange of reputation information which is unnecessary as long as node behaved well and is costly also.
Deng H et al.	Advertise if particular route exist or not by further RREQ and RREP to next node.	Cooperative Black Holes cannot be prevented, routing overhead.
Ming-Yang et al.	ABM-Estimate suspicious value by calculating RREQ and RREP difference.	Time delay and have to maintain extra data base for training data and its updations.

4. Conclusion

This paper shows various works related to black hole attack for detecting and preventing in AODV routing protocol. Various security issues of MANET are studied and as well as cooperative black hole attack is also studied. A malicious node can reduce the ratio of end to end delivery. When the suspicious value of node exceeds a thresh hold value, the detected IDS broad cast the Block message to all nodes to isolate the malicious node by all other nodes in the network. Although their exist many mechanisms for detecting black hole node but all have either more time delay or network overhead due to mathematical calculations or newly introduced packets in the network. Various proposals are given for detecting and preventing of black hole attacks by various authors. But at the same time, every proposal has its own draw back.

References

[1] Hongmei Deng, Wei Li, and Dharma P.Agarwal. Routing Security in Wireless Ad Hoc network. IEEE Communication Magzine, vol 40, no.10, October 2002.

[2] Marti S, Giuli TJ,Lai K, Baker M. Mitigating routing misbehavior in mobile ad hoc networks. International conference on mobile computing and networking, August 2000. Pp 255-265

[3] Semih Dokurer, Y.M. Erten, Can Erkin. Performance analysis of ad hoc networks under black hole attacks. IEEE Conference,pp 148-153 March 2007.

[4] Buchegger S,Boudec Le J. Performance analysis of the CONFIDANT protocol, in dynamic ad-hoc networks. ACM International symposium on mobile ad hoc networking and computing (MobiHoc'02); June 2002.pp.202-236.

[5] Ming- Yang Su, Kun- Lin Chiang, Wei Cheng Liao. Mitigation of Black Hole Nodes in Mobile Ad Hoc network. Parallel and Distributed Processing with Applications (ISPA) pp.162-167, September 2010.

[6] Venkat Balakrishnan, Vijay Varadharajan, Phillip Lues, Udaya Kiran Tupakula. Trust Enhanced Secure Mobile Ad-hoc Network Routing. 21st IEEE International Conference on AINA W 2007, Niagara Falls, Canada, pp. 27-33, May 2007.

[7] Alem, Y.F.; Zhao Cheng Xuan.Preventing black hole attack in mobile ad-hoc networks using Anomaly Detection. Future Computer and Communication (ICFCC), 2010 2nd International Conference on , vol.3, no., pp.V3-672-V3-676, 21-24 May 2010.

[8] Medadian, M., Mebadi, A., Shahri, E. Combat with Black Hole attack in AODV routing protocol. Communications (MICC), 2009 IEEE 9th Malaysia International Conference on, vol., no., pp.530-535, 15-17, Dec.2009.

[9] Lalit Himral, Vishal Vig, Nagesh Chand. Preventing AODV Routing Protocol from Black Hole Attack. International Journal of Engineering Science and Technology (IJEST) Vol. 3 No. 5 May 2011.

[10] Michiardi, P. and Molva, R. Core. A collaborative reputation mechanism to enforce node cooperation in mobile ad hoc Networks. In Proceeding of IFIP TC6/TC I I Sixth Joint Working Conference on Communication and Multimedia Security, 2002, 107-121.

A Modified SVD-DCT Method for Enhancement of Low Contrast Satellite Images

G.Praveena[1], M.Venkatasrinu[2],

[1] M.tech student, Department of Electronics and Communication Engineering, Madanapalle Institute of Technology and Science (MITS), MADANAPALLE-517325,

[2] Asst.professor, Department of Electronics and Communication Engineering, Madanapalle Institute of Technology and Science (MITS),MADANAPALLE-517325

Abstract:
During the last decade, several techniques are proposed for contrast enhancement of a low-contrast satellite images. In this a new technique has been proposed based on the Singular Value Decomposition (SVD) and Discrete Cosine Transform (DCT). The proposed technique modified SVD-DCT converts the image into the SVD-DCT domain after normalizing the singular value matrix. Then the modified image is reconstructed by using inverse DCT. In the enhancement procedure Adaptive Histogram Equalization (AHE) has been used.. The perceptual and quantitative results of the proposed modified SVD-DCT method clearly indicates increased efficiency and flexibility over the exiting methods such as Linear Contrast Stretching technique, GHE technique, DWT-SVD technique, DWT technique, De-correlation Stretching technique, Gamma Correction method based techniques.

Keywords: Adaptive Histogram Equalization (AHE), Contrast enhancement, Discrete Cosine Transform (DCT) and Singular Value Decomposition (SVD).

1. Introduction

Satellite images are used in many applications such as geosciences studies, astronomy and geographical information systems. One of the problem occurs in satellite images while capturing image with a huge amount of distance, is the dark light and contrast of image.

Suppose, if an image has been taken in very dark or a very bright situation, the information may be lost in those areas which are excessively and uniformly dark or bright. Satellite images are low contrast and dark images, which has complete information but is not visible. The problem is how the contrast of an image can be improved from the input satellite images.

Image contrast enhancement is one of the most important issues in low-level image processing. Its purpose is to improve the quality of low contrast images. There have been several techniques to overcome this issue for the contrast analysis of satellite image such as General Histogram Equalization (GHE), Gamma correction and Linear contrast stretching. These techniques are very simple and effective for the contrast enhancement. But these techniques are not efficient as the information laid on the histogram of the image, which is totally lost.

The proposed technique based on the singular value decomposition (SVD) and discrete cosine transform (DCT) has been proposed for enhancement of low-contrast satellite images. In the enhancement procedure **A**daptive **H**istogram **E**qualization (AHE) has been used.

The enhancement mapping cannot improve image contrast satisfactorily since the contrast of an object is interfered by the whole image. Naturally, it is difficult to find enhance the whole image. AHE stretches the local contrast to improve the visibility of satellite images, while preserving information as it is. For this purpose, we are using **A**daptive **H**istogram **E**qualization (AHE). Here unlike **G**eneral **H**istogram **E**qualization (GHE), AHE operates on small data regions rather than the entire image. It is therefore suitable for improving the local contrast of an image and bringing out more detail [1]. The result shows that visibility improvement of specific objects is successfully enhanced using SVD-DCT method by incorporating AHE.

2. Proposed Methodology

Satellite images are low contrast and dark images, which has complete information but is not visible. The problem is how the contrast of an image can be improved from the input satellite images.

For this reason, we propose a new contrast enhancement. Basically two parts involve in the enhancement of the satellite images. The first one is **S**ingular **V**alue **D**ecomposition (**SVD)** and second one is **D**iscrete **C**osine **T**ransform (**DCT)**. The result shows that visibility improvement of specific objects is successfully enhanced using SVD-DCT method by incorporating AHE.

2.1 Adaptive Histogram Equalization:

Adaptive histogram equalization (AHE) is a computer image processing technique used to improve contrast in images. It differs from ordinary histogram equalization in the respect that the adaptive method computes several histograms, each corresponding to a distinct section of the image, and uses them to redistribute the lightness values of the image. It is therefore suitable for improving the local contrast of an image and bringing out more detail.

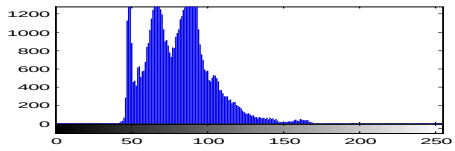
Ordinary histogram equalization uses the same transformation derived from the image histogram to transform all pixels. This works well when the distribution of pixel values is similar throughout the image. However, when the image contains regions that are significantly lighter or darker than most of the image, the contrast in those regions will not be sufficiently enhanced.

Adaptive histogram equalization (AHE) improves on this by transforming each pixel with a transformation function derived from a neighborhood region.

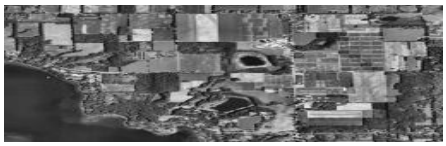
Adaptive histogram equalization is an extension to traditional Histogram Equalization technique. It enhances the contrast of images by transforming the values in the intensity image using contrast-limited adaptive histogram equalization (CLAHE).Unlike Histogram equalization, it operates on small data regions (tiles), rather than the entire image. Each tile's contrast is enhanced, so that the histogram of the output region approximately matches the specified histogram. The neighboring tiles are then combined using bilinear interpolation in order to eliminate artificially induced boundaries. The contrast, especially in homogeneous areas, can be limited in order to avoid amplifying the noise which might be present in the image.

Adaptive histogram equalization of an original image and its histogram as shown in below figure 2.1

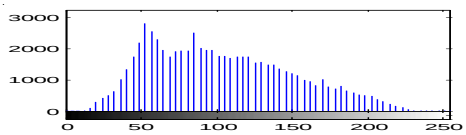
2.1.1: original



2.1.2: histogram



2.1.3: Adaptive equalized image



2.1.4: histogram equalization

Fig 2.1: Adaptive equalized image and its histogram

2.2 singular Value Decomposition (SVD)

SVD is based on a theorem from linear algebra which says that a rectangular matrix A, which is a product of three matrices that is (i) an orthogonal matrix UA, (ii) a diagonal matrix ΣA and (iii) the transpose of an orthogonal matrix VA. The singular-value-based image equalization (SVE) technique is based on equalizing the singular value matrix obtained by singular value decomposition (SVD). SVD of an image, can be interpreted as a matrix, is written as follows:

$$A = U_A \sum_A V^T{}_A \quad \text{..... (1)}$$

Where *UA* and *VA* are orthogonal square matrices and ΣA matrix contains singular values on its main diagonal [1].

Basic enhancement occurs due to scaling of singular values of the DCT coefficients. The singular value matrix represents the intensity information of input image and any change on the singular values change the intensity of the input image.

The main advantage of using SVD for image equalization, *ΣA* contains the intensity information of the image.

In the case of singular value decomposition the ratio of the highest singular value of the generated normalized matrix, with mean zero and variance of one, over a particular image can be calculated using the equation as:

$$\xi = \frac{\max(\Sigma N\ (\mu=0, var=1))}{\max(\Sigma A)} \quad \text{..... (2)}$$

By using this coefficient to regenerate an equalized image using:

$$E\ equalized_A = U_A(\xi \sum_A) V^T{}_A \quad \text{..... (3)}$$

Where E *equalized A* is used to denote the equalized image. The equalization of an image is used to remove the problem of the illumination.

2.3 Discrete Cosine Transform (DCT)

The DCT transforms or converts a signal from spatial domain into a frequency domain. DCT is real-valued and provides a better approximation of a signal with few coefficients. This approach reduces the size of the normal equations by discarding higher frequency DCT coefficients.

Important structural information is present in the low frequency DCT coefficients. Hence, separating the high-frequency DCT coefficient and applying the illumination enhancement in the low–frequency DCT coefficient, it will collect and cover the edge information from satellite images. The enhanced image is reconstructed by using inverse DCT and it will be sharper with good contrast [1].

In the proposed technique, initially the input satellite image 'A' for processed by AHE to generate $\hat{A}$. After getting this, both of these images are transformed by DCT into the lower frequency DCT coefficient and higher–frequency DCT coefficient. Then, the correction coefficient for the singular value matrix can be calculated by using:

$$\xi = \frac{max(\sum \hat{D})}{max(\sum D)} \quad \ldots\ldots (4)$$

Where $\left(\sum \hat{D}\right)$ is the lower-frequency coefficient singular matrix of the satellite input image, and $\left(\sum D\right)$ is the lower-frequency coefficient singular matrix of the satellite output image of the Adaptive Histogram Equalization (AHE).
The new satellite image (D) is determined by:

$$\left(\sum \bar{D}\right) = \xi\left(\sum D\right) \quad \ldots\ldots (5)$$

$$\bar{D} = U_D \sum \hat{D} V_D \quad \ldots\ldots (6)$$

$\bar{D}$ is the lower DCT frequency component of the original image that is reconstructed by applying the inverse operation (IDCT) to produce equalized image is

$$\bar{A} = IDCT(\bar{D}) \quad \ldots\ldots (7)$$

The performance of this method is measured in terms of following significant parameters:

$$Mean(\mu) = \frac{1}{MN}\sum_{x=1}^{M-1}\sum_{y=1}^{N-1} I(x,y) \quad \ldots\ldots (8)$$

$$standard\ deviation(\sigma) = \sqrt{\frac{1}{MN}\sum_{x=1}^{M-1}\sum_{y=1}^{N-1}\{I(x,y)-\mu\}^2} \quad \ldots\ldots (9)$$

Mean (μ) is the average of all intensity value. It denotes average brightness of the image, where as standard deviation is the deviation of the intensity values about mean. It denotes average contrast of the image. Here I(x, y) is the intensity value of the pixel (x, y), and (M, N) are the dimension of the Image.

2.4. Flow Chart:

The following flowchart shows the proposed scheme:

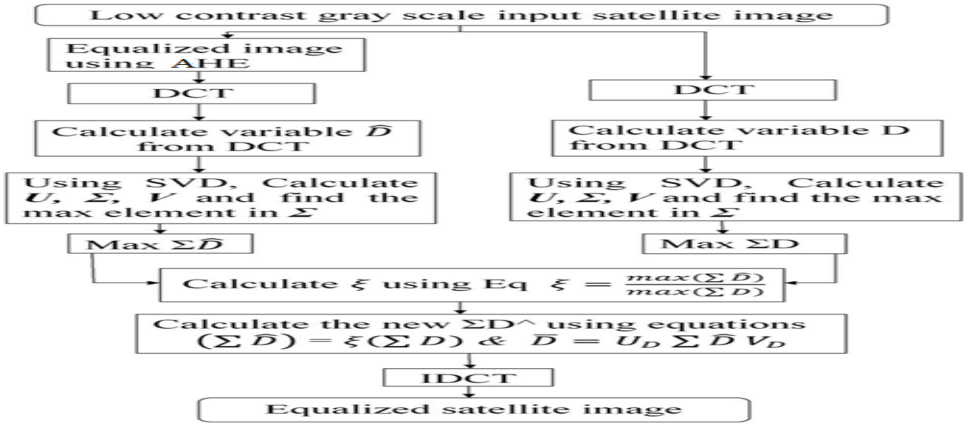


Fig 2.2 Flow chart of the proposed methodology

3. Results

The performance of this method is measured in terms of following significant parameters:

$$Mean(\mu) = \frac{1}{MN}\sum_{x=1}^{M-1}\sum_{y=1}^{N-1} I(x,y) \quad \ldots\ldots\ldots\ldots (8)$$

$$standard\ deviation(\sigma) = \sqrt{\frac{1}{MN}\sum_{x=1}^{M-1}\sum_{y=1}^{N-1}\{I(x,y)-\mu\}^2} \quad \ldots\ldots\ldots\ldots (9)$$

Mean (μ) is the average of all intensity value. It denotes **average brightness of the image**, where as standard deviation is the deviation of the intensity values about mean. It denotes **average contrast** of the image. Here I(x, y) is the intensity value of the pixel (x, y), and (M, N) are the dimension of the Image [3].

The results for the enhancement of satellite images are given. The proposed method is compared with other existing methods such as Linear Contrast Stretching technique, GHE technique, DWT-SVD technique, DWT technique, De-correlation Stretching technique, and Gamma Correction method shown in figure3.1. The visual and quantitative result shows that the proposed method has increased efficiency and flexibility.

The resultant images for the enhancement of satellite images are given below fig 3.1, the following resultant images of DCT-SVD gives the better contrast as well as high image quality. The images are:

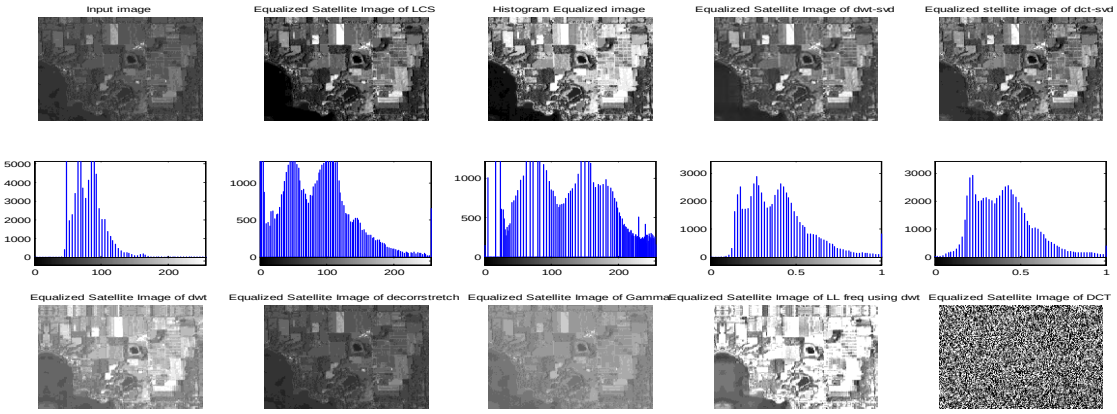


Fig 3.1: Various resultant images using existing techniques and proposed technique

The quality of the visual results indicates that the proposed technique is sharper and brighter than existing technique as compared. After obtaining mean and standard deviation, it is found that the proposed algorithm gives better results in comparison with the existing techniques. Mean (μ) represent the intensity of the image and the standard deviation represent (σ) the contrast present in the images. The proposed DCT method represents the better contrast as well as better brightness with appropriate contrast. However, the estimated mean (μ) and standard deviation (σ) in Fig 3.1 of the proposed method covers a good range of gray level and this is the cause of the better illumination. Therefore the observation of the proposed DCT gives the better result. In order to exhibit the superiority of the proposed methodology three different images have been taken for analysis. The singular values denote luminance of each image layer after decomposition using DCT methodology.

The **Mean** (μ) & **standard deviation** (σ) values are given below for analysis of this result. Here we can observe that the proposed method DCT-SVD gives better contrast as well as better brightness.

	Mean(μ)	**Standard deviation(□)**
Input image	80.1919	22.0798
DWT image	158.5717	42.1936
De-correlation Stretch Tech.	80.1919	22.0798
Gamma correction	102.1564	43.7705
LCS image	82.5653	53.7372
GHE image	129.2727	42.8887
DWT-SVD image	102.1848	46.4943
Proposed DCT-SVD image	**103.4778**	**48.0550**

Table1: Comparison of the results between different proposed methodology and already existing techniques.

4. Conclusion

In this paper, a new technique has been proposed based on the **S**ingular **V**alue **D**ecomposition (SVD) and **D**iscrete **C**osine **T**ransform (DCT) that means SVD-DCT domain for enhancement of low-contrast satellite images. The basic enhancement occurs due to scaling of singular values of the DCT coefficients. Performance of this technique has been compared with existing contrast enhancement techniques like histogram equalization, gamma correction and DWT-SVD based techniques.

From the above experimental results, it can be concluded that the proposed algorithm is effective in naturally enhancing low contrast images and the visibility improvement of specific objects compared to other existing methods. The results show that the proposed technique gives better performance in terms of contrast (variance) as well as brightness (mean) of the enhanced image as compared to the other existing techniques. Thus, this technique can be considered suitable for enhancement of low contrast satellite image.

5. Future Scope

Image enhancement of low contrast satellite images using discrete cosine transform and singular value decomposition can be implemented by using Adaptive histogram equalization is extended to color images. In case of grayscale images, DCT-SVD can be replaced by contrast stretching method and de-correlation stretch methods.

References

[1]. A. K. Bhandari, A. Kumar and P. K. Padhy,"Enhancement of Low Contrast Satellite Images using Discrete Cosine Transform and Singular Value Decomposition" in World Academy of Science, Engineering and Technology 79 2011.

[2]. Sulochana S Vidhya R, "Satellite Image Contrast Enhancement using Multi wavelets and Singular Value Decomposition(SVD)" in International Journal of Computer Applications (0975 – 8887) Volume 35– No.7, December 2011.

[3]. Hasan Demirel, Cagri Ozcinar, and Gholamreza Anbarjafari, "Satellite Image Contrast Enhancement Using Discrete Wavelet Transform and Singular Value Decomposition" in IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 7, NO. 2, APRIL 2010.

[4]. C. M. Pun, and H. M. Zhu, "Image Segmentation Using Discrete Cosine Texture Feature", International Journal of Computers, Vol. 4, No. 1, 2010, pp. 19-26.

[5]. Mr. Salem Saleh Al-amri1, Dr.N.V.Kalyankar2, Dr.S.D.Khamitkar3, "Linear and Non-linear Contrast Enhancement Image" in IJCSNS International Journal of Computer Science and Network Security, VOL.10 No.2, February 2010.

[6]. H. Demirel, G. Anbarjafari, and M. N. S. Jahromi, "Image Equalization Based On Singular Value Decomposition", Proceeding of IEEE Conference on Computer and Information Sciences, 2008, pp. 1-5.

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

[8]. T. A. Khaleel, "Enhancement of Spatial Structure of an Image by Using Texture Feature Extraction", Al-Rafidain Engineering, Vol.15, No.1, 2007, pp. 27-37.

[9]. P. S. Murty, and K.P. Rajesh, "A Robust Digital Image Watermarking Scheme Using Hybrid DWT-DCT-SVD Technique", International Journal of Computer Science and Network Security, Vol.10, No.1, October 2010, pp. 185-192.A. Sverdlovsk, S. Dexter, and A. M. Eskicioglu, "Robust DCT-SVD Domain Image Watermarking for Copyright Protection: Embedding Data in All Frequencies", Proceeding of 13th European Conference on signal processing, September 3-5, 2005, pp. 1-4.

[10]. A. Sagheer, N. Tsuruta, R. I. Taniguchi, and S. Maeda, "Hyper-Column Model vs. Fast DCT for Feature Extraction in Visual Arabic Speech Recognition", Proceeding of IEEE Conference on Signal Processing and Information Technology, 2005, pp. 761–766.

[11]. R. Reeves, and K. Kubik, "Benefits of Hybrid DCT Domain Image Matching. International Archives of Photogrammetric and Remote Sensing", Vol. 33, Part B3. Amsterdam 2000, pp. 771-778.

[12]. G. M. Hemes, S. Danaher, and A. Murray, "Characterization of Forestry Species - A Comparison Using Singular Value Decomposition (SVD) and Artificial Neural Networks (ANN)", Proceeding of IEEE Conference on image Processing and its Applications, 4-6 July 1995, pp. 815-819.

[13]. http://lisamccluremaps.blogspot.com/2008_07_01_archive.html

Video Steganography by LSB Substitution Using Different Polynomial Equations

A. Swathi [1], Dr. S.A.K Jilani, Ph.D[2]

[1](M.tech Student, [2]Professor,) Electronics and communication Engineering, Madanapalli Institute of Technology and science

Abstract:

Video Steganography is a technique to hide any kind of files into a carrying Video file. The use of the video based Steganography can be more eligible than other multimedia files, because of its size and memory requirements. The least significant bit (LSB) insertion is an important approach for embedding information in a carrier file. Least significant bit (LSB) insertion technique operates on LSB bit of the media file to hide the information bit. In this project, a data hiding scheme will be developed to hide the information in specific frames of the video and in specific location of the frame by LSB substitution using polynomial equation.

Keywords: least significant bit, Steganography

1. Introduction:

Currently, internet and digital media are getting more and more popularity. So, requirement of secure transmission of data also increased.For this reason various good techniques are proposed and already taken into practice. In this project, we use the steganography process for the secure data transmission from the sender to receiver through the internet.

2. Steganography Introduction:

Steganography is the process of secretly embedding information inside a data source without changing its perceptual quality.Steganography comes from the Greek word ***steganos*** which literally means "covered" and ***graphia*** which means "writing", i.e. covered writing. The most common use of steganography is to hide a file inside another file.

Generally, in data hiding, the actual information is not maintained in its original format.The format is converted into an alternative equivalent multimedia files like images, video or audio. Which in turn is being hidden within another object[7].

3. Block diagram:

The basic block diagram representation for steganography mechanism is shown in the below figure.

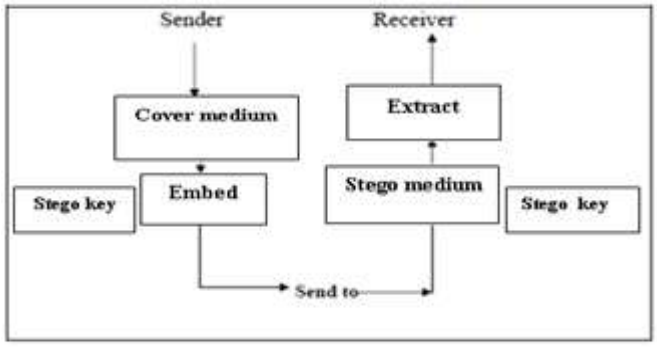


Fig 3. 1: Steganography mechanism

3.1 Block Diagram Explanation:

The above figure shows a simple representation of the generic embedding and extraction process in steganography. In this example ,a secret data is being embedded inside a cover image to produce the stego image. A key is often needed in the embedding process.The embedding procedure is done by the sender by using the proper stego key. The recipient can Extract the stego cover image in order to view the secret data by using the same key used by the sender. The stego image should look almost identical to the cover image.

4. Least Significant Bit Insertion Method:

Least Significant Bit (LSB) insertion is a common, simple approach to embedding information in a cover video. Video is converted into a number of frames, and then convert each frame in to an image[6]. After that, the Least Significant Bit (in other words the 8 bit) of some or all of the bytes inside an image is changed to a bit of each of the Red, Green and Blue colour components can be used, since they are each represented by a byte. In other words one can store 3 bit in each pixel. An 800 x 600 pixel image can thus store a total amount of 1,440,000 bits or 180,000 bytes of embedded data. We implemented our project such that it can accept and video of any size.

For example a grid for 3 pixels of a 24 bit image can be as follows:

(00101101 00011100 11011100)
(10100110 11000100 00001100)
(11010010 10101101 01100011)

When the letter A, which binary representation is 01000001 and is embedded into the least significant bits of this part of the image, the resulting grid is as follows:

(00101100 00011101 11011100)
(10100110 11000100 00001100)
(11010010 10101101 01100011)

Although the letter was embedded into the first 8 bytes of the grid, only the 2 highlighted bits need to be changed according to the embedded message. On average only half of the bit in an image will need to be modified to hide a secret message using the maximum cover size.

First we read the original video signal and text. We have to embed the text into the video signal. Then we have to convert the text data into the binary format. Binary convertion is done by taking the ASCII

Value of the character and converting those ASCII values into binary format. We take the binary representation of samples of cover signal and we insert the binary representation of text into that cover signal

The LSB bits of video signals are replaced by the binary bits of data and this encoded signal is called stego signal is ready for transmission through internet. For the steganography the important video format is Audio Video Interleave (AVI). The message which we want to hide is converted into ASCII and then converted into its binary representation with each word consist of 8bits.These bits are substituted in the Least Significant Bits of binary representation of each image ample. Suppose if we want to hide letter A (01000001) in LSBs of binary signal.

Here the polynomial equations are used to find the location of insertion of message binary bit in the video file. Below process shows finding the location by using polynomial equations. After finding the location in audio file Least Significant Bits are replaced as below process.

Y=X	M1=3X+5	M2=2X+5
123	if X=1→M1= 8	M2=7
124	X=2→M1= 11	M2=9
126	:	:
128	:	:
135	X=10→M1= 35	M2=25
144		
135	-------→10111000-----→10111000	
156	11111110 **AND** operation by 254th bit	
173	----------	
192	10111000	
203	1 **OR** operation by message bit	
257	----------	
269	10111001	
288	If the message is **mits**., ASCII(m)-129 →(10000001)	

If we get same locations by two polynomials
Then 2 will be added to location.
M1=22 M2=22 then M=22+2=24

01110001 10011000 11001000 00110110
10001101 11000111 01010011 10100011

becomes

01110000 10011001 11001000 00110110
10001100 11000110 01010010 10100011

5. Flow Chart for encoding:

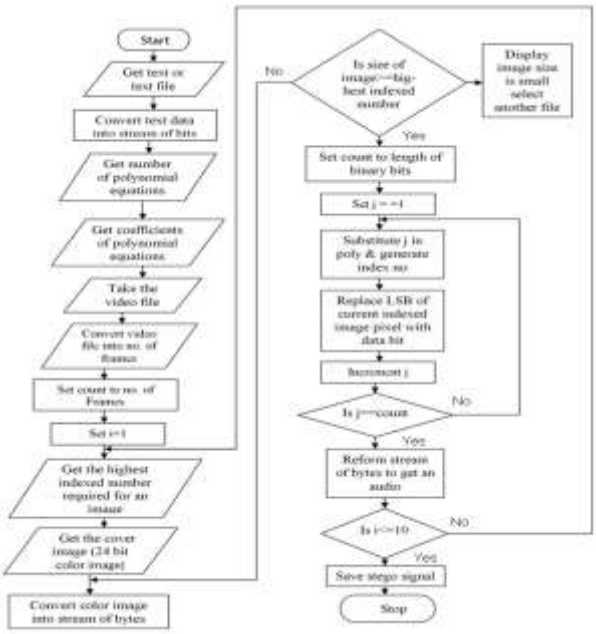


Fig 5.1 : Flow chart for encoding

Description about Encoding:

First we will take the original video and text file in which we have to embed into original image. Then we have to convert the video file into number of frames, we consider each frame as an image. Here we set the counter value to frames. Then we have convert the text data into binary format. Binary convertion is done by taking the ASCII value of each character and converting those ASCII values into binary format. We are going to set the counter value to the length of the binary message, so that the loop re[peats that much times. The LSB bit of the image pixel is replaced by the binary data. This encoded image called as stego video is ready for transmission through the internet.

5.1 Flow Chart for decoding:

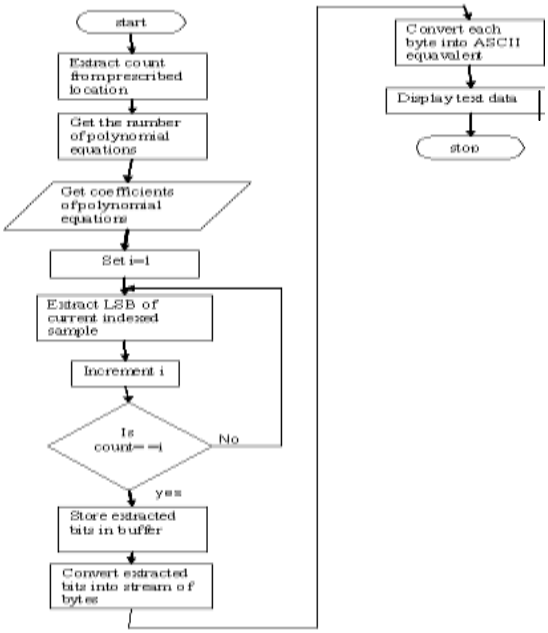


Fig 5.2 : Flow chart for Decoding

Description about Decoding:

First we take the LSB encoded image. We will set the counter to the length of the binary data. Then we will extract the binary data from the LSB encoded image by extracting the LSB bits of the image pixels. In order to form the text file from the binary data we have to group all the binary bits.

Cover Video

Stego Video

Description:

By observing the above two figures we can simply say that both stego and cover videos are identically equal.

6. Conclusion:

There are many kinds of steganography techniques are available among them hiding data in video by LSB substitution is a simple method . Here the information will be embedded based on the stego key. Key is used in the form of polynomial equations with different coefficients. By using this the capacity of embedding bits into the cover image can be increased.

References

[1] Ping Wah Wong and Edward J. Delp, editors. Security and Watermarking of Multimedia Contents, volume 3657.Society of Photooptical Instrumentation Engineers, 1999.

[2] Ping Wah Wong and Edward J. Delp, editors.Security and Watermarking of Multimedia Contents II, volume 3971.Society of Photo-optical Instrumentation Engineers, 2000.

[3] W. Bender, D. Gruhl , N. Morimoto , A. Lu : Techniques for data hiding, IBM SYSTEMS JOURNAL, VOL 35, NOS 3&4, 1996

[4] R. Anderson, F. Petitcolas : On the limits of the steganography, IEEE Journal Selected Areas in Communications, VOL .16, NO. 4 ,MAY 1998.

[5] FABIEN A. P. PETITCOLAS, ROSS J. ANDERSON, AND MARKUS G. KUHN, Information Hiding—A Survey, PROCEEDINGS OF THE IEEE, VOL. 87, NO. 7, JULY 1999.

[6] Deshpande Neeta, KamalapurSnehal, Daisy Jacobs: Implementation of LSB Steganography and Its Evaluation for Various Bits ,2001.

[7] NedeljkoCvejic, TapioSeppben : Increasing the capacity of LSB-based video steganography ,IEEE 2002.

[8] Dr D Mukhopadhyay, A Mukherjee, S Ghosh, S Biswas, P Chakarborty: An Approach for Message Hiding using Substitution Techniques and video Steganography, IEEE 2005.

[9] Zhou Lin-na, Liu Cui-qing, Ping Xi-jian, Wang Yun-he : Information Hiding Method in Digital Audio Signal Based on Perfect Codes , Information Hiding and Multimedia Signal Processing, IEEE 2006.

High Speed Efficient Data Transmission in MPLS Ring Network

Amit, [1] Mr. Shamsher Singh[2]

[1]Student, [2]Assistant Professor, University Institute of Engineering & Technology,
Maharshi Dayanand University, Rohtak

Abstract:
MPLS Transport Profile is gaining importance as it becomes a dominant solution for a converged transport network.In a high speed network, any disturbance over the network cause heavy data loss. When data is transferring a topology like bus or ring. The chance of data loss increases more in same ratio. In this present the discussion is being performed on high speed MPLS Ring network. As some fault occurs over the network it will return heavy data loss. In this work we are defining some new low cost hardware called save points. Save points are placed at equal distance between two nodes. The network will maintain two paths for reliable data transmission. One path is the actual data path and other is backup path. As the signal disruption detected, it will inform the nearest save point and the save point will direct the communication to the backup path. The present article describes the implementation in detail.

Keywords: MPLS, Save Point, Broken Link, Recovery

I Introduction

Based on the exponential growth of data traffic, which is mainly IP packet traffic, it is generally considered that the next-generation transport network will be a pure packet-based network.MPLS stands for Multiprotocol Label Switching, Multiprotocol because it can be used for any Layer 3 protocol. MPLS is about glueing connectionless IP to a connection oriented Network. MPLS is something between L2 and L3. MPLS is also called Tag Switching MPLS provides a mechanism to avoid hop-by-hop routing decision making (Notice that IP makes a hop-by-hop routing decisions) by setting up a Layer 2 fast path using Labels (hence the name Label Switching) to move the packets quickly along pre-established paths without examining each packet at the IP level at each router. This is very similar to ATM and FR Layer 2 routing and switching operations. Remember VPs and VCs in ATM and DLCI #s in FR operate at Layer 2, have only local significance and established before hand. In MPLS enabled Network, the packet needs to be examined only once at the edge of the MPLS Network (Network entry point). After that, the packet forwarding is based on simple tagging scheme rather than on more complex and variable IP header. The most important advantage of MPLS is that it is independent of L2 and L3 protocols, and so

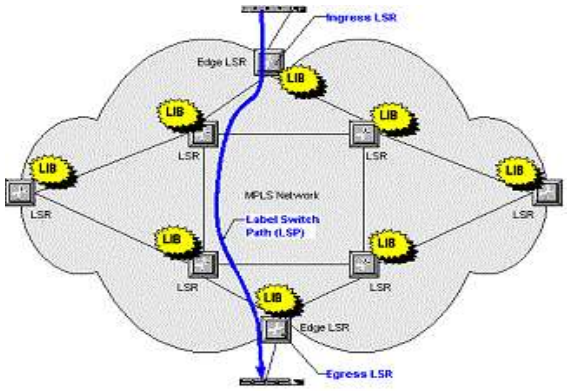


Figure 1: MPLS Network

It can be adopted any where in the world to any L2 or L3 infrastructure. The Canadian Post Office seems to work on this principle. MPLS is now increasingly being deployed in Optical Networking as well. Here figure 1 represents the MPLS network.

A) Route Selection: In MPLS network the route selection is done using basic 2 approaches

I) Hop-by-Hop Routing: In this case each LSR independently chooses the next hop for each FEC. The LSRs use existing OSPF protocol to make the routing decisions. This method allows rapid switching by labels and support for differentiated service. How ever it does not support traffic Engineering and policy Routing (Policy Routing refers to defining Routes on some policy related to QoS, Security, Queuing and discard mechanisms etc).

Ii) Explicit Routing:

In this kind of Routing, a single LSR, usually the ingress or egress LSR, specifies some or all of the LSRs in the LSP for a given FEC. Explicit Routing provides all the benefits of MPLS, including Traffic Engineering and policy Routing.

It is widely recognized [2] that survivability is crucial in MPLS-TP networks to deliver availability and reliability traffic to support the requirements of a service level agreement (SLA). As MPLS-TP technology seeks its way to popularize carrier transport networks, MPLS-TP has driven the need for a powerful set of protection tools.

B) MPLS Header:

- MPLS sticks a 32 bit Header as shown below:
- Label (20 Bits): Carries information about setting up the LSP based on a given FEC. Either Manual configuration or inserted on the basis of static or dynamic routing protocol
- COS (3 bits):Defines the required Class of Service. With 3 bits, 8 possible classes. Affects the Queuing and discard algorithms as the packet travels through the Network. Some times it is copied from the IP Header TOS field.
- S (STACK) (1 BIT): MPLS supports multiple Labels. The processing of labelled stack is always based on the top stack. We shall study this feature of MPLS in more detail shortly.
- TTL (Time To Live) (8 bits): Same function as TTL in the IP Header. The packet is killed after travelling the specified number of hops and if the packet has still not reached is destination.

In today's network world the advancement of the network infrastructure is the sound base that networks must rely on to compete in society. MPLS offers that sound base at a lower cost and with enhanced flexibility. It not only can use pre-existing equipment used for technologies such as ATM and Frame Relay, but it allows the fine tuning of more advanced technologies such as VoIP or video conferencing. Versatility and redundancy is what makes a network reliable and flexible, and MPLS can offer those solutions. It is not a matter of consideration; it is a matter of time and necessity in world today.

C) MPLS-TP

Tomorrow's network will mostly carry packets. As a result, an evolution of existing time-division multiplexing (TDM)-based transport networks is taking place, and new architectures optimized to carry packets are being defined. The function of a transport network is to carry information between service edge devices. These devices could be Digital Subscriber Line Access Multiplexers (DSLAMs), gateways, T1/E1 aggregators, broadband remote access servers (BRAS), etc. Traditional transport systems based on SDH/SONET platforms provide low-speed bandwidth granularity network services as well as high-speed long-haul transmission services. Circuit-switched transport network services with fixed bandwidth granularity (64 Kbps, 1.5 Mbps, 2 Mbps, 50 Mbps, 150 Mbps, 600 Mbps, etc.) were emulated using connection-oriented, packet-switched (CO-PS) technologies and similar managed-bandwidth services. However, in the access/aggregation and metro domains, there is a desire by carriers to simplify packet transport networking in order to reduce capital expenditures (CapEx) and operational expenses (OpEx) in their next-generation networks.[5] A protected domain is defined betweentwo or more protection reference endpoints whichare located at the edges of the protected domain and bounds the element on which protection can be provided. This element can be a span, a section, or an end-to-end path. The protected domains locate between a pair of neighboring MPLS-TP nodes in
the same layer network under span-level configuration.

Ii Literature Survey

MPLS-TP ring protection mechanisms as a novel protection mechanism, which aggregates merits of both wrapping and steering approaches, and which reduces packet loss significantly in case of in order delivery. The present article describes the implementation in detail and shows that the new approach highly reduces packet loss compared to the current mechanism[1]. A new architecture for MPLS-based micro mobility management. Presented proposal called Optimized Integrated-Multi-Protocol Label Switching (Optimized I-NMLPLS). Optimized I-NMPLS is a hierarchical approach to support micro-mobility. This approach integrates the traffic engineering and QoS capabilities of MPLS with Mobile IP Fast Authentication protocol (MIFA) as a mobility management framework. The integration between N1LPLS and MIFA is achieved through integration of MIFA control messages with MPLS signalling traffic[2]. MPLS is a next generation backbone architecture, it can speed up packet forwarding to destination by label switching. However, if there exists no backup LSP when the primary LSP fails, MPLS frames cannot be forwarded to destination. Therefore, fault recovery has become an important research area in MPLS Traffic Engineering[3]. A MPLS based routing algorithm is present for reducing the number of MPLS labels to N + M without increasing any link load. Presented explicit N + M bound makes it easy to limit the table size requirement for a planed network, and the linearity allows for tables implemented in fast memory. For differentiated services with K traffic classes with different load constraints, Presented bound increases to K(N+M). Presented stack-depth is only one, justifying implementations of MPLS with limited stack-depth[4].

This paper describes design, implementation, and capability of a MPLS simulator, which suppaas label swapping operation, LDP, CR-LDP, and various sorts of label distribution function. It enables researchers to simulate how a U P is established and terminated, and how the labelled packets act on the LSP. In order to show MPLS simulator's capability, the basic MPLS function defined in MPLS standards is simulated; label distribution schemes, flow aggregation, ER-LSP, and LSP Tunnel[5]. They introduces a T-MPLS simulation tool developed using OPNET modeler 11.5. Author proposed a

simulation node structure including four layers, i.e. Client Layer, Adaptation Layer, Control Layer and the Switching & Forwarding Layer. All the L2 (PDH, SONET/SDH, ETH, FR, ATM, etc) and L3 (IP) payload could be mapped into T-MPLS tunnel through either MPLS encapsulated method or pseudo wire mechanism[6]. Author provide some performance measurements comparing the prototype to software routers. The measurements indicate that the prototype is an appropriate tool for achieving line speed forwarding in testbeds and other experimental networks[7] A work on Mobile MPLS with Route Optimization is presented by the author and proposes a route optimization protocol to overcome this problem in Mobile MPLS. By adding a correspondent agent function to Mobile MPLS's edge routers, the mobility binding of a mobile node can be cached by the edge routers and the packet routing to the mobile node can be route-optimized[8]. Another work present an innovative multi-service testbed for MPLS-TP networks to demonstrate the transport performance of interconnected rings. By validating dual-label and two extended protection schemes, hardware experimental results are given to show their influences on transport[9].

III. Research Methodology

The proposed work is about to analyze the network for the network fault and provide the network restoration scheme by defining the alternative path. The work also includes the distribution of substituted path notification among all nodes over the network. The network reliability depends on how accurately the fault information is defined and the QOS depends on the efficient distribution of notification over the network.

In the existing methodology as the packet loss is detected in such network it is detected by the last save point that the information is lost. Now the save point will regenerate the packet and send it in the opposite direction and from the other path the communication will be performed. But the main drawback of this approach is the complete re-routing process in worst case. It means if the data will be lost near to the destination node and save point performs the communication from reverse direction. Now the packet has to travel the complete cycle again. The overall algorithm for the proposed work is presented in the form of a flowchart given as under.

Figure 2: Proposed Methodology

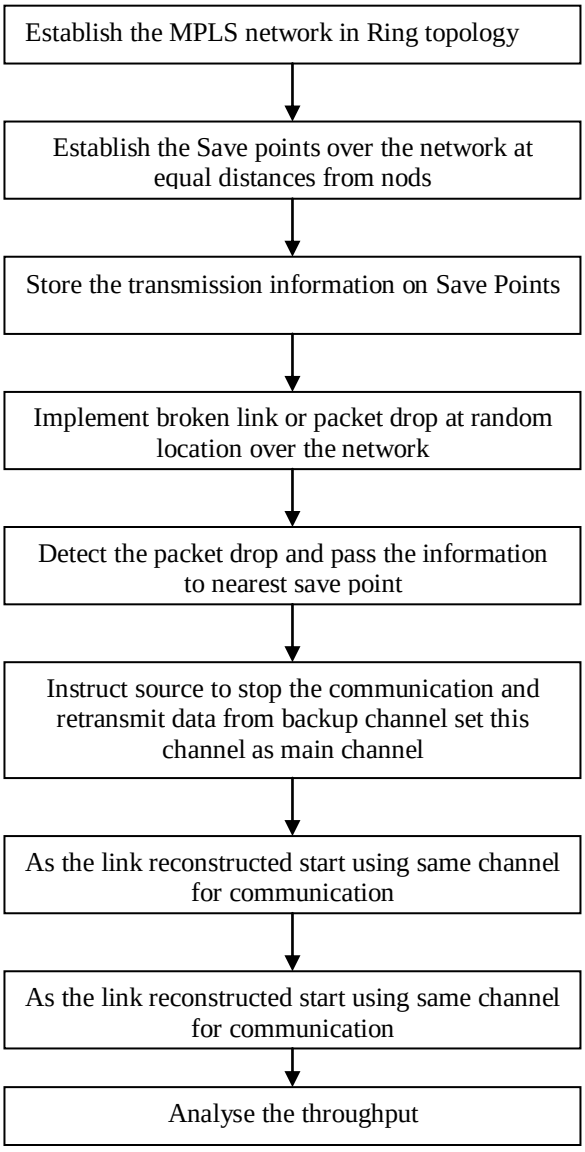


Figure 2 shows the complete architecture of MPLS network. In this proposed work we are defining a 2 Lane System for the network communication. The properties of proposed system are

- The proposed system is a 2 Lane system.
- Each Channel can send or receive data.
- Both channels are independent to the communication.
- Second channel is defined as the backup channel.
- The main communication will be performed using main channel.

- As the packet loss detected by the save point. Channel 2 will be selected for the data communication.
- Same Save point will control both channel.
- If the packet loss occur in second channel also then it will chose the traditional approach to move in opposite direction and perform the communication

IV. Conclusion

The proposed work is the improvement over the existing MPLS network with an dual path approach. The system will use extra hardware as well as extra backup path and gives the better reliability and speed over the network as packets will be delivered in lesser time with lesser packet loss.The proposed system is an intelligent system that will work on ring based high speed networks.

References

[1]. Wikipedia, "Internet backbone". Free encyclopedia of information [Online]. Available: http://en.wikipedia.org/wiki/Internet_backbone [Accessed: March 25th, 2009]

[2]. Juniper Networks, "Technical Documentation: JUNOS 5.3 Internet Software Configuration Guide: MPLS Applications," Juniper Networks Inc, 1999-2009. [Online]. Available: http://www.juniper.net/techpubs/software/junos/junos53/swconfig53-mpls-apps/html/swconfig53-mpls-appsTOC.htm. [Accessed: April. 17, 2009].

[3]. InetDaemon, "Tutorials by InetDaemon: BGP Update Message" InetDaemon", 1996. [Online]. Available: http://www.inetdaemon.com/tutorials/internet/ip/routing/bgp/operation/messages/update/nlri.shtml. [Accessed: April. 20, 2009].

[4]. H. Jonathan Chao, Xiaolei Guo. Quality of service control in high-speed networks, illustrated ed. Reading, MA: Wiley-IEEE, 2001. [E-book] Available: Google e-book.

[5]. Cisco White Paper, "Understanding MPLS-TP and Its Benefits, © 2009 Cisco Systems, Inc. All rights reserved. This document is Cisco Public Information.

[6] Wenjun Xie, Shanguo Huang, Wanyi Gu, "AN IMPROVED RING PROTECTION METHOD IN MPLS-TP NETWORKS", *Proceedings of IC-NIDC2010*, 978-1-4244-6853-9/10 ©2010 IEEE

[7] Ali Diab, Rene Boringer, "Optimized I-MPLS: A Fast and Transparent Micro- Mobility-Enabled MPLS Framework", 1-4244-0398-7/06 §2006 IEEE

[8] Yimin Qiu, "A Research of MPLS-based Network Fault Recovery", 2010 Third International Conference on Intelligent Networks and Intelligent Systems 978-0-7695-4249-2/10 © 2010 IEEE

[9] David Applegate, "Load optimal MPLS routing with N + M labels", 0-7803- 7753-2/03 (C) 2003 IEEE

[10] Gaeil Ahn and Woojik Chun, " Design and Implementation of MPLS Network Simulator Supporting LDP and CR-LDP", 0-7695-0777-8/00 *0* 2000 IEEE

[11] Bin Li, Yongjun Zhang, Shanguo Huang, Wanyi G, "Cascaded Packet Transfer Schemes to Improve Wireless T-MPLS Network Bandwidth Efficiency", 978-1-4244-3693-4/09 ©2009 IEEE

[12] James Kempf, Scott Whyte, Jonathan Ellithorpe, Peyman Kazemian, "OpenFlow MPLS and the Open Source Label Switched Router", 978-0-9836283-0-9 c 2011 ITC

[13] Savinya Polvichai, "Mobile MPLS with Route Optimization: The Proposed Protocol and Simulation Study", 20ll Eighth International Joint Conference on Computer Science and Software Engineering (JCSSE) 978-1-4577-0687-5/11 ©2011 IEEE

[14] Zhihui Zhang Yongjun Zhang Wanyi G, "Demonstration of Transport and Protection Schemes in a Multi-service Testbed for MPLS-TP Networks", 978-1- 4244-6554-5/11 ©2011 IEEE

[15] IEEE. IEEE standard 802.17: Resident Packet Ring *2*004.

Application of Multiobjective Particle Swarm Optimization to maximize Coverage and Lifetime of wireless Sensor Network

[1]Deepak Kumar Chaudhary, [1]Professor Rajeshwar Lal Dua
[1] M.Tech Scholar, [2] Professors, HOD,
Department of Electronics & Communication Engineering, JNU, Jaipur, India

Abstract

The autonomous nodes used for monitoring environments are known as Wireless Sensor Networks. The major use of Multi-Objective particle swarm optimization for maximizing coverage & life time using wireless sensor networks is described. Wireless sensor networks are being tremendous popular due to increasing number of applications for these networks. This paper introduces a proposal to extend heuristic called "Particle Swarm Optimization "(PSO) to deal with Multi-objective optimization problems.

This paper describes a new approach for energy efficient layout of wireless sensor network in which sensors communicate with each other to transmit their data to a high energy communication node which acts as an interface between data processing unit and sensors. Optimization of sensor locations is essential to provide communication for a longer duration. It discusses an energy efficient layout with good coverage based on Multi-objective Particle Swarm Optimization algorithm.

Keywords: Heuristic, Multi-objective Optimization (MOO), Particle Swarm Optimization (PSO), Wireless Sensor Network (WSN).

I. Introduction

The use and development of heurioptististic based multiobjective optimization a technique has significantly grown in last few years. Particle swarm optimization (PSO) is a reactively recent heuristic inspired by the choreography of a bird flack. Despite its current success in diverse optimization tasks, PSO remains as one of the heuristics for which not much work on multiobjective optimization has been done so far.

The main purpose of the sensor placement problem is to determine optimal locations of sensors for maximizing the information collected from the sensor network. But equally important are factors such as the energy consumption of the network. This results in a tradeoff between multiple objectives such as coverage and energy consumption.

In this paper, we address the optimization of the multiple objectives described above by adopting a multi-objective optimization framework. The approach is very simple to implement, it is population based, it uses an external memory called "repository" and a geographically-based approach to maintain diversity-MOPSO.

II. Particle Swarm Optimization

A. The original version

The particle swarm concept originated as a simulation of a simplified social system. The original intent was to graphically simulate the graceful but unpredictable choreography of a bird flock. These methods were proposed by Kennedy and Eberhart (1995). At some point in the evolution of the algorithm, it was realized that the conceptual model was, in fact, an optimizer.

PSO is similar to a genetic algorithm (GA) in that the system is initialized with a population of random solutions. Each particle keeps track of its coordinates in the problem space which are associated with the best solution, it has achieved so far. This fitness value is called pbest. Another "best" value that is tracked by the global version of the particle swarm optimizer is the overall best value, and its location, obtained so far by any particle in the population. This location is called gbest. The particle swarm optimization concept consists of, at each time, step, changing the velocity (accelerating) each particle toward its pbest & gbest locations (global version of PSO).

B. Implementation of global version of PSO

The (original) process for implementing the global version of PSO is as follows:

1. Initialize a population (array) of particles with random position & velocities (on d dimensions in the space).
2. For each particle, evaluate the desired optimization fitness function in d variables.
3. Compare particle's fitness evaluation with particle's pbest. If current value is better than pbest, the set pbest value equal to the current value, and the pbest location equal to the current location in d-dimensional space.

4. Compare fitness evaluation with the population's overall previous best. If current value is better than gbest, then reset gbest to the current particles array index and value.
5. Change the velocity and position of the particle according to equation (1) & (2) respectively.

$$V_{id} = V_{id} + C1*\text{rand}(\,) * (P_{id} - X_{id}) + C2 * \text{Rand}(\,) * (P_{gd} - X_{id}) \quad (1)$$

$$X_{id} = X_{id} + V_{id} \quad (2)$$

6. Loop to step (2) until a criterion is met, usually a sufficiently good fitness or a maximum number of iterations.
   Particles velocities on each dimension are clamped to maximum velocity V_{max} . If the sum of accelerations would cause the velocity on that dimension to exceed V_{max}, then the velocity on that dimension is limited to Vmax.
   If V_{max} is too high, particles might fly past good solutions. If V_{max} is too small, particles may not explore sufficiently beyond locally good regions. The acceleration constants C1 & C2 in equation (1) represent the weighting of the stochastic acceleration terms that pull each particle toward pbest and gbest positions.

III. Problem formulation

A. Modeling of Wireless Sensor Network

It is assumed that each node knows its position in the search space and all sensor nodes are homogeneous. High energy communication node (HECN) is assumed to be more powerful than sensor nodes. A flat square surface is considered in which HECN is placed at the center for convenience. The sensing area of each node is assumed to have a circular shape with radius R_{sens}. The communication range of each node is defined by the area of a circle with radius R_{comm} .Initially nodes are assumed to have equal energy. It is assumed that for every data transmission, the energy decreases by one unit. The co-ordinates of the sensor nodes (x_1,y_1), (x_2,y_2). . . . . are considered as design variables. Sensor nodes are assumed to have certain mobility.

B. Objectives of MOPSO

Two major objectives of MOPSO are considered.

• Maximize the total coverage of the sensor network: f1

• Maximize the lifetime of the sensor network: f2

Coverage

It is one of the measurements of quality of service (QoS) of a sensor network. The coverage of each sensor can be defined either by a binary sensor model is shown in Figure .1 or a stochastic sensor model is shown in Figure 2. In the binary sensor model shown in Figure 1, the detection probability of the event of interest is one within the sensing range; otherwise, the probability is zero. Coverage of a network using binary sensor model is determined by finding the union of sensing areas defined by location of each sensor and Rsens. Although the binary sensor model is simpler, it does not take the uncertainty factor in sensor measurement into consideration. The binary sensor model is given by

$$C_{ij}(x,y) = \begin{cases} 1 & \text{for } d_{ij}(x,y) \le R_{sens} \\ 0 & \text{for } d_{ij}(x,y) > R_{sens} \end{cases} \quad (3)$$

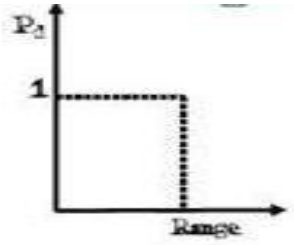


Figure 1: A binary sensor coverage model

The sensor field is represented by an mxn grid. $d_{ij}(x,y)$ denotes the Euclidean distance between a sensor node at (x, y) and any grid point at (i, j). The distances are measured in units of grid points. Equation 3 expresses the coverage $C_{ij}(x,y)$ of a grid point at (i, j) by a sensor at (x, y). The coverage for the entire grid is calculated as the fraction of grid points covered. In reality, sensor measurements are imprecise; hence the coverage needs to be expressed in probabilistic terms. In the stochastic sensor model shown in Figure 2, the probability of detection follows an exponential decaying function of distance from the sensor. The stochastic sensor model is given in Equation 4.

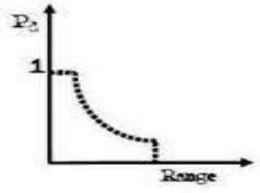


Fig 1: A binary sensor

$$C_{ij}(x,y) = \begin{cases} 1 & \text{for} \quad d_{ij}(x,y) \le (R_{sens} - R_e) \\ e^{(-\lambda\alpha)^\beta} & \text{for} \quad (R_{sens} + R_e) < d_{ij}(x,y) < R_{sens} \\ 0 & \text{for} \quad d_{ij}(x,y) \le (R_{sens} - R_e) \end{cases} \qquad (4)$$

R_e ($R_e < R_{sens}$) is a measure of the uncertainty in sensor measurement, a=$d_{ij}(x,y) - (R_{sens} - R_e)$, and λ and β are parameters that measure the detection probability when there is an uncertainty in sensor detection. The coverage for the entire sensor field is calculated as the fraction of grid points that exceeds the threshold C_{th}. So the first objective is maximization of coverage. This objective can be calculated by the following expression:

$$\textbf{Max Coverage (f1)} = \frac{U_{i=1,\ldots,N}}{A} A_1 \qquad (5)$$

where A_i is the area covered by the ith node, N is the total number of nodes and A is the area of the region of interest.

Lifetime

The second objective considered is maximization of lifetime. Lifetime is defined as the time until one of the participating nodes run out of energy. This objective can be calculated by the subsequent expression:

$$\textbf{Max Lifetime (f2)} = \frac{T_{failure}}{T_{max}} \qquad (6)$$

where $T_{failure}$ is the maximum number of sensing cycles before failure of any node and Tmax is the maximum number of possible sensing cycles. In every sensing cycle, the data from every node is routed to HECN through a route of minimum weight. Dijkstra algorithm is used to find out the route with minimum weight. These two objectives are competing with each other. The coverage objective will try to spread out the nodes for maximizing coverage while resulting in high energy loss and small lifetime. The lifetime objective will try to arrange the nodes as close as possible to the HECN for reducing loss in energy which results in poor coverage.

III. Proposed Multi-objective particle swarm optimization

A. Description of the proposed approach

A Pareto ranking scheme could be the straightforward way to extend the approach to handle multiobjective optimization problems. We implemented a mechanism such that each particle may choose a different guide. Our mechanism is based on the generation of hypercube which are produced dividing the search space explored.

In order to construct a direct relationship between the problem domain and the PSO particles for this problem, every particle represents coordinates of N number of nodes. So each particle represents a network layout. The proposed MOPSO algorithm is composed of the following steps:

1. Initialize the population pop:
   (a) For i=0 to Max /*Max=number of particles*/
   (b) Initialize pop[i]
2. Initialize the speed of each particle:
   (a) For i=0 to Max
   (b) VEL[i] = 0
3. Evaluate each of the particles in pop.
4. Store the positions of the particles that represent nondominated vectors in the repository REP.
5. Generate hypercubes of the search space and locate the particles using these hypercubes as coordinate system.
6. Initialize the memory of each particle. This memory is also stored in the repository.
   (a) For i=0 to MAX (b) PBESTS[i] = pop[i]
7. WHILE maximum number of cycles has not been reached DO
   (a) Compute the speed of each particle using the following expression.
   VEL[i] = W×VEL[i] + R1 × (PBESTS[i]-pop[i]) + R2× (REP[h]-pop[i])
   Where W (inertia weight) takes a value of 0.4. R1 & R2 are random numbers in the range (0 to 1); PBESTS[i] is the best position that the particle i has had; REP[h] is a value that is taken from the repository; the index h is

selected in the following way: those hypercubes containing more than one particle are assigned a fitness equal to the result of dividing any number x>1 (we used x=10 in our experiments) by the number of particles that they contain. Pop[i] is the current value of particle i.

(b) Compute the new position of the particles adding the speed produced from various step:

$$POP[i] = POP[i] + VEL[i] \quad (i)$$

(c) Maintain the particles within the search space in case they go beyond its boundaries.

(d)Evaluate each of the particles in POP.

(e) Update the contents of REP together with the geographical representation of particles within the hypercubes.

(f) When the current position of the particle is better than the position contained in its memory, the particle's position is updated using:

$$PBEST[i] = POP[i] \quad (ii)$$

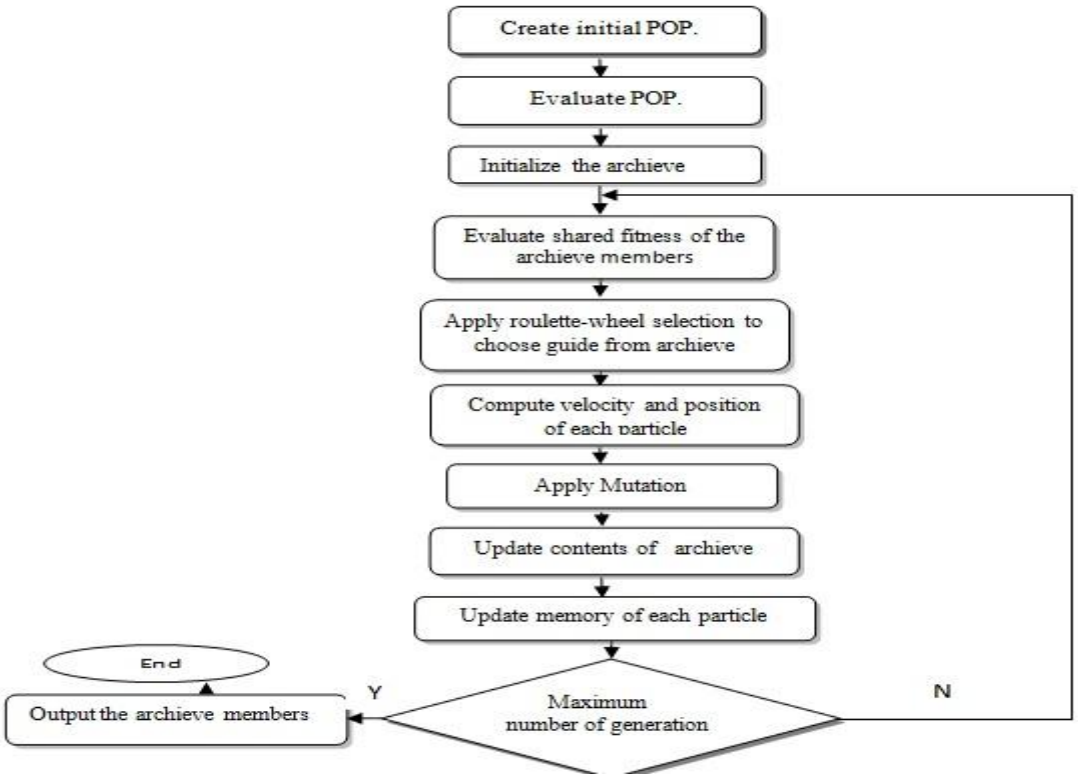
The criterion to decide what position from memory should be retained is simply to apply Pareto.

(g) Increment the loop counter.

(h) END WHILE

B. Flowchart of Multi Objective Particle Swarm Optimization

MOPSO that used in this study is dominance-based method that was proposed by Coello et al. In this algorithm the best nondominated solutions have ever been visited are stored in a memory space calling archieve. The main objective of every multi-objective optimization algorithm is to find Pareto-optimal set. Traditional type of assigning fitness function is aggregation–based method, where the fitness function is a weighted sum of the objective functions. However this classic approach can be very sensitive to precise aggregation of goals and tend to be ineffective and inefficient. The flowchart of this algorithm is shown in Fig. 3.



V. Results

The MOPSO algorithm starts with a "swarm" of particles randomly generated where each particle represents a network layout represented by sensor co-ordinates. The coverage and lifetime of the particles are then calculated. The archive containing non-dominated Pareto optimal set of solutions is developed according to the Pareto optimal dominance developed by Coello and Lechuge.

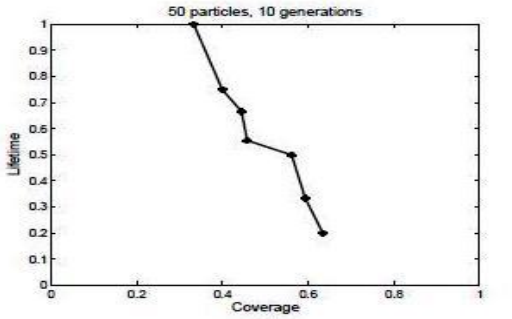


Figure 4: Pareto front for a WSN with 10 sensors

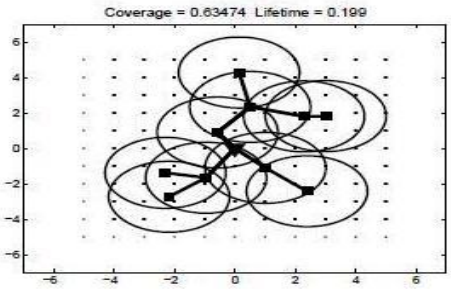


Figure 5: Pareto-optimal layout with best coverage for WSN with 10 sensors, 50 particles,10 generations

Hence sensors acting as communication relay loss more energy. The layout shown in Figure 6 is example of another pareto optimal layout available to the user. It is more interesting to look at the Pareto fronts obtained using two different sensor models as shown in Figure 7.

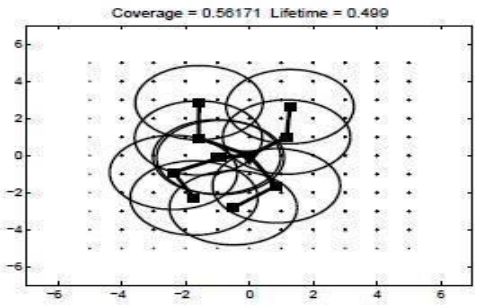


Figure 6: Example of another pareto-optimal layout for a WSN with 10 sensors, 50 particles, 10 generations

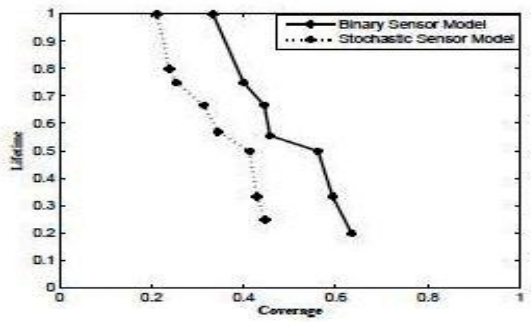


Figure 7: Pareto fronts obtained with different sensor models for a WSN with 10 sensors, 50 particles, 10 generations

The input parameters taken for simulation are as follows: grid size = 10*10, number of nodes = 10, number of particles = 50, number of generations = 10, R_{sens} = 2, R_{comm} = 2, Re = 1, λ = 0.5, β = 0.5, C_{th} = 0.7. Finally a well-populated Pareto front is obtained, from the external archive, which gives a solution set of layouts for optimization. The Pareto front obtained for a binary sensor model is shown in Figure 4. Two Pareto optimal layouts are shown in Figure 5 and 6 to illustrate the variety of layouts available. The effect of different sensor models on the Pareto front is shown in Figure 7. The improvement in Pareto front with increase in number of generations of MOPSO algorithm is also shown in Figure 8. The layout shown in Figure 5 is the layout with best coverage. For getting more coverage the particles spread out to minimize the overlapping region. Many sensors transmit their own data as well as act as communication relay for other far away sensors.

VI. Conclusion

In this paper, an energy efficient layout with good coverage for a WSN is considered. The application of MOPSO to maximize coverage and lifetime simultaneously is discussed. Thus, the aim of the proposed algorithm is to locate good non-dominated solutions under time pressure. In this paper we have also considered the deployment problem for mobile wireless sensor networks. It is important to indicate that PSO is an unconstrained search technique. Therefore, it is necessary to develop an additional mechanism to deal with constrained multiobjective optimization problems. For this purpose, we have proposed a multi objective approach for the deployment of nodes to improve upon an irregular initial deployment of nodes. Coverage and

lifetime are taken as the two conflicting objectives for achieving a set of layouts. Depending on the application, the user can choose a layout from the set of solutions. The performance of the MOPSO algorithm is determined by the computation time and uniformity of the solutions on the Pareto front. Simulation results show that the MOPSO algorithm obtains a better set of solutions as compared to single objective algorithms and other deployment algorithms.

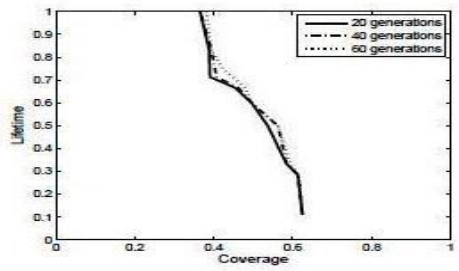


Figure 8: Pareto fronts obtained for a WSN with 10 sensors and 50 particles

VII. Future Work

More general data distribution, and also more sophisticated co-ope[illegible]f nearby nodes, are useful extensions and will be studied in future work. We studied the LMS implementation operating with Gaussian signals. Other strategies can be studied using the formulation presented here, such as the distributed normalized LMS (dNLMS), the distributed affine projection algorithms (dAPA) and distributed RLS implementations. If both the desired and the input data are corrupted by impulsive noise, then the Huber Prior Error Feedback-Leat Square Lattice (H-PEF-LSL) algorithm gives a good performance. In practice, a WSN is divided into multiple sub-regions for easy layout, organization and management. In future work, we will also take energy consumption due to sensor movement into account.

References

[1] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in Pro-ceedings of the 3rd International Symposium on Information processing in sensor networks, April 2004, pp. 20–27.
[2] Kennedy, J. and R. Eberhart, "Particle swarm optimization," *Proceedings of IEEE International Conference on Neural Network*, Piacataway, NJ, 1995.
[3] Engelbrecht, A. P., *Fundamentals of Computational Swarm Intelligence*, John Wiley & Sons, 2005.
[4] Coello, C. A. C., G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Trans. Evolutionary Computat.*, Vol. 8, 256–279, 2004.
[5] S. Haykin, Adaptive filter theory. Englewood Cliffs, NJ: Prentice-Hall, 2001.
[6] S. Dhillon, K. Chakrabarty, and S. Iyengar, "Sensor placement for grid coverage under imprecise detections," in Proceedings of Fifth International Conference on Information Fusion, vol. 2, Seol, Korea, 2002, pp. 1581–1587.
[7] E. Zitzler and L. Thiele. Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. IEEE Transactions on Evolutionary Computation, 3(4):257-271, 1999.

VII. Author Biographies

A. RAJESHWOR LAL DUA

Head Deptt. Of Electronics & Communication Engineering

Professor Rajeshwar Lal Dua a Fellow Life Member of IETE and also a Life member of I.V.S & I.P.A, former "Scientist F" is the most well known scientists in India in the field of Vaccum Electronic Devices for over three and half decades. His professional achievements span a wide area of vaccum microwave devices ranging from crossed-field and linear-beam devices to present-day gyrotrons. He was awarded a degree of M.Sc (Physics) and Sc Tech (Electronics) from BITS Pilani. He started his professional career in 1966 at Central Electronics Engineering Research Institute (CEERI), Pilani. During this period, indigenous know how was developed for several types of fixed frequency and tunable magnetrons of conventional and coaxial type. He also has several publications and a patent to his credit.

B. DEEPAK KUMAR CHAUDHARY

Pursuing M.TECH, Communication & Signal Processing System

Deepak Kumar Chaudhary was awarded a Bachelor of Electronics & Communication Engineering degree from the Pokhara University, Nepal, in 2008. During this period, his contribution was in Simulation of Digital Audio Broadcasting (DAB). Digital Audio Broadcasting (DAB) is a digital radio technology for broadcasting radio stations. He is currently working towards the M.TECH degree in the Department of Electronics & Communication Engineering at the Jaipur National University, Jaipur, India. His current research interests include: "Application of Multiobjective particle swarm optimization to maximize coverage and lifetime using wireless sensor network".

Power Management in At- Speed Scan Based Testing Applied to SOC

[1]M.P. Bhagya Lakshmi
Department of ECE, Sudharsun Engineering College,
Chennai Anna University, Pudukkotai.

Abstract:
The focal goal of this paper is to shrink power consumption for the duration of at speed scan based testing. In scan-based tests, power consumptions in both shift and capture phase may be drastically privileged than that in regular approach, which threaten circuits' trustworthiness through manufacturing test. High power consumed during shift and capture phases upshot in structural smash up to silicon or flawed data transmit during manufacturing test. Prior X filling techniques diminish either shift power or capture power however not together. Work of fiction Proposed X filling technique that can dwindle both shift- and capture-power for the period of speed scan based testing. Further still more reduce power consumption by adding dynamic voltage scaling method with X-filling method.

Keywords---At-Speed Scan-Based Testing, Low-Power Testing, Dynamic voltage scaling, X-Filling Technique

I. INTRODUCTION

With the advent of deep sub-micron technology and tight yield and reliability constraints, in order to perform a non-destructive test for high performance VLSI circuits power dissipation during test application should not exceed the power constraint set by the power dissipated during functional operation of the circuit. This is because excessive power dissipation during test application caused by high switching activity may lead to the following two problems namely
Destructive testing& Manufacturing yield loss

The power dissipation of integrated circuits (ICs) in scan-based testing can be significantly higher than that during normal operation [1]. This will threaten the reliability of the circuits under test (CUT), because: (i) the elevated average power dissipation adds to the thermal load that must be transported away from the CUT and can cause structural damage to the silicon, bonding wires, or the package; (ii) the excessive peak power dissipation is likely to cause a large voltage drop that may lead to erroneous data transfer in test mode only, especially in at-speed testing, thus invalidating the testing process and leading to yield loss.

In this effort, we focus on one of the most widely-used software-based solutions for test power reduction, which tries to reduce the CUT's switching activities by filling the don't-care bits (i.e., *X-bits*) in given test cubes intelligently, known as the *X-filling* technique.

In the this approach X-filling technique, first, we try to fill as few as possible X-bits to keep the capture-power under the peak power limit of the CUT to avoid test overkills, and then use the remaining X-bits to reduce shift-power as much as possible to cut down the CUT's average power consumption. Moreover, the X-filling technique is able to reduce power consumptions in both shift-in and shift-out processes, thus leading to significant shift-power reduction.

Dynamic voltage scaling is a power management technique in computer architecture, where the voltage used in a component is increased or decreased, depending upon circumstances. Dynamic voltage scaling to increase voltage is known as **overvolting**; dynamic voltage scaling to decrease voltage is known as **undervolting**. Undervolting is done in order to conserve power, particularly in laptops and other mobile devices, where energy comes from a battery and thus is limited. Overvolting is done in order to increase computer performance, or in rare cases, to increase reliability.

By Dynamic Voltage Scaling technique with X-filling, further we reduce power consumption in CUT's.

II. BACKGROUND

A. *Power Consumption in At-Speed Scan-Based Testing*

At-speed scan-based tests facilitate to detect speed-related defects of the CUTs and have been widely utilized in the industry in recent years, which typically involve a long low-frequency shift phase and a short at-speed capture phase. There are mainly two types of at-speed scan-based testing approaches: Launch-on-Shift (LoS) and Launch-on-Capture (LoC). LoC scheme is more widely utilized because it does not need the expensive high-speed scan-enable signal required by the LoS scheme. As shown in Fig. 1, there are three types of clock signals in LoC scheme: "SCLK" represents the shift clock signal, under which the test vectors are shift-in/out of the scan chains; "ACLK" is the at-speed clock in the CUT to be

applied in the capture phase; "TCLK" is the clock signal that the sequential elements on the scan chain will receive, by MUXing "SCLK" and "ACLK" signals. Typically two capture cycles (C1 and C2) are used to detect defects. We denote the initial state of the scan cells and the nodes in combinational portion of the circuit before capture as S1. The first capture C1 launches the state S2 into the CUT, while the second capture C2 store the circuit state S3 after the CUT is applied in functional mode.

The data input *DI* of each scan cell is connected to the output of the combinational logic as in the original circuit. To form a scan chain, the scan inputs *SI* of SFF_2 and SFF_3 are connected to the outputs Q of the previous scan cells, SFF_1 and SFF_2, respectively. In addition, the scan input *SI* of the first scan cell SFF_1 is connected to the primary input *SI*, and the output Q of the last scan cell SFF_3 is connected to the primary output *SO*. Hence, in shift mode, *SE* is set to 1, and the scan cells operate as a single scan chain, which allows us to shift in any combination of logic values into the scan cells. In capture mode, *SE* is set to 0, and the scan cells are used to capture the test response from the combinational logic when a clock is applied

Timing diagram to illustrate how the full-scan design is utilized to test the circuit shown in Figure 2.14a for stuck-at faults. During test, the test mode signal *TM* is set to 1, in order to turn on all test-related fixes.Two test vectors, V1 and V2, are applied to the circuit. In order to apply V1, *SE* is first set to 1 to operate the circuit in shift mode, and three clock pulses are applied to the clock *CK*. As a result, the PPI portion of V1, denoted by V_1:PPI, is now applied to the combinational logic. A **hold cycle** is introduced between the shift and capture operations. During the holdcycle, *SE* is switched to 0 such that the muxed-D scan cells are operated in capturemode, and the PI portion of *V*1, denoted by *V*1:PI, is applied. The purpose of thehold cycle is to apply the PI portion of *V*1 and to give enough time for the globally routed *SE* signal to settle from 1 to 0. At the end of the hold cycle, the complete test vector is now applied to the combinational logic, and the logic values at the primary outputs PO are compared with their expected values. Next, the capture operation is conducted by applying one clock pulse to the clock *CK* in order to capture the test response of the combinational logic to V1 into the scan cells. A second hold cycle is added in order to switch *SE* back to 1 and to observe the PPO value of the last scan cell at the *SO* output. Next, a new shift operation is conducted to shift out the test response captured in the scan cells serially through *SO*, while shifting in V2:PPI, which is the PPI portion of the next test pattern V_2.

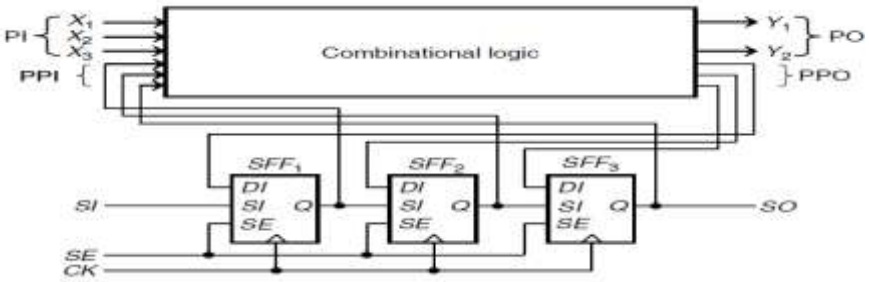


Fig-1: Example Circuit

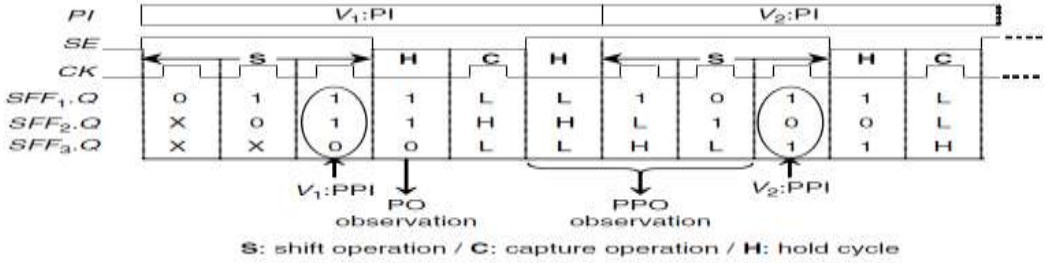


Fig-2: Timing Diagram for At-Speed Test operation

The shift process in scan-based testing not only dominates the test time of the CUT, but also determines the CUT's accumulated power dissipation. The main objective in shift-power reduction is thus to decrease it *as much as possible*, so that higher shift frequency and/or increase test parallelism can be applied to reduce the CUT's test time and hence cut down the test cost, under the average power constraint of the CUT.

In capture mode, on the other hand, since the duration is very short, it has limited impact on the CUT's accumulated test power consumption. However, because test vectors are usually generated to detect as many faults as possible, the excessive Transitions in capture phase may cause serious IR-drop and prolong circuit delay, thus causing false rejects (i.e., good chips fail the test) in at-speed tests. Consequently, the main objective in capture- power reduction is to keep it under a safe threshold to avoid test overkill. As long as this requirement is fulfilled, there is *no* need to further reduce capture-power.

B. Dynamic voltage scaling

MOSFET-based digital circuits operate using voltages at circuit nodes to represent logical state. The voltage at these nodes switches between a high voltage and a low voltage during normal operation—when the inputs to a logic gate transition, the transistors making up that gate may toggle the gate's output.

At each node in a circuit is a certain amount of capacitance. Capacitance can be thought of as a measure of how long it takes for a given current to effect a given voltage change. The node; since currents are related to voltage, the time it takes depends on the voltage applied. By applying a higher voltage to the devices in a circuit, the capacitances are charged and discharged more quickly, resulting in faster operation of the circuit and allowing for higher frequency operation. In this block diagram there are three main components. The first component is a performance sensor that monitors the main capacitance arises from various sources, mainly transistors (primarily gate capacitance and diffusion capacitance) and wires (coupling capacitance).

Here, we present a closed loop configuration based on fuzzy logic (FL) controllers to adaptively adjust the supply voltage of the processor compatible with workload variations. FL controllers are well suited for low-cost implementations based on cheap sensors and low-resolution analog-to-digital converters. Such systems can be easily upgraded by adding new rules to improve performance or add new features. Furthermore, in closed loop adaptive voltage scaling most of the power is dissipated in the time interval between changing the workload and acting to change the supply voltage. Consequently predicting the processor's workload compensates this drawback. Also, FL controller can be used to enhance power savings in comparison to other traditional controller systems, especially when the rate of workload variations is high.

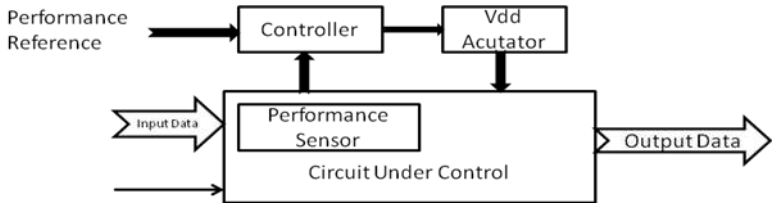


Fig 2: Block Diagram of Dynamic Voltage Scaling

C. Overview of Prior Work on Low-Power Testing

Adjacent fill is a simple yet effective technique targeting shift-power reduction. However, it can reduce the shift-in power only. We first addressed the low capture-power solution with X-filling. They considered the transitions at the output of SFFs during X-filling, which, however, does not necessarily have a good correlation with the total capture power of the whole circuit. Later, they took the above into consideration and introduced a new method to select the X-filling target based on a so-called set-simulation technique, which is proved to be a more effective X-filling method with experimental results on ISCAS'89 circuits. One of the main limitations of is that their computational time is quite high. This is because: (i). they are incremental filling approaches, that is, they fill the X-bits in the test cube one by one; (ii). Forward implications and backward justification are extensively used in their methodologies. In fact, the complexity of the set-simulation techniques proposed is quite high and it is difficult, if not impossible, to be applicable for two-pattern tests in industrial designs.

In developed an efficient probability-based X-filling technique, called *preferred fill*, which tries to fill all X-bits in the test cube in one step, instead of using incremental fill and logic simulation. Their technique, however, is inherently less effective as the available information for the probability calculation in their single-step filling is quite limited. Also, only transitions at the SFFs are considered while the transitions at logic gates are ignored in their work.

III. IMPACT ANALYSIS FOR X-FILLING

The impact of different X-bits on the CUT's shift- and capture-power (namely *S-impact* and *C-impact*, respectively).

A. Impact of X-Bits on Shift and Capture-Power

Test cube generally contains multiple X-bits, and as many-bits in the test response are likely to become determined values after filling one single X-bit in the test stimulus**,** their filling order significantly affects the CUT's test power dissipation. We therefore try to model the impact of an X-bit on a CUT's shift- and capture-power (namely *S-impact* and *C-impact*), and use them to guide the X-filling.

During the scan shift phase, the test stimuli are shifted in scan cells with previous test responses shifted out concurrently. To model the impact of an X-bit in the test stimuli on shift-power dissipation, we need to define a completely different cost factor because shift-power mainly concerns transitions between adjacent scan cells instead of switching activities in the entire CUT. Therefore, we first identify the scan cells in *S3* that are possibly affected by filling an X-bit in *S1* (denoted as S3i *affected S*), by tracing its fan-out logic network.

IV. PROPOSED X-FILLING TECHNIQUE

In this section, we detail our proposed *iFill* X-filling solution, including *C-filling* for capture-power reduction, *S-filling* for shift-power reduction, and the overall algorithm to achieve both shift- and capture-power reduction objectives. In addition, we show how to improve the runtime efficiency of the proposed solution.

B. C-Filling for Capture-Power Reduction

It is not necessary to reduce capture power as much as possible. Instead, we only need to control it to be smaller than a safe threshold and we wish to use as few as possible X-bits to achieve this objective, so that the remaining X-bits can be filled for shift-power reduction. Therefore, we need to fill X-bits with higher C-impact earlier.

The transition probability for a logic node in the capture cycle when filling X*i* is calculated as follows:

$$\mathbf{TP}i = \mathbf{P_{1i}}' \times \mathbf{P_{0i}} + \mathbf{P_{0i}} \times \mathbf{P_{1i}}'$$

Where P'1i(P'0i) is its probability to be "1" ("0") in S1, and P1i(P0i) is its probability to be "1" ("0") in S2.

The *Capture Transition Probability* (*CTP*) of the CUT caused by filling an X-bit X*i* in test stimuli can be calculated as sum of transition probabilities of X-bits in its fan-out part

$$CTPi = \sum_{j \in fan-out\ xi} TP\ j$$

Then we can decide the logic value to be filled for the target X-bit, which will cause less in fan-out portion of this scan cell.

C. S-Filling for Shift-Power Reduction

Prior work on X-filling for shift-power reduction considers shift-in power only. This is unfortunate, because filling these unspecified bits has impact on both shift-in and shift-out power consumption.

To model the shift transition probability caused by logic differences between X*i* and its adjacent scan cells in the test stimuli, we calculate the *shift-in transition probability* (*SITP*) caused by filling this X-bit as follows:

$$\text{S-impact}_i = \text{pi} + \sum_{j\text{es affected}} (l\ scj - P\ j$$

where P1s(P0s) represents the probability of X*i* to be 1 (0),*i* is the position this X-bit resides, which relates to transition number it may cause during the shift-in.

Where j ranges among all the X-bits affected by X*i*, and P1r (P0r) represents the probability of the response X-bit X*j* to be "1" ("0"). It should be noted that these X-bits in test responses can be in different scan chains.

Now we can determine the total *shift transition probability* (*STP*) when filling with "1" and "0", respectively. It can be simply calculated as the sum of its SITP and SOTP.

To reduce shift-power, we should fill X*i* with the logic value that will cause fewer transitions in both shift-in and shift-out phases. As shown in Fig. 4, if STPi(1)<STPi(0) , it means filling X*i* with logic "1" is likely to generate less circuit transitions during the scan shift phase than that of filling it with logic "0", so we should fill X*i* with logic "1". Otherwise, we should fill with logic "0".

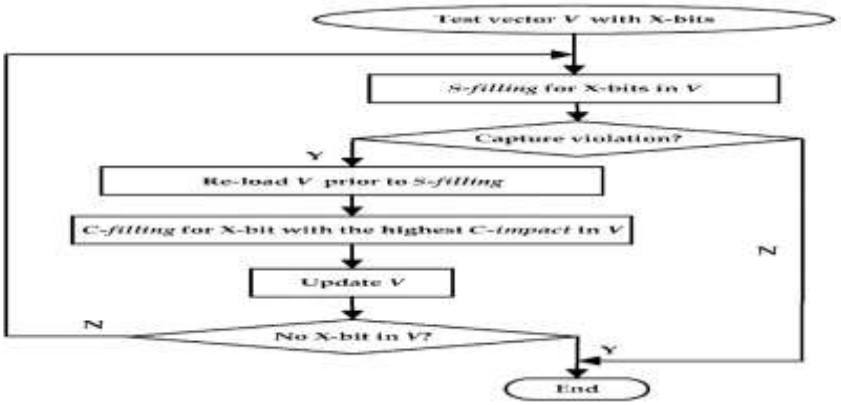


Fig-3 : S-Filling & C-Filling Flow for Power Reduction

D. Overall Flow

The objective of the proposed *X*-filling technique for simultaneous shift- and capture-power reduction is to keep the capture transitions under threshold and reduce shift transitions as much as possible. To meet this target, we proposed the overall flow as outlined in Fig. 4. First, we try to conduct *S-filling* to use all the X-bits in test vectors for shift-power reduction and check whether the capture-power violates the constraint after the *S-filling* process. If it does, we need to re-load the initial test cube, and fill one X-bit with the highest *C – impact* value for capture-power reduction. After filling every X-bit for capture- power reduction, the test vector will be updated, and we will apply *S-filling* procedure one more time to fill the remaining X-bits and then the capture-power will be checked again to see whether this test vector still has capture-power violation. When there is no power violation, we have completed filling the vector; Otherwise, *C-filling* procedure will be called again to reduce capture transitions. The above steps iterate themselves until there is no peak power violation or all X-bits have been utilized to reduce Capture-power. If the capture transitions still violates the limit after all X-bits have been filled, this test pattern need to be discard. After X-filling for all the test patterns in give test set, new test patterns need to be generated for the faults the test pattern violating the capture power limit covered.

E. DYNAMIC VOLTAGE SCALING

The *switching power* dissipated by a chip using static CMOS gates is $C \cdot V^2 \cdot f$, where C is the capacitance being switched per clock cycle, V is voltage, and f is the switching frequency,[1] so this part of the power consumption decreases quadratically with voltage. The formula is not exact however, as many modern chips are not implemented using 100% CMOS, but also uses pseudo nMOS gates, domino logic etc. Moreover, there is also a static leakage current, which has become more and more accentuated as feature sizes have become smaller (below 90 nanometres) and threshold levels lower.

Accordingly, dynamic voltage scaling is widely used as part of strategies to manage switching power consumption in battery powered devices such as cell phones and laptop computers. Low voltage modes are used in conjunction with lowered clock frequencies to minimize power consumption associated with components such as CPUs and DSPs; only when significant computational power is needed will the voltage and frequency be raised.

Some peripherals also support low voltage operational modes. When leakage current is a significant factor in terms of power consumption, chips are often designed so that portions of them can be powered completely off. This is not usually viewed as being dynamic voltage scaling, because it is not transparent to software.

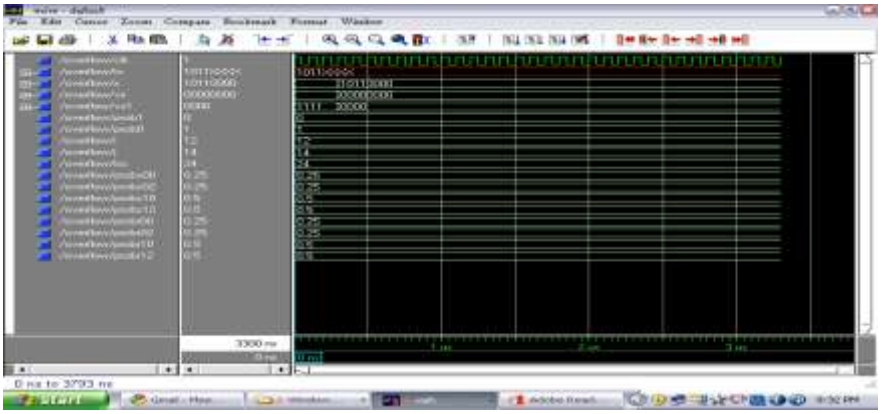


Fig-4:.Simulation Result for Overall

V. CONCLUSIONS

This paper presents an effective and efficient impact-oriented X-filling method, which is able to keep the CUT's capture-power within its peak power rating while reduce the CUT's shift-power as much as possible. Another contribution of the proposed technique is that it is able to cut down power consumptions in both shift-in and shift-out processes. In this work, all X-bits in given test cubes are used for test power reduction. Hardware implementation of Dynamic Voltage Scaling method with X-filling, we can reduce Power consumption of Entire CUT's.

REFERENCES

[1.] Dabholkar, S. Chakravarty, I. Pomeranz, and S. M. Reddy, "Techniques for minimizing power dissipation in scan and combinational circuits during test application," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 17, no. 12, pp. 1325–1333, Dec. 1998.

[2.] J. Li, Q. Xu, Y. Hu, and X. Li, "iFill: An impact-oriented X-filling method for shift- and capture-power reduction in at-speed scan-based testing," in *Proc. Des., Autom., Test Eur. (DATE)*, 2008, pp. 1184–1189.

[3.] J. Li, X. Liu, Y. Zhang, Y. Hu, X. Li, and Q. Xu, "On capture poweraware test data compression for scan-based testing," in *Proc. Int. Conf.Comput.-Aided Des. (ICCAD)*, 2008, pp. 67–72.

[4.] J.-L. Yang and Q. Xu, "State-sensitive X-filling scheme for scan capture power reduction," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 27, no. 7, pp. 1338–1343, Jul. 2008.

[5.] K. M. Butler, J. Saxena, A. Jain, T. Fryars, J. Lewis, and G. Hetherington, "Minimizing power consumption in scan testing: Pattern generation and DFT techniques," in Proc. IEEE Int. Test Conf. (ITC), Oct. 2004, pp. 355–364.

[6.] K. Miyase and S. Kajihara, "XID: Don't care identification of test patterns for combinational circuits," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 23, no. 2, pp. 321–326, Feb. 2004.

[7.] N. Z. Basturkmen, S. M. Reddy, and I. Pomeranz, "A low power pseudo-random BIST technique," in Proc. Int. Conf. Comput. Des. (ICCD), 2002, pp. 468–473.

[8.] P. Girard, "Survey of low-power testing of VLSI circuits," IEEE Des. Test Comput., vol. 19, no. 3, pp. 80–90, May-Jun. 2002.

[9.] P. Girard, X. Wen, and N. A. Touba, "Low-power testing," in System-on-Chip Test Architectures: Nanometer Design for Testability, L.-T. Wang, C.E. Stroud, and N.A. Touba, Eds. San Francisco, CA: Morgan Kaufmann, 2007, ch. 7.

[10.] P. M. Rosinger, B.M. Al-Hashimi, and N. Nicolici, "Scan architecture with mutually exclusive scan segment activation for shift- and capture- power reduction," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 23, no. 7, pp. 1142–1153, Oct. 2004.

[11.] Q. Xu, D. Hu, and D. Xiang, "Pattern-directed circuit virtual partitioning for test power reduction," in Proc. IEEE Int. Test Conf. (ITC), 2007, p. 25.2.

[12.] R. M. Chou, K. K. Saluja, and V. D. Agrawal, "Scheduling tests for VLSI systems under power constraints," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 5, no. 2, pp. 175–184, Jun. 1997.

[13.] R. Sankaralingam and N. A. Touba, "Controlling peak power during scan testing," in *Proc. IEEE VLSI Test Symp. (VTS)*, 2002, pp. 153–159.

[14.] R. Sankaralingam, B. Pouya, and N. A. Touba, "Reducing power dissipation during test using scan chain disable," in Proc. IEEE VLSI Test Symp. (VTS), 2001, pp. 319–324.

[15.] R. Sankaralingam, R. R. Oruganti, and N. A. Touba, "Static compaction techniques to control scan vector power dissipation," in *Proc. IEEE VLSI Test Symp. (VTS)*, 2000, pp. 35–40.

[16.] S. Remersaro, X. Lin, S. M. Reddy, I. Pomeranz, and J. Rajski, "Low shift and capture power scan tests," in *Proc. Int. Conf. VLSI Des.*, 2007, pp. 793–798.

[17.] S. Remersaro, X. Lin, Z. Zhang, S. Reddy, I. Pomeranz, and J. Rajski, "Preferred fill: A scalable method to reduce capture

[18.] S. Wang and S. K. Gupta, "ATPG for heat dissipation minimization during test application," *IEEE Trans. Comput.*, vol. 47, no. 2, pp. 256–262, Feb. 1998.

[19.] T. C. Huang and K. J. Lee, "An input control technique for power reduction in scan circuits during test application," in *Proc. IEEE Asian Test Symp. (ATS)*, 1999, pp. 315–320.

[20.] V. Iyengar and K. Chakrabarty, "Precedence-based, preemptive, and power-constrained test scheduling for system-on-a-chip," in *Proc. IEEE VLSI Test Symp. (VTS)*, Marina del Rey, CA, May 2001, pp. 368–374.

a. vation for shift- and capture- power reduction," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 23, no. 7, pp. 1142–1153, Oct. 2004.

[21.] W. Li, S. M. Reddy, and I. Pomeranz, "On test generation for transition faults with minimized peak power dissipation," in *Proc. ACM/IEEE Des. Autom. Conf. (DAC)*, 2004, pp. 504–509.

[22.] X. Wen, K. Miyase, S. Kajihara, T. Suzuki, Y. Yamato, P. Girard, Y. Ohsumi, and L. T. Wang, "A novel scheme to reduce power supply noise for high-quality at-speed scan testing," in Proc. IEEE Int. Test Conf. (ITC), 2007, p. 25.1.

[23.] X. Wen, Y. Yamashita, S. Kajihara, L.-T. Wang, K. K. Saluja, and K.Kinoshita, "On low-capture-power test generation for scan testing," in Proc. IEEE VLSI Test Symp. (VTS), May 2005, pp. 265–270.

Determination of Ultimate Lateral Loads in Deep Foundation in Multiple Layers of Cohessionless Soils

[1]B.S.Chawhan, [2]S.S.Quadri, [3]P.G.Rakaraddi,
[1]Asst. Prof, Govt. Engineering College, Haveri-581110,
[2] Prof.and Head, BVBCET, Hubli-580031,
[3] Associate Professor,BEC, Bagalkot,

Abstract:
This paper presents an experimental investigation on the lateral load carrying capacity of model piles of different flexural stiffness embedded in loose sand between dense sand and dense sand between loose sand layered soil strata. Attempts has been made to study the variation of lateral stiffness, eccentricity and soil layer thickness ratio and the effect of various parameters on the magnitude of N_h. The result of a model tested for the piles embedded in Tungabhadra river sand at Harihar taluk, Davangere dist, Karnataka State under monotonic lateral loadings. Experimental results are used for the load-deflection curves (p-y curves) for laterally loaded piles. The proposed p-y curves were compared to the existing curves with N_h and were evaluated with the experimental data. The ultimate lateral soil resistance and subgrade modulus were investigated and discussed.

Key words: Subgrade modulus, flexural stiffness, ground line deflection, model tests, piles, soil-pile interaction.

Introduction:

Pile foundations are the most popular form of deep foundations used for both onshore and offshore structures .They are often used to transfer large loads from the superstructures into deeper, competent soil layers particularly when the structures is to be located on shallow, weak soil layers. Piles are frequently subjected to lateral forces and moments; for example, in quay and harbor structures, where horizontal forces are caused by the impact of ships during berthing and wave action; in offshore structures subjected to wind and wave action; in pile-supported earth retaining structures; in lock structures, in transmission-tower foundations, where high wind forces may act; and in structures constructed in earthquake areas such as Japan or the West Coast of the United States.

The ultimate capacity of flexible piles and small pile groups in homogeneous and layered sand has been reported by Meyerhof and Ghosh 1989. But the state of art does not indicate a definite methodology by which the values of N_h can be obtained. Most of the investigators agree that N_h depends on soil and pile properties and value decreases with the increase in lateral load. Palmer et.al. (1954) indicated that width of pile has an effect on deflection, pressure and moment along pile even when flexural stiffness (EI) of pile is kept constant. Murthy (1992) has developed some relationship between N_h and other parameters like soil properties, flexural strength and lateral load. Dewaikar and Patil (2006) studied the analysis of laterally loaded pile in cohesionless soil and the Byung Tak Kim, Nak-Kyung Kim, Woo Jin Lee, and Young Su Kim studied the experimental Load Transfer Curves of Laterally Loaded Piles (April 2004).

This paper presents the experimental investigation of lateral load carrying capacity of model piles of various materials in homogeneous sand (both in loose and dense state), loose between dense and dense between loose sand layers with

horizontal loads acting at various eccentricities. In all the tests, the outer diameter (d) and depth of embedment (D) of piles are kept constant.

Experimental Set-Up and Model Tests

The test were carried in Tungabhadra river sand at Harihar taluk, Davangere dist, Karnataka State and having placement density of 13.35kN/m^3 and $\Phi=31^0$ for loose soil condition and 15.89kN/m^3, $\Phi=39^0$ for dense condition. The tests were conducted in two steps. a) The soil condition is loose sand layer between dense sand layer with H/D ratio of 0.25, 0.50, 0.75, and 0.90. b) The soil condition is dense sand layer between loose sand layer with H/D ratio of 0.25, 0.50, 0.75,and 0.90, where H is the thickness of middle layer and D is the total depth of embedment of the pile(=594mm). In both the cases the eccentricities of 0, 50, 100, 150 and 200mm are conducted in homogeneous loose sand layer and dense sand layer. The outside diameters of the piles are 25mm for solid Steel and Wooden Piles. Hollow Aluminium pile with 25mm outside and 0.3mm thickness. The embedment depths of all the piles are 594mm. The flexural rigidity of steel, wooden and Aluminium piles were 642.106Nm^2, 506.12Nm^2 and 51.041Nm^2 respectively. The relative stiffness factors for steel, wooden and Aluminium were 0.1192, 0.939 and 0.0094 respectively for Loose sand layer and 0.0263, 0.0207 and 0.0020 for Dense sand layer. The horizontal displacement and rotation of pile cap are recorded by L.V.D.T. and two dial gauges. The stabilized 'Rainfall-Technique' this standard technique is available in standard literature and this technique was used to pour the sand in the testing steel tank. Figure.1 shows schematic sketch of experimental setup.

Fig.1 Schematic sketch of experimental setup

The ultimate load bearing capacity of model piles are obtained from load deflection curves by the following criteria.

(A). Single tangent method
(B). Double tangent method
(C). Load corresponding to ground line deflection equal to 10% pile diameter
(D).Load corresponding to ground line deflection equal to 20% pile diameter
(E). Log-Log method.

It is observed that different criteria yield different ultimate load (vide Table-1). For the present analysis, the average of first three criteria is taken as ultimate pile capacity.

Method Of Analysis

Reese and Matlock (1956) have developed a set of equations based on dimensional analysis for computing deflection, slope, moment etc, along the pile. These equations are very useful for predicting the non-linear behavior of laterally loaded piles provided the magnitude of N_h is known at each load level. For deflection and slope of free head pile at ground level, the following equations are given by Reese and Matlock (1956).

$$Y_g = \frac{2.435PT^3}{EI} + \frac{1.62MT^2}{EI} \quad (1)$$

$$S_g = \frac{1.62PT^2}{EI} + \frac{1.75MT}{EI} \quad (2)$$

where,Relative Stiffness factor $T = \left(\frac{EI}{N_h}\right)^{\frac{1}{n+4}}$ (3)

P=Lateral load at pile head; M=Moment at pile head (=P*e); e=Eccentricity of horizontal load measured from ground level; and EI=Flexural stiffness of the model pile.

From the observed lateral resistance and corresponding ground line deflection and rotation, the value of coefficient of soil modulus variation N_h is estimated for different types of model piles by using the above equations (1) and (2).

Murthy .V.N.S (1976) has proposed the equations for determining N_h in cohesionless soil at each stage of loading as

$$N_h = \frac{A}{P_t^m} \quad (4)$$

where P_t= Lateral load at pile head, m is a constant equal to 0.8 and A is a factor which is a function of the effective unit weight γ of the soil and flexural stiffness EI of the pile.

$$N_h = \frac{156C_f\gamma^{1.5}(EIB)^{1/2}}{P_t} = \frac{A_s}{P} \quad (5)$$

where, P_t=Lateral load; A_s=Constant for pile in sand; P=Pt(1+0.67e/T); and C_f =Correction factor for the angle of friction = $3*10^{-5}(1.315)^{\Phi}$, where Φ is in degrees.

Results And Discussions

The experimental results were carried out and tabulated in following Table-1 and Table-2.

Table-1 presents ultimate loads of model Aluminium piles (embedded in Dense between loose sand layer) estimated from observed load deflection curves using different criteria mentioned earlier. It can be noted that ultimate load of a pile is not unique but depends on different methods or criteria. Table-2 presents ultimate loads of different pile materials in Loose between Dense layers estimated from observed experimental values. It can be noted that ultimate lateral resistance of pile decreases with the increase of H/D ratio in Loose between dense sand layer where as it increases with the increase of H/D ratio in Dense between Loose sand layer. Fig.2 shows typical load deflection curves of steel, wooden and aluminium piles embedded in loose sand between dense sand layer with e=50mm, H/D=0.25. In fig.3 the lateral load deflection curves of aluinium pile embedded in dense sand between loose sand layer having H/D=0.9 with varying eccentricity. The figure reveals that ultimate lateral resistance of pile decreases with increase in eccentricity. This phenomena is observed in all types of

model piles irrespective of all condition (i.e loose sand layer, dense sand layer, loose sand layer between dense sand layer and dense sand layer between loose sand layer).

In fig.4 the variation of coefficient of soil modulus v/s flexural stiffness(EI) curve of three model piles in dense between loose sand layer having h/D=0.90 with varying eccentricity. The figure reveals that flexural stiffness of pile increases with increase in variation of coefficient of soil modulus. This phenomena is observed in all conditions of soil layers.

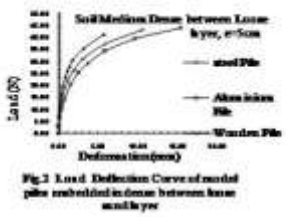
In fig.5 indicates the variation ultimate load of model piles with flexural stiffness EI when they are embedded in dense sand layer between loose sand layer having various H/D ratio and e=50mm. this reveals that ultimate load increases with increase in flexural stiffness of pile when all other conditions are kept constant.

Table-1.Comparison of ultimate load (N) by various methods (Aluminium pile, H/D=0.5, Dense between loose sand layer).

Eccentricity, e in mm	Methods (Different Criteria)				
e (mm)	A	B	C	D	E
50	20	14	16	32	33
100	19	19	15	26	33
150	12	09	12	24	26
200	10	08	08	22	24

Table-2. Ultimate load (N) of different pile (Steel, Wooden, and Aluminium pile, H/D=0.5, Dense between Loose sand layer).

Eccentricity, e in mm	H/D	Steel	Wooden	Aluminium
50	0.50	16.96	29.66	14.98
100	0.50	13.04	22.82	11.52
150	0.50	11.74	20.54	10.37
200	0.50	10.43	18.25	9.22



Dense sand layer between loose sand layer, H/D=0.9

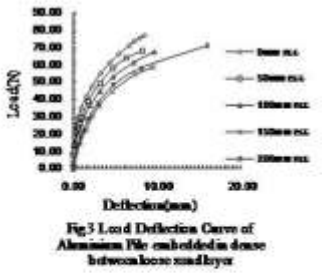


Fig 3 Load Deflection Curve of Aluminium Pile embedded in dense between loose sand layer

Dense sand layer between loose sand layer, H/D=0.9

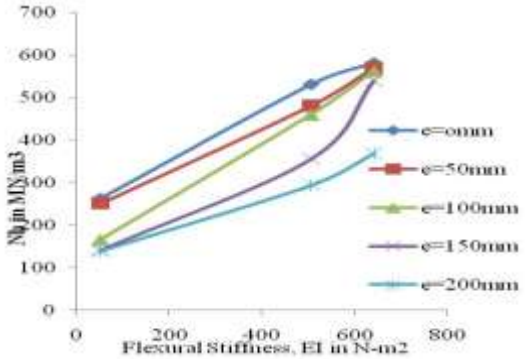


Fig.4. Variation of soil modulus N_h with Flexural stiffness of piles embedded in dense between loose sand layer, H/D=0.90, n=2/3.

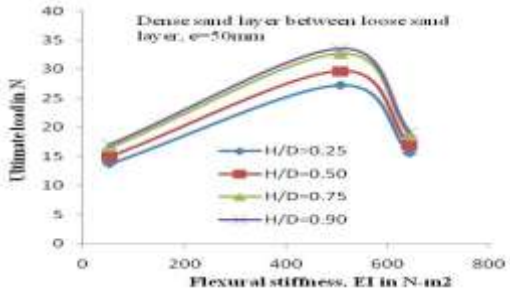


Fig.5. Variation of ultimate load with Flexural stiffness of piles embedded in dense between loose sand layer.

Conclusions

The following conclusions are made based on the experimental investigations.

(i) The ultimate lateral resistance of single pile decreases with the increase in eccentricity of load, it is about 8 to 11%.

(ii) The ultimate resistance of single pile subjected to horizontal load decreases with increase in eccentricity of load on the same pile provided the depth of embedment remains constant for homogeneous loose and dense layers, also loose between dense and dense between loose layered soils, and it is about 10 to 12%.

(iii) The ultimate lateral resistance of pile increases with increased value of flexural stiffness of pile and it is about 8 to 13% and the magnitude of N_h decreases with the increase in magnitude of horizontal load irrespective of flexural stiffness of pile and soil condition.

(iv) In dense sand, the magnitude of N_h increases with the increase in value of flexural stiffness of pile where as in case of loose sand the value decreases with the increase in EI value of piles and the ultimate lateral load carried is more (10 to 12%) in dense between loose sand layer and vice versa.

The tests may be conducted in multiple layers of Loose sand layer and Dense sand layer with constant and variable thickness of layers and also the Variation of Coefficient of soil modulus (N_h) in a different soil layers along the depth can be studied.

REFERENCES

[1.] Byung Tak Kim, nak-Kyung Kim, Woo Jin, Lee
[2.] and Young Su Kim. (2004), Experimental Load- Transfer Curves of Laterally Loaded Piles In Nak-Dong River Sand, Journal of Geotechnical and Geoenvironmental Engineering, 130(4),416-425.
[3.] Dewaikar D.M and Patil P.A. (2006), Analysis
[4.] of a Laterally Loaded Pile in Cohesionless Soil, IGC 2006,14-16 December 2006, Chennai, INDIA, 467-4.
[5.] Dewaikar D.M and Patil, D.S.(2001), Behaviour
[6.] of laterally loaded piles in cohesion-less soil under oneway cyclic loading, The New Millennium Conference, 14-16 December-2001.
[7.] Ghosh,D.P and Meyerhof,G.G.(1989), The
[8.] ultimate bearing capacity of flexible piles in layered sand under eccentric and inclined loads, Indian Geotech.J,(19)3, 187-201.
[9.] Murthy.V.N.S. (1992), Nonlinear Behaviour of
[10.] Piles Subjected to Static Lateral Loading.
[11.] P.Bandopadhyay and. B.C.Chattopadhyay. (1989), Ultimate Lateral Resistance of Vertical Piles, (2)4, 165-168.
[12.] Rees.L.C and Metlock.H. (1956), Non-
[13.] dimensional solutions proportional to depth, Prooceedings 8th Texas conference on Soil Mechanics and Foundation Engineering, Special publication no.29, Bureau of Engineering Research, University of Texas, Austin.
[14.] Terzaghi.K.(1955), Evaluation of coeeficient of subgrade reaction, Geotechnique, (5)4, 297-326.
[15.] V.Chandrashekaran, prof.S.Prakasan, and
[16.] Prof.S.Bhargava.(1975), Free Vibration Characteristics of Piles in Cohesionless Soils". Proceedings of the fifth Asian Regional Conference on Soil Mechanics and Foundation Engineering, December 19-22, 1975, Bangalore, India, pp-311-314.

Study of Transient Temperature Distribution in a Friction Welding Process and its effects on its Joints.

Sirajuddin Elyas Khany[1], K.N.Krishnan[2], Mohd Abdul Wahed[3]
[1]Associate Professor, [2]Professor, [3] Graduate Student,
MED, MJCET, Hyderabad, India.

Abstract:
In Materials like Stainless Steel 304 and Eutectoid Steel, microstructure is affected during the process of friction welding leading to variation in Hardness and strength. In the present Study, the transient temperature distribution in a friction welded Joint of similar materials in one dimension (uniaxial direction) is studied. A numerical method based on thermal networking is used to estimate the transient temperature distribution. Microstructure variation of specimen is also studied. The preliminary predictions are compared to actual experimental data from welding conducted under identical conditions. The results are shown to be in fair agreement. The numerical method proposed in this work provides guidance in weld parameter development and allows better understanding of the friction welding process.

Key Words: Friction Welding, Numerical Method, Temperature Distribution, Microstructure and Hardness.

Introduction:

In the present study, the friction welding technology used is one in which the rotational motion of the work pieces is stopped after the pressure has been applied for a very short period of time. During the process, the frictional heat is generated in the interface till plastic deformation stage is reached. Then, the rotation is stopped to let the welded joint cool down under free convection.

Peak joint temperature and the temperature profile in the region near the joint have a significant impact on the flash formation, the extend of heat effected zones and the joint strength. The thermal effects of the friction welding are observed to have lowered the hardness of the welded materials compared to the parent materials. Heating and cooling rates are closely related to the joint temperature and directly influence the residual stresses developed in the joint. Therefore, it is vital to have a means of rapidly and accurately estimating the peak joint temperature and cooling rates based on input parameters. A numerical method is developed, simulating the friction welding to quickly visualize the process specific responses to define the quality of the weld. An analysis of these phenomena is helpful in improving both the method and strength of the bonds in the welding process.

Mechanical evaluation and thermal modeling of friction welding of mild steel and aluminium is described by Hazman Seli et al (1).A new friction law of modeling of continuous drive friction welding and application to1045 steel welds is presented by Balsubramaniam et al (2). Thermically evaluation and modeling of friction welding is presented by Ahmet Can et al (3). Estimation of heat generation at the interface of cylindrical bars during friction process is presented by Wen-lih Chen et al (4). Experimental and numerical analysis of the friction welding process for the 4340 steel and mild steel combinations is described by Akbari Mousavi et al (5). Numerical simulation of linear friction welding of titanium alloy: Effects of processing parameters is presented by Wen-ya li et al (6). Thermo-mechanical and diffusion modeling in the process of ceramic-metal friction welding is done by Jolanta Zimmerman et al (7). Dynamic simulation of the temperature field of stainless steel laser welding is described by Han GuoMing et al (8). Temperature and stress distribution in ultrasonic metal welding—An FEA-based study is presented by S. Elangovan et al (9).

The focus of the present paper is to understand the thermal effects on the weld work piece and analyze the transient temperature profiles in the welded rods during friction welding. The numerical method is applied in obtaining the solution and the resulting temperature fields are compared with experimental data.

Materials and Method:

A continuous drive friction welding as shown in fig. 1 was used to produce a weld between stainless steel- stainless steel and eutectoid steel- eutectoid steel rods of diameter 15mm and length 80mm.

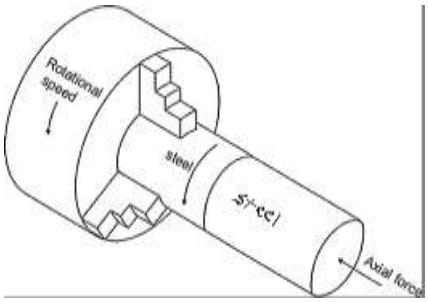


Fig. 1 Scheme of continuous drive friction welding

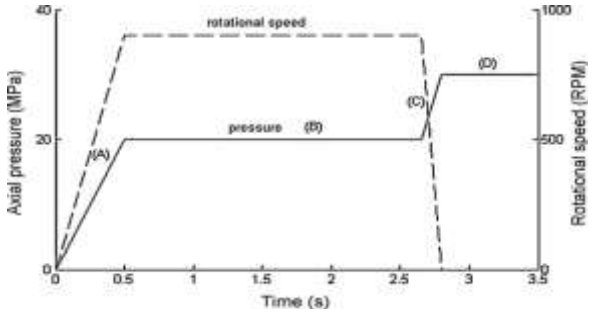


Fig. 2. Axial pressure and rotational velocity used in the experiment.

The changes of load and rotational speed in time during the process are shown in Fig. 2. The characteristic phases are marked. In phase (A), axial load and rotational speed of the steel rod were applied and the two rods were brought near the joined surfaces. Then in phase (B), the two rods underwent friction heating at the interface. When the rotational speed was stopped in phase (C), the frictional heated surface cooled down and at the same time, the upsetting under increased pressure occurred. Finally in phase (D), the pressure remained constant until the axial pressure is released.

The chemical composition of the Stainless steel and Eutectoid Steel employed in this study are summarized in Table 1. The temperature changes during the friction welding process are measured at a time by two k-type thermocouples attached to the stationary rod at a distance of 7mm, 14mm, 21mm and 28mm from the interface. The axial pressure is applied from the fixed end gradually. The temperature profiles are recorded at specific time intervals, till the welding is complete. The locations of the thermocouple is shown in fig.3.

Fig. 3. The locations of the thermocouple placement 7, 14, 21 and 24mm from the interface, respectively.

For metallurgical examination the welded sample is sectioned perpendicular to the weld interface. The joint sample is ground and polished. Microstructure is obtained using optical microscope. The Rockwell hardness is measured at the point of highest temperature from the interface (i.e. at 7mm from the interface).

Table 1 (Chemical composition of stainless steel 304 and eutectoid steel by wt %)

Element	C	Mn	P	Cu	Mo	S	Cr	N	Si	Fe	Ni
S.S(304)	0.08	20	0.045	0.75	0.75	0.05	20	0.10	1.00	Bal	10.5
Eutectoid steel	0.80	0.7	0.04	------	------	0.05	-----	------	------	-----	---------

3. Mathematical Modeling:

3.1. General assumptions for the analysis of friction welding:

- Heat generation and pressure across the interface are assumed to be uniform.
- There is no heat exchange between the end faces of the steel rods and the environment.
- The steel rods are assumed to be homogenous and isotropic.
- The heat input is considered constant throughout the process.

3.2. Thermal analysis:

During friction welding, the temperature in the weld region rises sharply due to extreme frictional heat and plastic deformation of the steel rod within a very short time. To calculate the temperature profile, the heat transfer analysis is undertaken by considering the frictional heat generation at the interface and heat loss to the surroundings by conduction and convection.

3.2.1. Frictional heat generation model:

Based on the assumption that force distribution remains constant, the frictional heat is deduced through the following analytical method.

First, a micro annulus with an inner radius r and a width dr in the friction surface is defined as in Fig. 4.
The constant pressure acting on the entire surface is given by p. The area of the micro annulus is $dA = (2\P r)\, dr$. Transforming the pressure equation into a differential characterization of the area dA, the following equation is obtained for the differential force, dF, acting on the area dA.

$$dF = p\, dA = 2p\P r\, dA \quad (1)$$

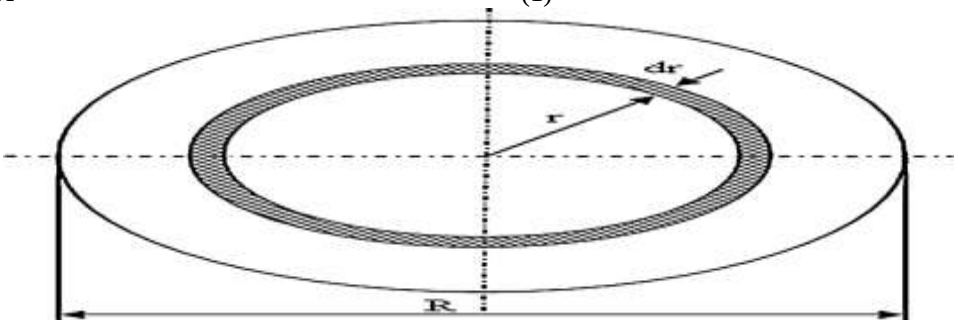


Fig. 4. Friction interface sketch.

Since dF is equivalent to the equal and opposite normal force acting on dA, the differential frictional force, dF_f, can be specified as

$$dF_f = \mu dF = 2\mu p\P r\, dr \quad (2)$$

Where μ is friction coefficient. It is known that the tangential velocity, v_T at any point on the element is the same,

$$v_T = r \quad (3)$$

Where ω is angular velocity.
The differential power that is exerted to rotate the annulus area dA is

$$dP = dF_f(v_T) = 2\mu p\P r^2\omega\, dr \quad (4)$$

Therefore, by integrating Eq. (4) with respect to r, the definition for the frictional heating power generation is obtained as

$$P = \int_0^R 2\mu p\P r^2\omega\, dr = (2/3)\, \mu p\P R^3\omega \quad (5)$$

The heat flux (q) generated by friction at the annulus is given as

$$q(r) = dP/dA = 1.6\mu p\omega r \quad (6)$$

3.3. Heat transfer:

The fundamental non-steady equation of Fourier's heat conduction in the coupled thermo-mechanical problem can be expressed as follows:

$$\partial T/\partial t + u\,(\partial T/\partial x) = 1/\rho C_p\,\partial/\partial x\,(k\partial T/\partial x) - hP/\rho C_p A\,(T - To) + q^{\cdot} \quad (7)$$

where T is temperature, T_o is the ambient temperature around the rod, u is the shortening velocity, ρ is the material density, C_p is the specific heat capacity, A is the cross-sectional area, P is the perimeter of the rod, k is the thermal conductivity, h is the convection coefficient, x is the distance from the weld interface and t is time.For simplicity, Eq. (7) can be rewritten as

$$\partial T/\partial t + u\,(\partial T/\partial x) = \alpha\,(\partial^2 T/\partial x^2) - \beta\,(T - To) + q^{\cdot} \quad (8)$$

Where $\alpha = k/\rho C_p$ and $\beta = hP/\rho C_p A$.

In this one-dimensional equation, the convection term on the right-hand side of the equation accounts for heat conduction and heat lost through convection along the lateral surfaces of the two components. It is assumed that there is no heat lost through radiation at the frictional interface. The problem of heat conduction in the whole process of friction welding determined by means of Eq. (8) was, thus, simplified to the calculation of temperature field, T = T(x,t). The calculation of the temperature of the friction welding process is carried out in two stages. The first stage is the heating part while the second is the cooling process. The initial and boundary conditions when solving Eq. (8) are expressed based on the two stages.

3.3.1. Heating stage:

The temperature distribution is calculated for heating separately for the rod of length(L) by assuming common average heat generated at the interface. For the heating stage, the initial and boundary conditions for the simplified equation (8) are derived as

$$T(x, t_h) = T_o, \text{ for } t_h = 0 \quad (9)$$

$$-k\,(\partial T/\partial x) = q,\ x = 0 \text{ for } t_h > 0 \quad \text{and} \quad (10)$$

$$-k\,(\partial T/\partial x) = h\,(T - T_o),\ x = L, \text{ for } t_h > 0 \quad (11)$$

Where t_h is the heating or frictional time, T_o is the initial temperature of the specimen, q is the surface heat flux generated at the friction surface (x = 0).

3.3.2. Cooling stage (welded):

At this stage the rods have been joined and considered as one new rod for the calculation, where the initial and boundary conditions for the simplified equation (8) are derived as

$$T\,(x_j, t_c) = T_n, \text{ for } t_c = 0 \quad \text{and} \quad (12)$$

$$K_s\,(\partial T/\partial x) = h_a\,(T_s - T_o),\ x_j = 0, \text{ for } t_c > 0 \quad (13)$$

Where x_j is the distance from left end of the joined rods, t_c is the cooling time, k is the thermal conductivity of the steel, and T is the temperatures of the free surfaces of the steel. T_n is the last temperature profile from the previous heating stage. The only unknown in the equation presented above is the shortening velocity (u) which is assumed to be zero. The calculation of the temperature is carried out using FORTRAN.

The properties of Stainless steel 304 and Eutectoid steel are listed in Table 2. As shown below.

Properties	Stainless steel 304	Eutectoid steel
Density(kg/m^3)	7900	7840
Thermal conductivity(w/mk)	15	46
Specific heat(j/kg k)	477	470
Melting Point	1371 - 1532 ^oC	

4. Results and discussion:

The experimental study is done on a continuous drive friction welding machine. The friction welding parameters were 490 and 790rpm rotational speed, 5-8 bar friction pressure and 11-15 bar upsetting pressure. Hardness profiles of parent and welded samples are determined. A one-dimensional numerical model for temperature profiles of this process is proposed and solved using FORTRAN program developed in the data processing center of Muffakham – Jah – College of Engg. & Tech. Hyderabad

4.1 Microstructure Variation:

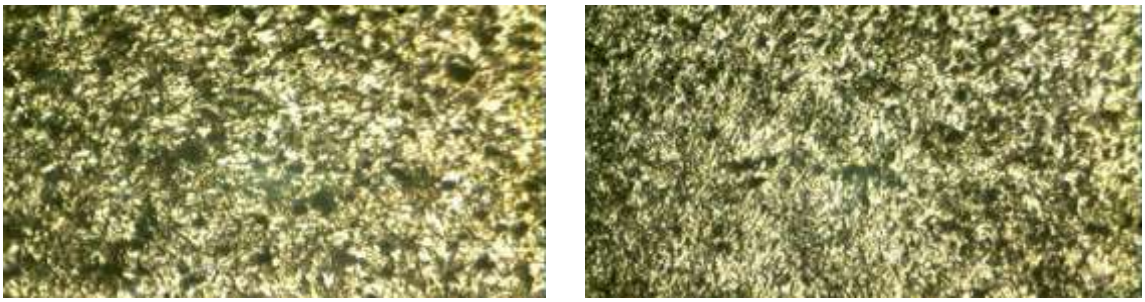


Fig.5 Microstructure variation of S.S 304 before and after friction welding at 675x.

The microstructure consists of single phase austenite. The microstructure is completely austenite of equiaxed uniform grains within the grains, slip bands are seen and the grain boundaries are clear. Precipitate of alloy carbides is noticed. Due to chemical composition deposition alloy carbides are seen which are not refined before friction welding.

Subjected to friction welding at 613^{o}C (i.e. at a distance of 7mm, the highest peak temperature from the interface) it is observed that there is grain refinement along with the precipitation of carbides.

Before friction welding	After friction welding
Grain size ASTM: 6	Grain size ASTM: 7

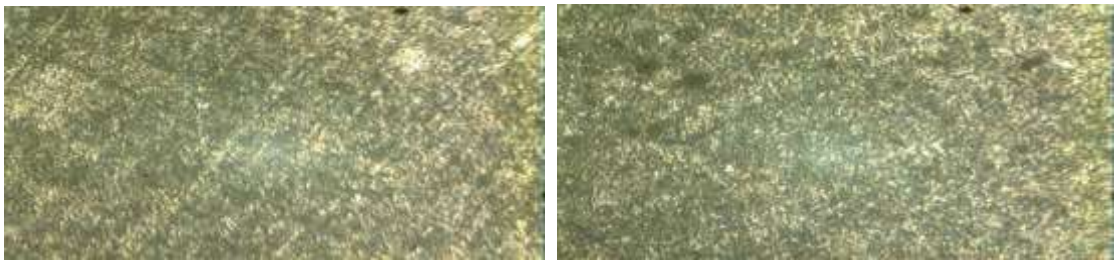


Fig.6 Microstructure variation of Eutectoid steel before and after friction welding at 675x.

Microstructure reveals 100% pearlite in the form of lamella structure (alternate bands of cementite and ferrite) before friction welding. Dark areas are pearlite and light areas are ferrite.

Microstructure reveals 100% pearlite which is more refined after friction welding and can be seen in the form of lamella structure (alternate bands of cementite and ferrite).

In this study the highest temperature reached is 600^{o}C (i.e. at a distance of 7mm from the weld interface) in both the steel rods. More changes in the microstructure are not expected because the temperature is within the range of eutectoid temperature (i.e. below 723^{o}C).

4.2 Hardness Profile:

Fig. 7a and 7b shows the result of the Rockwell hardness test for S.S 304 and Eutectoid steel. The hardness value is checked for both the steel rods at distance of 7, 14, 21 and 28mm of welded part and is found to be lower at the first node (i.e. at 7mm) than that of its parent part. This reduction could be due to the thermal effect in the welded steel. Due to normalizing the hardness is increasing as the distance from the weld is increasing. In stainless steel 304, hardness is increasing gradually and in eutectoid steel rapidly this, difference in both the steel rods is due to the refinement of grain structure and homogenization of the microstructure due to normalizing.

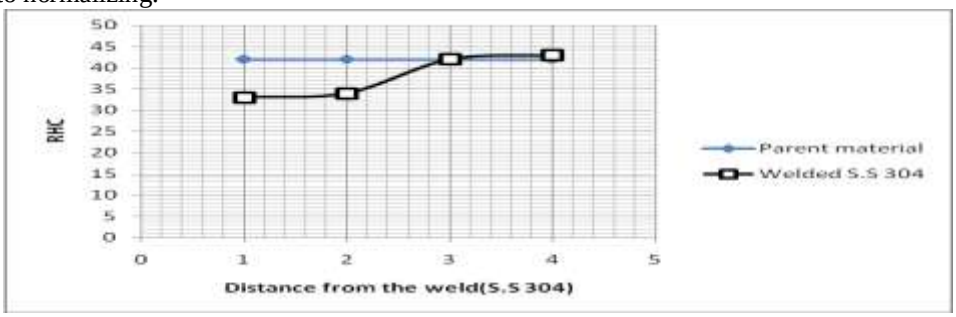


Fig.7a Results of the Rockwell hardness test for Stainless steel 304

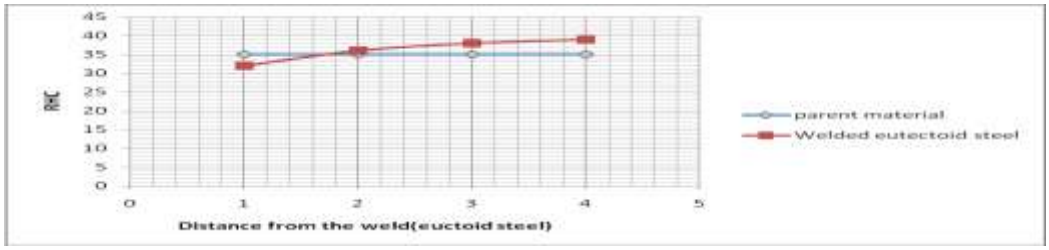


Fig.7b Results of the Rockwell hardness test for Eutectoid steel.

Due to normalizing, hardness is increasing in both the steel rods, and as the hardness is increases, strength also increases.

4.3 Heating and cooling stage:

The transient heating and cooling temperature distribution for steel rods are shown in Fig. (9a – 9p).From the temperature distribution, it is seen that the calculated peak temperature in heating for stainless steel at 90sec and 40sec is about 613^{o}C and 378^{o}C and for eutectoid steel is 578^{o}C and 270^{o}C at 200sec and 70sec, peak temperatures in heating for steel rods are obtained at a distance of 7mm from the interface at 490 and 790rpm with constant heat generated. Cooling starts of as the entire process, where the last heating temperature profile has been utilized. Similarly transient heating and cooling temperatures are obtained for similar parameters at 14, 21 and 28mm from the interface. The temperature increases rapidly at the interface, and gradually towards the end of the steel rods. The transient heating and cooling temperature distribution in eutectoid steel is faster. This is due to the higher thermal conductivity in eutectoid steel compared to that in stainless steel. Due to the lower thermal conductivity stainless steel is taking less time for welding when compared to eutectoid steel. As the speed (rpm) is increasing temperature as well as time is decreasing in both the steel rods, but the variation is nearly similar. The difference can be attributed to accomodating the heat losses at the interface in the analytical model.

The thermal heating profile likely exhibits the interaction between the frictional heating power and the frictional characteristics on the surface. However, in a real situation, the pressure distribution is not uniform with time as the two work pieces move in sinusoidal fashion. While the axial force remains constant, the area of contact between the two work pieces changes with movement and leads to the oscillation of the axial pressure at the interface. Therefore, altering the axial pressure during every cycle owing to the variation of contact area causes the frictional heat input to fluctuate as well. The friction coefficient varies widely with temperature. The increase in the temperature softens the ductile material and brings about deep ploughing on the contact surfaces. Then, the contact seems to have been altered into a polishing-like action and the friction coefficient is dramatically reduced. The inaccuracy in the calculation was attributed to the assumption of constant coefficient of friction and pressure for the analytical constant heat generation.

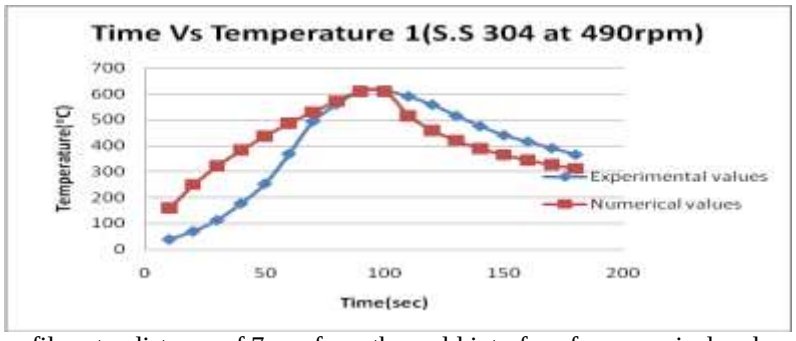


Fig.9a Time–temperature profiles at a distance of 7mm from the weld interface for numerical and experimental data of stainless steel at 490rpm.

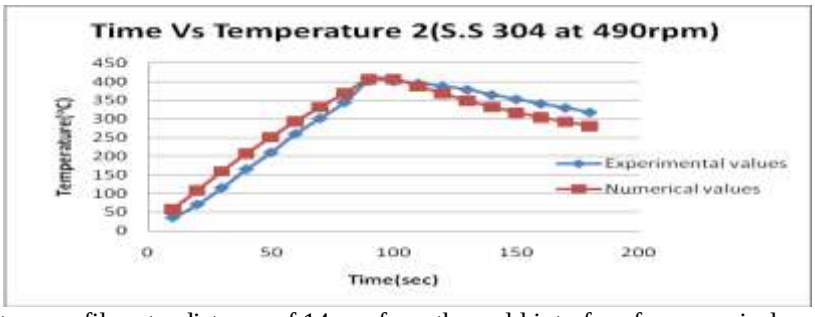


Fig.9b Time–temperature profiles at a distance of 14mm from the weld interface for numerical and experimental data of stainless steel at 490rpm

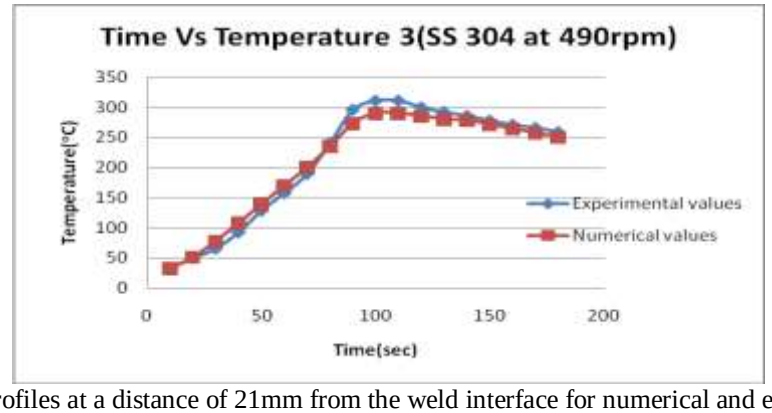


Fig.9c Time–temperature profiles at a distance of 21mm from the weld interface for numerical and experimental data of stainless steel at 490rpm

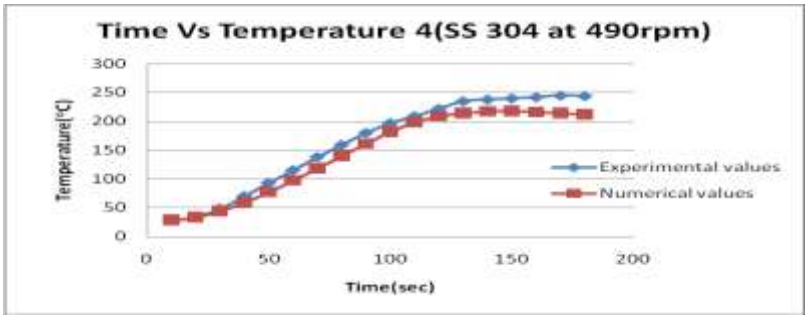


Fig.9d Time–temperature profiles at a distance of 28mm from the weld interface for numerical and experimental data of stainless steel at 490rpm

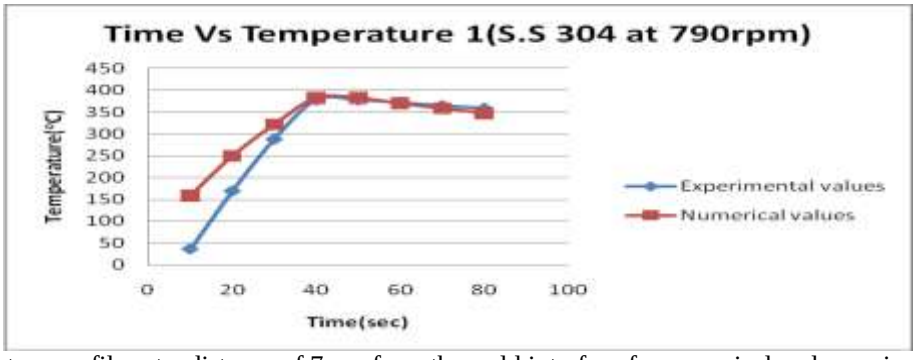


Fig.9e Time–temperature profiles at a distance of 7mm from the weld interface for numerical and experimental data of stainless steel at 790rpm

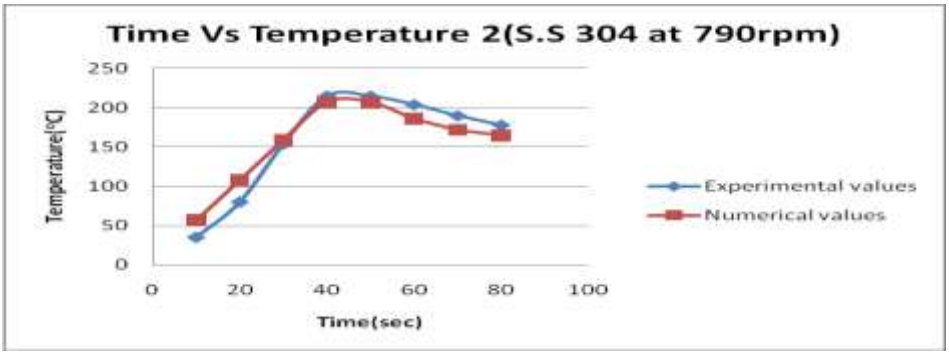


Fig.9f Time–temperature profiles at a distance of 14mm from the weld interface for numerical and experimental data of stainless steel at 790rpm

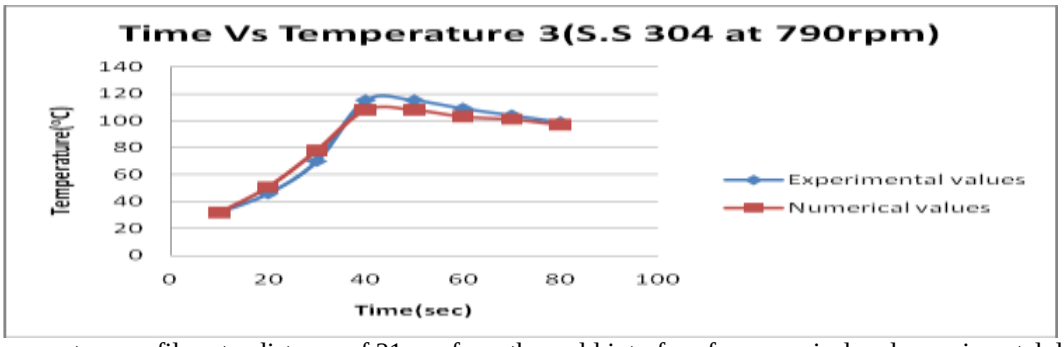


Fig.9g Time–temperature profiles at a distance of 21mm from the weld interface for numerical and experimental data of stainless steel at 790rpm

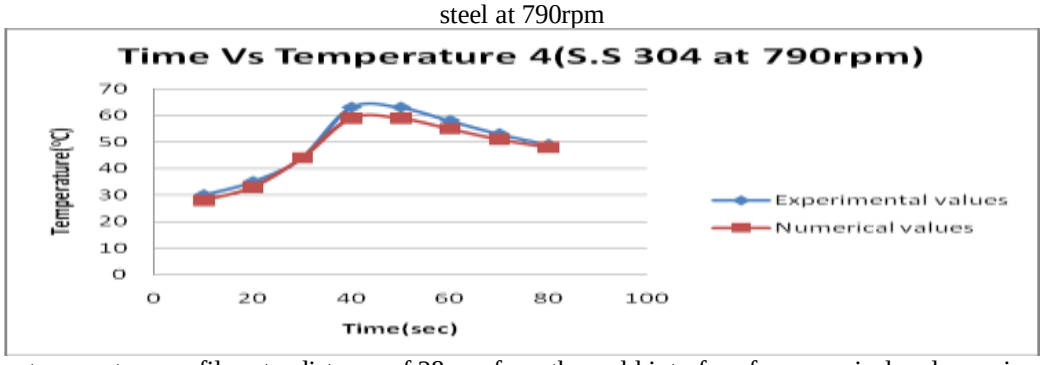


Fig.9h Time–temperature profiles at a distance of 28mm from the weld interface for numerical and experimental data of stainless steel at 790rpm

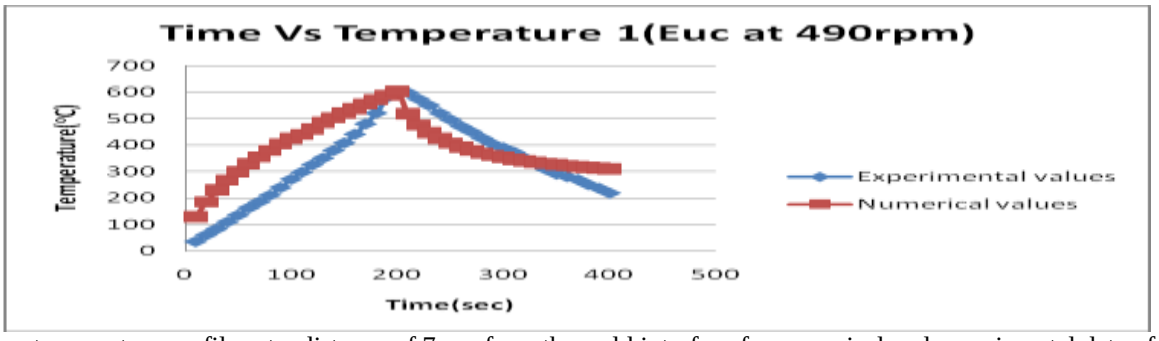


Fig.9i Time–temperature profiles at a distance of 7mm from the weld interface for numerical and experimental data of eutectoid steel at 490rpm

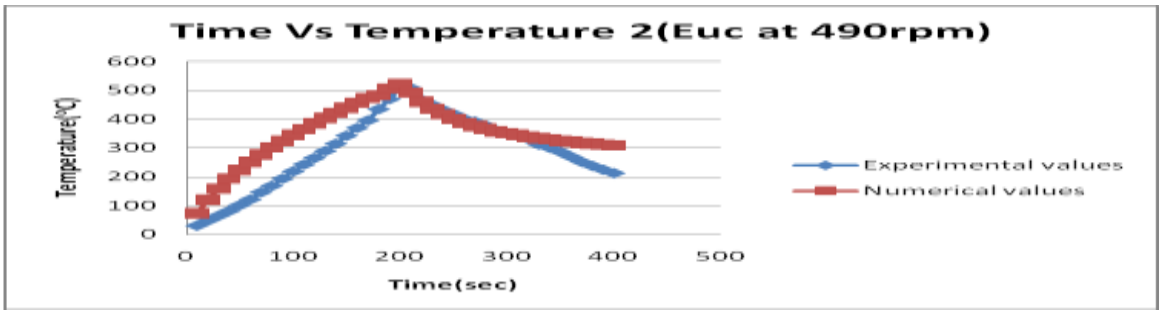


Fig.9j Time–temperature profiles at a distance of 14mm from the weld interface for numerical and experimental data of eutectoid steel at 490rpm

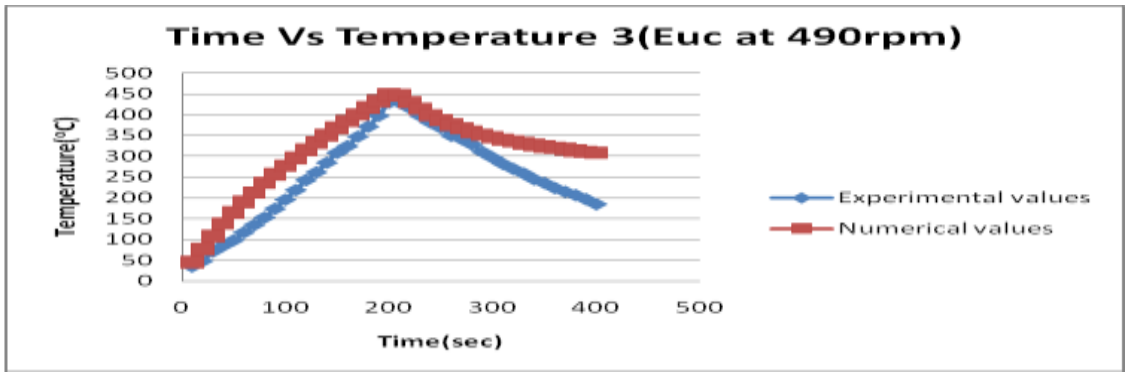


Fig.9k Time–temperature profiles at a distance of 21mm from the weld interface for numerical and experimental data of eutectoid steel at 490rpm

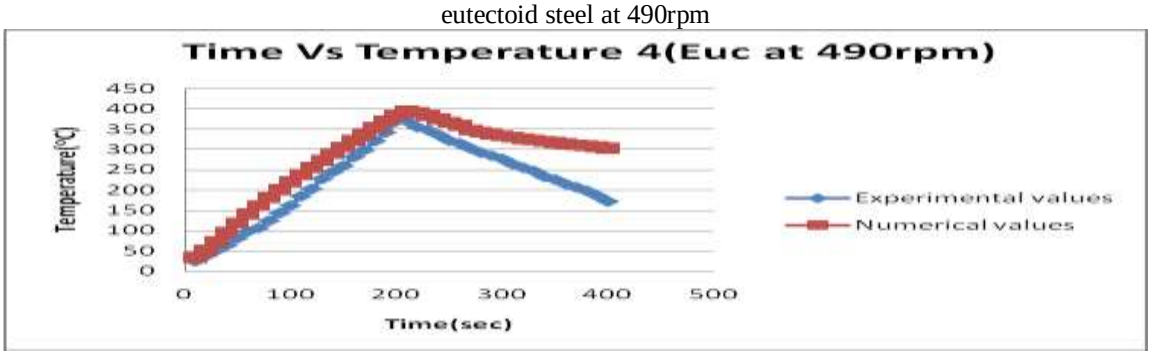


Fig.9l Time–temperature profiles at a distance of 28mm from the weld interface for numerical and experimental data of eutectoid steel at 490rpm

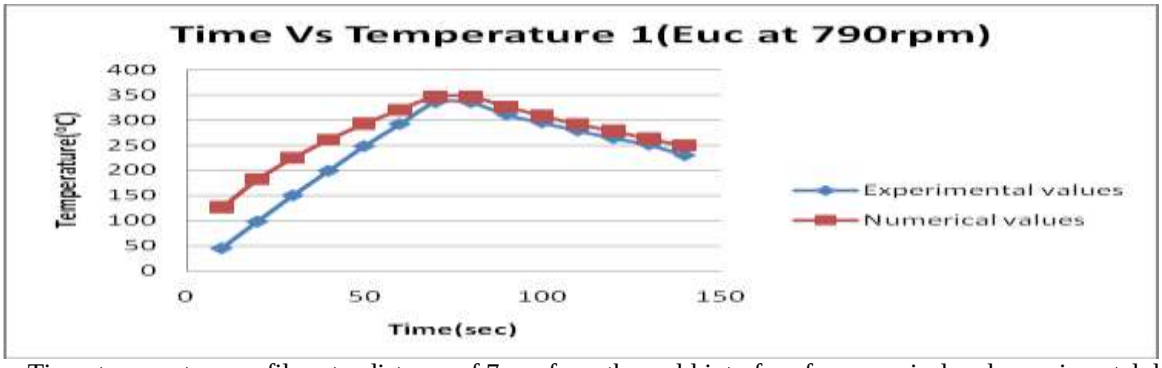


Fig.9m Time–temperature profiles at a distance of 7mm from the weld interface for numerical and experimental data of eutectoid steel at 790rpm

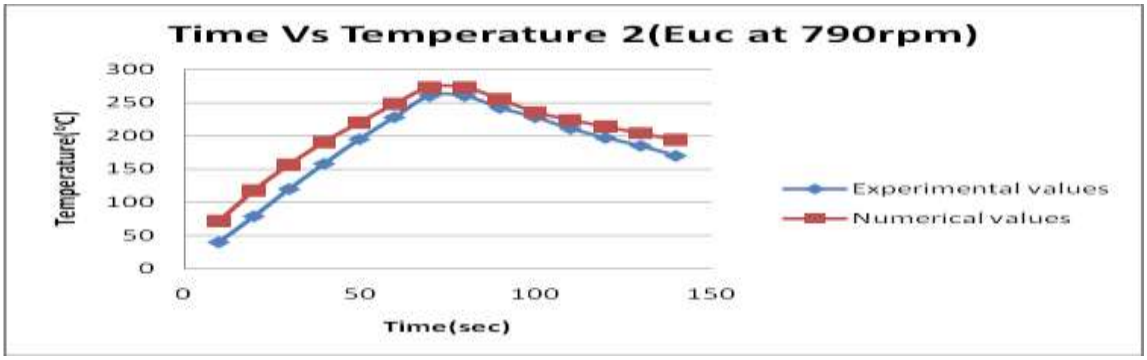


Fig.9n Time–temperature profiles at a distance of 14mm from the weld interface for numerical and experimental data of eutectoid steel at 790rpm

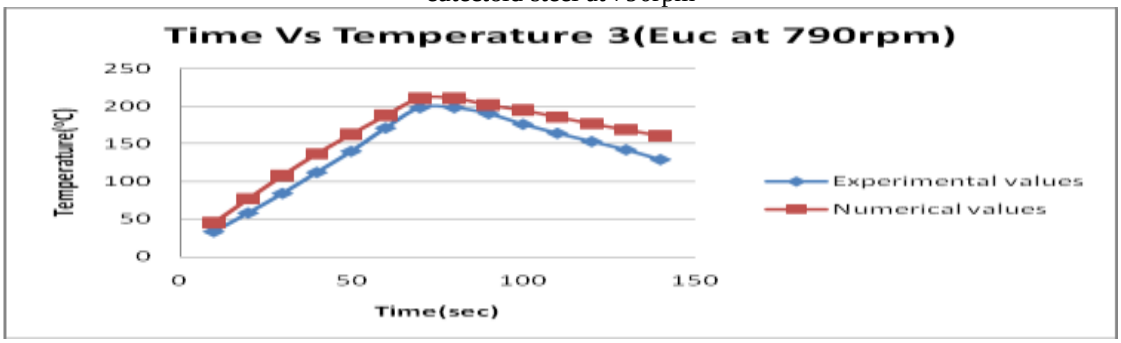


Fig.9o Time–temperature profiles at a distance of 21mm from the weld interface for numerical and experimental data of eutectoid steel at 790rpm

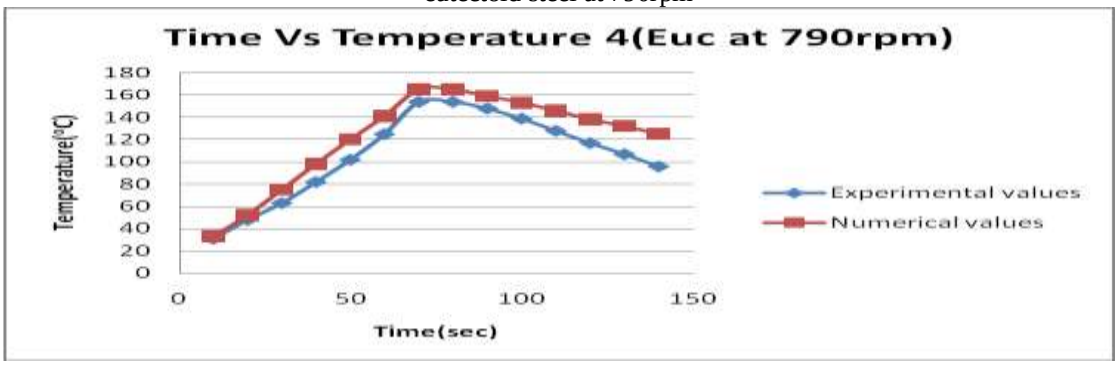


Fig.9p Time–temperature profiles at a distance of 28mm from the weld interface for numerical and experimental data of eutectoid steel at 790rpm

4.4. Verification of the predicted profile with experimental data:

The verification of the predicted temperature distribution of the welded rods are presented in Fig.9. for locations 7, 14, 21 and 28mm from the interface. The computed temperature profile does not exactly match with the experimental data. However, the trend of the profiles shown is similar.

The predicted heating and cooling temperature profiles of the friction welding are in fair agreement with the experimental temperature profiles. The heating and cooling rates of the process are approximately calculated around 138 ◦C/s and 45 ◦C/s, respectively. After the rotation of the steel rod is stopped, the temperature profile increases and decreases with lower gradients compared to the measured temperature profile. The discrepancy of the two temperature profiles is mostly attributed to the static nature of the model. After 10 s, the difference is quite significant, most probably because the proposed model does not consider the entire work pieces for calculation. In the friction welding process, at every moment of the friction phase, the plasticized material is expelled out of the faying surface due to the mutual movement of the mating surfaces. In contrast, the numerical model is simply based on the static analysis in which the heat generated at the interface is assumed to be totally transferred to the base materials with no flash formed (zero axial shortening assumption). Therefore, the relative movement (rubbing) of the two parts, which is the main reason for ejecting softened material from the rubbing surface, is not considered in the model. Apart from that, the errors could also come from the temperature measurement since operation is manual. Even though inaccurate, these details are still useful in friction welding process study, especially in explaining the heat affected zone which eventually affects the strength of the joint. Hence, the proposed model provides some parameter development and allows better understanding of the friction welding process.

5. Conclusions:

Bonds of stainless steel vs stainless steel and eutectoid steel vs eutectoid steel are achieved in the friction welding process. The welded materials have lower hardness measured at a distance of 7mm from the interface compared to their parent materials due to thermal effects of the friction welding. But as the distance from the interface is increasing hardness is increasing due to normalizing and this has been also compared with CCT diagram. A one-dimensional numerical model for continuous drive friction welding was developed according to the characteristics of the friction welding process. It is introduced to enable a better understanding of the process. The predicted transient heating and cooling temperature profile of the friction welding is in fair agreement with the experimental temperature distribution.

Acknowledgements:

The authors gratefully acknowledge the co-operation and help extended by Mr. Mansur from metal cutting laboratory and Mr. Hashmi from metallurgy laboratory in carrying out the experiment during this research work.

References:

[1.] Hazman Seli, Ahmad Izani Md.Ismail, Endri Rachman, Zainal Arifin Ahmad.(2010) Mechanical evaluation & thermal modeling of friction welding of mild steel & aluminium. Journal of materials processing technology 210 (1209-1216).

[2.] Vikram Balasubramaniana; Youlin Lib; Tim Stotlerc; Jeff Cromptonc; Alfred Soboyejoa; Noriko Katsubea;Wolé Soboyejo.(1999). A New Friction Law for the Modelling of Continuous Drive Friction Welding: Applications to 1045 Steel Welds, materials & manufacturing process, 14:6, (845-860).

[3.] ahmet CAN, Mumin SAHIN & Mahmut KUCUK(2009).Thermically Evaluation & modeling of friction welding,strojarstvo 51(1) (5-13).

[4.] Wen-lih chen, Yu-ching yang, Shao-shu chu(2009).Estimation of heat generation at the interface of cylindrical bars during friction process), Applied thermal Engineering 29 (351-357).

[5.] s.a.a.akbari mousavi & a. rahbar kelishami(2008), Experimental & numerical analysis of the friction welding process for the 4340 steel & mild steel combinations. Welding research 178-186s(vol.87).

[6.] Wen-ya li, Tiejun Ma, Jinglong Li(2010), Numerical simulation of linear friction welding of titanium alloy: Effects of processing parameters, Materials & Design 31(1497-1507).

[7.] Jolanta Zimmerman, wladyslaw wlosinski, zdzislaw R Lindemann Thermo-mechanical & diffusion modeling in the process of ceramic-metal friction welding(2009), Journal of materials processing Technology 209, (1644-1653).

[8.] Han GuoMing *, Zhao Jian, Li JianQang(2007), Dynamic simulation of the temperature field of stainless steel laser welding, Material & Design 28 (240-245).

[9.] *S. Elangovan, S. Semeer, K. Prakasan*Temperature and stress distribution in ultrasonic metal welding—An FEA-based study (2009),Journal of material processing Technology 209,(1143–1150).

E-Governance web services for web seniors

[1]Prof. K.KailasaRao, [2]Sravanthi

Abstract
Now a day's service of any sector can be online this has tremendous in banking sectors, gradually the services of government now become online in every department and provide optimized results .

Web services are added to every application to access data across the web. This has been largely the Result of the many incremental efforts to describe, advertise, discover, and invoke web services. These e-government services gradually covers almost several application domains, including social programs, Healthcare, voting, tax filing, etc.

The primary objective of web services is to provide the interoperability among different software and data applications running on a variety of platforms .
In this paper we mainly focuses on web services provided by government departmental applications especially for senior citizens called web senior.

Key Word: e-government, e-commerce, virtual organization, digital administration, administrative interface, e-governance

Introduction

service-oriented digital government applications focused on efficiently providing customized services to senior citizens. We designed and developed a Web Service Management System (WSMS), called Web Senior, which provides a service-centric framework to deliver government services to senior citizens. The proposed WSMS manages the entire life cycle of third-party web services. These act as proxies for real government services. Due to the specific requirements of our digital government application, we focus on the following key components of Web Senior: service composition, service optimization, and service privacy preservation. These components form the nucleus that achieves seamless cooperation among government agencies to provide prompt and customized services to senior citizens.
Governments worldwide are faced with the challenge of transformation and the need to reinvent government systems in order to deliver efficient and cost effective services, information and knowledge through information and communication technologies.

Development of Information and communication technologies catalyzed and led up to E-government.What is E-government? In this paper,E-government is defined as a way for governments to use the most innovative information and communication technologies, particularly web-based Internet applications, to provide citizens and businesses with more convenient access to government information and services, to improve the quality of the services and to provide greater opportunities to participate in democratic institutions and processes.

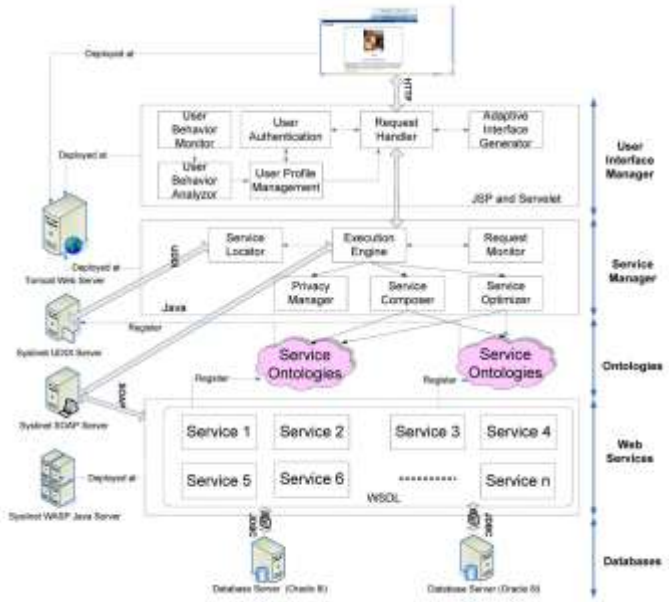
E-government presents a tremendous impetus to move forward in the 21st century with higher quality, cost-effective, government services and a better relationship between citizens and government.One of the most important aspects of e-government is how it brings citizens and businesses closer to their governments. This paper outlines eight different potential types or models in an e-government system that is useful to define scope of E-government studies: Government-to-Citizen (G2C);Citizen-to-Government (C2G); Governmentto-Business (G2B); Business-to-Government(B2G); Government-to-Government (G2G);Government-to-Nonprofit (G2N); Nonprofitto-Government (N2G); and Government-to-Employee (G2E). This paper also examines some examples in E-government practicesand presents a generally-applicable framework for analysis of challenges and problems in E-government development.Emerging with E-government, theories and practices of public administration have stepped into a new digital era. This paper proposed that contemporary issues related toE-government in public administration are administrative interface, i.e, peoplecomputer interface in management_ digital administration, i.e., digital process or procedures and system in management, and virtual organization, i.e., government online system, etc.

Web services provide an efficient vehicle for users to access the functionalities available on the web [30]. The development of web services has so far mostly been the result of standardization bodies usually operating on a consensus basis and driven by market considerations. In this context, innovation and long-term deployment issues are not usually of primary concern. Because of the global nature of the web, the standardization process has so far been very fragmented, leading to competing and potentially incompatible web service standards. Governments and commercial organization have meanwhile invested very heavily in web services technologies. These investments have resulted in a fast-growing number of web services being made available. The prevalent business model will, in
all likelihood, include a autonomous and competing communities of web service providers vying for consumers attentions. It is, however, important that these investments produce the expected results and soon for the area to really make an impact. One-key impediment has been the lack of any rigorous and systematic methodology for managing the entire life cycle of web services that would include delivering, selecting, optimizing, and composing services.

This needs to take place within a secure, trustworthy, and privacy protecting environment. We call the resulting system a Web Service Management System (WSMS). In summary, a WSMS is a comprehensive framework that provides an integrated view on the management of web services , including automatic service composition, service query optimization, service privacy preservation, service trust, and change management. Service composition is concerned with the automatic selection and integration of individual web services to provide value-added and personalized composite services . Service query and optimization are concerned with the ability to model queries and provide anoptimization framework suited for web services.

Service security and privacy preservation are concerned with ensuring that interactions with web services are conducted in a secure fashion, while sensitive informationcan be preserved as required.

Overview



Service centric architecture

In this paper, we describe a WSMS called WebSenior which is a service-oriented digital government system that

aims at providing services to senior citizens. Adopting webservices in DG enables government agencies to: outsource from other DG services, compose existing services to provide value-added services, effectively handle privacy

issues, and provide adaptive web-based user interface WebSenior is a collaborative effort between Virginia Tech

and the Virginia Department for the Aging (VDA). The VDA subcontracts with Area Agencies on Aging (AAAs) to offer services to communities throughout the state. AAAs work with public and private organizations to help seniors and their families find the services and information they need. One of the biggest barriers to providing services to customers is the lack of integrated information systems.The VDA and its partners offer many and diverse services.

The challenge is how to get these varied systems to cooperate and share information. WebSenior wraps the legacy systems by using web service technologies. It provides an integrated service framework that achieves a seamless cooperation among government agencies to provide prompt and customized services to senior citizens.

We summarize the key components of WebSenior, which constitute the major contributions of this paper as follows:

.***Service composition:*** This component performs three tasks: checking composability, checking soundness, and automatically composing services. The first issue when defining a composite service is whether its component services are composable. We present a set of composability rules to ensure the feasibility of the composed services. Another important issue is whether a composite service provides an added value.We present the concept of composition templates,based on which the soundness of a composite service can be evaluated. Finally, we propose a matchmaking algorithm for automatic composition of webservices.. Service Optimization: This component performs a"user-centric" optimization that selects the composition plan with the best quality based on users'preferences. First, we define a score function to evaluate a composition plan. In the score function, a weighting mechanism is adopted to express users' preferences over different quality parameters. A set of aggregation functions is presented to compute the quality of a composition plan that consists of multiple operations. Second, we present two optimization approaches to find the best plan: exhaustive search and greedy search.

Existing work

Present System is manually providing services to citizens and Senior citizens. They have to go service center to know some particular information. As part of an effort to improve government-citizen interactions, government agencies are now providing a wide spectrum of online services that implement the long awaited DG. These e-government services usually span several application domains, including social programs, healthcare, voting, tax filing, etc. DARPA Agent Markup Language-Service (DAML-S) is a major effort aimed at enabling the semantic description of web services (www.daml.org/services). However, DAML-S gives little support for the business process semantics of web services. WSMO uses the Web Service Modeling Framework as a starting point and refines it by developing a formal ontology and language. SWORD (a developer toolkit for web service composition) is a technique that uses a rule-based expert system to automatically determine whether a desired composite service can be constructed using existing services.

Disadvantages:

- Data integration problem, involves combining data residing in different sources and providing users with a unified view of these data.
- One of the biggest barriers to providing services to customers is the lack of integrated information systems.
- DAML-S and WSMO provides little or no support for specifying interoperation relationships.
- SWORD doesn't focus on relationships between web services or customized composite web service generation based on user profiles.

PROPOSED Models

The proposed WSMS manages the entire life cycle of third-party web services. These act as proxies for real government services. Due to the specific requirements of our digital government application,we focus on the following key components of WebSenior: service composition, service optimization, and service privacy preservation. These components form the nucleus that achieves seamless cooperation among government agenciesto provide prompt and customized services to senior citizens

Advantages:

- It provides an integrated service framework that achieves a seamless cooperation among government agencies to provide prompt and customized services to senior citizens.
- The specific needs of senior citizens by providing customized access to the services with their desired quality.
- Successful deployment of DG applications provides an efficient framework that accommodates autonomy, bridges heterogeneity, and integrates data and applications in the most useful and homogeneous way.

a.1. User Interaction Management:

The functionality of the user interaction tier is to authenticate users, present service information, and record and analyze user's behavior. After identifying the user, the user interface will change adaptively according to user's static profile and his or her current operation behaviors. There are six components in this level: Adaptive User Interface Generator, User Authentication, User Behavior Monitor, User Behavior Analyzer, Request Handler, and User Profile Management.

a.2. User Behavior Analyzer(In Graph):

These are component residents on the client side. Its main functionality is to track the user's operations on the computer. The component records the data, and aggregates and delivers it to the User Behavior Analyzer. After gathering the data from the User Behavior Monitor, User Behavior Analyzer analyzes the data and generates conclusions based on the behavior evaluation metrics.

a.3. Service composer:

Different senior citizens may make different service requests based on their particular situations. The service composition component enables customized delivery of services, which is particularly suitable for e-government applications. Service composer composes several web services to fulfill users' requests with value-added services. It needs to contact the service locator to get the required service descriptions. It also needs to contact the ontology manager to get the ontology information.

a.4. Service optimizer:

Service optimizer aims to provide the best quality of service through execution of a set of web services. It first selects web services (single services and composite services) to generate an execution plan. The selection is based on a request quality of service model to optimize the execution plan. The service optimizer then accepts the request of the user.

Some are the fallowing services provided by the e-gov. applications

Services for citizen Service

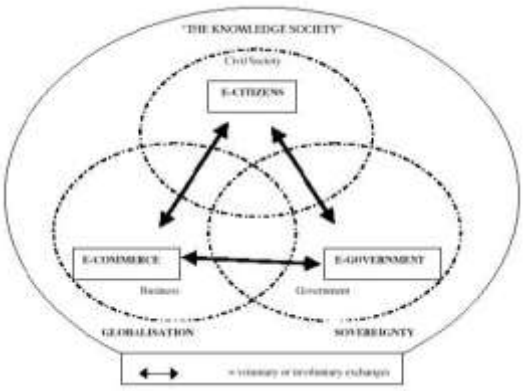
1. View movies-The citizen will see the available movies.
2. Book movies-The citizen will book the movies.
3. View trains- The citizen will see the available trains.
4. Book trains- The citizen will book the available movies.
5. Feedback- The citizen will send the feedback to admin.

Services for senior citizens Service

1. View List Of meals—this service provides list of meals to senior citizens
2. Order the Meals—this service provides order the meals to admin
3. Feed Back —this service is used to send the feed back to admin regarding the meals
4. View Medicines-this service provide the information about medicines
5. Order- Medicine-this service is used to send the medicine orders to admin

A Triangle Relationship Model among Government, Business and Citizens

The following diagram illustrates the relationship among E-Government, EBusiness, and E-Citizens in the context of the emergence of the so called "knowledge society", globalization, and sovereignty:Given the scale, scope, multi-portfolio nature, and transformational potential of egovernment, it has been advocated that it should be treated as a holistic system adjunct to the area of e-commerce in the E-society.



Conclusion

We provide the survey details of weservices provided by e-government applications . we describe the service oriented framework called WebSenior,to provide services to senior citizens. We implemented actual social services using web services it manages the life cycle of government social services.

We have shown a set of key service components in WebSenior, consisting of service composition, optimization, and privacy preservation. In particular, WebSenior provides the dynamic services, select service providers based on their quality attributes, and enforce the privacy of citizens when requesting and receiving government services.

Finally we conclude that framework is suitable for government and non profitable ngo's sites for providing service to the citizens.

References

[1.] Blake Harris, E-governance, 2000
[2.] (http://www.iadb.org)
[3.] Brendan Boyle, Electronic Government
[4.] for New Zealand: Managing the
[5.] Transition., May 8, 2000. Thesis of
[6.] MBA.
[7.] David C. Wyld, "The Auction Model:
[8.] How the Public Sector Can Leverage
[9.] the Power of E-Commerce Through
[10.] Dynamic Pricing". The PricewaterhouseCoopers
[11.] Grand Paper. October
[12.] 2000.
[13.] Electronic Government Task Force of
[14.] State of Texas, TexasOnline: A
[15.] Feasibility Report on Electronic
[16.] Government, 2001. (http://www.dir.state.
[17.] tx.us/egov/)
[18.] Jeffrey Roy, E-Government: Enabling &
[19.] Empowering Tomorrow's Public
[20.] Service, January 31, 2000. (www.govern
[21.] ance.uottawa.ca).
[22.] Jim Melitski, The World of Egovernment
[23.] and E-governance, 2001.
[24.] http://www.aspanet.org/solutions/The
[25.] World of E-government and Egovernance.
[26.] htm.
[27.] Mark Cleverley, *e-Government Symposium*
[28.] *a Great Success*, May 10, 2001,
[29.] NYSFIRM.
[30.] Maria A. Wimmer, Knowledge Management
[31.] in e-Government, 2001.
[32.] http://falcon.ifs.uni-linz.ac.at/
[33.] Rob Atkinson, "Creating a Digital Federal
[34.] Government", *Information Impact*,
[35.] October, 2000, http://www.cisp.org/
[36.] imp/october_2000/Creating a Digital
[37.] Federal Government.htm.
[38.] Service-Centric Framework for a
[39.] Digital Government Application IEEE TRANSACTIONS ON SERVICES COMPUTING, VOL. 4, NO. 1, JANUARY-MARCH 2011

Authors profile

CH.SRAVANTHI

ph:9059173083
Areas of Interest:Web Technologies,Data mining,Dataware Housing

K.KailasaRao,

Professor,DEPT Of CSE.
MALLA REDDY COLLEGE OF ENGINEERING & TECHNOLOGY,

A Novel Design of Fractal Antenna with EBG-GP

Yogesh[1], Aijaz Ahmed[2], Sagar Erande[3]
[1, 2, 3](Amity School of Engineering and Technology, Amity University Uttar Pradesh, India)

Abstract:
A new microstrip antenna is developed with a fractal patch and an electromagnetic bandgap ground plane (EBG-GP). This antenna was designed using Ansoft HFSS. A prototype was built and measured. The measured resonant frequency for this antenna is lower than that for a typical rectangular patch antenna with inset-fed, resulting in a size reduction of about 33.9 % at 2.45 GHz.

Keywords: Fractal Antenna, Electromagnetic bandgap, microstrip antenna, rectangular, patch, wireless, EBG-GP

1. Introduction

Fractal curves are used in the design of a patch antenna with an electromagnetic bandgap ground plane (EBG-GP), in order to get size reduced antennas [1]. The use of an EBG ground plane in the development of a fractal antenna level 1 allows to increasing the antenna bandwidth without change the overall antenna performance when compared to the typical microstrip antenna geometry. The proposed fractal antenna with EBG-GP has shown good radiation properties with good return loss results (<-20 dB) and impedance matching condition using the inset-feed technique.

2. Antenna Design

In the development of the fractal antenna level 1 a scaling factor of ¼ L and ¼ W was applied in a rectangular microstrip conducting patch with L and W dimensions, resulting in the fractal geometry shown in Figure 1(a). Observe that an inset-feed technique was used to match the antenna. The EBG-GP geometry is shown in Figure 1(b). The periodic array in the ground plane was obtained after a series of simulations in order to improve the antenna bandwidth. The antenna was built on a FR-4 substrate with ε_r equal to 4.4 and height equal to 1.5 mm.

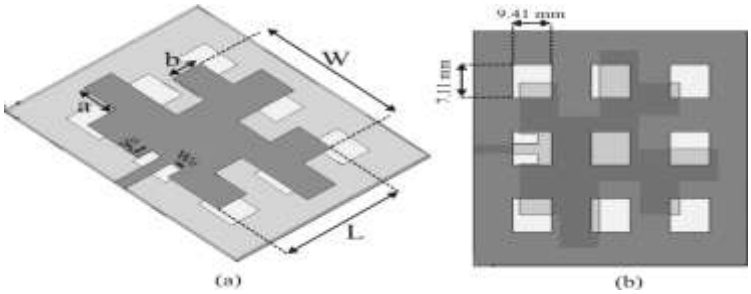


Fig. 1: Fractal microstrip Antenna with EBG-GP: (a) perspective view and (b) bottom view.

The proposed antenna was designed using Ansoft HFSS that implements the finite element method. A prototype was built and measured using a vector network analyzer.

3. Simulated Results & Discussion

Two resonant frequencies were measured and simulated: 1.62 GHz and 2.0 GHz. Figure 2(a) shows the measured and simulated results in HFSS for the return loss.

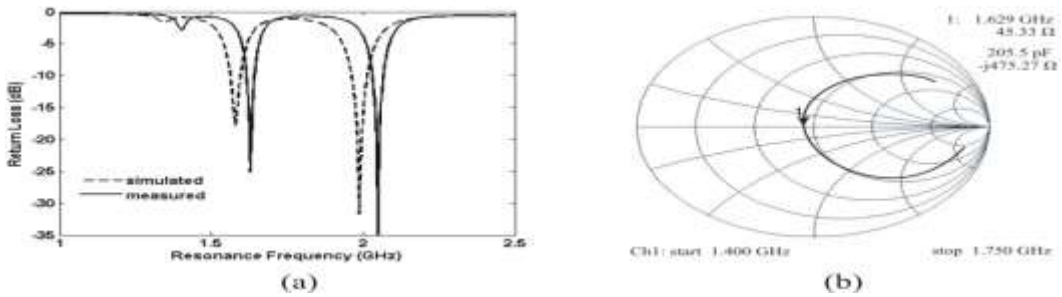


Figure 2: (a) Return loss and (b) input impedance results.

The input impedance measured at 1.62 GHz is 45.33 Ω. The antenna dimensions and parameters are given in Table below.

Table: Proposed antenna dimensions and parameters.

Measured resonant frequency	1.62 GHz / 2 GHz
Measured return loss	25.1 dB / -41.2 dB
Bandwidth	31 MHz / 40 MHz
Bandwidth (%)	1.91% / 2%
VSWR	1.13
Relative permittivity (εr)	4.4
Loss tangent (tan δ_e)	0.02
Patch length (L)	29.094 mm
Patch width (W)	37.2343 mm
y_0	3.4 mm
W_0	2.87 mm
a	7.27 mm
b	7.27 mm

4. Conclusion

This work described the development of an EBG-GP fractal antenna level 1 in order to reduce its resonant frequency, when compared to that of a rectangular microstrip patch antenna, and to improve its bandwidth for wireless communication system applications. Good agreement between measured and simulated results was obtained.

References

[1] Vinoy, K. J., "Fractal shaped antenna elements for wide and multi-band wireless applications," Pennsylvania, Aug. 2002.

[2] Garg, Bhatia, Bahl, Ittipiboon, "Microstrip Antenna Design Handbook", Artech House, London, 2000.

[3] Yahui Zhao, Jinping Xu, and Kang Yin, "A Miniature Coplanar Waveguide-Fed Ultra-Wideband Antenna", State Key Laboratory of Millimeter Waves, University, Nanjing, Jiangsu, P.R.China, 210096

[4] Masahiro Yanagi, Shigemi Kurashima, Takashi Arita, Takehiko Kobayashi, "A Planar UWB Monopole Antenna Formed on a Printed Circuit Board" W. Cheney and D. Kincaid.

[5] B. B. Mandelbrot, The Fractal Geomety of Nature, New York, W. H. Freeman, 1983.

[6] D. L. Jaggard, "On Fractal Electrodynamics," in D. L. Jaggard (eds.), Recent Advances in Electromagnetic Theory, New York, Springer-Verlag, 1990, pp. 183-224.

GSM Modem Based Data Acquisition System

Vandana Pandya[1] Deepali Shukla[2]
[1,2],(Asst.Professor) ,
[1,2],Medicaps Institute of Technology & Management, Indore (M.P.), India

Abstract

GSM Modem Based Data Acquisition is a process control system .it help in collecting data from various processes present at distant places. It helps us to monitor parameters like temperature ,rainfall ,humidity etc. the Service personnel need not to visit distant sites for data collection .the data is collected automatically formed a data base and stored in a PC. It can be used in industries as well as in home automation.

Key Words: GSM Modem ,SIM card, Arduino Board, Serial cable, PC.

1. Introduction

This system uses AVR microcontroller ATmega 644P. the inbuilt ADC receives analog data from sensors and converts it to digital data and passes it to the microcontroller . the sensors continuously sends data from the distant site. This system is interfaced with a GSM modem. this system senses the conditions continuously and a message is sent to a mobile no. using SMS on LCD every 10 minutes. Using this system, the operator can monitor the signals from any where. The GSM modem is connected to microcontroller using RS232 interface. Whenever an SMS is sent to the GSM modem, the GSM modem receives the data and sends to microcontroller. After receiving the signal from the microcontroller it processes the data and sends the read data to mobile number through GSM modem. The collected data is formed a database and stored in a pc. The received data is displayed on the LCD. 16X 2 LCD is provided for user interface.

2. Method

2.1 Block Diagram

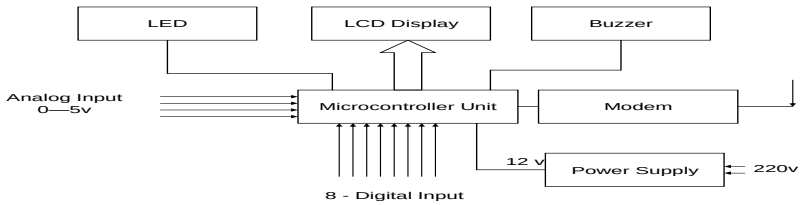


Figure 1 GSM Modem Based Data acquisition System

2.2 Working

- This system uses AVR microcontroller ATmega 644P.
- Eight channel ADC receives analog data from various sensors connected to it and converts it to digital.
- This digital data is sensed by the microcontroller and the system continuously monitors the data condition.
- This system is interfaced with a GSM modem.
- The system monitors the conditions continuously and sends message to mobile numbers using SMS every ten minutes.
- This data is displayed on LCD every second. Using this system, the operator can monitor the signals from anywhere in the world as it works on GSM network.
- At the receiving end the SMS is received by the modem and processed by the microcontroller and the data is collected by a central server and formed a data base and stored in a pc.

- Mobile Switching Centre is used for short messages originated with a mobile on that network. The SMS-GMSC role is similar to that of the GMSC, whereas the SMS-IWMSC provides a fixed access point to the Short Message Service Centre.

3. GSM Modem

A GSM modem is a specialized type of modem which accepts a SIM card, and operates over a subscription to a mobile operator, just like a mobile phone. From the mobile operator perspective, a GSM modem looks just like a mobile phone. When a GSM modem is connected to a computer, this allows the computer to use the GSM modem to communicate over the mobile network. While these GSM modems are most frequently used to provide mobile internet connectivity, many of them can also be used for sending and receiving SMS and MMS messages.GSM modem must support an "extended AT command set" for sending/receiving SMS messages. GSM modems can be a quick and efficient way to get started with SMS, because a special subscription to an SMS service provider is not required. In most parts of the world, GSM modems are a cost effective solution for receiving SMS messages, because the sender is paying for the message delivery.

To begin, insert a GSM SIM card into the modem and connect it to an available USB port on your computer.

3.1 Features of SIM300 GSM Module

- Designed for global market, SIM300 is a Tri-band GSM engine
- Works on frequencies EGSM 900 MHz, DCS 1800 MHz and PCS 1900 MHz.
- SIM300 features GPRS multi-slot class 10/ class 8 (optional) and supports the GPRS coding schemes.
- CS-1, CS-2, CS-3 and CS-4.With a tiny configuration of 40mm x 33mm x 2.85mm.
- SIM300 can fit almost all the space requirements in your applications, such as smart phone, PDA phone and other mobile devices.
- This GSM modem is a highly flexible plug and play quad band GSM modem
- interface to RS232.
- Supports features like Voice, Data/Fax, SMS,GPRS and integrated TCP/IP stack.
- Control via AT commands(GSM 07.07,07.05 and enhanced AT commands)
- Use AC – DC Power Adaptor with following ratings · DC Voltage : 12V /1A
  Current Consumption in normal operation 250mA, can rise up to 1Amp while transmission

3.2 Interfaces

RS-232 through D-TYPE 9 pin connector,

- Serial port baud rate adjustable 1200 to115200 bps (9600 default)
- BRK connector for MIC & SPK, SIM card holder
- Power supply through DC socket
- SMA antenna connector and Murata Antenna (optional)
- LED status of GSM / GPRS module

3.3 AT Commands

AT Commands are used to control a modem. AT means Attention. Every command line starts with "AT".These are of two types : Basic and Extended.

- ATEO – Echo off
- ATE1- Echo on
- ATD –Call a dial no.
  Syntax : ATD 9479555640
- ATDL- Redial last telephone no.
- ATA- Answer an incoming call
- ATH-Disconnect existing connection
- AT+CMGS-To send SMS
  Syntax: AT+CMGS="9479555640" Press enter

Type text and press ctrl+z

- AT+CMGR – To read SMS
  Syntax : AT+ CMGR=1 ; reads first SMS in sim card
- AT+CMGD – To delete SMS
  Syntax : AT+CMGD = 1 ; deletes first SMS in sim card

4. Hardware- The Arduino

Arduino is a tool for the design and development of embedded computer systems, consisting of a simple open hardware design for a single-board microcontroller, with embedded I/O support and a standard programming language. An Arduino is a tool for making computers that can sense and control more of the physical world.
Arduino can sense the environment by receiving input from a variety of sensors and can affect its surroundings by controlling lights, motors, and other actuators. The microcontroller on the board is programmed using the Arduino programming language(based on Wiring) and the Arduino development environment (based on Processing).

4.1 RS 232 Circuit

Since RS232 is not compatible with microcontrollers we need a voltage converter to convert the RS232's signals to TTL voltage levels. These are acceptable to the microcontroller's TxD and RxD pins. The MAX 232 converts the RS232 voltage levels to TTL voltage levels and vice versa.
The chip uses +5v power source which is the same as the power source for the microcontroller. It provides 2-channel RS232C port and requires external 10uF capacitors.

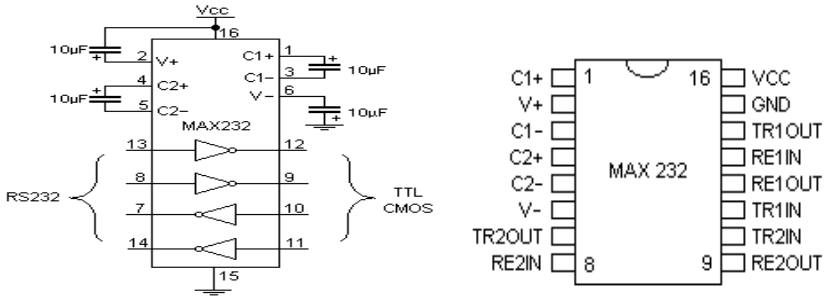
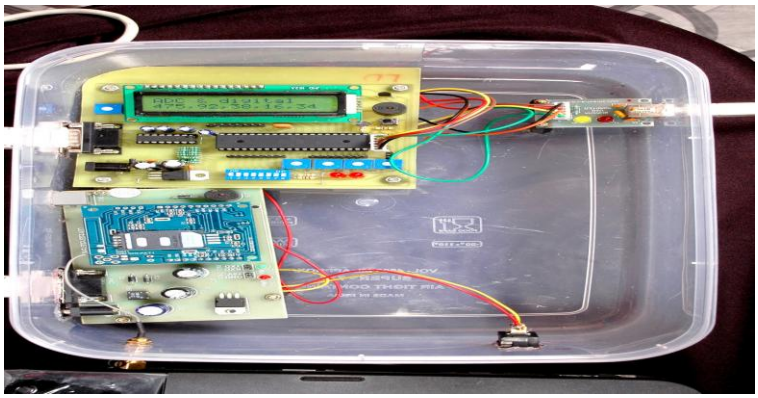


Figure 2

5. Arduino Software

The Arduino IDE is a cross-platform application written in Java which is derived from the IDE made for the Processing programming language and the Wiring project. It is designed to introduce programming to artists and other newcomers unfamiliar with software development. It includes a code editor with features such as syntax highlighting, brace matching, and automatic indentation, and is also capable of compiling and uploading programs to the board with a single click. There is typically no need to edit Makefiles or run programs on the command line. The Arduino IDE comes with a C / C++ library called "Wiring" (from the project of the same name), which makes many common input/output operations much easier. Arduino programs are written in C/C++, although users only need to define two functions in order to make a runnable program:
setup() – a function run once at the start of a program which can be used for initializing settings, and
loop() – a function called repeatedly until the board is powered off.
The code written in Arduino not be seen by a standard C++ compiler as a valid program, so when the user clicks the "Upload to I/O board" button in the IDE, a copy of the code is written to a temporary file with an extra include header at the top and a very simple main() function at the bottom, to make it a valid C++ program.
Since the IDE is pin oriented, we can quickly achieve our desired logic and build a working model



5.1 Functions used in Arduino

- **pinMode()**

Configures the specified pin to behave either as an input or an output.
pinMode(pin, mode)
pin : the number of the pin whose mode you wish
to set .
mode: either input or output.

- **digitalRead()**

Reads the value from a specified digital pin, either HIGH or LOW.
digitalRead(pin)
pin : the number of the digital pin you want to read (*int*)

- **digitalWrite()**

Write a HIGH or a LOW value to a digital pin

- **analogRead()**

Reads the value from the specified analog pin.

The Arduino board contains a 6 channel ,10-bit analog to digital converter.
This means that it will map input voltages between 0 and 5 volts into integer values between 0 and 1023. This yields a resolution between readings of: 5 volts / 1024 units or, 4.9 mV per unit.

- **analogWrite()**

Writes an analog value (PWM wave) to a pin.
After a call to **analogWrite()**, the pin will generate a steady square wave of the specified duty cycle.
The frequency of the PWM signal is approximately 490 Hz .

5.2 Serial Communication

The Serial port (USART) is used for communication between the Sanguino board and a computer or other devices.
The Sanguino has two serial ports :
Serial1 on pins 19 (RX) and 18 (TX)
Serial2 on pins 17 (RX) and 16 (TX)

- Serial.begin(int baud)
- Serial.available( )
- Serial.read( )
- Serial.println( )

6. Results

The design was successfully dissected into various modules. Each module with the exception of the database system has been implemented. The integration of the individual modules was achieved successfully. All the results were categorically measured. All the GUI functionality discussed in the report is operational .

- The GSM Modem is tested using hyperterminal on pc and run the AT commands.
- Read four analog channels and 8 bit digital and display it on serial port, LCD and send it through SMS.
- Receive SMS and display it on LCD.
- Received SMS display on pc using VisualC#.net and form a database(MS Access) to store it.

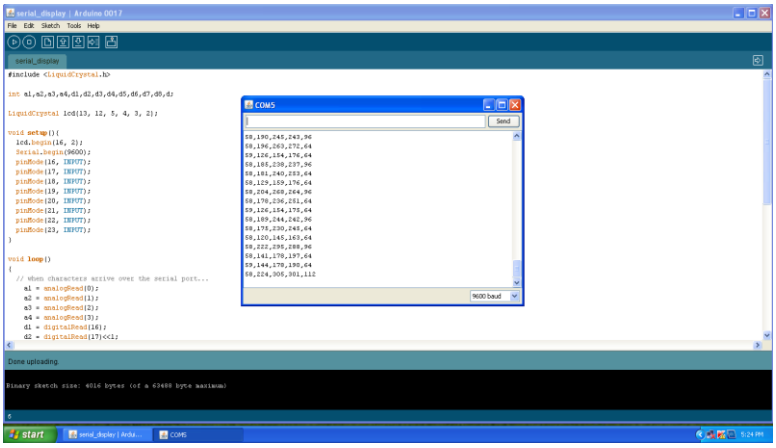


Figure 3: read four analog channels and 8 bit digital and display it on serial port

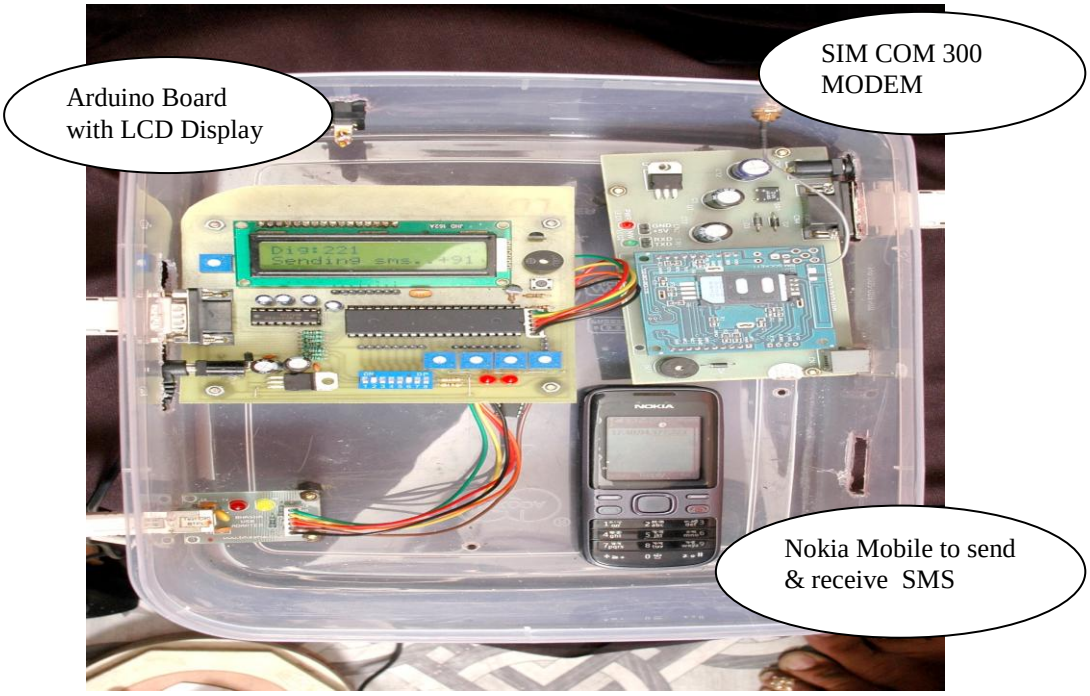


Figure 4: The System Designed

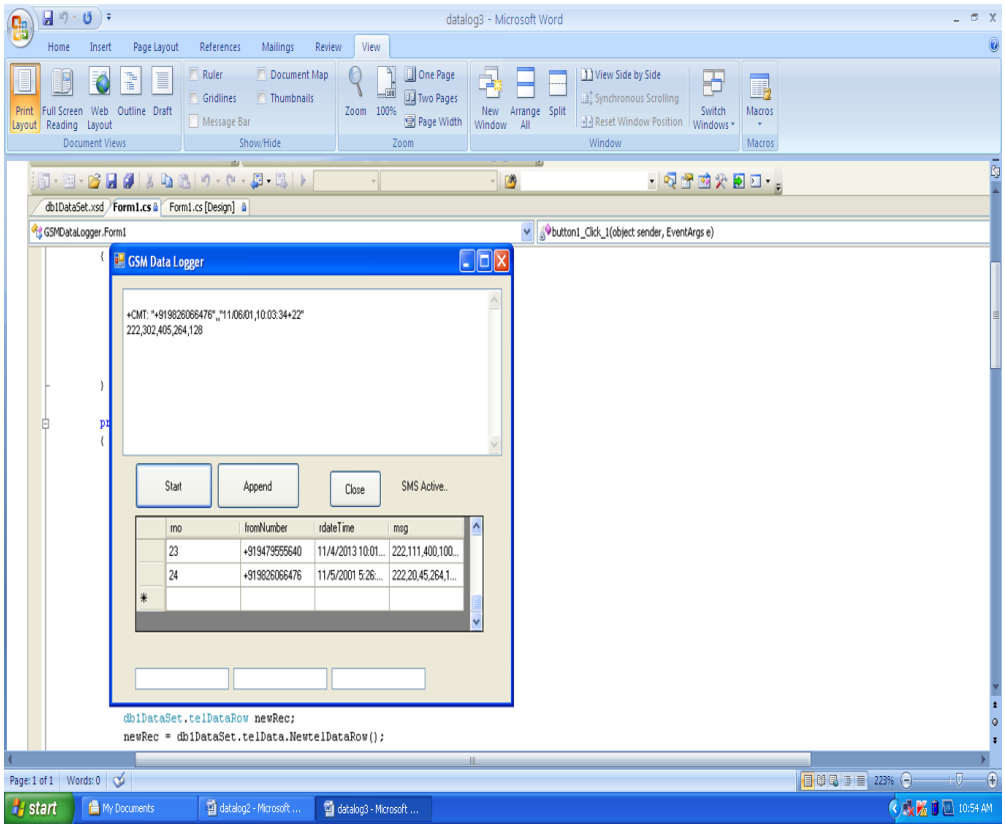


Figure 5: GSM Data Logger

REFERENCES

[1] .S.C.S. Jucá, P.C.M. Carvalho and F.T. Brito, "A low cost concept for data acquisition systems applied to decentralized renewable energy plants", Sensors, 2011, vol.11, pp. 743-756

[2]. S. Rosiek and F. Batlles, "A microcontroller-based dataacquisition system for meteorological station monitoring" Energy Conversion and Management, 2008, vol. 49, pp. 3746-3754.

[3]. M. Benghanem, "A low cost wireless data acquisition system for weather station monitoring", Renewable Energy, 2010, vol. 35, pp.862–872.

[4]. H. Belmili, S. Cheikh, M. Haddadi and C. Larbes, "Design and development of a data acquisition system for photovoltaic modules characterization". Renewable Energy, 2010, vol. 35, pp.1484–1492.

[5] Dr. Aditya Goel & Ravi Shankar Mishra," Remote Data Acquisition Using Wireless - Scada System"International Journal of Engineering (IJE), Volume (3) : Issue (1) 58

[6] SMS Tutorial –How to use Microsoft hyperterminal to send AT commands

[7] Hardware description of GSM modem simcom300 reference manual

Magnetohydrodynamic Free Convection Boundary Layer Flow past A Vertical Cone with Uniform Heat And Mass Flux

S. Gouse Mohiddin[1], O. Anwar Bég[2] and S. Vijaya Kumar Varma[3]

[1]Department of Mathematics, Madanapalle Institute of Technology & Science, Madanapalle – 517325, AP, India
[2]Biomechanics and Biotechnology Research, Aerospace Engineering Program, Mechanical Engineering Subject Group, Sheaf Building, Sheffield Hallam University, Sheffield, S1 1WB, England, UK
[3]Department of Mathematics, Sri Venkateswara University, Tirupati – 517502

Abstract

A numerical solution of unsteady laminar free convection in a viscoelastic fluid flow past a vertical cone with transverse magnetic field applied normal to the surface with uniform heat and mass flux is presented. The Walters-B short-memory liquid model is employed to simulate medical creams and other rheological liquids encountered in biotechnology and chemical engineering. This rheological model introduces supplementary terms into the momentum conservation equation. The dimensionless unsteady, coupled and non-linear partial differential conservation equations for the boundary layer regime are solved by an unconditionally stable implicit finite difference scheme of Crank-Nicolson type. The velocity, temperature and concentration fields have been studied for the effect of viscoelasticity parameter, Prandtl number, magnetic parameter, Schmidt number, and buoyancy ratio parameter and semi vertical angle. The local skin friction, Nusselt number and Sherwood number are also presented and analyzed graphically. The numerical results are validated by comparisons with previously published work and are found to be in excellent agreement.

Key words: cone; finite difference method; flux; Walters-B short-memory model; AMS Subject Classification : 35Q35, 76R10, 76W05. 80A20.

1. Nomenclature

x, y	coordinates along the cone generator and normal to the generator respectively
X, Y	dimensionless spatial coordinates along the cone generator and normal to the generator respectively
u, v	velocity components along the x- and y- directions respectively
U, V	dimensionless velocity components along the X- and Y- directions respectively
Gr_L	Grashof number
g	gravitational acceleration
t'	time
t	dimensionless time
Nu_x	local Nusselt number
Nu_X	non-dimensional local Nusselt number
Pr	Prandtl number
r	local radius of cone
R	dimensionless local radius of the cone
T'	temperature
T	dimensionless temperature
C'	concentration
C	dimensionless concentration
D	mass diffusion coefficient
L	reference length
Sc	Schmidt number
k_0	Walters-B viscoelasticity parameter

Γ dimensionless Walters-B viscoelasticity parameter
N buoyancy ratio parameter
M magnetic parameter
B_0 magnetic field strength
Sh_x local Sherwood number
Sh_X dimensionless local Sherwood number
$f''(0)$ local skin-friction in Ref. [6]

Greek Symbols

β volumetric thermal expansion
μ dynamic viscosity
ν kinematic viscosity
Δt dimensionless time-step
ρ density
α thermal diffusivity
ϕ semi vertical angle of the cone
ΔX dimensionless finite difference grid size in X-direction
ΔY dimensionless finite difference grid size in Y-direction
τ_x local skin friction
τ_X dimensionless local skin-friction
η dimensionless independent variable in [6]
θ temperature in Ref. [6]

Subscripts

w condition on the wall
∞ free stream condition

2. Introduction

Heat and mass transfer in non-Newtonian fluids is of great interest in many operations in the chemical and process engineering industries. Many geometrical configurations have been addressed including flat plates, channels, cones, spheres, wedges, inclined planes and wavy surfaces. The Walters-B viscoelastic model [14] was developed to simulate viscous fluids possessing short memory elastic effects and can simulate accurately many complex polymeric, biotechnological and tribological fluids. The Walters-B model has therefore been studied extensively in many flow problems. One of the first mathematical investigations for such a fluid was presented in [12]. Flat plate thermal convection boundary layer flow of a Walters-B fluid using numerical shooting quadrature was studied in [11]. Exact solutions were obtained in [10] for the combined nonsimilar hydromagnetic flow, heat, and mass transfer phenomena in a conducting viscoelastic Walters-B fluid. Recently theoretical studies on laminar free convection flow of axi-symmetric bodies have received wide attention especially in case of uniform and non-uniform surface heat and mass flux. Similarity solutions for the laminar free convection from a right circular cone were presented in [6]. Reference [8] focused the theoretical study on the effects of suction or injection on steady free convection from a vertical cone with uniform surface heat flux condition. Non-similarity solutions were studied in [5] for the free convection from a vertical permeable cone with non-uniform surface heat flux. Saturated porous media combined heat and mass transfer effects over a full cone was studied in [15].
Magnetohydrodynamic (MHD) flow and heat transfer is of considerable interest because it can occur in many geothermal, geophysical, technological, and engineering applications such as nuclear reactors and others. The geothermal gases are electrically conducting and are affected by the presence of a magnetic field. Vajravelu and Nayfeh [13] studied hydromagnetic convection from a cone and a wedge with variable surface temperature and internal heat generation or absorption. The above studies did not consider combined viscoelastic momentum, heat and mass transfer from a vertical cone, either for the steady case or unsteady case. Owing to the significance of this problem in chemical and medical biotechnological processing (e.g. medical cream manufacture) we study the transient case of such a flow in the present paper using the Walters-B viscoelastic rheological material model.

3. Constitutive Equations For The Walters-B Viscoelastic Fluid

$$p_{ik} = -pg_{ik} + p^*_{ik} \quad (1)$$

$$p^*_{ik} = 2\int_{-\infty}^{t} \Psi\left(t-t^*\right) e^{(1)}_{ik}(t^*)dt^* \quad (2)$$

$$\Psi\left(t-t*\right) = \int_0^{\infty} \frac{N(\tau)}{\tau} e^{\left(-\left(t-t^*\right)/\tau\right)} d\tau \quad (3)$$

where p_{ik} is the stress tensor, p is arbitrary isotropic pressure, g_{ik} is the metric tensor of a fixed coordinate system x_i, $e^{(1)}_{ik}$ is the rate of strain tensor and $N(\tau)$ is the distribution function of relaxation times, τ . The following generalized form of (2) has been shown by Walters [14] to be valid for all classes of motion and stress.

$$p^{*ik}(x,t) = 2\int_{-\infty}^{t} \Psi\left(t-t^*\right) \frac{\partial x^i}{\partial x^{*m}} \frac{\partial x^k}{\partial x^{*r}} e^{(1)mr}\left(x^* t^*\right) dt^* \quad (4)$$

in which $x^*_i = x^*_i(x,t,t^*)$ denotes the position at time t^* of the element which is instantaneously at the position, x_i, at time, t. Liquids obeying the relations (1) and (4) are of the Walters-B' type. For such fluids with short memory i.e. low relaxation times, eqn (4) may be simplified to:

$$p^{*ik}(x,t) = 2\eta_0 e^{(1)ik} - 2k_0 \frac{\partial e^{(1)ik}}{\partial t} \quad (5)$$

in which $\eta_0 = \int_0^{\infty} N(\tau)d\tau$ defines the limiting Walters-B' viscosity at low shear rates, $k_0 = \int_0^{\infty} \tau N(\tau)d\tau$ is the Walters-B' viscoelasticity parameter and $\frac{\partial}{\partial t}$ is the convected time derivative. This rheological model is very versatile and robust and provides a relatively simple mathematical formulation which is easily incorporated into boundary layer theory for engineering applications.

4. Mathematical Model

An axi-symmetric unsteady laminar free convective flow of a viscoelastic fluid past a vertical cone with uniform surface heat and mass flux is considered. Unsteady incompressible flow and non-reactive mass diffusion in a free convective viscoelastic flow driven by species buoyancy force occurs upwards along the cone. Also implicit in our analysis is the assumption that the cone surface and the surrounding fluid which are at rest possess the same temperature T'_∞ and concentration C'_∞ . At time t' >0, heat supplied from the cone surface to the fluid, concentration level near the cone surface are raised at uniform rate and are sustained as constant thereafter. In addition, the formulation of mathematical equations is based on the following assumptions:

- The concentration C' of the diffusing species in the binary mixture is assumed to be very less in comparison to the other chemical species, which are present. This leads to the assumption that the Soret and Dufour effects are negligible.
- The viscous dissipation effects and pressure gradient along the boundary layer are negligible.
- The magnetic Reynolds number is very small so that the induced magnetic field is negligible.
- The magnetic field is not strong enough to cause Joule heating so that the term due to electrical dissipation is neglected in energy equation.

The co-ordinate system chosen (as shown in Fig.1) is such that the *x*- coordinate is directed along the surface of the cone from the apex ($x = 0$) and the *y*- coordinate is orientated perpendicular to this i.e. at right angles to the cone surface, outwards. Here, ϕ designates the semi-vertical angle of the cone and *r* is the local radius of the cone. In compliance with the Boussinesq approximation, all fluid properties are assumed constant except for density variations, which induce *buoyancy* forces, these contributing a dominant, driving role in the free convection regime.

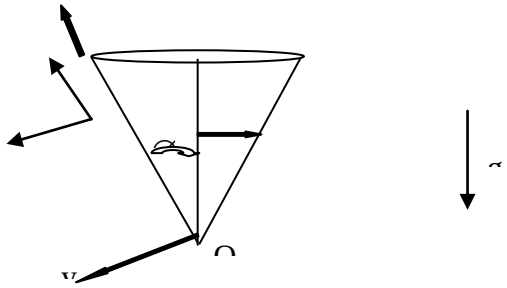


Figure 1: Physical Scenario with Coordinate Axes

Under the above assumptions, implementing the shear-stress strain tensor for a Walters-B liquid, the appropriate unsteady incompressible conservation equations, (neglecting convective inertial terms in the momentum equation) may be shown to take the form: $\frac{\partial(ur)}{\partial x}+\frac{\partial(vr)}{\partial y}=0$

(6) $$\frac{\partial u}{\partial t'}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=\nu\frac{\partial^2 u}{\partial y^2}-k_0\frac{\partial^3 u}{\partial y^2\partial t'}-\frac{\sigma B_0^2}{\rho}u+g\beta\cos\phi(T'-T'_\infty)+g\beta^*\cos\phi(C'-C'_\infty) \quad (7)$$

$$\frac{\partial T'}{\partial t'}+u\frac{\partial T'}{\partial x}+v\frac{\partial T'}{\partial y}=\alpha\frac{\partial^2 T'}{\partial y^2} \quad (8)$$

$$\frac{\partial C'}{\partial t'}+u\frac{\partial C'}{\partial x}+v\frac{\partial C'}{\partial y}=D\frac{\partial^2 C'}{\partial y^2} \quad (9)$$

The initial and boundary conditions are prescribed as:

$$t'\le 0: u=0, v=0, T'=T'_\infty, C'=C'_\infty \quad \text{for all x, y,}$$
$$t'>0: u=0, v=0, T'=T'_w, C'=C'_w \quad \text{at y = 0,}$$
$$u=0, T'=T'_\infty, C'=C'_\infty \quad \text{at x = 0,} \quad (10)$$
$$u\to 0, T'\to T'_\infty, C'\to C'_\infty \quad \text{as } y\to\infty.$$

where all parameters are defined in the nomenclature. Proceeding with analysis, we implement the following non-dimensional quantities to facilitate a numerical solution to the boundary value problem defined by Eqns. (6) to (9) under conditions (10):

$$X=\frac{x}{L}, \qquad Y=\frac{y}{L}(Gr_L)^{\frac{1}{4}}, \qquad R=\frac{r}{L}, \text{where } r=x\sin\phi,$$
$$V=\frac{vL}{\nu}(Gr_L)^{\frac{-1}{4}}, \qquad U=\frac{uL}{\nu}(Gr_L)^{\frac{-1}{2}}, \qquad t=\frac{\nu t'}{L^2}(Gr_L)^{\frac{1}{2}}, \quad (11)$$

$$T=\frac{T'-T'_\infty}{T'_w(L)-T'_\infty},\ Gr_L=\frac{g\beta\left(T'_w(L)-T'_\infty\right)L^3}{\nu^2},\quad \Pr=\frac{\nu}{\alpha},\ M=\frac{\sigma B_0^2L^2}{\mu}Gr_L^{\frac{-1}{2}}$$

$$C=\frac{C'-C'_\infty}{C'_w-C'_\infty},\qquad \Gamma=\frac{k_0Gr_L^{\frac{1}{2}}}{L^2},\quad N=\frac{\beta^*(C'_w-C'_\infty)}{\beta(T'_w-T'_\infty)},\quad Sc=\frac{\nu}{D}$$

Equations (6), (7), (8) and (9) are reduced to the following non-dimensional form

$$\frac{\partial(UR)}{\partial X}+\frac{\partial(VR)}{\partial Y}=0 \tag{12}$$

$$\frac{\partial U}{\partial t}+U\frac{\partial U}{\partial X}+V\frac{\partial U}{\partial Y}=\frac{\partial^2U}{\partial Y^2}-\Gamma\frac{\partial^3U}{\partial Y^2\partial t}-MU+T\cos\phi+NC\cos\phi \tag{13}$$

$$\frac{\partial T}{\partial t}+U\frac{\partial T}{\partial X}+V\frac{\partial T}{\partial Y}=\frac{1}{\Pr}\frac{\partial^2T}{\partial Y^2} \tag{14}$$

$$\frac{\partial C}{\partial t}+U\frac{\partial C}{\partial X}+V\frac{\partial C}{\partial Y}=\frac{1}{Sc}\frac{\partial^2C}{\partial Y^2} \tag{15}$$

The corresponding non-dimensional initial and boundary conditions are given by:

$$\begin{aligned} &t\le 0: U=0, V=0, T=0, C=0 && \text{for all X, Y,} \\ &t>0: U=0, V=0, T=1, C=1 && \text{at Y = 0,} \\ &U=0, \quad T=0, \ C=0 && \text{at X = 0,} \\ &U\to 0, \quad T\to 0, \ C\to 0 && \text{as } Y\to\infty. \end{aligned} \tag{16}$$

Where again all parameters are given in the nomenclature. The dimensionless local values of the skin friction (surface shear stress), the Nusselt number (surface heat transfer gradient) and the Sherwood number (surface concentration gradient) are given by the following expressions:

$$\tau_x=-\left(\frac{\partial U}{\partial Y}\right)_{Y=0} \tag{17}$$

$$Nu_x=-X\left(\frac{\partial T}{\partial Y}\right)_{Y=0} \tag{18}$$

$$Sh_x=-X\left(\frac{\partial C}{\partial Y}\right)_{Y=0} \tag{19}$$

We note that the dimensionless model defined by Eqns. (12) to (15) under conditions (16) reduces to *Newtonian* flow in the case of vanishing viscoelasticity i.e. when $\Gamma\to 0$.

5. Numerical Solution

In order to solve these unsteady, non-linear coupled equations (12) to (15) under the conditions (16), an implicit finite difference scheme of Crank-Nicolson type which is discussed by many authors Muthucumaraswamy and Ganesan [7], Ganesan and Rani [4], Ganesan and Loganathan [3], Prasad et al [9] and Bapuji et al. [1]. The finite difference scheme of dimensionless governing equations is reduced to tri-diagonal system of equations and is solved by Thomas algorithm as discussed in Carnahan et al. [2]. The region of integration is considered as a rectangle with $X_{\max}=1$ and $Y_{\max}=22$ where $Y_{\max}$ corresponds to $Y=\infty$ which lies very well out side both the momentum and thermal boundary layers. The maximum of Y was chosen as 22, after some preliminary investigation so that the last two boundary conditions

of (16) are satisfied within the tolerance limit 10^{-5}. The mesh sizes have been fixed as $\Delta X = 0.05, \Delta Y = 0.05$ with time step $\Delta t = 0.01$. The computations are carried out first by reducing the spatial mesh sizes by 50% in one direction, and later in both directions by 50%. The results are compared. It is observed in all cases, that the results differ only in the fifth decimal place. Hence, the choice of the mesh sizes seems to be appropriate. The scheme is unconditionally stable. The local truncation error is $O(\Delta t^2 + \Delta Y^2 + \Delta X)$ and it tends to zero as $\Delta t, \Delta X$ and ΔY tend to zero. Hence, the scheme is compatible. Stability and compatibility ensure the convergence.

6. Results and Discussion

Only selective figures have been reproduced here for brevity. Default values of the parameters are as follows: viscoelasticity parameter (Γ) = 0.005, buoyancy ratio parameter (N) = 1.0, Schmidt number (Sc) = 0.6 (oxygen diffusing in the viscoelastic fluid), magnetic parameter (M) = 1.0, Prandtl number (Pr) = 0.7 (water-based solvents) and semi-vertical angle of the cone (ϕ) = 20^0. All graphs therefore correspond to these values unless specifically otherwise indicated. In order to prove the accuracy of our numerical results, the present results in steady state at X = 1.0 are obtained by considering the modified Grashof number $Gr_L^* = Gr_L \cos\phi$, (i.e. the numerical solutions obtained from the equations (12) – (15) are independent of semi vertical angle of the cone ϕ)are compared with the available similarity solutions in the literature. The numerical values of local skin friction τ_X, temperature T, for different values of Prandtl number (Pr) are compared with those of Lin [6] in steady state using suitable transformation (i.e. $Y = (20/9)^{1/5}\eta$, $T = (20/9)^{1/5}(-\theta(0))$, $U = (20/9)^{3/5} f'(\eta)$, $\tau_X = (20/9)^{2/5} f''(0)$). It is observed that the results are in good agreement with each other. In addition, the values of local skin friction and temperature are compared with Bapuji et al. [1] for the case of Newtonian fluid in the absence of mass transfer in Table 1 and Table 2.

Table 1 Comparison of steady state local skin-friction values at X = 1.0 with those of Lin [6] and Bapuji et al. [1]

Pr	Lin[6]		Bapuji et al.[1]	Present Results
	$f''(0)$	$\left(\frac{20}{9}\right)^{2/5} f''(0)$	τ_x	
0.72	0.88930	1.2240	1.2154	1.22185
1	0.78446	1.0797	1.0721	1.07476
2	0.60252	0.8293	0.8235	0.82688
4	0.46307	0.6373	0.6328	0.63344
6	0.39688	0.5462	0.5423	0.54362
8	0.35563	0.4895	0.4859	0.48627
10	0.32655	0.4494	0.4460	0.44856
100	0.13371	0.1840	0.1813	0.18291

Table 2 Comparison of steady state temperature values at X = 1.0 with those of Lin [6] and Bapuji et al.[1]

Pr	Lin[6]		Bapuji et al.[1]	Present Results
	$-\theta(0)$	$-\left(\frac{20}{9}\right)^{1/5}\theta(0)$	T	
0.72	1.52278	1.7864	1.7796	1.78084
1	1.39174	1.6327	1.6263	1.62863
2	1.16209	1.3633	1.3578	1.35989
4	0.98095	1.1508	1.1463	1.14895
6	0.89195	1.0464	1.0421	1.04339
8	0.83497	0.9796	0.9754	0.97729
10	0.79388	0.9314	0.9272	0.92832
100	0.48372	0.5675	0.5604	0.56590

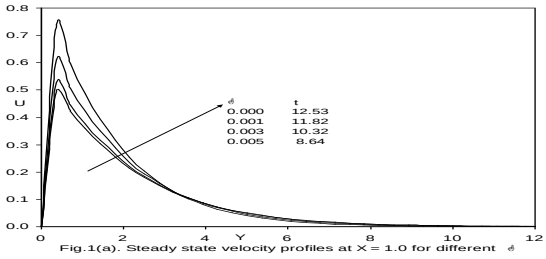


Fig.1(a). Steady state velocity profiles at X = 1.0 for different ϕ

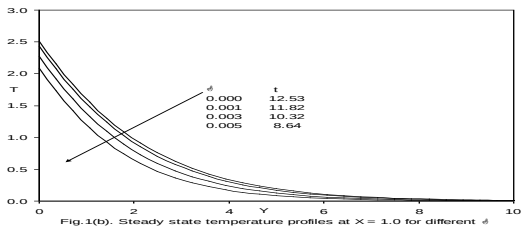


Fig.1(b). Steady state temperature profiles at X = 1.0 for different ϕ

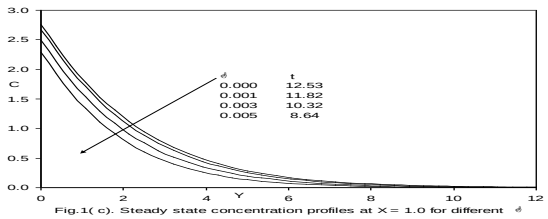


Fig.1(c). Steady state concentration profiles at X = 1.0 for different φ

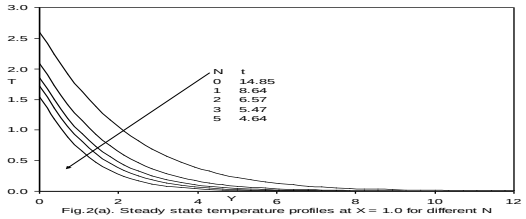


Fig.2(a). Steady state temperature profiles at X = 1.0 for different N

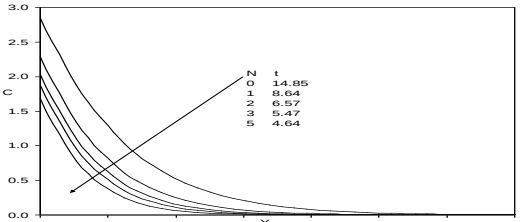


Fig.2(b). Steady state concentration profiles at X = 1.0 for different N

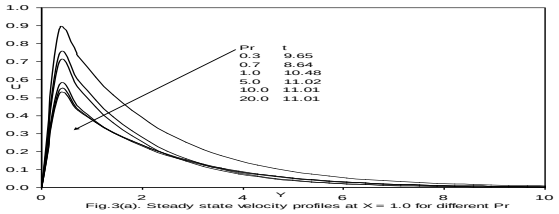


Fig.3(a). Steady state velocity profiles at X = 1.0 for different Pr

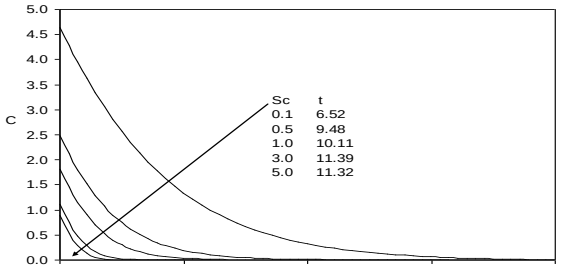


Fig.4.Steady state concentration profiles at X = 1.0 for different Sc

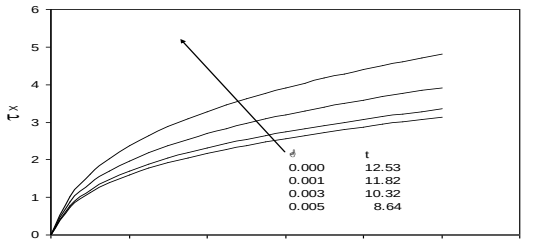


Fig.5(a).Effect of φ on local skin friction

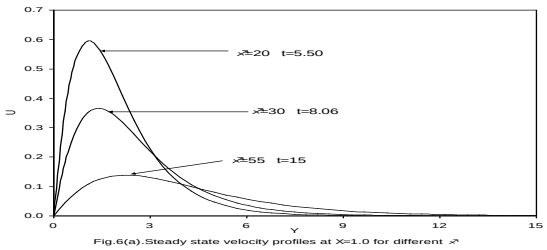


Fig.6(a).Steady state velocity profiles at X=1.0 for different x

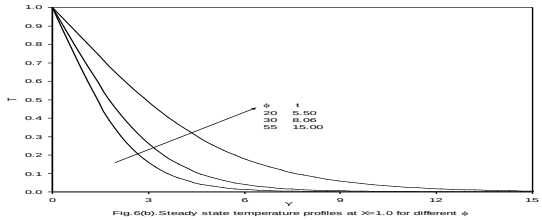


Fig.6(b).Steady state temperature profiles at X=1.0 for different φ

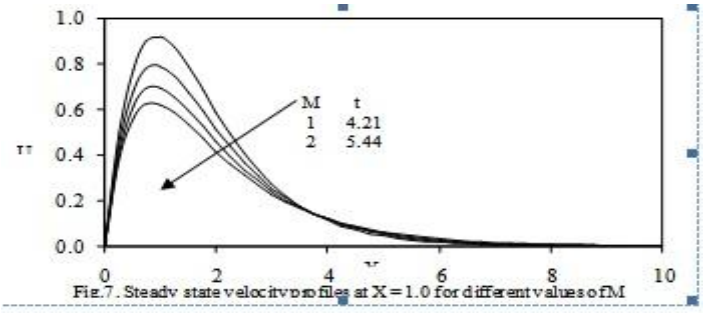


Fig.7. Steady state velocity profiles at X = 1.0 for different values of M

In figures 1a to 1c,

we have presented the variation of streamwise velocity (U), temperature function (T) and concentration (C) versus spanwise coordinate (Y) with collective effects of viscoelasticity (Γ) and time (t), close to the leading edge (i.e. cone apex) at X = 1.0. An increase in Γ from 0 to 0.001, 0.003 and the maximum value of 0.005, as depicted in figure 1a, clearly enhances the streamwise velocity, U which ascends sharply and peaks in close vicinity to the cone surface (Y = 0). In figure 1b increasing viscoelasticity Γ is seen to decrease temperature throughout the boundary layer. The graphs show therefore that increasing viscoelasticity cools the flow. With progression of time, however the temperature, T is consistently enhanced i.e. the fluid is heated as time progresses. A similar response is observed for the concentration field, C, in figure 1c. Increasing viscoelasticity again reduces concentration, showing that species diffuses more effectively in Newtonian fluids (Γ=0) than in strongly viscoelastic fluids. Once again with greater elapse in time the concentration values are reduced throughout the boundary layer regime ($0 < Y < 10$).

In figures 2a and 2b, the distributions of streamwise temperature function (T) and concentration (C) versus spanwise coordinate (Y) for various buoyancy ratio parameters (N) and time (t), close to the leading edge (i.e. cone apex) at X = 1.0, are shown. An increase in N from 0 through 1, 2, 3, 4 to 5 clearly shows that increasing N decrease both temperature (T) and species concentration (C) throughout the boundary layer transverse to the cone surface and that an increase in time has the opposite effect.

Figure 3a illustrate the spanwise spatial response of the velocity U to Prandtl number, Pr and time, t. Increasing Pr from 0.3 through 0.7, 1.0, 5.0, 10.0 to 20.0, clearly reduces strongly streamwise velocity, U (figure 3a) both in the near-wall regime and the far-field regime of the boundary layer. An increase in time, t, also serves to strongly retard the flow.

Figure 4 shows that increase in Sc from 0.1 (low weight diffusing gas species) through 0.5 (oxygen diffusing) to 1.0 (denser hydrocarbon derivatives as the diffusing species), 3.0 and 5.0, strongly *suppresses* concentration levels in the boundary layer regime. All profiles decay monotonically from the cone surface (wall) to the free stream. With a decrease in molecular diffusivity (rise in Sc) concentration boundary layer thickness is therefore decreased.

In **figure 5a** the variation of dimensionless local skin friction (surface shear stress), τ_x, versus streamwise coordinate (X) for various viscoelasticity parameters (Γ) and time (t) are illustrated. Shear stress is clearly enhanced with increasing viscoelasticity.

Figures 6a and 6b

present the spanwise spatial distributions of the U, T variables to semi-apex cone angle, ϕ and time, t. With wider cone angles i.e. a rise from 20 through 30 to 55 degrees (cone apex angle = 2ϕ =110 degrees), the velocity profiles are considerably reduced close to the cone surface i.e. the flow is strongly decelerated near the wall. With further transverse locations from the wall into the boundary layer, however this trend is reversed and streamwise velocity, U, is marginally greater than for smaller cone angles. With grater elapse of time from t = 5.0 through 8.06 to 15, the flow is also found to decelerate. Conversely in figure 6b, an increasingly wider cone apex, induces greater temperatures continuously throughout the boundary layer i.e. heats the boundary layer regime. With progression of time, T values are also increased.
Finally in **Figures 7** the influence of magnetic parameter (M) versus spanwise spatial distributions of velocity U are depicted. Application of magnetic field normal to the flow of an electrically conducting fluid gives rise to a resistive force that acts in the direction opposite to that of the flow. This force is called the Lorentz force. This resistive force tends to slow down the motion of the fluid along the cone and causes an increase in its temperature and a decrease in velocity as M increases. An increase in M from 1 though 2, 3, 4 clearly reduces streamwise velocity U both in the near-wall regime and far-field regime of the boundary layer.

7. Conclusions

A two-dimensional unsteady laminar incompressible boundary layer model has been presented for the external flow, heat and mass transfer in viscoelastic buoyancy driven convection regime past a vertical stationary cone. The Walters-B viscoelastic model has been employed which is valid for short memory polymeric fluids. The dimensionless conservation equations have been solved with the well-tested implicit crank-Nicolson finite difference numerical method. Present results are compared with the available results from the open literature and found to be in very good agreement. The following conclusions are drawn.
Increasing viscoelasticity accelerates the streamwise velocity and enhances shear stress (Local skin friction). Local Nusselt number and local Sherwood number but reduce temperature and concentration values transverse to the cone surface (i.e. inside the boundary layer regime).

- An increase in cone apex angle is found to strongly decelerate the flow near the cone surface but to increase temperature in the boundary layer regime.
- The flow is considerably accelerated close to the wall with a positive increase in buoyancy ratio parameter, corresponding to the case where both thermal and species buoyancy forces assist each other in the regime. The steady state response for the flow is also studied in detail.
- An increase in Schmidt number is observed to significantly decrease concentration.
- The time taken to reach steady state increases with increasing Prandtl number.
- An increase in Prandtl number is observed to reduce both velocity, temperature but increase concentration. The Momentum, Thermal boundary layers becomes thin when Pr is increased.

References

[1] Bapuji Pullepu, Ekambavanan, K., Chamkha, A. J., Unsteady laminar free convection from a vertical cone with uniform surface heat flux, Nonlinear Analysis: Modelling and Control. 13 (2008) 47-60.

[2] Carnahan, B., Luther, H. A., Wilkes, J.O., Applied Numerical Methods, John Wiley and Sons, New York (1969).

[3] Ganesan, P., Loganathan, P., Unsteady natural convection flow past a moving vertical cylinder with heat and mass transfer, Heat Mass Transf., 37 (2001) 59-65.

[4] Ganesan, P., Rani, H. P., Unsteady free convection MHD flow past a vertical cylinder with heat and mass transfer, Int. J. Therm. Sci., 39 (2000) 265-272.

[5] Hossain, M. A., Paul, S. C., Mandal, A. C., Natural convection flow along a vertical circular cone with uniform surface temperature and surface heat flux in a thermally stratified medium, International Journal of Numerical Methods for Heat and Fluid Flow, 12 (2002) 290-305.

[6] Lin, F. N., Laminar convection from a vertical cone with uniform surface heat flux, Letters in Heat and Mass Transfer, 3 (1976) 49-58.

[7] Muthucumaraswamy, R., Ganesan, P., Unsteady flow past an impulsively started vertical plate with heat and mass transfer, Heat Mass Transf., 34 (1998) 187-193.

[8] Pop, I., Watanabe, T., Free convection with uniform suction or injection from a vertical cone for constant wall heat flux, Int. Comm. Heat Mass Transfer, 19 (1992) 275-283.

[9] Prasad, V. R., Bhaskar Reddy, N., Muthucumaraswamy, R., Radiation and mass transfer effects on two dimensional flow past an impulsively started infinite vertical plate, Int. J. Thermal Sciences., 46 (2007) 1251-1258.

[10] Rajagopal, K., Veena, P. H., Pravin, V. K., Nonsimilar Solutions for heat and mass transfer flow in an electrically conducting viscoelastic fluid over a stretching sheet saturated in a porous medium with suction/blowing, J. Porous Media, 11 (2008) 219–230.

[11] Raptis, A. A., Takhar, H. S., Heat transfer from flow of an elastico-viscous fluid, Int. Comm. Heat and Mass Transfer, 16 (1989) 193–197.

[12] Soundalgekar, V. M., Puri, M., On fluctuating flow of an elastico-viscous fluid past an infinite plate with variable suction, J. Fluid Mechanics, 35 (1969) 561–573.

[13] Vajravelu K & Nayfeh L, Hydromagnetic convection at a cone and a wedge, Int Commun Heat Mass Transfer, 19 (1992) 701–710.

[14] Walters, K., Non-Newtonian effects in some elastico-viscous liquids whose behaviour at small rates of shear is characterized by a general linear equation of state, Quart. J. Mech. Applied. Math., 15 (1962) 63–76.

[15] Yih, K. A., Coupled heat and mass transfer by free convection over a truncated cone in porous media: VWT/VWC or VHF/VMF, Acta Mechanica, 137 (1999) 83-97.

A Value of E-Service in Local Government: A Fuzzy Approach Evaluation

Zuleaizal Sidek[1], Noor Hasimah Ibrahim Teo[2]
[1] Centre of Integrated Information System Universiti Teknologi MARA Selangor, Malaysia
[2] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Selangor, Malaysia

Abstract:
Usability and utility is a necessary condition for survival on the web. It determines the effectiveness of particular website in delivering its expected use. This paper reports the results of an evaluative study of the Malaysian e-Government portal (MyEG) from the usability and utility perspectives, using a common set of performance metrics from user participation. To deal with the uncertainty and vague expressions that surround the answers options, fuzzy approach will be applied. The result is categories based on demographic study and five main criterion of usability evaluation [9]. The results shows that MyEG have an average level of usability. Therefore, enhancements need to be done on the interface of the MyEG which at the same time would increase utility as well.

Keywords: e-service, e-Government, usability, utility, fuzzy logic, rule based approach, Evaluation.

1. Introduction.

The advancement of technology provides opportunities for government to address problems in many conventional services. E-Government has taken initiative to embraces the use of modern ICT to ensure e-Government delivers for all citizens. Malaysian Electronic Government (MYEG) initiatives provide information about public services and enable the citizen to conduct online government transactional services for Malaysian Community.The issue that becomes a major concern pertinent to this e-service portal is how usable is this portal in delivering services to the citizen. User tend to leave the portal if the information dos not answer users key questions, fails to clearly state what the portal offers, time consuming and etc. This paper provides a report on the usability evaluation of MyEG portal which address on the problem mentioned. as a medium for government to enhance their services.The major objectives for the existence of MyEG is to ease public to access services provide by the government. As a mediator between government service and public, effectiveness of the portal is an important issue to consider. These report gathered feedback by end-user which are Malaysian community. Future enhancement can be arranged for the benefit of the citizen and the nation.

2. E-Service In Public Sector: An Overview.

E-services are a component essentially provides the electronic link between the Government and citizen or businesses[2]. There are criteria have been defined[3] to identify usable e-services such as perceivable, comprehended, clearly expressed main topic, provide feedback, support error handling etc. Usable e-service on government portal is essential as it will assist reduce the effort of government in providing their services to its citizen.
Therefore, Malaysian E-Government has taken this opportunity to provide e-services by outsourcing the services to My E.G. service Berhad as a concessionaire. MyEG enable Malaysian to interact with numerous agencies within Federal, State and the local Government. The service provided ranging from information searches to license application.

3. Utility And Usability Evaluation Of E-Services.

Utility measure the functionality of the website[4]. It's simply means how practical and useful it is in term of the fitness of use. It can be measure by based on the number of user that has used the website. Utility actually depend on the usability of the usability of the website interface. If the user fined the website easier to use, it will increase the possibility of user to use the service provided by the website. Usability is a necessary condition for evaluating web portal to ensure its usefulness[1]. The word "usability" refers to quality attribute that assesses how easy user interfaces are to use[1][4]. The evaluation will enable the web portal in achieving the utility function.There is no clear consensus how to measure usability that deal with the uncertainty and vague expressions[3]. However, the uncertainty and vagueness in usability evaluation is presented in [5][6]. Other related works presented for evaluating usability are presented using statistical approach[1][4][7][8].According to [9], there are five criterions for evaluation usability on web portal which are attractiveness, control efficiency, helpfulness and learnability. In contrast with [4], the criterions are learnability, efficiency, memorability, satisfaction and errors. The criterion defined by [4] is use to evaluate regular user for particular portal as it has memorability criteria to be evaluated.

Thus, this paper proposes usability evaluation based on criterion presented by [9] as it is suitable for non-regular user of MyEG portal.

4. A Malaysia Local Government Case Study.

The usability attributes contribute to quality of use includes the style and properties of the user interface, the dialogues structure and the functionality [1]. The evaluation uses a set of criteria to determine whether the design of MyEG portal is successful in achieving usability. The criteria are measured by conducting random online survey to government servants. The collection of gathered survey respond is further analyzed using fuzzy linguistic approach.

Questionnaire has been designed that enable usability evaluation on web portal. It is to enable analysis for all five main quality criteria mention earlier. The questionnaire is divided into two parts which are demographic and web portal usability type of question [9][10].Demographic questions such as gender, age group, MyEG service commonly use and frequency of internet access will be analyzed and use as a categories of overall result in this research. Table 1 shows the sample questions for demographic.

Table 1. Demographic Sample Questions

Category	Questions
Gender	□ Male □ Female
Age	□ 18 – 30 □ 31 – 45 □ 46 - 65
Frequency of internet access	□ Daily □ Less than 4 times a week □ When necessary

Second part of the questionnaire focused on how far the usability involved. It is determine by calculating usability and score. Questions for each criterion are adopted from the combination of [6][9]. This score are obtained by using fuzzy approach. Table 2 shows sample questions design for this part.

Table 2. Sample Questions (source:[6] [9])

Evaluation Criterion	Sample Questions
Attractiveness	This web site is presented in an attractive way. You can learn a lot on this web site.
Control	Going from one part to another is easy on this web site. I feel in control when I'm using this web site.
Efficiency	You can find what you want on this web site right away. This web site works exactly how I would expect it to work.
Helpfulness	This web site has not been designed to suit its users. All the parts of this web site are clearly labeled.
Learnability	All the material is written in a way that is easy to understand. It will be easy to forget how to use this web site.

5. Usability Evaluation Using Fuzzy Approach.

The analysis is conducted using fuzzy approach[11] due to uncertain and vague terminologies in questionnaire answers' options[12][13]. The result is obtained with the aid of Matlab tool. Following are the steps undertake to obtain usability score.

5.1. Fuzzification.

Answers' options are in the form of linguistic variables which range from strongly disagree to strongly agree. Empirical scale for each of these variables is defined. The value obtain depend on user choice. It is then fuzzified into fuzzy value using sigmoid membership function. Figure 1, Figure 2 and Figure 3 show the membership function representation in Matlab for user answer option, number of occurrence of each answer's option and the usability score respectively.

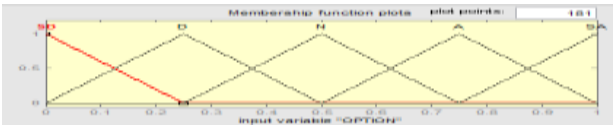


Figure 1. Memberships function for Answer Option.

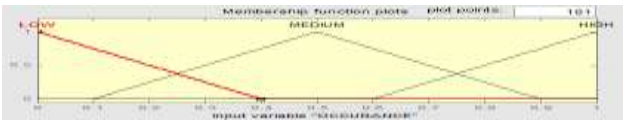


Figure 2. Membership function for user answer number of occurrence

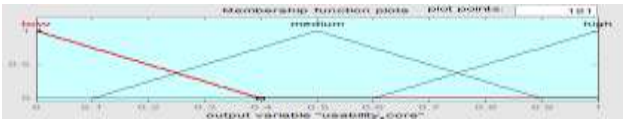


Figure 3. Membership function for usability score.

5.2. Rule Based Evaluation.

Rule base is equipped with the expert knowledge defined for every possible criterion. This rule base is store and inferred using fuzzy value in previous step. Figure 4 shows the rules defines for evaluating usability.

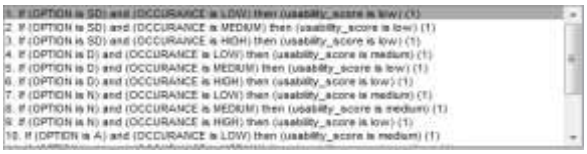


Figure 4. Rules defined for Usability Score

5.3. Aggregation.

Aggregation will calculate score for each of the rules inferred in the previous step. The result is based on the rules define in the knowledge base. Figure 5 shows the aggregation calculation. Figure 5 shows that only rules 7, 8 and 9 were inferred for 'Option'; rules 2, 5, 8, and 11 for 'Occurrence'. However, only rules 8 have usability score because both 'option' and 'occurrence' are inferred. If more than one rules are inferred, maximum value of usability score will be taken as final result in aggregation steps.


Figure 5: Rules aggregation for obtaining final fuzzy value.

Defuzzification: The value is defuzzified using Centroid of Gravity (COG) to form a single real value which is usability score. COG takes the average value obtain from aggregation steps. The score range from 0 – 100, which indicate usability degree; low(0-39); medium(40-75);high(76-100). Figure 6 shows the final result of the process which is obtaining crisp value. The crisp value in this case represents usability score.

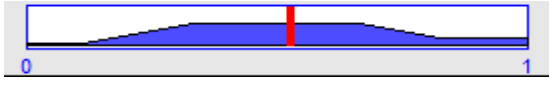


Figure 6: Crisp value obtains from defuzzification using COG.

6. Analysis of Results.

A typical questionnaire consists of a number of different type of questions are distributed and analyzed. The discussion on the analysis is divided into five main quality criterions to evaluate usability which are learnability; efficiency; attractiveness; helpfulness; control. This section will discuss the result in two-fold.

6.1. Result of Utility Evaluation.

Utility evaluation is conducted by using statistical approach. Data are collected from respondents for this survey by using online survey tool. Only 60 percent of the respondents have used MyEG and the rest are unaware and prefer to use traditional method. 75 percent from the respondents who used MyEG are female. In addition, the most popular services that use by the community are insurance renewal, road tax renewal and PDRM Summons alert.

6.2. Result of Usability Evaluation.

Usability evaluation is evaluated from the 60 percent of the respondents that used MyEG. Results for usability score is presented by each of the user and by criterion. Table 4 shows usability score by and Table 5 shows criterion the results of usability score for each user.

Table 4.Usability score by criterion

Criterion	Usability Score
Attractiveness	53.0
Control	53.0
Efficiency	53.0
Helpfulness	53.0
Learnability	53.0

Table 5.Usability score for each user.

USER	Usability Score
User 1	53.3
User 2	55.6
User 3	75.1
User 4	13.5
User 5	63.3
User 6	63.3
User 7	50.0
User 8	48.1
User 9	53.0
User 10	42.7
User 11	75.1
User 12	55.6

The result in Table 4 shows that all criterion having the same usability score of MyEG which are 53.0. The results in Table 5 shows most of the user evaluate usability under medium level. This has shown that MyEG have medium range of usability. Therefore, more improvement and adjustment need to be done on these criterions.

The result in Table 6 and Table 7 shows usability evaluation by demographic studies which are gender, age and Frequency of internet access.

Table 6. Usability Score by Gender

Gender	Usability Score
Male	53.3
Female	53.0

Table 6 shows that both male and female evaluating usability also on average score. Therefore, it shows that MyEG having an average score of usability. More improvement and enhancement need to be perform in order to maximize the usability and increase the probability of used by Malaysia community.

7. Conclusion

The Malaysian Government introduced e-Government under Multimedia Super Corridor (MSC). One of the flagships under eGovernment initiative is e-Services. The main idea of this paper is to reports the usability of MyEG web portal from user perspectives. The evaluation based on the attractiveness, control, efficiency, helpfulness and learnability. The evaluation is done through the questionnaire disseminated among government servants. Future work will is to evaluate the user interfaces in detail, utility and its reliability.

Acknowledgement

This research has been financially supported by Universiti Teknologi MARA, Malaysia.

References

[1] N. Bevan, "Usability is Quality of Use," In proceedings of the 6th International Conference on Human Computer Interaction, Yokohama, July 1995. Anzai & Ogawa (eds), Elsevier.

[2] A. Ancarini,"Towards quality e-service in public sector: The evolution of web sites in the local public service sector," Managing service Quality, Vol. 15 (1), pp 6 - 23

[3] A. Rstlinger and S. Cronholm, "Design Criteria for Public E-Services," ECIS 2008 Proceedings. Paper 28

[4] J. Neilsen, "Usability Engineering," 13th ed. San Francisco : Morgan Kaufmann Publishers Inc., 1994. P.362. ISBN 0-125184-06-9.

[5] M. Hub and M. Zatloukal, "Model of Usability Evaluation of Web Portals Based on the Fuzzy Logic," WSEAS Transaction on Information Science and Applications. 2010. Vol 7, Issue 4, pp.522-531

[6] M. Hub and M. Zatloukal, "Methodology of Fuzzy Uability Evaluation of Information Systems in Public Administration," WSEAS T ransactions on Information Science & Applications. 2008, Vol 5, Issue 11, pp. 1573-1583.

[7] J.Withrow, T. Brinck and A. Speredelozzi, "Comparative Usability Evaluation for an e-Government Portal," Diamond Bullet Design Report #U1-00-2, Ann Arbor, MI. Dec 2000

[8] Xin C. Wang, B. Yadamsuren, A. Paul, D. Adkins, G. Laur, A. Tawfik and S. Erdelez,"Iterative Usability Evaluation for an Online Educational Web Portal," International Journal of Multimedia Data Engineering and Management (IJMDEM, 2010. Vol 1, issue 4, pp. 19

[9] J. Kirakowski, N. Claridge and R. Whitehand, "Human Centered Measures of Success in web Site Design," In Proceeding of 4th Conference on the Human Factors on the web, 1998.

[10] M. Hub and M. Zatloukal, "Methodology of Fuzzy Usability Evaluation of Information Systems in Public Administration," WSEAS Transactions on Information Science & Applications. 2008, Vol 5, Issue 11, pp. 1573-1583.

[11] L. Zadeh, " Fuzzy Sets," Information and Control. 1965, Issue 8, pp. 333-353.

[12] M. Zatloukal, "Usability evaluation of information Systems in Public Administration based on the fuzzy approach," Department of System Engineering and Informatics, University of Pardubice, 2009. P. 100, Master Thesis.

[13] Siler, William and Buckley, James J. "Fuzzy Expert System and Fuzzy Reasoning." 1st ed. Hoboken : John Wiley & Sons, Inc., 2005. P. 424. ISBN 0-471-38859-9 [9] W. Cheney and D. Kincaid. Numerical Mathematics and Computing.Brook-s/Cole, 1985.

Vanet Based Traffic Management System Development And Testing Using Aodv Routing Protocol.

PATIL V.P.
Smt. Indira Gandhi college of Engineering, New Mumbai, INDIA

Abstract:
Vehicular Ad-Hoc network (VANET) is a type of Mobile Ad-Hoc (MANET) network where the nodes are constrained to move along the road. Vehicles in VANET are equipped with a radio device to communicate with each other and also with the road side units (base stations). Vehicular networks aims to make the driving experience safer, efficient and enjoyable. Vehicle traffic congestion is reflected as delays while traveling. Traffic congestion has a number of negative effects and is a major problem in today's society. Several techniques have been deployed to deal with this problem. In this paper, it has been proposed an innovative approach to deal with the problem of traffic congestion using the characteristics of vehicular ad-hoc networks (VANET). The system is developed and tested using AODV protocol of ad hoc mobile network to deal with the problem of vehicle traffic congestion in vehicular networks. The performance is measured in terms of no. of packets broadcasted, percentage of packets delivered, and percentage of traffic diverted and overhead to manage the problem of data traffic congestion in computer networks. Using simulations the applicability of the algorithm in the domain of vehicle traffic congestion in a VANET is demonstrated.

Keywords: VANET, Aodv, Traffic management, Java, Testing, proactive protocol, traffic congestion.

I. Introduction

Vehicular ad hoc networks (VANETs) [1, 2] are attracting the interest of a great number of academicians and industrials. One of the most interesting features is the possibility to use a spontaneous and inexpensive wireless ad hoc network between vehicles to exchange useful information such as warning the drivers of an accident or a danger. One key limitation of many existing tools that integrate vehicular traffic and network simulation is the lack of dynamic interactions between the two domains. Thus transportation simulators would use pre-computed and aggregate network level delay and packet loss computations whereas network simulators would use pre-scripted mobility data. The shortcoming of these approaches is the lack of dynamic interaction between an event (e.g. accident) as it unfolds in the transportation simulator and its dissemination to the vehicles using the network as embedded within the vehicles in its vicinity and the feedback to the transportation simulator the change in velocities and positions of the vehicles as they react in real-time to the information conveyed to them by the communication network. The lack of the above type of dynamic interaction between the transportation and the network simulator reduces the level of realism that can be achieved for key ITS applications like active safety and traveler information systems which in reality influence the vehicles' movements significantly.
Incorporating accurate network simulations into transportation simulators is proven to have significant impact on the predicted performance of applications under realistic operating conditions of the VANET network. The penetration ratio required for Dynamic Route Planning to achieve sufficient performance gain in scenarios of high channel bandwidth is determined by the relative roadway capacity on alternative routes. In these cases, tipping point increases as the relative capacity enlarges. With limited channel bandwidth, higher penetration ratio beyond the tipping point greatly impairs the application performance due to channel saturation [3, 4].In this paper Next section reviews about literature of VANET and scope of traffic management using vehicular network communication is given in the section III, in section IV the implementation of the proposed system is done, section V describes about result analysis and testing and at the end section VI concludes bout the findings of the implemented system.

II. Review of Literature

2. 1 About VANET

Vehicular Ad hoc Networks (VANET) is the subclass of Mobile Ad Hoc Networks (MANETs). VANET is one of the influencing areas for the improvement of Intelligent Transportation System (ITS) in order to provide safety and comfort to the road users. VANET assists vehicle drivers to communicate and to coordinate among themselves in order to avoid any critical situation through Vehicle to Vehicle communication e.g. road side accidents, traffic jams, speed control, free passage of emergency vehicles and unseen obstacles etc. Besides safety applications VANET also provide comfort applications to the road users. [4, 5, 6] Each node within VANET act as both, the participant and router of the network as the nodes

communicates through other intermediate node that lies within their own transmission range. VANET are self organizing network. It does not rely on any fixed network infrastructure. Although some fixed nodes act as the roadside units to facilitate the vehicular networks for serving geographical data or a gateway to internet etc. Higher node mobility, speed and rapid pattern movement are the main characteristics of VANET. This also causes rapid changes in network topology. VANET is a special type of MANET, in which vehicles act as nodes. Vehicular ad hoc networks present a promising way to build up a decentralized parking guidance system. Designing such an application can be decomposed into major issues: (1) which information on a parking place needs to be known by the vehicles and thus has to be distributed in the vehicular ad hoc network? And finally, (2) how can this information be used to maximize the benefit for the driver? [9, 10, 11].

2.2 Routing protocols in VANET

A routing protocol governs the way that two communication entities exchange information; it includes the procedure in establishing a route, decision in forwarding, and action in maintaining the route or recovering from routing failure. This section describes recent unicast routing protocols proposed in the literature where a single data packet is transported to the destination node without any duplication due to the overhead concern. Some of these routing protocols have been introduced in MANETs but have been used for comparison purposes or adapted to suit VANETs' unique characteristics. Because of the plethora of MANET routing protocols and surveys written on them, we will only restrict our attention to MANET routing protocols used in the VANET context. VANET routing protocols can be classified as topology-based and geographic (position-based) in VANET.[2-5]

2.2.1 Topology-based Routing Protocols

These routing protocols use links' information that exists in the network to perform packet forwarding. They can further be divided into proactive (table-driven) and reactive (on-demand) routing.

2.2.2 Proactive (table-driven) Routing protocols

Proactive routing carries the distinct feature: the routing information such as the next forwarding hop is maintained in the background regardless of communication requests. Control packets are constantly broadcast and flooded among nodes to maintain the paths or the link states between any pair of nodes even though some of paths are never used. A table is then constructed within a node such that each entry in the table indicates the next hop node toward a certain destination. The advantage of the proactive routing protocols is that there is no route discovery since route to the destination is maintained in the background and is always available upon lookup. Despite its good property of providing low latency for real-time applications, the maintenance of unused paths occupies a significant part of the available bandwidth, especially in highly mobile VANETs.

2.3 Ad hoc On demand Distance Vector routing protocol (AODV)

Unlike traditional wire line networks, ad hoc networks don't rely on existing infrastructure to support communication. Each mobile node acts as an end node when it is the source or destination of a communication and forwards packets for other nodes when it is an intermediate node of the route. This makes multi hop communication possible. Ad hoc networks are easier to deploy than wire line networks and have found many applications, such as in rescue, battlefields, meeting rooms etc., where either a wire line network is unavailable or deploying a wire line network is inconvenient. [6, 7]

AODV: AODV has the merits of DSR and DSDV protocol. DSDV maintains routes to all destinations with periodical route information flooding and uses sequence numbers to avoid loops. AODV inherits the sequence numbers of DSDV and minimizes the amount of route information flooding by creating routes on-demand, and improves the routing scalability and efficiency of DSR, which carries the source route in the data packet. In AODV protocol, to find a route to the destination, the source broadcasts a route request packet (RREQ). Its neighbors relay the RREQ to their neighbors until the RREQ reaches the destination or an intermediate node that has fresh route information. Then the destination or this intermediate node will send a route reply packet (RREP) to the source node along the path from which the first copy of the RREQ is received. AODV uses sequence numbers to determine whether route information is fresh enough and to ensure that the routes are loop free. In AODV protocol, the route is built on demand and is not updated until the route breaks or times out. The route can't adapt to topology changes and breaks frequently in the case of high mobility. AODV uses local repair to restrict the route discovery zone so as to reduce overhead. If the source moves away and causes the route to break, it can re-initiate route discovery to the destination. In case an intermediate link breaks and the broken link is near the destination, the upstream node of this broken link may choose to repair the route. It initiates a route request with a fresh destination sequence and the RREQ will be flooded in a zone with a radius no less than the original hop count between this node and the destination. If the upstream node of the

broken link decides not to repair the broken route, it will send a route error packet (RERR) upwards to the source node. The source node will re-initiate route discovery in an even bigger zone than that of the local repair if the route is still needed.

2.4. A Robust AODV Protocol

In AODV protocol, usually the RREQ that first arrives at the destination determines the route. The route may contain neighboring nodes with long distances, which leads to low stability. And the node's mobility makes the route more susceptible to breaking. Our purpose is to make AODV adaptive to mobility and more robust. The Robust AODV protocol contains three parts: active route discovery, backup route building, and route maintenance. [6, 7].

A. Active Route Discovery.

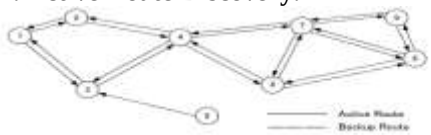


Fig. 1 Active route and backup route

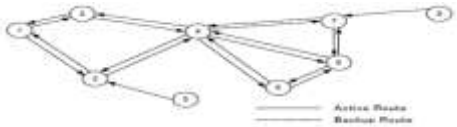


Fig.2 Switch to better route with lower hop count

In our Robust AODV protocol, the route discovery process is almost the same as in the original AODV except that all routes are built as backup routes. Only when data packets pass through will the route flag be changed to ACTIVE. From then on, the active route information is broadcasted to 1-hop neighbors, so that backup routes can be built. Here only the active route is maintained. If no packets pass through within a certain time, the active route will time out and be switched to a backup route then it's not maintained.

Fig. 1 is an example: the active route between node 1 and node 8 is built as 1-3-4-7-8. Node 1 has two routes for node 8, the route whose next hop is node 3 is an active route and the route whose next hop is node 2 is a backup route.

B. Backup Route Building

In Robust AODV protocol, the AODV hello message is substituted by the ad hoc route update message, which is broadcasted only to 1-hop neighbors by the node who is on the active route or who can hear the active route information. For each destination, at most one route is added to the route update message. The route can be either an active route when the node is on the active route or the highest priority backup route when the node can hear the active route information. The number of routes in the route update message is no more than that of the total active routes in the ad hoc network.

The broadcasted route update message contains all necessary route information, especially the destination sequence number and the next hop, which are important for avoiding loops. Multiple route update messages may be received from different nodes. The route with a fresh sequence number and lower hop count has the highest priority and is preferred, which is the same as in the original AODV. For incoming new route information, the sequence number is checked first. If this route has an old sequence number, it is discarded. If it has a fresh sequence number with bigger hop count, it must be stable before it is switched to become the new active route. Next, the hop count is checked. If the received route's hop count is lower than the current hop count +1, it is added to the route table as a backup route.

C. Route Maintenance

Route entries are updated when a route update packet is received. If a better route is available after updating, this route will be switched to become the new active route. Routes are monitored and maintained when they are used to forward data packets. When a link break is detected, if a backup route exists, the route switch is performed and the new route is checked; if no backup route exists, an RERR is sent toward the source node, just like in the original AODV.

1) Update Route Entry: Each time a route update message is received from an old sender, the corresponding route's next hop's next hop is updated. The backup route's lifetime also gets updated. The backup route will be deleted if it times out.

2) Switch to Better Route.: When the current active route is less preferred, the highest priority backup route is switched to become the new active route. And the old active route will be a backup route. In Fig. 1, node 8 moves near to node 4 and the shorter route 8-4-3-1 is switched to become the new active route, as is shown in Fig. 2. In this way, routes in Robust AODV adapt to topology variations. In the original AODV, when node 8 moves towards node 4, the route is not updated; and as node 8 continues moving, the link between node 8 and node 7 will break, and local repair may be conducted, which actually can be avoided.

Iii. Vanet For Traffic Management System

The current traffic management system in India is not complete and efficient for the public. There are billions of vehicles running on the roads. Therefore, it is highly desired to have a reliable and cost effective way to track the traffic or congestion on roads and thus choose the appropriate road that is congestion free. If there is base station in the square who got idea about congestion, then it transmit the signal containing information about the congestion, car receives the signal through trans

receiver and thus choose another road to go further. The car can transmit the signal within its range to alert incoming car about accident. With the advance and wide deployment of wireless communication technologies, many major car manufactories and telecommunication industries gear up to equip each car with the On Board Unit (OBU) communication device, which allows different cars to communicate with each other. Roadside infrastructure i.e. Roadside Units (RSUs), in order to improve not only road safety but also better driving experience. Therefore, it becomes possible to track the congestion on roads and then guide drivers to the available congestion free road, through vehicular communications.

3.1 System implementation Requirement

3.1.1 Software specification:

1) Operating System – Windows 7or XP.;2) Front end – Java ;3) Back-end database – Java
4) NetBeans IDE 7.0.

3.1.2 NetBeans Platform

The NetBeans Platform is a reusable framework for simplifying the development of Java Swing desktop applications. The NetBeans IDE bundle for Java SE contains what is needed to start developing NetBeans plugins and NetBeans Platform based applications; no additional SDK is required.Applications can install modules dynamically. Any application can include the Update Center module to allow users of the application to download digitally-signed upgrades and new features directly into the running application. Reinstalling an upgrade or a new release does not force users to download the entire application again. NetBeans IDE is a free, open-source, cross-platform IDE with built-in-support for Java Programming Language.

NetBeans IDE: NetBeans IDE is an open-source integrated development environment. NetBeans IDE supports development of all Java application types (Java SE (including JavaFX), Java ME, web, EJB and mobile applications) out of the box. Among other features are an Ant-based project system, Maven support, refactoring, version control (supporting CVS, Subversion, Mercurial and Clear case).

Modularity: All the functions of the IDE are provided by modules. Each module provides a well defined function, such as support for the Java language, editing, or support for the CVS versioning system, and SVN. NetBeans contains all the modules needed for Java development in a single download, allowing the user to start working immediately. Modules also allow NetBeans to be extended. New features, such as support for other programming languages, can be added by installing additional modules. For instance, Sun Studio, Sun Java Studio Enterprise, and Sun Java Studio Creator from Sun Microsystems are all based on the NetBeans IDE.

3.2 Operational feasibility

Operational feasibility is a measure of how well a proposed system solves the problems, and takes advantage of the opportunities identified during scope definition and how it satisfies the requirements identified in the requirements analysis phase of system development. Fig.4 shows data flow diagram (DFD) for Client to Client Communication and Fig.5 shows DFD for Client to Base Station Communication. Fig 3 shows the flow chart of the system.

3.2 FLOW CHART

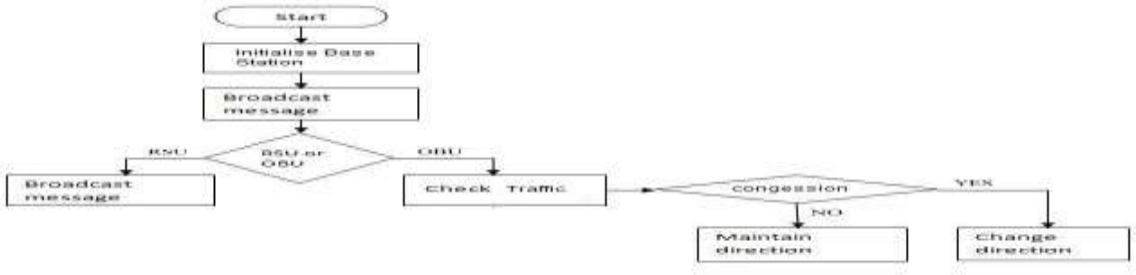


Fig.4: FLOW CHART

IV. Implementation of the Proposed System

4.1 Algorithm development

We have implemented AODV algorithm for the proposed system and its stages are as follows:
a)Active Route Discovery ;b)Backup Route Building ;c) Route Maintenance; i)Update Route Entry Switch to better route; ii) Switch route in the case of link break; iii) Route check; iv) Hop count restriction

4.2 Screen Snapshot & module description

The system is tested for its working under different situations and the working is satisfactorily observed as shown in the screen snapshots shown in fig 6 to 11.

User Interface*:* Fig.6 is first Main page of GUI of Our System shows the simulation scenario without Input panel.Fig.7 is first Main page of GUI of Our System shows the simulation scenario with input panel. Fig.8 is first Main page of GUI of Our System shows the simulation scenario with input parameters. As shown in Fig.9, In this scenario our system shows the accidental scenario and the start the broadcast of messages. In the scenario as shown in fig 10, our system shows Inter Vehicle Communication and broadcast of messages.Fig.11 shows our system shows inter vehicle communication and traffic diversion due to congestion of traffic.

As in Fig.12, The graphs show: Graph.1 Overhead on base station per second; Graph.2 Number of packets broadcasted; Graph.3 Percentage of packets delivered; Graph.4 Percentage of traffic diverted.

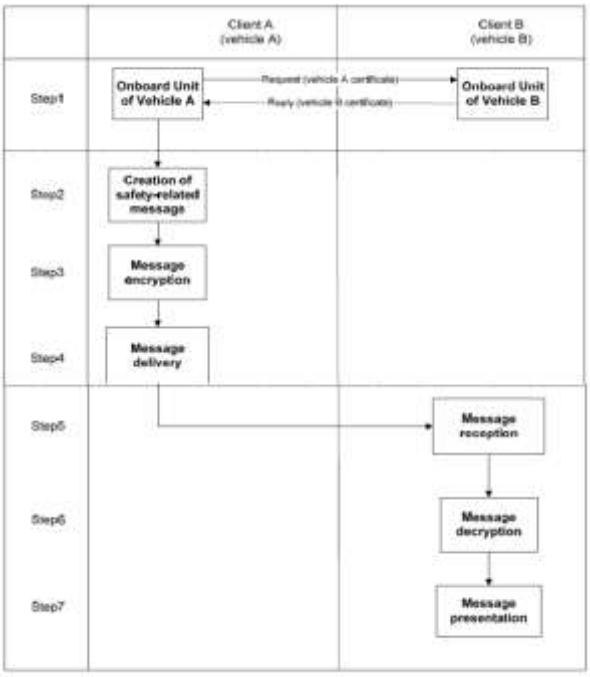


Fig.4: DFD for Client to Client Communication.

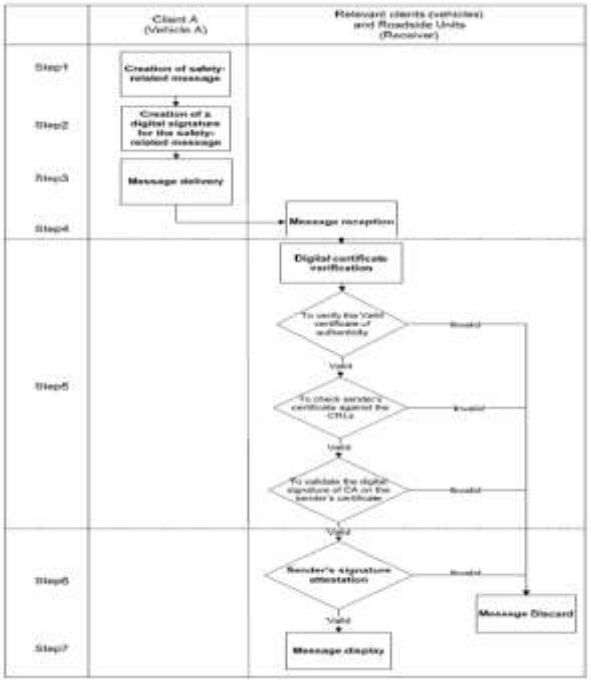


Fig.5: DFD for Client to Base Station Communication

Fig.6: Basic Design

Fig.7: Basic Design With Input Panel

Fig.8 Basic Design With Input Panel

Fig.9: Accident Scenario & Packet Broadcast

Fig.10: Inter Vehicle Communication

Fig.11: Traffic diversion

4.3 Output Graphs:

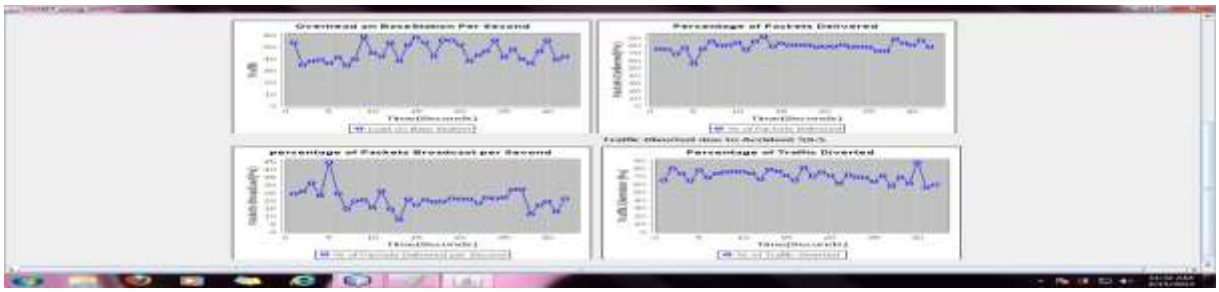


Fig.12: Output Graphs

V. Testing and Result Analysis

Testing is process of executing a program with the intent of finding an error. A good test case is one that has a high probability of finding an as yet undiscovered error. A successful test is one that uncovers previously undiscovered errors. The main objective of testing, thus, is to design tests that uncover different classes of errors, and to do so with minimum time & effort spent.

5.1. Testing methods:

Two classes of input are provided to the test process:

i) A software configuration that includes a Software Requirement Specification, Design Specification and Source code.
ii) A test configuration that includes a test plan and procedure, any testing tools to be used and testing cases with their expected results.

5.1.1 White Box Testing:

White box testing can be performed to validate whether code implementation follows intended design, to validate implemented security functionality, and to uncover exploitable vulnerabilities.

The general outline of the white box testing process is as follows:

i) Exercise all independent paths within the module at least once.
ii) Exercise all logical decisions for both true and false scenarios
iii) Execute all loops at their boundaries and within their operational loops
iv) Execute test cases and communicate results.

5.1.2 Black Box Testing

It is also known as functional testing. A software testing technique whereby the internal workings of the item being tested are not known by the tester. For example, in a black box test on software design the tester only knows the inputs and what the expected outcomes should be and not how the program arrives at those outputs.

Black Box testing attempts to find errors in the following categories:

i) Incorrect or missing functions.
ii) Interface errors.
iii) Errors in data structures or external database access.
iv) Performance errors
v) Initialization and termination errors.

5.2 Testing Vanet Based Traffic Management System.

Vanet Based Traffic Management System is tested using Black Box Testing. In this type the functional requirements of the system are tested. The system is given some input and the outputs of the system are studied to uncover the errors that might be present in the system.We test the our Project at two side i.e. Administrator Side, User Side

5.2.1 Testing VANET Based Traffic Management System at Administrator Side

Test Cases for Default Input:

i) Testing is done to check that when the user logs dosent provide any input.
The System takes default input for every variable parameter.

Test Cases for User Input:

i) The main case for the user input is that all fields are filled properly.
ii) To check whether each field is filled with valid input.

5.3 Running a Test

Once you have created a test or test suite, you can use the Run Test command to initiate execution of the test. Run Test commands are available on source nodes only. After you run a test, you can choose to rerun individual test methods executed during the test and displayed in the JUnit Test Results window.

a) To run tests for an entire project:

1. Select any node or file in the project you want to test in the Projects or Files window.
2. From the main menu, choose Run > Test *project_name* (Alt-F6).

The IDE executes all of the project's tests.

- If you want to run a subset of the project's tests or run the tests in a specific order, you can create test suites that specify
- the tests to run as part of that suite. After creating a test suite you run the suite in the same way you run a single test class.

b) To run a test for a single class:

1. Select the node of the class for which you want to run a test in the Projects or Files window.
2. From the main menu, choose Run > Run File > Test *classname* (Ctrl-F6).

- You can also run a class's test by selecting the test class node itself and choosing Run > Run File > Run
- testclassname (Shift-F6).

c) To run a single test method:

1. Run the test class or suite containing the test method.
2. In the JUnit Test Results window, right-click the test method and choose Run Again.

- To run a single test method the method must be listed in the JUnit Test Results window.

5.4 Working with JUnit Test Output

When you run a test, the IDE shows the test results in two tabs in the JUnit Test Results window:

- A summary of the passed and failed tests and the description of failed tests are displayed in the left pane of
- the window.
- The textual output from the JUnit tests themselves is displayed in the right pane of the window.
- The output from the Ant process that builds and runs the test is displayed in the Output window. Double-click any error to jump to the line in the code where the error occurred.

5.5 Result analysis:

It has been observed that almost all route request broadcasts reach the destination, only a few over long distances with middle car density failed. But the load on the network originating from the naive broadcast is tremendous. As a result it also leads to quickly growing delays and link failure. Several route replies do not come through because broadcasting is still going on. This is a critical problem, especially in city areas with high car density. It seems to be appropriate to replace the common broadcast system by another, more efficient, version. Another phenomenon that can often be observed is that a route breaks before the data packet can be successfully transmitted or even that the route reply does not find its way back. This is most critical for short transmission ranges and high mobility. During analysis of the trace files, we observed that such link breakages mostly come from the following situation: let us take the highway scenario and think of two cars driving in the same direction. They are driving 120 kilometers per hour and the distance between them is several kilometers (dozens of hops). If we are able to

establish a route between them where every intermediate node drives in the same direction, the route is more or less stable. But if only one intermediate node drives in the opposite direction we have a serious problem. Such a node covers about 35 meter per second or 70 meters per second relative to the cars driving in the opposite direction. This is about 20 to 50 percent of the simulated transmission ranges. Since we know from the simulations above that a transmission of a single data packet with route establishment over such a distance takes about a second, it is logically that this leads to a link breakage. If one wants to use AODV on highways, it is essential to extend it in a way that avoids such link breakages.

VI. Conclusion

The traffic management system using AODV protocol for VANET under java environment is simulated and tested satisfactorily. The testing was done for one of the event of accident at one place and due which it was observed that traffic automatically diverts to another alternative route. This was possible due to multi hop vehicle to vehicle communication and vehicle to road side base station communications. It has been observed that almost all route request broadcasts reach the destination, only a few over long distances with middle car density failed. But the load on the network originating from the naive broadcast is tremendous. As a result it also leads to quickly growing delays and link failure. Several route replies do not come through because broadcasting is still going on. This is a critical problem, especially in city areas with high car density. It seems to be appropriate to replace the common broadcast system by another, more efficient, version. Another phenomenon that can often be observed is that a route breaks before the data packet can be successfully transmitted or even that the route reply does not find its way back. This is most critical for short transmission ranges and high mobility. Since we know from the simulations and testing of this system that a transmission of a single data packet with route establishment over such a distance takes about a second, it is logically that this leads to a link breakage. If one wants to use AODV on highways, it is essential to extend it in a way that avoids such link breakages. In future the system can be implemented using network simulator NS-2 and using any other VANET protocol.

References

[1] Internet Engineering Task Force (IETF) Mobile Ad Hoc Networks (MANET) Working Group Charter, Chaired by Joseph Macker and Scott Corson, http://www.ietf.org/ html.Charters/manet -charter.html

[2] Charles E. Perkins, Pravin Bhagwat, Highly Dynamic Destination Sequenced Distance-Vector Routing (DSDV) for Mobile Computers, ACM SICOMM'94 Conference on Communications Architectures, Oct, 1994, pp234-244

[3] David B. Johnson, David A. Maltz, Yih-Chun Hu , The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks, draft-ietf-manet-dsr-09.txt, Apr.15, 2003

[4] Perkins, C.E.; Royer, E.M., Ad-hoc on-demand distance vector routing, Mobile Computing Systems and Applications, 1999. Proceedings, pp: 90 –100, 1999

[5] C. Perkins, E. Belding-Royer, S. Das**,** Ad hoc On-Demand Distance Vector (AODV) Routing, RFC3561, July, 2003

[6] Pearlman, M.R.; Haas, Z.J., Determining the optimal configuration for the zone routing protocol, Selected Areas in Communications, IEEE Journal on, Vol.17, Issue: 8, 1999, pp: 1395 - 1414

[7] Lee, S.-J.; Gerla, M.;, AODV-BR: backup routing in ad hoc networks, Wireless Communications and Networking Conference, pp:1311-1316, Vol.3, 2000

[8] D. K. Tokuda, M. S. Nakabayashi, M. M. Hamaguchi, and M. H. Tsutsui. Oki project. http://www.ece.osu. edu/hpcnl/okipublic/, 2004.

[9] IEEE Computer society LAN MAN Standards Committee, Wireless LAN Medium Access Protocol (MAC) and Physical Layer (PHY) Specification, IEEE Std 802.11-1997, IEEE, 1997

[10] Bonnmotion - a mobility scenario generation and analysis tool. http://web.informatik.uni-bonn.de/IV/ Mitarbeiter/dewaal/BonnMotion/.

[11] T. Kosch. Efficient message dissemination in vehicle ad-hoc networks. In Proceedings of the 11th World Congress on Intelligent Transportation Systems, October 2004.

[12] C. Schroth, F. D¨otzer, T. Kosch, B. Ostermaier, and M. Strassberger. Simulationg the traffic effects of vehicle-to-vehicle messaging systems. In Proceedings of the 5th International Conference on ITS Telecommunications, June 2005.

Author

PATIL V.P. is currently working as a faculty member in Electronics and Telecommunication Engineering department in smt. Indira Gandhi college of Engineering .New Mumbai. His is graduate in B.E. and post graduate in M.TECH (ELECTRONICS DESIGN AND TECHNOLOGY).He is having 25 years of experience in teaching in engineering colleges. His area of research is in computer communication networking and microwave engineering.

Dynamic Modelling Of Single-Phase Permanent Capacitor Induction Motor and Study of Non-Linear Phenomenon

[1]Mr. Animesh Karmakar [2]Mr. Nihar Ranjan Roy [3]Mr. Rajarshi Mukherjee
[4]Dr. (Prof) Pradip Kumar Saha [5]Dr. (Prof) Gautam Kumar Panda

[1, 2, 3] (PG Student, Electrical Dept., Jalpaiguri Govt. Engg. College),
[4,5,] (Professor, Electrical Dept., Jalpaiguri Govt. Engg. College),

Abstract:
This paper presents a non-linear dynamic model of a permanent capacitor single-phase induction motor. The D–Q axis model of the permanent capacitor induction motor, based on the state vector analysis of the system is conducted; revealing the periodic and chaotic phenomenon under different system parameters. Accordingly, a chaotic-speed fan can be achieved by appropriately varying the motors' internal parameters or the operation condition. Mathematical analysis and computer simulations are attached to testify the proposed non-linear model of the permanent capacitor single-phase induction motor.

Keywords: bifurcation, chaos, non-linear, permanent-capacitor, periodicity, phase-plot, single-phase.

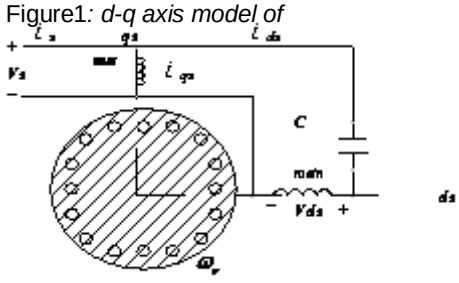
Introduction:

Permanent capacitor single-phase induction motors (SPCIMs) are commonly found in the drive systems of fans, compressors and pumps etc. Here in these particular modelling performance characteristics of the induction motors circuit of lumped parameters are used; offering simplicity and fast computation. Assumptions of the mathematical linearity of magnetic circuits and negligible slotting of the stator and rotor are considered. Chaos can be defined as an a-periodic long term behaviour in a deterministic system that exhibits sensitive dependence on initial condition. Power systems, power converters, motor drives, telecommunications and medical electronics are few of the electrical systems that have already been identified to exhibit chaotic behaviors; characterized by a noise-like spectrum which has a continuous broad-band nature. As the initial state of a practical system can only be specified with some tolerance, the long-term behaviour of a chaotic system can never be predicted accurately.

D-Q axis modeling of the motor:

The D-Q model a single phase induction machine can be considered to be an unsymmetrical two phase induction machine. Three-phase machines' modeling can be easily be defined by the D-Q axis modeling. The equivalent circuit thus obtained can be represented for the SPIM is shown in alongside figure, and the machine is modeled by the following equations:

Figure1*: d-q axis model of*



$$V_{qs}^{s} = V_{s}$$

$$V_{ds}^{s} = V_{s} - \frac{1}{C}\int i_{ds}\,dt$$

$$V_{ds} = i_{ds}R_{ds} + p\lambda_{ds} - \omega_r\lambda_{qs}$$
$$V_{qs} = i_{qs}R_{qs} + p\lambda_{qs} + \omega_r\lambda_{ds}$$
$$V_{dr} = i_{dr}R_{dr} + p\lambda_{dr}$$
$$V_{qr} = i_{qr}R_{qr} + p\lambda_{qr}$$

$$T_e = \frac{p}{2}\left(\lambda_{ds}i_{qs} - \lambda_{qs}i_{ds}\right)$$
$$T_e = T_L + J\frac{d\omega_r}{dt} + B_m\omega_r$$

$$\Rightarrow \frac{d\omega_r}{dt} = \frac{T_e}{J} - \frac{T_L}{J} - \frac{B_m}{J}\omega_r$$

where,

R_{dr} = Direct axis rotor resistance,

R_{qr} = Q-axis rotor resistance,

L_{lds} = Direct axis stator leakage inductance,

L_{lqs} = Q-axis stator leakage inductance,

L_{md} =Direct axis mutual inductance,

L_{ldr} = Direct axis rotor leakage inductance,

L_{lqr} = Q-axis rotor leakage inductance,

L_{mq} = Q-axis mutual inductance,

ω_r = Rotor angular speed and

ω = Speed of the reference frame,

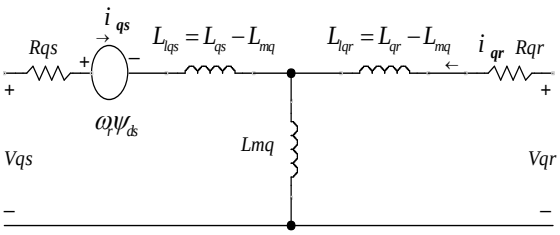


Figure 3: *Equivalent Q-axis model of the induction motor*

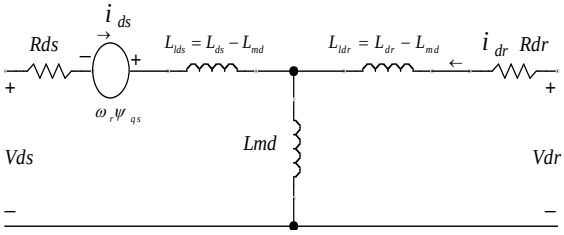


Figure 2: *Equivalent D-axis model of the induction motor*

MATLAB/SIMULINK Modelling:

For a system possessing more than one unique behavior, as a parameter is varied, an abrupt change in the steady-state behavior of the system is called a bifurcation. Here we vary the amplitude of applied voltage. The voltage is varied from zero to 160 volt (r.m.s). The Two (main and auxiliary) stator windings are displaced 90 degree in space. Normally the capacitor is connected with auxiliary winding and the value is selected in such a way that the auxiliary current leads the main winding current by 90 degree for balanced operation. The modeling of the single phase permanent capacitor induction motor projects the non-linear model of the system.

Here to inject non linearity the capacitor is connected to the main winding instead of auxiliary winding and the inductance value of the auxiliary winding is chosen in such a way that the phase difference of currents in the two winding is very less and also different in magnitude. It can be found that the chaotic speed waveforms offer the well-known chaotic properties, namely random-like but bounded oscillations. Also, these waveforms are a-periodic and very sensitive to the initial condition. Physically, these chaotic motions reveal the unbalanced status of the interaction between the magnetic fields by the main winding and the auxiliary winding.

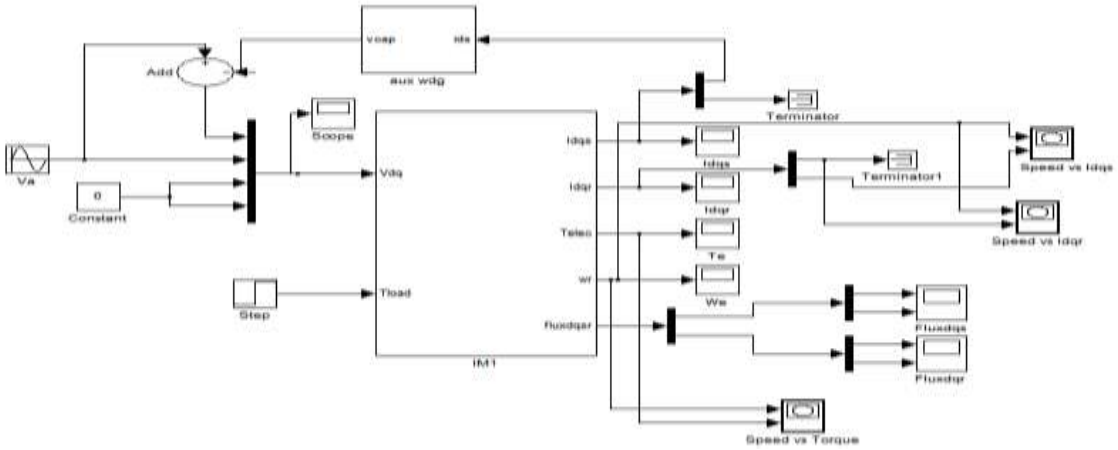


Figure 4: *MATLAB/SIMULINK model of dynamic model of a permanent capacitor single-phase induction motor.*

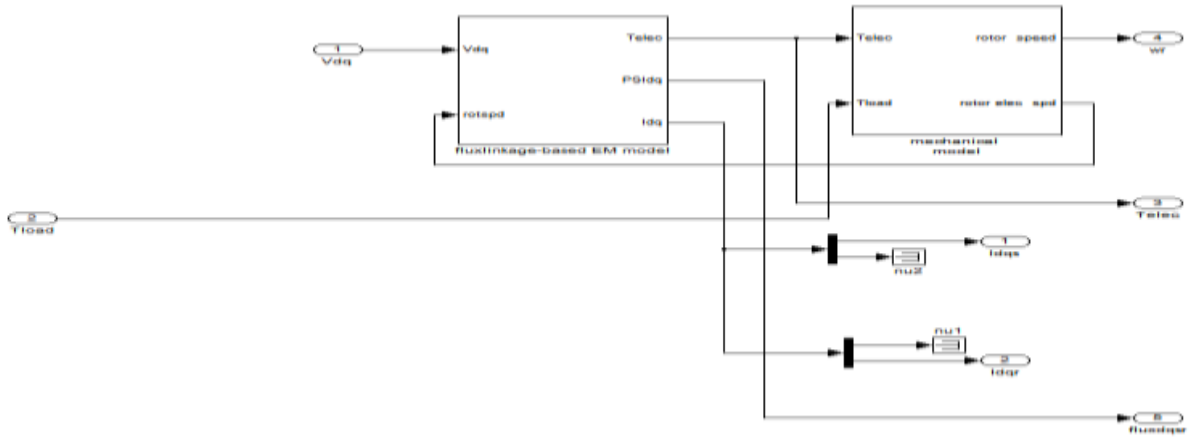


Figure 5: *MATALB/SIMULINK sub-systems for the induction motor*

Parameters	Value	Parameters	Value
Pair Of Poles	1	L_{lar}	0.000482 H
R_{sa}	20 Ω	L_{mo}	0.12 H
R_{ar}	20 Ω	J	0.00007Kg-m^2
L_{sa}	0.0211 H	B_m	0.001e-2 N.m.s
L_{lsa}	0.0117 H	C_{mm}	211.5μF
L_{ma}	0.2356 H	R_{sa}	9.5 Ω
L_{lar}	0.000482 H	R_{ar}	4.5Ω
Frequency (f)	40Hz		

Table 1: *Motor parameters*

Results and simulations:

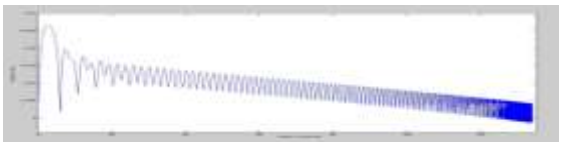


Figure 6: *Speed-Torque Curve for 10 Volt*

Figure 7: *Speed Waveform for 10 Volt (R.M.S)*

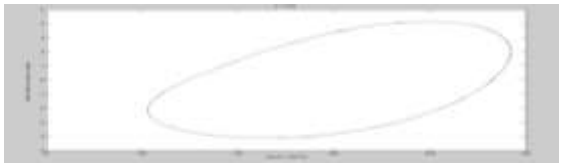


Figure 8: *Idr vs speed for 80 Volt (R.M.S)*

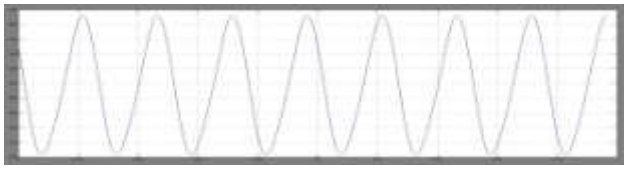


Figure 9: *Speed waveform (Period One) for 80 Volt (R.M.S)*

Figure 10*: Idr Vs Speed for 100 Volt (R.M.S)*

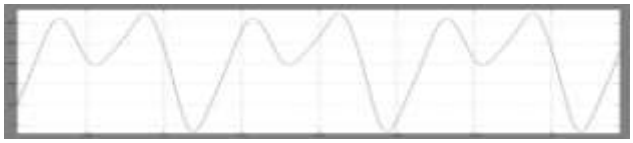


Figure 11: *Speed waveform (Period Two) for 100 Volt (R.M.S)*

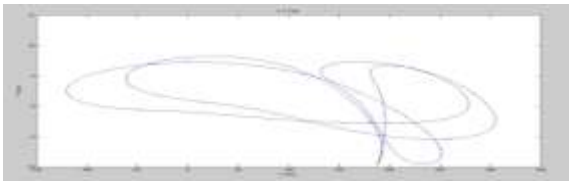


Figure 12: *Idr Vs Speed for 140 Volt (R.M.S)*

Figure 13: *Speed waveform (Period Five) for 140 Volt (R.M.S)*

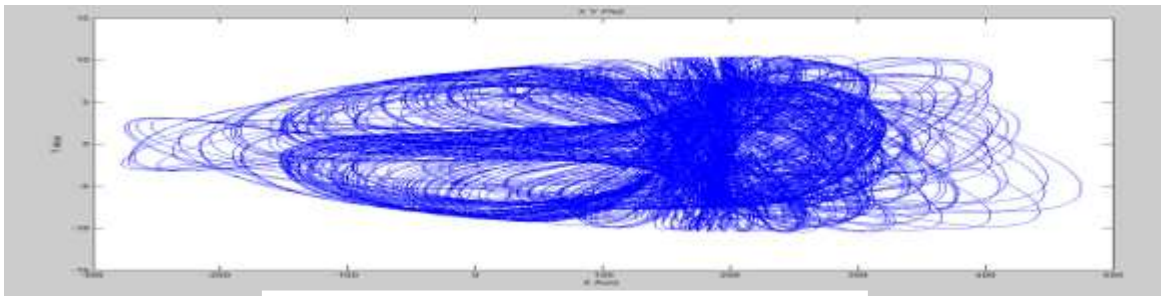


Figure 14: *Idr vs Speed for 160 Volt (R.M.S)*

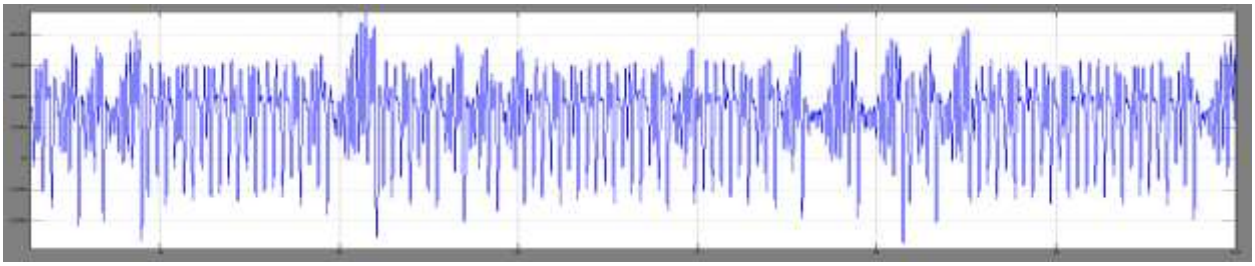


Figure 15: *Chaotic Speed for 160 Volt (R.M.S)*

Conclusions:

The non-linear phenomenon in permanent capacitor single-phase induction motors (SPCIMs) are observed as the input voltage is varied from 0 to 160 Volts. Figures have shown above represents the speed waveforms and the corresponding phase-portraits, at various periodic-speed operations, namely the period-1, period-2 and period-6 operations. These waveforms consist with the well-known phenomenon of inevitable torque pulsation. It should be noted that once the operating condition are known, the motor parameter whose variation bringing chaos may be more than one. Possible candidates are the ratio of *B*m to *J*, main winding resistance, split capacitor and so on.

References:

[1] J.H. Chen, K.T. Chau and C.C. Chan, "Analysis of chaos in current mode controlled dc drive systems," IEEE Transactions on Industrial Electronics, vol. 47, pp. 67-76, 2000.

[2] K.T. Chau and J.H. Chen, "Modeling, analysis and experimentation of chaos in a switched reluctance drive system," IEEE Transactions on Circuits and Systems - I, vol. 50, pp. 712-716, 2003.

[3] Y. Gao and K.T. Chau, "Design of permanent magnets to avoid chaos in PM synchronous machines," IEEE Transactions on Magnetics, vol. 39, no. 5, pp. 2995-2998, 2003.

[4] Gao, Y; Chau, KT; Ye, S "A novel chaotic-speed single-phase induction motor drive for cooling fans", The 40th I A S Annual Meeting Industry Applications Conference Record, Hong Kong, China, 2-6 October 2005, v. 2, p. 1337-1341.

[5] Krzysztof MakowskI, Marcin J. Wilk, Wroclaw University of Technology, Institute of Electrical Machines, Drives and Measurements "Determination of dynamic characteristics of the single-phase capacitor induction motor" Przegland Elektrotechniczny (Electrical Review), ISSN 0033-2097, R.87 NR 5/2011

[6] Dynamic simulation of electric machinery using MATLAB/SIMULINK by Chee-Mun Ong. Prentice Hall PTR 1998, ISBN 0-13-723785-5

[7] Analysis of Electric Machinery and Drives Systems by Paul C. Krause, Oleg Wasynczuk and Scott D. Sudhoff. Sudhoff. Second (IEEE Press Series on Power Engineering) 05-Mar-2002

An Identity-Based Broadcast Encryption Scheme for Mobile Ad Hoc Networks

Sharad Kumar Verma[1], Dr. D.B. Ojha[2]
[1]Research Scholar, Department of CSE, Mewar University, Chittorgarh, Rajasthan, India
[2]Professor, Department of Mathematics, Mewar University, Chittorgarh, Rajasthan, Indi)

Abstract:
A Mobile Ad-hoc Network (MANET) is an autonomous collection of mobile users that communicate over relatively bandwidth constrained wireless links. Since the nodes are mobile, the network topology may change rapidly and unpredictably over time. The network is decentralized, where all network activity including discovering the topology and delivering messages must be executed by the nodes themselves, i.e., routing functionality will be incorporated into mobile nodes. Such networks are more vulnerable to security attacks than conventional wireless networks. In this paper, we propose a secure identity-based ad hoc protocol for mobile devices to construct a group key for a setup of a secure communication network in an efficient way and propose a collision-free method for computing such keys. Unlike group key management protocols, we use identity-based keys that do not require certificates which simplify key management. In contrast to other interactive protocols, we only need one broadcast to setup the group key and member removal is also highly efficient.

Keywords: MANET, Network Topology, Identity-based, Key Management.

1. Introduction

In the next generation of wireless communication systems, there will be a need for the rapid deployment of independent mobile users. Significant examples include establishing survivable, efficient, dynamic communication for emergency/rescue operations, disaster relief efforts, and military networks. Such network scenarios cannot rely on centralized and organized connectivity, and can be conceived as applications of Mobile Ad Hoc Networks. A MANET is an autonomous collection of mobile users that communicate over relatively bandwidth constrained wireless links. Since the nodes are mobile, the network topology may change rapidly and unpredictably over time. The network is decentralized, where all network activity including discovering the topology and delivering messages must be executed by the nodes them selves, i.e., routing functionality will be incorporated into mobile nodes. Such networks are more vulnerable to security attacks than conventional wireless networks. Identity-based encryption (IBE) is a form of public-key cryptography in which a third-party server uses a simple identifier, such as an e-mail address, to generate a public key that can be used for encrypting and decrypting electronic messages. Compared with typical public-key cryptography, this greatly reduces the complexity of the encryption process for both users and administrators. An added advantage is that a message recipient doesn't need advance preparation or specialized software to read the communication. In a broadcast encryption scheme a broadcaster encrypts a message for some subset S of users who are listening on a broadcast channel. A user in S can

use his private key to decrypt the broadcast. Any user outside the privileged set S should not be able to recover the message. The concept of Broadcast Encryption (BE) was introduced by Fiat and Naor. BE is the problem of sending an encrypted message to a large user base such that the message can only be decrypted by a privileged subset. In an ad hoc network, the privileged subset is changing and dynamic. Hence efficiency in transmission cost has been considered to be a critical problem. In addition, the efficiency of a broadcast encryption scheme is also measured by user storage cost and computational cost at a user's device. We provide a general framework for constructing identity-based and broadcast encryption systems. In particular, we construct a general encryption system called spatial encryption from which many systems with a variety of properties follow. The cipher text size in all these systems is independent of the number of users involved and is just three group elements. Private key size grows with the complexity of the system. One application of these results gives the first broadcast HIBE system with short cipher texts. Broadcast HIBE solves a natural problem having to do with identity-based encrypted email.

2. Preliminaries

The following computational problem and complexity assumption are used in the security analysis of our schemes

2.1 Bilinear Maps

Let G_1, G_2, and G_t be cyclic groups of the same order.

2.1.1 Definition

A bilinear map from $G_1 \times G_2$ to Gt is a function
e : $G_1 \times G_2 \rightarrow G_t$ such that for all u ϵ G_1, v ϵ G_2, a, b ϵ Z,

$$\mathbf{e(u^a, v^b) = e(u, v)^{ab}}.$$

Bilinear maps are called pairings because they associate pairs of elements from G_1 and G_2 with elements in G_t. Note that this definition admits degenerate maps which map everything to the identity of G_t.

3. General Diffie-Hellman Exponent Problem

Let p be an integer prime and let s, n be positive integers. Let P,Q ϵ $Fp[X_1, \ldots, X_n]^s$ be two s-tuples of n-variate polynomials over F_p and let f ϵ $F_p[X_1, \ldots, X_n]$. Thus, P and Q are just two ordered sets containing s multi-variate polynomials each. We write P = (p_1, p_2, . . . , p_s) and Q = (q_1, q_2, . . . , q_s). We require that the first components of P,Q satisfy $p_1 = q_1 = 1$; that is, the constant polynomials 1. For a set Ω, a function h : $F_p \rightarrow \Omega$, and a vector $x_1, \ldots, x_n$ ϵ F_p, we write

$$\mathbf{h(P(x_1, \ldots, x_n)) = (h(p_1(x_1, \ldots, x_n)), \ldots, h(p_s(x_1, \ldots, x_n))) \epsilon \Omega^s.}$$

We use similar notation for the s-tuple Q. Let G_0,G_1 be groups of order p and let e : $G_0 \times G_0 \rightarrow G_1$ be a non-degenerate bilinear map. Let g ϵ G_0 be a generator of G_0 and set g_1 = e(g, g) ϵ G_1. We define the (P, Q, f)-Diffie-Hellman Problem in G as follows: Given the vector

$$\mathbf{H(x_1, \ldots, x_n) = (g^{P(x1,\ldots,xn)}, g_1^{Q(x1,\ldots,xn)} \epsilon G_0^{\,s} \times G_1^{\,s},}$$
$$\mathbf{\text{compute } g_1^{\,f(x1,\ldots,xn)} \epsilon G_1}$$

To obtain the most general result, we study the decisional version of this problem. We say that an algorithm B that outputs b ϵ {0, 1} has advantage ε in solving the Decision (P, Q, f)-Diffie-Hellman problem in G_0 if

$$\mathbf{| Pr [B(H(x_1, \ldots, x_n), g_1^{f(x1,\ldots,xn)}) = 0] - Pr[B(H(x_1, \ldots, x_n), T) = 0] | > \varepsilon}$$

where the probability is over the random choice of generator g ϵ G_0, the random choice of $x_1, \ldots, x_n$ in F_p, the random choice of T ϵ G_1, and the random bits consumed by B.

4. Joux's 3-Party Diffie-Hellman

Let G be a group with prime order q, e : G × G→G^t be a bilinear map, and g be a generator of G. Let ˆg = e(g, g) ε G^t.

- Aman picks a $\leftarrow^R Z_q$, Anuj picks b $\leftarrow^R Z_q$, and Sharad picks c $\leftarrow^R Z_q$.
- Aman, Anuj, and Sharad broadcast g^a, g^b, and g^c respectively.
- Aman computes $e(g^b, g^c)^a = \hat{}g^{abc}$, Anuj computes $e(g^c, g^a)^b = \hat{}g^{abc}$, and Sharad computes $e(g^a, g^b)^c = \hat{}g^{abc}$.

4.1 Boneh and Franklin's IBE Scheme

Let G be a group with prime order q, e : G × G→Gt be a bilinear map, and g be a generator of G. Let ˆg = e(g, g) ε Gt. Let h1 : $\{0, 1\}^* \rightarrow G$ and h2 : Gt →$\{0, 1\}^*$ be hash functions. These are all public parameters.

4.1.1 Setup

PKG picks s $\leftarrow^R$ Zq. Then g^s is the public key of PKG.

4.1.2 Encryption

If Aman wants to send a message m to Anuj,

he picks r $\leftarrow^R$ Zq then computes the following.

Encrypt (g, g^s, "Anuj",m)
$= (g^r, m \oplus h_2(e(h_1(\text{"Anuj"}), g^s)^r)$
$= (g^r, m \oplus h_2(e(h_1(\text{"Anuj"}), g)^{rs})$

4.1.3 Making a Private Key

PKG may compute the private key of **Anuj** as follows.

MakeKey (s, "Anuj") = $h_1(\text{"Anuj"})^s$

4.1.4 Decryption

Given an encrypted message (u, v) = $(g^r, m \oplus h_2(e(h_1(\text{"Anuj"}), g)^{rs})$ and a private key w = $h_1(\text{"Anuj"})^s$, Anuj may decrypt as follows.

Decrypt (u, v,w) = $v \oplus h_2(e(w, u))$
$= m \oplus h_2(e(h_1(\text{"Anuj"}), g)^{rs}) \oplus h_2(e(h_1(\text{"Anuj"})^s, g^r))$
$= m \oplus h_2(e(h_1(\text{"Anuj"}), g)^{rs}) \oplus h_2(e(h_1(\text{"Anuj"}), g)^{rs})$
$= m$

How to understand this?

- Let t be the discrete log of h_1("Anuj") base g
- We don't know what it is, but it is well defined
- Now the situation is like 3-party Diffie-Hellman
  - Aman has public g^r, private r
  - PKG has public g^s, private s
  - Anuj has public g^t, unknown (!) t
- $e(h_1(\text{"Anuj"}), g)^{rs} = e(g^t, g)^{rs} = \hat{}g^{rst}$ is like session key for encryption
- Aman and PKG could compute $\hat{}g^{rst}$ just like in Joux's scheme

- But what about Anuj?
- ➢ PKG helps him over previously authenticated, secure channel
- ➢ PKG computes $(g^t)^s = g^{st}$ and sends it to Anuj
- ➢ Anuj can now compute $e(^{gst}, g^r) = \hat{}g^{rst}$
- The point is that Anuj gets g^{st} rather than $\hat{}g^{st}$
- ➢ With g^{st}, still one cheat left
- ➢ If it was $\hat{}g^{st}$ (which anyone can compute), couldn't apply e anymore

5. Adaptive Security Model

A broadcast encryption scheme is said to be secure if given any S, the subscribers not in S as well as the non-subscribers are not able to extract the message from its cipher-text, meant for the subscribers in S, even through collusion. Formally, the security can be defined using the following game between a challenger A and an attacker B:

i. **Setup:** A runs Setup(λ, n) to generate the public parameters which it passes onto B.
ii. **Query Phase 1 :** B adaptively queries about the secret keys of subscribers i_1; i_2; : : : ; i_l and A responds with the keys K_{i1} ;K_{i2} ; : : : ;K_{il} .
iii. **Challenge:** B decides on a set $S^* \subseteq \{1,2,\ldots\ldots,n\} \setminus \{i_1, i_2,\ldots..,i_l\}$ of subscribers it wants to attack. It chooses a pair of distinct messages (M_0,M_1) and gives it to A along with S^*. A chooses a random b ε {0,1} and runs Encrypt (S^*, PP,M_b) to obtain the cipher-text C^* which it gives to B.
iv. **Query Phase 2 :** B continues to adaptively query about the secret keys of other sub- scribers i_{l+1}, $i_{l+2},\ldots\ldots i_{l+m}$, who are not in S^*, and gets the keys K_{il+1},$Ki_{l+2}\ldots\ldots$,K_{il+m}.
v. **Guess:** B guesses b' ε {0,1} for b and wins the game if b = b'.

The broadcast encryption scheme against CPA if for all attacks

$$Pr^{(b = b')} = ½ + e(\lambda)$$

Where e(λ) is a negligible function in λ.

6. Identity Based Security Framework

It is based on the Bohio-Miri scheme and consists of several parts: secure symmetric communication, group broadcast, encryption and signature to support privacy, authentication, integrity, no repudiation, and free key-escrow. Even though the framework is based on the Bohio-Miri scheme, the authors propose a few modifications to reduce its vulnerabilities. It provides two versions of pair wise key agreement: static and dynamic. The static one uses the same static pair-wise key as the Bohio-Miri scheme, providing the most efficient performance.

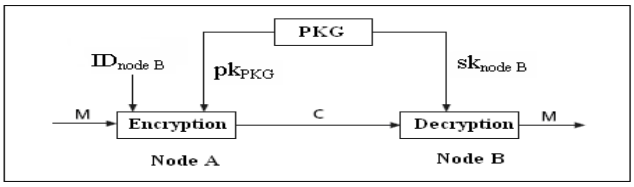


Figure 1: Identity-based encryption scheme.

However, it is fully ID-based, not requiring support structures or online servers. The dynamic pair wise key agreement provides a fresh and distinct key for each session; following the same principles as the static pair-wise key agreement. It also provides a tripartite key agreement to set up secure communication among three entities, and it is used as a primitive for group key management.

7. Conclusion

Security is one of the major issues in MANETs. Their natural characteristics make them vulnerable to passive and active attacks, in which misbehaving nodes can eavesdrop or delete packets, modify packet contents, or impersonate other nodes. We have also presented a description of application fields for ID-based key management. It is important to point out that the major problem with ID-based cryptographic schemes is that they yield only level 2 trust; that is, the private key of users must be known by the key management authority. In conventional networks this is not a major problem, but in MANETs, in which the authority is distributed among online servers or emulated by an arbitrary entity, this may be an issue.

References

[1] D. Boneh and M. Franklin, "Identity- Based Encryption from The Weil Pairing," *SIAM J. Computing*, vol. 32, no. 3, 2003, pp. 586–615.
[2] M. J. Bohio and A. Miri, "Efficient Identity-Based Security Schemes for Ad Hoc Network Routing Protocols," *Ad Hoc Networks*, vol. 2, no. 3, 2004, pp. 309–17.
[3] W. Liu, "Securing Mobile Ad Hoc Networks with Certificateless Public Keys," *IEEE Trans. Dependable Secure Computing*, vol. 3, no. 4, 2006, pp. 386–99.
[4] H. Deng, A. Mukherjee, and D. P. Agrawal, "Threshold and Identity-Based Key Management and Authentication for Wireless Ad Hoc Networks," *Proc. Int'l. Conf. Info. Tech.: Coding and Computing*, vol. 2, 2004, p. 107.
[5] N. Saxena, G. Tsudik, and J. Hyun Yi, "Identity-Based Access Control for Ad Hoc Groups," *Proc.*

Int'l. Conf. Info. Security and Cryptology, 2004.

[6] B. Park and W. Lee, "ISMANET: A Secure Routing Protocol Using Identity-Based Signcryption Scheme for Mobile Ad-Hoc Networks," *IEICE Trans. Commun.*, vol. 88, no. 6, 2005, pp. 2548–56.

[7] J. Pan *et al.*, "Identity-Based Secure Collaboration in Wireless Ad Hoc Networks," *Computer Networks*, vol. 51, no. 3, 2007, pp. 853–65.

[8] A. Khalili, J. Katz, and W. A. Arbaugh, "Toward Secure Key Distribution in Truly Ad-Hoc Networks," *Proc. 2003 Symp. Apps. and the Internet Wksps.*, 2003.

[9] K. Hoeper and G. Gong, "Bootstrapping Security in Mobile Ad Hoc Networks Using Identity-Based Schemes with Key Revocation," tech. rep., Centre for Applied Cryptographic Research, Univ. of Waterloo, 2006.

[10] H. Chien and R. Lin, "Improved ID-Based Security Framework for Ad Hoc Network," *Ad Hoc Networks*, vol. 6, no. 1, 2008, pp. 47–60.

Image Interpolation Algorithm for Edge Detection Using Directional Filters and Data Fusion

B.HIMABINDU
Asst.professor, Dept. of E.C.E, Chalapathi Institute of Technology, Guntur.

Abstract:
Preserving edge structures is a challenge to the image interpolation algorithms to reconstruct a high resolution image from a low resolution counterpart. We propose a new guide edge linear interpolation technique via address filter and data fusion. For a pixel to be interpolated, two sets of observation are defined in two orthogonal directions, and each set produces an estimated value of the pixel. These estimates of direction, following the model the different measures of the lack of noisy pixels are fused by linear least mean square estimation error (LMMSE) technique in a more robust estimate, and statistics two sets of observations. It also presents a simplified version of Based LMMSE interpolation algorithm to reduce computational cost without sacrificing much the interpolation performance. Experiments show that the new interpolation techniques can preserve sharp edges and reduce artifacts call.

Keywords: Bicubical convolution interpolation, Data Fusion, Edge preservation, Image interpolation, Laplacian, Linear Mean Square Estimation Error (LMMSE), optimal weight cubic interpolation.

1. Introduction

Many users of digital images desire to improve the native resolution offered by imaging hardware. Image Interpolation aims to reconstruct a higher resolution (HR) image of the associated low-resolution (LR) capture. You have medical imaging applications, remote sensing and digital Photos [3] - [4], etc. A number of image interpolation methods have been developed [1], [2], [4], [5], [7] - [15]. While commonly used linear methods such as duplication, pixel bilinear interpolation and bicubic interpolation convolution, have advantages in simplicity and fast implementation [7] who suffers from some inherent flaws, including the effects of block, blur the details and call artifacts around the edges. With the prevalence of low cost and relatively digital image LR devices and computing power increasingly interests and the demands of high quality, image interpolation algorithms have also increased. The human visual systems are sensitive to edge structures, to transmit a large part of the semantics of the picture, so that a re-key requirement for image interpolation algorithms to reconstruct faithfully the edges of the original scene. The traditional linear interpolation methods [1] - [3], [4] [5], does not work very well under the criterion of preserving the advantage. Some linear interpolation techniques [7] - [14] have been proposed in recent years to maintain Total sharpness. The interpolation scheme of Jensen and Anastassiou [7] detects the edges and adapts them for some templates to improve the visual perception of large images. Li and Orchard [8] uses the covariance of the estimate LR image covariance HR image, which represents the edge direction information to some extent, and proposed a Wiener filter-as the interpolation scheme. Since this method requires a relatively great window to calculate the covariance matrix for each offense sample, we can introduce some artifacts due to local structures statistics shows the change and, therefore, incorrect estimation of covariance. The interpolator of the image and Tenze Carrato [9] first replicates the pixels and then corrected by the use of some March 3 pre-edge patterns and optimizing the parameters of the operator. Muresan [14] detected no advantage in diagonal and diagonal addresses, and then recovered samples missing along direction detected by one-dimensional (1-D) polynomial interpolation.Some linear interpolation methods try to extend a image by predicting the fine structures of the image of human resources LR counterpart. For this, a multi-resolution representation image is needed. Takahashi and Taguchi [10] represents a Laplacian pyramid image, and with two empirically determine the parameters, it is estimated that the unknown high frequency components of the detail signal LR Laplacian. in the the last two decades, the wavelet transform (WT) theory [16] has been well developed and provides a good framework for multiresolution for the representation of the signal. WT decomposes signal different scales, along the sharp edges which have a signal correlation. Carey, et al. [11] exploits the Lipschitz property sharp edges of the scales of wavelets. Module is used Maximum thicker scales information to predict the unknown wavelet coefficients at the finer scale. Then the HR image is constructed by reverse WT. Muresan and Parks [13] extended this strategy through the influence of a full cone sharp edge in the wavelet scale space, rather than just the top module, for estimation of the best coefficients of scale through an optimal recovery theory. The wavelet interpolation method by Zhu et col. [12] uses a discrete time parametric model to characterize major edges. With this model in the wavelet domain lost information on the edge of the finest scale is recovered via minimum linear mean square estimation error (LMMSE). The previous schemes used, implicitly or explicitly, an isolated sharp edge model as an ideal step edge or softened,in the development of algorithms. For

real images, however, the wavelet coefficients of a sharp edge can be interfered with by the neighboring edges. In general, linear interpolation methods better advantage in preserving the linear methods. In [15], Guichard Malgouyres and analyzed some linear and nonlinear expanding imaging methods theoretically and experimentally. Compared with the discontinuities in the signals of 1-D, Ringer images of two dimensional (2-D) has an additional property: the direction. In the methods of linear interpolation, filtering is 1-D alternatively be made in horizontal and vertical directions without pay attention to the local edge structures. In the presence of a strong advantage if a sample is interpolated fault rather than artifacts from the direction of the edge, large and visually disturbing be introduced. A conservative strategy to avoid more serious devices is to use a 2-D isotropic filter. This, however, reduces the sharpness of the edges. A more "assertive" approach is to interpolate estimated edge in one direction. The problem with the latter is it worth the image quality is high if the estimated edge address is incorrect, which may occur due to difficulty in determining the direction of the edge of the paucity of data provided image by LR.In this paper we propose a new balanced approach to the problem. A sample is interpolated fault in not one but two orthogonal directions. The results are treated as two estimates of the sample and using the statistics fused adaptively a local window. Specifically, the partition of the neighborhood of each sample is missing in two orthogonal oriented subsets directions. The hope is that the observation of two sets exhibit different statistics, since the sample has missing higher correlation with its neighbors in the direction of the edge. Each oriented subset produces an estimate of the missing pixel. The finally pixel is interpolated by combining the two directional estimates on the principle of LMMSE. This process can discriminate the two subgroups according to their consistency absence of the sample, and make the subset perpendicular to the edge contribute less to address the LMMSE estimate of the missing shows. The new approach over a significant improvement the linear interpolation methods in preserving edge sharpness while the suppression of artifacts, by adapting the local interpolation gradient image. A drawback of the interpolation method proposed computational complexity is relatively high. Also interpolation algorithm to develop a simplified greatly reduced computing requirements, but without significant degradation in performance.

2. Edge-Guided Lmmse-Based Interpolation

We take a picture LR image F_t decreased from an image directly associated through human resources$F_h, F_l\ (n,m) = F_h\ (2n-1,2m-1), 1 \le n \le N, 1 \le m \le M$. Concerning the fig. 1, the black dots represent the samples available F_l and white dots represent samples missing from$F$.The problem of the interpolation is to estimate the missing samples F_h in HR image, whose size is 2N X 2M in the samples in LR F_l image whose size NXM.

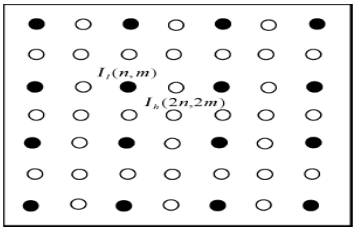


Fig. 1. Formation of an LR image from an HR image by directly down sampling.
The black dots represent the LR image pixels and the white dots represent the missing HR samples.

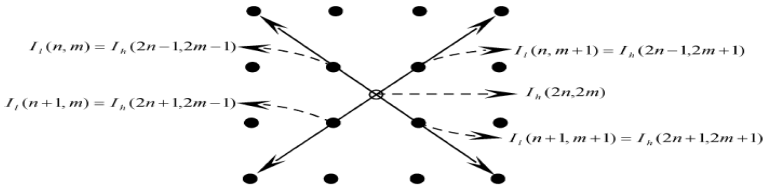


Fig. 2. Interpolation of the HR samples I_h *(2n, 2m)*.

Two estimates of I_h *(2n, 2m)* are made in the 45 and 135 directions as two noisy measurements of I_h *(2n, 2m)*.
The focus of the image interpolation is how to infer and use information about the shows that need to be hidden in neighboring pixels. If the sign of the sub-sampled LR image exceeds the Nyquist sampling, convolution-Methods based on interpolation will be affected by the alias problem in image reconstruction of human resources. This is the cause of artifacts such as ringing effects of image interpolation that are common to the linear interpolation methods. Given that the human visual system is very sensitive to the edges, especially in its spatial location is crucial to suppress interpolation artifacts, while retaining the sharpness

of the edges and geometry.The edge direction information is most important for the interpolation process. To extract and use this information, partitions of neighboring pixels of each sample of lack in two directional subsets are mutually orthogonal. In subset, a directional interpolation is done, and then the two interpolated values are merged to arrive at an estimate of LMMSE the sample is missing. We recover the HR image in two steps. First, the missing samples $F_h(2n, 2m)$ in the center locations surrounded by four samples are interpolated LR. Secondly, the other samples that is missing and is interpolated with the help of samples and recovered.

2.1. Interpolation of Samples $\mathbf{F_h\ (2n, 2m)}$

Referring to Fig. 2, we can interpolate the missing HR sample $F_h\ (2n, 2m)$ along two orthogonal directions: 45 diagonal and 135 diagonal. Denote by $\hat{F}_{45}\ (2n, 2m)$ and the $\hat{F}_{135}\ (2n, 2m)$ two results of interpolation direction from some of the linear methods, as bilinear interpolation, bicubic interpolation convolution [1] - [5]. Note the direction of interpolation results as noisy measurements of the sample failure HR

$$\hat{F}_{45}\ (2n, 2m) = F_h\ (2n, 2m) + u_{45}\ (2n, 2m)$$

$$\hat{F}_{135}(2n,2m) = F_h\ (2n, 2m) + u_{135}\ (2n, 2m) \quad (1)$$

where the random noise variables u_{45} and u_{135} represent the interpolation errors in the corresponding direction.
To fuse the two directional measurements $\hat{F}_{45}$ and $\hat{F}_{135}$ into a more robust estimate, we rewrite (1) into matrix form

$$Z = 1.F_h + U \quad (2)$$

Where

$$Z = \begin{bmatrix} \hat{F}_{45} \\ \hat{F}_{135} \end{bmatrix}, 1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} u_{45} \\ u_{135} \end{bmatrix}$$

Now, the interpolation problem is to estimate the unknown sample F_h from the noisy observation Z. This estimation can be optimized in minimum mean square-error sense. To obtain the minimum mean square-error estimation (MMSE) of F_h, i.e., $\hat{F}_h = E[F_h|Z] = \int F_h p(F_h|Z) dF_h$, we need to know the probability density function $p(F_h|Z)$. In practice, however, it is very hard to get this prior information or $p(F_h|Z)$ cannot be estimated at all. Thus, in real applications, LMMSE is often employed instead of MMSE. To implement LMMSE, only the first and second order statistics of F_h and Z are needed, which may be estimated adaptively.
From (2), the LMMSE of $F_h$can be calculated as [18]

$$\hat{F}_h = \mu_h + cov(F_h, Z)\big(Var(Z)\big)^{-1}(Z - E[Z]) \quad (3)$$

Where $\mu_h = E[F_h], Con(A, B) = E[(A - E[A])(B - E[B])^T]$ is the co-variance operator, and we abbreviate $Cov(A, A)$ as $Var(A)$, the variance operator. Through the LMMSE operation, $\hat{F}_h$ fuses the information provided by directional measurements $\hat{F}_{45}$ and $\hat{F}_{135}$.
Let $\mu_{45}^u = E[U_{45}], \mu_{135}^u = E[U_{135}]$. Through intensive experiments on 129 images, including outdoor and indoor images, portraits, MRI medical images, and SAR images, etc., we found that $\mu_{45}^u \approx 0$ and $\mu_{135}^u \approx 0$. Thus, noise vector U can be considered to be zero mean. Denote by n_1 and n_2 the normalized correlation coefficients of u_{45} and u_{135} with F_h.

$$n_1 = \frac{E[(u_{45} - \mu_{45}^u).(F_h - \mu_h)]}{\sqrt{E[(u_{45} - \mu_{45}^u)^2]E[(F_h - \mu_h)^2]}}$$

$$n_2 = \frac{E[(u_{135} - \mu_{135}^u).(F_h - \mu_h)]}{\sqrt{E[(u_{135} - \mu_{135}^u)^2]E[(F_h - \mu_h)^2]}}$$

Our experiments also show that the values of n_1 and n_2 are very small. Thus, we consider u_{45} and u_{135} and, consequently, U to be nearly uncorrelated with F_h. With the assumption that U is zero mean and uncorrelated with F_h, it can be derived from (3) that

$$\hat{F}_h = \mu_h + \sigma_h^2 1^T(1.\sigma_h^2.1^T + R_v)^{-1}(Z - \mu_h) \quad (4)$$

Where $\sigma_h^2 = var(F_h)$ and $R_V = Var(V)$.To implement the above LMMSE scheme for F_h, parametersμ_h, σ_h^2, and R_V need to be estimated for each sample $F_h(2n, 2m)$ in a local window.

First, let us consider the estimation of μ_h andσ_h^2. Again, referring to Fig. 2, the available LR samples around $F_h(2n, 2m)$ are used to estimate the mean and variance of $F_h(2n, 2m)$. Denote by W a window that centers at $F_h(2n, 2m)$ and contains the LR samples in the neighborhood of $F_h(2n, 2m)$. For estimation accuracy, we should use a sufficiently large window as long as the statistics is stationary in W. However, in a locality of edges, the image exhibits strong transient behavior. In this case, drawing samples from a large window will be counterproductive. To balance the conflicting requirements of sample size and sample consistency, we propose a Gaussian weighting in the sample window W to account for the fact that the correlation between $F_h(2n, 2m)$ and its neighbors decays rapidly in the distance between them. The further an LR sample is from$F_h(2n, 2m)$, the less it should contribute to the mean value of $F_h(2n, 2m)$. We compute μ_h as

$$\mu_h = \sum_l \sum_k W(l,k)\, G(l,k) \qquad (5)$$

Where $G(x,y) = (1/2\pi\zeta^2)\exp\left(-(x^2+y^2)/2\zeta^2\right)$ is a 2-D Gaussian filter with scale ζ. The variance of $I_h(2n, 2m)$ is computed as

$$\sigma_h^2 = \sum_l \sum_k (W(l,k) - \mu_h)^2\, G(l,k) \qquad (6)$$

Next, we discuss the estimation of R_V, the co-variance matrix of U. Using (1) and the assumption that u_{45} and u_{135} are zero mean and uncorrelated with F_h, it can be easily derived that

$$Var(u_{45}) = Var\left(\hat{F}_{45}\right) - \sigma_h^2$$
$$Var(u_{135}) = Var\left(\hat{F}_{135}\right) - \sigma_h^2 \qquad (7)$$

Since σ_h^2 has been estimated by (6), we need to estimate $var(\hat{F}_{45})$ and $var(\hat{F}_{135})$in a local window to arrive at $Var(u_{45})$ and $Var(u_{135})$ For this, we associate $\hat{F}_{45}$ with a set of its neighbors in the 45^0 diagonal direction. Denote by Y_{45} the vector that centers at $\hat{F}_{45}(2n, 2m)$

$$Y_{45} = \{\ldots, \hat{F}_{45}(2n+2, 2m-2), F_l(n+1, m), \hat{F}_{45}(2n, 2m)\, F_l(n, m+1), \hat{F}_{45}(2n-2, 2m+2), \ldots\} \qquad (8)$$

Set Y_{45} encompasses $\hat{F}_{45}$ and its neighbors, i.e., the original samples and the directional (45^0 diagonal) interpolated samples. Symmetrically, we define the sample set Y_{135} for $\hat{F}_{135}$ associated with interpolated results in the 135^0 diagonal

$$Y_{135} = \{\ldots, \hat{F}_{135}(2n-2, 2m-2), F_l(n, m), \hat{F}_{135}(2n, 2m) F_l(n+1, m+1), \hat{F}_{135}(2n+2, 2m+2), \ldots\} \qquad (9)$$

The estimates of $var(\hat{F}_{45})$ and $var(\hat{F}_{135})$ are computed as

$$var\left(\hat{F}_{45}\right) = \sum_k (Y_{45}(k) - \mu_h)^2 g(k)$$

and

$$var(\hat{F}_{45}) = \textstyle\sum_k (Y_{135}(k) - \mu_h)^2 g(k) \qquad (10)$$

Where $g(x) = (1/\sqrt{2\pi\xi})\exp\left(-x^2/2\xi^2\right)$ is a 1-D Gaussian filter with scaleξ.

Now, $Var(u_{45})$ and $Var(u_{135})$ can be computed by (8), and finally the co-variance matrix R_V can be estimated as

$$R_V = \begin{bmatrix} Var(u_{45}) & c_3 \cdot \sqrt{Var(u_{45}) Var(u_{135})} \\ c_3 \cdot \sqrt{Var(u_{45}) Var(u_{135})} & Var(u_{135}) \end{bmatrix} \qquad (11)$$

Where c_3 is the normalized correlation coefficient of u_{45} with u_{135}

$$c_1 = \frac{E[(u_{45} - \mu_{45}^u).(u_{135} - \mu_{135}^u)]}{\sqrt{E[(u_{45} - \mu_{45}^u)^2] E[(u_{135} - \mu_{135}^u)^2]}}$$

Although u_{45} and u_{135} are nearly uncorrelated with F_h, they are somewhat correlated to each other because $\hat{F}_{45}$ and $\hat{I}_{135}$ have some similarities due to the high local correlation. We found that the values of n_3 are between 0.4 and 0.6 for most of the test images. The correlation between u_{45} and u_{135} varies, from relatively strong in smooth areas to weak in active areas. In the

areas where sharp edges appear, which is the situation of our concern and interests, the values of n_3 are sufficiently low, and we can assume that u_{45} and u_{135} are uncorrelated with each other without materially affecting the performance of the proposed interpolation algorithm in practice. In practical implementationn_3, the correlation coefficient between u_{45} andu_{135}, can be set as 0.5 or even 0 for most natural images. Our experiments reveal that the interpolation results are insensitive ton_3. Varying n_3 from 0 to 0.6 hardly changes the PSNR value and visual quality of the interpolated image.If a sharp edge presents in F_h in or near one of the two directions (the 45^0 diagonal or the 135^0 diagonals), the corresponding noise variances $Var(u_{45})$ and $Var(u_{135})$ will differ significantly from each other. By the adjustment of R_V in (4), the interpolation value $\hat{F}_{45}$ or$\hat{F}_{135}$, whichever is in the direction perpendicular to the edge, will contribute far less to the final estimation result$\hat{F}_h$. The presented technique removes much of the ringing artifacts around the edges, which often appear in the interpolated images by cubic convolution and cubic spline interpolation methods.

2.2. Interpolation of samples $\mathbf{F_h(2n-1,2m)}$ and $\mathbf{F_h(2n,2m-1)}$

After the missing HR samples $\hat{F}_h(2n, 2m)$ are estimated, the other missing samples $F_h(2n-1,2m)$and $F_h(2n, 2m-1)$can be estimated similarly, but now with the aid of the just estimated HR samples. Referring to Fig. 3(a) and (b), the LR image pixels $F_t(n, m)$ are represented by black dots "●" while the estimated samples by symbols "⊗" The samples that are to be estimated are represented by white dots "○" As illustrated in Fig. 3, the missing sample $F_h(2n-1,2m)$or $F_h(2n, 2m-1)$ can be estimated in one direction by the original pixels of the LR image, and in the other direction by the already interpolated HR samples. Similar to (2), the two directional approximations of the missing sample are considered as the noisy measurements of $F_h(2n-1,2m)$and$F_h(2n, 2m-1)$, and then the LMMSE of the missing sample can be computed in a similar way as described in the previous section. Finally, the whole HR is reconstructed by the proposed edge-guided LMMSE interpolation technique.

3. Simplified Lmmse Interpolation Algorithm

In interpolating the HR samples, the LMMSE technique of (4) needs to estimateμ_h, σ_h^2, R_V, and compute the inverse of a 2X2 matrix. This may amount to too heavy a computation burden for some applications that need high throughput. Specifically, if we set μ_h be the average of the four nearest LR neighbors of F_h to reduce computation, then computing μ_h needs three additions and one division and computing σ_h^2 needs seven additions, four multiplications, and one division. By setting the size of vector Y_{45} and Y_{135} as 5 and setting $R_V = diag\{Var(u_{45})$ and $Var(u_{135})\}$, i.e., $n_3 = 0$, in (10) to reduce the computational cost, we still need 20 additions and 20 multiplications to compute R_V. The remaining operations in (4) include nine additions, eight multiplications, and one division. In total, the algorithm needs 39 additions, 32 multiplications, and three divisions to compute a $\hat{F}_h$ with (4).One way to reduce the computational complexity of the algorithm is invoked judiciously LMMSE only for pixels where high local activities are detected, and use a simple linear interpolation method in smooth regions. Since edge pixels represent the minority of the total population of the sample, this will result in significant savings in the calculations. Furthermore, a simplified version of the algorithm based on LMMSE interpolation while only slightly decreasing the performance.We can see that the LMMSE estimate of HR sample I_h is actually a linear combination of $\hat{F}_{45}$, $\hat{F}_{135}$ and μ_h. Referring to (4) and let $\Gamma = \sigma_h^2 1^T(1.\sigma_h^2.1^T + R_V)^{-1}$, then Γ is a 2-D vector and we rewrite (4) as

$$\hat{F}_h = \Gamma_1.\hat{F}_{45} + \Gamma_2.\hat{F}_{135} + (1 - \Gamma.1)\mu_h \quad (12)$$

Where Γ_1 and Γ_2 are the first and second elements of Γ. We empirically observed that $(1 - \Gamma.1)$ is close to zero, and, hence, μ_h has a light effect on $\hat{F}_h$. In this view, $\hat{F}_h$ can be simplified to a weighted average of $\hat{F}_{45}$ and $\hat{F}_{135}$, while the weights depend largely on the noise covariance matrix R_V.

Instead of computing the LMMSE estimate of F_h, we determine an optimal pair of weights to make $\hat{F}_h$ a good estimate of F_h. The strategy of weighted average leads to significant reduction in complexity over the exact LMMSE method. Let

$$\hat{F}_h = w_{45}.\hat{F}_{45} + w_{135}.\hat{F}_{135} \quad (13)$$

Where $w_{45} + w_{135} = 1$. The weights w_{45} and w_{135} are determined to minimize the mean square-error (MSE) of $\hat{F}_h$:

$\{w_{45}, w_{135}\} = arg\ \min_{w_{45}+w_{135}=1} E[(\hat{F}_h - F_h)^2]$.

Although the measurement noises of $\hat{F}_{45} and\ \hat{F}_{135}$, i.e., u_{45} and u_{135} , are Correlated to some Extent, Their correlation is Sufficiently low to Consider u_{45} and u_{135} as being approximately uncorrelated. This assumption holds better in the areas of Edges That Are Critical to the human visual system and of interests to us. In fact, if u_{45} and $u_{135}$are highly Correlated, That is to say, the two estimates $\hat{F}_{45} and\ \hat{F}_{135}$ are close to each other, then F_h varies little in w_{45} and w_{135}anyway. With the assumption That u_{45} and $u_{135}$are approximately uncorrelated, we can show the optimal weights are That

$$w_{45} = \frac{var(u_{135})}{var(u_{45})+var(u_{135})},\ w_{135} = 1 - w_{45} \quad (14)$$

It is quite intuitive why weighting system works. For example, for an edge in or near the 135 ° diagonal direction, the variance Var (u_{45}) is greater than var (u_{135}). Of (14), which is smaller than $w_{45}$and w_{135}will therefore have less influence F_{45} in F. $\hat{F}_{135}$, and vice versa. To calculate var (u_{45}) and var (u_{135}) as described in Section II, however, we still have 30 additions, 24 multiplications and divisions, two. In order to simplify and speed up the calculation of $w_{45}$and w_{135}, we use the following approximations: $var(u_{45}) \cong (\sum_{k=1}^{5}|Y_{45}(k) - \mu_h|)^2$

$$var(u_{135}) \cong \left(\sum_{k=1}^{5}|Y_{135}(k) - \mu_h|\right)^2 \quad (15)$$

where "≅" is almost equivalent. With the above simplification, we only need 23 additions, multiplications and divisions to two, two to get w_{45} and w_{135}. Finally, with (13), only need 24 additions, multiplications and divisions of four, two to get $\hat{F}$. This results in significant savings compared with computational (4), which requires 39 additions, 32 multiplications and divisions, three. Table I shows the counts of the algorithm and the algorithm simplified LMMSE.

TABLE I
Operations needed for the LMMSE algorithm
And the simplified algorithm

Operation	Addition	Multiplication	Division
LMMSE algorithm	39	32	3
Simplified algorithm	24	4	2

4. Experimental Results

The proposed interpolation algorithms were implemented and tested, and their performance was compared to some existing methods. We welcome some images of human resources for the corresponding LR images, of which the original images were reconstructed human resources by the proposed methods and competitive. Since the original images of HR are known in the simulation, we compare the results with real images interpolated, and measure the PSNR of the interpolated images. The interpolator based LMMSE introduced was compared with bicubic convolution interpolation, bicubic spline interpolator, the subpixel edge detection based Anastassiou interpolator and Jensen [7], and the Wiener-like filter interpolator Orchard Li and [8]. To assess the sensitivity of the proposed interpolation algorithms for different initial estimates of management before the merger, which were tested when combined with interpolators bicubic and bilinear convolution, respectively. In the figure legends, the LMMSE method developed in Section II is labeled LMMSE_INTR_linear LMMSE_INTR_cubic or, depending on whether or bilinear bicubic convolution interpolation is used to obtain an initial estimate of direction. Similarly, the simplified method of Section III is labeled OW_INTR_cubic (OW represents the optimal weight) or OW_INTR_linear. In the experiments, sets the scale of 2-DG ζ Gaussian filter [referring to (5)] around 1 and ξ scale of1-D Gaussian filter g [referring to (9)] in about 1.5. Our experimental results also draw attention to a fact that the proposed methods are insensitive to the choice of initial directional interpolators. Even with bilinear interpolation, which normally get significantly worse results than bicubic interpolation, the end result is merged very close to that of bicubic interpolation, especially in terms of visual quality. Figs. 3 and 4 show the interpolated images butterfly Lena and LMMSE_INTR_cubic and LMMSE_INTR_linear methods. In visual effects, the two methods are almost indistinguishable. This shows the power of LMMSE strategy based on data fusion in correcting much of interpolation errors of traditional linear methods.

(a)

(b)

Fig.3. Interpolated image Lena by (a) LMMSE_INTR_cubic and (b) LMMSE_INTR_linear.


(a)


(b)

Fig.4. Interpolated image *Butterfly* by
(a) LMMSE_INTR_cubic and (b) LMMSE_INTR_linear.


(a)


(b)

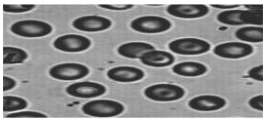

(c)

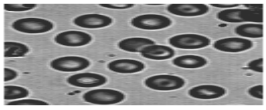

(d)

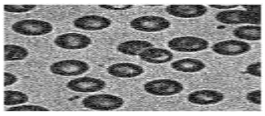

(e)

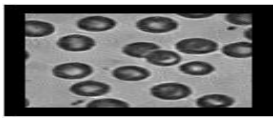

(f)

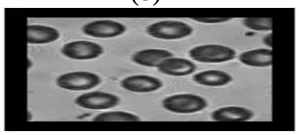
Fig.5. Interpolation results of the image *Lena*. (a)Original image, interpolated image by (b) the cubic convolution, (c)the method in [8], (d) the method in [9],(e)the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.


(a)


(b)


(c)


(d)


(e)

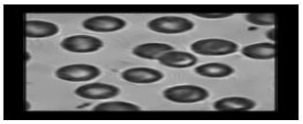
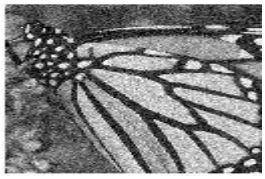

(f)

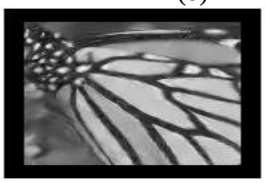
Fig.6. Interpolation results of the image blood. (a) Original image, interpolated image by (b) the cubic convolution, (c) the method in [8], (d) the method in [9],(e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.


(a)


(b)


(c)


(d)


(e)

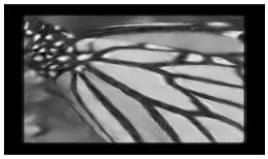
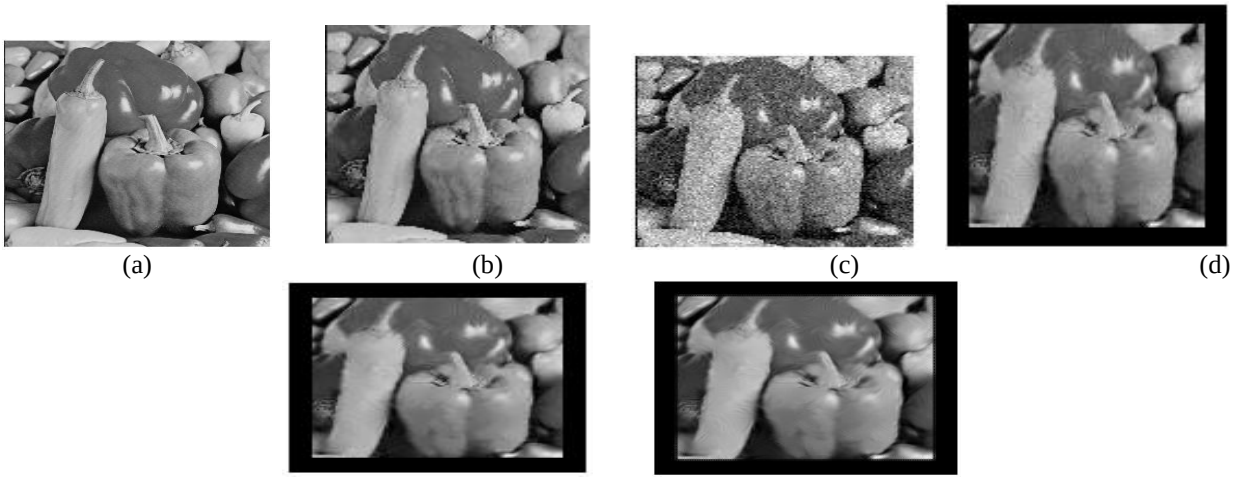

(f)

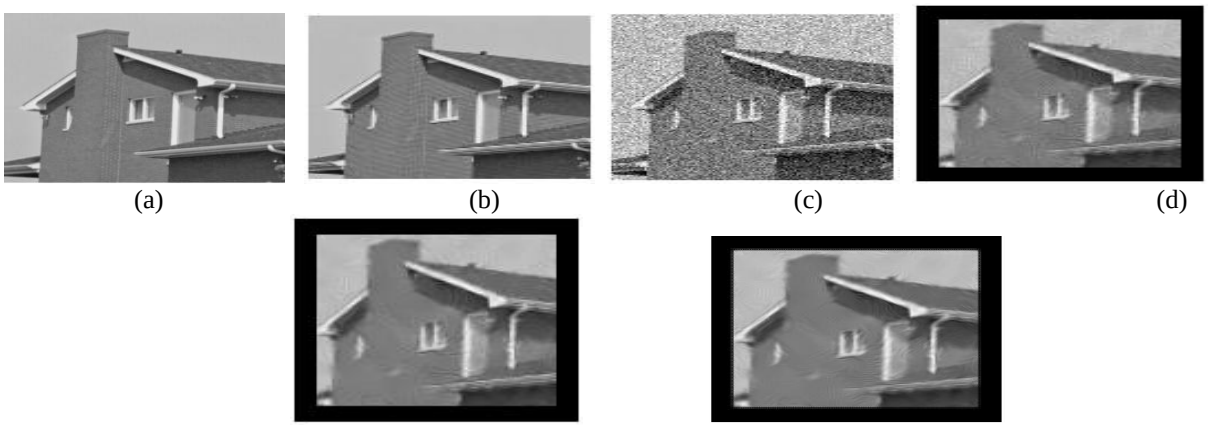
Fig.7. Interpolation results of the image butterfly. (a) Original image, interpolated image by (b) the cubic convolution, (c) the method in [8], (d) the method in [9],(e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.



(a) (b) (c) (d)

(e) (f)

Fig.8. Interpolation results of the image peppers. (a) Original image, interpolated image by (b) the cubic convolution, (c) the method in [8], (d) the method in [9],(e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.



(a) (b) (c) (d)

(e) (f)

Fig.9. Interpolation results of the image house. (a) Original image, interpolated image by (b) the cubic convolution, (c) the method in [8], (d) the method in [9],(e) the proposed LMMSE_INTR_cubic, and (f) the proposed OW_INTR_cubic.

In Figs. 5-9, we compare the visual quality of the test interpolation methods for natural images: Lena, blood samples, butterfly, Peppers, and house. The proposed methods remove many of the ringing and other visual artifacts of the other methods. The OW_INTR_cubic method is slightly inferior to the LMMSE_INTR_cubic method in reducing the ringing effects, but this is a small price to pay for the computational savings of the former. The interpolator of Jensen and Anastassiou [7] can reproduce very thin edges in the object contour because it contains a subpixel edge detection process, but it causes visible artifacts when the edge detector commits errors.This method leaves a considerable amount of ringing effects in the hat of Lena and the wing of the Butterfly. The interpolator of Li and Orchard [8] can preserve large edge structures well, such as those in Lena; however, it introduces artifacts in the finer edge structures, such as the drops of Splash and the head part of Butterfly. Another disadvantage of Li and Orchard's method is its high computational complexity. If an 8x8 window is used to compute the covariance matrix, this algorithm requires about 1300 multiplications and thousands of additions. In comparison, the proposed LMMSE_INTR_cubic algorithm requires only tens of multiplications and divisions. The down-sampling process considered in this paper, through which an LR image is generated from the corresponding HR image, is ideal Dirac sampling. An alternative model of LR images is that of low-pass filtering followed by down-sampling.

5. Conclusion

We have developed an edge type guided LMMSE image interpolation technique. For each pixel to be interpolated, the partition of their neighborhood into two subsets of observation in two orthogonal directions. Each subset observation was used to generate an estimate of the missing sample. These two directional estimates were processed as two sample noisy measurements missing. Using statistics and combination of the two subsets of observation merged the two measurements of noise in a more robust estimation through linear minimum mean square estimation error. To reduce the computational complexity of the proposed method was simplified to an optimal weighting problem and determines the optimum weights. The simplified method had a competitive performance with significant computational savings. The experimental results showed that the methods presented avoided interpolation against edge directions and, therefore, achieve remarkable reduction in timbre and other visual artifacts.

Referance

[1] H. S. Hou, "Cubic splines for image interpolation and digital filtering,"*IEEE Trans. Acoustic, Speech, Signal Process.*, vol. ASSP-26, no. 6,pp. 508–517, Dec. 1978.

[2] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoustic, Speech, Signal Process.*, vol.ASSP-29, no. 6, pp. 1153–1160, Dec. 1981.

[3] T. M. Lehmann, C. Gönner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imag.*, vol.18, no. 11, pp. 1049–1075, Nov. 1999.

[4] M. Unser, "Splines: A perfect fit for signal and image processing,"*IEEE Signal Process. Mag.*, no. 11, pp. 22–38, Nov. 1999.

[5] M. Unser, A. Aldroubi, and M. Eden, "Enlargement or reduction of digital images with minimum loss of information," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 247–258, Mar. 1995.

[6] B. Vrcelj and P. P. Vaidyanathan, "Efficient implementation of all-digital interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 11, pp.1639–1646, Nov. 2001.

[7] K. Jensen and D. Anastassiou, "Subpixel edge localization and the interpolation of still images," *IEEE Trans. Image Process.*, vol. 4, no. 3,pp. 285–295, Mar. 1995.

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

[9] S. Carrato and L. Tenze, "A high quality 2_image interpolator," *IEEE Signal Process. Lett.*, vol. 7, no. 6, pp. 132–135, Jun. 2000.

[10] Y. Takahashi and A. Taguchi, "An enlargement method of digital images with the predition of high-frequency components," in *Proc.Int. Conf. Acoustics, Speech, Signal Processing*, 2002, vol. 4, pp.3700–3703.

[11] W. K. Carey, D. B. Chuang, and S. S. Hemami, "Regularity –Preserving image interpolation," *IEEE Trans. Image Process.*, vol. 8, no. 9, pp.1293–1297, Sep. 1999.

[12] Y. Zhu, S. C. Schwartz, and M. T. Orchard, "Wavelet domain image interpolation via statistical estimation," in *Proc. Int. Conf. Image Processing*,2001, vol. 3, pp. 840–843.

[13] D. D. Muresan and T. W. Parks, "Prediction of image detail," in *Proc.Int. Conf. Image Processing*, 2000, vol. 2, pp. 323–326.

[14] D. D. Muresan, "Fast edge directed polynomial interpolation," in *Proc.Int. Conf. Image Processing*, 2005, vol. 2, pp. 990–993.

[15] F. Malgouyres and F. Guichard, "Edge direction preserving image zooming: A mathematical and numerical analysis," *SIAM J. Numer.Anal.*, vol. 39, pp. 1–37, 2001.

[16] S. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic,1999.

Evaluation of Thermal Properties of E-Glass/ Epoxy Composites Filled By Different Filler Materials

K.Devendra[1], T. Rangaswamy[2]
[1]Asst. Professor, Dept. of Mech. Engg., SKSVMACET, Laxmeshwar, KA, India,
[2] Professor, Dept. of Mechanical Engineering, GEC, Hassan, KA, India,

Abstract
This paper compares the values of the thermal and fire resistance properties of composites made using the hand layup technique. E-Glass fiber reinforced epoxy composites was fabricated by filling varying concentration of aluminum oxide (Al_2O_3), magnesium hydroxide($Mg(OH)_2$), silicon carbide (SiC), and hematite powder. The main aim of this work was to determine the thermal conductivity, thermal expansion coefficient, time to ignition and flame propagation rate of composites. Experimental results show that Al_2O_3 and Mg $(OH)_2$ filled composites exhibited low thermal conductivities. Composites filled by SiC particles exhibited low thermal expansion coefficient when compared with other filled composites. Fire test results indicated that increase the loading of Al_2O_3, Mg $(OH)_2$, and hematite powder increase the time to ignition and reduces the flame propagation rate of composites.

Key words: Composites, Fillers, Fire resistance, Properties, Thermal

1. Introduction
The primary field of application for fiber reinforced polymer composites is the aerospace industry. For several years, however, composite materials have been increasingly used for various other technical tasks, where it is beneficial to apply lightweight construction materials which have high strength and stiffness characteristics. The favorable specific properties of fiber reinforced polymer composites are based on the low density of the matrix resins used, and the high strength of the embedded fibers. Fabrication of fiber reinforced polymer composites relatively low cost. These composites are considered as replacements for metal materials. Polymer matrix reinforced by fiber is probably the most commonly used form of composites in structural application, such as air craft's, boats, automobiles[1] etc. For many materials applications, information is needed on their thermal properties [2]. In recent years there have been an increasing number of applications such as communication, satellites, high density electronics, and advanced aircraft requiring more effective and light weight thermal management materials [3]. The temperature fields in composites materials cannot be determined unless the thermal conductivity of the media is known and for any material a low thermal expansion is ideally required [2, 4]. Polymer composites are very flammable due to their chemical structure and they cannot satisfy some applications, which require high flame retardency. H. Dvir et al. [5] have studied the effect of additives such as pentabromobenzyleacryllate (PBBMA) and magnesium hydroxide on mechanical and flame retardant properties of polypropylene composites containing glass fibers. He concluded that addition of PBBMA and Mg (OH_2) has a positive effect on flame retardant properties and with minimal negative effect on mechanical properties and optimized the formulation of percentage of matrix and reinforcement. The best formulation obtained was with 38.50% matrix 22.31 % fibers and 15 % PBBMA and 17.99 % Mg (oH_2). It has good impact strength 90+.4 and modulus 6953mpa, Vo UL-94 rating and total flaming time 6sec. A. Shojaei et al. [7] have determined the thermal conductivity of rubber-based composites friction materials. Various thermally conductive fillers including copper, brass chips with two sizes, aluminum chips, aluminum oxide and talc have been selected and the effect of addition of these fillers on the thermal conductivity. It was found that the thermal conductivity of the materials increases continuously by increasing the content of the selected fillers. Maximum thermal conductivity is related to aluminum chips filled composites. Now a variety of inorganic fillers have been used to change the properties of composites. In this work E-glass/epoxy based composites filled with varying concentrations of aluminum oxide (Al_2O_3), magnesium hydroxide (Mg $(OH)_2$), silicon carbide (SiC), and hematite powder were prepared by hand layup technique. The objective of this work was to investigate the effect of fillers on thermal and fire resistance properties of E-glass fiber reinforced epoxy composites and comparison of results.

2. Materials and Fabrication
2.1. Materials
The composites were made from E-glass fiber and commercially available ARALDITE (L-12) along with hardener K-6. Al_2O_3, Mg $(OH)_2$, SiC and hematite powder was used as filler materials. Aluminum oxide particles is a ceramic powder commonly used filler, it is also used as an abrasive due to its hardness. Magnesium hydroxide is an inorganic compound and it is a white powder with specific gravity of 2.36, very slightly soluble in water; decomposing at 350^o C. Magnesium hydroxide is attracting attention because of its performance, price, low corrosiveness and low toxicity. Silicon carbide exhibits favorable mechanical and chemical properties at high temperatures for many applications. The benefits of using

SiC as reinforcement are improved stiffness, strength, thermal conductivity, wear resistance, fatigue resistance, reduced thermal expansion and dimensional stability. Hematite is an iron oxide with the same crystal structure as that of corundum (rubies and sapphires). Usually the color varies between a metallic grey and red. Hematite is one of the oldest stones mined in our history.

2.2 Fabrication of Composites

The E-glass /Epoxy based composites filled with varying concentrations (0, 10 and 15 Vol %) of aluminum oxide (Al_2O_3), magnesium hydroxide ($Mg(OH)_2$), silicon carbide (SiC), and hematite powder were prepared. The volume fraction of fiber, epoxy and filler materials are determined by considering the density, specific gravity and mass. Fabrication of the composites was done at room temperature by hand lay-up techniques.

2.3 Specimen Preparation

The prepared slabs of the composite materials were taken from the mold and then specimens were prepared from composite slabs for different thermal and fire resistance tests according to ASTM standards. The test specimens were cut by laminate by using different tools. Three identical test specimens were prepared for different tests.

Table 1 Designation of Composite Materials

Material Designation	Glass Fiber (%Volume)	Epoxy (%Volume)	Filler Materials (% Volume)
GE	50	50	Nil
GEA_1	50	40	10% Al_2O_3
GEA_2	50	35	15% Al_2O_3
GEM_1	50	40	10% $Mg(OH)_2$
GEM_2	50	35	15% $Mg(OH)_2$
$GESI_1$	50	40	10% SiC
$GESI_2$	50	35	15% SiC
GEH_1	50	40	10% Hematite
GEH_2	50	35	15% Hematite

3. Experimentation

3.1. Thermal Property Tests

Thermal conductivity and thermal expansion coefficient of prepared composites was determined according to standard methods.

Thermal Conductivity

Thermal conductivity measurements are carried out under steady state condition. According to ASTM E1530 disc shaped specimens with diameter of 50mm and thickness of 10mm are used in the instrument for thermal conductivity measurements. A known constant heat is applied from one side of the specimen. When the thermal equilibrium is attained and the system approaches to steady state situation, the temperature of top bottom surfaces were measured by using thermocouples installed on top and bottom of the specimen. Knowing the values of heat supplied, temperatures, and thickness, the thermal conductivity was determined by employing one-dimensional Fourier's law of conduction. All measurements are carried out approximately in the similar temperature range, i.e., 25-90 ^oC

$$Q = -KA\frac{dT}{dx} = \frac{KA(T1-T2)}{L} \qquad \text{---------- (1)}$$

Where

Q= Heat transfer rate (W)

K= Thermal conductivity (W/m ^oC)

A=Area (m^2)

T_1= temperature at bottom surface of the sample (^oC)

T_2= Temperature at top surface of the sample (^oC)

L= Thickness of the sample

dT/dx = Temperature gradient

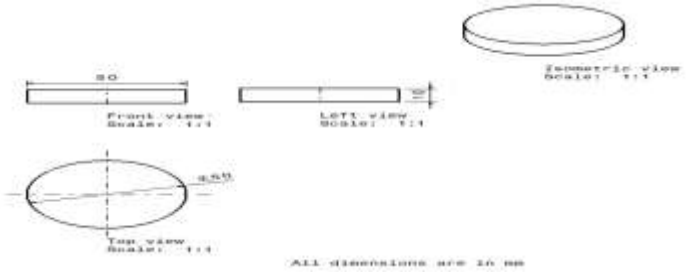


Fig. 1 Thermal conductivity test specimen

Thermal Expansion Coefficient Measurement

The specimens for the thermal expansion coefficient testing had length and thickness 90mm and 10mm respectively. The linear thermal expansion tests are performed over the temperature range of 25 to 90°C using electric furnace. The samples are slowly heated from 30 to 90 ^{0}C in the electric furnace and kept at 90^{0}C for 10 min. Thermal Expansion Coefficient is then given by the relationship

$$\alpha = \frac{\Delta L}{L \Delta T} \qquad \text{---------- (2)}$$

Where

α=Thermal Expansion coefficient
L= Original length of the sample
ΔL= Change in length of the sample
ΔT= Temperature change

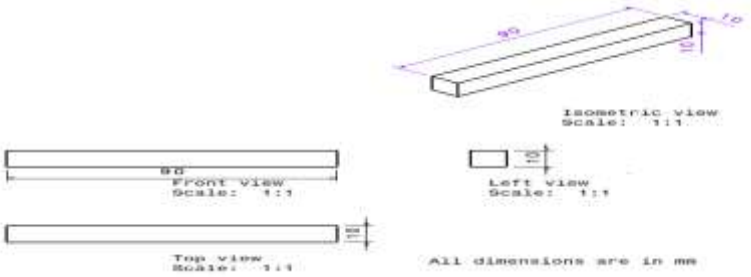


Fig. 2 Thermal expansion coefficient specimen

3.2. Fire Resistance Test

The fire properties of composite materials depend on several factors: the type of material, the fire conditions, and the test method used to measure the property.

Vertical UL94 (UL94V) Test

Rectangular shaped samples with dimensions of 127x12.7x3.2mm are exposed vertically to a methane gas burner flame as required by UL94V. The samples ignited at the bottom and burns upward. The time required for the flame to self extinguish after burner removal is measured and the occurrence of dripping on to a piece of cotton placed underneath the sample is recorded. Test is repeated for different samples. This test also classifies the materials as V-0, V-1 and V-2.
V-0 = specimens not burn more than 10 Sec and the drip do not ignite the cotton.
V-1= specimens not burn more than 30 Sec and the drip do not ignite the cotton.
V-2= specimens not burn more than 30 Sec and the drip ignite the cotton.
If the entire sample is consumed the material is classified as non-rated (NR)

Time-to-Ignition

Time to ignition is the period of time that a combustible material can with stand exposure to a constant heat flux before igniting and undergoing sustained flaming combustion, more simply, it is the time taken for a material to start burning. The ignition time can be used as a rough measure of the flammability resistance of material. Obviously it is desirable to use material with long ignition times in high fire risk applications. Extending the time- to- ignition value reduces the fire hazard of composite material used in an aircraft. The unit of time-to-ignition is seconds (Sec).

Flame Propagation Rate

The rate of the movement of the flame front is defined as the flame propagation rate.

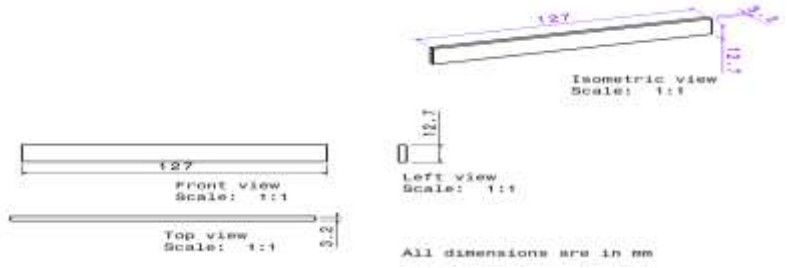


Fig. 3 UL-94 Vertical test specimen

4. Results and Discussion

The thermal conductivity, thermal expansion coefficient, time to ignition and flame propagation rate for different composition of composite materials are presented in tables 2- 5 and their variations shown in figures 4 to 7 respectively.

4.1 Thermal Conductivity

Thermal conductivity is the property describing a material's ability to transfer heat. It is well known that thermal conductivity of the composite is dependent on such factors as Polymer-filler interaction and filler characteristics, namely type and shape of filler. From the obtained results it is observed that composites filled by (10% Vol) Al_2O_3 and Mg $(OH)_2$ exhibited low thermal conductivity it may be due to that while heating of materials slow chemical reaction takes place in Al_2O_3 and Mg $(OH)_2$ and releases the small amount of water and this released water resist the heat flow and seems possible that there is a fundamental difficult in transferring heat from the matrix to the fibers. Composites filled by (10% Vol.) SiC exhibited maximum thermal conductivity (3.515 W/m ^oC). From the literature review we can observed that SiC particles having good thermal conductivity property. Hematite filled composites also exhibited high thermal conductivities when compared with Al_2O_3 and Mg $(OH)_2$ because hematite powder contains the iron particles and these particles enhances the heat transfer rate.

Table 2 Comparison of Thermal Conductivity

Composite materials	Thermal Conductivity(W/m ^oC)
GE	2.89
GEA_1	1.32
GEA_2	1.72
GEM_1	1.565
GEM_2	2.38
$GESI_1$	3.515
$GESI_2$	2.765
GEH_1	2.45
GEH_2	3.06

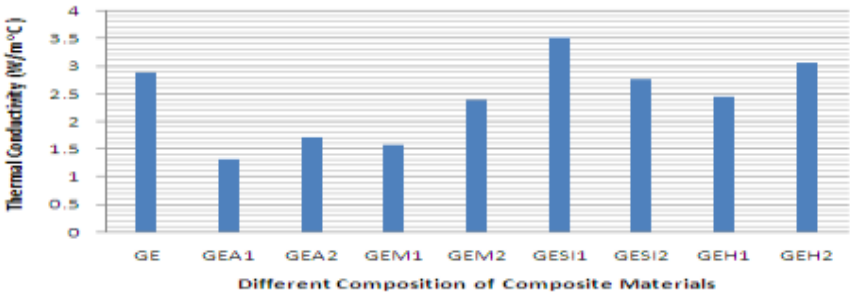


Fig. 4 Thermal conductivity for different composition of composite materials

4.2Thermal Expansion Coefficient

The experimental values of the thermal expansion coefficient of composites are presented in table-3

Table 3 Comparison of Thermal Expansion Coefficient

Composite materials	Thermal Expansion Coefficient (/^oC)
GE	1.96×10^{-5}
GEA_1	2.40×10^{-5}
GEA_2	1.66×10^{-5}
GEM_1	2.44×10^{-5}
GEM_2	1.11×10^{-5}
$GESI_1$	7.40×10^{-6}
$GESI_2$	3.70×10^{-6}
GEH_1	1.85×10^{-5}
GEH_2	1.85×10^{-5}

From the experimental results it is observed that increase the addition of Al_2O_3, Mg $(OH)_2$ and SiC to composites reduces the thermal expansion coefficient. It has been noticed that (15% Vol) Al_2O_3, Mg $(OH)_2$ and SiC filled composite exhibited less thermal expansion coefficient this may be by adding the more fillers in composite materials providing good filler matrix interaction in the system, the filler binds the matrix and prevents it from expanding as much as it would on its own. Subsequently, this would affect the thermal expansion of the composite system. The many studies have shown that materials with higher filler content leads to a lower thermal expansion coefficient. Composites filled by (15% Vol.) SiC exhibited low thermal expansion coefficient (3.70×10^{-6}) when compared with other filled composites this due to that SiC particle having reduced thermal expansion and good dimensional stability. Hematite filled composites exhibited good thermal stability.

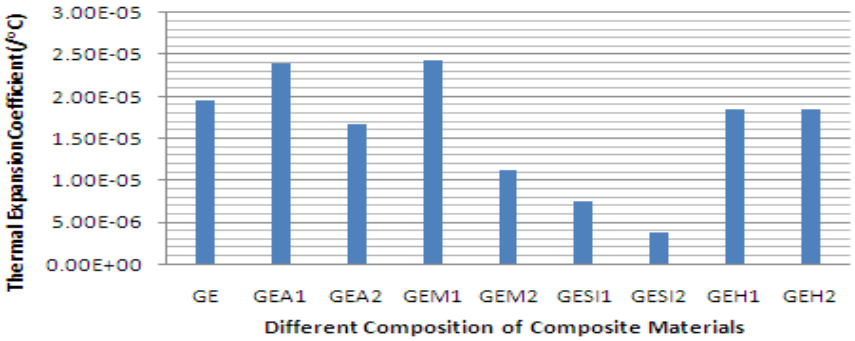


Fig.5 Thermal expansion coefficient for different composition of composite materials

4.3Time to Ignition and Flame Propagation Rate

A popular method to reduce the flammability of composites is the addition of inert fillers like silica or thermally active fillers like hydrated oxides to the polymer matrix.From the figs 6 and 7 it is observed that increase the loading of Al_2O_3, and Mg $(OH)_2$ increase the time to ignition and reduces the flame propagation rate because Al_2O_3 is active in both the condensed and gas phases of the combustion process and it is remarkably effective in suppressing flaming combustion and smoke. The main condensed phase mechanism of aluminum oxide is the absorption of heat when the filler decomposes. And this is a highly endothermic reaction that absorbs heat; another important aspect of the reaction is the creation of water vapor formed from the hydroxyl groups binded to the aluminum. This water is released into the flame where it hinders combustion by diluting the concentration of flammable gases evolved from the polymer matrix and restricting the access of oxygen to the composite surface. Magnesium hydroxide also acts as flame retardants in a similar manner to Al_2O_3 with several flame retardant mechanisms occurring in a fire. Magnesium compound undergo a highly endothermic decomposition reaction that slows the heating rate of the host material in a fire. In addition, the hydroxyl groups bonded to the magnesium are converted in the reaction into water vapor, which dilutes the concentration of flammable organic volatiles and H/Oh radicals in the flame. The decomposition of magnesium compounds also yields magnesia (MgO) that has good insulating properties. The flame retardant mechanisms of magnesium hydroxide are effective in prolonging the ignition time and reducing the amount of smoke produced by high temperature polymers. More addition of iron based hematite filler decreases the time to ignition and flame spread rate fe and derivatives are believed to decompose at elevated temperature which promotes the formation of char via a Lewis acid reaction process. The increased char yield reduces the amount of flammable organic volatiles released by decomposing polymer, which lowers the fuel load in the flame. From the obtained results we can concluded that Al_2O_3, Mg $(OH)_2$ are good flame retardant fillers. All the composites having flame extinguished time more than 30sec and failed to come under UL-94 rating.

Table 4 Comparison of Time to Ignition

Composite materials	Time To Ignition (Sec)
GE	08
GEA_1	11
GEA_2	14
GEM_1	12
GEM_2	15
$GESI_1$	09
$GESI_2$	11
GEH_1	10
GEH_2	12

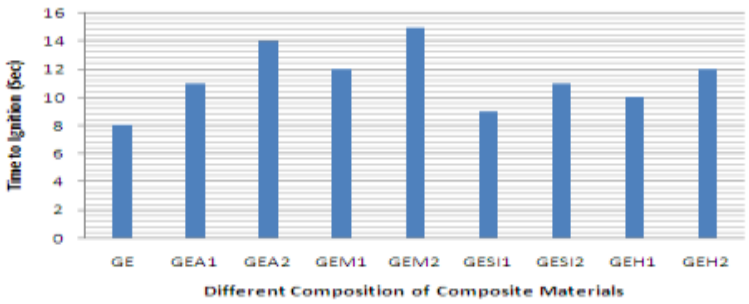


Fig.6 Time to ignition for different composition of composite materials

Table 5 Comparison of Flame Propagation Rate

Composite materials	Flame propagation Rate (mm/sec)
GE	0.60
GEA_1	0.47
GEA_2	0.44
GEM_1	0.49
GEM_2	0.45
$GESI_1$	0.512
$GESI_2$	0.52
GEH_1	0.537
GEH_2	0.514

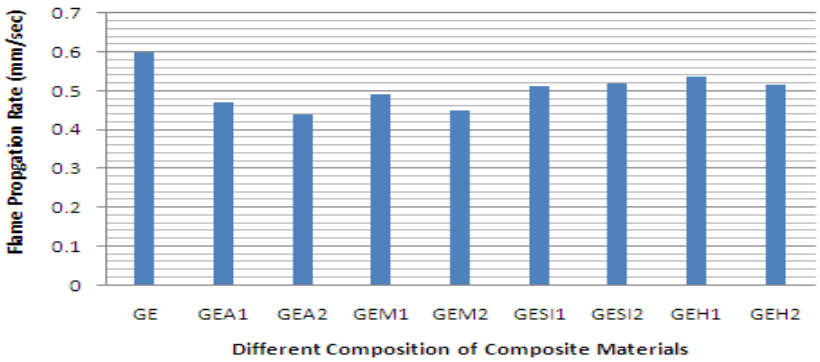


Fig. 7 Flame propagation rate for different composition of composite materials

5. Conclusions

1. From the obtained results it is observed that composites filled by (10% Vol) Al_2O_3 and Mg $(OH)_2$ exhibited low thermal conductivities. Composites filled by (10% Vol.) SiC exhibited maximum thermal conductivity (3.515 W/m ^oC). Hematite filled composites exhibited high thermal conductivities when compared with Al_2O_3 and Mg $(OH)_2$ filled composites.
2. when increase the adding of filler materials to composites reduces the thermal expansion coefficient it is seen from the composites filled by (15% Vol) Al_2O_3, Mg $(OH)_2$ and SiC when compared with (10% Vol) filled composites. Composites filled by SiC exhibited low thermal expansion coefficient when compared with other filled composites.
3. Fire test results show that increase the loading of Al_2O_3, Mg $(OH)_2$, and hematite powder increase the time to ignition and reduces the flame propagation rate of composites. From the obtained results we can concluded that Al_2O_3 and $Mg(OH)_2$ are good flame retardant fillers.
4. All the specimens having flame extinguish time more than 30 sec and failed to come under UL-94V rating.

References

[1] Maries Indicula, Abderrahim Boudenne. Thermopyhsical properties of natural fiber reinforced polyester composites. Composites science and technology, 66, (2006), 2719-2725

[2] Dilek Kumlutas, Ismail H Tavman. Thermal conductivity of particle filled polyethylene composite materials' Composites science and technology, 63 (2003), 113-117

[3] M T Assael, K D Antoniadis. The use of the transient hot-wires technique for measurement of the thermal conductivity of an epoxy-resin reinforced with glass fibers and/or corbon multiwall nanotubes. 68, (2008), 3178-3183

[4] Y LE Bozee , S Kaang , P J Hine. The thermal –expansion behavior of hot –compacted polypropylene and polyethylene composites. Composites science and technology, 60, (2003), 333-344

[5] H. Dvir, M. Gottlied. Optimization of a flame–retarded polypropylene composite. Composites science and technology', 63 (2003), 1865-1875

[6] P K Rohatgi, N. gupta and Simon Alaraj. Thermal expansion of aluminium-flyash cenosphere composites synthesized by pressure infiltration technique. Journal of composite materials, 40, (2006), 1163-1174

[7] A. Shojaei, M Fahimian. Thermally conductive rubber based composite friction materials for rail road brakes-Thermal conduction characteristics. Composites science and technology, 67, (2007), 2665-2674

[8] Hua Gui. Effect of dispersion of nano-magnesium hydroxide on the flammability of flame retardant ternary composites. Composites science and technology, 61 (2001), 475-490

[9] Suhreta Husic, Ivan Javni. Thermal and Mechanical properties of glass reinforced soy-based polyurethane composites. Composites science and technology, 65 (2005), 19-25

[10] Cedric sauder, Jacques Lamon. Thermo mechanical properties of carbon fibers at high temperatures (upto 2000^{o}c). Composites science and technology, 62 (2002) ,499-504

[11] A.P. Mouritz, Z. Mathys, 'Post-fire mechanical properties of glass-reinforced polyester composites', Composites science and technology, 61, (2001) ,475-490.

[12] Zhongfu Zhao, Jihua Gou. Fire retardancy of clay/carbon nano fiber hybrid sheet in fiber reinforced polymer composites. Composites science and technology, 69, (2009), 2081-2087

[13] A yasmin, J. J. Luo. Mechanical and thermal behavior of clay/epoxy nano composites. Composites science and technology, 66, (2006), 2415-2422

[14] A G Gibson, Y S Wu. Laminate theory analysis of composites under load in fire. Journal of composite materials, 639- 647.

[15] Li-ping Gao, De-Yiwang. A flame–retardant epoxy resin based on a reactive phosphorous-containing monomer of DODPP and its thermal and flame–retardant properties. Polymer Degradation and stability, 93, (2008), 1308-1315

[16] Terese E Glodek, Steven E Boyd. Properties and performance of a fire resistant eco-composites using polyhedral oligomeric silsesquioxane (poss) fire retardants. Composites science and technology, 68, (2008), 2994-3001

[17] James Giancaspro, Christos papakonstantinou. Fire resistance of inorganic saw dust bicomposite. Composites science and technology, 68, (2008), 1895-1902