

Interscience Research Network

## Interscience Research Network

---

Conference Proceedings - Full Volumes

IRNet Conference Proceedings

---

12-30-2011

## International Conference on Information , Computing and Telecommunications

Prof.Srikanta Patnaik Mentor

IRNet India, patnaik\_srikanta@yahoo.co.in

Follow this and additional works at: [https://www.interscience.in/conf\\_proc\\_volumes](https://www.interscience.in/conf_proc_volumes)



Part of the [Computational Engineering Commons](#), [Digital Communications and Networking Commons](#), and the [Systems and Communications Commons](#)

---

### Recommended Citation

Patnaik, Prof.Srikanta Mentor, "International Conference on Information , Computing and Telecommunications" (2011). *Conference Proceedings - Full Volumes*. 67.  
[https://www.interscience.in/conf\\_proc\\_volumes/67](https://www.interscience.in/conf_proc_volumes/67)

This Book is brought to you for free and open access by the IRNet Conference Proceedings at Interscience Research Network. It has been accepted for inclusion in Conference Proceedings - Full Volumes by an authorized administrator of Interscience Research Network. For more information, please contact [sritampatnaik@gmail.com](mailto:sritampatnaik@gmail.com).

*Proceedings of International Conference on*  
**INFORMATION , COMPUTING AND TELECOMMUNICATIONS**



(ICICT-2011)  
30<sup>th</sup> DECEMBER, 2011  
BANGALORE, India

Interscience Research Network (IRNet)  
Bhubaneswar, India

## Editorial

If we make a review of the 21<sup>st</sup> century generation, their planning and activities, their interest and involvement driven by some electronic device coupled with an advanced technical functioning. There is a spectacular focus on Information science and Communication technology as it drives the present socio-technical system of the global society. ICT appears as an effective tool for empowering all the civic and anti civic systems of the world. ICT is emerging as an investment area in the millennium development goals of various organizations like UNO, WTO, IBRD and other international apex bodies. It has become an integrated discipline in all the academic spheres as it increases global adaptability.

Information Economy report produced by the UNCTAD establishes the fact that developing nations are making a global pace by gaining mileage from ICT. Let me include some of the contents from report "In rich countries, broadband subscribers increased by almost 15% in the last half of 2005, reaching 158 million. Business broadband connectivity grew most significantly -- in the European Union, for example, from 53% of enterprises in 2004 to 63% in 2005. Broadband enables companies to engage in more sophisticated e-business processes and to deliver a greater range of products and services through the Internet, thus maximizing the benefits of information and communication technology (ICT). The use of broadband directly increases competitiveness and productivity, the report says -- and that, in turn, has an impact on macroeconomic growth. It is estimated that broadband can contribute hundreds of billions of dollars a year to the Gross Domestic Products (GDPs) of developed countries over the next few years.

The mushrooming growth of the IT industry in the 21<sup>st</sup> century determines the pace of research and innovation across the globe. In a similar fashion Computer Science has acquired a path breaking trend by making a swift in a number of cross functional disciplines like Bio-Science, Health Science, Performance Engineering, Applied Behavioral Science, and Intelligence. It seems like the quest of Homo Sapiens Community to integrate this world with a vision of Exchange of Knowledge and Culture is coming at the end. Apparently the quotation "Shrunken Earth, Shrinking Humanity" holds true as the connectivity and the flux of information remains on a simple command over an internet protocol address. Still there remains a substantial relativity in both the disciplines which underscores further extension of existing literature to augment the socio-economic relevancy of these two fields of study. The IT tycoon Microsoft addressing at the annual Worldwide Partner Conference in Los Angeles introduced Cloud ERP (Enterprise Resource Planning,) and updated CRM (Customer Relationship Management) software which emphasizes the ongoing research on capacity building of the Internal Business Process. It is worth mentioning here that Hewlett-Packard has been with flying colors with 4G touch pad removing comfort ability barriers with 2G and 3G. If we progress, the discussion will never limit because advancement is seamlessly flowing at the most efficient and state-of-the art universities and research labs like Laboratory for Advanced Systems Research, University of California. Unquestionably apex bodies like UNO, WTO and

IBRD include these two disciplines in their millennium development agenda, realizing the aftermath of the various application projects like VSAT, POLNET, EDUSAT and many more. 'IT' has magnified the influence of knowledge management and congruently responding to social and industrial revolution.

The conference is designed to stimulate the young minds including Research Scholars, Academicians, and Practitioners to contribute their ideas, thoughts and nobility in these two integrated disciplines. Even a fraction of active participation deeply influences the magnanimity of this international event. I must acknowledge your response to this conference. I ought to convey that this conference is only a little step towards knowledge, network and relationship.

I congratulate the participants for getting selected at this conference. I extend heart full thanks to members of faculty from different institutions, research scholars, delegates, IRNet Family members, members of the technical and organizing committee. Above all I note the salutation towards the almighty.

Editor-in-Chief

**Prof. (Dr.) Srikanta Patnaik**

President, IRNet India and Chairman IIMT

Interscience Campus, Bhubaneswar

Email: patnaik\_srikanta@yahoo.co.in

# Adoption of Distributed Approach to Traffic Advisory Services For Sufficient Segment Traces

Amit R. Welekar<sup>1</sup> & S.S. Dorle<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering, GHRCE, Nagpur (MH) -441110, India

<sup>2</sup>Department of Electronics Engineering, GHRCE, Nagpur (MH) -441110, India

---

**Abstract**—this paper proposed the Distributed Approach to traffic advisory services for bidirectional road and presents the expected result. Inter vehicle communication in VANET requires to handle transmission of data in sparse as well as in dense situation. In bidirectional situation due the increased density of vehicles the aggregation of data is required. We used the Cost-Based aggregation to aggregate the data. The main focus of this paper is on handling of data in case of opposite direction vehicles.

**Keywords:** *VANET, InterVehicle, Bidirectional Road, Distributed VANET.*

---

## I. INTRODUCTION

VANET is an enhancement in ad-hoc wireless network toward the application of vehicular system. In VANET different types of vehicles communicate with each other to share the speed, location and different parameters. The size of information shared and message to be sent at distance is varying according to the density of vehicles.

In today's life traffic congestions are unavoidable part. Drivers waste their most of time in traffic congestion. Intelligent transportation systems are accepted to avoid road accidents, to find nearest road segment, and traffic free movement of vehicles.

There are basically two approaches for intelligent transportation system (ITS). Centralised system in which vehicles communicate with road side unit and in Distributed system there is a vehicle to vehicle communication. In centralised system road side units are placed near road side at a density of 2 or 3 per kilometre. In distributed system vehicles are at any distance from each other. So the communication range in distributed should be large as compared to centralised approach.

At a dense situation in distributed VANET the media should carry maximum data. Consider if the average size of data of vehicles in 100 meters are 1K bytes, and 1M bytes in a range of 250 meters. So the media should carry that amount of data at a time. As a density increases the size of the data to be carried increased, so maximum bandwidth should be required.

In this paper, we present expected result for Distributed Approach for traffic advisory services at simulation time for multiple road scenario. Distributed

Approach defines a framework to disseminate and gather information about the vehicles on the road. Using such a system, a vehicle driver will be aware of the road traffic, which helps driving in situations like speed and location of surrounding vehicles, road accidents, and road side view several miles long.

## II. RELATED WORK

The research in Inter-Vehicle-Communication has emerged in the past couple of years; mainly because it is a good experimental platform for Mobile Ad Hoc Networks (MANETs), and has a great market potential [1].

Several major automobile manufactures and universities have begun to investigate in this field; GM research centre in CMU [2], BMW Research Labs [3] and Ford Research Labs [4], Rice University [5][6], and Harvard University [7] are a few to name. CarNet [8] project focuses on how the radio nodes in the vehicles get IP connectivity with the help of Grid [9]. In [10], a wireless traffic light system is presented. At the intersection, a static control unit periodically broadcasts the current light status, location of intersection, and a reference point, using which the vehicles approaching the intersection can check their relative position and make a decision accordingly. They also designed collision warning system [4] in which peer-to-peer beacon message exchange is used.

Architecture of the vehicular communication is described in [11]. It integrates inter-vehicle communication (IVC) with Vehicle-Roadside Communication (VRC), where both moving vehicles and base stations can be peers in the system. The peers

are organized into Peer Spaces for message exchange, in which flooding is the main method of delivery. Authors in [6] examine the feasibility of short range communication between fast moving vehicles using Bluetooth, and a mobile test-bed RUSH has been established in [5], composed of the fixed base station and mobile nodes on shuttle buses.

### III. PROBLEM DESCRIPTION

In "Traffic View: A Scalable Traffic Monitoring System" for the entire situation they assume the lane is straight [12]. For this paper we assume multiple road scenario situations. Note that we assume the real situation; a road might be bidirectional, where vehicles move in two opposite directions. In this case, a vehicle will need to examine the movement vector in a record received about another vehicle, and ignore it if that vehicle is moving in the opposite direction. This can also be applied in the case of an intersection where a vehicle might hear about different vehicles moving in different directions.

For all the simulations in this paper, we fixed the length of the road to be 22,000 meters with 4 cross lanes. We used 802.11b (with a data transmission rate of 11Mb) as the wireless media with a transmission range of 250m<sup>3</sup>. During a simulation, nodes broadcast messages periodically. The broadcast period is selected uniformly from [1.75, 2.25] seconds, and each node recalculates the next broadcast period after the current broadcast. For all the simulation runs, we use broadcast messages of size 2312 (the maximum payload size of 802.11b standards) and we fix the simulation time to 600 seconds.

We are using the diffusion mechanism in which each vehicle communicate with each other. In this each vehicle broadcast message about itself and other vehicles will received that message in multi-hop manner.

The aggregation mechanism we used is Cost-Based aggregation. Data aggregation is based on the date semantics. For example, the records from two vehicles can be replaced by a single record with little error if the vehicles are very close to each other, and they are moving with relatively the same speed. The way data aggregation contributes to the Traffic -View system is by delivering as many records as possible in one broadcast message. This way, more new records can be delivered in certain period of time and the overall system performance is improved.

### IV. SYSTEM VIEW

In this section we present the design of the implementing prototype of Distributed Approach system.

#### A. System Design

The scenario for Distributed Approach contains different road connectors. Road connectors are used to construct different road scenario with bidirectional road.

**Load Map:** To load the available previous available road map.

**Fix Map:** to fix the new or previous map.

**No. of vehicles:** To insert the density of vehicles

**Source Vehicles:** To select the source vehicle to which the data is required

**Communication Status:** To know the current communication between two vehicles

**Road Status:** To know the status of the road.

**Direction:** To find the direction of the vehicles.

**Zoom In:** To zoom in the map.

**Zoom Out:** To zoom out the map.

**Run:** To run the simulation.

**Pause:** To pause the simulation for taking snapshots

**Stop:** To end the simulation

**Message Box:** To show the messages like accidents, fuel stations etc.

**Scroll Bars:** To scroll the road map, and message box.

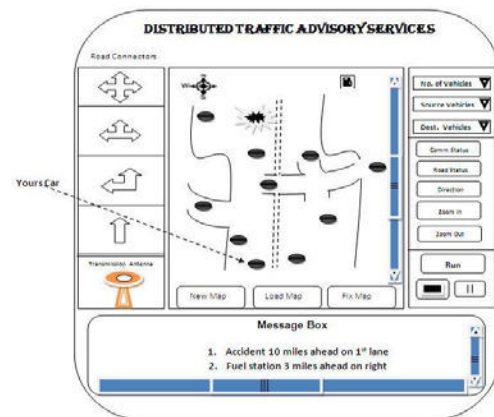


Fig 1: System Design

#### B. Location Aware Data Retrieval

We are using the grid computing to divide the long area of the road. In grid the location aware data retrieval is based on the grid computing. The grid computing divides the geographic location into the grids. By having a large collection of cars working together in a grid environment, drivers can easily determine the best route

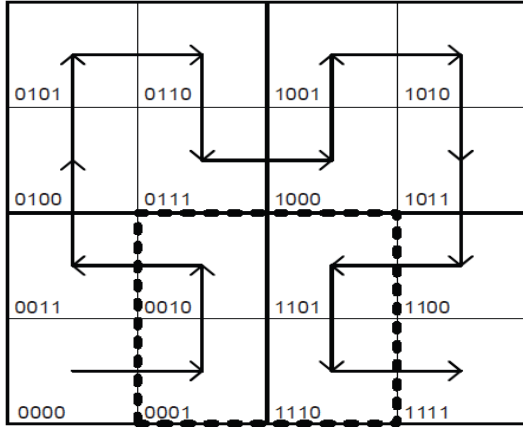
(route with minimal congestion) to reach their destination from their current location.

□ 2Mx2M grids

- M is the smallest exponent that covers the entire area.

□ E.g., size of the contiguous U.S. is approximated as 3000km X 3000km

- 213 X 213 fixed grids where R is given as 1km.



- ▶ Routing Semantics
- ▶ Geocasting to a single grid point
  - ▶ Simple rectangular area based addressing as  $\{(x1; y1); (x2; y2)\}$ .
- ▶ Concurrent Geocasting
- ▶ Due to overhead of sequential search, we utilize how the Hilbert curve is constructed.

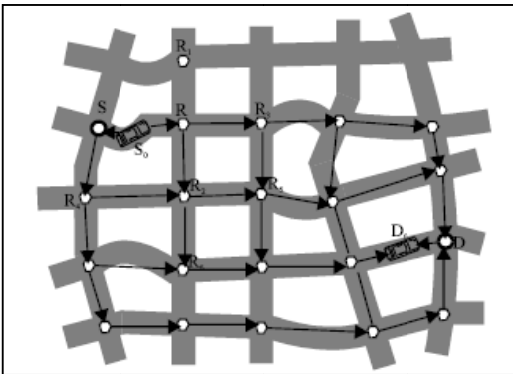


Fig 2: GRID Based Approach for Traffic Routing

Since VANET networks are characterized by rapid topology changes and frequent network disconnections, some existing topology-based routing algorithm, such as DSR (Johnson and Maltz, 1996; Lakshmi and

Sankaranarayanan, 2006) and AODV (Perkins and Royer, 1999; Hussain et al., 2007) are not suitable for VANETs. Recent years, position-based routing algorithms have been proposed as a promising routing scheme for VANETs which commonly require the auxiliary information about the geographic position of the participating nodes and do not need the establishment or maintenance of routes. Fig 2 shows the GRID based approach for Distributed approach.

GVGrid (Sun et al., 2006) and HarpiaGrid (Chen et al., 2008) are two grid-based routing algorithms which partition the geographic region into squares of equal-size called grids. GVGrid selects a grid sequence in which nodes are likely to move at similar speeds and toward similar directions as the network route with the best stability. In the route maintenance process, GVGrid does not reconstruct the previous grid sequence but recover route from waiting for nodes moving to the proper positions in the broken grid sequence. This increases the packet delay time severely. In contrast with GVGrid, in cases of low vehicle density, HarpiaGrid will trace back to the grid which is closest to the destination and is not on the same road with the fault grid, then remove all grids after the current grid and generate a new grid forwarding route [14].

Progressed modules for,

1. Simulation Result for Multiple paths Created for movement of vehicles is shown below fig 3.
2. Simulation Result for Route Selection Module for Distributed VANET shown fig 4.
3. Simulation Result for Message Passing between Vehicles are shown in fig 5.

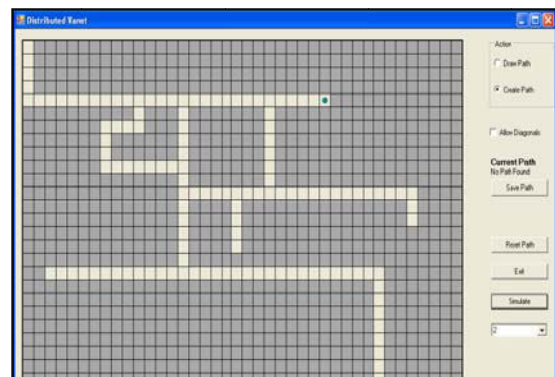


Fig 3: Path Creation Module in Distributed Approach



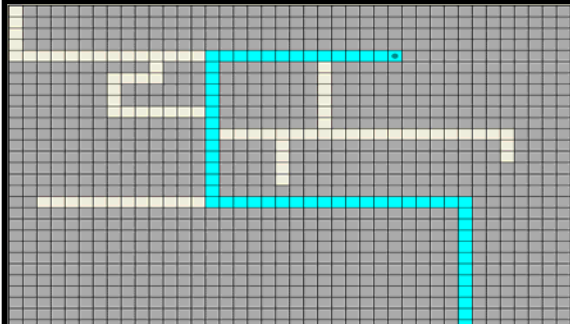


Fig 4: Route Selection Module in Distributed Approach

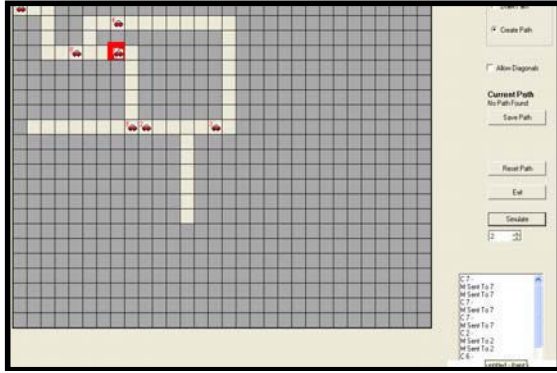


Fig 5: Simulation Result of Message Passing Between Vehicles

### C. Data Representation

The data shared between different vehicles are consist of the *ID* assign to each vehicle to identify their existence, *LOC* to identify the current position of vehicle, *SPD* to know the speed of vehicle, and *BT* to find the time at which the vehicle broadcast the packet.

Each record about another vehicle consists of fields:

- Identification (ID): Uniquely identify the records belonging to different vehicles.
- Location (LOC): The current estimated position of the vehicle.
- Speed (SPD): Used to predict the vehicle's position if no messages containing information about that vehicle is received.
- Broadcast Time (BT): The global time at which the vehicle broadcast that information about itself.

The fig 6 shows Data Flow Structure in Node in Distributed Framework.

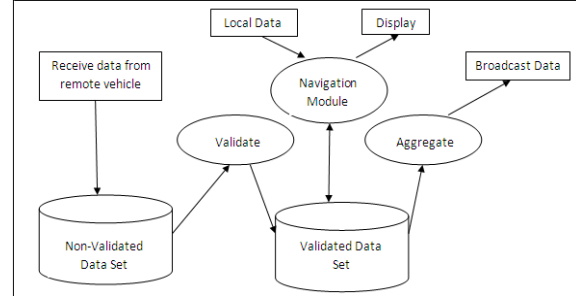


Fig 6: Data Flow Structure in Node in Distributed Framework

### A. Cost-Based Aggregation

Considering the Cost-based algorithm assigns a cost for aggregating each pair of records, and whenever it needs to aggregate two records, the two that correspond to the minimum cost are chosen. Assume two records storing aggregated information about *s1* and *s2* vehicles, with a relative distance of *d1* and *d2*, respectively [12]. The cost of aggregating the two records is calculated as follow:

$$\text{Cost} = \frac{|d1-d2| \times S1 + |d2-da| \times S2}{da}$$

In Cost-Based the aggregation ratios and message portion values are the inputs to the algorithm. For each aggregation ratio and the corresponding portion value, the algorithm starts by continuously selecting the two records that result in the minimum cost, and aggregating them until the number of records is reduced to the value needed by the factor of the aggregation ratio. Afterwards, it writes the contents of the first records in the sorted list to the beginning of the message until they fill the space allocated according to the corresponding portion value. In the next iteration, the same procedure of aggregation and writing is applied to the rest of the records that are not written yet. The aggregation ratio in each iteration is compared with the optimum aggregation ratio to avoid over-aggregation.

A problem that might happen is that as the algorithm proceeds, the number of records left decreases, and the distance between any two consecutive records increases. Hence there is a risk of combining two records that correspond to vehicles that are very far from each other. To avoid this, after the cost is calculated, the algorithm terminates whenever cost is greater than a threshold (cost-threshold.)



### B. Algorithm and Metrics

We will use the following metrics to assess the performance of the algorithms [12]:

**Accuracy:** The road in front of each vehicle is divided into regions of 500 meters long, and the average error in estimating the position of vehicles in each region is calculated. In the accuracy graphs, the average estimation error for each region is shown, averaged over all the nodes during the corresponding simulation.

**Visibility:** We define the visibility of a specific vehicle as the *average* relative distance to the vehicles it knows about. A point  $(d, p)$  on a visibility graph means that  $p\%$  of the vehicles have had a visibility of  $d$  meters or more.

**Knowledge Percentage:** The road in front of each vehicle is divided into regions of 200 meters long. For each region, the percentage of the vehicles in that region about which the current node knows is defined as the knowledge percentage of that node for that region. The knowledge percentage graph presents the knowledge percentage for each region, averaged over all the nodes during a simulation run.

A problem that might happen is that as the algorithm proceeds, the number of records left decreases, and the distance between any two consecutive records increases. Hence there is a risk of combining two records that correspond to vehicles that are very far from each other. To avoid this, after the cost is calculated, the algorithm terminates whenever cost is greater than a threshold (cost-threshold.)

### V. EXPECTED RESULT

The simulation result will show the communication between different vehicles on bidirectional road. We could construct different road scenario for bidirectional road including, cross road. We can increase or decrease the density of vehicles to evaluate the data dissemination complexity. The IEEE 802.11b standard would support communication range for dense as well sparse situation. The Cost-Based aggregation will aggregate the data to carry maximum data in dense situation. Different vehicles will share the data to know the speed, location, and road conditions. The accuracy will be maintained in different road scenario.

### REFERENCES

- [1] W. Kellerer, "(Auto)Mobile Communication in a Heterogeneous and Converged World," IEEE Personal Communications, Vol. 8(6), pp. 41-47, December 2001.
- [2] General Motors Collaborative Laboratory website available online at <http://gm.web.cmu.edu/>.
- [3] C. Schwingenschloegl and T. Kosch, "Geocast Enhancements for AODV in Vehicular Networks," in ACM Mobile Computing and Communications Review, vol. 6(3), pp. 96-97, July 2002.
- [4] R. Miller and Q. Huang, "An Adaptive Peer-to-Peer Collision Warning System," in IEEE Vehicular Technology Conference (VTC), Birmingham, AL, May 2002.
- [5] E. Welsh, P. Murphy, and P. Frantz, "A Mobile Testbed for GPS-Based ITS/IVC and Ad Hoc Routing Experimentation," in International Symposium on Wireless Personal Multimedia Communications (WPMC), Honolulu, HI, October 2002.
- [6] P. Murphy, E. Welsh, and P. Frantz, "Using Bluetooth for Short-Term Ad-Hoc Connections Between Moving Vehicles: A Feasibility Study," in IEEE Vehicular Technology Conference (VTC), pp. 414-418, Birmingham, AL, May 2002.
- [7] Z. D. Chen, H. T. Kung, and D. Vlah, "Ad Hoc Relay Wireless Networks over Moving Vehicles on Highways," in Proceeding of the ACM International Symposium on Mobile Adhoc Networking and Computing, pp. 247-250, October 2001.
- [8] R. Morris, J. Jannotti, F. Kaashoek, J. Li, and D. Decouto, "CarNet: A Scalable Ad Hoc Wireless Network System," in 9th ACM SIGOPS European Workshop, Kolding, Denmark, September 2000.
- [9] J. Li, J. Jannotti, D. S. J. De Couto, D. R. Karger, and R. Morris, in Proceedings of the Sixth Annual International Conference on Mobile Computing and Networking, Boston, MA, August 2000.
- [10] Q. Huang and R. Miller, "The Design of Reliable Protocols for Wireless Traffic Signal Systems". Technical Report WUCS-02-45, Washington University, Department of Computer Science and Engineering, St. Louis, MO.
- [11] I. Chisalita, and N. Shahmehri, "A peer-to-peer approach to vehicular communication for the support of traffic safety applications," in 5<sup>th</sup> IEEE Conference on Intelligent Transportation
- [12] T. Nadeem, S. Dashtinezhad, C. Liao, and L. Iftode, "Traffic View: A Scalable Traffic Monitoring System" "2004 IEEE International Conference on Mobile Data Management (MDM'04)".
- [13] T. Nadeem, S. Dashtinezhad, C. Liao, and L. Iftode, "Traffic View: Traffic Data Dissemination using Car-to-Car Communication" "2004.
- [14] Meimei Huo, Zengwei Zheng, Xiaowei Zhou and Jing Ying PCAR: A Packet-delivery Conditions Aware Routing Algorithm for Vanet Networks



# Adoption of Shortest Path Algorithm (Dijkstra's) in Distributed VANET Using GRID Computing

Amit R. Welekar<sup>1</sup>, Prashant Borkar<sup>2</sup> & S.S. Dorle<sup>3</sup>

<sup>1 & 2</sup> Department of Computer Science and Engineering, GHRCE, Nagpur (MH) -440015, India

<sup>3</sup> Department of Electronics Engineering, GHRCE, Nagpur (MH) -441110, India

---

**Abstract**—this paper proposed approach provides assistance in traffic routing using Grid computing with shortest path algorithm in VANET. VANET is an enhancement in ad-hoc wireless network toward the application of vehicular system. In VANET different types of vehicles communicate with each other to share the speed, location and different parameters in hoping manner. By using shortest path algorithm appropriate nearest path can be selected in VANET.

**Keywords**- GRID, Shortest Path Algorithm, Dijkstra's, Distributed VANET

---

## I. INTRODUCTION

VANET is an enhancement in ad-hoc wireless network toward the application of vehicular system. In VANET different types of vehicles communicate with each other to share the speed, location and different parameters. The size of information shared and message to be sent at distance is varying according to the density of vehicles.

In today's life traffic congestions are unavoidable part. Drivers waste their most of time in traffic congestion. Intelligent transportation systems are accepted to avoid road accidents, to find nearest road segment, and traffic free movement of vehicles.

There are basically two approaches for intelligent transportation system (ITS). Centralised system in which vehicles communicate with road side unit and in Distributed system there is a vehicle to vehicle communication. In centralised system road side units are placed near road side at a density of 2 or 3 per kilometre. In distributed system vehicles are at any distance from each other. So the communication range in distributed should be large as compared to centralised approach.

At a dense situation in distributed VANET the media should carry maximum data. Consider if the average size of data of vehicles in 100 meters are 1K bytes, and 1M bytes in a range of 250 meters. So the media should carry that amount of data at a time. As a density increases the size of the data to be carried increased, so maximum bandwidth should be required.

In this paper we have proposed the idea of shortest path algorithm in which the nodes of the group will be running a distributed Shortest Path algorithm on their

grid environment [1]. This shortest path algorithm will compute shortest paths

from a single source to all the destinations. By having a large collection of cars working together in a grid environment, drivers can easily determine the best route (route with minimal congestion) to reach their destination from their current location.

## II. RELATED WORK

VGrid [2],[3] integrates Vehicular Ad hoc network along with Grid Computing technology to solve a large number of traffic related problems like lane merging, ramp metering etc.

Grid-based approach [4] this paper extended the idea of [2], [3] and proposed a novel approach to traffic routing by establishing a grid environment between cars. Grid Computing is also an emerging discipline and is defined as a flexible, secure and coordinated resource sharing among virtual organizations (individuals, institutions, organizations) [5].

## III. PROBLEM DESCRIPTION

In GRID Computing (Fig. 1).based distributed VANET number of vehicles forming a computational grid to determine the routes for nodes based on any distributed shortest path algorithm [1].



size called grids. GVGrid selects a grid sequence in which nodes are likely to move at similar speeds and toward similar directions as the network route with the best stability. In the route maintenance process, GVGrid does not reconstruct the previous grid sequence but recover route from waiting for nodes moving to the proper positions in the broken grid sequence. This increases the packet delay time severely. In contrast with GVGrid, in cases of low vehicle density, HarpiaGrid will trace back to the grid which is closest to the destination and is not on the same road with the fault grid, then remove all grids after the current grid and generate a new grid forwarding route [6].

The nodes of the group will be running a distributed Shortest Path algorithm on their grid environment [1]. This shortest path algorithm will compute shortest paths from a single source to all the destinations.

#### IV. PROPOSED APPROACH

In this section we present the shortest path algorithm (Dijkstra's algorithm) for finding nearest node (vehicles) by computing from source to all nearby nodes.

For example, if you have this grid, where a \* = obstacle and you can move up, down, left and right, and you start from S and must go to D and 0 = free position:

S	0	0	0
*	*	0	*
*	0	0	*
0	0	*	*
*	0	0	D

You put S in your queue, then "expand" it:

S	1	0	0
*	*	0	*
*	0	0	*
0	0	*	*
*	0	0	D

Then expand all of its neighbours:

S	1	2	0
*	*	0	*
*	0	0	*
0	0	*	*
*	0	0	D

And all of those neighbours' neighbours:

S	1	2	3
*	*	3	*
*	0	0	*
0	0	*	*
*	0	0	D

And so on, in the end you'll get:

S	1	2	3
*	*	3	*
*	5	4	*
7	6	*	*
*	7	8	9

So the distance from S to D is 9. The running time is  $O(NM)$ , where  $N$  = number of lines and  $M$  = number of columns. I think this is the easiest algorithm to implement on grids, and it's also very efficient in practice. It should be faster than a classical Dijkstra's, although Dijkstra's might win if you implement it using heaps.

Dijkstra's algorithm is a graph search algorithm that solves the single-source shortest path problem for a graph with nonnegative edge path costs, producing a shortest path tree. This algorithm is often used in routing and as a subroutine in other graph algorithms [7].

For a given source node, the algorithm finds the path with lowest cost (i.e. the shortest path) between that node and every other node. It can also be used for finding costs of shortest paths from a single node to a single destination node by stopping the algorithm once the shortest path to the destination node has been determined. For example, if the vertices of the graph represent nodes (vehicles) and edge path costs represent driving distances between pairs of nodes connected by a direct road, Dijkstra's algorithm can be used to find the shortest route between one node and all other nodes

#### D. Forward Search algorithm

Forward search find the shortest paths from a given SOURCE node to ALL other nodes, by developing the paths in order of increasing path length. The algorithm Proceeds in stages.

- $N$  = set of nodes in the network;
- $S$  = source node;
- $M$  set of nodes so far incorporated by the algo.

- $l(i,j)$  = link cost from node  $i$  to  $j$ ;  $l(i,i)=0$ ;  $l(i,j)=\infty$  if the two nodes are not directly connected;  $l(i,j)>0$  if they are directly connected;
- $C1(n)$  = least-cost paths from  $S$  to  $n$

**Forward Search Algorithm:**

1. Set  $M=\{S\}$ ; for each node  $n \in N-S$ , set  $C1(n) = l(S, n)$ ;
2. Find  $W \in N-M$  so that  $C1(w)$  in minimum, add  $W$  to  $M$ . Set  $C1(n) = \min \{ C1(n) \text{ and } C1(w) + l(w, n) \}$
3. Repeat 2. Until  $M=N$

**Algorithm Description:**

Let the node at which we are starting be called the initial node. Let the distance of node  $Y$  is the distance from the initial node to  $Y$ . Dijkstra's algorithm will assign some initial distance values and will try to improve them step by step. [7]

1. Assign to every node a tentative distance value: set it to zero for our initial node and to infinity for all other nodes.
2. Mark all nodes except the initial node as unvisited. Set the initial node as current. Create a set of the unvisited nodes called the unvisited set consisting of all the nodes except the initial node.
3. For the current node, consider all of its unvisited neighbors and calculate their tentative distances. For example, if the current node  $A$  is marked with a distance of 6, and the edge connecting it with a neighbor  $B$  has length 2, then the distance to  $B$  (through  $A$ ) will be  $6+2=8$ . If this distance is less than the previously recorded distance, then overwrite that distance. Even though a neighbor has been examined, it is not marked as visited at this time, and it remains in the unvisited set.
4. When we are done considering all of the neighbors of the current node, mark the current node as visited and remove it from the unvisited set. A visited node will never be checked again; its distance recorded now is final and minimal.
5. The next current node will be the node marked with the lowest (tentative) distance in the unvisited set.
6. If the unvisited set is empty, then stop. The algorithm has finished. Otherwise, set the unvisited node marked with the smallest tentative distance as the next "current node" and go back to step 3.

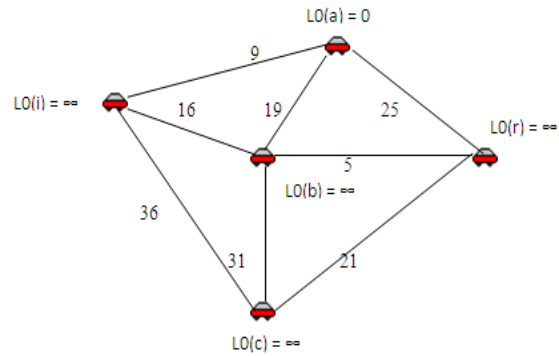


Fig 3: Dijkstra's Shortest Path Finding Algorithm used in VANET

Consider a scenario where 5 vehicles in neighbored with each other. Node 'a' is finding the nearest path to 'c'.

Set initially node 'a' to zero and all other to infinity. Mark all nodes as unvisited except the node 'a' as it is the initial node. Calculate the tentative distance from current node to destination node via intermediate nodes. Here distance from node 'a' to node 'i', 'r', and 'b' are 9, 25, and 19 respectively. Then distance from all intermediate nodes 'i', 'r', and 'b' to destiny node 'c' are 45, 46 and 50 respectively. The minimum value is 45 so the shortest distance from node 'a' to node 'c' is 45 as shown in fig 4.

So vehicle 'a' will send the message to destiny node 'c' via node 'i' as it is the closest and shortest node to node 'c'.

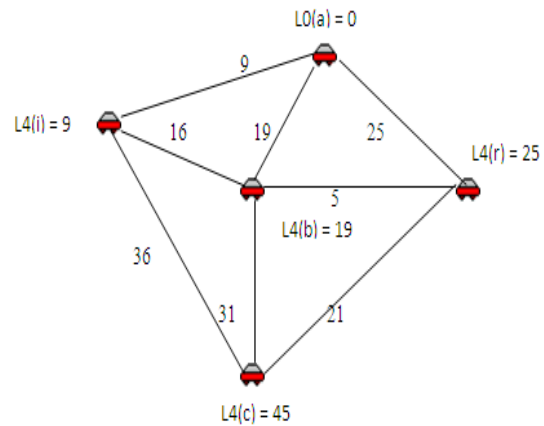


Fig 4: Shortest Path Finding from node 'a' to node 'c'.

## V. CONCLUSION

In this paper we have proposed a shortest path finding algorithm in distributed VANET using Grid Computing based.

The proposed approach of Dijkstra's algorithm is scalable, and for any implementation of node set  $Q$  the running time is,  $O(|E| \cdot dkQ + |V| \cdot emQ)$  where  $dkQ$  and  $emQ$  are times needed to perform decrease key and extract minimum operations in set  $Q$ , respectively. Worst case performance of Dijkstra's algorithm  $O(|E| + |V| \log |V|)$ .

## REFERENCES

- [1] K. M. Chandy and J. Misra, "Distributed Computation on Graphs: Shortest Path Algorithms", Communications of the ACM, November 1982, Volume 25, Number 11
- [2] Jason LeBrun, Joey Anda, Chen-Nee Chuah, Michael Zhang, Dipak Ghosal, "VGrid: Vehicular AdHoc Networking and Computing Grid for Intelligent Traffic Control", IEEE 61<sup>st</sup> Vehicular Technology Conference VTC 2005 Spring, 29th May - 1st June, Stockholm, Sweden
- [3] Andrew Chen, Behrooz Khorashadi, Chen-Nee Chuah, Dipak Ghosal, and Michael Zhang, "Smoothing Vehicular Traffic Flow using Vehicular-based Ad Hoc Networking & Computing Grid (VGrid), IEEE ITSC 2006, September 17-20, 2006, Toronto, Canada
- [4] Noman Islam, Zubair A. Shaikh, Shahnawaz Talpur Center of Research in Ubiquitous Computing National University of Computer & Emerging Sciences Karachi, Pakistan "Towards a Grid-based approach to Traffic Routing in VANET"
- [5] JETS, [http://www.dsg.cs.tcd.ie/dynamic/?category\\_id=-28](http://www.dsg.cs.tcd.ie/dynamic/?category_id=-28), last accessed on January 19, 2008
- [6] Meimei Huo, Zengwei Zheng, Xiaowei Zhou and Jing Ying PCAR: A Packet-delivery Conditions Aware Routing Algorithm for Vanet Networks
- [7] [http://en.wikipedia.org/wiki/Dijkstra's\\_algorithm](http://en.wikipedia.org/wiki/Dijkstra's_algorithm)



# Adaptive Resource Allocation For Wireless Multicast MIMO-OFDM Systems

Shanmugavel G. & Prellly K. E

Department of ECE, DMI College of Engineering, Chennai , India

**Abstract** - Multiple antenna orthogonal frequency division multiple access (OFDMA) is a promising technique for the high downlink capacity in the next generation wireless systems, in which adaptive resource allocation would be an important research issue that can significantly improve the performance with guaranteed QoS for users. Moreover, most of the current source allocation algorithms are limited to the unicast system. In this paper, dynamic resource allocation is studied for multiple antenna OFDMA based systems which provide multicast service. The performance of multicast system is simulated and compared with that of the unicast system. Numerical results also show that the proposed algorithms improve the system capacity significantly compared with the conventional scheme.

**Key words** - Adaptive resource allocation, MIMO, multicast service, OFDM, water-filling.

## I. INTRODUCTION

The next-generation wireless networks are expected to provide broadband multimedia services such as voice, web browsing, video conference, etc. with diverse Quality of Service (QoS) requirements [3]–[5], [7]. Multicast service over wireless networks as in Fig. 1 is an important and challenging goal oriented to many multimedia applications such as audio/video clips, mobile TV and interactive game [3]–[5], [7], [11].

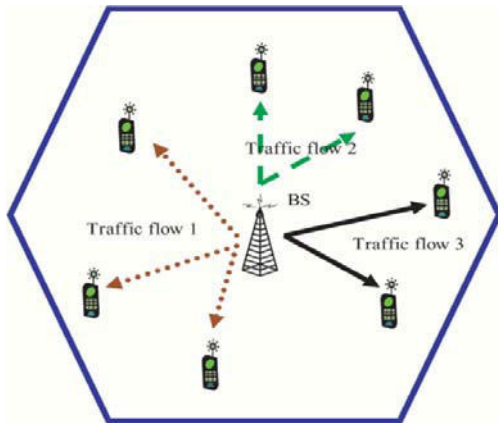


Fig. 1: Cellular structure of multicast transmission system.

There are two key traffics, namely, unicast traffics and multicast traffics, in wireless multimedia communications. Current studies mainly focus on unicast traffics. In particular, dynamic resource allocation has been identified as one of the most efficient techniques to achieve better QoS and higher

system spectral efficiency in unicast wireless networks. Furthermore, more attention is paid to the unicast OFDM systems. Orthogonal Frequency Division Multiplexing (OFDM) is regarded as one of the promising techniques for future broadband wireless networks due to its ability to provide very high data rates in the multi-path fading environment [16]. Orthogonal Frequency Division Multiple Access (OFDMA) is a multiuser version of the popular OFDM scheme and it is also referred as multiuser OFDM. Multiple input multiple output (MIMO) technologies have also received increasing attentions in the past decades. Many broadband wireless networks have now included MIMO technology in their protocols including the multicast system [3]. Compared to single input single output (SISO) system, MIMO offers the higher diversity which can potentially lead to a multiplicative increase in capacity. In multiuser OFDM or MIMO-OFDM systems, dynamic resource allocation always exploits multiuser diversity gain to improve the system performance [8]–[10], [6], [15] and it is divided into two types of optimization problems: 1) to maximize the system throughput with the total transmission power constraint [9]; 2) to minimize the overall transmit power with constraints on data rates or Bit Error Rates (BER) [15]. To the best of our knowledge, most dynamic resource allocation algorithms, however, only consider unicast multiuser OFDM systems. In wireless networks, many multimedia applications adapt to the multicast transmission from the base station (BS) to a group of users. These targeted users consist of a multicast group which receives the data packets of the same traffic flow. The simultaneously achievable transmission rates to these users were investigated in [12] and [13]. Recently



scientific researches of multicast transmission in the wireless networks have been paid more attention. For example, proportional fair scheduling algorithms were developed to deal with multiple multicast groups in each time slot in cellular data networks [2]. The dynamic resource allocation for OFDM based multicast system was researched in [1], however it focused on SISO system and can not be applied to MIMO system directly.

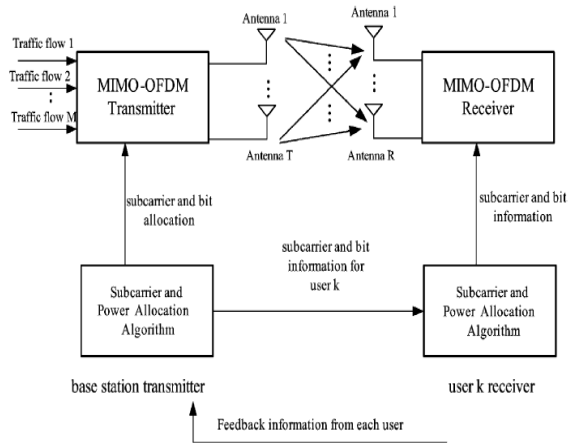


Fig. 2 : Block diagram of multiple antenna OFDM multicast system.

On the other hand, the conventional scheme in current standards such as IEEE 802.16 or 3GPP LTE for multicast service considers the worst user very much, which may waste the resource. In this paper, we propose dynamic subcarrier and power allocation algorithms for MIMO OFDMA-based wireless multicast systems. In the proposed algorithms, the subcarriers and powers are dynamically allocated to the multicast groups. Our aim is to maximize the system throughput given the total power constraint. Let us assume that there are multiple multicast groups in a cell and each multicast group may contain a different number of users. The users included in the same multicast group are called co-group users and these can be located in different places in the cell.

This paper is organized as follows. Section II introduces the multiple antenna OFDMA based multicast system model and presents the optimization objective function. In Section III, the proposed resource allocation algorithm is described. Simulation results are illustrated in Section IV and conclusions are drawn in Section V.

## II. SYSTEM MODEL

The block diagram of multiuser MIMO-OFDM downlink system model is shown in Fig. 2. It shows that in the base station channel state information of each

couple of transmit and receive antennas are sent to the block of subcarrier and power algorithm through the feedback channels. The resource allocation information is forwarded to the MIMO-OFDM transmitter. The transmitter then selects the allocated number of bits from different users to form OFDMA symbols and transmits via the multiple transmit antennas. The spatial multiplexing mode of MIMO is considered. The resource allocation scheme is updated as soon as the channel information is collected and also the subcarrier and bit allocation information are sent to each user for detection.

The following assumptions are used in this paper. The transmitted signals experience slowly time-varying fading channel, therefore the channel coefficients can be regarded as constants during the subcarrier allocation and power loading period. Throughout this paper, let the number of transmit antennas be  $T$  and the number of receive antennas be  $R$  for all users. Denote the number of traffic flows as  $M$ , the number of user as  $K$  and the number of subcarriers as  $N$ . Thus in this model downlink traffic flows are transmitted to users over subcarriers. Assume that the base station has total transmit power constraint  $Q$ . The objective is to maximize the system sum capacity with the total power constraint. We use the equally weighted sum capacity as the objective function. The system capacity optimization problem for multicast MIMO-OFDM system can be formulated to determine the optimal subcarrier allocation and power distribution:

$$\max C = \frac{1}{N} \sum_{k=1}^K \sum_{n=1}^N p_{k,n} \left( \sum_{i=1}^{M_{k,n}} \log \left( 1 + \frac{\lambda_{k,n} q_{k,n}}{N_0} \right) \right) \quad (1)$$

$$\text{subject to : } \sum_{n=1}^N \max_k (q_{k,n}) \leq Q$$

$$q_{k,n} \geq 0 \text{ for all } k, n$$

$$p_{k,n} \in \{0,1\} \text{ for all } k, n$$

Where  $C$  is the system sum capacity which can be derived based on [16] and the above assumptions;  $Q$  is the total available power;  $q_{k,n}$  is the power assigned to user in the subcarrier  $n$ ;  $p_{k,n}$  can only be the value of 1 or 0 indicating whether subcarrier  $n$  is used by user  $K$  or not.  $M_{k,n}$  is the rank of  $H_{k,n}$  which denotes the MIMO channel gain matrix ( $R \times T$ ) on subcarrier  $n$  for user  $K$  and  $\{\lambda_{k,n}^{(i)}\}$   $i = 1:M_{k,n}$  are eigenvalues of  $H_{k,n} H_{k,n}^H$ ;  $K_n$  is the allocated user index on subcarrier  $n$ ;  $N_0$  is the noise power in the frequency band of one subcarrier.

The different point of multicast optimization problem in (1) compared to the general unicast system is

that there is no constraint of  $\sum_{k=1}^K P_{k,n} = 1$  for all  $n$ , which means that many users can share the same subcarrier in multicast system because they may need the same multimedia contents. The capacity for user  $K$ , denoted as  $R_k$ , is defined as

$$R_k = \frac{1}{N} \sum_{n=1}^N P_{k,n} \left( \sum_{i=1}^{M_{k,n}} \log \left( 1 + \frac{\lambda_{k,n}^{(i)} q_{k,n}}{N_0} \right) \right) \quad (2)$$

### III. PROPOSED SUBOPTIMAL SUBCARRIER ALLOCATION AND POWER DISTRIBUTION

The optimization problem in (1) is generally very hard to solve. It involves both continuous variables and binary variables. Such an optimization problem is called a mixed binary integer programming problem. Furthermore, since the feasible set is not convex the nonlinear constraints in (1) increase the difficulty in finding the optimal solution. Ideally, subcarriers and power should be allocated jointly to achieve the optimal solution in (1). However, this poses a prohibitive computational burden at the base station in order to reach the optimal allocation. Furthermore, the base station has to rapidly allocate the optimal subcarrier and power in the time varying wireless channel. Hence, low-complexity suboptimal algorithms are preferred for practical implementations. Separating the subcarrier and power allocation is a way to reduce the complexity, because the number of variables in the objective function is almost reduced by half. In an attempt to avoid the full search algorithm in the preceding section, we devise a suboptimum two-step approach. In the first step, the subcarriers are assigned assuming the constant transmit power of each subcarrier. This assumption is used only for subcarrier allocation. Next, power is allocated to the subcarriers assigned in the first step. Although such a two-step process would cause suboptimality of the algorithm, it makes the complexity significantly low. In fact, such a concept has been already employed in OFDMA systems and also its efficacy has been verified in terms of both performance and complexity. However, the algorithm proposed in this paper is unique in dealing with MIMO-OFDM based multicast resource allocation.

Before we describe the proposed suboptimal resource allocation algorithm, we firstly show mathematical simplifications for the following subcarrier allocation. It is noticed that in large SNR region, i.e.  $\lambda_{k,n}^{(i)} q_{k,n} / N_0 \gg 1$ , we get the following approximation:

$$\arg \min_k \sum_{i=1}^{M_{k,n}} \log \left( 1 + \frac{\lambda_{k,n}^{(i)} q_{k,n}}{N_0} \right)$$

$$= \arg \min_k \prod_{i=1}^{M_{k,n}} \left( 1 + \frac{\lambda_{k,n}^{(i)} q_{k,n}}{N_0} \right)$$

$$\approx \arg \min_k \prod_{i=1}^{M_{k,n}} \left( \frac{\lambda_{k,n}^{(i)} q_{k,n}}{N_0} \right)$$

$$= \arg \min_k \prod_{i=1}^{M_{k,n}} \lambda_{k,n}^{(i)} \text{ when } M_{1,n} = \dots = M_{K,n} = M \quad (3)$$

Where  $\arg \min_k \prod_{i=1}^{M_{k,n}} \lambda_{k,n}^{(i)}$  is named as product-criterion which tends to be more accurate when the SNR is high. On the other hand, in small SNR region, i.e.  $\lambda_{k,n}^{(i)} q_{k,n} / N_0 \ll 1$ , using  $\log(1+x) = x$ , we get

$$\arg \min_k \sum_{i=1}^{M_{k,n}} \log \left( 1 + \frac{\lambda_{k,n}^{(i)} q_{k,n}}{N_0} \right)$$

$$\approx \arg \min_k \sum_{i=1}^{M_{k,n}} \left( \frac{\lambda_{k,n}^{(i)} q_{k,n}}{N_0} \right)$$

$$= \arg \min_k \sum_{i=1}^{M_{k,n}} \left( \sum_{i=1}^{M_{k,n}} \lambda_{k,n}^{(i)} \right) \frac{q_{k,n}}{N_0}$$

$$= \arg \min_k \sum_{i=1}^{M_{k,n}} \lambda_{k,n}^{(i)} \quad (4)$$

Where  $\arg \max_k \sum_{i=1}^{M_{k,n}} \lambda_{k,n}^{(i)}$  is named as sum-criterion which is more accurate when the SNR is low. These two approximations will be used in the suboptimal algorithm for the high SNR and low SNR cases, respectively. In this way, we can reduce the complexity significantly with minimal performance degradation.

The steps of the proposed suboptimal algorithm are as follows:

- Step 1 Assign the subcarriers to the users in a way that maximizes the overall system capacity;
- Step 2 Assign the total power to the allocated subcarriers using the multi-dimension water-filling algorithm.

#### A. Step 1—Subcarrier Assignment

For a given power allocation vector  $q = (q_1, q_2, \dots, q_n)$  for each subcarrier, RA optimization problem of (1) is separable with respect to each subcarrier. The subcarrier problem with respect to subcarrier is

$$\text{Max} R(n) = \sum_{k=1}^K P_{k,n} \left( \sum_{i=1}^{M_{k,n}} \log \left( 1 + \frac{\lambda_{k,n}^{(i)} q_{k,n}}{N_0} \right) \right)$$

$$\text{Subject to: } \max_k \{q_{k,n}\} \leq q_n$$

$$P_{k,n} = \{0,1\} \text{ for all } k,n \quad (5)$$

Then the multicast subcarrier allocation algorithm based on (3) for each subcarrier is given as follows.

- 1) For the  $t$ th subcarrier, calculate the current total data rate when the  $t$ th user is selected as the user who has lowest eigenvalue product

$$R(n) = N_{k,n} \sum_{i=1}^{M_{k,n}} \log \left( 1 + \frac{\lambda_{k,n}^{(i)} q_n}{N_0} \right) \quad (6)$$

- 2) For the  $n$ th subcarrier, select the user index  $kn$  which can maximize

$$Kn = \arg \max_k R(n) \quad (7)$$

Then we have

$$P_{k,n} = \begin{cases} 1, & \prod_{i=1}^{M_{k,n}} \lambda_{k,n}^{(i)} \geq \prod_{i=1}^{M_{k,n}} \lambda_{k,n}^{(i)} \\ 0 & \text{other wise} \end{cases} \quad (8)$$

For the low SNR case, the product-criterion (3) is changed into the sum-criterion (4) for this step's subcarrier allocation.

#### B. Step 2 Power Allocation

The subcarrier algorithm in step 1 is not optimum because equal power distribution for the subcarriers is assumed. In this step, we propose an efficient power allocation algorithm based on the subcarrier allocation in step 2. Corresponding to each subcarrier, there may be several users to share it for the multicast service. In this case, the lowest user's channel gain on that subcarrier among the selected users in step 1 will be used for the power allocation. The multi-dimension water-filling method is applied to find the optimal power allocation as follows. The power distribution over subcarriers is

$$q_n^* = \max(0, q_n)$$

where  $q_n$  means the power assigned to each antenna of subcarrier  $n$  and it is the root of the following equation,

$$\sum_{i=1}^{M_{K_n,n}} \frac{\lambda_{k,n}^{(i)}}{\lambda_{k,n}^{(i)} q_n + N_0} + \alpha = 0, n = 1, 2, \dots, N, \quad (9)$$

Where  $K_n$  is the allocated user index on subcarrier  $\alpha$ ; is the water-filling level which satisfies  $\sum_{n=1}^N M_{k,n} q_n^* = Q$  where  $Q$  and  $N$  are the total power and the number of subcarriers, respectively.

In case of  $T = R = 1$ , that is, a single antenna system, the optimal power distribution for the subcarriers is transformed into the standard water-filling solution:

$$q_n^* = \left( -\frac{1}{\alpha} - \frac{N_0}{\lambda_{k,n}^{(1)}} \right)^+ = \left( -\frac{1}{\alpha} - \frac{N_0}{H_{k,n}^{(1)}} \right)^+ \quad (10)$$

Where  $(x)^+ = \max(0, x)$  and  $\lambda_{k,n}^{(1)}$  is the same as  $H_{k,n}^{(1)}$  for a single antenna. The multi-dimension water-filling algorithm is an iterative method, by which we can find the optimal power distribution to realize the maximum of system capacity.

#### IV. SIMULATION RESULTS

In this section, simulation results are presented to demonstrate the performance of the proposed algorithm. The simulation parameters of the proposed system are given in

Table I.

The Simulation Parameters for the MIMO-OFDM Systems.

Number of subcarriers	64
Number of transmitter antenna	2
Number of receive antenna	2
Number of users	4
Bandwidth	1 MHz
Max. transmit power	1W
AWGN PSD	-100dBW/Hz - 80 dBW/Hz
BER	1e-3
Number of multipaths	6

wireless channel is modeled as a frequency selective channel consisting of six independent Rayleigh multipaths. Each multipath component is modeled by Clarke's flat fading model. The number of users is 4 and the number of antennas is  $T = R = 2$ . Each couple of transmit antenna and receive antenna is assumed to be independent to the other couples. Total transmit power is 1 W and AWGN power spectral density varies from 85 dBW/Hz to 60 dBW/Hz. The total bandwidth  $B$  is 1 MHz, which is divided into 64 subcarriers. The capacities in the following figures are averaged over 10000 channel realizations.

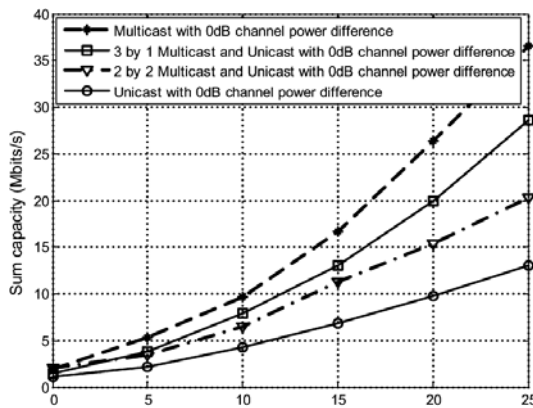


Fig. 3 : Sum capacity comparison of multicast and unicast systems.

#### A. Comparison of Multicast and Unicast Systems:

In Fig. 3, the sum capacities of multicast and unicast schemes are shown for multiple antenna OFDM systems. Here it is supposed that there is no channel power difference between the users. In the multicast system, it is supposed that 4 users receive the same contents, while in the unicast system the contents of users are different from each other. 3 by 1 multicast and unicast system means that 3 users receive the same contents as one group and the left one user receives different content. And 2 by 2 multicast and unicast system means that 2 users receive the same contents as one group and the left two users are unicast users. It is noticed that the multicast scheme with the proposed method can achieve higher capacity than the unicast scheme or the mixed cases. The more multicast users exit, the higher system capacities can be achieved. For the Fairness of comparison, on each subcarrier the user with the highest eigenvalue product is selected to use it in the unicast system.

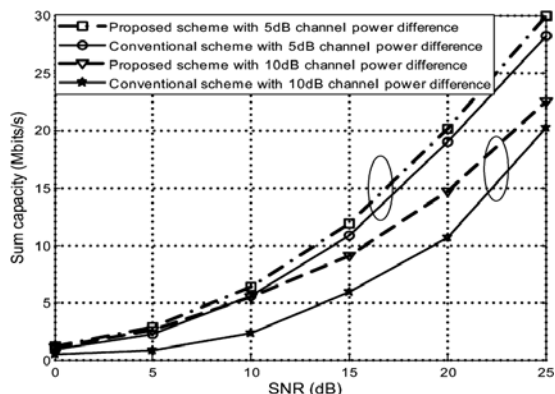


Fig. 4 : Sum capacity comparison of proposed scheme and conventional one in multicast system when the SNR is high.

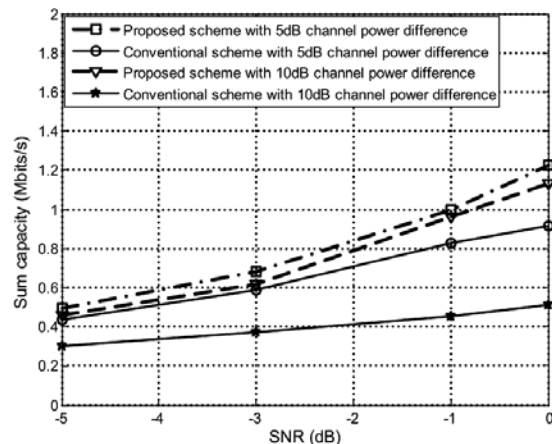


Fig. 5 : Sum capacity comparison of proposed scheme and conventional one in multicast system when the SNR is low.

#### B. Comparison of Proposed Scheme and Conventional One

In this subsection, the sum capacities of the proposed scheme and conventional scheme for the multicast system are shown in Figs. 4 and 5 for the high SNR and low SNR cases, respectively. It is supposed that 4 users receive the same contents and there is 5 dB or 10 dB average channel power difference between the users. In the conventional scheme, on each subcarrier the user with the lowest eigenvalue product is selected as the baseline to transmit the information. From both Figs. 4 and 5, it is noticed that the proposed method can achieve higher capacity than the conventional one. The more average channel power difference between the users, the larger gains can be obtained. This means that the proposed adaptive subcarrier and power allocation algorithm is more effective in the presence of higher channel or link difference. This is because the drawback of the conventional scheme is more evident in the case of higher channel or link difference.

## V. CONCLUSION

This paper presented a new method to solve the subcarrier and power allocation problem for multi-user MIMO-OFDM based multicast system. The optimization problem was formulated to maximize the system capacity with a total transmit power constraint. Due to the complexity of optimal algorithm, two step suboptimal algorithm was proposed. The proposed subcarrier allocation algorithm determined the number of users for each subcarrier based on the maximization criteria, in which the capacity of each subcarrier can be maximized. Then the proposed power allocation scheme adopted multi-dimension water-filling method in order to maximize the system capacity. Simulation results

showed that the system capacity of the proposed scheme is significantly improved as compared with the conventional one.

## REFERENCES

- [1] J. Liu, W. Chen, Z. Cao, and K. B. Letaief, "Dynamic power and subcarrier allocation for OFDMA-based wireless multicast systems," in *Proc. IEEE ICC 2008*, May 2008, pp. 2607–2611.
- [2] H. Won and H. Cai, "Multicast scheduling in cellular data networks," in *Proc. IEEE Infocom 2007*, May 2007, pp. 1172–1180.
- [3] A. Correia, J. Silva, N. Souto, L. Silva, A. Boal, and A. Soares, "Multi-resolution broadcast/multicast systems for MBMS," *IEEE Trans. Broadcasting*, vol. 53, no. 1, pp. 224–234, Mar. 2007.
- [4] F. Hartung, U. Horn, J. Huschke, M. Kampmann, T. Lohmar, and M. Lundevall, "Delivery of broadcast services in 3G networks," *IEEE Trans. Broadcasting*, vol. 53, no. 1, pp. 188–199, Mar. 2007.
- [5] M. Chari, F. Ling, A. Mantravadi, R. Krishnamoorthi, R. Vijayan, G. Walker, and R. Chandhok, "FLO physical layer: An overview," *IEEE Trans. Broadcasting*, vol. 53, no. 1, pp. 145–160, Mar. 2007.
- [6] Y. Ben-Shimol, I. Kitroser, and Y. Dinitz, "Two-dimensional mapping for wireless OFDMA systems," *IEEE Trans. Broadcasting*, vol. 52, no. 3, pp. 388–396, Sep. 2006.
- [7] S. Y. Hui and K. H. Yeung, "Challenges in the migration to 4G mobile systems," *IEEE Commun. Mag.*, vol. 41, pp. 54–56, Dec. 2003.
- [8] T. C. Alen, A. S. Madhukumar, and F. Chin, "Capacity enhancement of a multi-user OFDM system using dynamic frequency allocation," *IEEE Trans. Broadcasting*, vol. 49, no. 4, pp. 344–353, Dec. 2003.
- [9] M. Ergen, S. Coleri, and P. Varaiya, "QoS aware adaptive resource allocation techniques for fair scheduling in OFDMA based broadband wireless access systems," *IEEE Trans. Broadcasting*, vol. 49, no. 4, pp. 362–370, Dec. 2003.
- [10] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 171–178, Feb. 2003.
- [11] U. Varshney, "Multicast over wireless networks," *Communications of the ACM*, vol. 45, pp. 31–37, Dec. 2002.
- [12] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.
- [13] L. Li and A. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels: Part I: Ergodic capacity," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1083–1102, Mar. 2001.
- [14] I. Emre Telatar, "Capacity of Multi-antenna Gaussian channels," *European Trans. Telecommunications*, vol. 10, pp. 585–595, Nov. 1999.
- [15] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit and power allocation," *IEEE J. Select. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.
- [16] L. J. Cimini and N. R. Sollenberger, "OFDM with diversity and coding for advanced cellular internet services," in *Proc. IEEE Globecom'1997*, Nov. 1997, pp. 305–309.



# Design And Analysis of Fractal Planar Hexagonal Array With Low Side Lobes

L. Rambabu<sup>1</sup>, B.Ramarao<sup>2</sup> & P. V. Sridevi<sup>3</sup>

<sup>1&2</sup>Dept of ECE, AITAM Tekkali, A.P., India

<sup>3</sup>Dept. of ECE, A.U. College of Engineering, Visakhapatnam, A.P., India

---

**Abstract** - The hexagonal arrays are becoming increasingly popular, especially for their applications in the area of wireless communications. The overall objective of this paper is to use the theoretical foundation developed for the analysis of radiation patterns and design of the hexagonal arrays. A technique has been developed for the analysis of radiation patterns from concentric ring arrays. A family of function, known as generalized Weierstrass functions, has been shown to play a key role in the theory of fractal radiation pattern analysis.

**Key words** - fractal antenna arrays, hexagonal array, fractal antenna radiation patterns, low side lobe antennas radiation patterns.

---

## I. INTRODUCTION

The name "fractal", from the Latin 'fractus' meaning broken, was given to highly irregular shapes by Benoit Mandelbrot in his foundational essay in 1975 [1]. Since then, fractal geometry has attracted widespread, and some times controversial, attention. The subject has grown on two fronts: on the one hand, many "real fractals" of science and nature have been identified. On the other hand, the mathematics that is available for studying fractal sets, much of which has its roots in geometric measure theory, has developed enormously with new tools emerging for fractal analysis. This paper concerned with the mathematics of fractals and application to antenna theory [2].

## II. THEORY

### *Concentric circular ring array:*

The technique is very general, and consequently provides much more flexibility in the design of fractal arrays when compared to other. The generator, in this is based on a concentric circular ring array.

The generating array factor for the concentric circular ring array may be expressed in the form [2]

$$GA(\theta, \phi) = \sum_{y=1}^Y \sum_{x=1}^X I_{yx} e^{j\alpha_{yx}(\theta, \phi)} \quad (1)$$

$$\alpha_{yx}(\theta, \phi) = kr_y \sin\theta \cos(\phi - \phi_{yx}) + \beta_{yx} \quad (2)$$

$$K = 2\pi / \lambda$$

Y = Total number of concentric rings

X<sub>y</sub> = Total number of elements on the y<sup>th</sup> ring

R<sub>y</sub> = Radius of the y<sup>th</sup> ring

I<sub>yx</sub> = Excitation current amplitude of the x<sup>th</sup> element on the y<sup>th</sup> ring located at  $\phi = \phi_{yx}$

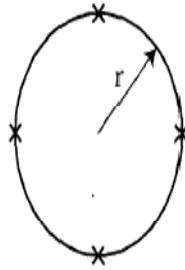
$\alpha_{yx}$  = Excitation current phase of the x<sup>th</sup> element on the y<sup>th</sup> ring located at  $\phi = \phi_{yx}$

A wide variety of interesting as well as practical fractal array designs may be constructed using a generating subarray of the form given in equation (1). The fractal array factor for a particular stage of growth P may be derived directly from equation (1) by following a procedure similar to that outlined in the previous section. The resulting expression for the array factor was found to be

$$AF_P(\theta, \phi) = \prod_{p=1}^P \left\{ \sum_{y=1}^Y \sum_{x=1}^{X_y} e^{j\delta^{p-1} \alpha_{yx}(\theta, \phi)} \right\} \quad (3)$$

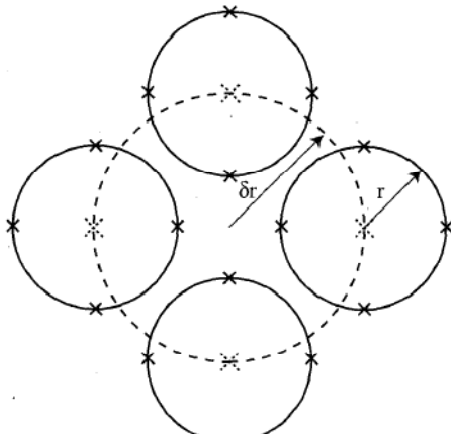
where  $\delta$  represents the scaling or expansion factor associated with the fractal array. A graphical procedure, which is embodied in equation (3), can be used to conveniently illustrate the construction process for fractal arrays. For example, suppose we consider the simple four element circular array of radius r, shown in Figure 2.1(a). If we regard this as the generator (stage 1) for a fractal array, then the next stage of growth (stage-2) for the array would have a geometrical configuration of the form shown in Figure 2.1(b). Hence the first step

in the construction process, as depicted in Figure 2.1 is to expand the four-element generator array by a factor of  $\delta$ .



Stage 1

Fig. (a)



Fig(b).

Stage 2

Figure 2.1: Construction of four-element circular subarray generator: (a) Stage 1 ( $P = 1$ ) and (b) stage 2 ( $P = 2$ ).

This is followed by replacing each of the elements of the expanded array by an exact copy of the original unscaled four-element circular subarray generator. The entire process is then repeated in a recursive fashion, until the desired stage of growth for the fractal array is reached.

It is convenient for analysis purposes to express the fractal array factor, above equation in the following normalized form.

$$AF_p(\theta, \phi) = \prod_{p=1}^P \left\{ \frac{\sum_{y=1}^Y \sum_{x=1}^X I_{yx} e^{j\delta^{p-1} \alpha_{yx}(\theta, \phi)}}{\sum_{y=1}^Y \sum_{x=1}^X I_{yx}} \right\} \quad (4)$$

$$|AF_p(\theta, \phi)| = \prod_{p=1}^P \left| \frac{\sum_{y=1}^Y \sum_{x=1}^{X_y} I_{yx} e^{j\delta^{p-1} \psi_{yx}(\theta, \phi)}}{\sum_{y=1}^Y \sum_{x=1}^{X_y} I_{yx}} \right| \quad (5)$$

Which has a corresponding representation in terms of decibels given by

$$|AF_p(\theta, \phi)|_{dB} = 20 \sum_{p=1}^P \log \left| \frac{\sum_{y=1}^Y \sum_{x=1}^{X_y} I_{yx} e^{j\delta^{p-1} \psi_{yx}(\theta, \phi)}}{\sum_{y=1}^Y \sum_{x=1}^{X_y} I_{yx}} \right| \quad (6)$$

For the special case when  $\delta=1$  eqs.5 and 6 reduce to

$$|AF_p(\theta, \phi)| = \left| \frac{\sum_{y=1}^Y \sum_{x=1}^{X_y} I_{yx} e^{j\psi_{yx}(\theta, \phi)}}{\sum_{y=1}^Y \sum_{x=1}^{X_y} I_{yx}} \right|^P \quad (7)$$

$$|AF_p(\theta, \phi)|_{dB} = 20P \log \left| \frac{\sum_{y=1}^Y \sum_{x=1}^{X_y} I_{yx} e^{j\psi_{yx}(\theta, \phi)}}{\sum_{y=1}^Y \sum_{x=1}^{X_y} I_{yx}} \right| \quad (8)$$

Another unique property of equation (4) is the fact that the conventional co-phasal excitation

$$\alpha_{yx} = -kr_y \sin \theta_0 \cos(\phi - \phi_{yx}) \quad (9)$$

Where  $\theta_0$  and  $\phi_0$ , are the desired main-beam steering angles, can be applied at the generating subarray level. From this, the position of the main beam produced by equation (4) is independent of the stage of growth  $P$ , since it corresponds to a value of  $\psi_{yx} = 0$  (i.e., when  $\theta = \theta_0$  and  $\phi = \phi_0$ ). In other words, once the position of the main beam is determined for the generating subarray, it will remain invariant at all higher stages of growth.



### EXAMPLES :

There are several different examples of recursively generated arrays which have common fact that they may be constructed via a concentric circular ring subarray generator. Hence, the mathematical expressions that describe the radiation patterns of these arrays are all special cases of above equation (4).

### III. HEXAGONAL ARRAYS

One type of planar array configuration in common use is the hexagonal array. These arrays are becoming increasingly popular, especially for their applications in the area of wireless communications. The standard hexagonal arrays are formed by placing elements in an equilateral triangular grid with spacings  $d[3]$ . These arrays can also be viewed as consisting of a single element located at the center, surrounded by several concentric six-element circular arrays of different radii. This property has been used to derive an expression for the hexagonal array factor [3].

Hexagonal arrays may be realized via a construction process based on the recursive application of a generating subarray. Consider the uniformly excited six-element circular generating subarray of radius  $r=\lambda/2$ . This particular value of radius was chosen so that the six elements in the array correspond to the vertices of a hexagon with half-wavelength sides ( $d=\lambda/2$ ). Consequently, the array factor associated with this six-element generating subarray may be shown to have the following representation

$$AF_p(\theta, \phi) = \frac{1}{6^p} \prod_{p=1}^P \sum_{x=1}^6 e^{j\delta^{p-1}[\pi \sin \theta \cos(\phi - \phi_x) + \beta_x]} \quad (10)$$

$$\text{Where } \phi_n = (x-1)\frac{\pi}{3}$$

$$\beta_x = -\pi \sin \theta_0 \cos(\phi_0 - \phi_x)$$

The array factor expression given in equation (10) may also be written in the form

$$AF_p(\theta, \phi) = \frac{1}{6^p} \prod_{p=1}^P \sum_{x=1}^6 e^{j\delta^{p-1}\alpha_x(\theta, \phi)} \quad (11)$$

$$\alpha_x(\theta, \phi) = \pi[\sin \theta \cos(\phi - \phi_x) - \sin \theta_0 \cos(\phi_0 - \phi_x)] \quad (12)$$

The special case where the expansion factor of the recursive hexagonal array is assumed to be unity (i.e.

$\delta=1$ ) Under these circumstances, equation (11) reduces to

$$AF_p(\theta, \phi) = \left[ \frac{1}{6} \sum_{x=1}^6 e^{j\alpha_x(\theta, \phi)} \right]^P \quad (13)$$

These arrays increase in size at a rate that obeys the relationship

$$X_p = 3P(P+1) + (1 - \delta_{p1}) \quad (14)$$

Where  $\delta_{p1}$  represents the Kronocker delta function, defined by

$$\delta_{p1} = 1 \quad P=1 \quad 0 \quad P \neq 1$$

In other words, every time this fractal array evolves from one stage to the next, the number of concentric hexagonal subarrays contained in it increases by one.

The second special case of interest to be considered in this results when a choice of  $\delta = 2$  is made. Substituting this value of  $\delta$  into equation (15) yields an expression for the recursive hexagonal array factor given by

$$AF_p(\theta, \phi) = \frac{1}{6^p} \prod_{p=1}^P \sum_{x=1}^6 e^{j2^{p-1}\alpha_x(\theta, \phi)} \quad (15)$$

where

$$X_p = 3[2^p(2^p - 1) - 2^{p-1}(2^{p-1} - 1)]$$

Clearly, by comparing equation (15) with equation (14), it comes to conclusion that these recursive arrays will grow at a much faster rate than those generated by a choice of  $\delta = 1$ . The first four stages in the construction process of these arrays the element locations correspond to the vertices of the hexagons and the element locations and associated current distributions for each of the four hexagonal arrays. The hexagonal arrays that result from the recursive construction process with  $\delta = 2$  have some elements missing, i.e., they are thinned.

This is a potential advantage of these arrays from the design point of view, since they may be realized with fewer elements. Another advantage of these arrays is that they possess low sidelobe levels, as indicated by the set of radiation pattern slices for  $\phi = 90^\circ$ . Finally, we note that the compact product form of the array factor given in equation (15) offers a significant advantage in terms of computational efficiency, especially for large arrays. This is a direct consequence of the recursive nature of these arrays, and may be exploited to develop rapid beam-forming algorithms.

When an element with two units of current is added to the center of the hexagonal generating subarray, the expression for the array factor given in equation (10) must be modified in the following way:

$$AF_p(\theta, \phi) = \frac{1}{8^p} \prod_{p=1}^P \left\{ 2 + \sum_{x=1}^6 e^{j\delta^{p-1} \alpha_x(\theta, \phi)} \right\}$$

By taking different values of  $\delta$  (i.e  $\delta=1,2,3,4..$ ) we will get different cases of hexagonal array.

#### IV. RESULTS

The radiation pattern for the different stages ( $P=1,2,3,4$ ) the radiation pattern plots for the hexagonal arrays and fully populated hexagonal arrays (two cases i.e.  $\delta=1$  and  $\delta=2$ ) are obtained by MATLAB are shown in figure 1,2 and 3. The radiation pattern indicate that a further reduction in side lobe levels may be achieved by including a central element in the generating sub array. Finally, plots of the radiation intensity for this array are shown in figure where the phasing of the generating sub array has been chosen so as to produce a main-beam maximum at  $\theta_0 = 45^\circ$  and  $\phi_0 = 90^\circ$

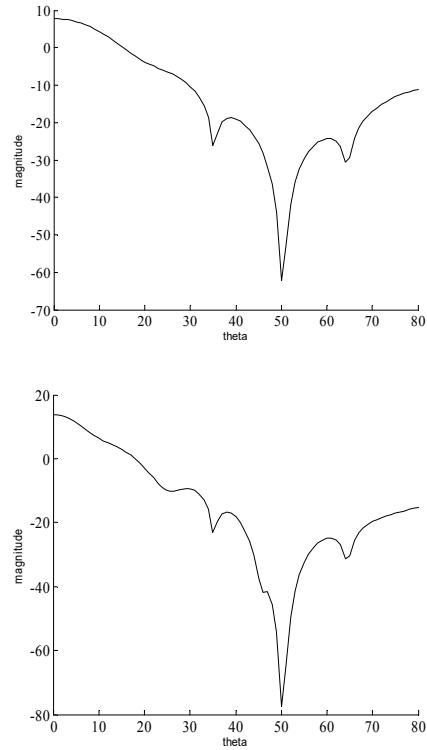
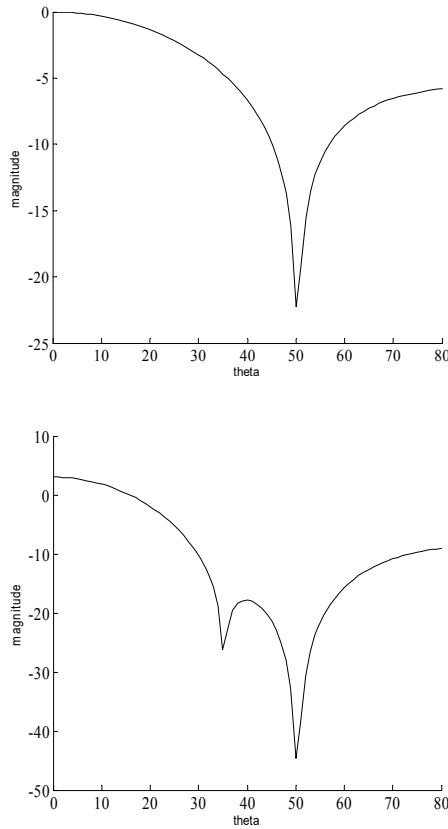
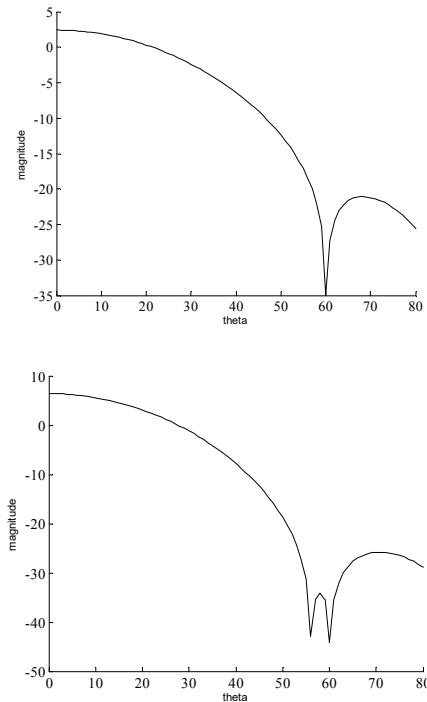


Fig. 1 : Radiation patterns of hexagonal array for  $p=1, p=2, p=3, p=4$ .



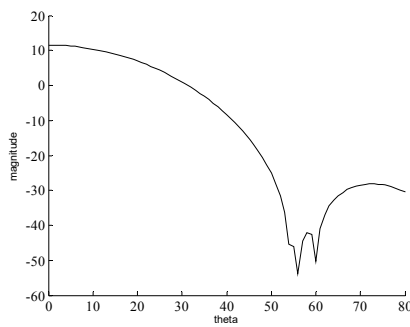


Fig. 2 : Radiation patterns of fully populated hexagonal array for  $p=1,2,3\&4$ .

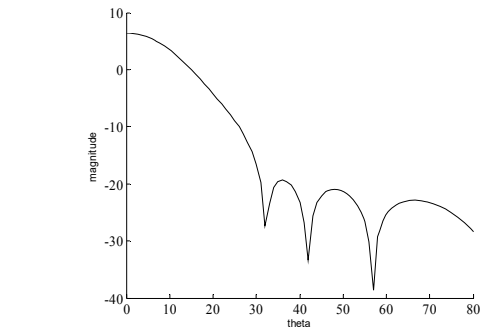
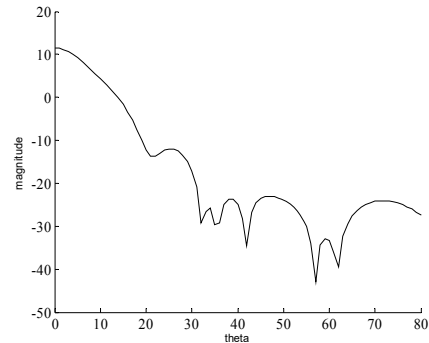
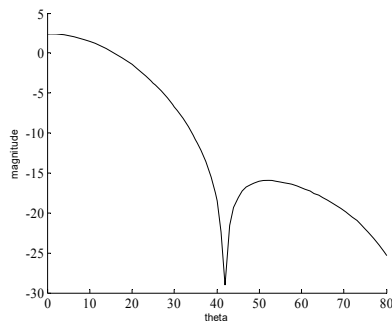
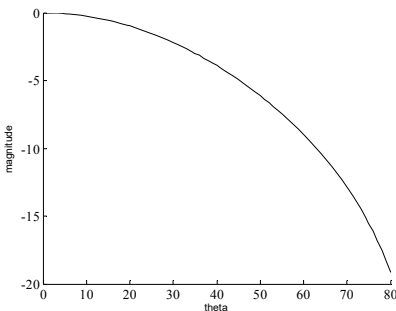
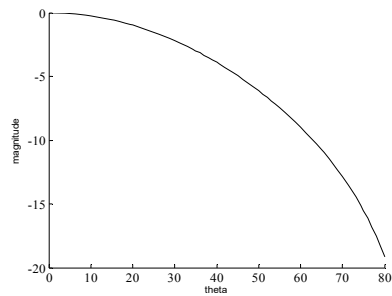


Fig. 3 : Radiation patterns of fully populated hexagonal array for  $p=1,2,3,p=4\&\delta=2$

## V. CONCLUSION

The research in the area of fractal antenna has recently yielded a rich class of new designs for antenna elements as well as arrays.

The essential property of hexagonal arrays have some elements missing, i. e., they are thinning. This is a potential advantage of these arrays from the design point of view, since they may be realized with fewer elements.

Another property of these arrays is that they possess low side-lobe levels at  $2/\pi\phi$ . Finally, it should be noted that the compact product form of the array factor for some particular cases. This offers a significant advantage in terms of computational efficiency, especially for large arrays, and may be exploited to develop rapid beam-forming algorithms

## REFERENCES

- [1] Mandelbrot B. B. "The fractal geometry of nature", San Francisco: 1982
- [2] D. H. Werner, R. L. Haupt and P.L. Werner "Fractal Antenna Engineering: The theory and design of fractal Antenna Arrays" IEEE Antennas

- and propagation Magazine, 1999, vol.41, NO. 5, pp. 7-59.
- [3]. J.Litva and T.K.Y.Lo digital beam forming in wireless communication, Boston, Massachusetts, Artech house, 1996.
- [4] Kravchenko, V.F and Masyuk, V. M. "analysis and synthesis of multiband antenna arrays" Electromagnetic waves and electronic systems, 2004, vol.9, nos. 3-4
- [5] Kravchenko, V.F and Masyuk, V. M. "The ring fractal antenna arrays" Electromagnetic waves and electronic systems, 2004, vol.9, nos. 5
- [6] Werner D. H., and Mittra R., "Frontiers in Electromagnetics", IEEE Press, 2000
- [7] Stutzman W. L. and Thiele G. A., "Antenna Theory and Design", 2d ed., John Wiley & Sons: NY, 1998.
- [8] Balanis C.A. "Antenna theory: Analysis and Design", 2d ed., Wiley, 1997.
- [9] Peitgen H. O., Jurgens H., and Saupe D., "Chaos and Fractals: New Frontiers of Science", New York, Springer-Verlag, Inc., 1992.



# A Comprehensive Approach For The Literature Of Mobile Operating System

Supriya S. Thombre & Sampada S. Wazalwar

Department of Computer Science & Engineering, G.H. Rasoni College of Engineering, Nagpur, India

---

**Abstract** - In the recent years we are witnessing a radical change driven by the introduction and further development of smartphones, where we see emergence of new business models, including 'device centric models', where the user can get access to new applications and services by connecting to the device manufactures' app stores. One of the main drivers of this change is the advanced capabilities of the smartphones enabling the mobile devices to reap the advantages of the convergence process and bring advanced internet applications and services to the mobile devices. The aim of this paper is to give a comparative analysis of different mobile operating systems and identify their strengths and weakness seen from user and market perspective.

**Key words** - Android OS, Symbian OS, Apple's iOS, mobile OS, smartphone, mobile OS comparison, mobile OS features.

---

## I. INTRODUCTION

A smartphone is a high-end mobile phone built on a mobile computing platform, with more advanced computing ability and connectivity than a contemporary feature phone. A feature phone is a mobile phone that combines the functions of a personal digital assistant (PDA) and a mobile phone. As the capabilities of the mobile devices increase, they are not simple voice centric handsets anymore; rather they provide mobile computing power that can be used for several purposes. Especially smartphones represent a possibility of moving appropriate applications from the PC to mobile devices, as they mostly provide large bandwidth wireless network access, office tools and the possibility of installing third party programs [9].

A mobile operating system is the software platform on top of which other programs can run on mobile devices. The operating system is responsible for determining the functions and features available on the device.

The fragmentation of technological platforms and standards are on the one hand create competition which typically is booster for advances and progress and on the other hand as a barrier for development of content and services, which ties the users to specific technologies or puts an extra load on the content and service providers to adopt their contents and services to multiple platforms [1]. The existing mobile operating systems are Android OS, Bada, BlackBerry OS, iPhone OS / iOS, MeeGo OS, Palm OS, Symbian OS, webOS, Windows Mobile. In this paper the main concentration is on Android OS, Symbian OS, and Apple's iOS. Section II contains the

mobile OS history, section III contains version history, section IV contains the technical environment, and section V contains the application environment.

## II. HISTORY

In the beginning, there were cell phones and personal digital assistants (or PDAs). Cell phones were used for making calls and not much else while PDAs, like the Palm Pilot, were used as personal, portable organizers. A PDA could store your contact info and a to-do list, and could sync with your computer. Eventually, PDAs gained wireless connectivity and were able to send and receive e-mail. Cell phones, meanwhile, gained messaging capabilities, too. PDAs then added cellular phone features, while cell phones added more PDA-like (and even computer-like) features. The result was the smartphone. All the mobile phones having mobile OS as Android, Symbian, and Apple etc are the smartphones.

### A. Android OS

The Android mobile operating system is Google's open and free software stack that includes an operating system, middleware and also key applications for use on mobile devices, including smartphone. The OS is fully open source; Updates for the open source Android mobile operating system have been developed under "dessert-inspired" codenames (Cupcake, Donut, Eclair, Gingerbread, Honeycomb, and Ice Cream Sandwich) with each new version arriving in alphabetical order with new enhancements and improvements.

### B. Symbian OS

Symbian is a mobile operating system (OS) targeted at mobile phones that offers a high-level of integration with communication and personal information management (PIM) functionality[.]. Symbian OS combines middleware with wireless communications through an integrated mailbox and the integration of Java and PIM functionality. It has roots in systems that were developed in the 1990's and its debut was in 2001. Symbian OS has its roots in handheld devices and has seen rapid development through several versions. Symbian OS has evolved from a handheld operating system to an operating system that specifically targets real-time performance on a smartphone platform

### C. iPhone OS / iOS

Apple's iPhone OS was originally developed for use on its iPhone devices. Now, the mobile operating system is referred to as iOS. The iOS mobile operating system is available only on Apple's own manufactured devices as the company does not license the OS for third-party hardware. Apple iOS is derived from Apple's Mac OS X operating system.

The operating system was released in June 2007. On March 6, 2008, Apple released the first beta, along with a new name for the operating system: "iPhone OS". In June 2010, Apple rebranded iPhone OS as "iOS". The trademark "IOS" had been used by Cisco for over a decade for its operating system, IOS, used on its routers. To avoid any potential lawsuit, Apple licensed the "IOS" trademark from Cisco. The iOS operating system is a fully featured modern operating system, which is extensively documented and fully supported by Apple Inc. [10]

## III. VERSION HISTORY

### A. Android OS

1) The *first version* of Android was released in 2009 and it comes with the basic features like GPS, Bluetooth, Multitasking and integration of Google services. There are 1669 android apps available for this version.

2) The *second version* was released in April 2009 with the name CUPCAKE. Android 1.5 CUPCAKE is based on Linux Kernel 2.6.27 On 30 April 2009; the official 1.5 (Cupcake) update for Android was released. There were several new features and UI updates included in the 1.5 update and they are as like a new soft-keyboard with text-prediction, Bluetooth A2DP and AVRCP support, Ability to automatically connect to a Bluetooth headset within a certain distance. There are 3453 apps are available for this version on android market.

3) The *third version* is released September 2009 with name DONUT. Android 1.6 (Donut) is based on Linux Kernel 2.6.29. On 15 September 2009, the 1.6 (Donut) SDK was released. Included in the update were: Gallery now enables users to select multiple photos for deletion, Updated Voice Search, with faster response and deeper integration with native applications. Updated technology support for CDMA/EVDO, 802.1x, VPNs, and a text-to-speech engine. There are 3453 apps available on android market for this version.

4) The *fourth version* was released on October 2009 with the name Eclair. Android2.0/2.1 (Eclair) is based on Linux Kernel 2.6.29. On 26 October 2009 the 2.0 (Eclair) SDK was released. Among the changes were: Optimized hardware speed, Support for more screen sizes and resolutions, Digital Zoom, Improved virtual keyboard. There are 14,294 applications available in android market for this version.

5) The *fifth version* was released on May 2010 with the name Froyo (most successful Version). Android 2.2 (Froyo) is based on Linux Kernel 2.6.32. On 20 May 2010 the 2.2 (Froyo) SDK was released. Changes included are General Android OS speed, memory, and performance optimizations, USB tethering and Wi-Fi hotspot functionality, Added an option to disable data access over mobile network, voice dialing and contact sharing over Bluetooth. Total apps available in the android market for this version are 97,703.

6) The *sixth version* is named as Gingerbread. Android 3.0 (Gingerbread) is based on Linux Kernel 2.6.33 or 34. It includes support for WebM video playback, improved copy-paste functionalities, and support for bigger screens with up to Wide XGA (1366×768) resolution.

7) *2.3 Gingerbread* refined the user interface, improved the soft keyboard and copy/paste features, improved gaming performance, added SIP (Session Initiation Protocol)[13] support (VoIP calls), and added support for Near Field Communication[14].

8) *3.0 Honeycomb* was a tablet-oriented release which supports larger screen devices and introduces many new user interface features, and supports multi-core processors and hardware acceleration for graphics. The first device featuring this version, the Motorola Xoom tablet, went on sale in February 2011.

9) *3.1 Honeycomb*, released in May 2011, added support for extra input devices, USB host mode for transferring information directly from cameras and other devices, and the Google Movies and Books apps.

10) *3.2 Honeycomb*, released in July 2011, added optimization for a broader range of screen sizes, new "zoom-to-fill" screen compatibility mode, loading media

files directly from SD card, and an extended screen support API. Huawei MediaPad is the first 7 inch tablet to use this version.

11) *4.0 Ice Cream Sandwich*, announced on October 19, 2011, brought Honeycomb features to smartphones and added new features including facial recognition unlock, network data usage monitoring and control, unified social networking contacts, photography enhancements, offline email searching, and information sharing using NFC. Android 4.0.1 Ice Cream Sandwich is the latest Android version that is available to phones. The source code of Android 4.0.1 was released on November 14, 2011.

#### B. Symbian OS

1) *Symbian^1*, being the first release, forms the basis for the platform. It incorporates Symbian OS and S60 5th Edition (which is built on Symbian OS 9.4) and thus it was not made available in open source. S60 consists of a suite of libraries and standard applications, such as telephony, personal information manager (PIM) tools, and Helix-based multimedia players. It is intended to power fully featured modern phones with large color screens, which are commonly known as smartphones.

2) *Symbian^2* was the first royalty-free version of Symbian. While portions of Symbian^2 are EPL (Eclipse Public License) licensed, most of the source code is under the proprietary SFL (Standard Function Library) license and available only to members of the Symbian Foundation.

3) *Symbian^3* was announced on 15 February 2010. It was designed to be a more 'next generation' smartphone platform. The Symbian^3 release introduced new features like a new 2D and 3D graphics architecture, UI improvements, and support for external displays via HDMI. It has single tap menus and up to three customizable home screens. The Symbian^3 SDK (Software Development Kit) was released September 2010.

4) *Symbian^4* was released in the first half of 2011. Instead, many of the UI enhancements planned for Symbian^4 will be released as updates to Symbian^3.

5) *Symbian Anna* is an update to Symbian^3, released by Nokia in April 2011. Symbian Anna includes such improvements as a new browser, a virtual keyboard in portrait orientation, new icons and real-time homescreen scrolling.

6) *Symbian Belle* is an update to Symbian Anna. It was released in 26<sup>th</sup> October 2011. Symbian Belle adds to the Anna improvements with a pull-down status/notification bar, deeper near field communication integration, free-form re-sizable homescreen widgets.

7) *Symbian Carla* will include a new browser application, enhanced NFC features, and support for Dolby Surround sound processing.

8) *Symbian Donna* will be the first Symbian release to support dual-core processors.

#### C. Apple's iOS

On June 29, 2007, Apple released the first version of what became iOS - concurrently with the first iPhone. The final 1.x series release was 1.1.5.

1) *OS 2.0* was released in July 11, 2008 with upgrades through version 2.2 made available to all models.

2) *iOS 3.0* was released June 17, 2009 with updated through version 3.1.3.

3) *iOS 4.0* was released On June 21 to the public and was made available only to the iPod touch and iPhone. 4.0 were announced to have over 1500 new APIs for developers, with the highly anticipated multitasking feature. The second generation iPod touch and iPhone 3G have iOS 4.2.1 as their latest version available. iOS 4.2 is the first version to bring major feature parity to the iPhone and iPad. The 4.2 version sequence continued for the CDMA phone while 4.3 was released for all other products.

4) *iOS 5* was previewed Apple On June 6, 2011, Apple TV 4.4 beta and the iOS SDK 5 beta along with iCloud beta among other products. This update introduced iMessage chat between devices running iOS 5, a new notification system, and Newsstand subscriptions, Twitter integrated into iOS, Reminders app, Enhancements to AirPlay, full integration with iCloud and over 200 new features.

### IV. TECHNICAL ENVIRONMENT

One issue of critical importance of mobile computing is the design, development and deployment of software solutions for mobile users. These users use such applications in a highly dynamic environment and such systems need to perceive and demonstrate responsiveness to environmental changes [5]. Developers would be interested in supported platform because if their application has the potential to be deployed on more phones, then they get more exposure. To compare the phones in their basic criteria, we analyze the number of phones that deploy the particular operating systems, the number of manufacturers using the operating system, and the number of providers using the operating system [2].

In order to improve software speed, it's needed to increase the cost of space in exchange for time; in order to optimize the code size; we need to increase the running time. Therefore, it should be based on specific



needs to develop appropriate optimization objective, under conditions of limited processor resources, primary consideration should be optimized speed; in the case of storage resource constraints, it should give priority to code size optimization [8].

iOS device displays are high resolution, but then only a small fraction of the image is visible. Everything feels slightly confining [7]. iOS hardware provides a unique platform for sensor applications. iOS hardware includes a complete microcontroller system, with wireless and wire capable networking. Development tools are rigidity controlled and well documented [10].

TABLE I. TECHNICAL ENVIRONMENT

Criteria	Mobile Operating System		
	Android OS	Symbian OS	Apple iOS
Public domain or private domain	Public domain	private domain	private domain
Manufacturer-specific/independent	manufacturer independent	manufacturer specific	manufacturer specific
Audio Playback Format	AAC LC/LTP 3GPP, HE-AACv1, HE-AACv2 AMR-NB, AMR-WB, MP3, MIDI, Ogg Vorbis, PCM/WVE, WAVE	All	AAC, Protected AAC, HE-AAC, MP3, MP3 VBR, AIFF, WAV
Video Playback Format	H.263, H.264 AVC, MPEG-4 SP, DivX, XviD, VP8	H.263, H.264, WMV, MPEG4, MPEG4 @ HD 720p 30fps, MKV, DivX, XviD	H.264 AVC, MPEG-4, M-JPEG
Turn-by-turn GPS	Google Maps Navigation or 3rd party software	free global Nokia Ovi Maps and 3rd party software	3rd party software

Official SDK platform	Linux, Mac OS X and Windows	Windows using Symbian SDK or Linux, Mac OS X and Windows using Nokia Qt SDK	Mac OS X using iOS SDK
GPS Support	Provides GPS & Location Manager APIs allowing Development that is not HW specific	GPS, A-GPS and network-based positioning	Technology uses GPS, cell, or Wi-Fi signals to triangulate the position
Language Supported	JAVA	C++	C++
Application installation method	Unknown; installation on emulator is not reflective of production devices	Direct, PC suite	Via iCloud
Underlying architecture	Linux	Symbian	iOS
iCloud Support	No	No	Yes
Kernel type	Monolithic (Linux Kernel)	Micro kernel	Hybrid (XNU)
Default user interface	Graphical	Avkon	Cocoa Touch (Multi-Touch, GUI)
OS Family	Linux	Embedded OS	Mac OS X/BSD/Unix-Like
Supported Platform	ARM, MIPS, Power Architecture, x86	ARM, x86	ARM
License	Apache License 2.0 Linux Kernel patches under GNU GPL V2	Proprietary	Proprietary EULA except for open source component

<i>Package Manager</i>	APK	Nokia OVi Suite	iTunes
<i>Processor Speed</i>	200 MHz online	1 GHz	866 MHz
<i>Memory</i>	32 MB RAM and 32 MB Flash	512 MB RAM	512 MB RAM
<i>Multitasking</i>	Yes	Yes	Very Limited (versions before 4+)
<i>Multituser</i>	No	No	No

## V. APPLICATION ENVIRONMENT

Apple App Store has introduced a substantially different model for digital good distribution to mobile users. [6]. An Android OS smartphone receives the data assuring intermediate processing, graphical user interface and data synchronization. [3]. Android is based on a Linux kernel with the user space and the JVM for Android (Dalvik) being written in C [1].

Use space and time optimization strategy on the code side to solve the contradiction between storage capacity and efficiency of running in resource-constrained devices; use double buffering and direct screen access technology on the display side to solve the problem of graphics flashing when refreshing the screen [4].

The device can be hooked up to an iOS device when the control process can accept digital commands, to exchange data and be configured and controlled. The device gains all the functionality offered by the iOS operating system [10].

TABLE II. APPLICATION ENVIRONMENT

Criteria	Mobile Operating System		
	<i>Android OS</i>	<i>Symbian OS</i>	<i>Apple iOS</i>
<i>Official Application Store</i>	Android market	Symbian Horizon, Ovi store	App store
<i>Unified Inbox</i>	No	Yes	Yes
<i>Offline Voice Command</i>	No	Yes	Yes
<i>Photo/video import from memory cards</i>	NO	Yes	With Camera Connection Kit

<i>Tethering [11]</i>	Mobile Wi-Fi Hotspot, USB, Bluetooth	USB, Bluetooth, Mobile Wi-Fi Hotspot (with 3rd party software)	Bluetooth, USB, Personal Hotspot (Wi-Fi Tethering)
<i>Text/Document Support</i>	Microsoft Office 2003/2007, PDF, Images, TXT/RTF	Microsoft Office Mobile, PDF	Microsoft Office, iWork, PDF, Images, TXT/RTF, VCF
<i>Printer support</i>	No built-in function, but have apps use Google Cloud Print available for 2.0+	No	Yes (AirPrint)

## VI. CONCLUSION

The Android OS and Apple OS have their own plusses and minuses. Both are equally strong contenders and are bound to rule the app market place with their own strength and positives. Android OS and Apple OS are much advance and provide more value to the end users while Symbian needs to update its features. Android OS and Apple OS are the only two major Operating Systems available in the smartphone market until 2010. With regards to mobile operating systems, Symbian has for long time been the dominating technology, however, it seems that in the transition to the smartphones other operating systems like iPhone and Android are taking the lead. However, there are several possibilities that Symbian OS will give tough competition to them in the year 2012 onwards.

## REFERENCES

- [1] A. Hammershoj, A. Sapuppo, R. Tadayoni, "Challenges for Mobile Application Development", Center for Communication, Media and Information Technologies (CMI), Aalborg University Lautrupvang 15, 2750 Ballerup, Copenhagen, Denmark, 978-1-4244-7445-5/10/\$26.00 c 2010 IEEE.
- [2] M. Wei, A. Chandran, Hsin-Ping Chang, Jui-Hung Chang, C. Nichols, "Comprehensive Analysis of SmartPhone OS Capabilities and Performance".

- [3] O. Postolache, P. S. Girão, M. Ribeiro, M. Guerra, J. Pincho, F. Santiago, A. Pena “Enabling telecare assessment with pervasive sensing and Android OS smartphone”, 978-1-4244-9338-8/11/\$26.00 ©2011 IEEE.
- [4] D. Zhu, J. He and Jinxiang Li, “Study on Optimization of Applications Based on Symbian Platform”, 978-1-61284-459-6/11/\$26.00 ©2011 IEEE
- [5] R. Abidar, K. Moummadi, H. Medromi, “Mobile device and Multi agent systems An Implemented platform of real time data communication and synchronization”,978-1-61284-732-0/11/\$26.00 ©2010 IEEE.
- [6] K. Kimbler, “App Store Strategies for Service Providers”, 978-1-4244-7445-5/10/\$26.00 ©2010 IEEE.
- [7] M. P. Rogers, “There and back again: leveraging ios development on MAC OS X” ,CCSC: Central Plains Conference.
- [8] Scott Meyers, “Effective C++”, BEIJING: Publishing House of Electronics Industry, 2006.
- [9] A.D. Schmidt, F. Peters, F. Lamour, S. Albayrak, “Monitoring Smartphones for Anomaly Detection”, Mobilware’08 February 12-15, 2008, Innsbruck, Austria, Copyright 2008 ACM 978-1-59593-984-5/08/02.
- [10] D. Brateris, D. Bedford, D. Calhoun, A. Johnson, N. Kowalski, K. Martino, T. Mukalian, J. Reda, A. Samaritano and R. R. Krchnavek, “iOS Hardware as a Sensor Platform: DMM Case Study”, 978-1-4244-8064-7/11/\$26.00 ©2011 IEEE.
- [11] <http://mobileoffice.about.com/od/phonesformobileworkers/f/what-is-tethering.htm>
- [12] <http://socialcompare.com/en/comparison/android-versions-comparison>
- [13] <http://www.sipcenter.com/sip.nsf/html/What+Is+SIP+Introduction>
- [14] <http://google.about.com/od/socialtoolsfromgoogle/g/Nfc-Near-Field-Communication.htm>



# Study of Object Extraction from Color Images in Frequency and Spatial Domain

Saiyed Umer & Kalyani Mali

Department of Computer Science & Engineering, University of Kalyani, Kalyani, West Bengal, India

---

**Abstract**— A correlation model for image retrieval is proposed. This model captures the semantic relationships among images in a database from simple statistics of user provided relevance feedback information. It is applied in the post-processing of image retrieval results such that more semantically related images are returned to the user. The algorithm is easy to implement and can be efficiently integrated into an image retrieval system to help improve the retrieval performance.

**Keywords**- Fast Fourier Transformation (FFT), Correlation Filter, Sum of Squared Difference (SSD).

---

## I. INTRODUCTION

The main idea behind the proposed model is the assumption that two images represent similar semantics. Correlation is an important measure is used to find the positions of a target object in a given image. The main objective of this paper is the study of extraction of target object using correlation theorem in spatial domain as well as frequency domain from a given gray scale and color images. One of the spatial domain approach SSD (Sum of Squared Difference) method and FFT are used for correlation theorem in spatial and frequency domain approach respectively.

## II. PREVIOUS WORK

“On Spatial Quantization of Color Images” by Jan Puzicha, Marcus Held, Jens Ketterer, Joachim M. Buhmann, *Member, and IEEE* in 2000 proposed a new model to simultaneously quantize and halftone color images. The method is based on a rigorous cost-function approach which optimizes a quality criterion derived from a simplified model of human perception which incorporates spatial and contextual information into the quantization and thus overcomes the artificial separation of quantization and halftoning.

“Directional processing of color images Theory and Experimental result” by Prof. Joel C. De Goma in 2007 had proven that the addition of noise suppression techniques in images increased the accuracy of the match returned by the Normalized Cross Correlation algorithm of template matching.

“Hyper complex Auto and Cross Correlation of Color Images” by Stephen J. Sangwine, define correlation application for color images, based on quaternions or hyper complex numbers. They have

devised a visualization of the result using the polar form of a quaternion in which color denotes quaternion Eigen axis and phase, and a greyscale image represents the modulus.

“Video Based Moving Object Tracking by Particle Filter” by Md. Zahidul Islam, developed the video based object tracking deal with non-stationary image stream that changes over time using color information, distance transform (DT) based shape information, nonlinearity and template from Particle filtering. They want to develop Robust and Real time moving object tracking system in computer vision research area.

“Object Recognition Based On Template Correlation In Remote Sensing Image” by Xiyuan Zhou proposed the analysis of the conventional methods for information extracting from the remote sensing image, a method of extraction particular object in remote sensing image based on feature template correlation. So, they applied methods to several high-resolution example images to extract and recognize.

“Face Detection Through Template Matching and Color Segmentation” by Scott Tan Yeh Ping describes a process for face detection, which involves multi-resolution template matching, region clustering and color segmentation, works with high accuracy, and gives good statistical results with training images.

“Template Matching Through Subsequent Analysis of an Image” by Prof. Joel C. De Goma in 2007 aims to incorporate noise suppression and image enhancement through the use of image restoration algorithms. It is expected that, utilizing the algorithms stated, an accurate match can be decided by the system. His study had proven that the addition of noise suppression techniques

increased the accuracy of the match returned by the Normalized Cross Correlation algorithm.

Template matching has obvious applications in computer and robot vision, stereography, image analysis and motion estimation. Template matching in the context of an image processing is a process of locating the position of a sub image within an image of the same, or more typically, a larger size. Template matching can also be described as a process to determine the similarity between two images. The sub image is referred to as the template image and the larger image is referred to as the search area (main image). The template matching process involves shifting the template over the search area and computing a similarity between the template image and the window of the search area over which the template lies. These shifting and computing processes operate simultaneously and do the repetition until the template image lies on the edge of the search area. So, the main objective of this paper is to find positions of a particular object in the given image database and to discuss the comparative study of template matching for correlation method in frequency domain as well as spatial domain based on gray scale and color images.

### III. COLOR IMAGES

Color is a sensation created in response to excitation of our visual system by electromagnetic radiation known as light. More specific, color is the perceptual result of light in the visible region of the electromagnetic spectrum, having wavelengths in the region of 400nm to 700nm, incident upon the retina of the human eye. Physical power or radiance of the incident light is in a spectral power distribution (SPD), often divided into 31 components each representing a 10nm band.

Hence Color imaging systems are used to capture and reproduce the scenes that humans see. Imaging systems can be built using a variety of optical, electronic or chemical components. However, all of them perform three basic operations, namely: (i) image capture, (ii) signal processing, and (iii) image formation. Color-imaging devices exploit the tri chromatic theory of color to regulate how much light from the three primary colors is absorbed or reflected to produce a desired color.

### IV. RGB COLOR SPACES

It is the classical Computer Color space with three different colors which are Red (R), Green (G), Blue (B). If R, G, B have the same energy we perceive a shade of White (gray, black). A single pixel consists of three components: [0, 255]. Each pixel is a vector. Sometimes pixels are not stored as vectors but in Image Bands. First, the complete red-component is stored, then the

complete green, then blue. For an example for a pixel value.

Fig. 1(A) and 1(B) represent RGB color spaces and one pixel value of a color image. Fig. 1(C) is the original image. Fig. 1(D), 1(E) and 1(F) are red, green and blue band of Fig. 1(C) respectively.

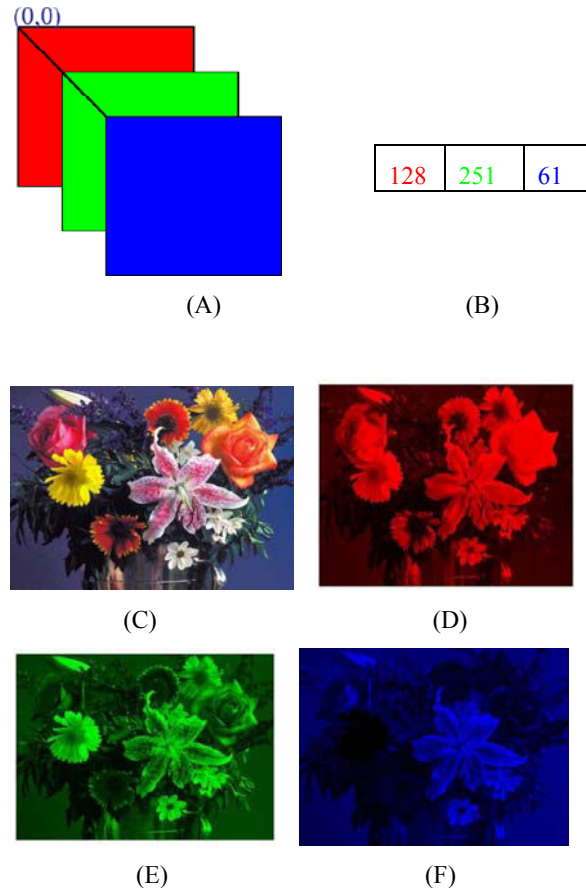


Fig.1: A) RGB COLOR SPACES. B) ONE PIXEL VALUE. C) ORIGINAL IMAGE. D) RED COMPONENT. E) GREEN COMPONENT. F) BLUE COMPONENT.

### V. CONVERT COLOR TO GREY SCALE IMAGES

RGB color image is converted to Gray Scale image using the following equation

$$I = (R + G + B) / 3 \quad \text{----- (1)}$$

Where I, R, G and B represents Intensity component, Red Component, Green Component and Blue Component respectively of the color image.

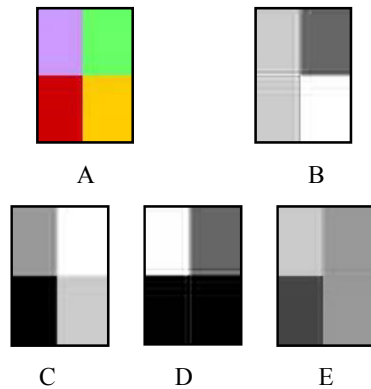


FIG 2: A) ORIGINAL IMAGE. B) RED COMPONENT. C) GREEN COMPONENT. D) BLUE COMPONENT. E) INTENSITY COMPONENT.

Fig 2.A. is an original color image. Fig2B, Fig2C, Fig2D are red, green and blue component which are obtained by reading each pixel of the color image Fig2A. Fig2E is intensity component which is obtained by using EQ. (1).

## VI. THE CORRELATION THEOREM IN SPATIAL DOMAIN

In spatial domain, SSD (Sum of Squared Distance) is a pre-processing technique for correlation method. It is one of the approaches which gives satisfactory result in template matching. Here the template must have same size as in the provided search image.

This method is normally implemented by first picking out a part of the search image to use as a template. We will call the search image  $S(x, y)$ , where  $(x, y)$  represent the coordinates of each pixel in the search image. We will call the template  $T(x_t, y_t)$ , where  $(x_t, y_t)$  represent the coordinates of each pixel in the template. We then simply move the centre (or the origin) of the template  $T(x_t, y_t)$  over each  $(x, y)$  point in the search image and calculate the sum of squares of distance between the coefficients in  $S(x, y)$  and  $T(x_t, y_t)$  over the whole area spanned by the template. As all possible positions of the template with respect to the search image are considered, the position with the highest score is the best position. This method is sometimes referred to as 'Linear Spatial Filtering' and the template is called a filter mask.

For example, one way to handle translation problems on images, using template matching is to compare the intensities of the pixels, using the SSD (Sum of Squared differences) measure.

A pixel in the search image with coordinates  $(x_s, y_s)$  has intensity  $I_s(x_s, y_s)$  and a pixel in the template

with coordinates  $(x_t, y_t)$  has intensity  $I_t(x_t, y_t)$ . Thus the squared difference in the pixel intensities is defined as

$\text{Diff}(x_s, y_s, x_t, y_t) = (I_s(x_s, y_s) - I_t(x_t, y_t))^2$  which is shown in EQ. (3).

$$\text{SSD}(x, y) = \sum_{i=0}^{T_{\text{rows}}-1} \sum_{j=0}^{T_{\text{cols}}-1} \text{Diff}(x+i, y+j, i, j) \quad (2)$$

The mathematical representation of the idea about looping through the pixels in the search image as we translate the origin of the template at every pixel and take the SSD measure is the Eq. (4):

$$\sum_{i=0}^{S_{\text{rows}}-1} \sum_{j=0}^{S_{\text{cols}}-1} \text{SSD}(x, y) \quad (3)$$

Where **Srows** and **Scols** denote the rows and the columns of the search image and **Trows** and **Tcols** denote the rows and the columns of the template image, respectively. In this method the lowest SSD score gives the estimate for the best position of template within the search image. The method is simple to implement and understand. Fig.3 shows the experimental results using grey scale images and grey scale templates.

So, Fig. 3(A) is the original image of hand gestures. Fig. 3(B) is a template which must be same size and shape if it presents in Fig. 3(A) image. Now, Fig. 3(C) is the result of SSD using EQ. (2) and (3). In Fig. 3(C), the darkest point is the position of the template at Fig. 3(B). Fig. 3(D) shows the matched template.

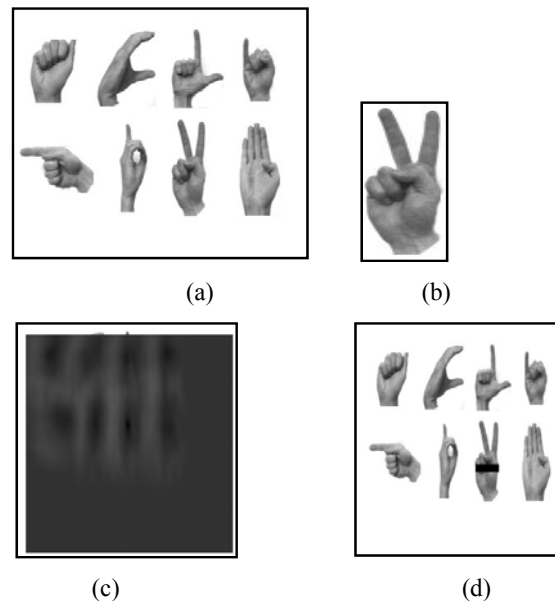


FIG.3: A) ORIGINAL IMAGE. B) TEMPLATE. C) CORRELATION RESULT. D) BEST MATCHED TEMPLATE IMAGE.

## VII. THE CORRELATION THEOREM IN FREQUENCY DOMAIN

Template matching using correlation basically uses the Correlation theorem:

$$f(x, y) \circ h(x, y) \Leftrightarrow F(u, v) H^*(u, v) \text{-----(4)}$$

Where  $f(x, y)$  is the search image having  $N \times M$  sizes and  $h(x, y)$  is the template image having  $K \times L$  sizes.

It consists of the following steps:

1. Multiply the search and the template image by  $(-1)^{x+y}$  centre transformed.
2. Compute  $F(u, v)$ , the FFT of the variable padded search image from (1).
3. Compute  $H(u, v)$ , the FFT of the variable padded template from (1).
4. Multiply  $F(u, v)$  by  $H^*(u, v)$  (Conjugate of  $H(u, v)$ ).
5. Compute the inverse FFT of the result in (4).
6. Obtain the real part of the result in (5).
7. Multiply the result in (6) by  $(-1)^{x+y}$ .

Here the size of the variable padded image equals to  $N+K-1$ ,  $M+L-1$ , where  $N$ ,  $M$  are the size of the original image and  $K$ ,  $L$  are the size of the template image. The resulting image will have the highest value at the matching point. An example of correlation in frequency domain is discussed in FIG.4 which shows a simple illustration of image padding and correlation. Fig. 4 (A) is the image of ten hand gestures and Fig. 4(B) is a hand gesture taking as a template. The padded images are shown in Fig 4(C) and 4(D) respectively. The correlation of the two padded images is displayed as an image in Fig 4(E). The correlation function defined in EQ. (4) was obtained by computing the transforms of the padded images, taking the complex conjugate of one of them (we chose the test image), multiplying the two transforms and computing the inverse FFT. As expected, we see in Fig. 4(E) that the highest value of the correlation function occurs at the point.

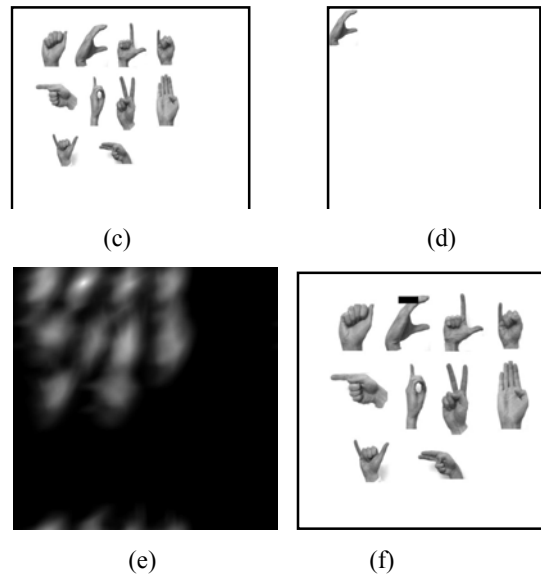
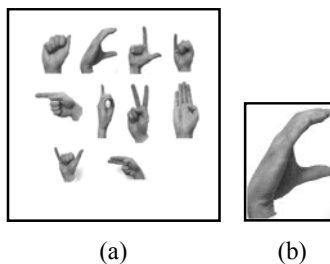


FIG.4: A) ORIGINAL IMAGE. B) TEST IMAGE. C) PADDED IMAGE OF A. D) PADDED IMAGE OF B. E) CORRELATION RESULT. F) BEST MATCHED TEMPLATE IMAGE.

## VIII. EXPERIMENTAL RESULTS

Fig. 5 represents correlation result of color image in spatial domain. Here Fig. 5(A) is an image of group of people. Fig. 5(B) is a face as test image to find in Fig. 5(A). Fig 5(C), 5(D), 5(E) are red, green, blue component of Fig. 5(A). Fig. 5(F) and Fig. 5(G) are the intensity image of Fig. 5(A) and 5(B) respectively which are obtained by using EQ. (1). Now SSD algorithm discussed in section (VI) is applied for correlation theorem between Fig. 5(F) and 5(G) which produces result in Fig. 5(H). Then, pixel wise difference between Fig. 5(H) and 5(F) is performed to obtain Fig. 5(I). Now, Fig. 5(I) contains brightest point where the test image will be found. Fig. 5(J) is the matched template image.





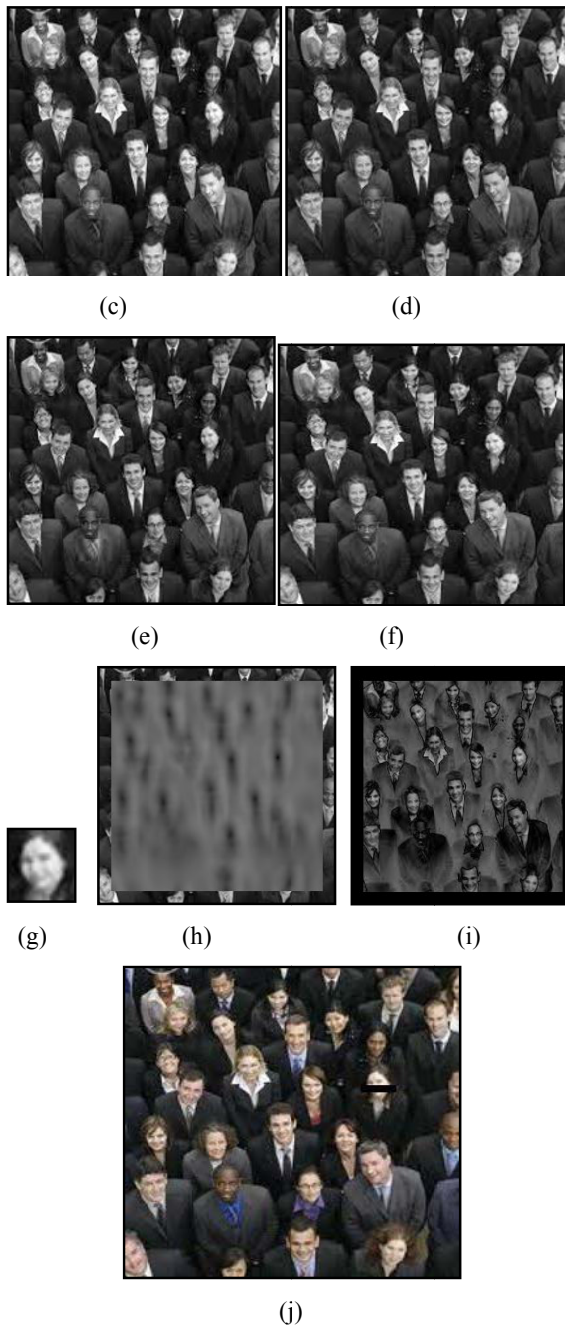


Fig.5: A) Original Image. B) Template. C) Red Component. D) Green Component. E) Blue Component. F) Intensity Component Of A. G) Intensity Component. Of B. H) Correlation Result. I) Image Difference Result. J) Matched Template Image.

Fig. 6 shows another example for color image. Fig. 6(A) is an image of different color boxes. Fig. 6(B) is an intensity component of Fig. 6(A) which is obtained as

discussed in EQ. (1). Now, Fig. 6(C) and 6(D) are test image and its intensity component respectively. Fig. 6(E) is SSD result using images at Fig. 6(B) and 6(D). The darkest point of Fig. 6(E) is shown in Fig. 6(F) as matched test image position.

The objective of example at Fig. 6 is that the object of same size and shape but different in color can also be recognized.

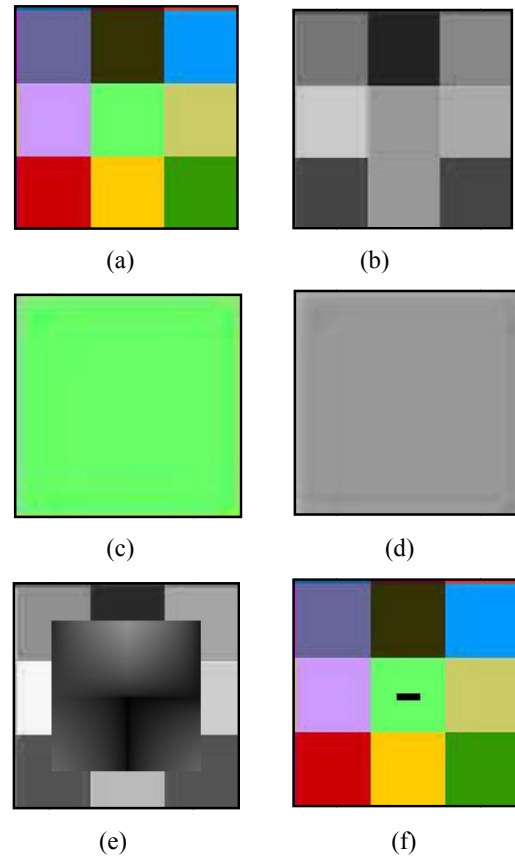
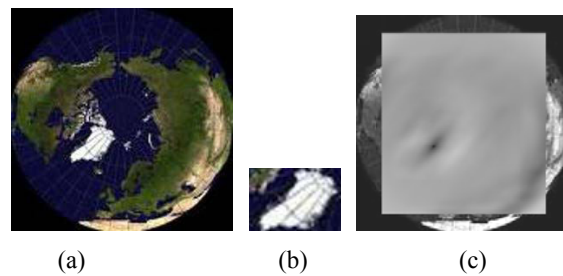


Fig.6: A) Box Image. B) Intensity Of A. C) Test Image. D) Intensity Of C. E) Ssd Result. F) Matched Test Image Shown.



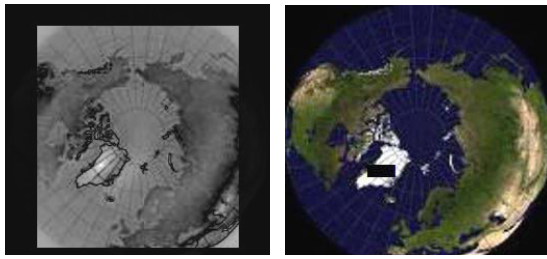


Fig.7: a) Original Image. b) Test IMAGE. C) CORRELATION RESULT. D) IMAGE DIFFERENCE RESULT. E) MATCHED TEMPLATE IMAGE

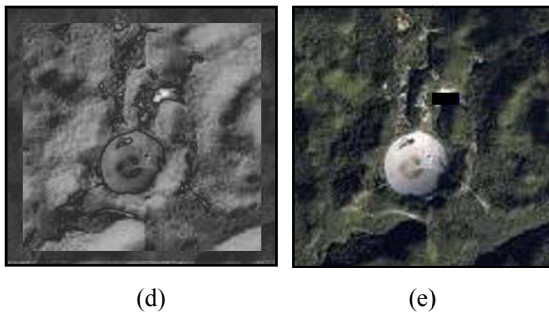
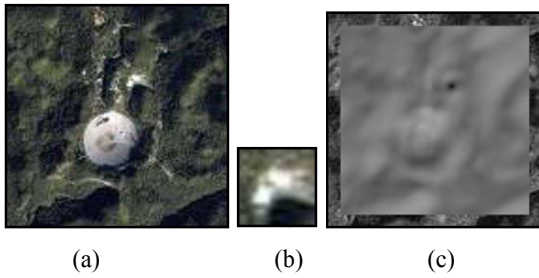


FIG.8: A) ORIGINAL IMAGE. B) TEST IMAGE. C) CORRELATION RESULT. D) IMAGE DIFFERENCE RESULT. E) MATCHED TEMPLATE IMAGE.

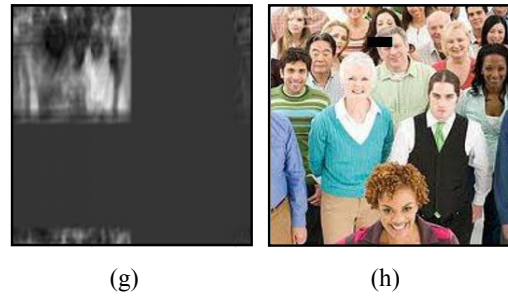
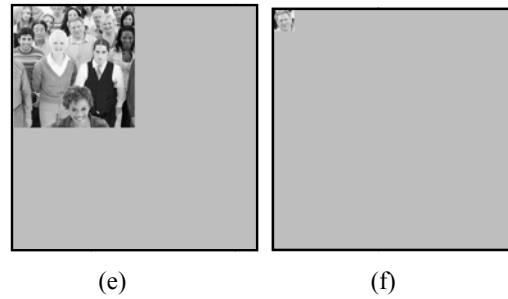
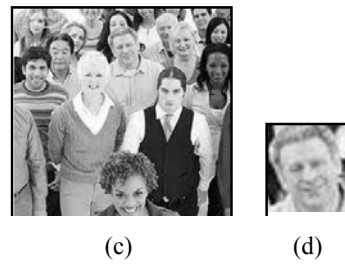
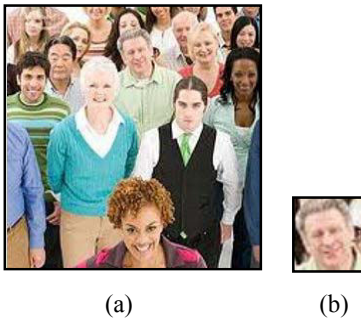
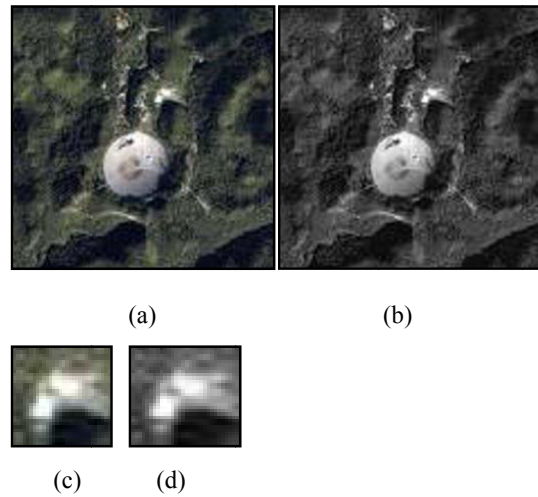


Fig.9: A) Original Image. B) Test Image. C) Intensity Component Of (A). D) Intensity Component Of (B). E) Padded Image Of (C). F) Padded Image Of (D). G) Correlation Result. H) Matched Template Image.



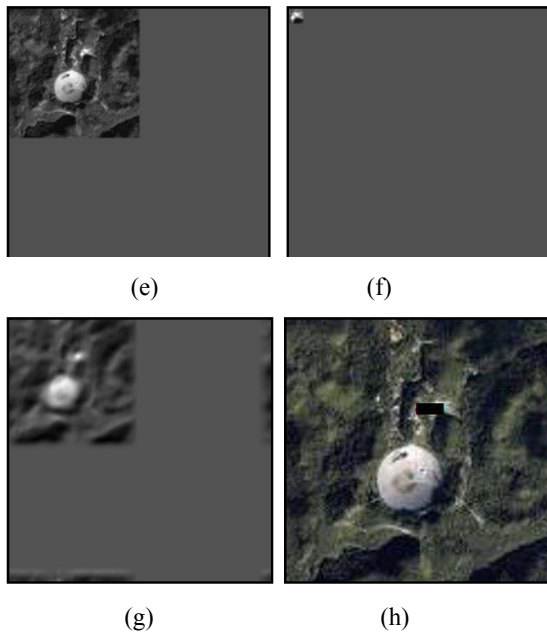


Fig.10: A) Original Image. B) Intensity Component Of (A). C) Test Image. D) Intensity Component Of (C). E) Padded Image Of (C). F) Padded Image Of (D). G) Correlation Result. H) Matched Template Image.

Fig.9 and Fig.10 shows the experimental results for correlation of color images in frequency domain. These results have been done using algorithm defined in section (VII). The cause of successful of these results depend on selecting the padding values iteratively which are used here to perform correlation theorem in more robust manner. TABLE 1 shows padding values range for images used in Fig. 9 (A) , 9 (B) and Fig. 10 (A) and 10 (B) for correlation theorem in frequency domain.

Table-1

Image Name	Padding Range
FIG. 9.(A), 9 (B)	120 - 190
FIG. 10.(A), 10 (B)	80 - 140

## IX. CONCLUSIONS

The correlation is best method for template matching. Both, the spatial and frequency domains approaches have been used here for both greyscale and colour images with satisfactorily experimental results. The results presented here have been for position

invariance. The spatial domain template matching processes are slower one and dependent on template shape and size where as frequency domain template matching is faster one and independent of size of template. The matching accuracy equation is totally reliable and result a perfect score in this project. For this project, even though the results obtained so far are encouraging, more investigations need to be done in more detail on both theoretical and practical side for further study on matching process. The image difference process which is discussed previously is needed for spatial domain approach but there is no need of it in frequency domain. Selection of padding value matters in frequency domain but not in spatial domain approach.

In future work we will develop more robust technique to extract an object which may be variation in size of the template and independent of predefined test pattern from a given image database.

## REFERENCES

- [1] Yongmin Kim, Shijun Sun, HyunWook Park, "Template Matching Using Correlative Auto predicative Search". University of Washington, 2004.
- [2] W. Krattenthaler, K.J. Mayer, M.Zeiler, "Point correlation: a reduced cost template matching technique", International conference on image processing, pp. 208-212, 1994.
- [3] J. Kim, J. A. Fessler, "Intensity-Based Image Registration Using Robust Correlation Coefficients", IEEE Trans. Med. Imag., vol. 23, no. 11, pp. 1430-1444, Nov. 2004.
- [4] Burrus, C. S. and Parks, T. W., (1985), DFT/FFT and Convolution Algorithms, WileyInterscience, NY, USA.
- [5] R. C. Gonzalez and P. Wintz. Digital Image Processing. Addison Wesley, Menlo Park, Ca. second edition, 1987 .
- [6] G. Shakhnarovich and B. Moghaddam. Face recognition in subspaces. In S. Z. Li and A. K. Jain, editors, Handbook of Face Recognition, page 35. Springer-Verlag, December 2004.
- [7] Ming-Hsuan Yang, D. J. Kriegman and N. Ahuja, "Detecting faces in images: a survey," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 24, pp. 34-58, 2002.
- [8] M. J. Atallah, "Faster image template matching in the sum of the absolute value of differences measure," Image Processing, IEEE Transactions on, vol. 10, pp. 659-663, 2001.

- [9] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. IEEE*, vol. 83, pp. 705–741, 1995.
- [10] Lieven N. and Ewins, D., « Spatial correlation of mode shapes, the coordinate modal assurance criterion (COMAC) », 6th Intl. Modal Anal. Conf., Kissimmee, Florida, pp 605-609, 1988.
- [11] COOLEY( J . W.), TUKEY( J. W.), *Mathematics of Computation*, 1965, 19, 296.
- [12] W. K. Pratt, "Correlation techniques of image registration," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-10, pp. 353–358, 1974.
- [13] P. J. Burt, C. Yen, X. Xu, "Local Correlation Measures for Motion Analysis: a Comparative Study", *IEEE Conf. Pattern Recognition Image Processing* 1982, pp. 269-274.
- [14] Harley R. Myler, Arthur R. Weeks, "The Pocket Handbook of Imaging Processing Algorithms in C". Department of Electrical & Computer Engineering, University of Central Florida, Orlando Florida. 1993.



# Effect of RFID Technology on Software Architecture

Kaveh Ahamadi Niar, Fereshte Jadidi, Maliheh Shirvanian & Radin Zahedi

Department of Computer Engineering, Shahid Beheshti University, Tehran, Iran

Department of Computer Engineering, Islamic Azad University, Tehran, Iran

---

**Abstract**—Radio frequency had been an efficient real-time method for collecting the surroundings data and converting it to the digital data for several years. Using this method tracking of objects, animals and human beings become feasible through RFID. However considering a new technology without noticing its gradual influence on system components and processes, questions the effectiveness of the technique. Therefore in this paper initially requirements of RFID software architecture in abstract level and afterward the proposed software architecture are defined. To conclude, the influence of RFID on software architecture and software technology is explained.

**Keywords**—component; RFID, Software Architecture, Real-time Distributed Event Driven Architecture

---

## I. INTRODUCTION

Radio Frequency Identification (RFID) is a reliable technology in automatic identification of objects, animals and people. RFID is used in transportation and distribution, retail and consumer packaging, industrial and manufacturing, security and access control and a lot more. Nowadays RFID is on the peak, the same way its predecessor - barcode - used to be on top one day [1].

Software architecture is a high level abstraction of a software system that can determine valuable information such as system quality and hidden risks even before a system is developed. Software architecture consists of component and their relationship and has different level of abstraction [3]. Software architecture can be described as an architecture style.[9] This paper defines the highest level of abstraction of an RFID system where only RFID system components interaction is considered.

In this paper RFID is introduced briefly. Then we propose a real-time distributed even driven software architecture. To conclude effect of RFID on software architecture and in general on software technology is described.

## II. AUTO IDENTIFICATION AND RFID

Automatic Identification and Data Capture (AIDC) refers to the methods of automatically identifying objects, collecting data about them, and entering that data directly into computer systems without human involvement. Technologies typically considered as part of AIDC include bar codes, Radio Frequency Identification (RFID), biometrics, magnetic stripes, Optical Character Recognition (OCR), smart cards, and

voice recognition. The aim of most of AIDC techniques is to increase the performance, decrease human errors and provide personnel with more time to serve customers.

Among the named techniques, RFID is relatively the newest one and is more noticed due to its live tracking ability and the great improvement it created in different markets [1].

RFID provides an important platform to identify objects and collect their information to track and manage them. The mentioned platform is a combination of products and technologies that collects and transfers data between tagged items and an information management system through radio frequency. RFID tags are designed based on a frequency range and system requirements such as read distance and environmental characteristics. RFID Reader is usually attached to a computer and has a job similar to barcode scanner.

Radio frequency elements can be classified to three different categories:

### A. Hardware

An RFID tag which is a microchip combined with an antenna in a compact package; the packaging is structured to allow the RFID tag to be attached to an object to be tracked.

The tag's antenna picks up signals from an RFID reader or scanner and then returns the signal, usually with some additional data (like a unique serial number or other customized information).

RFID tags can be very small - the size of a large rice grain. Others may be the size of a small paperback book.

An RFID reader is a device that is used to interrogate an RFID tag. The reader has an antenna that emits radio waves; the tag responds by sending back its data. The data is transferred to a host computer.

The host system is a computer which connects to RFID readers and usually the middleware or database resides on it.

### *B. Middleware*

The middleware sends the information contained in the tags to whatever systems need that information [4]. Middleware is a platform for managing RFID data and routing it between tag readers or other auto identification devices and enterprise systems. In fact middleware is the interface between the hardware and software. Middleware processes the data collected from tags and send the required information to each front-end system [2].

Middleware should support different protocols to connect to different readers. Besides it should be able to filter and group the data and remove the redundancy [4]. The main responsibility of a middleware is:

- Capturing, securing, and accessing EPC related data.
- Obtaining filtered, aggregated data from several sources. (This standard is sometimes referred to as the ALE standard.)
- Exchanging data and commands between hosts and readers to do things like read tags, write to tags, and kill tags.
- Configuring, provisioning, and monitoring individual readers.

A RFID system should be able to utilize different RFID Readers and protocols. In order to support a multi protocol environment middleware is required. In this manner the host system is not dependent on a special reader or protocol. Middleware should use standard protocols and be able to communicate with different applications at the same time.

Most of the middleware available today is commercial-based and has been developed by the commercial players, for example, Microsoft BizTalk RFID [5], Oracle Fusion [6] and Sun RFID Middleware [7]. There are also some middlewares which have been developed for research purposes, such as the Accda [8].

RFID has not been completely standardized yet and there is some dissimilarity between different

manufacturers. Therefore one of the main responsibilities of middleware is to fill this gap.

### *C. Software*

The host computer requires software and a databank to store and retrieve the information. Software can be a part of ERP, WMS, SCM, and CRM or any other application that require the information.

In another word the most important user of the identification systems is enterprise systems. Identification systems process the raw data collected from the backend systems and prepare the information top management require for enterprise decision making [2].

## **III. RFID SOFTWARE CHARACTERISTICS**

Considering the distribution feature of RFID systems, the software should also be designed distributed to handle the system requirement. Besides business partners should be able to share data in a supply chain system to improve the quality of service all along the supply chain. Therefore the capability of sharing data all along the supply chain should be considered. The main barrier in fast growth of RFID is high installation and maintenance cost. This should be considered in design phase to reduce cost. From the software development point of view, the simpler and more reliable the software is, the less implementation and maintenance cost.

RFID hardware mechanism is real-time. It gathers all the data and sends it to software almost in no time. Hence the software should also have this real time feature and the software architecture should be a real-time structure.

## **IV. AVAILABLE SOFTWARE ARCHITECTURE**

### *A. Layered Rfid Software Architecture*

In this architecture a layered structure is described and RFID system components rely on of the layers. Fig. 1 shows the defined layers. Each layer includes some elements that perform the layer job [2].

Layer one is the communication network layer in lowest level. In an RFID system some components such readers and PCs are connected through a computer network which is the IT infrastructure. In this layer all this hard ware infrastructure is define. Active and passive network equipments, telecommunication equipments and network software and services is defined is this layer.

Layer two includes RFID equipments such as readers, antenna and tags Readers collect the information and send it to the middleware through the communication networks.



The third layer is the middleware and the forth layer is the Information systems and enterprise applications.

#### B. Sun Microsystems Software Architecture

The Java System RFID Software architecture is built with Java Enterprise System software and technology. Fig 2. demonstrates this architecture. At the bottom of the stack are tag readers or sensors that are responsible for reading tagged items. Each reader continuously reads many tagged items and sends that data to the next layer in the architecture stack [11].

Java System RFID Software consists of two major components:

- The Java system RFID event manager
- Java system RFID information server

The Java System RFID Event Manager is designed to process streams of tag or sensor data (event data) coming from one or more reader devices. The RFID Event Manager has the capability to filter and aggregate data prior to sending it to a requesting application. Sun's RFID Event Manager filters can be programmed to throw out any data that shows the RFID-tagged

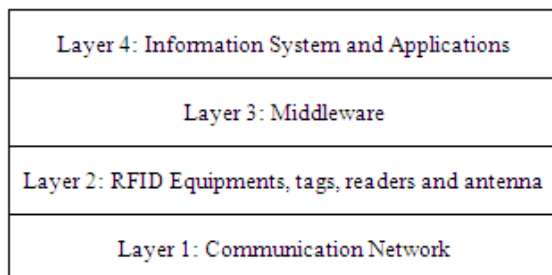


Figure 1. Layered RFID Software Architecture [2].

object is in the same place and trigger an action or event only when there is a change in state for the object. The RFID Event Manager can also be programmed with other types of filters to enforce specific business rules.

To localize reader traffic, an enterprise may have numerous instances of Sun's RFID Event Manager at each geographically remote site, such as a store, distribution center, or warehouse. Given the amount of network traffic from readers, it is important to localize data by enabling the RFID Event Manager servers to filter the tag data at each site instead of sending it over the Internet. In addition, it is good practice to isolate the readers from the Internet for security reasons.

The other major component is the Java System RFID Information Server. Sun advocates that integration technologies be used to connect the RFID Event

Manager layer to enterprise information systems (EIS) - such as legacy, enterprise resource planning (ERP), warehouse management systems (WMS), supply chain management (SCM), and customer relationship management (CRM) systems - as well as other applications that might want to use tag information. These include components and technologies that comprise the Java Enterprise System, such as Web, application, communication, and security services, and Java technologies such as the Java Message Service (JMS) and the Java 2 Platform, Enterprise Edition (J2EE™) Connector Architecture to enterprise information systems.

The topmost layer in the architecture stack is comprised of EIS systems such as ERP, WMS, legacy systems, and proprietary enterprise systems.

#### C. Real-Time Distributed Software Architecture

In this section a real-time distributed RFID software architecture is surveyed. First the event is defined and then the architecture is explained [12].

Details of anything that is happening at a specific time and place is an event. Event can be described with a set of dimensions:

- source of event
- event granularity
- location of event
- time at which the event occurred

There are two categories of event happening in RFID:

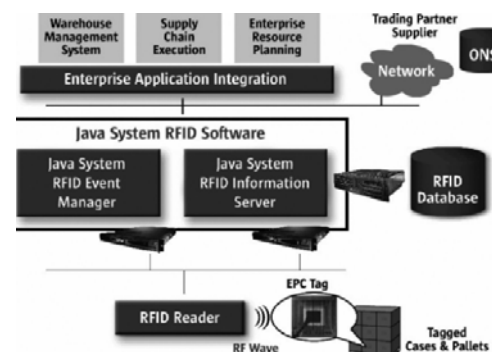


Fig. 2 : Sun Java System RFID Software Architecture [11].

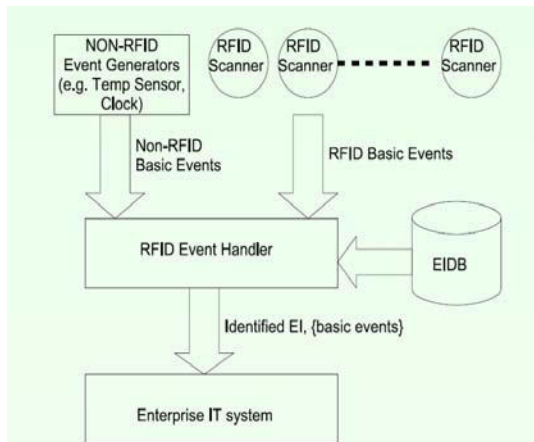


Figure 3. Event Driven Architecture [12].

Basic Events: that are generated by one or more source and is defined as (L, S, T); L for event label or description, S for location and T for time of happening. The base event source can be finding an object, a clock, environmental conditions triggering a sensor, or a part of application generating the event such as a part of the program produces the event.

Non-basic events: that can be monitored, controlled or reactivated. The events can be single or a combination of more than one event. Event causes one or more predefined reactions. Fig 3. Shows the event driven architecture. Events are generated by several sources. EIDB is the database of the events and stores the rules related to (L, S, T). EH uses the EIDB and events to identify the event and send it to Enterprise IT system then the system shows the appropriate reaction.

In the EIDB all the events are saved using a R-Tree based on priority and hierarchy. To recognize an event the Event handler uses the following algorithm:

- 1- Finding all the main events from the tree.
- 2- Calculating a set of base events.
- 3- Finding all the combinational events related to the base event.
- 4- Identifying the conditions and evaluating the results came from the above.
- 5- State graph tracing.
- 6- Sending the result to the IT systems.

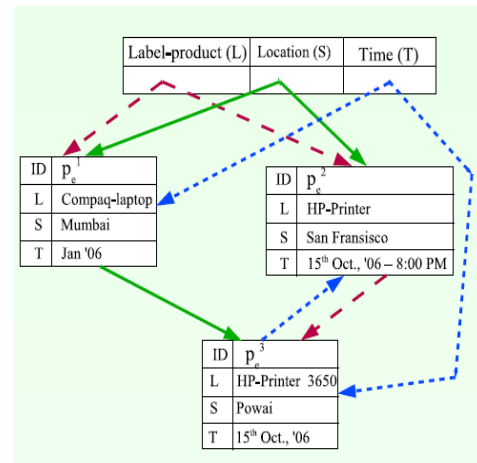


Figure 4. Storing EI [12].

Fig. 4 shows how the EI are stored in the database.

## V. PROPOSED REAL-TIME DISTRIBUTED EVENT DRIVEN ARCHITECTURE

Real-time response to requests and sending the real-time information gathered from the reader to the middleware and from middleware to the database is a necessary requirement of any RFID software. Therefore our proposed software architecture is both real-time and distributed because of the distributed nature of RFID systems.

According to the above facts we are proposing a real-time distributed architecture. This RTDED architecture is based on a five-layer real-time architecture. The mentioned technology is based on JAVA, hence the classes exist in the definition.

Five-layer software architecture is an exceptional efficient architecture for general structure of many embedded and real-time systems and is a special adoption of layered architecture. The structure is presented in Fig. 5 [13].

In Fig.6 RTDED Architecture method is presented. In hardware abstract layer tags, readers and in fact the hardware interface class which is responsible for data collection is presented. In operating system layer the operating system installed on server is defined. In communication layer the middleware is located and database is right next to the middleware. The application layer is a domain for Work Flows and Enterprise systems definition. User interface layer as the name suggests is associated to classes related to user and system interaction. Following a detail description of each layer is described.

<sup>1</sup> RTDED



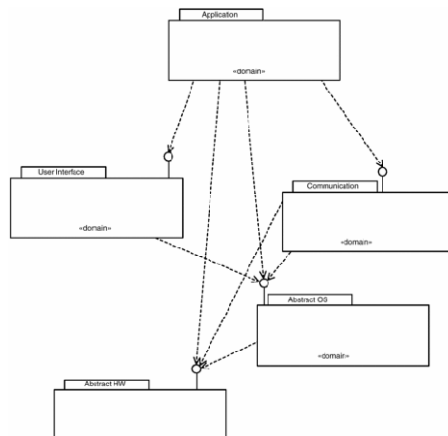


Figure 5. Five-Layered Real Time Architecture Structure [13].

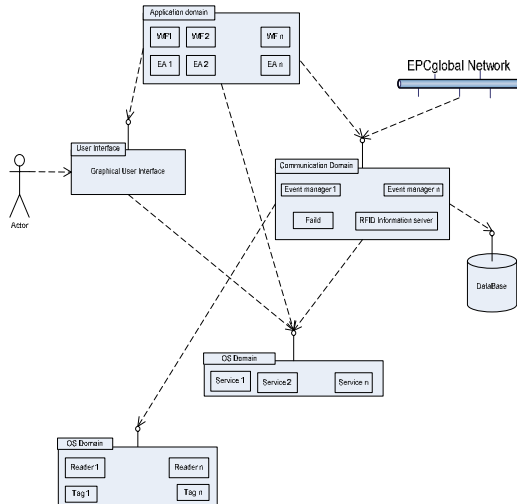


Figure 6. Proposed Architecture

#### A. Application Domain

Fig. 7 presents the Application Domain layer. This layer includes application classes and Work Flows defined based on the requirements. Also it includes classes related to the Enterprise System. In fact in this layer software and applications mentioned in user requirements are laid. This separation leads to extensibility of the system so that the addition of more requests to the system is easily managed. The gradual modification of system requirements is the main motivation of having this feature.

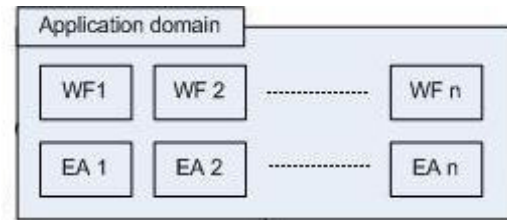


Figure 7. Application Layer.

Direct connection between the Application Domain layer and the database might be considered, depending on how large the system is. In large systems the direct connection between these two is not required as the Communication layer facility can be utilized; however in smaller systems there might be direct connections. (We would discuss this in more details in Communication layer chapter)

Pipelining can be used to handle fast response to the generated events of parallel Work Flows in order to get closer to real-time reactions.

In large systems This layer has the realization relationship:

- To the Operating System layer due to the fact that application resides on the operating system.
- To the User Interface layer to provide interaction to the client.
- And to the Communication layer to collect the data and events required by the applications.

And in small systems the relationship to the database is required.

Here we present an example to make the topic clearer. In a chemical material extraction and production, assets and personnel real-time tracking is implemented to locate each and also to define and control access to different part of the factory or to determine the warranty deadline of each product. Therefore a tag is given to each personnel and attached to the assets. As it is presented in Fig. 8:

- As an assumption the Entrance Work Flow is defined so that: if a person have the card with id X wants to enter a zone name L1, is he/she permitted or not.
- The Asset Location class is able to determine where exactly each asset is located in the real-time.
- In an event driven approach, Expired Cloth class can detect if an asset warranty has expired or not. And also notify the owner Y, to proceed to the related department for replacement or

maintenance. In case the asset is a personal protective equipment (PPE) i.e. clothing, helmets, goggles, or other garment designed to protect the wearer's body from injury, a routine can be called to warn the owner that she must change her cloth otherwise she cannot enter any zone in the factory until she returns the cloth and get a new one.

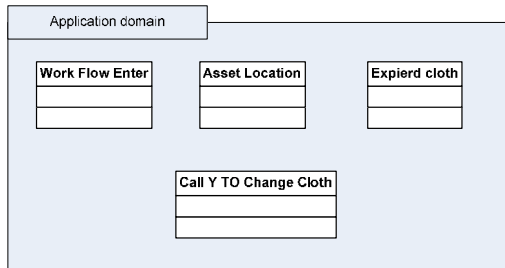


Figure 8. An application layer example.

### B. Communication Domain

Middleware is implemented through Communication Domain classes. Some middleware have a feature to define Work Flows. However this feature is defined in Application Domain to have better organized structure.

As it is presented in Fig. 9 the essential class required in every RFID system regardless of the size, is RFID Event Manager Class. The mentioned class is responsible for creation of an interface between the readers and other part of the system. This interface is required to transfer collected data from readers to the system. In this layer data collected from different readers with different formats should become compatible and interaction between them become possible. These are provided through Compatibility Domain subclasses.

Besides it should have the ability to filter duplicate data – the data gathered previously from another reader or the data gathered several times from one reader about on tag – and noises. This duty can be done by Filtering class. To manage detect and report events a subclass is needed called Event Handler. The Event Handler tries to remove duplication. However to have more reliability the Filtering subclass is also utilized. Finally a subclass is needed to insert data and events to the database which is called Insert DB.

There is another class in this layer called Information Server. It provides Application layer access to important events. Existence of this class gives more flexibility to the system and help to gather information faster and more easily. Especially in large systems where the business requirement of the system modifies. Nevertheless in small systems it would have a converse

result and produce overhead costs that threatens the real-time feature of the system in this scale. It is good to mention EPCglobal Network can access data and events through the same class.

Distribution of the architecture is available with implantation of several RFID Event Manager class for a group of readers on different PCs. Since the distributed systems should be able to adapt to the network modification a class called Failed is included in this layer. The Fail class verifies availability and functionality of each Event Manager on computational systems and in case any of them stops functioning the task can be handled by the closest Event Manager to reduce system failure.

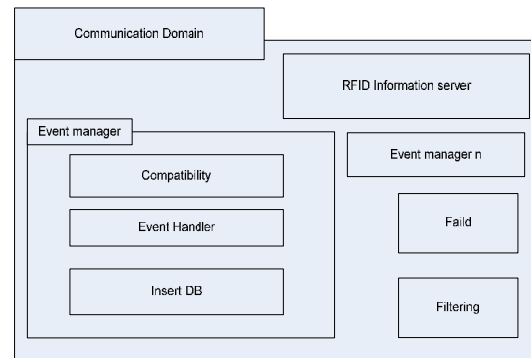


Figure 9. Communication Layer.

### C. OS Domain

As it is shown in Fig. 10 in this layer an operating system resides on each server is mentioned and it includes classes related to operating system services. Separation of this layer results in more portability and reusability. A part of distribution is clear in this layer because of several Event Managers. In large systems only required services can be implemented and have more efficacy and real-time reaction.

### D. User Interface Domain

The main responsibility of this domain is to provide interaction with the user. One side of this domain is the client and the other side is the operating system. Consequently it should include classes related to the user such as Window and Scrollbar and a middleware demonstration class to show a graphical view of the middleware and an interface class get user requests.

### E. Hardware Domain

In this domain Reader Driver classes for different readers such as hand-held and fixed readers from different manufacturers and sensors exist. Also the communication hardware infrastructure is defined in this layer.

## VI. EFFECT OF RFID ON SOFTWARE ARCHITECTURE

Due to the fast growth of RFID in industry and all along the supply chain and its real-time ability to response to system requirements, the software should also have the real-time feature. Therefore as it has been proposed in chapter V, the software architecture should have real-time nature and also be distributed due to distributed modality of RFID systems.

Since this technology is real-time it can answer to a lot of requests that without RFID were not even cost effective to be implemented. An example is the scenario we presented in chapter V about the personnel protective equipments being tagged in a chemical extraction and production factory and warn in case the PPE requires maintenance or replacement. These kind of solutions are now feasible with a reasonable cost and make improvement in business processes and finally customer satisfaction.

RFID is a pioneer in helping some solutions to be put into practice that were not feasible without RFID. As a result there are a lot more situations to be considered when designing a software. Answering to all these conditions is a value added to our software.

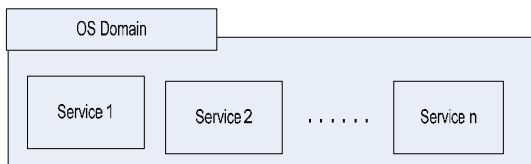


Figure 10. Operation System Layer.

In another word using this technology almost all the requests can be fulfilled. As an example a car manufacturer can tag all parts. When the customer visits an authorized service center to get a after sale support, it can be detected in no time if an original part has been replaced by a counterfeit one in a non-authorized repair shop or whether the warranty of each part is expired or not; only by looking into the manufacturing date of the part and the distance the vehicle has traversed. The maintenance history and task codes can be encoded in the tag for further references. Implementing such scenario makes the maintenance faster and more efficient.

## VII. EFFECT OF RFID ON SOFTWARE TECHNOLOGY

Another issue is the numerous data collected in RFID technology. Exploring these data, leads to useful results both from management point of view and quality aspects. In supply chain application, sharing these

numerous and Real-time data can improve EPCglobal Network targets in supply chain. Therefore there should be some facilities to handle such data in the software.

In addition to real-time response to the requests, the RFID technology brings more quality to the software. As now there are some requirements that can be fulfilled in presence of RFID that were not cost effective to be implemented without RFID.

## VIII. CONCLUSION

In this paper, the effect of RFID on the software technology especially on the software architecture was described, and a new architecture was proposed. The software should be Real-time and it should be distributed according to modality of RFID system.

RFID leads to implementing some requirements that without RFID are not even cost effective to be considered. Therefore while designing the architecture we should think about some new scenarios which are now easily implementable with RFID systems. Numerous data is another issue which is obtained from this technology. Research in these data, causes useful results in management and quality feature. In supply chain application, sharing these numerous and Real-time data can improve EPCglobal Network targets in supply chain.

## REFERENCES

- [1] Shirvanian, M., Fotovat Ahmadi, A. Modiri, N., "RFID effect on organization structure and information technology infrastructure.", 2nd International RFID Conference, 2009.
- [2] E. Karshenas, F. Sobhan Manesh, "RFID Auto ID System Architecture", 2nd International RFID Conference, 2009.
- [3] Y. Liue, T. Lin, S. Ram, and X. Su., "A Non-Invasive Software Architecture Style for RFID Data Provisioning", IJAL, vol. 1, 2010, pp 1-15, doi:10.4018/jal.2010090201.
- [4] L.Fengqun, C.Bocheng, C.Y Chan, Wu.C.H, Ip.C.H, A.Mai, H.Wang, W.Liu., "The Design of a Lightweight RFID Middleware", IJEBM, vol. 1, 2009, pp 73-78.
- [5] A Crimson Consulting Group, BizTalk Server Developer Productivity Study, Available from: [http://www.microsoft.com/biztalk/en/us/r\\_d.aspx](http://www.microsoft.com/biztalk/en/us/r_d.aspx), 2007. Knifsend, Oracle Fusion Middleware and Microsoft Interoperability: Addressing Enterprisewide Needs, Available from: <http://www.oracle.com>, 2005.

- [6] RFID Update, Sun Releases 2.0 Upgrade to RFID Middleware, Available from: [http://www.rfidupdate.com / articles/index.php?id=85](http://www.rfidupdate.com/articles/index.php?id=85), 2005.
- [7] C.Floerkemeier, C.Roduner , M.Lampe, "RFID Application Development with the Accada Middleware Platform", Systems Journal IEEE, vol. 1, Dec. 2007, pp 82-94, doi: 10.1109/JSYST.2007.909778.
- [8] \*P.Clement , L.Bass, R.Kazman, software architecture in practice, 2th ed., vol. 1. Addison-Wesley Professional, 2003.
- [9] J.Sun, & Q.Chen, "The RFID Middleware Development in the World and China", Logistics & Material Handling, Vol. 2, pp. 98-101, ISSN: 1007-1059, 2007.
- [10] sun.com, "The Sun Java System RFID Software Architecture", 2005.
- [11] A.Nagargadde, S.Varadarajan, K. Ramamritham, "Real-time Event Handling in an RFID Middleware System", 5<sup>th</sup> international conference on Databases in networked information systems, 2007.
- [12] Bruce Powel Douglass, Real-Time Design Patterns: Robust Scalable Architecture for Real-Time Systems, First ed, Addison Wesley, 2002, pp 99-102.



# The Repercussions of Technology Based Active Learning In Computer Science Education

Khamis Faraj Alarabi, P. F. Xavier Patrick Kishore & P. Sagaya Aurelia

Department of Department of Computer Science,  
Faculty of Education, Baniwalid, Libya

**Abstract-** This paper emphasizes on active learning in technology based Computer Science Education. The advantages of active learning are state to focus our demand. The existing active learning continuum is stated and there by the necessity of special and separate continuum for Computer Science Education is proposed. Next generation computer teachers must be highly research oriented therefore active pedagogical learning for next generation Computer teachers is a must. Specifically in our context, technology is used to enable the instructor to selectively show student work on a public display. An outcome from the study offers suggestions of how appropriate e-Technologies can be incorporated into learning environments for effective teaching.

**Keywords:** active learning, technology, computer based education

## I. INTRODUCTION

For many students, learning by doing is more effective than learning by listening. This approach is supported in activity-based instruction by having students work through examples, and then reviewing and discussing the results, instead of having the instructor work through the examples while students passively observe.

The requirement to learn how to use information technologies is the specialty of the IT professional however; in the field of education there are teacher training techniques which are the specialty of the teacher. These two fields need to be linked, and active learning is a perfect teaching model that can be utilized by both the IT professional and the Computer Science teachers.

## II. AN OVERVIEW OF ACTIVE LEARNING

In active learning, a few labeled instances are typically provided together with a large set of unlabeled instances. The underlying system is then rerun to improve the performance. This continues in an iterative fashion for convergence which typically is a threshold on the achievable performance before exhausting all the unlabeled data set. Table 1 states Kolb's four stages for learning styles.[4]

The outcomes suggest that students' participation in active learning activities steadily increase as students mature, with later year students more actively engaged than their first year counterparts.

The objective is to rank a set of instances in an optimal way for an external oracle to label them.

TABLE 1. KOLB'S FOUR STAGES FOR LEARNING STYLES (KOLB, 1984).

Stage	Dimension
The activist learning style	Doing and feeling
The reflector learning style	Feeling and watching
The theorist learning style	Thinking and watching
The pragmatist learning style	Thinking and doing

## III. ADVANTAGES OF ACTIVE LEARNING

- Student motivation is increased (especially for adult learners)
- Students are involved more than passive listening
- There is greater emphasis placed on the exploration of attitudes and values
- Students are engaged in additional activities (e.g., reading, discussing, programming)
- There is less emphasis placed on information transmission and greater emphasis placed on developing student skills.
- Students are involved in higher order thinking analysis, synthesis, (evaluation)
- Students can receive immediate feedback from their instructor [3]

## IV. THE ACTIVE LEARNING CONTINUUM FRAMEWORK

The Active Learning Continuum framework developed by Bonwell and Sutherland details the use of

four continuums to measure variables associated with the process of selecting an appropriate activity. The continuums are [1]

1. Task Complexity Continuum
2. Course Objectives Continuum
3. Classroom Interaction Continuum
4. Continuum of Student Experiences

## V. ACTIVE LEARNING CONTINUUM IN SPECIFIC FOR COMPUTER SCIENCE

Even though the active learning continuum framework is good still Computer Science education is a field, which is subjected to maximum up gradation. Computer Science education requires more précised and in depth framework of active learning Why? The development in Computer Science doesn't only affect the respective field but also have impact in all the other areas. Therefore we feel that active learning continuum should go much more beyond the framework, for the current technology based Computer Education system. The following are the active learning continuum framework proposed for the technology based Computer Science education.

### 1. Creative Continuum

Creativity is indeed thinking up new concepts. Creativity is dreaming up a new invention. When you bring something new into existence you can say you created it. This continuum has the originality of the source is one extreme and on the other end extreme of creativity. Even though time limit is not taken under consideration in this continuum importance is given to creative thinking. All new emerging technologies are the output of this type of creativity only. This makes the learner to think beyond the limits from the available input or resource. The continuum is shown in figure 1.

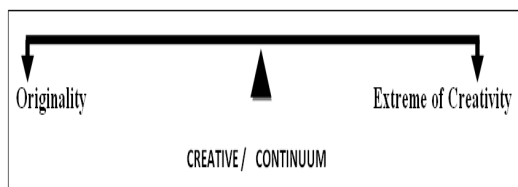


Fig. 1 : Creative thinking continuum

For example designing web page activities not only requires the originality of the source content (just the same program or similar HTML tags) but extreme creative and rational thinking. Fig (1.a) is an example of this activity.



Fig. (1.a). Example for Continuum / Creative / Rational

### 2. Motivation / Influence continuum

This continuum refers the impact of the concepts, ideologies or theories to be implemented in a technology based concept. This continuum also deals with the inter relation or the collection of ideas. For this motivation/influence continuum learners interested in relating concepts and exploring things using the concepts will involve more. This is the beginning stage for transforming learners to researchers. Figure 2 shows this continuum.

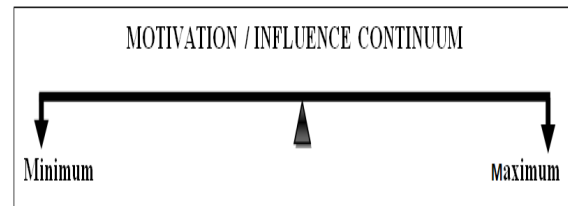


Fig. 2 : Motivation/Influence Continuum

For example troubleshooting in Computer Science education requires motivation and influence of the basic concepts and implementation of those concepts to an extended version.

This Motivation/Influence continuum should be irrespective of theory or practical learning. An example is shown in Fig. 2.a for maximum value and example for activity having poor motivation / Influence is shown in Fig. 2.b.

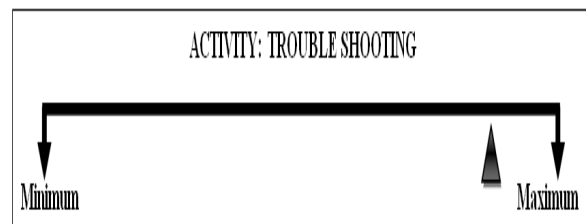


Fig. (2.a) High Motivation/Influence Continuum

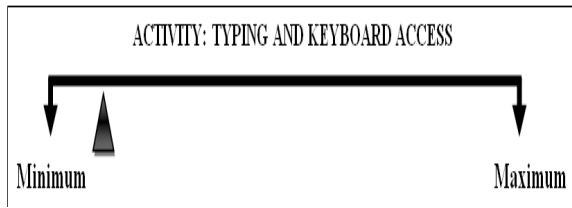


Fig (2.b) Low Motivation/Influence Continuum

### 3. Innovative continuum

Innovation is giving a practically improved shape to creativity. Innovation is about making creativity real. Innovation is making this new concept practical in a novel way. When you improve something that already exists you cannot say you created it but you can say you innovated it. Innovation plays a vital role in the field of Computer Science education. Innovation is the only reason for the tremendous growth in technologies. This continuum is shown in figure 3.

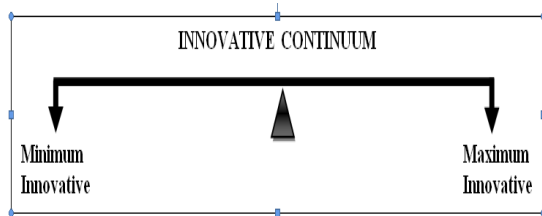


Fig. 3 Innovative continuum

For example in programming (irrespective of language) when we have to swap two numbers it can be done either using a third variable or not. Innovation is very important here in which program is written without using third variable which in turn reduces the memory space, NOL number of line in the program will decrease, and speed will increase. Examples are shown in Fig.(3.a) and Fig. (3.b)

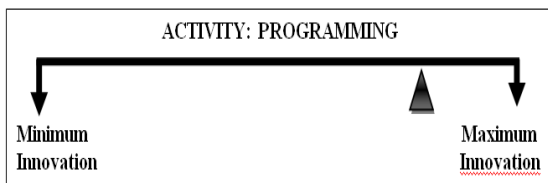


Fig (3.a) High innovative continuum

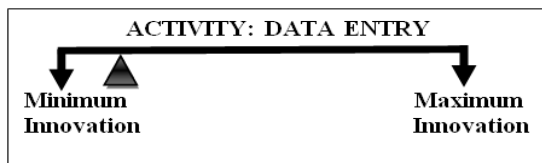


Fig (3.b) Low innovative continuum

### 4. Informative/Knowledge continuum

This continuum requires complete subject knowledge. This can be implemented in lecture as cross word puzzle, select from multiple choice, present paper from a given topic (one minute paper), and finding errors in the given program. It is a high correlation of knowledge, skills and abilities. Fig. 4 shows the continuum and Fig. (4.a) an example shows for this continuum.

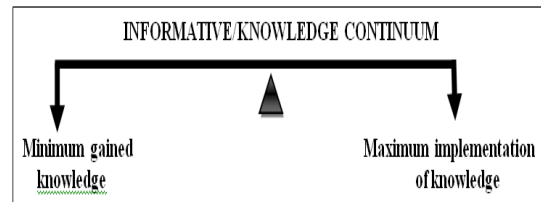


Fig. 4 Informative/Knowledge continuum

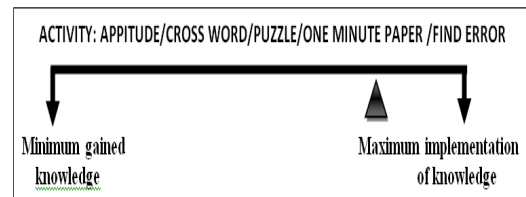


Fig. (4.a) Example for Informative/Knowledge continuum

In Computer Science education world which is based on technologies, creativity, motivation, innovation and Informative / knowledge are the most important criteria to make learning more active and effective.

### VI. SOME COMPUTER ACTIVITIES

Computer activities should be implemented as a part of learning process to make it more efficient. It should not be considered as a theory or practical part in academic but as an activity which will increase the learning ability of the student and make the lecture more interactive. Few sample activities are

- Desktop activity
- Email activity
- Internet activity
- Trouble shooting activity
- Typing and keyboard activity
- Word processing and spreadsheet activity
- Pc tuning activity
- Data entry
- Group discussion about last trends
- Exercises, pace and level of content
- Crossword/puzzle/aptitude /finding errors



- Article preparing
- Communicating through social networking

## VII. ACTIVE LEARNING PEDAGOGY FOR NEXT GENERATION COMPUTER TEACHERS

An active pedagogy means that students are active in their own learning and the classroom becomes a problem solving environment rather than a one way delivery or teacher centered environment. Active pedagogy coupled with active learning means creating a learning environment, or a classroom, where the students are encouraged to do something.

Another way at looking at this is “thinking about thinking”. What this means is a student has to think about what the meaning of the lesson is before the lesson is acted upon. [2]

- A. Active Research
- B. Cognitive Strategy.
- C. Active learning techniques, tools etc.,
- D. Evaluation

### A. Active Research:

Active research has three components: (1) Asking a question, (2) Acquiring data, and (3) Interpreting that data. If the teacher is going to ask the question about some aspect of computer in the classroom learning environment then an understanding of which aspect of the students learning behavior is being questioned must be understood. All the parts of the active learning pedagogy are in place they act together in a cycle whereby all the stages fit together and work simultaneously (see fig 5). Evaluation at the very center of the cycle because change is occurring very fast in the learning environment will as the data that is being added to each discipline which the student is studying. [2]



Fig 5. : Active learning pedagogy cycle

### B. Cognitive Strategy:

Cognitive activators are activities whereby students are asked to imagine how an idea or a fact, or even a procedure can be understood? As an example the student could be asked to demonstrate the shape of a car using a Lego block set or even a set of wooden blocks. The point is that one medium is used to describe another medium.

What the teacher is trying to do at this stage of the cycle is connecting the learner with the task using a pre-designed activity that has a similar cognitive routine as the one that is useful to complete the task. The cognitive activators are strategies that can be easily used to complete a task before the actual task is given. There are many different simple cognitive activators that can be used in preparing a student to learn a task.[2]

### C. Active learning technique, tools:

Pedagogically speaking the teacher in an active learning classroom should have predesigned steps prepared for the exercises a student will participate in, which in turn will challenge the student to increase their own learning skills. An easy exercise can be followed by a more difficult exercise until the teacher has fulfilled the design criterion for an active learning technique and the student has a working knowledge of the exercise.[2]

### D. Evaluation:

The teacher is always observing the outcomes of the student's actions. The cognitive activators will give immediate feedback as the student tries to complete a task not knowing that it is related to the learning goals the teacher has in mind. In this evaluation there is a lot of play and relaxation because it appears the tasks are easy. However, some students will prefer to do one thing one way and some another way. It is the goal to have the students feels as if they have been able to find the answer to their problems using their own ability. Why are we evaluating? The main purpose in evaluation is to improve projects by judging them and this comes by observation and this helps in modifying or changing a particular activity or program.

## VIII. ROLE OF THE TECHNOLOGY

The use of networked pen-based computers changes the logistics and facilitates the integration of activities into the classroom discussion in ways that would be difficult to achieve without the technology. Specifically in our context, technology was used to:

- Allow the instructor to privately preview (on a tablet) submitted student work;



- Collect the students' answers (as annotated slides) to an assigned activity and deliver them to the instructor as soon as students submit their responses;
- Distribute examples (on slides) to students in real time over a network;

The efficiency and flexibility of the submission process allow students to send in their responses as they finish an activity, which enables the instructor to evaluate student answers and gain an immediate impression of how students are doing on an activity. While the instantaneous distribution of activities in digital form to students is important for allowing efficient incorporation of activities into a lecture, it is the other direction where students can instantaneously submit their work back to the instructor that is very different from relying on paper and where some of the crucial benefits of using the technology lie.[5]

## IX. E-TECHNOLOGIES IN ACTIVE LEARNING

For the modern generation of students, technologies that they are familiar with and are used extensively in their day-to-day lives away from study seem to be the key to engaging them with their studies. In the analysis of the 12 interviews conducted for faculty had clear ideas about what they were trying to achieve and the pedagogical value of adopting and using particular eTechnologies to meet the learning needs of the students and make learning more enjoyable.

This is a study of faculty at an Australian university who are using eTechnologies in their teaching and discusses their perceptions of how these technologies help students to learn. From the above table we see that following are the advancement in active learning because of the growth in technology and it can also be used in various areas to improve the learning skills.[6]

- Active Learning through Podcasts and Vodcasts
- Active Learning through Virtual Worlds
- Active Learning through Synchronous Communication
- Active Learning through Shared Documents
- Active Learning through Wikis
- Active Learning through Social Networking

## X. CONCLUSION

The teacher is responsible for creating the activities the student will be exposed to and it is up to the teacher to monitor and evaluate the student's progress in every technique that is being used in the classroom.

**TABLE 2 : FACULTY RATING OF E-TECHNOLOGIES WITH GOOD PRACTICE PRINCIPLES\***

eTechnology Category	eTechnology Tool	Application and Context	Principles						
			1	2	3	4	5	6	7
Assessment and Survey Tools	Quiz	#01 Online Quizzes Science				√	√		
	Peer Assessment	#02 Peer Group Assessment Architecture		√		√			
Podcasts, Vodcasts and Streaming Virtual Worlds	Podcasts/ Vodcasts	#03 Professional Practice Public Relations	√	√	√*				
	Online Roleplay	#04 Video Scenarios Nursing			√*	√			√
Synchronous Communication	ESimulations	#05 Esimulations Information Systems	√	√	√*				
	ELive	#06 Student role play Psychology	√*	√	√				√
		#07 Online classes Property & Management	√		√*	√			
Shared Documents	Google Docs	#08 Team collaboration Marketing	√	√	√*				√
Wiki	Wiki	#09 Team collaboration Architecture		√	√	√	√		
		#10 Team collaboration Accounting & Finance		√	√				
Photosharing	Gallery 2	#11 External peer collaboration Arts	√			√		√	√
Social Networking	Facebook	#12 Transition & Support Law			√	√			

If all the stages are addressed simultaneously in The New-Generation of Teachers Project, then the ease to manage each stage will be made easier. The tools being developed are changing rapidly and knowledge accumulation is growing exponentially. If techniques are developed early in a students learning cycle then access to information using digital tools will be easier.

However, incorporating new technologies in innovative ways into a teaching portfolio requires time and effort on the part of the innovator. Thus it is important that sufficient time is allocated for staff to be able to experiment and implement these effectively and successfully. Further, the speed with which technologies come and go suggests that the study could be repeated quite regularly with differing outcomes.

## REFERENCES

- [1] <http://www.webster.edu/fdc/alhb/alhb2006.pdf>
- [2] Willard G. Van de Bogart "Active Learning Pedagogy" A new teaching methodology for a new generation of teachers
- [3] Charles C. Bonwell, Ph.D. "Active Learning: Creating Excitement in the Classroom "Active Learning Workshops
- [4] Wael Ibrahim, Rasha Morsi ,, and Theresa Tuttle," Concept Maps: An active learning and assessment tool in Electrical and Computer Engineering "Illinois-Indiana and North Central Joint Section Conference American Society for Engineering Education March 31-April 1, 2006 – Indiana University Purdue University Fort Wayne (IPFW) 2006

- [5] James A. Newell, Survivor: “ A Method for Active Learning in the Classroom that Addresses Student Motivation “, Proceedings of the 2004 American Society for Engineering Education Annual Conference & Exposition Copyright © 2004, American Society for Engineering Education
- [6] Felder, R. M., “The Effects of Personality Type on Engineering Students Performance and Attitude,”



# Predictive Traffic Light Control System : Existing Systems and Proposed Plan for Next Intersection Prediction

Prashant Borkar<sup>1</sup>, Amit R. Welekar<sup>2</sup>, Sanjeevani Jenekar<sup>3</sup> & S.P. Karmore<sup>4</sup>

<sup>1&2</sup>Department of CSE, GHRCE, Nagpur (MH) -440016, India,

<sup>3</sup>Department of EXTC, PCE, Nagpur (MH) -440019, India,

<sup>4</sup>Department of CSE, GHRCE, Nagpur (MH) -440016, India,

---

**Abstract**— This paper discusses a proposed systems for predicting the next intersection timing and generating the required speed at current intersection to cross next intersection without stopping at it. The system is speed module for next intersection prediction embedded in intelligent traffic light control system at intersection. It can also be designed for GPS based navigation system. For efficiently predicting the time and speed required for crossing next intersection without stopping at it centralized static approach is taken into account, the distance between current intersection and next intersection and traffic signal timings of next intersection are considered as input to the system. The traffic signal timings are more on highway than on city road. System then generates the required speed in range to cross next intersection without stopping at it. Speed generated by the system is in specified range like 32Km/Hr to 40 Km/Hr. Also it can't exceed the speed limit of road.

**Keywords:** Traffic congestion, Intersection, Traffic signal, required speed.

---

## I. INTRODUCTION

As the number of vehicle in urban areas is ever increasing, it has been a major concern of city authorities to facilitate effective control of traffic flows in urban areas [6]. Especially in rush hours, even a poor control at traffic signals may result in a long time traffic jam causing a chain of delays in traffic flows and also CO<sub>2</sub> emission. Vehicle's engine idling time consumes more fuels and releases more CO<sub>2</sub> than the vehicle in motion states. Thus if vehicle's idling time can be reduced, the amount of CO<sub>2</sub> emission can be decreased [5]. Poor traffic lights control is believed to account for longer intersection waiting time, thus suitable traffic lights control scheme is crucial for reducing CO<sub>2</sub> emissions. The total amount of accumulated delay time in a city due to waiting at signal stops is enormous if it is counted on an annual basis.

Also no one like to stop waiting at intersection, spending too much time at intersection may leads to driving stress. In many cities, these rising demands cannot be counteracted by further extending the existing road infrastructure giving a special importance to the efficient use of the existing network. In this respect, traffic lights are a vital factor since good control strategies are often capable of improving the network-wide traffic flows. To reduce the waiting time of vehicles at traffic signals is to reduce consumption of fuel and man-hours, thus it is significant to control the traffic signals in an effective manner. In the remainder of paper, an effective time management plan is proposed where the timings of next intersection are predicted at current intersection and accordingly required speed is

generated in range to cross next intersection without stopping at it, so that the effective waiting time at the intersection and corresponding traffic congestion can be reduced too much extent. In ideal case no one would have to wait at intersection.

## II. EXISTING INTELLIGENT TRAFFIC LIGHT CONTROL SYSTEMS

Before moving on to the proposed plan it is necessary to consider some of the existing systems and architectures. In general the systems are centralised control and dynamic controls also there are some systems which are Hybrid.

### A. 'Organic' Traffic Lights Sense Traffic and Adjust Light Timing Accordingly

Holger Prothmann developed organic computing approach to develop a decentralized traffic control system and compared its impact on traffic flow with a conventional system. "The organic approach is based on industry-standard traffic light controllers," These have been adapted to have an observer/controller architecture that allows the traffic light to respond to traffic flow and to pass on information to the other traffic lights on neighbouring roads. In the case of an urban traffic system, the sensors would be closed-circuit TV cameras mounted on road gantries and other places while the controllers, or actuators, would be traffic lights, which can effectively start and stop the flow of traffic [1].

### B. RHODES: A REAL-TIME TRAFFIC SIGNAL CONTROL SYSTEM

An adaptive real-time traffic signal control system referred to as RHODES. RHODES is a "dynamic network loading" model that captures the slow-varying characteristics of traffic. These characteristics pertain to the network geometry (available routes including road closures, construction, etc.) and the typical route selection of travellers. In general the system takes as input detector data for real-time measurement of traffic flow, and "optimally" controls the flow through the network. The system utilizes a control architecture that (1) decomposes the traffic control problem into several sub problems that are interconnected in an hierarchical fashion, (2) predicts traffic flows at various levels (individual vehicles and platoons) (3) allows various optimization modules for solving the hierarchical sub problems, and (4) utilizes a data structure and computer/communication approaches that allow for fast solution of the sub problems [2].

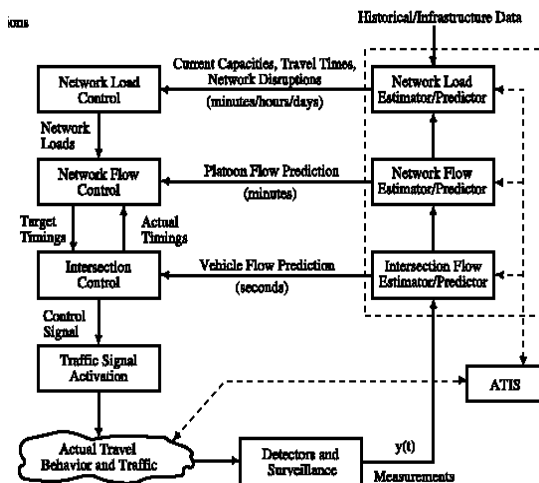


Fig 1. RHODES hierarchical architecture

Based on the slow-varying characteristics of the network, estimates of the load on each particular link, in terms of vehicles per hour, can be calculated. Which then allow RHODES to allocate "green time" for each pattern and each phase. These decisions are made at the middle level of the hierarchy, referred to as "network flow control". Traffic flow characteristics at this level are measured in terms of platoons of vehicles and their speeds. Given the approximate green times, the "intersection control" at the third level selects the appropriate phase change epochs based on observed and predicted arrivals of individual vehicles at each intersection.

### C. Urban Arterial Traffic Two-direction Green Wave Intelligent Coordination Control Tech.

The architecture includes two layers - the coordination layer and the control layer. Public cycle time, splits, inbound offset and outbound offset are calculated in the coordination layer. Public cycle time is adjusted by fuzzy neural networks (FNN) according to the traffic flow saturation degree of the key intersection. Splits are calculated based on historical and real-time traffic information. Offsets are calculated by the real-time average speeds. The control layer determines phase composition and adjusts splits at the end of each cycle. The target of this control strategy is to maximize the possibility for vehicles in each direction along the arterial road to pass the local intersection without stop while the utility efficiency of the green signal time is at relatively high level. [3]

Green wave control is a kind of arterial traffic coordination control method [3], that coordinates traffic signals of adjacent intersections on arterial road to make vehicles driving by a certain speed meet no or less red lights. In other word, traffic signals of adjacent intersections become green one by one according to a certain time sequence in a direction, like a rolling "green wave".

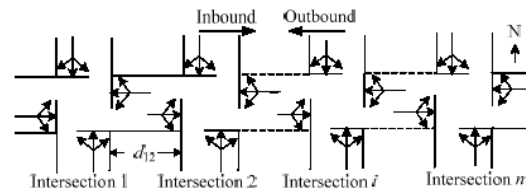


Fig 2: Sketch map of urban arterial road

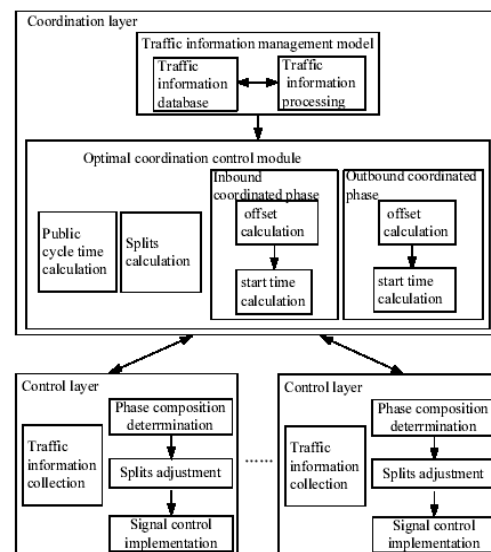


Fig 3: system architecture (Structure of hierarchical control system)

### III. PROPOSED PLAN OF WORK

This section describes the proposed system that is Next intersection prediction module for intelligent traffic light control system. The proposed plan of work is centralized static approach in which all the intersections are centrally connected and static time (time for which lights at intersection are green or red) is provided as input to the intelligent traffic light control system in advance and according to that required speed is generated in range at current intersection to cross next intersection without stopping at it. All the traffic signal timings at particular intersection are considered according to signal timings of Nagpur city, India. The traffic signal timings at intersection may vary from city to city also all the calculations are done according to intersection timings of Nagpur city.

#### A. Traffic Signal Timings

This is the very important aspect to be considered while designing proposed system where in the timings for which particular traffic signal is green or red are preloaded into traffic signal control box (refer to Figure 4 and 5) and these timings are fixed throughout the day and we are assuming here that the traffic signals are centrally connected that is once the main system is initialised the entire traffic signals of particular city are start at a time. If there are any updates in traffic light timings then new timings are to be loaded into traffic signal control box for particular intersection or more than one. Generally traffic signal timings are to be updated because of several festivals or certain events such as cricket match etc.

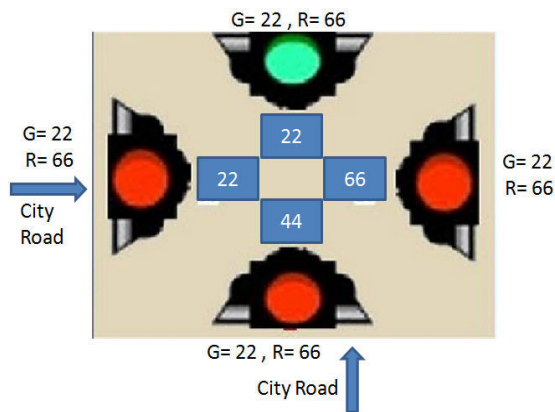


Fig 4: Normal Traffic signal timings (Intersections on city road)

The figure 4 refers to intersection timings of city intersection which come on city road. Here we see that all have to wait at intersection for equal time duration and also green timings for all 4 sides are same that is 22 sec.

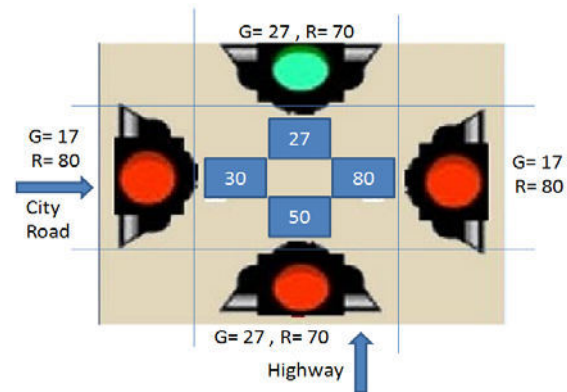


Fig 5: Increased traffic signal timings on highway

The fig. 5 refers to intersection which comes on highway; from the figure we can see that for traffic on highway the green timing i.e. 27 sec and waiting time that is red= 70 sec. Compared to green timing = 17 sec and red time = 80 sec on city road. This is because due to more percentage of traffic on highway road as compared to city road. Also the timing provided for various traffic signals are may be different or we can even provide the timings according to rush for particular time of the day. Ex. Traffic intensity is very much high during 9 to 11 am and 4-8 pm as it office, school, college timings etc. So for this particular time period the traffic light timings can be increase by few more seconds to avoid congestion.

#### B. Neighbouring Intersections and Approximate calculations for generating required Speed to cross next intersection

Until now we considered only traffic signal timings for particular intersection or it can be generalised to more than one. This section of the paper provides the approximate calculations of required speed so as to cross next intersection without stopping at it. Also the required speed is in specified range so that if we follow the speed in range we don't have to stop at intersection. Here we are considering five intersections of a particular city denoted by A, B, C, D and E. (refer to fig. 6) Traffic flow is indicated by arrows. The circle and star represents the traffic light control system. At any given intersection there are 4 traffic light controls system and every traffic light system controls the flow of traffic coming towards it as indicated by an arrow. The traffic flows from intersection A to intersection B are on highway and traffic flows from intersection C and intersection D are on normal city road. The timings for intersection A, B and F are increased as compared to intersection C and D, because they are on highway.

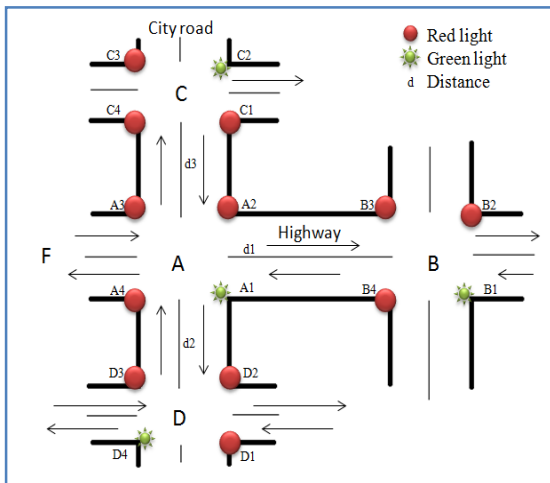


Fig. 6: View of 5 intersections of a particular city

Let us consider the intersection A as shown in Figure 6. Consider the flow of traffic coming towards intersection as indicated by arrow from intersection F, also  $d_1$ ,  $d_2$ , and  $d_3$  are the distances from intersection A to intersection B, D and C respectively. In proposed plan every traffic signal light at intersection knows the traffic light timings of next signal which is ahead of it and to the right of it and to the left of it. From the fig.6 the node A.1 of intersection A knows the timings of node B.1 of intersection B and node D.4 of intersection D and node C.2 of the intersection C, here the term node refers to the traffic light control box (every intersection has 4 traffic light control box).

Moving on to the approximate calculations for next intersection prediction at current intersection and generating required speed in range to cross next intersection at current one without stopping at it. All the calculations are done according to traffic signal timings of Nagpur city, these calculations may vary from city to city.

As all the traffic light control system at intersection of entire city are centrally connected that is once main switch on the entire traffic signals for that city are on at a time. Let at particular instance of time  $t_1$  all the traffic signals are on. As all light box cannot be green at time anyone can be green at a time, from the fig. 6 consider node A.1 of intersection A, node B.1 of intersection B, node D.4 of intersection D, node C.2 of intersection C and so on for entire city are becomes green at time  $t_1$ . Consider the timings for intersection A and B equal to the timings shown in fig. 5. Let  $d_1$  be the distance between intersection A and B which is equal to 1KM. at time instance  $t_1$  both intersections timings will look like this.

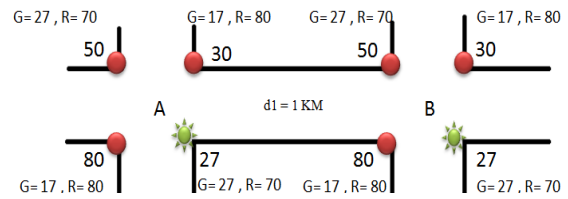


Fig. 7: signal timings of intersection A and B

From the figure we see that at time  $t_1$  the node A1 and node B1 are on i.e. green and duration for which it remains on is 27 sec. for the traffic flow from A to B, and other timings such as 30, 50 and 80 seconds are the red timings for the traffic flow in other directions as from A to C, from B to A, and from A to D respectively similar for B. The person at A going towards B the distance he/she has to cover is 1 KM. and the speed limit of this highway is 45 KM/hr. aim is to cross intersection B without stopping at it. Now consider following cases:

*Case 1:* scenario where person at intersection A has to cover a distance 1 KM in 27 sec. for this the speed required is  $1000/27 * 18/5 = 133.333$  KM/hr which exceed the speed limit of road. Not possible

*Case 2:* person just reaches intersection B when it just on so he/she still have 27 sec to cross that intersection. Here person at intersection A has to cover a distance 1 KM in  $27+17+27+17 = 88$ sec. for this the speed required is  $1000/88 * 18/5 = 40.9$  KM/hr. Is possible and this is the upper limit of range.

*Case 3:* person just reach intersection B when it going to become red (here he/she don't have to wait at intersection B. or just for 3 to 6 sec if he/she reach earlier) here person at intersection A has to cover a distance 1 KM in  $27+17+27+17+22 = 110$ sec. here last 22 sec. are considered for safely crossing intersection. For this the speed required is  $1000/110 * 18/5 = 32.73$  KM/hr. This is possible and is lower limit of range. So in general if person follows speed in range 32.73 Km/hr to 40.9 KM/hr, that person doesn't have to wait at intersection B.

Now consider the traffic flow from intersection A to intersection C and D. let us assume the distance  $d_2$  and  $d_3$  be 0.8 KM. and at time  $t_1$  the node D4 and node C2 is on i.e. green. First consider the traffic flow from A to D. at  $t_1$  timings of intersection D will look like this.

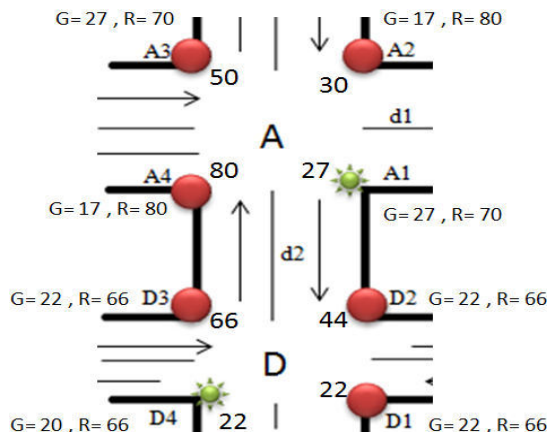


Fig. 8: signal timings of intersection D

The person at A going towards D the distance he/she has to cover is 0.8 KM. and the speed limit of this city road be 40 KM/hr. Aim is to cross intersection D without stopping at it. Now consider following cases:

*Case 1:* scenario where person at intersection A has to cover a distance 0.8 KM in 22 sec. for this the speed required is  $800/22 * 18/5 = 130.9$  KM/hr which exceed the speed limit of road. Not possible

*Case 2:* person just reaches intersection D when it just on so he/she still have 22 sec to cross that intersection. Here person at intersection A has to cover a distance 0.8 KM in  $22+22+22+22 = 88$ sec. for this the speed required is  $800/88 * 18/5 = 32.72$  KM/hr. Is possible and this is the upper limit of range.

*Case 3:* person just reach intersection D when it going to become red (here he/she don't have to wait at intersection B. or just for 3 to 6 sec if he/she reach earlier) here person at intersection A has to cover a distance 0.8 KM in  $27+22+22+22+17 = 105$ sec. here lat 17 sec. are considered for safely crossing intersection. For this the speed required is  $1000/105 * 18/5 = 27.42$  KM/hr. This is possible and is lower limit of range. So in general if person follows speed in range 27.42 Km/hr to 32.72 KM/hr, that person doesn't have to wait at intersection D. similar calculations for intersection C.

#### IV. PREDICTIVE TRAFFIC LIGHT CONTROL SYSTEM AT INTERSECTION

The above mentioned system can be look like this at any intersection.

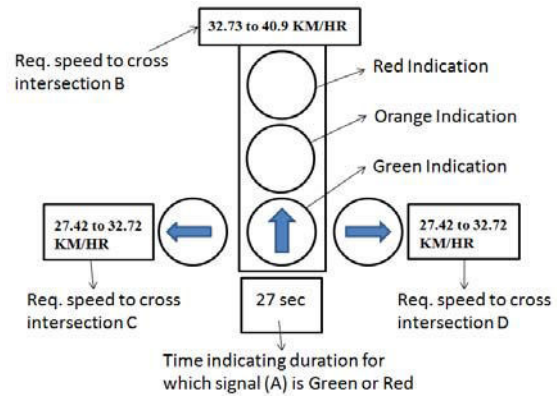


Fig. 9: Predictive traffic light control system at intersection A

Here the green indication is symbol that intersection is on. Orange is for warning generally it is of 3 sec. the red one is the indication to wait at intersection. The box 27 sec. indication the intersection is on for 27 sec. the box 27.42 to 32.72 KM/HR indicates the required speed in range to cross side intersection such as intersection D and C without stopping at it with respect to intersection A. the box 32.73 to 40.9 KM/HR indicates the required speed in range to cross next intersection B without stopping at it.

#### Advantages:

- ❖ In ideal case if everyone follows the system then no one has to wait at intersection.
- ❖ There will be sufficient reduction in average waiting time at intersection.
- ❖ There are fewer accidents because the platoons of vehicles arrive at each signal when it is green, thereby reducing the possibility of red signal violations.
- ❖ By using above mentioned system vehicle's idling time can be reduced, the amount of CO2 emission can be decreased
- ❖ Traffic flow are more smoothly, often with an improvement in capacity
- ❖ Vehicle speeds are more uniform because there is no incentive to travel at excessively high speeds to reach a signalized intersection within a green interval that is not in step. Also, the slow driver is encouraged to speed up in order to avoid having to stop for a red light.



### **Limitations:**

- ❖ The system only considers the static traffic signal timing data. Therefore for certain events such as during festivals, sport matches etc. the system may not work properly due to rush on path.
- ❖ Here traffic intensity is not taken into consideration, if we consider the traffic intensity on particular road then accordingly it becomes dynamic approach and in such case continuously changing signal timings are to be fed to the traffic light control box to generate the required speed.

### **V. CONCLUSION**

The improvement of urban traffic condition is largely dependent on the various modern techniques of traffic management and control. Advanced traffic signal controllers and control system contribute to the improvement of the urban traffic problem. The proposed plan for next intersection prediction at current one and then generating the required speed to cross next intersection without stopping at it is discussed in this paper. Speed generated is in range. The primary goal is to minimize the waiting time at the intersection and alternately having smooth flow of traffic. By using above mentioned system vehicle's idling time can be reduced, the amount of CO<sub>2</sub> emission can be decreased.

For the future work we will consider the scenario more complicated also traffic light timings are continuously changing according to traffic flow and the generating the required speed to cross next intersection. We can also implement this approach at GPS based navigation device.

### **REFERENCES**

- [1] Holger Prothmann, Jürgen Branke and Hartmut Schmeck "Organic traffic light control for urban road networks " Int. J. Autonomous and Adaptive Communications Systems, Vol. 2, No. 3, 2009.
- [2] Pitu Mirchandani, Fei-Yue Wang "RHODES to Intelligent Transportation Systems" 1541-1672/05/ © 2005 IEEE. IEEE INTELLIGENT SYSTEMS
- [3] Xiangjie Kong, Guojiang Shen, Feng Xia, and Chuang LinUrban. "Arterial Traffic Two-direction Green Wave Intelligent Coordination Control Technique and Its Application" International Journal of Control, Automation, and Systems (2011) 9(1):60-68
- [4] W. Wen. "A dynamic and automatic traffic light control expert system for solving the road congestion problem" 2007 Elsevier Ltd. All rights reserved. doi:10.1016/j.eswa.2007.03.007. science direct a Expert Systems with Applications 34 (2008) 2370–2381
- [5] Chunxiao LI, Shigeru SHIMAMOTO. "A Real Time Traffic Light Control Scheme for Reducing Vehicles CO<sub>2</sub> Emissions" The 8th Annual IEEE Consumer Communications and Networking Conference - Emerging and Innovative Consumer Technologies and Applications.
- [6] Chen Xiao-feng, Shi Zhong-ke, Zhao Kai "Research on an Intelligent Traffic Signal Controller" 0-7803-8125-4/03 © 2003 IEEE





# Bi-Objective Optimization of Distributed Power Systems

## Design using Neuro-Swarm Technique

T. Ganesan<sup>1</sup>, I. Elamvazuthi<sup>2</sup>, P. Vasant<sup>3</sup>, M.K.A. Ahamed Khan<sup>4</sup>,  
Rajendren Sinadurai<sup>5</sup> & M. Amudha<sup>6</sup>

<sup>1,2&3</sup>Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak Darul Ridzuan, Malaysia

<sup>4&5</sup>University Industri Selangor, Kampus Bestari Jaya - 45600 Berjuntai Bestari, Selangor Darul Ehsan, Malaysia

<sup>6</sup>Cosmopoint, KLMU, Malaysia

---

**Abstract** - The rapid development of renewable energy sources in power systems is due to the general apprehension concerning the environmental impacts and increasing energy costs involved with the usage of fossil fuels as an energy source. One recent development in the direction of producing greener and cost effective sources is the idea of distributive generation (DG) power systems. Issues such as cost effectiveness as well as pollutant emissions arise during the development of these power systems. The optimization of the DG power system with alternative power sources was carried out effectively in the previous work using Particle Swarm Optimization (PSO) algorithm. In this work, the optimization was performed further using two techniques. These were the PSO and the Hybrid Hopfield Neural Networks - PSO (HPSO) algorithms. Some comparative analysis was then done to seek out the effectiveness of the application of the algorithm in this problem.

**Key words** - *bi-objective, optimization strategy, hybrid algorithms Particle swarm optimization (PSO), hybrid Hopfield neural networks-PSO (HPSO), alternative energy, distributed generation (DG).*

---

### I. INTRODUCTION

DG power systems have been rapidly being applied over the past decade. This is due to the arising needs concerning the environmental impacts as well as elevated energy costs in this era of increasing population growth [1], [2]. The concept of DG is a power source which is connected directly to a distribution network [2].

In the advent of diminishing fossil fuel and increasing demand in the power consumption, many alternative energy power sources have been integrated into the DG systems. Some alternative energy sources are; wind turbines, PV arrays, fuel cells, and fossil fuel-fired generators. However, although renewable energy sources are highly favourable in DG systems, reliability issues are frequently encountered. This is mainly, due to the fact that the sources of alternative energies are dependent on factors such as climate conditions (eg: cloud cover, average wind speed, irradiation rates, etc, [2-4]). The integration of alternative power sources into the DG systems is called hybrid power generation systems [5-7]. A detail review on cost effectiveness, reliability and pollutant emissions can be seen in [8] and [9]. Wang *et al.* [10], applied the particle swarm optimization (PSO) methods in the design of DG systems. Other works on the design and sizing of hybrid

power systems with solar and wind power sources include Dehghan *et al.* [11] and Chedid *et al.* [12].

Optimization methods such as Hopfield Neural Networks and PSO were used in this work. These strategies were applied to the design of the DG system with alternative energy fuel source. Comparative studies as well as algorithmic performance benchmarking were then performed to identify the best optimization strategy that achieves all the objectives and obeys all the power balance and design constraints.

The Hopfield Artificial Neural Nets (HNN) was developed by Hopfield [13] and [14]. Neural nets observed to have many applications in multi-objective optimization problems (for instance, see Guillermo Jimenez *et al* [15], Guillermo Jimenez *et al* [16] and Chayakulkheeree *et al* [17]). Unlike regular neural nets that may produce divergent solutions, this property enhances the convergence capabilities of the HNN algorithm. The HNN uses reinforced-type learning (modified Hebbian learning) to update the weights in each recursion.

Particle Swarm Optimization (PSO) is an optimization algorithm developed based on the swarming behaviours of certain types of organisms. PSO was developed by James Kennedy and Russell Eberhart [18]. Recently, PSO has been applied to

various fields in optimization especially in electrical power systems (see [19], [20] and [21]).

This paper is organized as the following: Section II introduces the PSO and the HPSO approaches; Section III presents the problem description for the optimization of the DG system. The computational results are included in Section IV and the analysis and discussion of the results are presented in section V. The paper ends with concluding remarks and recommendations for future research work.

## II. OPTIMIZATION TECHNIQUES

### A. Particle Swarm Optimization (PSO)

The PSO algorithm works by searching the search space for candidate solutions and evaluating them with respect to a specified fitness function. The velocity and position updating rule provides the recursive optimization capabilities of the PSO algorithm. The velocity of each particle in motion or 'swarming' is updated using the following equation.

$$v_i(n+1) = wv_i(n) + c_1r_1[\hat{x}_i(n) - x_i(n)] + c_2r_2[g(n) - x_i(n)] \quad (1)$$

Where each particle is identified by the index  $i$ ,  $v_i(n)$  is the particle velocity and  $x_i(n)$  is the particle position with respect to iteration ( $n$ ). The parameters  $w$ ,  $c_1$ ,  $c_2$ ,  $r_1$  and  $r_2$  are initialized by the user.

After the computations of the particle velocity, the particles positions are then calculated as follows:

$$x_i(n+1) = x_i(n) + v_i(n+1) \quad (2)$$

The iterations are continued until all candidate solutions are at their fittest positions in the fitness landscape and some stopping criterion which is set by the user is met. For more PSO based applications, refer to [22] and [23].

For the PSO algorithm, the fitness criterion is defined such that ; If during the iteration process, the position of all the particles converges to some constant, no further optimization occurs in the objective function, no constraints are broken and all the decision variables are non-negative then the program halts. The PSO algorithm is as follows:

- Step 1: Set no of particles,  $i$  and the algorithm's initial parameters.
- Step 2: Specify initial position  $X_i(n)$  and velocity  $V_i(n)$
- Step 3: Compute individual and social influence.
- Step 4: Compute position  $X_i(n+1)$  and velocity  $V_i(n+1)$  at next iteration

Step 5: If the swarming time,  $n > n_o + T$ , update position  $X_i$  and velocity  $V_i$  and go to Step 3, else proceed to step 6

Step 6: Initialize DG system coefficients and design parameters

Step 7: Evaluate fitness of the design parameters.

Step 8: If fitness function is satisfied, halt and print solutions, else go to step 3.

where  $n_o$  is some constant,  $n$  is the swarm iteration and  $T$  is the overall program iteration.

### B. Hybrid Neuro - Particle Swarm Optimization

In this work the Hopfield Neural Net and the PSO algorithms was hybridized and used as an optimization algorithm. The hybrid algorithm HPSO is as follows:

Step 1: Initialize no of particles,  $i$  and the algorithm parameter setting.

Step 2: Set initial position  $X_i(n)$  and velocity  $V_i(n)$

Step 3: Compute individual and social influence

Step 4: Compute position  $X_i(n+1)$  and velocity  $V_i(n+1)$  at next iteration

Step 5: If the swarm evolution time,  $n > n_o + T$ , update position  $x_i$  and velocity  $v_i$  and go to Step 3, else proceed to step 6

Step 6: Initialize power system coefficients and design parameters

Step 7: Evaluate fitness of the design parameters.

Step 8: If fitness criterion satisfied, halt and print solutions, else go to step 3.

where  $n_o$  is some constant,  $n$  is the swarm iteration,  $m$  is the network recursion and  $T$  is the overall program iterations.

## III. MATERIALS & METHODS

The setup of the DG system is limited to wind turbine generators (WTGs), photovoltaic cell panels (PVs), storage batteries (SBs) and the fuel-fired generators (FFGs). During the design of a DG power system of this type, issues such as cost and impact on the environment is taken into account. In this work cost and pollutant emission are considered as the bi-objective functions and the range of the design parameters and power balance equations are taken as constraints. Since wind and solar power are relatively lower in terms of reliability as compared to fossil fuel as a energy source, the inclusion of storage batteries into the DG system is highly desirable.

This energy storage method can thus filter-off power fluctuations and give a constant amount of power supply for any given time [24]. The storage batteries are analogous to a voltage/current regulator that balances the variation in supply and demand. Individual power sources cater differently for cost effectiveness and pollutant emissions.

This is a nonlinear problem that involves 9 constraints and 67 decision variables. The problem statement is formulated as the following:

$$\text{Min} \rightarrow \text{COST} (\$/\text{yr})$$

$$\text{Min} \rightarrow \text{Pollutant Emissions (PE)} (\text{ton/yr})$$

subject to *power balance* and *design* constraints

For the definitions and the notations used in this paper, please refer to the nomenclature section below. The objective functions (refer to [6], [7] and [12]), for the overall cost, COST (\$/yr) is as the following:

$$\text{COST} = \frac{\sum_{i=w,s,b} (I_i - S + OM_{P_i})}{N_p} + C_g \quad (3)$$

The objective function for the pollutant emissions which was quadratically approximated (see [25]) is as the following:

$$PE = \Omega + \Phi \times \sum_{t=1}^T (P_{g,t}(t)) + \Gamma \times \left[ \sum_{t=1}^T (P_{g,t}(t)) \right]^2 \quad (4)$$

The detail calculations of the output power by the PV and the WTG are given in [7].

*A. Design constraints.*

$$A_{w\min} \leq A_w \leq A_{w\max} \quad (5)$$

$$A_{s\min} \leq A_s \leq A_{s\max} \quad (6)$$

$$P_{b\min} \leq P_{bsoc} \leq P_{bcap} \quad (7)$$

$$0 \leq P_{bcap} \leq P_{bcap\max} \quad (8)$$

$$P_b \leq P_{b\max} \quad (9)$$

$$P_{g\min} \leq \sum_{t=1}^T P_{g,t} \leq P_{g\max} \quad (10)$$

$$0 \leq \kappa \leq 1 \quad (11)$$

*B. Power balance constraints:*

$$P_b(t) + P_w(t) + P_s(t) + P_g(t) \geq (1 - R)P_d(t) \quad (12)$$

$$P_b(t) + P_w(t) + P_s(t) + P_g(t) - P_{dump}(t) \leq P_d(t) \quad (13)$$

The data used as input design parameters and the hourly input of the insolation, wind speed patterns and the hourly load demand in this work was obtained from [10].

#### IV. RESULTS & ANALYSIS

In this work, the algorithms were programmed using the C++ language on a personal computer with an Intel dual core processor running at 2 GHz. The objective functions, cost (\$/yr), and pollutant emissions, (PE) in tonnes/yr was minimized using the PSO and the hybrid HPSO approaches. The results were then compared against each other as well with the results obtained in [10]. In the results, design 1 and design 2 were the numerical results obtained from [10] while PSO and the HPSO are the results of the algorithms developed in this work.

The comparisons of the values of the objective functions are as in Table I. The comparisons of the design parameters for each of the methods are provided in Table II. functions (cost (\$/yr), EIR, emissions (ton/yr)).

TABLE I. COMPARISON OF VALUE OF OBJECTIVE FUNCTIONS

	<i>design 1</i>	<i>design 2</i>	<i>PSO</i>	<i>HPSO</i>
Cost (\$/year)	5323	6802	5340.95	5326.2
Emissions (ton/year)	12.460	65.381	4.9157	6.2916
Execution time (msec)	-	-	67	127

TABLE II. COMPARISON OF THE OPTIMIZED DESIGN PARAMETERS FOR EACH METHOD

	<i>design 1</i>	<i>design 2</i>	<i>PSO</i>	<i>HPSO</i>
--	-----------------	-----------------	------------	-------------

$A_w(m^2)$	420	640	432.406	434.777
$A_s(m^2)$	50	40	43.2406	43.4777
$P_{bcap}(kWh)$	16	16	16.2571	15.3477
$\kappa$	0.2	0.58	0.12225	0.136145

The lowest pollutant emission rate (PE) was achieved using the design produced by the PSO method while the highest is given by design 2. The second lowest PE was produced by the HPSO method followed by the design 1 method. It can be observed the PSO method was the greenest design option since it minimizes the PE effectively.

In the design proposed by the PSO method, it can be seen that although it provides the lowest emissions among all the other designs, it was not as cost effective as the design proposed by design 1 and the HPSO methods. The design by the HPSO method was more cost effective as compared to the design by the PSO and design 2 methods, however more costly than the design 1 method. In terms of emissions, the design by the HPSO method has lower PE as compared to all except the PSO method. Thus, it can be said that the HPSO method was the second greener alternative design after the design by the PSO method. The Pareto front of the results obtained by the PSO and the HPSO algorithms are as the following:

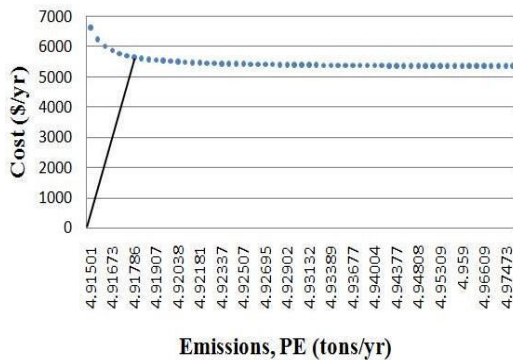


Fig. 1 : Pareto front of the results obtained by the PSO algorithm

In Figs. 1 and 2, the black lines indicates the solution that has the most minimal distant with respect to the origin (0,0). In essence the solution with the most minimal distance is considered to be the most dominant solution achieved by each of the algorithm in this work. The position of the particles versus the number iterations ( $t$ ) is shown in Fig. 3.

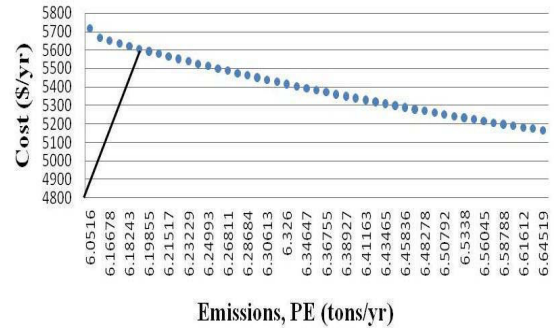


Fig. 2 : The Pareto front of the results obtained by the HPSO algorithms

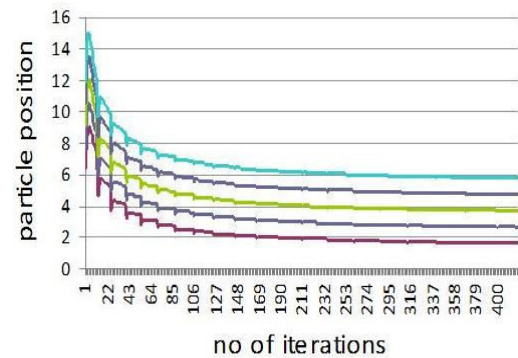


Fig. 3 : The position of the particles versus the number iterations ( $t$ )

The list of parameters initialized in the HPSO and the PSO algorithms prior to execution is as in Table III and Table IV respectively.

TABLE III. LIST OF INITIALIZED PARAMETERS

Parameters	Values
Initial weight for the network	1.5
Initial network inputs ( $x_1, x_2, x_3, x_4, x_5$ )	(0.111, 0.026, 0.004, 0.0009, 0.0003)
Number of particles	5
Learning rate coefficient	0.0001
$k$ , scalar constant in the energy function	0.5
Initial parameter ( $c_1, c_2, r_1, r_2, w$ )	(0.05, 0.05, 0.35, 0.45, 0.4)
Number of particles	5

initial social influence ( $s_1, s_2, s_3, s_4, s_5$ )	(1.1, 1.05, 1.033, 1.025, 1.02)
initial personal influence ( $p_1, p_2, p_3, p_4, p_5$ )	(3, 4, 5, 6, 7)

It can be observed that slight instabilities are present at the initial iterations. As the iterations advances, the positions of all the particles stabilize and converge to a fixed value at approximately 241 iterations onwards.

## V. DISCUSSION ON FINDINGS & RESULTS

The HPSO method can be considered robust and applicable for a wide range of optimization problems (convergence capabilities and the recursive nature of the HNN component). The PSO method on the other hand, can be said to be very suitable for handling nonlinearities. Both techniques in this work stable while performing the computations. From this work it can also be seen that the inclusion of other optimization methods (as hybrids) as well as constraint handling features may improve all three approaches.

However, it is important to take note in the sense of minimization of emissions, the PSO method seems supreme to all other methods including the HPSO method due to its capabilities in handling nonlinearities. The PSO method is the most efficient in program execution as compared to the HPSO methods. The reason for this can be attributed to the difference in terms of algorithmic complexity. Due to the network's energy component in the HNN segment of the HPSO algorithm, the stability and the convergence of the computations are assured in both these methods. As for the PSO method, the oscillations dampen out with the number of iterations due to the balancing of the social and the personal interaction mechanisms for each particle. Thus, the stability of the computations and the convergence of the solutions are also assured in the PSO method (see Fig. 3).

## VI. CONCLUSIONS & RECOMMENDATIONS

In this work, the bi-objective optimization was done well by the application of the HPSO and the PSO techniques. Comparing with the obtained solutions given in design 1 and design 2 [7], it was seen that the HPSO and PSO methods indeed produced superior results. The PSO method seems to minimize the emissions (PE) very well although not as well in the optimization of the other objectives. In this work, a new local optimum was reached for the objective functions using the PSO and the HPSO method. All the methods satisfied the constraints almost on an equal level. As an enhancements to the techniques used in this work, other optimization techniques such as; genetic algorithms [26]

and simulator annealing [27] should be implemented and tested in this problem.

## NOMENCLATURE

<i>Variables</i>	<i>Definition</i>
$COST (\$/yr)$	total cost
$w, s, b$	wind, solar and battery storage
$I_i, S_{pi}, OM_{pi}$	initial cost, present worth of salvage value, present worth of operation and maintenance cost
$N_p (yr)$	lifespan of the project
$C_g$	annual cost of purchasing power from the utility grid
$A_w, A_s$	swept area of WTG and PV panels
$N_p, N_w, N_s, N_b (yr)$	Lifespan of project, WTG, PV and storage batteries
$\eta_s, \eta_w, \eta_b$	efficiency of PV, WTG, and storage batteries
$P_{g,t}(kW)$	purchased power from the utility at hour $t$
$\psi (\$/kWh)$	grid power price
$EIR$	Energy Index of Reliability
$EENS (kWhr/yr)$	Expected Energy Not Served
$E$	Total power demand per annum
$K$	ratio of purchased power with respect to the hourly insufficient power
$PE$	pollutant emission
$\Omega, \varphi, \Gamma$	coefficients approximating the generator emission characteristics
$P_{bcap} (kW)$	capacity of storage batteries
$P_{bsoc} (kW)$	state of charge of storage batteries
$P_{bmax} (kW)$	maximum conversion capacity
$P_{bmin} (kW)$	minimum permissible storage level
$P_{bcapmax} (kW)$	allowed storage capacity
$P_{br} (kW)$	rated battery capacity
$P_b(t) (kW)$	discharge power from the storage batteries

$P_{gmax}(kW)$	maximum annual power allowed to be purchased from the utility grid
$P_{gmin}(kW)$	minimum annual power allowed to be bought from the utility grid
$T (hr)$	period under observation, 8760 hr (per year)
$P_{bsup}(t) (kW)$	surplus power at hour $t$
$P_d(t) (kW)$	load demand during hour $t$
$P_{total}(t) (kW)$	total power from WTG, PV and FFG
$P_g(kW)$	power from the FFG
$P_w(kW)$	power from the WTG
$P_s(kW)$	power from the PV
$R$	ratio of maximum permissible unmet power
$V, V_{ci}, V_r, V_{co} (m/s)$	wind speed, cut-in wind speed, rated wind speed, cut-off wind speed
$P_r(kW)$	rated WTG power
$A_{wmax}, A_{wmin} (m^2)$	maximum and minimum swept area of WTGs
$A_{smin}, A_{smax} (m^2)$	minimum and maximum swept area of PVs

## REFERENCES

- [1] CIGRE technical brochure on modeling new forms of generation and storage, CIGRE, 2000.
- [2] N. Acharya, P. Mahat, N. Mithulanathan, "An analytical approach for DG allocation in primary distribution network," *International Journal of Electrical Power and Energy Systems*, vol. 28, pp. 669–678, 2006.
- [3] Impact of increasing contribution of dispersed generation on the power systems, CIGRE, Working Group 37.23, 1999.
- [4] D. Robb, "Standing up to transmission reliability standards," *Power Engineering International*, vol. 12, pp. 20–22, Feb 2004.
- [5] M. S. Kandil, S.A. Farghal, and A.E. EL-Alfy, 'Optimum operation of an autonomous energy system suitable for new communities in developing countries', *Electric Power Systems Research*, Vol. 21, pp.137–146, 1991.
- [6] W. Kellogg, M.H. Nehrir, G. Venkataramanan, and V. Gerez, 'Optimal unit sizing for a hybrid wind/photovoltaic generating system', *Electric Power Systems Research*, Vol. 39, pp. 35–38, 1996.
- [7] W. Kellogg, M.H. Nehrir, G. Venkataramanan, and V. Gerez, "Generation unit sizing and cost analysis for stand-alone wind, photovoltaic, and hybrid wind/PV systems", *IEEE Transactions on Energy Conversion*, Vol. 13, No. 1, pp. 70–75, 1998.
- [8] H. Lund, "Large-scale integration of wind power into different energy systems, *Energy*", Vol. 30, pp. 2402–2412, 2005.
- [9] G. Notton, C. Cristofari, P. Poggi, and M. Muselli, "Wind hybrid electrical supply system: Behaviour simulation and sizing optimization", *Wind Energy*, Vol. 4, pp. 43–59, 2001.
- [10] L. Wang, and C. Singh, "PSO-based multi-criteria optimum design of a grid-connected hybrid power system with multiple renewable sources of energy". *Proceedings of the IEEE Swarm Intelligence Symposium*, 2007.
- [11] S. Dehghan, B. Kiani, A. Kazemi, A. Parizad, "Optimal Sizing of a Hybrid Wind/PV Plant Considering Reliability Indices", *World Academy of Science, Engineering and Technology*, Vol. 56, pp 5267-535, 2009 .
- [12] R. Chedid and S. Rahman, "Unit sizing and control of hybrid wind-solar power systems", *IEEE Transactions on Energy Conversion*, Vol.12, No. 1, March, pp. 79–85, 1997.
- [13] J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of National Academy of Sciences*, Vol. 29, USA, pp.2554-2558, 1982
- [14] D.W. Tank, and J.J. Hopfield, "Simple 'Neural' Optimization Network: an A/D Converter, Signal Decision Circuit and a Linear Programming Circuit", *IEEE Trans. Circuits and Systems*, CAS-33, pp. 533 -541. 1986.
- [15] Guillermo Jimenez de la C., Jose A. Ruz-Hernandez, Evgen Shelomov, (2009), 'Optimization of an Oil Production System using Neural Networks and Genetic Algorithms', *IFSA-EUSFLAT*, pp 1815-1820.
- [16] Guillermo Jimenez de la C., Jose A. Ruz-Hernandez, Evgen Shelomov, (2008), 'Obtaining an Optimal Gas Injection Rate for an Oil Production System via Neural Networks', 2008

- International Joint Conference on Neural Networks (IJCNN 2008), pp. 3359-3365.
- [17] Keerati Chayakulkheeree and Weerakorn Ongsakul, 'Optimal Power Flow Considering Non-Linear Fuzzy Network and Generator Ramp-rate Constrained', *International Energy Journal*, Vol. 8, pp 131-138, 2007.
- [18] J. Kennedy, and R. Eberhart, "Particle Swarm Optimization", *IEEE Proceedings of the International Conference on Neural Networks: Perth, Australia*, pp. 1942-1948, 1995
- [19] Assareh, E., Behrang, M.A. , Assari, M.R., Ghanbarzadeh, A., 'Application of PSO (particle swarm optimization) and GA (genetic algorithm) techniques on demand estimation of oil in Iran', *Energy*, Vol. 35, pp. 5223-5229, 2010.
- [20] Amjadi M.H., Nezamabadi-pour, H., Farsangi M.M., ' Estimation of electricity demand of Iran using two heuristic algorithms', *Energy Conversion and Management*, Vol.51, pp.493-497, 2010.
- [21] Behranga, M.A., Ghanbarzadehb, A., Assareha E., 'Comparison of the Bees Algorithm (BA) and Particle Swarm Optimisation (PSO) abilities on natural gas demand forecasting in Iran's residential-commercial sector' *Proceedings of Conference of Innovation Production Machines and Systems*, 2009.
- [22] N. Phuangpornpitak, W. Prommee, S.Tia and W. Phuangpornpitak, "A Study of Particle Swarm Technique for Renewable Energy Power Systems", *PEA-AIT International Conference on Energy and Sustainable Development: Issues and Strategies*, Thailand, pp.1-7, 2010.
- [23] Frans van den Bergh, "An Analysis of Particle Swarm Optimizers". PhD thesis, University of Pretoria, 2001.
- [24] J.P. Barton, and D.G. Infield, "Energy storage and its use with intermittent renewable energy", *IEEE Transactions on Energy Conversion*, Vol. 19, No. 2, pp. 441-448, 2004.
- [25] J.H. Talaq, F. El-Hawary, and M.E. El-Hawary, "A summary of environmental/economic dispatch Algorithms", *IEEE Trans. Power Syst.* Vol. 9, pp.1508-1516, 1994.
- [26] Deb, K., 'Optimization for Engineering Design: Algorithms & Examples', Prenticehall of India Private Limited, New Delhi, 1995.
- [27] Bennage, W. A., Dhingra, A. K., "Single and multiobjective structural optimization in discrete-continuous variables using simulated annealing" *International Journal of Numerical Methods in Engineering*, Vol.38, pp 2753-73, 1995.



# Simulation And Statistical Analysis of Network Delays For Controller Area Network

Joseph Mainoo & R. Kolla

Electronics and Computer Technology, Bowling Green State University, Bowling Green, OH 43403, USA

---

**Abstract** - Fieldbus networks such as Controller Area Network (CAN), Foundation Fieldbus and Profibus, are widely used in recent control system implementations. It is important to study the performance of these networks via simulation before implementing them in hardware. This paper uses a software simulation tool called CANoe for the development and performance analysis of CAN fieldbus network. It also provides a general overview of CAN protocol. Statistical analysis using SPSS software is performed on the network delays data for an automobile system simulation. The analysis of the collected data shows that various CAN parameters have an effect on network delays.

**Key words** - Controller Area Network (CAN); CANoe; Analysis of Delays; Simulation of CAN; Automobile System.

---

## I. INTRODUCTION

Fieldbus networks, such as Controller Area Network (CAN), Profibus, Foundation Fieldbus, are all-digital, two-way, multi-drop communication systems that are used to connect field devices such as sensors, actuators, and controllers in control systems [1]. The increasing popularity of fieldbus networks in automobile systems, factory automation, and power grid can be attributed to a host of advantages. These include: greater system functionality, simplicity, accuracy, less cost of purchase and expansion, interoperability, and other savings [1]. CAN-based networks are more accepted in automotive applications [2]. Recently many simulation tools are being used to study the performance of these fieldbus-based systems before implementing them in hardware [3-8]. One of the important performance measure considered is the network induced delays as they have significant impact on the systems [9, 10].

A Foundation Fieldbus simulation platform and its remote access interface using LabVIEW software was designed by Mossin et al. [3]. Özçelik et al. [4] designed a model of a hybrid Profibus network to obtain the performance of the Profibus medium access protocol using CACI Network II.5 software based on discrete event simulation. Rasheed [5] provided a simulation framework for Wireless Mesh Networks using OMNeT++ software. Pescaru and Dobrescu [6] used OPNET simulation tool to study various factors such as ring, mesh, star and bus topologies that affect the design of fieldbus networks. Bayilmis et al. [7] also used OPNET to model and simulate CAN. Chen et al. [8] simulated transmission of data flows in a fieldbus used for a ship power station by designing a set of CAN bus

models by OPNET simulator. Though the literature indicated use of general purpose simulation tools such as OPNET, OMNeT++, Network II, and LabVIEW, there exist other simulation tools that are designed specifically to address a particular type of fieldbus such as CAN due to its importance in the automobile field. These simulation tools help to analyze network delays and other performance measures.

Wang et al. [11] provided a model for computing the maximum and expected delays for CAN. Network induced time delay was stochastically modeled by Morales-Menendez et al. [12] for CAN. A method to calculate CAN message response times was given by Tindell et al. [13]. A probabilistic approach to determine response time distribution for messages in CAN was given by Kumar et al [14]. Schedulability analysis for CAN was discussed by Davis et al. [15]. Li et al. [16] investigated the delays associated with the use of Profibus-PA networks within control loops. The existing delay analyses used analytical and stochastic methods to establish relationships for delays. This paper uses statistical analysis methods to study the effect of various CAN parameters on network delays.

This paper employs a simulation tool called CANoe developed by Vector CANTech to generate data for the statistical analysis of network delays in a CAN based automobile system. CANoe simulation software was also used by other researchers [17, 18]. Section II gives a general overview of CAN fieldbus. Section III outlines CANoe software and describes its use in an automobile system simulation. Different types of network delays and factors affecting these delays are explained in Section IV. Results from statistical analysis of the



network delays data simulated for different cases are presented in Section V. Section VI offers concluding remarks.

## II. CONTROLLER AREA NETWORK

CAN was developed by Robert Bosch GmbH for automotive applications [19]. The CAN specification initially became an ISO 11898 standard in 1993, and extended in 1995 to permit longer device identifiers as CAN 2.0B. The CAN protocol is based on the Open Systems Interconnection/International Organization for Standardization (OSI/ISO) 7-layer reference model [1] of digital communication, shown in Figure 1. CAN implements only Physical and Data Link layers as shown in Figure 1.

No. of Layer	ISO/OSI Reference Model	CAN Protocol Specification
7	<i>Application</i>	Application Specific
6	<i>Presentation</i>	Optimal: Higher Layer Protocol (HLP)
5	<i>Session</i>	
4	<i>Transport</i>	
3	<i>Network</i>	CAN protocol (with free choice of medium)
2	<i>Data Link</i>	
1	<i>Physical</i>	

Fig. 1 : ISO/OSI reference model and CAN protocol.

CAN utilizes twisted pair cable wires as physical medium to interconnect network nodes. It is a serial, multi-master, multi-cast protocol. CAN messages are assigned static priorities, and a transmitting node will remain a transmitter until the bus becomes idle or it is superseded by a node with a higher priority message through arbitration. Bit rates up to 1 Mbps are possible in short networks that are below 40 m length. Also, longer network distances reduce the available bit rate to 125 kbps at 500 m.

There are four categories of CAN frames. They are Data Frame, Remote Frame, Error Frame, and Overload Frame. The CAN data frame is shown in Figures 2 [19]. It consists of seven fields: start of frame (SOF), arbitration, control, data, cyclical redundancy checks (CRC), acknowledge (ACK), and end of frame (EOF). CAN message bits are referred to as “dominant” (0) or “recessive” (1). CAN 2.0B support both 11 bit (standard) and 29 bit (extended) identifiers. A typical CAN message may contain up to 8 bytes of data. A message identifier describes the data content and is used by the receiving nodes to determine the destination on the network. As a result of its popularity, some higher layer protocols have been developed on top of CAN in recent years. These include CAN Kingdom, CANopen,

DeviceNet, Local Interconnect Network (LIN), Media Oriented Systems Transport (MOST), FlexRay, and J1587 [2]. Also tools to simulate CAN systems have been developed. In the next section one such simulation tool is described.

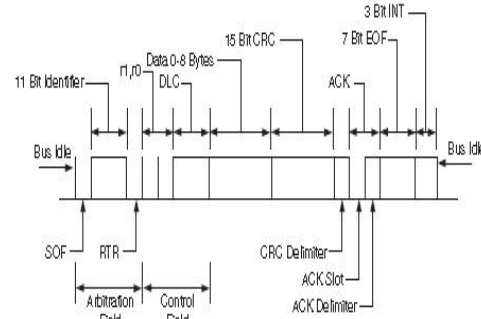


Fig. 2 : CAN data frame.

## III. CANOE SIMULATION SOFTWARE

CAN open environment (CANoe) software provides a universal development, testing, and analysis environment for CAN bus systems [20]. It was created by Vector Informatik GmbH, and permits simultaneous network development of systems in the CAN, LIN, MOST, FlexRay, J1587 and many other CAN-based protocols. It was designed primarily to model both electronic control unit (ECU) nodes on a network as well as the network that integrates them. CANoe provides two major windows. These are the “measurement setup” window, shown in Figure 3, used to configure data and statistics that are needed to be monitored, and “simulation setup” window used to configure and run the entire simulation.

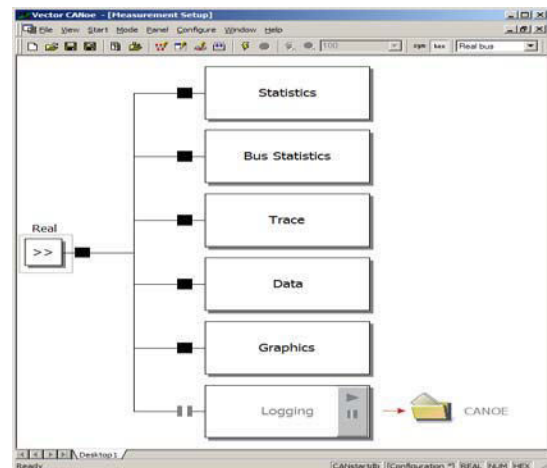


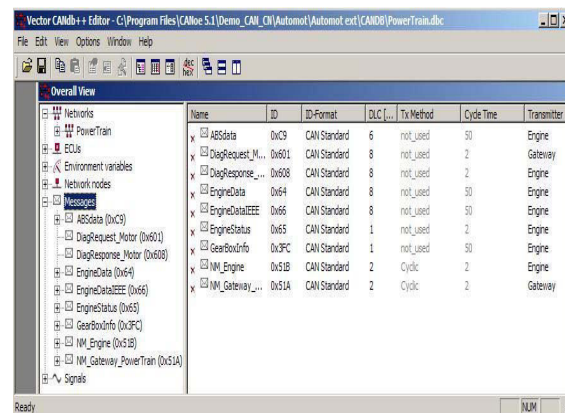
Fig. 3 : Measurement setup in CANoe.

A typical CANoe system model can be created in the “simulation setup” using the following steps [20]:

1. Create the database with messages, signals and environment variables. CANdb++ editor is used to create these databases.
2. Create the network node periphery, which includes control panels. Panel Editor is used to create these panels.
3. Create the network node model in C-like language called CAPL.

The database provides the environmental variables for describing the I/O interface among the nodes and their peripheries. Each periphery element is then “wired” to an environmental variable. It is also connected to the CAPL program for the network node. CANoe differentiates between discrete and continuous variables. Switches positions can be represented as discrete variables. Using continuous variables, dimensions such as temperature, and engine RPM, can be represented. The control panels provide a user-friendly interface to the environment variables. The user can create the panels separately with the help of the Panel Editor. During the simulation run, values of environment variables can be displayed and interactively modified.

The above three-step process of CANoe is followed in an automobile system simulation [20]. The partial list of messages in the database listed with their Name, ID, DLC, etc. is shown in Figure 4. These messages include ABSdata with identifier c9 and length 6 bytes, EngineData with identifier 64 and length 8 bytes, and GearBoxInfo with identifier 3fc and length 1 byte. The EngineData message consists of signals such as EngineSpeed with 16 bits length, EngineTemp with 7 bits length, IdleRunning with 1bit length, PetrolLevel with 8 bits length, EngForce with 16 bit length, and EngPower with 16 bit length. Similar signals are linked to the other messages c9 and 3fc. The panels of the automobile dashboard and various other consoles are shown in Figure 5 and part of the network nodes in the simulation setup are shown in Figure 6. This automobile simulator is run with several conditions to generate the data required for the performance analysis of the system. The results of the analysis are presented in Section V.



Name	ID	ID-Format	DLC	Tx Method	Cycle Time	Transmitter
ABSdata	0xc9	CAN Standard	6	not_used	50	Engine
DiagRequest_M...	0x601	CAN Standard	8	not_used	2	Gateway
DiagResponse_...	0x608	CAN Standard	8	not_used	2	Engine
EngineData	0x64	CAN Standard	8	not_used	50	Engine
EngineDataEEE...	0x66	CAN Standard	8	not_used	50	Engine
EngineStatus	0x65	CAN Standard	1	not_used	2	Engine
GearBoxInfo	0x3fc	CAN Standard	1	not_used	50	Engine
NM_Engine	0x51B	CAN Standard	2	Cyclic	2	Engine
NM_Gateway_...	0x51A	CAN Standard	2	Cyclic	2	Gateway

Fig. 4 : Partial database of automobile simulation network.



Fig. 5 : Automobile simulation panel.

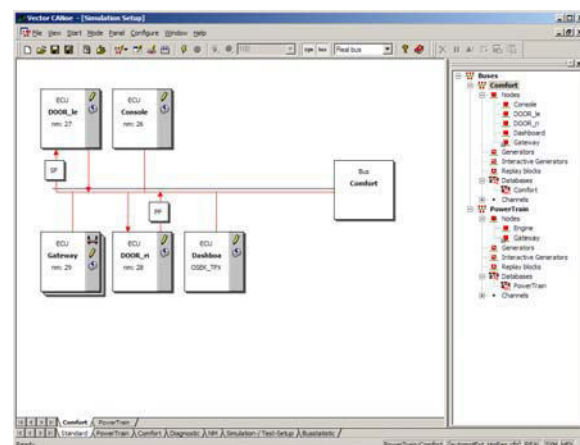


Fig. 6 : Automobile simulation network nodes.

#### IV. NETWORK DELAYS

The use of fieldbus networks such as CAN result in several advantages which include reduced system wiring and interoperability. The introduction of these networks, however, introduces different time delays in communicating among devices. Zhang [9] indicated that the existence of time delays degrades control performance in network control systems. Depending on the medium access control (MAC) protocol of the control network, these delay can be constant, time varying, or random. MAC protocols generally fall into either random access or scheduling [9]. Carrier sense multiple access (CSMA) is used in random access networks. Scheduling networks use Token passing (TP) and time division multiple access (TDMA). The control networks such as DeviceNet based on CAN and Ethernet use CSMA protocols.

Figure 7 shows two nodes that are repeatedly transmitting messages with respect to a fixed time line for random access networks during different types of conditions [9]. A node on a CSMA network monitors the network before each transmission. As shown in Type 1 of Figure 7, a node begins transmission immediately when the network is idle; otherwise it waits until the network is not busy [9]. A collision occurs when two or more nodes try to transmit concurrently. The approach to resolve the collision depends on the protocol used. CAN utilizes CSMA with a bitwise arbitration (CSMA/BA) protocol. Since CAN messages are prioritized, the message with the highest priority is transmitted without interruption when a collision occurs, and transmission of the lower priority message is terminated and will be retried when the network is idle as shown in Type 2 of Figure 7. Ethernet also utilizes a CSMA with collision detection (CSMA/CD) protocol. All the affected nodes will retreat when there is a collision, wait a random time and retransmit as shown in Type 3 of Figure 7. CSMA networks are considered nondeterministic; but higher priority messages have a better chance of timely transmission when messages are prioritized as in CAN [9].

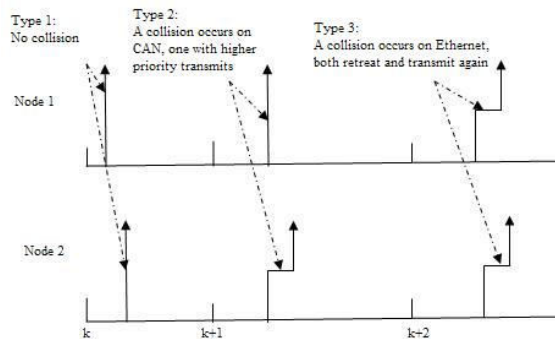


Fig. 7 : Timing diagram for two nodes on a network.

#### V. STATISTICAL ANALYSIS OF NETWORK DELAYS

In this section the effect of CAN parameters such as baud-rate (transmission speed), bus-load (% of bus activity) and message-length (0 to 8 bytes) on the delay time is studied. This is accomplished by gathering data from simulation of automobile system discussed in Section III. This simulation allows various data sources in CANoe to place information or messages on the bus “cyclically” and by “event driven”. Each individual message has a unique identifier. Messages received by the controller gets the attribute  $R_x$  and a time stamp from the card’s clock when they are received. The driver returns the time of the transmit request assigned to the CAN microcontroller. The message to be transmitted is assigned an attribute  $T_xR_q$ . After the successful transmission, the message is returned with the actual time of transmission and the attribute  $T_x$ , so that the transmit messages can be displayed and logged in the Trace windows. The time between the message’s  $T_x$  attribute and  $T_xR_q$  attribute is the Delay [20]. This is the time that the CAN controller needs to place a message completely on the CAN bus. The time delays are observed for all the messages at different baud rates, bus loads, and message lengths.

Figure 8 shows CAN bus activity for the automobile system executed at 100 kbps. For message identifier 64 (Hex) and c9 (Hex) the data is highlighted for one sample. The delay time ( $T_x - T_xR_q$ ) for message with identifiers 64 and c9 are observed to be 0.001141s (1.141 ms) and 0.00097s (0.97 ms), respectively. Samples of delay times are collected for all the messages at different CAN parameters values such as baud rates, bus loads and message lengths during simulation runs. The data is analyzed using statistical methods [21] to study the effects of various CAN parameters on the delays.

Time	Dir	ID	Name	Dir	DLC	Data
18.280309	1	65	EngineStatus	TxRq	1	01
18.280369	2	65		Rx	1	01
18.280879	1	65	EngineStatus	Tx	1	01
18.273700	1	66	EngineDataIEEE	TxRq	8	4e 63 9f 45 00 80 3b 45
18.274819	2	66		Rx	8	4e 63 9f 45 00 80 3b 45
18.274829	1	66	EngineDataIEEE	Tx	8	4e 63 9f 45 00 80 3b 45
18.278465	1	41b	NH_DOORLeft	TxRq	4	1d 12 01 ff
18.279269	2	41b		Rx	4	1d 12 01 ff
18.279279	1	41b	NH_DOORLeft	Tx	4	1d 12 01 ff
18.279309	1	c9	ABSdata	TxRq	6	10 01 00 00 49 27
18.280269	2	c9		Rx	6	10 01 00 00 49 27
18.280279	1	c9	ABSdata	Tx	6	10 01 00 00 49 27
18.275459	1	64	EngineData	TxRq	8	ec 13 3f 24 b8 0b 14 37
18.276599	2	64		Rx	8	ec 13 3f 24 b8 0b 14 37
18.276599	1	64	EngineData	Tx	8	ec 13 3f 24 b8 0b 14 37
18.270957	1	1a0	Console_1	TxRq	4	41 87 35 01
18.271749	2	1a0		Rx	4	41 87 35 01
18.271759	1	1a0	Console_1	Tx	4	41 87 35 01

Fig. 8 : Trace window showing bus activities executed at 100 kbps.

### A. Effect of Baud Rate on Network Delays

In this study 208 samples of delays for engine data signal with identifier 64 are collected at three different baud rates (50 kbps, 100 kbps, 200kbps) and analyzed using statistical package SPSS [21]. The histogram for 100 kbps is shown in Figure 9 with delays given in milliseconds (ms), and similar data is observed for the other two baud rates. The descriptive statistics obtained for this data is shown in Table I. From this table it is clear that the mean values of the delays decrease with the increase in baud rate. The one-way Analysis of Variance (ANOVA) [21] gives the results shown in Table II. The variability of mean delays between 50 kbps, 100 kbps, and 200 kbps baud rate groups, and variability of sample delays within each of these groups are shown in that table. Based on this information ( $F(2, 621) = 9027.746$ ,  $p = 0.000$ ), there is a statistically significant difference between the mean delay time for the different baud rates [21].

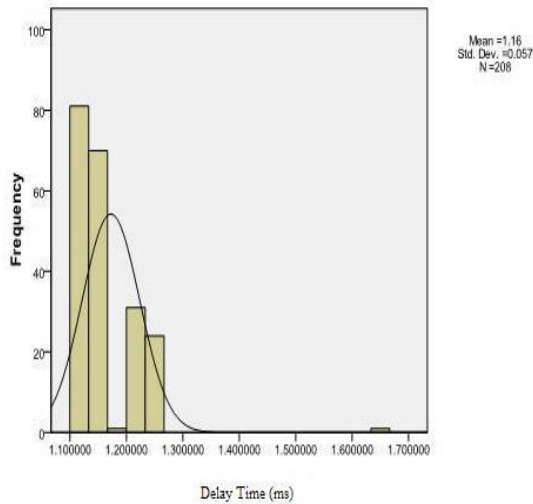


Fig. 9 : Histogram for 100 kbps baud rate data.

TABLE I. DESCRIPTIVE STATISTICS OF DELAY TIME FOR BAUD RATES.

	N	Mean	Std. Deviation
BaudRate_50	208	2.30536	0.02689
BaudRate_100	208	1.16284	0.05714
BaudRate_200	208	0.76261	0.20078

TABLE II. ANOVA ANALYSIS OF BAUDRATE DELAYS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	266.629	2	133.315	9027.746	.000
Within Groups	9.170	621	.015		
Total	275.800	623			

### B. Effect of Message Length on Network Delays

For this study also 208 samples of delays for three different signals that have different message lengths are collected. The delays are in milliseconds (ms). The signals are engine data with identifier 64 (Hex) which has a message length of 8 bytes, ABS data with identifier c9 (Hex) which has a message length of 6 bytes, and gearbox info with identifier 3fc (Hex) which has a message length of 1 byte. Baud rate selected for this study is 200 kbps. The data is analyzed using statistical package SPSS [21]. The histogram for gearbox info (c9) signal that has the message length of 6 bytes is shown in Figure 10, and similar data is observed for the other two message length signals. The descriptive statistics obtained for this data is shown in Table III. From this table it is clear that the mean values of the delays increase with the increase in message length. The one-way ANOVA analysis gives the results shown in Table IV. The variability of mean delays between 8 bytes, 6 bytes, and 1 byte data length groups, and variability of sample delays within each of these groups are shown in that table. Based on this information ( $F(2, 621) = 71.938$ ,  $p = 0.000$ ), there is a statistically significant difference between the mean delay time for the different message lengths considered [21].

### C. Effect of Bus Load on Network Delays

For this study 120 samples of delays for the signal 3fc (Hex) at two different bus loads are collected. The delays are in milliseconds (ms). The bus loads considered are 17% and 22%. The baud rate selected for this study is 100 kbps. The data is analyzed using statistical package SPSS [21]. The histogram for 22% is shown in Figure 11, and similar data is observed for the other busload. The descriptive statistics obtained for this data is shown in Table V. From this table it is clear that the mean values of the delays increase with the increase in busload. The one-way ANOVA analysis gives the results shown in Table VI. The variability of mean delays between 22% and 17% bus load groups, and variability of sample delays within each of these groups are shown in that table. Based on this information ( $F(1, 238) = 18.483$ ,  $p = 0.000$ ), there is a



statistically significant difference between the mean delay time for the different bus loads considered [21].

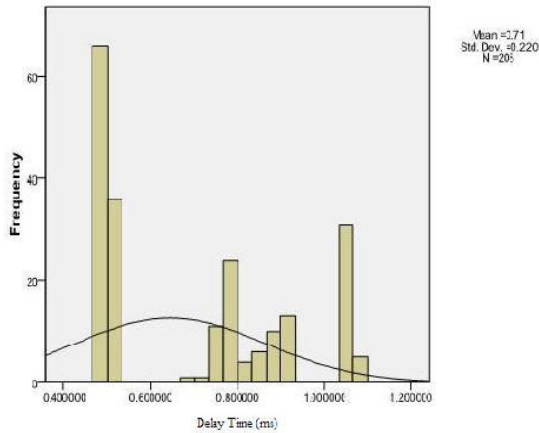


Fig. 10 : Histogram for 6 bytes message length c9 data.

TABLE III. DESCRIPTIVE STATISTICS OF DELAY FOR MESSAGE LENGTH

	N	Mean	Std. Deviation
DataLength 8_64	208	0.76261	0.20078
DataLength 6_C9	208	0.70553	0.22023
DataLength 1_3FC	208	0.52079	0.22310

TABLE IV. ANOVA ANALYSIS OF MESSAGE LENGTH DELAYS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6.647	2	3.323	71.938	.000
Within Groups	28.688	621	.046		
Total	35.334	623			

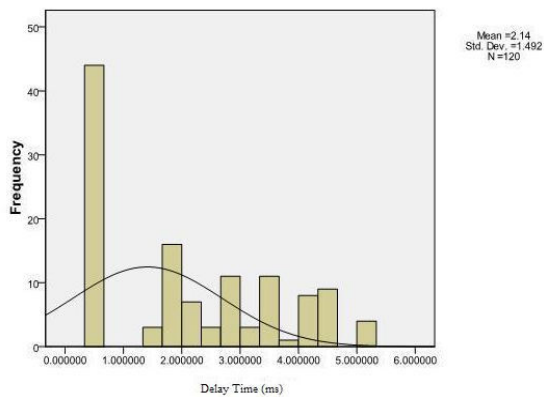


Fig. 11 : Histogram for 22% bus load data.

TABLE V. DESCRIPTIVE STATISTICS OF DELAY FOR BUS LOAD

	N	Mean	Std. Deviation
High Busload (22%)	120	2.13697500	1.49216
Low Busload (17%)	120	1.44580833	0.93544

TABLE VI. ANOVA ANALYSIS OF BUS LOAD DELAYS

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	28.663	1	28.663	18.483	.000
Within Groups	369.088	238	1.551		
Total	397.751	239			

## VI. CONCLUSION

The purpose of this paper was to present a software simulation technique and statistical analysis of network delays for CAN fieldbus. An overview of the CAN protocol and its operation were initially explained. CANoe simulation software for CAN was described in detail with an automobile system example. The data obtained from the automobile system simulation for network delays was analyzed using statistical methods with SPSS program. Based on this data it was observed that the mean delay decreases with the increase in baud rate. The mean delay also increases with the increase in message length of the transmitted signal. It was further observed that increase in busload increases the mean value of the network delay. One-way ANOVA analysis on the data collected confirmed that differences exist between these mean values of delays for different parameters. The statistical information obtained through these simulations is useful to evaluate the performance of the hardware system.

## ACKNOWLEDGMENT

The authors thank Mr. Bruce Emaus of Vector CANtech, Novi, MI, USA for providing support to Bowling Green State University through "CAN Goes to College" grant program, and for the use of CANoe material in this paper. The authors also thank Dr. David Border of Electronics and Computer Technology, BGSU for his support to this work.

## REFERENCES

- [1] J. Berge, *Fieldbuses for Process Control: Engineering, Operation, and Maintenance*, Research Triangle Park, NC: ISA Press, 2002.
- [2] U. Keskin, *In-vehicle Communication Networks: A Literature Survey*. Eindhoven: Technische Universiteit Eindhoven, VII, 2009. 1-53, Retrieved from <http://alexandria.tue.nl/repository/books/652514.pdf>.
- [3] E. A. Mossin, R. P. Pantoni, and D. Brandao, "A Fieldbus Simulator for Training Purposes," *ISA Trans.*, vol. 48, 2008, pp. 132-141.
- [4] İ. Özçelik, and H. Ekiz, "Designing and Modeling a Hybrid PROFIBUS Network using Discrete Event Simulation Technique," Retrieved from [http://ipac.kacst.edu.sa/eDoc/2005/156206\\_1.pdf](http://ipac.kacst.edu.sa/eDoc/2005/156206_1.pdf), 2005.
- [5] T. Rasheed, "Wireless Mesh Network Simulation Framework for OMNeT++." Create-Net Technical Report, CN-TR-200700016, 2007, Retrieved from [http://www.wing-project.org/\\_media/publications:cn-tr-200700016.pdf](http://www.wing-project.org/_media/publications:cn-tr-200700016.pdf).
- [6] A. Pescaru, and R. Dobrescu, "Improving Fieldbus Design with OPNET Simulation," *U.P.B.Sci.Bull., Series C*, vol. 72 (4), 2010, pp. 25-34.
- [7] C. Bayilmis, I. Erturk, C. Ceken and I. Ozcelik, "Modeling Controller Area Networks Using Discrete Event Simulation Technique," *Proc. Complex Computing Networks*, 2006, vol. 104, part II, pp. 353-358.
- [8] D. Chen, L. Xia, and H. Wang, "Modeling and Simulation of Monitor-Control Network in Ship Power Station," *Proc. IEEE Workshop on Power Electronics and Intelligent Transportation System*, August 2008, pp. 384-388.
- [9] W. Zhang, M. S. Branicky, and S. M. Phillips, "Stability of Networked Control Systems," *IEEE Control Systems Magazine*, vol. 21, 2001, pp. 84-99.
- [10] P. Antsaklis and J. Baillieul (eds.), "Special Issue on Technology of Networked Control Systems," *Proceedings of IEEE*, vol. 95, 2007, pp. 5-8.
- [11] Z.Wang, H. Lu, G. E. Hedrick and M. Stone, "Message Delay Analysis for CAN Based Networks," *Proc. ACM/SIGAPP Symp. Applied Computing (SAC 92)*, pp. 89-94, 1992.
- [12] R. Morales-Menendez, L. Garza-Castanon, R. Vargas-Rodriguez and R. Ramirez-Mendoza, "Time Delay in Controller Area Network (CAN)-Based Networked Control Systems," *NAMRI/SME Trans.*, vol. 37, 2009, pp. 221-228.
- [13] K. Tindell, A. Burns, and A. J. Wellings, "Calculating Controller Area Network (CAN) Message Response Times," *Control Engineering Practice*, vol. 3, no. 8, pp. 1163, 1995.
- [14] M. Kumar, A. K. Verma, and A. Srividya, "Response-Time Modelling of Controller Area Network (CAN)" in *Distributed Computing and Networking*, V. G. Wattenhofer, and K. Kothapalli, Eds. Berlin: Springer-Verlag, 2009, pp. 163-174.
- [15] R. I. Davis, A. Burns, R. J. Bril and J. J. Lukkien, "Controller Area Network (CAN) Schedulability Analysis: Refuted, Revisited and Revised," *Real-Time Systems*, vol. 35, issue. 3, 2007, pp. 239-257.
- [16] Q. Li, J. Jiang and D. J. Rankin, "Evaluation of Delays Induced by Profibus PA Networks," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 8, 2011, pp. 2910-2917.
- [17] C.-J. Zhou, H. Chen, Y.-Q. Qin, Y.-F. Shi and G.-C. Yu, "Self-organization of Reconfigurable Protocol Stack for Networked Control Systems," *International Journal of Automation and Computing*, vol. 8, issue. 2, 2011, pp. 221-235.
- [18] F. Zhou, S. Li, and X. Hou, "Development Method of Simulation and Test System for Vehicle Body CAN Bus Based on CANoe," *7th World Congress on Intelligent Control and Automation, WCICA 2008*, pp. 7515 – 7519.
- [19] R. Bosch, "CAN Specification Version 2.0," Retrieved from [esd.cs.ucr.edu/webres/can20.pdf](http://esd.cs.ucr.edu/webres/can20.pdf), 1991.
- [20] Vector[CANoe], *The Professional Development and Test Tool for CAN, LIN, MOST, FlexRay, Ethernet and J1708*, 2011. Retrieved from [http://www.vector.com/vi\\_canoe\\_en.html](http://www.vector.com/vi_canoe_en.html).
- [21] M. J. Norusis, *SPSS 17.0 Statistical Procedures Companion*, New Jersey: Prentice Hall, 2009.



# Analysis of Timing Jitter Effect on High Speed OFDM Systems

T. Surendra Kumara Swamy & V. Sumalatha

Department of ECE, JNTUACE, Anantapur, A.P., India

---

**Abstract** - Timing jitter is emerging as an important factor in determining the performance of high-speed Orthogonal Frequency Division Multiplexing (OFDM) systems, particularly in optical OFDM systems where bit rates reach 100Gbps and beyond. In this paper we first formulate a timing jitter matrix which is used to study the impact of timing jitter on high speed OFDM systems. This work shows that the rotational and intercarrier interference (ICI) effects caused by timing jitter greatly the system performance. Further ICI is the dominant effect. Next, we show that both fractional oversampling and integral oversampling methods can reduce the effect of timing jitter. The theoretical results compared with the simulation results. Oversampling results in a 3db reduction in jitter noise power for every doubling of the sampling rate.

**Key words** - Timing jitter, OFDM, ICI, Oversampling.

---

## I. INTRODUCTION

Today, Orthogonal Frequency Division Multiplexing (OFDM) has grown to be the most popular communication system in high-speed communications. It is used in many wireless broadband communications systems because it is a simple and scalable solution to intersymbol interference caused by a multipath channel. Very recently the use of OFDM in optical systems has attracted increasing interest [1-5]. Data rates in optical fiber systems are typically much higher than in RF wireless systems. For example transmission of 121.9 Gbit/s within an optical bandwidth of 22.8 GHz has been reported [6, 7].

At these very high data rates, timing jitter is emerging as an important limitation to the performance of OFDM systems. Transmission at high data rates requires high speed analog-to digital converters (ADCs) using an accurate sampling clock. Timing jitter is the deviation of the clock signal edge from its ideal position. A major source of jitter is the sampling clock in high speed ADCs. Timing jitter is also emerging as a problem in high frequency bandpass sampling OFDM radios [8]. Very high speed ADCs uses a parallel pipeline structure [9]. This introduces different forms of impairment, where slight errors in the delays in the different paths results in a form of timing jitter. However this 'jitter' is in the form of a repeated pattern, which can potentially be corrected within an OFDM receiver. The effect of timing jitter analyzed in [10], [11] focus on the coloured low pass timing jitter which is typical of systems using phase lock loops (PLL). They consider only integral oversampling.

In OFDM, fractional oversampling can be achieved by leaving some band-edge subcarriers unused. In this letter we investigate both fractional and integral oversampling to reduce the effect of timing jitter. Before that we introduce a novel timing jitter matrix to investigate in detail the effect of timing jitter in OFDM systems. This timing jitter matrix can be de-composed into a rotational matrix and an intercarrier interference (ICI) matrix. We extend the timing jitter matrix to analyze the detail of the ICI in an oversampled system, which we consider as the dominant effect. For Very high speed ADCs the white jitter which is the focus of this paper is a more appropriate model. The analysis of this paper is completely based on [12], [13].

The rest of the paper is organized as follows: Section II states the timing jitter problem, and derives the timing jitter matrix, rotational matrix and ICI matrix with the help of a simplified high-speed OFDM system model which includes timing jitter. Section III covers the effect of the timing jitter matrix and average ICI power. Section V covers the effect of Oversampling. The simulations are presented in Section V. Finally, Section VI gives the conclusion.

## II. SYSTEM MODEL

Consider the high-speed OFDM system shown in Fig. 1. There are  $N$  subcarriers and the OFDM symbol period is  $T$ . At the transmitter, in each symbol period,  $N$  complex values representing the constellation points are used to modulate  $N$  subcarriers. In this paper, we consider timing jitter introduced at the receiver ADC.

We now derive the mathematical model for this system. In the following  $n$  represents the time index represents the subcarrier index at the transmitter and  $l$  represent the subcarrier index at the receiver.

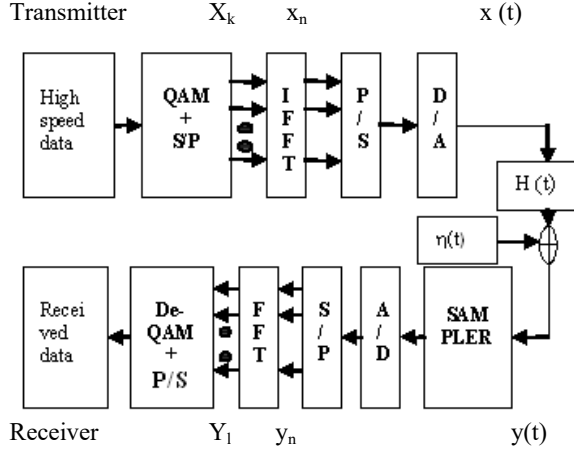


Fig. 1 : OFDM block diagram with timing jitter.

The samples at the output of the transmitter inverse fast Fourier transform (IFFT) are given by

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=-N/2+1}^{N/2} X_k e^{j\frac{2\pi}{N}kn} \rightarrow (1)$$

The analog signal  $x(t)$  at the output of the digital to analog converter (DAC), is given by

$$x(t) = \frac{1}{\sqrt{N}} \sum_{k=-N/2+1}^{N/2} X_k e^{j2\pi f_k t} \rightarrow (2)$$

Where  $f_k = k/T$ . At the receiver, the received signal  $y(t)$  is affected by the channel. So we write:

$$y(t) = \frac{1}{\sqrt{N}} \sum_{k=-N/2+1}^{N/2} H_k X_k e^{j2\pi f_k t} + \eta(t) \rightarrow (3)$$

Where  $\eta(t)$  is the additive Gaussian white noise (AWGN).

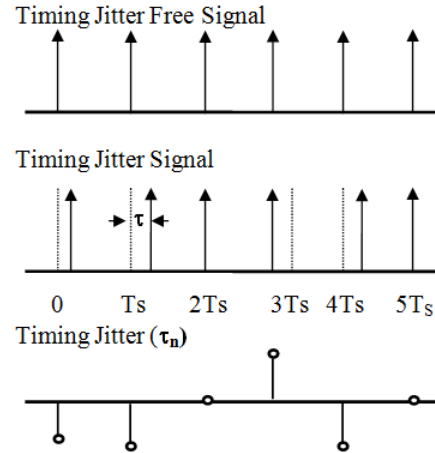


Fig.2. Definition of timing jitter.

Fig. 2 shows how timing jitter is defined. The first axis shows a signal sampled at regular timing intervals of  $T/N$ . This is the ideal case. However the effect of timing jitter is to cause deviation in the actual sampling times from the ideal sampling instants. This deviation is  $\tau$  shown in second axis. The third axis demonstrates the timing jitter  $\tau_n$  for constant time intervals.

After passing it through the ADC block, the digitalized signal is given by

$$\begin{aligned} y_n &= y\left(n\left(\frac{T}{N}\right) + \tau_n\right) + \eta_n \\ &= \frac{1}{\sqrt{N}} \sum_{k=-N/2+1}^{N/2} H_k X_k e^{j2\pi f_k \left(n\left(\frac{T}{N}\right) + \tau_n\right)} + \eta_n \rightarrow (4) \end{aligned}$$

Finally, the QAM symbol is  $Y_l$  recovered by the fast Fourier transform (FFT) block,

$$\begin{aligned} Y_l &= \frac{1}{\sqrt{N}} \sum_{k=-N/2+1}^{N/2} y_n e^{-j\frac{2\pi}{N}nl} + N_l \\ &= \frac{1}{\sqrt{N}} \sum_{n=-N/2+1}^{N/2} \sum_{k=-N/2+1}^{N/2} H_k X_k e^{j2\pi f_k \frac{\tau_n}{T}} e^{-j\frac{2\pi}{N}(k-l)n} + N_l \rightarrow (5) \end{aligned}$$

From (5) it can be seen that  $Y_l$  depends on the timing jitter  $\tau_n$  in the exponential form and the details will be explained in the following section. In order to analysis timing jitter, rewrite (5) as a compact matrix form:

$$\mathbf{Y} = \mathbf{W}\mathbf{H}\mathbf{X}^T + \mathbf{N} \rightarrow (6)$$



Where  $\mathbf{X}$ ,  $\mathbf{Y}$  and  $\mathbf{N}$  are the transmitted, received and additive white Gaussian noise (AWGN) vectors respectively,  $\mathbf{H}$  is the channel response matrix and  $\mathbf{W}$  is the timing jitter matrix

Where,

$$\mathbf{Y} = \begin{bmatrix} Y_{-N/2+1} \\ \vdots \\ Y_0 \\ \vdots \\ Y_{N/2} \end{bmatrix} \quad \mathbf{H} = \begin{bmatrix} H_{-N/2+1} \\ \vdots \\ H_0 \\ \vdots \\ H_{N/2} \end{bmatrix} \quad \mathbf{X}^T = \begin{bmatrix} X_{-N/2+1} \\ \vdots \\ X_0 \\ \vdots \\ X_{N/2} \end{bmatrix}$$

And

$$\mathbf{W} = \begin{bmatrix} w_{-N/2+1, -N/2+1} & \cdots & w_{-N/2+1, 0} & \cdots & w_{-N/2+1, N/2} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{0, -N/2+1} & \cdots & w_{0, 0} & \cdots & w_{0, N/2} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{N/2, -N/2+1} & \cdots & w_{N/2, 0} & \cdots & w_{N/2, N/2} \end{bmatrix}$$

By observing timing jitter matrix  $\mathbf{W}$ , the diagonal entries denote the factors which are used to multiply the transmitted signal in order to obtain the received signal. That is, each element of  $\mathbf{Y}$  (received signal) is equal to the product of each element of  $\mathbf{X}$  (transmitted signal) and each element of diagonal entries accordingly. These entries cause the rotation of the QAM signal which is explained in detail below. Then define  $\text{diag}(\mathbf{W})$  as timing jitter rotational matrix  $\mathbf{W}_r$

$$\mathbf{W}_r = \text{diag}(\mathbf{W})$$

$$= \begin{bmatrix} w_{-N/2+1, -N/2+1} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & w_{0, 0} & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & w_{N/2, N/2} \end{bmatrix} \rightarrow (7)$$

Where the  $l$ th elements are given by

$$w_{l,l} = \frac{1}{N} \sum_{n=-N/2+1}^{N/2} e^{j2\pi k \frac{\tau_n}{T}} (k=l) \rightarrow (8)$$

And  $l$  is the subcarrier index of the demodulate subcarrier. The other entries of  $\mathbf{W}$  introduce the interference from other  $N-1$  subcarriers onto the current subcarrier, which called intercarrier interference (ICI), and we define  $\mathbf{W}_{\text{ICI}}$  a timing jitter ICI matrix which has the following entries.

$$\mathbf{W}_{\text{ICI}} = \begin{bmatrix} 0 & \cdots & w_{-N/2+1, 0} & \cdots & w_{-N/2+1, N/2} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{0, -N/2+1} & \cdots & 0 & \cdots & w_{0, N/2} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{N/2, -N/2+1} & \cdots & w & \cdots & 0 \end{bmatrix} \rightarrow (9)$$

Where,

$$w_{l,k} = \frac{1}{N} \sum_{n=-N/2+1}^{N/2} e^{j2\pi k \frac{\tau_n}{T}} e^{j\frac{2\pi}{N}(k-l)n} (k \neq l) \rightarrow (10)$$

From the above equations (8) and (10) we can write for all the elements of  $l, k$  as

$$w_{l,k} = \frac{1}{N} \sum_{n=-N/2+1}^{N/2} e^{j2\pi k \frac{\tau_n}{T}} e^{j\frac{2\pi}{N}(k-l)n} \rightarrow (11)$$

Timing jitter causes an added noise like component in the received signal. From (6)

$$\mathbf{Y} = \mathbf{H}\mathbf{X}^T + (\mathbf{W} - \mathbf{I})\mathbf{H}\mathbf{X}^T + \mathbf{N} \rightarrow (12)$$

where  $\mathbf{I}$  is the  $N \times N$  identity matrix. The first term in (12) is the wanted component while the second term gives the jitter noise.

### III. EFFECT OF TIMING JITTER

In this section, we use the timing jitter matrix to derive the interference caused by timing jitter.

$$w_{l,k} = \frac{1}{N} \sum_{n=-N/2+1}^{N/2} e^{j2\pi k \frac{\tau_n}{T}} e^{j\frac{2\pi}{N}(k-l)n} \rightarrow (13)$$

By using the approximation  $e^{j\theta} = 1 + j\theta$  for small  $\theta$ , then

$$w_{l,k} = \frac{1}{N} \sum_{n=-N/2+1}^{N/2} \left( 1 + \frac{j2\pi k \tau_n}{T} \right) e^{j\frac{2\pi}{N}(k-l)n} \rightarrow (14)$$

In [14] it has been shown that when the timing jitter is white, the variance of the weighting coefficients is given by

$$E\{w_{l,k}^2\} = \frac{1}{N} \left( \frac{2\pi k}{T} \right)^2 E\{\tau_n^2\}, \quad k \neq l \rightarrow (15)$$

And

$$E\{w_{l,k}^2\} = 1 + \frac{1}{N} \left( \frac{2\pi k}{T} \right)^2 E\{\tau_n^2\}, \quad k = l \rightarrow (16)$$

where  $E\{\cdot\}$  denotes the expectation operator. The variance of the coefficients given in (15) characterizes the ICI and in (16) represents an interference term that has rotational effect on subcarriers. In both cases, the interference depends on the timing jitter and the transmitted subcarrier index  $k$  not on  $l$ .

From (15) it can also be seen that higher frequency subcarriers *cause* more ICI but the ICI *affects* all subcarrier equally. Combining (15) and (16), then

$$E\{w_{l,k} - I_{l,k}\} = \frac{1}{N} \left( \frac{2\pi k}{T} \right)^2 E\{\tau_n^2\} \rightarrow (17)$$

From (17) we can see that the variance of the weighting coefficients depends on the subcarrier index  $k$  and the timing jitter  $\tau_n$ .

We now calculate the average jitter noise power for each subcarrier for the case of white jitter. From (12) we have

$$Y_l = H_l X_l + \sum_{k=-N/2+1}^{N/2} (w_{l,k} - I_{l,k}) H_k X_k + N_l \rightarrow (18)$$

where the second term represents the jitter noise. In the following, we consider a flat channel with  $H_k = 1$ . Then the average jitter noise power

$$\begin{aligned} P_j(l) &= E \left\{ \left| \sum_{k=-N/2+1}^{N/2} (w_{l,k} - I_{l,k}) H_k X_k \right|^2 \right\} \\ &= \sum_{k=-N/2+1}^{N/2} E \left\{ (w_{l,k} - I_{l,k})^2 |X_k|^2 \right\} \rightarrow (19) \end{aligned}$$

Substituting (17) into (19), gives the following

$$P_j(l) = \frac{1}{N} \left( \frac{2\pi}{T} \right)^2 E\{\tau_n^2\} \sum_{k=-N/2+1}^{N/2} k^2 |X_k|^2 \rightarrow (20)$$

Assume that the transmitted signal has variance

$$\sigma_s^2 = \frac{1}{2} E|X_k|^2$$

and white timing jitter has normalized variance

$$\sigma_j^2 = \left( \frac{\text{stdev}(\tau_n)}{T/N} \right)^2$$

Then the average jitter noise power,  $P_{l(ICI)}$  to received signal power of  $l$ th subcarrier is given by

$$\frac{P_{l(ICI)}}{\sigma_s^2} = \frac{1}{3} \pi^2 (\sigma_j)^2 \rightarrow (21)$$

From (21), average ICI power for white jitter is only proportional to  $\sigma_j$  and independent of  $l$ th subcarrier index.

#### IV. EFFECT OF OVERSAMPLING

We now analyze the effect of both fractional and integral oversampling in OFDM and show that either or both can be used to reduce the degradation caused by timing jitter. To achieve integral oversampling, the *received* signal is sampled at a rate of  $MN/T$ , where  $M$  is an integer. For fractional oversampling some band-edge subcarriers are unused in the *transmitted* signal. When all  $N$  subcarriers are modulated, the bandwidth of the baseband OFDM signal is  $N/2T$ , so sampling at intervals of  $T/N$  as shown in Fig. 2 is Nyquist rate sampling. If instead, only the subcarriers with indices between  $-NL$  and  $+NU$  are non zero, the bandwidth of the signal is  $(NL + NU)/2$ . In this case sampling at intervals of  $T/N$  is above the Nyquist rate. the degree of oversampling is given by  $(NL + NU)/N$ . In the general case, where both integral and fractional oversampling are applied, the signal samples after the ADC in the receiver are given by

$$\begin{aligned} y_{n_M} &= y \left( \frac{n_M T}{NM} \right) \\ &= \frac{1}{\sqrt{N}} \sum_{k=-N_L}^{N_U} H_k X_k e^{j \frac{2\pi k}{T} \times \frac{n_M T}{NM}} + \eta \left( \frac{n_M T}{NM} \right) \rightarrow (22) \end{aligned}$$

where  $n_M$  is the oversampled discrete time index and  $\eta$  is the AWGN. With integral oversampling, the  $N$ -point FFT in the receiver is replaced by an 'oversized'  $NM$ -point FFT. The output of this FFT is a vector of length  $NM$  with elements.

$$Y_{l_M} = \frac{1}{\sqrt{M}} \frac{1}{\sqrt{NM}} \sum_{n_M=-NM/2+1}^{NM/2} y_{n_M} e^{\left(\frac{-j2\pi n_M l_M}{NM}\right)} \rightarrow (23)$$

where  $l_M$  is the index at the output of the  $NM$  point FFT. Then combining (13), (22) and (23), we obtain the modified weighting coefficients for the oversampling case,

$$w_{l_M,k} = \frac{1}{NM} \sum_{n_M=-NM/2+1}^{NM/2} e^{j2\pi k \frac{\tau_{n_M}}{T}} e^{j\frac{2\pi}{NM}(k-l_M)n_M} \rightarrow (24)$$

By using the approximation  $e^{j\theta} = 1 + j\theta$  for small  $\theta$ , then

$$w_{l_M,k} \approx \frac{1}{NM} \sum_{n_M=-NM/2+1}^{NM/2} \left(1 + \frac{j2\pi k \tau_{n_M}}{T}\right) e^{j\frac{2\pi}{NM}(k-l_M)n_M} \rightarrow (25)$$

which can be simplified to give

$$w_{l_M,k} \approx \begin{cases} \frac{1}{NM} \sum_{n_M=-NM/2+1}^{NM/2} \frac{j2\pi k \tau_{n_M}}{T} e^{j\frac{2\pi}{NM}(k-l_M)n_M} & k \neq l_M \\ 1 + \frac{1}{NM} \sum_{n_M=-NM/2+1}^{NM/2} \frac{j2\pi k \tau_{n_M}}{T} & k = l_M \end{cases} \rightarrow (26)$$

From the previous section we can write,

$$E\{|w_{l_M,k}|^2\} \approx \left(\frac{1}{NM}\right) \left(\frac{2\pi k}{T}\right)^2 E\{\tau_{n_M}^2\} \quad k \neq l_M \rightarrow (27)$$

From (27) it can be seen that white timing jitter  $E\{\tau_{n_M}^2\}$  is inversely proportional to  $M$  so increasing the integer oversampling factor reduces the intercarrier interference (ICI) due to timing jitter.

And rotational effect is given by

$$E\{|w_{l_M,k}|^2\} \approx 1 + \left(\frac{1}{NM}\right) \left(\frac{2\pi k}{T}\right)^2 E\{\tau_{n_M}^2\} \quad k = l_M \rightarrow (28)$$

We now calculate the average jitter noise power for each subcarrier for the case of white jitter. From (12), (23) and (26),

$$Y_{l_M} = H_{l_M} X_{l_M} + \sum_{k=-N_L}^{N_U} (w_{l_M,k} - I_{l_M,k}) H_k X_k + N(l) \rightarrow (29)$$

where the second term represents the jitter noise. In the following, we consider a flat channel with,  $H_k = 1$ , and assume that the transmitted signal power is distributed equally across the used subcarriers so that for each used subcarrier  $E\{X_k^2\} = \sigma_s^2$ .

## V. SIMULATION RESULTS

We now present simulation results for 2000 OFDM symbols,  $N = 512$  and  $\sigma_j$  varies from 0.03 UI (unit interval) to 0.3 UI. So in Fig. 1 we plot the average ICI power of all subcarriers Fig. 3 shows the variance of the noise due to jitter as a function of received subcarrier index when band-edge subcarriers are unused. It shows that the power of the jitter noise is not a function of subcarrier index and that removing the band-edge subcarriers reduces the noise equally across all subcarriers. Fig. 2 shows both the theoretical and simulation results for average jitter noise power as a function of the oversampling factor. There is close agreement between theory and simulation. Increasing the sampling factor gives a reduction of  $10\log_{10}(N_U/NM)$  in jitter noise power, so every doubling of the sampling rate reduces the jitter noise power by 3 dB.

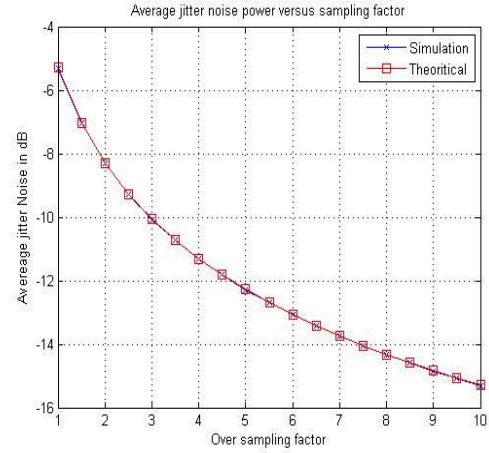


Fig. 1 : Average ICI power P versus  $\sigma_j$

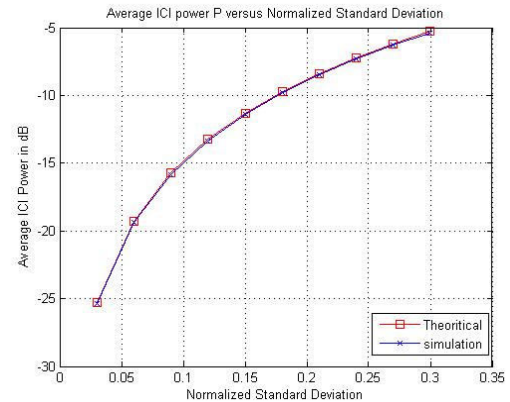


Fig. 2 : Average jitter noise power versus sampling factor.

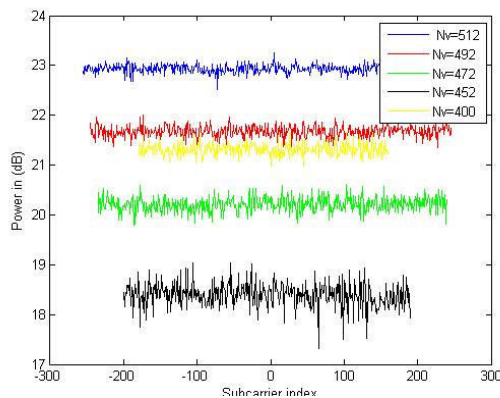


Fig. 3 : Average jitter noise power versus subcarrier index when leaving band-edge subcarrier unused.

## VI. CONCLUSION

This paper investigates the impact of timing jitter on high-speed OFDM systems using a novel timing jitter matrix and simulation. The timing jitter matrix highlights that ICI is the dominant effect. It is shown that for the case of white timing jitter, that the ICI in a received subcarrier is independent of the subcarrier index, but that higher index subcarriers cause more ICI. This has implications for the design of very high speed OFDM systems. Then it has been shown both theoretically and by simulation that oversampling can reduce the degradation caused by timing jitter in OFDM systems. Two methods of oversampling were used: fractional oversampling achieved by leaving some of the band-edge subcarriers unused, and integral oversampling implemented by increasing the sampling rate at the receiver. For the case of white timing jitter both techniques result in a linear reduction in jitter noise power as a function of oversampling rate. Thus oversampling gives a 3 dB reduction in jitter noise power for every doubling of sampling rate.

## REFERENCES

- [1] O. Gonzalez, R. Perez-Jimenez, S. Rodriguez, J. Rabadan, and A. Ayala, "Adaptive OFDM system for communications over the indoor wireless optical channel," *IEE Proceedings-Optoelectronics*, vol. 153, pp. 139-44, 2006.
- [2] J. Armstrong and A. J. Lowery, "Power efficient optical OFDM," *Electronics Letters*, vol. 42, pp. 370-2, 2006.
- [3] A. J. Lowery and J. Armstrong, "10 Gbit/s multimode fiber link using power-efficient orthogonal-frequency-division multiplexing," *Optics Express*, vol. 13, 2005.
- [4] W. Shieh and C. Athaudage, "Coherent optical orthogonal frequency division multiplexing," *Electronics Letters*, vol. 42, pp. 587-9, 2006.
- [5] J. Armstrong and B. J. C. Schmidt, "Comparison of asymmetrically clipped optical OFDM and DC-biased optical OFDM in AWGN," *IEEE Communications Letters*, vol. 12, pp. 343-5, 2008.
- [6] S. L. Jansen, I. Morita, and H. Tanaka, "10x121.9-Gb/s PDM-OFDM Transmission with 2-b/s/Hz Spectral Efficiency over 1,000 km of SSMF," *National Fiber Optic Engineers Conference*, 2008, p. PDP2.
- [7] S. L. Jansen, I. Morita, T. C. W. Schenk, N. Takeda, and H. Tanaka, "Coherent optical 25.8-Gb/s OFDM transmission over 4160-km SSMF," *Journal of Lightwave Technology*, vol. 26, pp. 6-15, 2008.
- [8] V. Syrjala and M. Valkama, "Jitter mitigation in high-frequency bandpass sampling OFDM radios," in *Proc. WCNC 2009*, pp. 1-6.
- [9] L. Sumanen, M. Waltari, and K. A. I. Halonen, "A 10-bit 200-MS/s CMOS parallel pipeline A/D converter," *IEEE Journal of Solid-State Circuits*, Stockholm, Sweden, 2001, pp. 1048-55.
- [10] K. N. Manoj and G. Thiagarajan, "The effect of sampling jitter in OFDM systems," in *Proc. IEEE Int. Conf. Commun.*, vol. 3, pp. 2061-2065, May 2003.
- [11] U. Onunkwo, Y. Li, and A. Swami, "Effect of timing jitter on OFDM based UWB systems," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 787-793, 2006.
- [12] L. Yang, P. Fitzpatrick, and J. Armstrong, "The Effect of timing jitter on high-speed OFDM systems," in *Proc. AusCTW 2009*, pp. 12-16.
- [13] L. Yang and J. Armstrong, "Oversampling to reduce the effect of timing jitter on high speed OFDM systems," *IEEE Communications Letters*, vol. 14, pp. 196-198.



# Scaling Apriori Algorithm For Generating Frequent Item Sets

S. Sreedhar & D. Kavitha

Department of CSE, G.Pulla Reddy College of Engineering, Andhra Pradesh, India

---

**Abstract** - Generating frequent item is a key approach in association rule mining. The Data mining is the process of generating frequent itemsets that satisfy minimum support and minimum confidence. Efficient algorithms to mine frequent patterns are crucial in data mining. Since the Apriori algorithm was proposed to generate the frequent item sets, there have been several methods proposed to improve its performance. But they do not satisfy the time constraint. However, most still adopt its candidate set generation-and-test approach. In addition, many methods do not generate all frequent patterns, making them inadequate to derive association rules. The Enhanced apriori algorithm has proposed in this paper requires less time in comparison to apriori algorithm. So the time is reducing.

**Key words** - Apriori, Item set, Frequent Item set, Support count, threshold, Confidence.

---

## I. INTRODUCTION

Data Mining is a promising and flourishing frontier in database systems and new database applications. Data mining [2] is the process of finding interesting trends or patterns in large data sets to guide decision about future activities. It is the analysis of dataset to find unsuspected relationships and to summarize the data in new ways that are both understandable and useful. Progress in digital data acquisition and storage technology has resulted in growth of huge database. Data is often noisy and incomplete, it is likely that many interesting patterns will be missed and reliability of detected patterns will be low. So knowledge Discovery in databases (KDD) and Data Mining (DM) helps to extract useful information from raw data. Frequent patterns are ones that occur at least a user-given number of times (minimum support) in the dataset. They allow us to perform essential tasks such as discovering association relationships among items, correlation, sequential pattern mining, and much more.

Association rules mining, introduced by Agrawal, has been traditionally applied in databases of sales transactions (referred to as market basket data). A transaction  $T$  is a set of items and contains an item set  $I$  if  $I \subseteq T$ . If  $I$  has  $k$  members, then  $I$  is called a  $k$ -itemset. An association rule is an implication  $X \rightarrow Y$  where  $X$  and  $Y$  are itemsets with no items in common i.e.  $X \cap Y = \emptyset$ . The intuitive meaning of such a rule is that the transactions (or tuples) that contain  $X$  also contain  $Y$ . The rule  $X \rightarrow Y$  holds with confidence  $c$  if  $c\%$  of transactions that contain  $X$  also contain  $Y$ . The rule  $X \rightarrow Y$  has a support  $s$  if  $s\%$  of the transactions in the database contains  $X \cup Y$ . Given a database, the problem of mining association rules is to generate all rules that have support and confidence greater than the user-

specified minimum thresholds, min-Support and min-Confidence.

Association rules processing is usually broken down into two sub problems:

1. Finding all frequent itemsets (whose supports are greater than the min-Support), also called covering or large itemsets in the literature.
2. Generating association rules derived from the frequent itemsets.

The association rules technique has also been applied to tabular data sets. An example of an association rule in tabular data is as follows : (  $A_1 = 2$  ) and (  $A_2 = 3$  ) and (  $A_4 = 5$  )  $\rightarrow$   $A_6 = 1$  support = 10%; Confidence= 60%

Where  $A_1, \dots, A_6$  are attributes.

### 1.1 KDD Steps used in Generating Frequent Item set

The knowledge discovery [7] in databases follows certain steps that are given below:

1. Domain Knowledge
2. Examining the entire raw dataset identifies finding target dataset- the target subset of data and the attributes of interest
3. Data cleaning, data reduction, and data transformation
4. Choosing data mining task and algorithms
5. Knowledge discovery

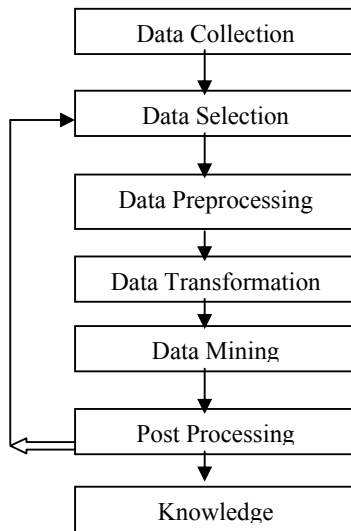


Fig. 1.1 : Steps of KDD

In applying various data mining techniques to the data source of Interest, different knowledge comes out as the mining result. This knowledge [9] is evaluated by certain rules, such as the domain knowledge or concepts. After the evaluation as shown in fig 1.1, if the result does not satisfy the requirements then we have to redo some processes until getting the correct results. The results can be projected as raw data, tables, and decision trees. The main objective of the KDD process is to make the data mining results easier to be used and more understandable.

## II. APRIORI ALGORITHM FOR GENERATING FREQUENT ITEMSETS

Different algorithms have been proposed for finding frequent item sets. The Apriori Algorithm is a well-known approach which is proposed by Agrawal & Srikant [1] (1994). It is an iterative approach and there are two steps in each iteration. The first step generates a set of candidate item sets. Then, in the second step we count the occurrence of each candidate set in database and prune all disqualified candidates (i.e. all infrequent item sets). Apriori uses two pruning technique, first on the bases of support count (should be greater than user specified support threshold) and second for an item set to be frequent, all its subset should be in last frequent item set. The iterations begin with size 2 item sets and the size is incremented after each iteration. The algorithm is based on the closure property [2] of frequent item sets: if a set of items is frequent, then all its proper subsets are also frequent.

### Apriori Algorithm

```

Initialize:  $k := 1$ ,  $C_1$  = all the 1- item sets;
read the database to count the support of  $C_1$  to determine  $L_1$ .
 $L_1 := \{\text{frequent 1- item sets}\}$ ;
 $k:=2$ ; //k represents the pass number//
while ( $L_{k-1} \neq \emptyset$ ) do
begin
 $C_k := \text{gen\_candidate\_itemsets}$  with the given  $L_{k-1}$ 
prune( $C_k$ )
for all transactions  $t \in T$  do
increment the count of all candidates in  $C_k$  that are
contained in  $t$ ;
 $L_k := \text{All candidates in } C_k \text{ with minimum support ;}$ 
 $k := k + 1$ ;
end
Answer :=  $\cup_k L_k$  ;
  
```

The first weakness of this algorithm is the generation of a large number of candidate item sets. The second problem is the number of database passes which is equal to the max length of frequent item set.

## III. ENHANCEMENT IN APRIORI ALGORITHM

The proposed algorithm shows the enhancement in apriori algorithm which reduces the time for generating the frequent item. The main problem in the apriori is that it takes the number of passes which is equal to the max length of frequent item set. In our approach time will be reduced by using the intersection method.

### Enhance Apriori algorithm

```

Initialize:  $K := 1$ ,  $C_1$  = all the 1- item sets;
read the database to count the support of
 $C_1$  to determine  $L_1$ .
 $L_1 := \{\text{frequent 1- item sets}\}$ ;
 $k:=2$ ; //k represents the pass number//
while ( $L_{k-1} \neq \emptyset$ ) do
begin
 $C_k := \text{gen\_candidate\_itemsets}$  with the given  $L_{k-1}$ 
Prune ( $C_k$ )
for all candidates in  $C_k$  do
  
```

count the number of transactions by using intersect method that are common in each item  $\in C_k$

$L_k :=$  All candidates in  $C_k$  with minimum support ;

$k := k + 1$ ;

end

Answer  $:= \cup_k L_k$  ;

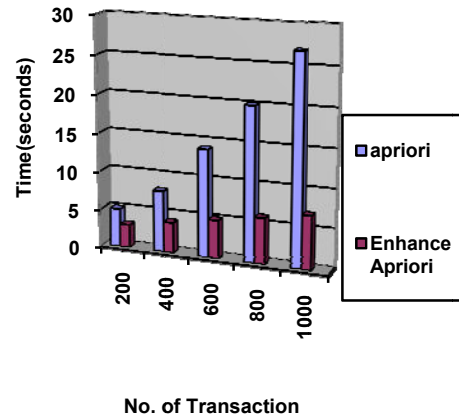
Although it is a first algorithm proposed in this field of frequent pattern mining [3] , but with the time a number of modified algorithm were designed to improve the efficiency of time, memory management and remove the complexity of process. Here we are presenting a different approach in Apriori algorithm to count the support of candidate item set. Basically this approach is more appropriate for vertical data layout, since Apriori basically works on horizontal data layout. In this new approach, we use the set theory concept of intersection. In Classical Apriori algorithm, to count the support of candidate set each record is scanned one by one and check the existence of each candidate, if candidate exists then we increase the support by one. This process takes a lot of time, requires iterative scan of whole database for each candidate set, which is equal to the max length of candidate item set. In modified approach, to calculate the support we count the common transaction that contains in each element's of candidate set, by using the intersect query of SQL. This approach requires very less time as compared to classical Apriori.

#### IV. TIME COMPARISON IN APRIORI AND ENHANCE APRIORI ALGORITHM INCREASING NUMBER OF TRANSACTION

For the comparative study of Apriori and Enhance Apriori algorithm we have taken a database of 2000 transaction of 20 items. In this analytical process we considered 400 transactions to generate the frequent pattern with the support count 15% .We have repeated the same process by increasing the transaction, after the experiment on both algorithms, we have designed a graph and summarized a result in the following table There we could see that in Apriori algorithm the time taken is directly proportion to the number of transactions where as in the Enhance apriori ( after a time period the consistency of the time was maintained) take less time. In the conclusion of the analysis we can say that for 2000 records classical Apriori take 100% time while Enhance apriori take only 25% time in comparison of classical Apriori. It means that 75% time is saving. Have designed a graph and summarized a result in the following table There we could see that in Apriori algorithm the time taken is directly proportion to the number of transactions where as in the Enhance

apriori ( after a time period the consistency of the time was maintained) take less time. In the conclusion of the analysis we can say that for 2000 records classical Apriori take 100% time while Enhance apriori take only 25% time in comparison of classical Apriori. It means that 75% time is saving.

Table 1.1



Transactions	Apriori	Enhancement in Apriori
200	5 Seconds	3 Seconds
400	9 Seconds	4 Seconds
600	14 Seconds	4 Seconds
800	18 Seconds	6 Seconds
1000	27 Seconds	7 Seconds

Figure 1.2 Time Comparison

#### V. CONCLUSION

Throughout the last decade, a lot of people have implemented and compared several algorithms that try to solve the frequent item set mining problem as efficiently as possible. For example, a very often used implementation of the Apriori algorithm is that by S. Sreedhar D. Kavitha, An Enhanced Scaling Apriori for Association Rule Mining Efficiency [13]. Nevertheless, when we compared his implementation [5] with ours, the performance of both algorithms showed immense differences. We can conclude that in this new approach, we have the key ideas of reducing time. As we have proved above how the Enhance apriori take less time than that of classical algorithms. That is really going to be fruitful in saving the time in case of large database. This key idea is surely going to open a new gateway for

the upcoming researcher to work in the filed of the data mining.

## REFERENCES

- [1] Agrawal, R., Imielinski, T., and Swami, A. N. 1993. Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, 207-216.
- [2] Agrawal, R. and Srikant, R. 1994. Fast algorithms for mining association rules. In Proc. 20th Int. Conf. Very Large Data Bases, 487-499.
- [3] Agarwal, R. Agarwal, C. and Prasad V., A tree projection algorithm for generation of frequent item sets. In J. Parallel and Distributed Computing, 2000.
- [4] J. Han, J. Pei, and Y. Yin, "Mining Frequent Patterns without Candidate Generation," Proceedings of ACM SIGMOD International Conference on Management of Data, ACM Press, Dallas, Texas, pp. 1-12, May 2000.
- [5] Vaibhav Kant Singh and Vinay Kumar Singh "Minimizing Space Time Complexity by RSTDB a new method for Frequent Pattern Mining" To be appeared in proceeding of the First International Conference on Intelligent Human Computer Interaction ,Allahabad,2009.
- [7] M. S. Chen, J. Han, and P. S. Yu. Data mining: An overview from a database Perspective. IEEE Trans. Knowledge and Data Engineering, 8:866-883, 1996.
- [8] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy. Advances in Knowledge Discovery and Data Mining. AAAI/MIT Press, 1996.
- [9] W. J. Frawley, G. Piatetsky-Shapiro and C. J. Matheus, Knowledge Discovery in Databases: An Overview. In G. Piatetsky-Shapiro et al. (eds.), Knowledge Discovery in Databases. AAAI/MIT Press, 1991.
- [10] Osmar R. Zaïane, 1999 CMPUT690 Principles of Knowledge Discovery in Databases. Han and M. Kamber. Data Mining: Concepts and Techniques. Morgan Kaufmann, 2009.
- [12] Arun K. Pujari.
- [13] S.Prakash, R.M.S.Parvathi., An Enhanced Scaling Apriori for Association Rule Mining Efficiency. European Journal of Scientific Research, ISSN 1450-216X Vol.39 No.2 (2010), pp.257-264
- [14] T. Imielinski and H. Mannila. A database perspective on knowledge discovery. Communications of ACM, 39:58-64, 1996.
- [15] G. Piatetsky-Shapiro, U. M. Fayyad, and P. Smyth. From data mining to knowledge discovery: An overview. In U.M. Fayyad, et al. (eds.), Advances in Knowledge Discovery and Data Mining, 1-35. AAAI/MIT Press, 1996.
- [16] G. Piatetsky-Shapiro and W. J. Frawley. Knowledge Discovery in Databases. AAAI/MIT Press, 1991.





# Hybrid Dynamic Routing On Distributed Systems With Security Considerations

S. Ramesh & K. Kalpana

Department of CSE. G. Pulla Reddy College of Engineering, Andhra Pradesh, India

---

**Abstract** - An ad hoc network is the cooperative engagement of a collection of mobile nodes without the required Intervention of any centralized access point or existing infrastructure. There is an increasing trend to adopt ad hoc networking for commercial uses; however, their main applications lie in military, tactical and other security-sensitive operations. In these and other applications of ad hoc networking, secure routing is an important issue. Most of the a secure routing protocols proposed in the literature are either proactive or reactive in nature. In this paper, we proposed for ad hoc network, called hybrid routing with security consideration, which is based on the concept of Zone routing protocol (ZRP). Different from the past work on the designs of cryptography algorithms and system infrastructures, we will propose a dynamic routing algorithm that could randomize delivery paths for data transmission. The algorithm is easy to implement and compatible with popular routing protocols, without introducing extra control messages

**Key words** - *Ad-hoc Routing protocols, RIP, Secure routing, DSDV.*

---

## I. INTRODUCTION

An ad hoc network is a collection of wireless computers(nodes), communicating among themselves over possibly multi-hop paths, without the help of any infrastructure are popularly such as base stations or access points [1],[2]. Unlike traditional mobile wireless networks, ad hoc networks greatly improve have no fixed infrastructure. Mobile nodes that are within each other's radio range communicate directly via wireless links, while those far apart rely on other nodes to relay messages as routers. In ad hoc network each node acts both as a host (which is capable of sending and receiving) and a router which forwards the data intended for some other node. Applications of ad hoc network range from military operations and emergency disaster relief, to commercial uses such as community networking and interaction between attendees at a meeting or students during a lecture. Most of these applications demand a secure and reliable communication. The objective of this work is to explore a security enhanced routing algorithm based on distributed routing information widely supported in existing wired and wireless networks. We aim at the randomization of delivery paths for data transmission to provide considerably small path similarity (i.e., the number of common links between two delivery paths) of two consecutive transmitted packets. These protocols shall not increase the number of control messages if the proposed algorithm is adopted.

In this paper, we proposed a secure hybrid ad hoc routing protocol, called Hybrid Routing with Security Consideration, which takes the advantage of both proactive and reactive approach. Our proposed protocol is based on zone routing protocol (ZRP) [18],[21]. The reasons for selecting ZRP as the basis of our protocol are as follows: (i) ZRP is based on the concept of routing zones, a restricted area, and it is more feasible to apply the security mechanisms within a restricted area than in a broader area that of the whole network, (ii) Since the concept of zones separate the communicating nodes in terms of interior (nodes within the zone) and exterior (nodes outside the zone) nodes, certain information like network topology and neighborhood information etc. can be hidden to the exterior nodes, (iii) In case of a failure, it can be restricted within a zone. we will use a dynamic routing algorithm to provide security enhanced data delivery without introducing any extra control messages algorithm to randomize the data delivery paths. An analytic study on the proposed algorithm is conducted. Section 4 summarizes our experimental results to demonstrate the capability of the proposed algorithm. Section 5 is the conclusion.

## II. RELATED WORK

Among many well-known designs for cryptography- based systems, the IP Security (IPSec) [12] and the Secure Socket Layer (SSL) supported and implemented in many systems and platforms. Though IPSec and SSL do the security level for data

transmission, they unavoidably introduce substantial overheads especially on gateway/host performance and effective network bandwidth. Another alternative for security-enhanced data transmission is to dynamically route packets between each source and its destination so that the chance for system break-in, due to successful interception of consecutive packets for a session, is slim. The intention of security-enhanced routing is different from the adopting of multiple paths between a source and a destination to increase the throughput of data transmission (see, e.g., [5] and [6]). In particular, Lou et al. [10], [12], [7] proposed a secure routing protocol to improve the security of end-to-end data transmission based on multiple path deliveries. The set of multiple paths between each source and its destination is determined in an online fashion, and extra control message exchanging is needed. Bohacek et al. [2] [7] proposed a secure stochastic routing mechanism to improve routing security. Similar to the work proposed by Lou et al. [10], [12], a set of paths is discovered for each source and its destination in an online fashion based on message flooding. Thus, a mass of control messages is needed. Yang and Papavassiliou [13] explored the trading of the security level and the traffic dispersion. They proposed a traffic dispersion scheme to reduce the probability of eavesdropped information along the used paths provided that the set of data delivery paths is discovered in advance. Although excellent research results have been proposed for security-enhanced dynamic routing, many of them rely on the discovery of multiple paths either in an online or offline fashion. For those online path searching approaches, the discovery of multiple paths involves a significant number of control signals over the Internet. On the other hand, the discovery of paths in an offline fashion might not be suitable to networks with a dynamic changing configuration.

### III. PROBLEM STATEMENT

The objective of this work is to explore a security-enhanced dynamic routing algorithm based on distributed routing information widely supported in existing networks. In general, routing protocols over networks could be classified roughly into two kinds: distance-vector algorithms and link state algorithms [11]. Distance-vector algorithms rely on the exchanging of distance information among neighboring nodes for the seeking of routing paths. Examples of distance-vector based routing algorithms include RIP and DSDV. Link-state algorithms used in the Open Shortest Path First protocol [19] are for global routing in which the network topology is known by all nodes. Our goal is to propose a distance-vector-based algorithm for dynamic routing to improve the security of data transmission.

Before we proceed with further discussions, our problem and system model shall be defined. A network could be modeled as a graph  $G = (N, L)$ , where  $N$  is a set of routers (also referred to as nodes) in the network, and  $L$  is a set of links that connect adjacent routers in the network. A path  $p$  from a node  $s$  (referred to as a source node) to another node  $t$  (referred to as a destination node) is a set of links  $(N_1, N_2) (N_2, N_3) \dots (N_i, N_{i+1})$ , where  $s = N_1$ ,  $N_{i+1} = t$ ,  $N_j \in N$  and  $(N_j, N_{j+1})$  for  $1 \leq j \leq i$ . Let  $P_{s,t}$  denote the set of all potential paths between a source node  $s$  and a destination node  $t$ . Note that the number of paths in  $P_{s,t}$  could be an exponential function of the number of routers in the network, and we should not derive  $P_{s,t}$  in practice for routing or analysis.

#### Definition 1 (Path Similarity) :

Given two paths  $P_i$  and  $P_j$ , the path similarity  $\text{Sim}(P_i, P_j)$  for  $p_i$  and  $p_j$  is defined as the number of common links between  $p_i$  and  $p_j$ :

$$\text{Sim}(P_i, P_j) = |\{(N_x, N_y) | (N_x, N_y) \in p_i \wedge (N_x, N_y) \in p_j\}|,$$

where  $N_x$  and  $N_y$  are two nodes in the network

The path similarity between two paths is computed based on the algorithm of Levenshtein distance [12].

#### Definition 2 (The expected value of path similarity for any two consecutive delivered packets) :

Given a source node  $s$  and a destination node  $t$ , the expected value of path similarity of any two consecutive delivered packets is defined as follows:

$$E[\text{Sims}, t] = \sum \text{Sim}(p_i, p_j) \cdot \text{Prob}(p_i | p_j) \cdot \text{Prob}(p_i) \cdot p_{j \in P_{s,t}}$$

Where  $P_{s,t}$  is the set of all possible transmission paths between a source node  $s$  and a destination node  $t$ .  $\text{Prob}(p_i | p_j)$  is the conditional probability of using  $p_j$  for delivering the current packet, given that  $p_i$  is used for the previous packet.  $\text{Prob}(p_i)$  is the probability of using  $p_i$  for delivering the previous packet.

The purpose of this research is to propose a dynamic routing algorithm to improve the security of data transmission. We define the eavesdropping avoidance problem as follows:

Given a graph for a network under discussion, a source node, and a destination node, the problem is to minimize the path similarity without introducing any extra control consecutive packets over a messages, and thus to reduce the probability of eavesdropping specific link.

#### IV. SECURITY- ENHANCED DYNAMIC ROUTING

##### 4.1 Notations and Data Structures

The objective of this section is to propose a distance-vector based algorithm for dynamic routing to improve the security of data transmission. We propose to rely on existing distance information exchanged among neighboring nodes (referred to as routers as well in this paper) for the seeking of routing paths. In many distance-vector-based implementations, e.g., those based on RIP, each node  $N_i$  maintains a routing table (see Table 1a) in which each entry is associated with a tuple  $(t, WN_{i,t}, Nexthop)$  where  $t$ ,  $WN_{i,t}$  and  $Nexthop$  denote some unique destination node, an estimated minimal cost to send a packet to  $t$ , and the next node along the minimal-cost path to the destination node, respectively.

With the objective of this work in the randomization of routing paths, the routing table shown in Table 1a is extended to accommodate our security-enhanced dynamic routing algorithm. In the extended routing table (see Table 1b), we propose to associate each entry with a tuple  $(t, WN_{i,t}, CN_i t, HN_i t)$ .  $CN_i t$  is a set of *node candidates* for the nexthop (note that the candidate selection will be elaborated in Procedure 2 of Section 3.2), where one of the nexthop candidates that have the minimal cost is marked.  $HN_i t$ , a set of tuples, records the history for packet deliveries through the node  $N_i$  to the destination node  $t$ . Each tuple  $(N_j, hN_j)$  in  $HN_i t$  is used to represent that  $N_j$  previously used the node  $hN_j$  as the nexthop to forward the packet from the source node  $N_j$  to the destination node  $t$ . Let  $Nbri$  and  $wN_i$ ,  $N_j$  denote the set of neighboring nodes for a node  $N_i$  and the cost in the delivery of a packet between  $N_i$  and a neighboring node  $N_j$ , respectively. Each node  $N_i$  also maintains an array (referred to as a link table) in which each entry corresponds to a neighboring node  $N_j$  —  $Nbri$  and contains the cost  $wN_i, N_j$  for a packet delivery.

The proposed algorithm achieves considerably small path similarity for packet deliveries between a source node and the corresponding destination node. However, the total space requirement would increase to store some extra routing information. The size of a routing table depends on the topology and the node number of a network under discussions. In the worst case, we have a fully connected network. For each entry in the routing table shown in Table 1b, the additional spaces required for recording the set of node candidates (as shown in the third column of Table 1b) and for

recording the routing history (as shown in the fourth column of Table 1b) are  $O(|N|)$ . Because there are  $|N|$  destination nodes at most in each routing table, the additionally required spaces for the entire routing table for one node are  $O(|N|^2)$ . Since the provided distributed dynamic routing algorithm (DDRA) is a distance-vector-based routing protocol for intra domain systems, the number of nodes is limited, and the network topology is hardly fully connected. Hence, the increase of the total space requirement is considerably small. However, the impact of the space requirement on the search time will be analyzed in the following section.

#### V. HYBRID ROUTING WITH SECURITY CONSIDERATION

##### 5.1 Notations and Data Structures

The objective of this section is to propose a hybrid routing algorithm to improve the security of data transmission. The hybrid Routing with security consideration Protocol is based on zone routing protocol (ZRP) [18],[21]. Like ZRP it performs intrazone [19] and interzone [20] routing; however, it differs from ZRP in security aspects. In ZRP where there is no security consideration, hybrid Routing with security consideration designed to address all measure security concerns like end to end authentication, message/packet integrity and data confidentiality during both intra and interzone routing. For end to end authentication and message integrity RSA digital signature mechanism [11] is employed, where as data confidentiality is ensured by an integrated approach of both symmetric and asymmetric key encryption [11]. Each communicating node has two pairs of private/public keys, one pair for signing and verifying and the other for encrypting and decrypting. We propose to rely on existing information exchanged among neighboring nodes (referred to as routers as well in this paper) for the seeking of routing paths. In ZRP implementation, each node  $N_i$  maintains a routing table in which each entry is associated with a tuple  $(t, WN_{i,t}, Nexthop)$ , where  $t$ ,  $WN_{i,t}$  and  $Nexthop$  denote some unique destination node, an estimated minimal cost to send a packet to  $t$ , and the next node along the minimal-cost path to the destination node, respectively. With the objective of this work in the randomization of routing paths, the routing table is extended to accommodate our hybrid routing with security consideration algorithm. In the extended routing table, we propose to associate each entry with a tuple

$(t, WN_{i,t}, CN_i t, HN_i t)$ .  $CN_i t$  is a set of node candidates for the nexthop . where one of the nexthop candidates that have the minimal cost is marked.  $HN_i t$  , a set of

tuples, records the history for packet deliveries through the node  $N_i$  to the destination node  $t$ .

Each tuple  $(N_j, hN_j)$  in  $HN_i t$  is used to represent that  $N_i$  previously used the node  $hN_j$  as the nexthop to forward the packet from the source node  $N_j$  to the destination node  $t$ . Let  $Nbr_i$  and  $\omega_{N_i, N_j}$  denote the set of neighboring nodes for a node  $N_i$  and the cost in the delivery of a packet between  $N_i$  and a neighboring node  $N_j$ , respectively. Each node  $N_i$  also maintains an array (referred to as a link table) in which each entry corresponds to a neighboring node  $N_j \in Nbr$  and contains the cost  $w_{N_i, N_j}$  for a packet delivery. The proposed algorithm achieves considerably small path similarity for packet deliveries between a source node and the corresponding destination node. However, the total space requirement would increase to store some extra routing information. The size of a routing table depends on the topology and the node number of a network under discussions. In the worst case, we have a fully connected network. For each entry in the routing table shown in Table 1b additional spaces required for recording the set of node candidates (as shown in the third column of Table 1b) are  $O(|N|)$ . Because there are  $|N|$  destination nodes at most in each routing table, the additionally required spaces for the entire routing table for one node are  $O(|N|^2)$ . Since the provided distributed dynamic routing algorithm (DDRA) is a distance-vector-based routing protocol for intradomain systems, the number of nodes is limited, and the network topology is hardly fully connected. Hence, the increase of the total space requirement is considerably small. However, the impact of the space requirement on the search time will be analyzed in the following section

TABLE 1

An Example of the Routing Table for the Node  $N_i$

Destination Node ( $t$ )	Cost ( $W_{N_i, t}$ )	Nexthop
$N_1$	7	$N_6$
$N_2$	8	$N_{21}$
$N_3$	9	$N_9$
$\vdots$	$\vdots$	$\vdots$

(a)

Destination Node ( $t$ )	Cost ( $W_{N_i, t}$ )	Nexthop Candidates ( $C_t^{N_i}$ )	History Record for Packet Deliveries to the Destination Node $t$ ( $H_t^{N_i}$ )
$N_1$	7	$\{\tilde{N}_6, N_{20}, N_{21}\}$	$\{(N_2, N_{21}), (N_3, N_6), \dots, (N_{31}, N_{20})\}$
$N_2$	8	$\{N_9, N_{21}\}$	$\{(N_1, N_9), (N_3, N_9), \dots, (N_{31}, N_{21})\}$
$N_3$	9	$\{\tilde{N}_9\}$	$\{(N_1, N_9), (N_2, N_9), \dots, (N_{31}, N_9)\}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

(b)

(a) The routing table for the original distance-vector-based routing algorithm. (b) The routing table for the proposed security-enhanced routing algorithm.

## 5.2 Distributed Dynamic Routing Algorithm

The DDRA proposed in this paper consists of two parts:

- 1) a randomization process for packet deliveries and
- 2) maintenance of the extended routing table.

### 5.2.1 Randomization Process

Consider the delivery of a packet with the destination  $t$  at a node  $N_i$ . In order to minimize the probability that packets are eavesdropped over a specific link, a randomization process for packet deliveries shown in Procedure 1 is adopted. In this process, the previous nexthop  $hs$  (defined in  $HN_i$  of Table 1b) for the source node  $s$  is identified in the first step of the process (line 1). Then, the process randomly pick up a neighboring node in  $CN_i t$  excluding  $hs$  as the nexthop for the current packet transmission. The exclusion of  $hs$  for the nexthop selection avoids transmitting two consecutive packets in the same link, and the randomized pickup prevents attackers from easily predicting routing paths for the coming transmitted packets. Procedure 1

Procedure 1: RANDOMIZEDSELECTOR ( $s, t, pkt$ )

- 1: Let  $hs$  be the used nexthop for the previous packet delivery for the source node  $s$ .
- 2: if  $hs \in CN_i t$  then
- 3: if  $|CN_i t| > 1$  then
- 4: Randomly choose a node  $x$  from  $\{CN_i t - hs\}$  as a nexthop, and send the packet  $pkt$  to the node  $x$ .
- 5:  $hs \leftarrow x$  and update the routing table of  $N_i$ .
- 6: else
- 7: Send the packet  $pkt$  to  $hs$ .
- 8: end if
- 9: else
- 10: Randomly choose a node  $y$  from  $CN_i t$  as a nexthop, and send the packet  $pkt$  to the node  $y$ .
- 11:  $hs \leftarrow y$ , and update the routing table of  $N_i$ .
- 12: end if

The number of entries in the history record for packet deliveries to destination nodes is  $|N|$  in the worst case. In order to efficiently look up the history record for a destination node, we maintain the history record for each node in a hash table. Before the current packet

is sent to its destination node, we must randomly pick up a neighboring node excluding the used node for the previous packet. Once a neighboring node is selected, by the hash table, we need  $O(1)$  to determine whether the selected neighboring node for the current packet is the same as the one used by the previous packet. Therefore, the time complexity of searching a proper neighboring node is  $O(1)$ .

### 5.2.2 Routing Table Maintenance

Let every node in the network be given a routing table and a link table. We assume that the link table of each node is constructed by an existing link discovery protocol, such as the Hello protocol in [8]. On the other hand, the construction and maintenance of routing tables are revised based on the well-known Bellman-Ford algorithm and described as follows:

Initially, the routing table of each node (e.g., the node  $N_i$ ) consists of entries  $\{N_j, W_{N_i, N_j}, C_{N_i, N_j} = \{N_j\}, H_{N_j}$

$N_i = \emptyset\}$ , where  $(N_j, N_{br})$  and  $W_{N_i, N_j} = w_{N_i, N_j}$ . By exchanging distance vectors between neighboring nodes, the routing table of  $N$  is accordingly updated. Note that the exchanging for distance vectors among neighboring nodes can be based on a predefined interval. The exchanging can also be triggered by the change of link cost or the failure of the link/node. In this paper, we consider cases when  $N_i$  receives a distance vector from a neighboring node  $N_j$ . Each element of a distance vector received from a neighboring node  $N_j$  includes a destination node  $t$  and a delivery cost  $W_{N_j, t}$  from the node  $N_j$  to the destination node  $t$ . The algorithm for the maintenance of the routing table of  $N_i$  is shown in Procedure 2, and will be described below.

Procedure 2: DVPROCESS( $t, W_{N_j, t}$ )

```

1: if the destination node  $t$  is not in the routing table
   then
2:   Add the entry( $t, (w_{N_i, N_j} + W_{N_j, t}), C_{N_i, t} = \{N_j\}, H_{N_i} = \emptyset$ )
3: else if  $(w_{N_i, N_j} + W_{N_j, t}) < W_{N_i, t}$  then
4:    $C_{N_i, t} \leftarrow \{N_j\}$  and  $N_j$  is marked as the minimal-cost nexthop.
5:    $W_{N_i, t} \leftarrow (w_{N_i, N_j} + W_{N_j, t})$ 
6:   for each node  $N_k \in N_{br}$  except  $N_j$  do
7:     if  $W_{N_k, t} < W_{N_i, t}$  then
8:        $C_{N_i, t} \leftarrow C_{N_i, t} \cup \{N_k\}$ 
9:     end if
10:  end for
11: Send ( $t, W_{N_i, t}$ ) to each neighboring node  $N_k \in N_{br}$ .

```

```

12: else if  $(w_{N_i, N_j} + W_{N_j, t}) > W_{N_i, t}$  then
13:   if  $(N_j \in C_{N_i, t})$  then
14:     if  $N_j$  was marked as the minimal-cost nexthop then
15:        $W_{N_j, t} = \min(w_{N_i, N_k} + W_{N_k, t})$ 
16:        $C_{N_i, t} \leftarrow \emptyset$ 
17:     for each node  $N_k \in N_{br}$  do
18:       if  $W_{N_k, t} < W_{N_i, t}$  then
19:          $C_{N_i, t} \leftarrow C_{N_i, t} \cup \{N_k\}$ 
20:       end if
21:     end for
22:   Send ( $t, W_{N_i, t}$ ) to each neighboring node  $N_k \in N_{br}$ 
23: else if  $W_{N_j, t} > W_{N_i, t}$  then
24:    $C_{N_i, t} \leftarrow C_{N_i, t} - \{N_j\}$ 
25: end if
26: else if  $(N_j \notin C_{N_i, t}) \wedge (W_{N_j, t} < W_{N_i, t})$  then
27:    $C_{N_i, t} \leftarrow C_{N_i, t} \cup \{N_j\}$ 
28: end if
29: end if

```

## VI. CONCLUSION

The simulation results for Hybrid Routing With Security Consideration under different mobility patterns and traffic scenarios show that the proposed protocol is as efficient as ZRP in discovering and maintaining routes. However, the impact of the overhead caused is almost insignificant and negligible as compared to the proposed degree of security, which provides compared to its other counterparts. The advantages of a multipath approach are clearly exemplified. We can argue that the multipath approach can increase confidentiality.

## REFERENCES

- [1] C. Siva Ram Murthy and B. S. Manoj, "AdHoc Wireless Networks, Architecture and Protocols", Prentice Hall PTR, 2004.
- [2] Stefano Basagni, Marco Conti, Silvia Giordano and Ivan Stojmenovic, "Mobile Ad Hoc Networks", IEEE press, A John Wiley & Sons, INC. publication, 2003
- [3] George Aggelou, "Mobile Ad Hoc Networks", 2<sup>nd</sup> edition, Mc Graw Hill professional engineering, 2004
- [4] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On Power- Law Relationships of the Internet

- Topology,” Proc. ACM SIGCOMM’99, pp. 251-262, 1999.
- [5] I. Gojmerac, T. Ziegler, F. Ricciato, and P. Reichl, “Adaptive Multipath Routing for Dynamic Traffic Engineering,” Proc. IEEE Global Telecommunications Conf. (GLOBECOM), 2003.
- [6] C. Hopps, Analysis of an Equal-Cost Multi-Path Algorithm, Request for comments (RFC 2992), Nov. 2000.
- [7] C. Kaufman, R. Perlman, and M. Speciner, Network Security—PRIVATE Communication in a PUBLIC World, second ed. Prentice Hall PTR, 2002
- [8] J.F. Kurose and K.W. Ross, Computer Networking—A Top-Down Approach Featuring the Internet. Addison Wesley, 2003.[10] V.I. Levenshtein, “Binary Codes Capable of Correcting Deletions,Insertions, and Reversals,” Soviet Physics Doklady, vol. 10, no. 8,pp. 707-710, 1966
- [9] S.-H. Liu, Y.-F. Lu, C.-F. Kuo, A.-C. Pang, and T.-W. Kuo, “ThePerformance Evaluation of a Dynamic Configuration Method over IPSEC,” Proc. 24th IEEE Real- Time Systems Symp.: Works in Progress Session (RTSS WIP), 2003.
- [10] W. Lou and Y. Fang, “A Multipath Routing Approach for Secure Data Delivery,” Proc. IEEE Military Comm. Conf. (MilCom), 2001.[12] W. Lou, W. Liu, and Y. Fang, “SPREAD: Improving Network Security by Multipath Routing,” Proc. IEEE Military Comm. Conf. (MilCom), 2003.
- [11]. M. O. Pervaiz, M. Cardei, and J. Wu, "Routing Security in Ad Hoc Wireless Networks," Network Security, S. Huang, D. MacCallum, and D. -Z. Du (eds.), Springer,2008
- [12] R. Thayer, N. Doraswamy, and R. Glenn, IP Security Document Roadmap, Request for comments (RFC 2411), Nov. 1998.
- [13] J. Yang and S. Papavassiliou, “Improving Network Security by Multipath Traffic Dispersion,” Proc. IEEE Military Comm. Conf. (MilCom), 2001.
- [14] 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 Proc. ACM MOBICOM, pages 85–97, Oct. 1998.
- [15] C. Waal and M. Gerharz, “BonnMotion: A Mobility Scenario Generation and Analysis Tool”, Communication Systems Group, Institute of Computer Science IV, University of Bonn, Germany
- [16] G. Malkin, Routing Information Protocol (RIP) Version 2 Carrying Additional Information, Request for comments (RFC 1723), Nov. 1994
- [17] K. Sadasivam, “Tutorial for Simulation-based Performance Analysis of MANET Routing Protocols in ns-2”,
- [18] Haas Z. J., Pearlman M. R., and Samar P., “The Zone Routing Protocol (ZRP)”,IETF Internet Draft, draft-ietfmanet- zone-zrp-04.txt, July 2002.
- [19] Haas, Zygmunt J., Pearlman, Marc R., Samar, P.: “Intrazone Routing Protocol (IARP)”, IETF Internet Draft,draft-ietf-manet-iarp-01.txt, June 2001
- [20] Haas, Zygmunt J., Pearlman, Marc R., Samar, P.:“Interzone Routing Protocol (IERP)”, IETF Internet Draft,draft-ietf-manet-ierp-01.txt, June 2001
- [21] Jan Schaumann, “Analysis of Zone Routing Protocol”,Course CS765, Stevens Institute of Technology Hoboken,New Jersey, USA, 8th December 2002
- [22] P. Jacquet, P. Muhlethaler, A. Qayyum, “Optimized Link State Routing Protocol”, Internet Draft, draft-ietfmanetolsr- 00.txt, November 1998



# Fault Diagnosis of Power Transformer Based on Dissolved Gas Analysis And Adaptive Neuro-Fuzzy Inference System

**Pallavi Patil & Vikal Ingle**

Bapurao Deshmukh College of Engineering, Sewagram

---

**Abstract** - Power Transformers are a vital link in a power system. Well-being of power transformer is very much important to the reliable operation of the power system. Dissolved Gas Analysis (DGA) is one for the effective tool for monitoring the condition of the transformer. To interpret the DGA result multiple techniques are available. IEC codes are developed to diagnose transformer faults. But there are cases of errors and misleading judgment due to borderline and multiple faults. Methods were developed to solve this problem by using fuzzy membership functions to map the IEC codes and heuristic experience to adjust the fuzzy rule. This paper proposes a neuro-fuzzy method to perform self learning and auto rule adjustment for producing best rules.

**Key words** - Dissolved Gas Analysis, Fault Diagnosis, Fuzzy Inference System, Gas Concentration.

---

## I. INTRODUCTION

Power transformer plays an important role in power system. Hence for the reliable operation of power supply system the health of transformers must be properly maintained. Power transformer is subject to various thermal, electrical and mechanical stresses. These stresses can cause incipient faults, deterioration or even failure of power transformer. To prevent this timely fault detection via a suitable standard technique is must. Dissolved Gas Analysis (DGA) is an important tool for online monitoring of transformers of transformers [1-3]. Different standards are available such as the IEC 599 ratio codes, Rogers's ratio and Triangle ratio [1-3]. These ratio methods are efficient and simple to use.

Fault type	Fault Description	$C_2H_2 / C_2H_4$	$CH_4 / H_2$	$C_2H_4 / C_2H_6$
HEDA_4	High Energy Discharge Arcing	2	0 or 2	X
HEDA_3		1	0 or 2	X
HEDA_2	Low Energy Discharge	2	1	X
HEDA_1		1	1	X
LED	Partial Discharge	0	1	X
Normal	Normal Aging condition	0	0	0

OH_T1	Low temp overheating	0	0	1 or 2
OH_T2	Thermal Fault T1	0	2	0
OH_T3	Thermal Fault T2	0	2	1
OH_T4	Thermal Fault T3	0	2	2

Table 2: Table of the details about the fault

There are certain limitations of the ratio methods. In case of more than one fault present, only the dominating faults are represented in these rules. Due to this some of the faults are left unrecognized. Another shortcoming is due to the structure of the IEC codes used. Known as the gas ratios these codes are quantized to define the crisp boundaries. Practically mostly these figures are fuzzy. Thus these codes could lead to errors. To find out the solution of this difficulty some of artificial neural network (ANN) methods were employed [1-4]. In these ANN based methods the key gas methods were not used which is very important for the correct fault diagnosis. Finally a fuzzy DGA method was developed which uses both IEC 60599 and the key Gas concentration.[5].

## II. IEC 60599 DGA CODES

During faults in the transformer due to electrical and thermal stresses, oil and paper decomposition occurs evolving gases that will decrease the heat dissipation capability and the dielectric strength of the

oil. These released gases get dissolved in the oil, which are known as dissolved gases.

Different fault are associated and are reflected by the different concentration of the gases in oil. IEC 50699 uses 3 gas ratios  $C_2H_2/C_2H_4$  (acetylene upon ethylene),  $CH_4/H_2$  (methane upon hydrogen) and  $C_2H_4/C_2H_6$  (ethylene upon ethane). Each ratio is quantized to a classification code 0,1 or 2. Thus there must be total 27 combinations but IEC 50699 defines only 11 combinations leading to non-decision diagnosis.

	$C_2H_2/C_2H_4$	$CH_4/H_2$	$C_2H_4/C_2H_6$
<0.1	0	1	0
0.1-0.25	1	1	0
0.25-1	1	0	0
1-3	1	2	1
>3	2	2	2

Table 1 : IEC gas ratios

The IEC codes are extended into the expert rules using experiences in the field by filling in the gaps created by IEC. The new knowledge base is given in the tables 1 and whose main advantage is elimination of the non-decision problem with all the 27 combinations included. It shows the IEC codes for different gas concentration ratios.

### III. FUZZY DIAGNOSIS SYSTEM

The fuzzy logic analysis involves three successive process namely: fuzzification, fuzzy inference and defuzzification. Fuzzification converts a crisp gas ratio into a fuzzy input membership. A chosen fuzzy inference system (FIS) is responsible for obtaining conclusions from the knowledge based fuzzy rules set of if – then linguistic statements. Defuzzification then converts the output values back into the crisp values.

#### 3.1. Input of the System

The input for our fuzzy diagnosis system are three gas ratios  $C_2H_2/C_2H_4$  (acetylene upon ethylene),  $CH_4/H_2$  (methane upon hydrogen) and  $C_2H_4/C_2H_6$  (ethylene upon ethane). The values of these ratios i.e Code 0, Code 1 or code 2 are each represented by a trapezoidal membership function. These inputs are given to a sugeno model for obtaining the output.

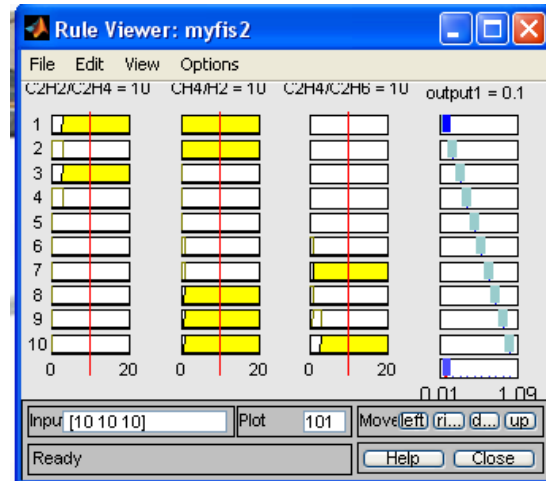


Fig. 1: Fuzzy (sugeno) analysis of the fuzzy rules.

#### 3.2. Fuzzy Rules Base

The three gas ratios constitutes the fuzzy inputs. The if-then statements is used to make decision on the fuzzy inputs derived from the three gas ratios. The fault s are described in the given table 2. For example if  $C_2H_2/C_2H_4$  is 0,  $CH_4/H_2$  is 2 &  $C_2H_4/C_2H_6$  is also 0 then the fault type corresponding to this combination of the ratios is OH\_T2 i.e thermal fault (overheating)  $150 < T < 300$ . Such ten fuzzy rules are created in the matlab environment.

#### 3.3. Fuzzy Inference System (FIS)

The fuzzy inference system can be of two types mamdani and the sugeno type. Here the fuzzy inference system used is of the sugeno type. It obtains its output from judging all the written fuzzy rules by finding the membership for the fault types as shown by the fuzzy rules. Depending upon the concentrations of the different gas elements of the DGA, the ratios for  $C_2H_2/C_2H_4$  (acetylene upon ethylene),  $CH_4/H_2$  (methane upon hydrogen) and  $C_2H_4/C_2H_6$  (ethylene upon ethane) are found out. depending on this value a code is assigned to each ratio, which can be 0, 1 or 2. The figure (1) shows the analysis of the 10 fuzzy rules graphically.

Thus fuzzy logic can be effectively used to obtain the output of the system. As the age of transformer increases there is a continuous change in the relationship between the gas ratios and the fault type. Hence the fuzzy logic fault diagnosis can cause error in the decision. Thus artificial neural networks can be engaged for self learning.

This paper proposes fault diagnosis of the transformer based on the DGA employing ANFIS (Adaptive Neuro-fuzzy inference system). It is an



small attempt to make the diagnosis intelligent by self learning.[5-6]

#### IV. ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS)

For implementing the current system ANFIS is coded in matlab environment.

Following given is the ANFIS model information:

Number of nodes- 78

Number of linear parameters- 27

Number of Non-Linear Parameters-27

Total number of Parameters-54

Numbers of training data pairs- 10

Number of checking data pairs-94

Number of fuzzy rules-27

#### V. ANFIS RESULTS

The anfis has been trained and tested for 94 different transformer DGA data. The results obtained are also comparable. Out of 94 different data, the proposed system is able to classify almost all the faults . only four combinations are left undetected. The results are clearly visible from the figure 4. The average error is 0.21. Here the anfis –window is shown. Here different colors (blue and red) dots are shown. These are actually the outputs of our system. The blue color dots represent the actual values while the red dot represent output calculated from existing membership functions. Firstly the system was trained with the available known data set and then it is tested for real transformer data. The classification rate is nearly 90% for the same.

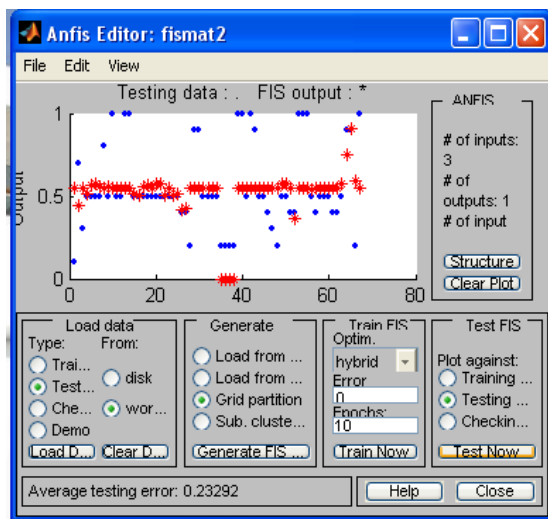


Fig. 2 : Anfis Results

#### VI. CONCLUSION AND FUTURE WORK

The Anfis has been trained and tested for 94 different transformer DGA data. The results obtained are also comparable. Out of 94 different data, the proposed system is able to classify almost all the faults . only four combinations are left undetected. The results are clearly visible from the figure 2. The average error is 0.23. Here the anfis –window is shown. Here different colors (blue and red) dots are shown. These are actually the outputs of our system. The blue color dots represent the actual values while the red dot represent output calculated from existing membership functions. Firstly the system was trained with the available known data set and then it is tested for real transformer data. The classification rate is nearly 90% for the same.

#### REFERENCES

- [1] A.singh and P Verma ,”A review of intelligent diagnostic methods for condition assessment of insulation system in power transformers” in Condition Monitorig and Diagnosis, 2008 CMD 2008 international Conference on 2008, pp.1354-1357.
- [2] Kelly,J.J, ‘Transformer Fault Diagnosis by dissolved gas analysis ”IEEE Trans. On Industry Applications, vol.16, no. 4. Pp 777-782, Dec.1980.
- [3] M.Hongzhong, L.Zheng, P.Ju.H Jingdong, “ Diagnosis of Power Transformer Faults on Fuzzy Three Ratio Method” in Power Engineering Conference, 2007 IPEC 2005 The 7<sup>th</sup> international 2005.pp 1-456
- [4] Rogers, R, “ IEEE and IEC Codes to interpret incipient faults in transformer, using gas in oil analysis .” IEEE Trans. On Electr. Insu. Vol. 13, no. 5, pp.349-354, October 1978.
- [5] Su,Q., Lai, L.L, and Austin p, “A fuzzy dissolved gas analysis method for the diagnosis of diagnosis of multiple incipient fault in a transformer” IEEE Transactions on Power System, vol, No May 2000 pp 593-598.
- [6] W. Zhenyuan, L. Yilu. And P.J.Griffin, “Neural net and expert systems diagnose transformer faults.” Computer Applications in Power IEEE vol 13, pp,50-55, 2000
- [7] S .Mofizul Islam T Wu and G Ledwich, “ A novel fuzzy logic approach to transformer fault diagnosis “Dielectric and Electrical insulation IEEE transaction on vol, pp. 177-186, 2000.
- [8] Wang, Z.Y., Liu, Y .L., and Griffin , P.J, “A combined ANN and expert system tool for

- transformer fault diagnosis.” Power Engineering Society Winter meeting, 2000. IEEE, Vol 2, 23-27 Jan 2000, pp;1261-1269 vol 2
- [9] M. Duval, “New Techniques for dissolved gas in oil analysis ,” Electrical insulation magazine, IEEE, vol 19 pp.6-15,2003
- [10] N.A. Muhamad. B.T. Phung and T.R.Blackburn , “Comparitive study and analysis of DGA methods for mineral oil using fuzzy logic,” in power engineering conference,2007,IPEC.2007 international 2007 pp1301-1306.



# Service Oriented Architecture Based Customer Relationship Management For Full Service Integration

K A Haseena & P P Abdul Haleem

Department of Computer Science and Engineering, MES College of Engineering, Kuttippuram, Kerala, India

---

**Abstract** -As customer started getting more and more importance in enterprises, it resulted in the development of Customer Relationship Management (CRM) module. Service Oriented Architecture (SOA) based CRM is introduced to enhance business based on sales, services etc., Initial focus of full fledged CRM was towards mobile telecom sector. Here a study of CRM that is designed to offer full telecom service, covering both fixed and mobile telephone services is done. In this CRM, the existing mobile CRM is reconstructed by service atomization method, to offer full telecom service. The service atomization layer concept basically introduces a new layer into existing mobile CRM architecture. The merits and demerits of this system are identified. The merits include the ability of such a system to accommodate all service requirements of telecom, its capability to support continuous release of new telecom products and services and its advantage of decreasing the maintenance complexity of business logic code. The demerit of the system is that such a system results in increased database storage since both mobile and landline systems data need to be stored. It is concluded that such a system will enhance the efficiency of telecom system.

**Key words** - CRM; SOA; Full Service Integration.

---

## I. INTRODUCTION

CRM is the front-end system which holds all the customer related data. It is the entry point. CRM started gaining more and more importance and several features got introduced into CRM. In recent years, there has been a market demand for landline + mobile services. In countries like China and Italy, the operators started thinking about accommodating the landline services into mobile CRM. If the mobile CRM can accommodate the landline services too, the operators will get benefited in many ways. The business requirement of accommodating landline services into the mobile CRM led to the research in SOA based CRM.

### A. CRM

CRM is a business strategy to acquire, grow and retain profitable customer relationships, with the goal of creating a sustainable competitive advantage. Product/price-based differentiation is waning because of four broad trends: maturing markets, global trade, efficient manufacturing and the Internet. Now CRM is emerging as a critical strategy simply because relationships are coming to the forefront of the competitive battleground. The beneficiaries of the competition being consumers, the telecom players in today's environment are required to design and deploy customer-centric strategies not only to grab a share in the market but also to sustain in the market in the long-run. The players have realized the importance of

constant service-quality delivery to the customers for long-run sustainability. Customer relationship signifies identifying the needs of the customers and stretching out ways and means to satisfy them. To be precise, it means achieving high customer profitability and customer revenues over and above customer costs, which demands matching customer expectations with customer satisfaction. The high cost of customer acquisition is making today's businesses understand the importance of retaining the customers for long-run sustainability. CRM aims at narrowing the gap between the company and its customers. In Telecom Sector, CRM plays a vital role in bringing the customers close to the company, and in identifying the changing behavioral pattern of the customers as well. In technology-dynamic markets like telecom, an efficient CRM system is essential, since the customer attrition is high due to the presence of close substitutes and near-zero switching costs.

### B. Service Oriented Architecture (SOA)

A SOA is essentially a collection of services. These services communicate with each other. The communication can involve either simple data passing or it could involve two or more services coordinating some activity. SOA is a flexible set of design principles used during the phases of systems development and integration in computing. A system based on a SOA will package functionality as a suite of interoperable services

that can be used within multiple, separate systems from several business domains.

CRM is built on business demand (or services). These services are to be loosely coupled from each other. Only then, it will be flexible enough to modify the services easily on business demand change. The code changes can be easily one without much impact on other services. For SOA based, the basic necessity is to identify the services.

Few of the services identified are [3]: (i) Customer Information Management, (ii) Business Management, (iii) Sales Management, (iv) Service Management, and (v) Call Centre.

These services, combined with a lot of other service modules are defined and the CRM is built on these modules. Thus the CRM becomes SOA based CRM. In SOA based, each service modules talk to each other through well-defined interfaces. The modules are loosely coupled from each other, making it flexible for any enhancements.

### *C. Integration of mobile + landline services*

A CRM strategy assists firms to earn advanced profits, increase customer perceived value, and acquire new customers. The existing mobile CRM caters the needs of customers to access, update and interact with customer data wherever they are. This functionally is to be extended to cover the landline users also without much complexity on the entire system. This can be achieved by the service atomization techniques supported by SOA.

The rest of the paper is organized as follows: Section II surveys the historical development of full service CRM in telecom. Section III describes full service integration in detail. The merits and demerits of full service integration systems are discussed in section IV. Section V concludes the work and mentions its future enhancements.

## **II. LITERATURE SURVEY**

Service providers are very much concerned about reducing cost per service and increasing the average revenue per user. Bundling of services in combination with different tariff schemes is adopted as a viable strategy by the service providers for this purpose. In [4] it was investigated how service oriented principles already introduced in the business domain can be extended to telecom domain, to achieve the service bundling. An approach for telecommunication service creation on SOA through service orchestration was detailed. The advantages of this include transformation of telecom services into loosely coupled reusable communication component. It was easy to develop,

deploy, execute, monitor and manage user centric applications in this flexible service environment. The cost of service delivery was also lowered.

In [3] a new approach to build a CRM based on SOA was investigated. The advantages include identification of service requirements needed for telecom manufacturing enterprises and introduction of architecture for CRM solution based on SOA to represent the interconnection of different value added activities which could bring profits with enterprise. Due to these advantages it was able to adapt to the ever-changing demands of the market to speed up the accomplishing process of information. In addition to this several case studies based on SOA embedded with CRM could be seen in the literature ([6] for instance).

The effect of a paradigm shift from Data Oriented Approach (DOA) to SOA in telecom CRM is the topic of study in [2]. The dependability between the existing three layers in DOA (application layer, middleware layer and database layer) resulted in the forced alteration of other layers if one layer is changed. An architecture of CRM system based on SOA for telecom was presented [2]. The advantages of SOA based CRM over DOA based CRM include, introduction of loose coupling and well defined interfaces.

In [1] the existing mobile CRM was reconstructed by a new technique called service atomization method to offer full telecom service including fixed and mobile telecom services. The service atomization layer concept basically introduces a new layer called the service layer into the existing mobile CRM architecture. The advantages of this new architecture are: (i) it helped to support continuous release of new telecom products and services by flexible combination and separation of telecom business, (ii) the maintenance complexity of business logic code of CRM was decreased, and (iii) it also smoothened the information gap and helped to promote the transaction processing throughput rate of business operating system by decreasing average business transaction handle time

It can be concluded that the research direction is towards providing SOA oriented architecture for CRM, as an alternative to existing DOA oriented architecture. SOA principles that existed in business domain are extended into telecom sector to meet the growing needs of the telecom customers. The SOA based CRM helped to meet the increasing demand of customers for full service by reconstructing existing mobile CRM, to offer full telecom service including mobile and fixed telephone services.

With a view to integrating mobile CRM with landline services, SOA based service atomization techniques can be applied. As the business identified the

importance of retaining customers in long-run, CRM system was designed to narrow the gap between company and its customers. Such a system is expected to meet the increasing demand of customers for full service by reconstructing existing mobile CRM, to offer full telecom service including mobile and fixed telephone services.

### III. FULL SERVICE INTEGRATION

Full Service integration in telecom deals with integrating fixed and mobile telecom services to offer full telecom service. Existing SOA based CRM for mobile system is modified using a method called service atomization for full service integration. Such a system can enhance the efficiency of telecom system by providing advantages like supporting continuous release of new telecom products and services, decreasing maintenance complexity of business logic code, and providing unified atomic service.

#### A. SOA based CRM for Full Service Integration

Mobile CRM can be easily enhanced to accommodate the future requirements like broadband, Internet Protocol Television (IPTV) etc., [1]. If the mobile CRM can accommodate the landline services too, the operators get benefited in the following business ways: (i) easy to maintain, (ii) easy to market, (iii) easy to bundle out plan with mobile + landline, and (iv) easy to give out flexible plans with mobile + landline (e.g.: credit limit set to xxx for all the connections together).

Apart from these, there is huge advantage of cost savings for the operators. They need not maintain two separate systems (one for landline and the other for mobile). It also makes their business easier, as they can feed the customer data just into one system (instead of feeding mobile customer data into one system and the landline customer data into other system). The customer also finds it much easier as they can just go to one outlet and demand for landline + mobile services. This enhances the customer satisfaction, and enhances the business for the operator, thus fulfilling one of the fierce competitions faced by them.

#### B. Service Atomization Method in SOA based CRM

The business requirement of accommodating landline services into the mobile CRM led to the research in SOA based CRM. This led to the introduction of service atomization method [1] which makes it flexible to easily accommodate the landline services into the existing mobile CRM, thereby enhancing the future flexibility of accommodating it with the future service requirements.

The service atomization layer concept basically introduces a new layer called the "service layer". Figure

1 shows the service layer introduced into SOA based CRM [1]. Functionalities addressed by each layer are [1]: (i) Presentation Layer (offer web based calling method of CRM business functions), (ii) Platform Component Layer (service calling and service loading from business function layer, promote interactive efficiency between presentation layer and business function layer), (iii) Business Function Layer ( compile and compose different atomized services into compound business logical function), (v) Atomic Service Layer (generate atomic services which are the basic component of business functions in the above layer), (vi) Technology Component Layer (loads balance and manage data exchange between atomic layer and data base management system), and (vii) Data Access and Data Storage Layer (manages database and contains process for accessing the data like stored procedure call for accessing huge data).

#### C. Merits

The merits of full service integration technique include the following [1] [3]:

- It can accommodate all service requirements of telecom (landline + mobile + any future services).
- This system can support continuous release of telecom products and services. Service atomization layer makes the code re-use easier in the business layer in case any of the services share same kind of business requirements. Thus the code modification in business layer also becomes easier.
- It is easy to perform resource modeling and service evaluation.
- Maintenance complexity of business logic code is less. The full service concept will enable the operators to remove the legacy landline CRM and rely on one single CRM for all the requirements. This eases the operators in business as there is no need to keep or feed data into different systems, and instead it can be done using one single system.
- Abstraction of business information and operations on different kind of telecom service into unified atomic service. This helps the operator to save the costs.

#### D. Demerits

The demerit of the system includes [5]: the additional overhead need for increased database storage in database archiving mechanism. Database archiving is the act of removing selected data objects from an operational database that are not expected to be referenced again and placing them in an archive data store where they can be accessed if needed. In telecom, data may need to be retained for both internal and

external reasons. Internal reasons are driven by company needs. If an organization business requires the data to conduct business and make money then that data will be retained. But external reasons, typically driven by the mandate to comply with legal and government regulations are another significant factor driving the need to store more data.

#### IV. CONCLUSION

CRM is an important application of today's convergent technology. Using SOA in CRM will decrease the maintenance complexity of business logic code of telecom CRM. In this paper a new framework based on real telecom CRM system reconstruction to accommodate full telecom service is studied, and its merits and demerits are identified. Such a system will increase the efficiency of telecom system by smoothening the information gap and will promote the transaction processing throughput rate of business operating system with decreasing average business transaction handle time. However it results in an increased requirement of database storage and archival mechanism. Future enhancements include adding more features to the CRM system to improve its applications.

#### REFERENCES

- [1] X. L. Deng, B. Wang, and X. L. Zhu, "Reconstruct and evaluate telecom CRM for full service" in Proceedings of the 2010 IEEE Asia-Pacific Services Computing Conference, ser. APSCC '10. Washington, DC, USA: IEEE Computer Society, 2010. pp 35-40. [Online]. Available: <http://dx.doi.org/10.1109/APSCC.2010.66>
- [2] F. Xiang-ling, Y. Ya-nan, S. Mao-qiang, and Q. Jia-yin, "Research of SOA based CRM in telecommunications industry," in Proceedings of 2010 International Conference on Management and Service Science (MASS). 24-26 Aug 2010, Washington DC, USA: IEEE Computer Society, 2010, pp. 1-4.
- [3] B. Li and O. Liang, "Research on customer relationship management system based on SOA," in Proceedings of the 2009 First International Workshop on Education Technology and Computer Science –Volume 02. 7-8 March 2009, Washington DC, USA: IEEE Computer Society, 2009, pp.508-511. [Online]. Available: <http://portal.acm.org/citation.cfm?id=1545020.1546420>.
- [4] T. Pollet, G. Maas, J. Marien and A. Wambecq, "Telecom services delivery in a SOA," in Proceedings of the 20<sup>th</sup> International Conference on Advanced Information Networking and Applications (AINA 2006), 18-20 April 2006, Vienna, Australia. IEEE Computer Society, 2006, pp. 529-533. [Online]. Available: <http://doi.ieeecomputersociety.org/10.1109/AINA.2006.322>.
- [5] R. Parker, "IBM case study," June 2008, Available: <http://public.dhe.ibm.com/commom/ssi/ecm/en/mc14049usen/IMC140490USEN.PDF/>.
- [6] G. Budzick, T. Dupre, "Case Study: SOA for CRM-integration at T-Mobile," 2008, Available: [http://www.cloudyintegration.com/White\\_Paper\\_Library.html](http://www.cloudyintegration.com/White_Paper_Library.html).

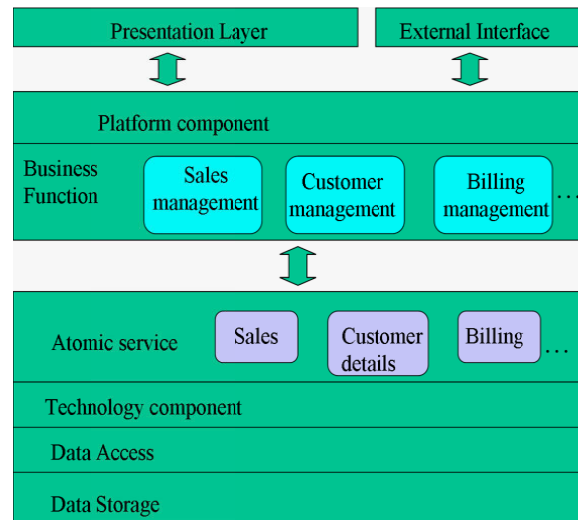


Fig. 1 : Service Layer introduced into SOA based CRM



# Channel Zapping Time Reduction Technique for IPTV

Himanshu Pandey<sup>1</sup>, Lokesh Gyanchandani<sup>2</sup>, Santosh Kumar Singh<sup>3</sup>

<sup>1</sup>Oracle Corporation, Bangalore, India, <sup>2</sup>Encora Inc., Gurgaon, India

<sup>3</sup>Division of Electronics and Communication Engineering, NSIT, Dwarka, Delhi

<sup>3</sup>M T N Ltd. (Govt. of India Enterprises), New Delhi, India

---

**Abstract** - Internet Protocol Television (IPTV) delivers television content over an IP infrastructure and has the potential to become a serious alternative to traditional broadcast schemes, such as cable, satellite or terrestrial, to deliver TV to the home. Channel Change/Zap time is the time it takes for the new TV channel to start playing from the instant a request to zap to that channel occurred.

In digital TV systems such as IPTV where TV channel content is transported as an MPEG-2 or MPEG-4 data stream encapsulated in IP packets, zapping latency is a significant concern. This delay in switching, which can anywhere from less than a second to a few seconds, can hinder the process of IPTV deployment at a large scale and pose scalability issues for IPTV.

To alleviate the problem to some extent, this paper presents a channel prefetching mechanism which can be implemented on the Digital Subscriber Line Access Multiplexer (DSLAM) or the Residential Gateway and also involves, though not necessary, extending the functionality of the set-top-box (STB). The idea is to fetch content of a number of channels before they are actually requested. This selection is done by maintaining rankings of atleast one of the available channels at the ranking engine connected to the DSLAM based on their popularity. The rank is atleast partly based on the data stored on the database residing at DSLAM.

**Key words** - DSLAM, IPTV, MPEG, QoE, Set top box, Zap time.

---

## I. INTRODUCTION

Since the birth of internet, its demand in every application is growing day by day. No wonder, like every other service, television is also getting attention. IPTV is not a new concept as such. Early days of internet did not have enough bandwidth to support entertainment quality media delivery. Hence broadcast over IP network did not become popular.

The recent advancement in communication network and media compression techniques has now enabled network delivery of high quality content. IPTV is a new form of television technology that uses the existing IP network to deliver entertainment grade audio-video content to consumers. It uses video compression techniques to reduce the data to be transported to the end user. The compressed digital media is then transported to end users over a standard IP network which is already in place for data services. Data, Voice and Video are three components of the triple play service. IPTV falls into the category of Video.

The Key promises of IPTV System are Integrated and interactive Service. Since, consumers prefer a reliable one-stop service from a trusted provider, thus IPTV is a preferred choice from the customer perspective.

Also, IPTV Service has been preferred choice from service provider perspective, mainly due to the following reasons:

- Flexibility: Since the underlying infrastructure is a general network, the upper layer can be modified and altered without fiddling with the base.
- Low Cost: The deployment cost of the network infrastructure is low as the IPTV Service being built on top of the existing data.

As the border between broadcasting and IP communication is indistinct, the convergence between broadcasting and IP communication is getting faster and faster. IP communication companies have recently started entering into the broadcasting market in collaboration with broadcasting or movie companies. Broadcasting and cable companies are coming into high-speed internet service business. IPTV can support not only passive broadcasting service but also active, intelligent and bidirectional (interactive) services such as VoD (Video on Demand) & Time shifted Tele Vision (TSTV); T (Television)-Commerce. Therefore IP communication companies expect that IPTV will be the representative service which can find a means of another huge earning.

The paper is organized as follows: In 2nd section, we have described the main factors affecting IPTV service. 3rd section discusses the parameters affecting channel zapping time. 4th section deals with the existing implementations to reduce zapping time. Finally, in the last one (5th section), we have discussed our proposed implementation to reduce zapping time.

## II. FACTORS AFFECTING IPTV SERVICE

A typical IPTV configuration from the DSLAM to the customer premises is shown in Figure 1. The video stream is delivered using ADSL2+ from the IP-based DSLAM to the user's ADSL2+ broadband router. The router, while supporting voice and Internet service, passes the video stream to the STB for decoding. The STB converts the video stream into required signals for displaying on the consumer's TV. In the whole processes of IPTV service delivery, following are factors affecting the Quality of Experience (QoE).

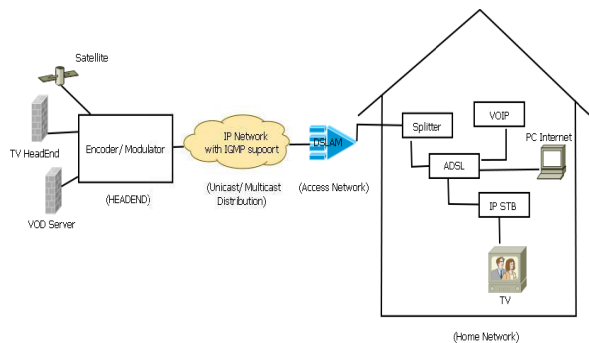


Fig. 1: Basic IPTV Architecture

**Encoding and Compression:** The quality of the video being distributed across the network can be affected right at the source; i.e., at the video head end. The encoding and compression process usually creates a trade-off between the quality of the video and the desired compression level. In addition to it, the amount of video information per IP packet will vary depending on the encoding and compression technique used. Therefore, an IP packet loss can represent a single unnoticeable missing point of the video sequence or a large period of degraded, pixelated or unavailable image.

**Jitter:** Jitter is defined as a short-term variation in the packet arrival time, typically caused by network or server congestion. If the Ethernet frames arrive at the STB at a rate that is slower or faster, as determined by the network conditions, buffering is required to help smooth out the variations. Based on the size of the buffer, there are delivery conditions that can make the

buffer overflow or underflow, which results in a degradation of the perceived video.

**Limited Bandwidth:** The total amount of video-stream data that can be sent is limited ultimately by the customer's actual ADSL/ADSL2+ rate. Core IP infrastructure is usually based on optical networks with a low level of congestion; therefore, bandwidth limitations are commonly located only within the access network or the customer's home network. When traffic levels hit the maximum bandwidth available, packets are discarded, leading to video quality degradation. ADSL2+ rates may be temporarily affected by external factors, which in turn can generate pixelization of the image. Another situation might occur when, in addition to the IPTV service, a high amount of data is downloaded simultaneously to a PC and the traffic priorities have not been assigned correctly by the service provider; in these cases, video streaming packets are lost. Bandwidth limitation is one of the main factors to be evaluated during the network design stage.

**Packet Loss:** Loss of IP packets may occur for multiple reasons — bandwidth limitations, network congestion, failed links, and transmission errors. Packet loss usually presents a bursty behavior, commonly related to periods of network congestion. Depending on the type of transport protocol used for the video streaming, a packet loss will have different impact on the quality of the perceived video. When UDP is used, the lost packets will directly affect the image, as the information cannot be recovered and the image will simply be corrupt or unavailable. When using TCP, a packet loss will generate a retransmission, which can produce a buffer underflow and, consequently, a possible frozen image.

**Channel Switching:** It refers to the delay between a new channel request and its display on the monitor. Though the acceptable delay between a channel request by the user and the appearance of the frame on the screen is only 200 ms, in the current scenario, it is as high as a few seconds. This will be discussed in more detail in the following section [1].

## III. IPTV CHANNEL ZAPPING TIME

Zapping time is the total duration from the time viewer presses the channel change button, to the point the picture of the new channel is displayed, along with the corresponding audio.

In an IPTV network, traditional broadcast television channels are delivered via IP multicasting. Broadcast TV makes use of IP Multicasts to deliver the programming efficiently through the IP system. A Multicast is designed to allow multiple users



simultaneous access to the session. At the request of the viewer, the selected programming is located from within the network (from a server) and a unique unicast is setup to deliver the program to the user. This is in effect a private network connection between the server and the viewer's STB.

This process of delivery of multicast traffic to interested and authorized users is done through Internet Group Management Protocol (IGMP). IGMP is used by IP hosts to register their dynamic multicast group membership. It is also used by connected routers to discover these group members.

In order to preserve bandwidth over the final link to the house, IPTV systems are designed to deliver only the requested channel to the STB. If the user switches the channel, the STB has to perform the process of sending IGMP multicasting tree, and using its complex protocol to tell the upstream equipment to stop sending ("leave") one channel or begin sending ("join") another channel. After waiting for the object video stream to come, STB waits more for a decodable frame, which is called an Intra-coded frame (I-frame). Then STB buffers some frames to avoid the unsmooth display caused by the delay jitter over the Internet. This process requires a finite time (Zapping Time) to complete and the time taken is heavily influenced by transmission delays in the network which in turn has a direct impact on the channel change timings of the system. In essence, in IPTV systems, the channel change is made in the network and not on the local STB. While preserving precious last mile bandwidth, this approach presents a number of challenges to the scalability and usability of the system.

Comparing it with traditional cable TV, Zapping latency is negligible in this case as all channels are present and changing channels meant switching to a new frequency. It is never a challenge and definitely not a source for user dissatisfaction. All the channels are delivered to the STB in the home (via Cable, Satellite or Terrestrial). There could be hundreds of channels, all of which are delivered simultaneously. The STB tunes to the desired channel in response to requests from the viewer's remote control. As a result of this local tuning the channel changes are almost instantaneous.

Typical Channel Switching Process - There are a number of sequential steps involved in the process of switching a channel [2].

- User initiates a request. (Usually in the form of a button press on the remote).
- If the channel is not already being transmitted to the STB, it translates the request into an IGMP Leave request and sends to the network.

- The End router responds with an IGMP Query or forwards the Leave message to upper level routers. This continues till the server end is reached.
- Next, the STB sends an IGMP Join request. It also propagates in the same fashion.
- Server node prepares a new stream for the requested channel and begins transmission of the stream.
- STB receives the stream from the network, performs buffering, decoding and sends the picture to the television.

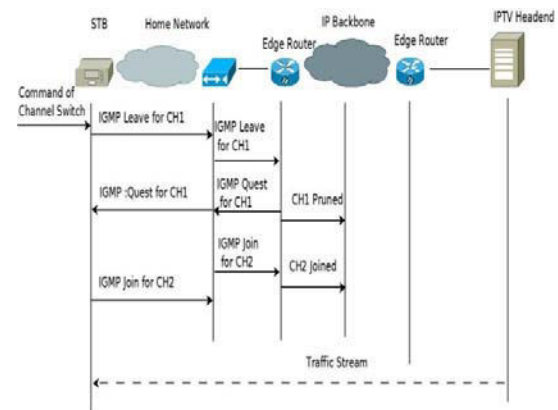


Fig. 2 : IGMP process of IPTV channel switch

One of the approaches to reduce the channel zapping time is for the STB to join the adjacent channels of the current channel in advance. If the user switches to an adjacent channel, the user can watch the selected adjacent channel without any channel zapping time because the stream of the adjacent channel is already being sent to the STB. This method can be expanded to include more channels that are highly likely to be selected next [3], [4]. This kind of method is called 'prejoining method' or 'predictive tuning'.

#### IV. EXISTING IMPLEMENTATIONS FOR ZAPPING TIME REDUCTION.

Important prediction algorithms are listed as follows. They all use information from user and try to predict the next channel to be switched to. The table lists describes the properties and merits/demerits of these algorithms [5].

TABLE 1: Zapping Time Reduction Techniques

Technique	Merits	Demerits
-----------	--------	----------

Channel Zapping Reduction Based on Rating Server [6]	Uses rating server to record activities, and predicts channels.	High bandwidth, Complex.
User modeling and Recommendation based on 3 different user models [7]	Uses three different sources to gather information about user's behavior.	Complex, High Processing needed, Expensive.
Recommendation System Based on Fuzzy Logic ( <i>Shi et al.</i> ) [8]	Evaluates degree of user's interest for particular channel using fuzzy logic.	Takes some time to predict sudden mood changes.

## V. PROPOSED ZAPPING TIME REDUCTION ALGORITHM

We propose a channel ranking algorithm to ‘guess’ the next most probable channel that the consumer could switch over to. It’s a ‘parameter’ based algorithm that assigns ‘weights’ to the parameters. We presume that the channels are categorized as per their genre. Following briefly describes all the parameters taken into consideration:

### (1) Channels of same category (G):

- As for the user, the next channel to be browsed has high probability to belong to same category/ genre as the current channel.
- For a channel, let  $W_G$  denote weight for this parameter. All the channels of same genre are given same fixed weight as per their ranking.

### (2) Channel Count (C):

- For each subscriber we maintain a list of channels and his visit count. We take top  $N$  channels for our calculations.
- For a channel, let  $W_C$  denote weight for this parameter. Channels are assigned weight  $W_C$  as per their ranking.

### (3) Time Spent on Channels (T):

- For each subscriber, we maintain a list of channels and the time spent by the subscriber on particular channel. We take top  $N$  channels for our calculations.
- For a channel, let  $W_T$  denote weight for this parameter. Channels are assigned weight  $W_T$  as per their ranking.

### (4) Favourites (F):

- We allow users to set some of their channels as Favourites.
- For a channel, let  $W_F$  denote weight for this parameter. All the ‘Favourite’ channels are given a fixed parameter weight.

### (5) Seasonal Channels (S):

- When some event occurs, TRP’s of a particular (or set of) channel increase. Like for example at time of cricket world cup people will watch Neo Cricket more. Similarly News Channels get more TRP when there is some Breaking News.
- For a channel, let  $W_S$  denote weight for this parameter. Weights for seasonal channels are variable and should be assigned by ISP.

### (6) Other Factors (O):

- This parameter is IPTV provider specific provided to take care of their particular requirements. This is a reserved parameter.
- Let the weight assigned to this parameter be ‘ $W_O$ ’.

### (7) Previous History (W):

- We take into account the previous history by storing weights of the channels calculated on last iteration.
- We multiply the weights from the “info” table with a coefficient ( $\alpha$ ) to get effective weight to be taken in to calculations. Let it be denoted by  $W_W$ .

## A. Algorithm

### Glossary:

$N$  : Number of top channels to be considered for prioritization.

$N_{\text{channels}}$  : Total number of channels subscribed by user

$N_{\text{buffer}}$  : Number of channels to be buffered

**info** : Table of channels sorted by “ID” Fields ( $P_{ID}$ ,  $P_{CID}$ ,  $P_W$ ,  $P_C$ ,  $P_T$ ,  $P_G$ ,  $P_S$ ,  $P_O$ ,  $P_F$ ) where each represents column index (“col\_index”) of the table corresponding to the parameter

**info** ( $x, y$ ) : Returns value of cell in “info” table corresponding to  $x^{\text{th}}$  row and  $y^{\text{th}}$  column.

**ID** : Primary key of the table.

**ID<sub>current</sub>** : ID of the current channel (presently being viewed by the user)

**BufferChannelList**

: Final list of channels to be buffered sorted by current weight of the channel (represented by column “ $P_W$ ”)

**list** <id> : List that stores “ID” of channels.

**Algorithm :****updateWeightsAndReturnChannelList()**

```

// Initialization
 $T_{lower} \leftarrow P_W$ 
// starting column-index
 $T_{upper} \leftarrow P_F$ 
// ending column-index

// Iterate through each column one by one
while ( $T_{lower} < T_{upper}$ ) do
  list  $\leftarrow$  getTopNChannelsByProperty( $T_{lower}$ )
  while (list  $\neq$  null) do
    // check if column being considered is  $P_W$ 
    if ( $T_{lower} = P_W$ )
      info (list.id,  $P_W$ )  $\leftarrow$  0

      info (list.id,  $P_W$ )  $\leftarrow$  info (list.id,  $P_W$ ) +
        getWeightByParameter ( $T_{lower}$ )
      list  $\leftarrow$  list.next()
      // next item in “list”
    end while
     $T_{lower} \leftarrow T_{lower} + 1$ 
  // next column
end while

// Get top  $N_{Buffer}$  channels by their Weight
BufferChannelList  $\leftarrow$  getTopChannels ( $P_W$ ,  $N_{Buffer}$ )
return BufferChannelList

end func

```

**getTopNChannelsByProperty (col\_index)**

```

if (col_index =  $P_S$  or col_index =  $P_O$  or col_index =  $P_F$ )
  id  $\leftarrow$   $N_{channels}$ 
  while (id > 0) do
    // Check if Favourite Flag is set to 1
    if (info (id, col_index) = 1)
      list.add (id)
      id  $\leftarrow$  id - 1
    end while
else if (col_index =  $P_G$ )
  id  $\leftarrow$   $N_{channels}$ 
  while (id > 0) do
    // Check if channel being considered is of
    // same
    // category as the currently viewed channel
    if (info (id,  $P_G$ ) = info ( $ID_{current}$ ,  $P_G$ ))
      list.add (id)
      id  $\leftarrow$  id - 1
    end while
else
  list.add (getTopChannels (col_index, N))
end func

```

**getTopChannels (col\_index, N)**

// Returns top “N” entries from the info table sorted by descending order of the column corresponding to “col\_index”.

**getWeightByParameter (col\_index)**

// Returns the weight assigned to the parameter represented by “col\_index”.

**B. Algorithm Details**

The algorithm aims to rank all the channels on basis of above seven parameters, and the channel with maximum total weight is declared as the most probable. As per the bandwidth available, ISP has an option of streaming additional channels to the consumer, thus reducing the zapping time, if the correct channels are guessed.

To calculate the correct weights, a table is maintained at DSLAM for each subscriber, containing data for all the channels the consumer has subscribed to. We have called it “**info**”. It has the following fields: **ID**, **C\_ID**, **C**, **T**, **G**, **S**, **O**, and **F**. On every request for new

channel from subscriber, the data for channel count and time spent on current channel is increased in the table. Accordingly, new weight calculated is also updated. The algorithm executes for each row of the table, and returns a list of channels which most probably would be selected next.

The biggest question to this algorithm is where the table for each subscriber should be stored and calculation done. The ranking of channels can be done at STB level and also at DSLAM servers. Both approaches have their merits and demerits but we prefer storing the tables at DSLAM servers with dedicated ranking engine. Among other advantages, this scheme is economically more feasible as not all the STB's would require upgradation, as it is done only for DSLAM - which is a onetime investment.

TABLE 2: Info Table

ID	C_ID	W	C	T	G	S	O	F
1	ch_101	4	4	2	1	0	0	1
2	ch_102	5	5	1	1	0	0	0
3	ch_103	9	80	30	1	0	0	1
4	ch_201	8.8	30	12	2	0	0	1
5	ch_202	0	0	0	2	0	0	0
6	ch_301	0	0	0	3	0	0	0
7	ch_302	3	4	2	3	0	0	0
8	ch_303	7	20	16	3	0	0	1
9	ch_304	0	0	0	3	0	0	0
10	ch_305	3	8	4	3	0	0	1
11	ch_401	3	3	3	4	0	0	0
12	ch_402	2.1	6	2	4	0	0	0
13	ch_403	5.1	18	6	4	0	0	1
14	ch_501	6.8	10	11	5	1	0	0
15	ch_502	1.4	15	2	5	0	0	0
16	ch_503	2	12	4	5	0	0	1
17	ch_504	0	0	0	5	0	0	0
18	ch_505	0	0	0	5	0	0	0
19	ch_601	2.2	8	3	6	0	0	0
20	ch_602	3.6	5	5	6	0	0	0

### Glossary:

C\_ID : Channel ID ; W : Weight ; C : Channel Count ; T : Time Spent on Each Channel ; G : Channels of same Genre/Category ; S : Seasonal Channels ; O : Other Factor ; F : Favourite Channel

### C. Example for Algorithm Implementation

To predict the next channel with the highest probability, we ran this data through our algorithm to update weights in the info table of the consumer, as well get top channels to be buffered as BufferChannelList, which will have the  $N_{buffer}$  sorted by weight. The Beauty of the algorithm is that, it does not matter how many channels does the subscriber subscribes to, the processing will be done only on a fraction of channels, which subscriber regularly switches to. So in above sample data, channels like ch\_602 would not ever get into processing, hence reducing unnecessary wastage of resources.

## VI. CONCLUSION

The paper has provided a channel ranking algorithm to predict the next most likely channel which could be selected by the subscriber. The algorithm can be implemented both at DSLAM level or STB level but preference for DSLAM level is proved. Further it is suggested that there is much scope for future work in this paper i.e. by enhancing the user profiling, hence improving the accuracy of weights that should be assigned to the channels in our algorithm.

## REFERENCES

- [1] Francisco Palacios, "IPTV TESTING OVER DSL", [http://documents.exfo.com/appnotes/anote\\_148-ang.pdf](http://documents.exfo.com/appnotes/anote_148-ang.pdf)
- [2] S. K. Mandal and M. Mburu, "Intelligent Prefetching to Reduce Channel Switching Delay in IPTV Systems".
- [3] Harald Fuchs, Nikolaus Färber, "Optimizing channel change time in IPTV applications", International Symposium on Broadband Multimedia Systems and Broadcasting, IEEE, Las Vegas, June 03, 2008, pp. 1-10.
- [4] Hyunchul Joo, Hwangjun Song, Dai-Boong Lee, Inkyu Lee, "An Effective IPTV Channel Control Algorithm Considering Channel Zapping Time and Network Utilization", IEEE Transactions on Broadcasting, IEEE, May 23, 2008, pp. 1-10.
- [5] Muhammad Zeeshan Ahmad, Nabil Ur Rehman, Junaid Qadir, Adeel Baig, Hammad Majeed, "Prediction-based Channel Zapping Latency

- Reduction Techniques for IPTV Systems - A Survey", 2009 International Conference on Emerging Technologies, IEEE, Islamabad, December 11, 2009, pp. 1-10.
- [6] J. Lee, G. Lee, S. Seok, and B. Unknown, "Advanced Scheme to Reduce IPTV Channel Zapping Time," in Managing Next Generation Networks and Services. Springer Berlin /Heidelberg, 2007, vol. Volume 4773/2007, pp. 235-243.
- [7] L. G. C. a. T. P. Ardissono, "User Modeling and Recommendation Techniques for Personalized Electronic Program Guides," in Personalized Digital Television – Targeting Programs to Individual Viewers, volume 6 of Human-Computer Interaction Series, chapter 1. Kluwer Academic Publishers, pp. 3-26.
- [8] S. Xiaowei, "An Intelligent Recommendation System Based on Fuzzy Logic," in Information in Control, Automation and Robotics I. Springer Netherlands, 2006, pp. 105-109.
- ◆◆◆

# Minimax Filtering In Wireless Sensor And Actor Networks

K. Ramesh & S.Vasundra

Department of CSE, JNT University, Anantapur, India

---

**Abstract** - In this paper to handle the mobility of actors a hybrid strategy that includes location updating and location prediction is used. The usage of Kalman Filtering in location prediction high power and energy consumptions. To avoid the drawbacks of Kalman Filtering in location prediction, we make use of Minimax filtering (also Known as  $H^\infty$  filtering). Minimax Filter has been used in WSANs by minimizing the estimation error and maximizing the worst case adversary noise. Minimax filtering will also minimize power and energy consumptions.

*Key words* - Wireless Sensor and actor networks,, minimax filtering, mobility, power efficient.

---

## I. INTRODUCTION

Wireless Sensor and Actor networks (WSANs) are distributed wireless networks of heterogeneous devices referred to as sensors and actors. In Wireless Sensor and Actor Networks (WSANs) the collaborative operation of sensors enables the distributed sensing of physical phenomenon, while actors collect and process sensor data and perform appropriate actions. Sensors are low-cost, low-power, multi functional devices that communicate in short distances. Actors are resource-rich devices equipped with high processing capabilities, high transmission power and long battery life. Actors collect the sensor data and process that data and consequently perform actions in the network.

WSANs are used in several applications. In some applications, actors are part of the network and perform actions based on the information gathered by sensors. For example, In Distributed Robotics the task is not completed by a single robot but a team of collaborating robotics. Information about the surrounding environment is usually gathered by onboard sensors and team members exchange sensor information to move or perform actions.

In WSANs, the sensors distribute the sensed data after detecting an event that is occurred in the environment. The event data are distributively processed and transmitted to the actors, which gather, process, and eventually reconstruct the event data. When an event occurred in the environment the data flows between sensors and actors, this process is referred to as sensor-actor coordination. Once an event has been detected, actors coordinate to reconstruct it, to estimate the event characteristics and make a collaborative decision on

how to perform the action. This process is referred to as actor-actor coordination.

Sensors and actors are movable devices in WSANs. In previous work on WSANs [1], uses location management scheme for location updates and location prediction. It uses Voronoi diagrams for location updation and Kalman Filtering for location prediction. Kalman Filtering used in WSANs will cause some problems. The main drawbacks of Kalman Filtering are it fails to identify the unknown noise and fails to minimize the estimation error. By replacing Kalman Filtering with Minimax Filtering we can minimize the estimation error and it can identify the worst case adversary unknown noise.

## II. RELATED WORK

Wireless Sensor and Actor Networks (WSANs) research challenges and several application scenarios are described along with challenges for effective sensor-actor coordination and actor-actor coordination in [7]. As discussed in [2] There are many challenges in Wireless Sensor and Actor Networks, especially due to resource constraints. In [5], [8] the authors considered the issue of real-time communication in sensor networks. The SPEED protocol [5] provides real-time communication services and is designed to be a stateless, localized algorithm with low control overhead. MMSPEED [8] is an extension of SPEED that can differentiate between flows with different delay and reliability requirements. SPEED and MMSPEED try to provide real-time delivery of individual flows from different sensors. None of these papers deals with sensor-actor coordination or with actor-actor coordination.

Some research papers like [9] deals with the problem of mutual exclusion in WSNs. In [6], the authors deal with the problem of “hazards” which consists of out-of-order execution of queries and commands due to the lack of coordination between sensors and actors. To enable a wide range of trade-offs between delay and energy consumption that controls the wake-up cycle of sensors based on the experienced packet delay [13] presents a delay-energy aware routing protocol (DEAP) for sensor and actor networks. However, the paper only deals with sensor-actor communication.

The mobility of WSNs has been handled in [1] with sensor-actor coordination and actor-actor coordination using power-controlled energy-delay adjustment and event preemption for multi actor task allocations respectively. It also provides a solution for multi actor task allocation problem by selecting the best actor team that minimizes energy consumption. However it uses Kalman filtering in location prediction, which is having two major limitations. First, Kalman Filtering assumes that the noise properties are known, if the system have unknown noise then it fails to identify that. Second, Kalman Filtering minimizes average estimation error and fails to minimize worst case estimation error.

The above limitations gave raise to Minimax Filtering also known as  $H_\infty$  filtering. The usage of Minimax Filtering in Wireless Sensor Networks has been discussed in [3]. The Minimax filter is a robust filter that minimizes the estimation error by considering the worst case noise.

### III. LOCATION MANAGEMENT

The network is composed of  $N_S$  sensors and  $N_A$  actors, with  $N_S \gg N_A$ . Each sensor is equipped with a low data rate radio interface. Actors are equipped with two radio transmitter i.e., a low data rate transmitter to communicate with the sensors and a high data rate wireless interface for actor-actor communication so that each sensor will route information to its closest actor, unless an alternative actor is preferable in case of congestion.

In general location management may follow two strategies: *location updating* and *location prediction*. Location updating is a passive strategy in which each actor periodically broadcasts its position to the neighboring sensors. Location prediction is a dynamic strategy in which sensors proactively estimate the location of their neighboring actors. *In this case we used Minimax Filtering for predicting the positions of the actors* for location updates we proposed location management scheme based on spatial and temporal domains. In spatial domain update messages sent by

mobile actors to sensors. Therefore, sensor-actor communications are localized. Hence in the spatial domain, broadcasts can be limited to Voronoi diagrams [11]. In the temporal domain, location updates can be limited to actor positions that cannot be predicted at the sensor side. Location updates are triggered at the actors when the actual position of the actor is far away from the predicated sensor based on past measurements. Therefore, actors that moves following predictable paths.

#### A. Voronoi diagrams in location updation

We use Voronoi diagrams in location updates. The Voronoi diagram of a set of discrete sites partitions the plane into a set of convex polygons such that all points inside a polygon are closed to only one site. For their properties and ease of computation, Voronoi diagrams are previously used in the area of sensor networks. In [12], Voronoi diagrams are used to measure how well an object is moving on an arbitrary path can be observed by the sensor network over a period of time. In [13], an optimal polynomial-time worst- and average-case algorithm for coverage calculation with homogeneous isotropic sensors is derived.

The Voronoi cell of an actor  $a_i$  contains all points of the plane that are closer to  $a_i$  than to any other actor in the network. A sensor  $s$  is said to be *dominated* by an actor  $a_i$  if its location lies in the Voronoi cell of  $a_i$ . Every actor is responsible for location updates to sensors in its Voronoi cell. Each sensor will thus expect to receive location updates from the actor which is dominated by that sensor.

The energy consumption for location updates will drastically reduced with respect to flooding. With a flooding like protocol each actor sends a message to its  $N$  neighboring sensors. We consider the link metric  $E = 2E_{elec} + E_{amp}d^\alpha$ , where  $\alpha$  is the path loss propagation exponent ( $2 \leq \alpha \leq 5$ ),  $E_{amp}$  is a constant [ $J/(bits.m^\alpha)$ ], and  $E_{elec}$  is the energy needed by the transceiver circuitry to transmit or receive one bit [ $J/bits$ ]. Each sensor, upon receiving the message, forwards it by broadcasting again. On this first hop only, the energy consumption is  $N_A \cdot (NE_{elec} + N(E_{elec} + E_{amp}d^\alpha + NE_{elec})) = N_A \cdot (2N E_{elec} + NE_{amp}d^\alpha + N^2 E_{elec})$ . At least we need a message from each actor to reach each sensor in the network, and the same message can potentially be relayed to each other node in the network before it is discarded. This is clearly a worst-case scenario, but it provides an indication of the scaling law for the energy consumption. Instead, provided that each actor can transmit data within its Voronoi cell, no forwarding is needed, and hence, the energy consumption is in the order of the number of sensors (energy needed to receive the update packets). Hence, the worst-case energy consumption of a flooding scheme increases as a function order of  $O(N_S^2 \cdot N_A)$ , and

most of the energy burden is on the sensors. Conversely, if the actor is able to reach all sensors in its Voronoi cell in one hop, which may be true in many practical cases, the energy consumption increases as a function order of  $O(N_s)$ , and most of the energy burden is on the actors.

#### B. Minimax Filtering in location prediction

Location updates can be triggered at the actors only when the actual position of the actor is “far” from what can be predicted at the sensors based on past measurements. Therefore, actors that move following predictable paths will need to update their position much less frequently than actors that follow temporally uncorrelated paths. In [4], the Kalman filtering is used for adaptively varying frequency of location updates based on sensor side previously received updates.

We further observe that Kalman filtering is used as a means of decentralized estimation of objects in sensor networks in [14], [15] and in wireless multimedia sensor networks in [16] where as in [4], Kalman filtering is used for object tracking with the design of a location management to enable geographic routing in WSANs. In [3], Minimax filtering is used to target tracking in sensor networks. In this paper we introduced Minimax filtering in WSANs, instead of Kalman filtering for location prediction.

Minimax filtering is used to estimate the states of a dynamic system based on the measurements related to the estimated states, the measurement model and the system model. It looks the same as other state estimators, such as Kalman filtering. However, the difference is that the system state model includes fictitious adversary disturbances, which includes some partially unknown noise. Minimax filtering is a robust filter that minimizes the estimation error by considering the worst case noise. Here actors are assumed to be endowed with an onboard localization system (e.g., GPS), while sensors predict the position of actors based on Minimax filtering of sparse measurements (taken at the actor and transmitted to the sensors).

The dynamic movement model for the  $i^{\text{th}}$  actor in two-dimensional coordinates can be described by continuous-time linear dynamic system. In WSANs multiple sensors and multiple actors are presented. The sensor nodes have to estimate the position of actor nodes. From [18] the actor position can be estimated with the discrete-time dynamic equation.

$$x_i^{t+1} = A x_i^t + B w_i^t + d_i^t \quad (1)$$

Equation (1) represents the state transition equation for the system describing the motion of actor  $i$  between  $t$  and  $t+1$ , where  $x_i^t = [x_i^t, y_i^t, \dot{x}_i^t, \dot{y}_i^t]^T$  represents position and velocity of actor  $i$  at step  $t$ ;  $w_i^t = [w_i^{t,x}, w_i^{t,y}]^T$  represents the system noise in the control input;  $d_i^t =$

$[d_i^{t,x}, d_i^{t,y}]^T$  represents the adversary disturbance.  $A, B$  are the matrices of the appropriate dimensions with bounded entities. Assume that the target actor is intelligent and can maximize the estimation error. Let  $\hat{x}_i^t$  denotes the estimated state, the estimation error is  $x_i^t - \hat{x}_i^t$ . The adversary disturbance is modeled as

$$d_i^t = L(C(x_i^t - \hat{x}_i^t) + n_t) \quad (2)$$

Where  $L$  is gain to be determined,  $n_t$  is Gaussian noise with zero mean diagonal covariance matrix  $S > 0$ .  $C$  is the position observed by the actor at step  $t$  is related to the state by the *measurement equation*

$$y_i^t = C_i x_i^t + v_i^t \quad (3)$$

where  $y_i^t = [y_i^{t,x}, y_i^{t,y}]$  represents the *observed position* of the actor at step  $t$ ,  $v_i^t = [v_i^{t,x}, v_i^{t,y}]$  represents the *Gaussian noise* with zero mean and covariance matrix  $R \geq 0$ . Thus, the observed position of the actor  $y_i^t$  is, the actual position of the actor affected by a Gaussian noise.

The Minimax filter provides computationally efficient set of recursive equations to estimate the state of such process. The joint use of Minimax filter at the sensor and actor sides enables reducing the number of necessary location updates. In fact, the filter is used to *estimate the position* at the actor based on measurements, which is a common practice in robotics, and to *predict* the position of the actors at the sensors, thus reducing the message exchange. The position of actor  $i$  can be estimated and predicted at the sensors in its Voronoi cell, based on the measurements  $y_i^t$  taken at the actor and broadcast by the actor. At step  $t$ , each sensor  $s$  in  $i$ 's Voronoi cell updates the state (that represents position and velocity of the actor) based on the equations

$$\hat{x}_{i,s}^{t+1} = A \hat{x}_{i,s}^t + K(y_i^t - C \hat{x}_{i,s}^t) \quad (4)$$

Where  $K$  is the gain of the Minimax filter, the estimation error is defined by

$$e_i^t = x_i^t - \hat{x}_i^t \quad (5)$$

The Minimax algorithm can be summarized as:

$$\Sigma_t^{-1} = P_{i,s}^{-1} + C' (R^{-1} - S^{-1}) C \quad (6)$$

$$P_{i,s}^{t+1} = A \Sigma_t A' + B Q B' \quad (7)$$

$$K_{i,s}^t = A \Sigma_t C' R' \quad (8)$$

The sensor  $s$  predicts the state of actor  $i$  before receiving the measurement (a priori estimate) with (4). After receiving the measurement from the actor  $y_i^t$ , sensor  $s$  updates the Minimax filter gain  $K_{i,s}^t$  and corrects the state estimate and covariance matrix according to the measurement, using (6), (7) and (8). In particular (7) updates the covariance matrix, (8) updates the Minimax gain and (4) calculates the new state.



### III. SIMULATION RESULTS

#### A. Simulation Specifications:

OS	: Fedora 9
Simulator	: NS2
Topology	: Wireless Topology
Number of Nodes	: 49
Maximum Transmission range	: 40 m
Simulation time	: 400s
Area of the network	: 100X100 m

#### B. Simulation Results:

Network Simulator (NS2) is used for simulating the existing and proposed systems. NS2 is an IEEE standardized simulator for simulating Networks.

In Fig.1, we show a comparison of the average power consumption in WSNs using Kalman Filtering and Minimax Filtering with increasing forwarding range. The power consumption in WSNs using Kalman Filtering is drawn with green line and the power consumption in WSNs using Minimax Filtering is drawn using red line. In all the cases the power consumed by Kalman Filtering is more than that of the power consumed by Minimax Filtering.

In Fig.2, we show a comparison of energy consumption in WSNs using Kalman Filtering and Minimax Filtering. The energy consumption for Kalman Filtering is more than the energy consumption for Minimax Filtering.

### IV. CONCLUSIONS

We discussed the drawbacks of Kalman Filtering in location prediction process of a WSN. The drawbacks of Kalman Filtering are over come by Minimax Filtering. Using Minimax Filtering the estimation error was minimized by maximizing the worst case noise. By replacing Minimax Filtering with Kalman Filtering in location prediction of WSNs reduce the power and enrgy consumptions.

### REFERENCES

- [1] Tommaso Melodia, Dario Pompili, and Ian F.Akyildiz "Handling Mobility in Wireless Sensor and Actor Networks", IEEE Trans, Mobile Computing, vol. 9, no.2, Feb.2010.
- [2] J.A.Stankovic, T.F. Abdelzaher, C.Lu, L.Sha, and J.Hou, "Real-Time Communication and Coordination in Embedded Sensor Networks," Proc.IEEE, Vol.91, no.7, pp.1002-1022, July 2003.
- [3] Dongbing GU "A Gametheory approach to Target tracing in sensor networks," IEEE Trans, Systems, MAN and Cybernetics-part B: Cybernetics, Vol.41, no1, Feb.2011.
- [4] T. Melodia, D. Pompili, and I.F. Akyildiz "A Commuication architecture for mobile Wireless Sensor and Actor Networks" Proc. IEEE Conf., Sensor, Mesh and Ad Hoc Comm.and Networks (SECON), Sept.2006.
- [5] T.He, J.Stankovic,C.LU, and T.Abdelzaher "SPEED: A Real-Time routing protocol for Sensor Networks," Proc. IEEE int'l Conf. Distributed Computing Systems(ICDCS),pp. 46-55,May 2003.
- [6] R.Vedantham, Z.Zhuang, and R.Sivakumar, "Hazard Avoidance in Wireless Sensor and Actor Networks",Computer Comm.,vol. 29,nos. 13/14,pp.2578-2598,Aug.2006.
- [7] I.F.Akyildiz and I.H.Kasimoglu, "Wireless Sensor and Actor Networks: Reasearch Challenges,"Ad Hoc Networks, vol. 2, no.4, pp.351-367, Oct. 2004.
- [8] E.Felemban, C.-g. Lee, E.Ekici, R.Boder, and S. Vural, "Probabilistic QoS Guarentee in reliability and Timeliness Domains in Wireless Sensor Networks," Proc. IEEE INFOCOM,Mar. 2005.
- [9] R.Vedantham, Z.Zhuang, and R.Sivakumar, "Mutual Exclusion in Wireless Sensor and Actor Networks," Proc. IEEE.Conf.Sensor, Mesh and Ad Hoc Comm. And Networks (SECON), Sept. 2006.
- [10] A.Durresi, V.Paruchuri, and L.Barolli, "Delay-Energy Aware Routing Protocol for Sensor and Actor Networks",Proc. Int'l. Conf. Parallel and Distributed Systems, vol.1, pp.292-298, July 2005.
- [11] F.Aurenhammer, "Voronoi Diagrams- A Survey of a Fundamental Geometric Data Structure," ACM Computing Surveys, vol.23, pp.345-405, 1991.
- [12] S.Meguerdichian, F.Koushanfar,G.Qu, and M.Potkonjak, "Exposure in Wireless Ad-Hoc Sensor Networks," Proc. ACM MobiCom,pp.139-150,2001.
- [13] S.Megerian, F. Koushanfar, M.Potkonjak, and M.Srivastava, "Worst and Best-Case Coverage in Sensor Networks",IEEE Trans. Mobile Computing, vol.4, no.1,pp.84-92,Jan./Feb.2005.

- [14] R.Olfati-Sabar, "Distributed Tracking for Mobile Sensor Networks With Information-Driven Mobility," Proc. AM. Control Conf., July 2007.
- [15] A.Ahmed, M. Gani, and F.Yang "Decentralized Robust Kalman Filtering for Uncertain Stochastic Systems over Heterogeneous Sensor Networks," Signal Processing, vol. 88, no.8, pp.1919-1928, 2008.
- [16] H.Medeiros, J. Park, and A.Kak, "Distributed Object Tracking using a Cluster-Based Kalman Filter in Wireless Camera Networks," IEEE J. Selected Topics in Signal Processing, vol.2, no.4, pp.448-463, Aug.2008.
- [17] L.Hu and D.Evans, "Localization for Mobile Sensor Networks," Proc. ACM MobiCom, Sept.2004.
- [18] Dongbing Gu and Huosheng Hu "Distributed Minimax Filtering for Tracking and Flocking", IEEE int'l Conf. on Intellegent Robotics and Systems.pp.18-22.Oct 2010.

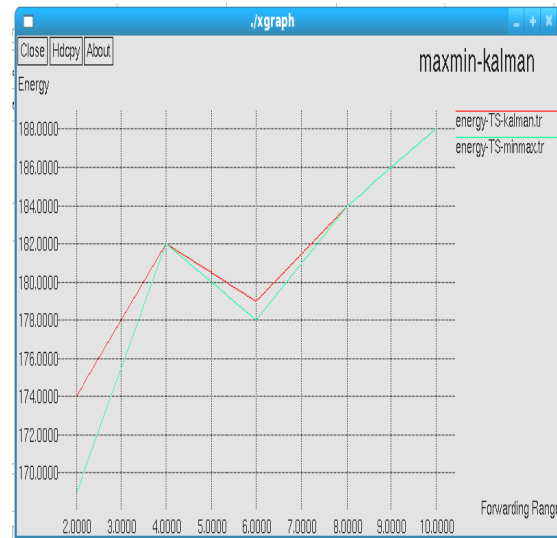


Fig. 2 : Energy consumption in WSANs using Kalman filtering and Minimax Filtering



Fig.1 : Average power consumption versus forwarding range in WSANs using Kalman filtering and Minimax Filtering.

# The Reduction of PAPR In OFDM By Using New Companding Transform

T. Nagaraju & K. Ramanaidu

JNTUACE, A.P., India

---

**Abstract** - The main drawback of the OFDM is its high peak to average power ratio(PAPR).There are several PAPR reduction techniques. Among the various PAPR reduction techniques, companding transform is attractive for its simplicity and effectiveness. This paper proposes a new companding algorithm. The proposed algorithm offers an improved bit error rate and minimized out-of-band interference while reducing PAPR effectively, compared with the others. Theoretical analysis and numerical simulation are presented.

**Key words** - OFDM, Companding, PAPR.

---

## I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is the multicarrier modulation technique, which supports high data rates. It is immune to the multipath fading [1].The applications of OFDM includes WiMAX, DVB/DAB and 4G wireless systems. Though, OFDM has several applications, it has one critical problem. The critical problem is its high peak-to-average power ratio (PAPR) [1]. High PAPR increases the complexity of analog-to-digital(A/D) and digital-to-analog (D/A) converters, and lowers the efficiency of power amplifiers. Several PAPR reduction techniques have been proposed Over the past decade, such as block coding, selective mapping (SLM) and tone reservation, just to name a few [2]. Among all these techniques the simplest solution is to clip the transmitted signal when its amplitude exceeds a desired threshold. Clipping is a highly nonlinear process, however. It produces significant out-of-band interference(OBI). A good remedy for the OBI is the so-called companding. The method was first proposed in [3], which employed the classical  $\mu$ -law transform and showed to be rather effective. Since then many different companding transforms with better performances have been published [4]-[7].This paper proposes and evaluates a new companding algorithm. The algorithm uses the special airy function and is able to offer an improved bit error rate (BER) and minimized OBI while reducing PAPR effectively. The paper is organized as follows. In the next section the PAPR problem in OFDM is briefly discussed. Section III presents the new algorithm and its theoretical analysis. Section IV shows the performance simulation. The last section describes the conclusion.

## II. PAPR IN OFDM

Let  $y(0), y(1), \dots, y(N-1)$  represent the data sequence to be transmitted in an OFDM symbol with  $N$  subcarriers. The baseband representation of the OFDM symbol is given by:

$$y(t) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} y(n) e^{j2\pi n t} \quad 0 \leq t \leq T$$

where  $T$  is the duration of the OFDM symbol. According to the central limit theorem, when  $N$  is large, both the real and imaginary parts of  $y(t)$  become Gaussian distributed, each with zero mean and a variance of  $E[y(t)^2]/2$ , and the amplitude of the OFDM symbol follows a Rayleigh distribution. Consequently it is possible that the maximum amplitude of OFDM signal may well exceed its average amplitude. Practical hardware (e.g. A/D and D/A converters, power amplifiers) has finite dynamic range; therefore the peak amplitude of OFDM signal must be limited. PAPR is mathematically defined as:

$$PAPR = 10 \log_{10} \frac{\max [|y(t)|^2]}{\frac{1}{T} \int_0^T |y(t)|^2 dt} \quad (\text{dB})$$

It is easy to see from the equation that PAPR reduction may be achieved by decreasing the numerator or increasing the denominator or both. The effectiveness of a PAPR reduction technique is measured by the complementary cumulative distribution function

(CCDF), which is the probability that PAPR exceeds some threshold, i.e.:

$$\text{CCDF} = \text{Probability}(\text{PAPR} > \rho_0),$$

where  $\rho_0$  is the threshold

### III. NEW COMPANDING ALGORITHM

OBI is the spectral leakage into alien channels. Quantification of the OBI caused by companding requires the knowledge of the power spectral density (PSD) of the companded signal. Unfortunately analytical expression of the PSD is in general mathematically intractable, because of the nonlinear companding transform involved. Here we take an alternative approach to estimate the OBI. Let  $f(y)$  be a nonlinear companding function, and  $y(t) = \sin(\omega t)$  be the input to the compander. The companded signal  $z(t)$  is:

$$z(t) = f[y(t)] = f[\sin(\omega t)].$$

Since  $z(t)$  is a periodic function with the same period as  $y(t)$ ,  $z(t)$  can then be expanded into the following Fourier series:

$$z(t) = \sum_{k=-\infty}^{+\infty} c(k) e^{jk\omega t}$$

where the coefficients  $c(k)$  is calculated as:

$$c(k) = c(-k) \\ = \frac{1}{T} \int_0^T z(t) e^{-jk\omega t} dt \quad T = \frac{2\pi}{\omega}$$

*Companding introduces minimum amount of OBI if the companding function  $f(x)$  is infinitely differentiable.* The functions that meet the above condition are the smooth functions. We now propose a new companding algorithm using a smooth function, namely the airy special function. The companding function is as follows:

$$f(y) = \beta \cdot \text{sign}(y) \cdot [\text{airy}(0) - \text{airy}(\alpha \cdot |y|)]$$

Where  $\text{airy}(\cdot)$  is the airy function of the first kind.  $\alpha$  is the parameter that controls the degree of companding (and ultimately PAPR).  $\beta$  is the factor adjusting the average output power of the compander to the same level as the average input power:

$$\beta = \sqrt{\frac{E[|y|^2]}{E[|\text{airy}(0) - \text{airy}(\alpha \cdot |y|)|^2]}}$$

where  $E[\cdot]$  denotes the expectation. The decomanding function is the inverse of  $f(x)$ :

$$f^{-1}(y) = \frac{1}{\alpha} \cdot \text{sign}(y) \cdot \text{airy}^{-1} \left[ \text{airy}(0) - \frac{|y|}{\beta} \right]$$

Where the superscript  $-1$  represents the inverse operation. Notice that the input to the decompander is a quantized signal with finite set of values. We can therefore numerically pre-compute  $f^{-1}(x)$  and use table look-up to perform the decomanding in practice. Next we examine the BER performance of the algorithm. Let  $z(t)$  denote the output signal of the compander,  $w(t)$  the white Gaussian noise. The received signal can be expressed as:

$$s(t) = z(t) + w(t).$$

The decomanded signal  $\hat{y}(t)$  simply is:

$$\hat{y}(t) = f^{-1}[s(t)] = f^{-1}[z(t) + w(t)]$$

### IV. PERFORMANCE SIMULATION

The OFDM system used in the simulation consists of 64 QPSK-modulated data points. The size of the FFT/IFFT is 256, meaning a  $4 \times$  oversampling. Given the compander input power of 3dBm, the parameter  $\alpha$  in the companding function is chosen to be 30. Consequently about 19.6 percent of  $s(t)$  is within the noise suppression range of the decomanding function. Two other popular companding algorithms, namely the  $\mu$ -law companding [3] and the exponential companding [5], are also included in the simulation for the purpose of performance comparison. The simulated PSD of the companded signals is illustrated in Fig. 2. The proposed algorithm produces OBI almost 3dB lower than the exponential algorithm, 10dB lower than the  $\mu$ -law. The result is in line with our expectation. The  $\mu$ -law function has a singularity in its second order derivative at  $x=0$  and therefore is expected to have the strongest OBI. Fig. 3 depicts the CCDF of the three companding schemes. The new algorithm is roughly 1.5dB inferior to the exponential, but surpasses the  $\mu$ -law by 2dB. The BER vs. SNR is plotted in Fig. 4. Our algorithm outperforms the other two. To reach a BER of  $10^{-3}$ , for example, the required SNR are 8.9dB, 10.4dB and 11.7dB respectively for the proposed, the exponential and the  $\mu$ -law companding schemes, implying a 1.5dB and 2.8dB improvement with the new algorithm. The amount of improvement increases as SNR becomes higher. One more observation from the simulation is: unlike the exponential companding whose performance is found almost unchanged under different degrees of companding, the new algorithm is flexible in adjusting

its specifications simply by changing the value of  $\alpha$  in the companding function.

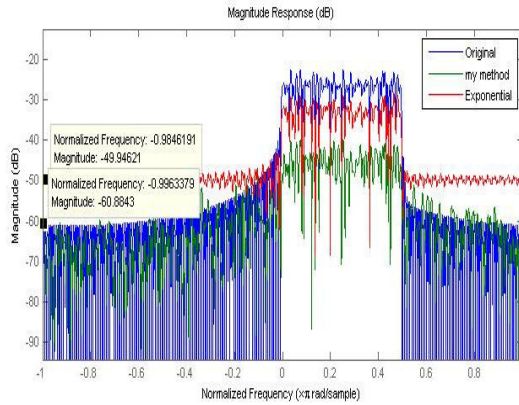


Fig.1 : Power spectral density of original and companded signals

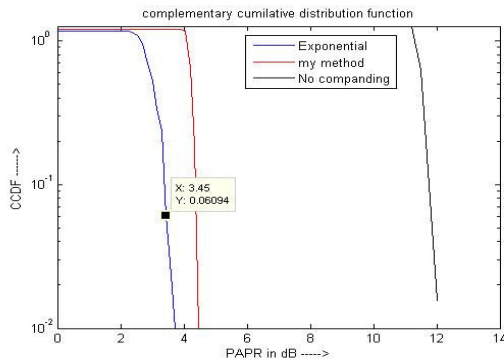


Fig. 2 : Complementary cumulative distribution of original and companded signals.

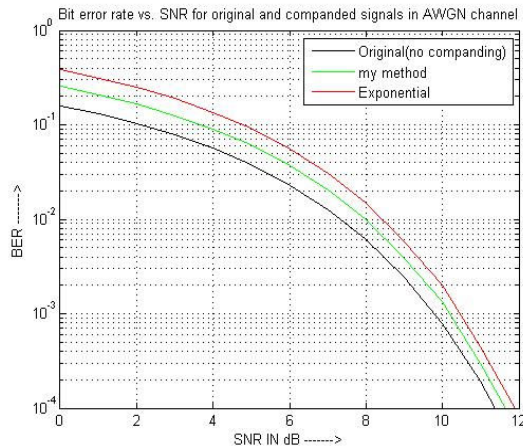


Fig. 3 : Bit error rate vs. SNR for original and companded signals in AWGN channel.

## V. CONCLUSION

In this paper, we have proposed a new companding algorithm. Both theoretical analysis and computer simulation show that the algorithm offers improved performance in terms of BER and OBI while reducing PAPR effectively.

## ACKNOWLEDGMENT

The author would like to thank the anonymous reviewers for their valuable comments and suggestions.

## REFERENCES

- [1] R. van Nee and R. Prasad, OFDM for Wireless Multimedia Communications. Boston, MA: Artech House, 2000.
- [2] S. H. Han and J. H. Lee, "An Overview of peak-to-average power ratio reduction techniques for multicarrier transmission," IEEE Wireless Commun., vol. 12, pp. 56-65, Apr. 2005.
- [3] X. Wang, T. T. Tjhung, and C. S. Ng, "Reduction of peak-to-average power ratio of OFDM system using a companding technique," IEEE Trans. Broadcast., vol. 45, no. 3, pp. 303-307, Sept. 1999.
- [4] T. Jiang and G. Zhu, "Nonlinear companding transform for reducing peak-to-average power ratio of OFDM signals," IEEE Trans. Broadcast., vol. 50, no. 3, pp. 342-346, Sept. 2004.
- [5] T. Jiang, Y. Yang, and Y. Song, "Exponential companding technique for PAPR reduction in OFDM systems," IEEE Trans. Broadcast., vol. 51, no. 2, pp. 244-248, June 2005.
- [6] D. Lowe and X. Huang, "Optimal adaptive hyperbolic companding for OFDM," in Proc. IEEE Second Intl Conf. Wireless Broadband and Ultra Wideband Commun., pp. 24-29, Aug. 2004.
- [7] T. Jiang and Y. Wu, "An overview: peak-to-average power ratio reduction techniques for OFDM signals," IEEE Trans. Broadcast., vol. 54, no. 2, pp. 257-268, June 2008.
- [8] I. N. Bronshtein, K. A. Semendyayev, G. Musiol, and H. Muehlig, Handbook of Mathematics, 5th ed. New York: Springer, 2007, p. 422.



# A NOVEL APPROACH TO CREATE RELIABLE STORAGE POOL

Leena Muddi<sup>1</sup>, Karan Tilak Kumar<sup>2</sup>, Taranisen Mohanta<sup>3</sup> & M.Jaya Bharata Reddy<sup>4</sup>

<sup>1,2 & 3</sup> Firmware Engineer, HP, Bangalore, India

<sup>4</sup>EEE Department, NIT, Trichy, TN, India

---

**Abstract** - In recent day's popularity of the Internet and storage has increased in multiple folds. For today's world minimum requirement for storage is reliability and cost efficiency.

The backbone of the mass storage is the hard disks, which finally preserves the data. So considering the cost, performance and availability, there are different kinds of hard disks available. However, the reliability and data availability from the hard disk is a crucial aspect. This paper describes an efficient and novel approach to select the hard disks from the bunch/pool of hard disks of different types so that the data will be preserved safely with high reliability. The hard disk vendors provide the failure reporting mechanism S.M.A.R.T (Self-Monitoring, Analysis, and Reporting Technology). In this approach, the essential/crucial S.M.A.R.T parameters are chosen, which potentially make the drive failure or performance impact. For each drive based on S.M.A.R.T parameters calculate and assign an overall score to the physical disks. The selection of the disks will be based on the weight of the value.

The approach can be used in the (DAS) Direct Attached Storage, AS (Network attached Storage) or any storage where the group of the drives will be used. Essentially different RAID levels can be created based on the drive reliability.

**Keywords:** Hard disk, S.M.A.R.T (Self-Monitoring, Analysis, and Reporting Technology), Direct Attached Storage(DAS), Network attached Storage(AS)

---

## I. INTRODUCTION

There are different kinds of storage devices available in the market based on the business need such as Solid State Device (SSD), SATA drives, SCSI drives, SAS drives, etc. Be it DAS (Direct Attached Storage), NAS (Network Attached Storage) or SAN (Storage Area Network) in all these mentioned cases the hard disks are the essential part and parcel of storage. For different applications and business needs the hard disks are used for a different purpose. Even though the prices of the hard disks are reduced, and the total cost per storage is less now as compared to early days. Tranquil, the reliability of the data is important and the hard disk reliability is very crucial. So the hard disk vendors provide the SMART (Self-Monitoring, Analysis, and Reporting Technology) parameters for the hard disks so that user will take appropriate action to preserve the data before the drive fails.

There are different SMART parameters available for the hard disk drives [1] which help in determining the quality of the disk. There are different kinds of methodology available to predict the failure of the FC disks, SCSI disks, SATA, and FATA, SAS disks etc. et al [2, 4, and 5]. In the Storage System, for failure

protection, different storage pools are created using the hard disks. The RAID level selection is done using the multiple devices/drives. The reliability and the data preservation depend on the health of the hard disks, which are part of the RAID group or the storage pool.

In this context, it is highly desirable to have a simple and accurate algorithm to choose the right kind of hard disks. Although there are researches and different attempts made to predict the different hard disk's failure prediction model [6, 7, 9] an efficient, simple and improved model is highly desirable.

## II. SMART PARAMETERS FOR HARD DISKS

The hard disk vendors support several SMART parameters. In this paper, we have selected the major SMART parameters, which influences the reliability of the disks as it becomes old. The important parameter, which selected as follows, and we can further add these variables to have the more accurate selection. These are Unrecoverable Read Errors, ECC-Corrected Read Errors, Unrecoverable Write Errors, Not Ready Failures, Seek Errors, Reallocated Sectors and Current Pending Sector Count and so on.

### III. RELIABLE STORAGE POOL OR RAID CREATION

The main goal is to have a group of disks, which will be the more reliable and thereafter selection of these hard disks to create the storage pool and RAID groups.

#### A. Storage Pool creation in the model

As already mentioned, our solution provides a way to consider the quality and performance of a disk while creating the logical units in the storage system. The solution provides an algorithm to create a logical unit based on the performance/reliability requirement of a user. Based on the user requirements, the algorithm selects the best disks for the logical unit creation. The

user inputs are basically the reliability and performance requirements of the application that will be run on the logical unit (refer the Table 2 given below for the user input). Effective usage of disks is possible by using this algorithm. The algorithm suggests to use S.M.A.R.T parameters to evaluate the best disks in the storage system. Most of the disks today, support smart parameters. However, they do not implement all the parameters. Hence, a limited set of S.M.A.R.T parameters are used to evaluate the disks. The Disk group will have all the Parameters that affect the Reliability of disk. Figure 1 shows the Classification of SMART parameters for reliability.

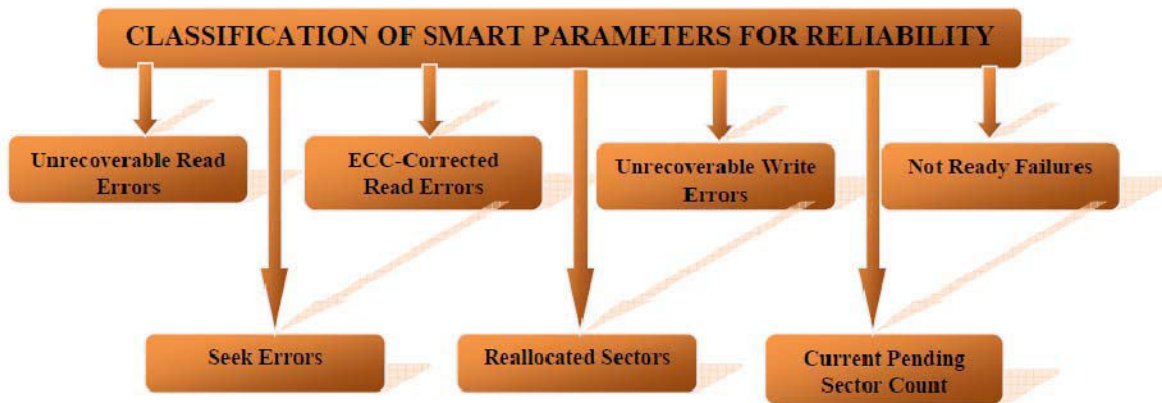


Figure 1: Classification of SMART parameters for reliability

#### B. Application of Storage Pool creation using SMART parameters

Using the SMART parameters the algorithm provides a rating to each disk in the group. Using this rating, the algorithm selects the best disks for the user requirement. The algorithm works in the following manner.

Step 1: For each S.M.A.R.T parameter, get the value from all disks.

Step 2: For this S.M.A.R.T parameter (SRm) calculates the average value (Avg).

Step 3: For each disk find the deviation.

$$Dn = SRm - Avg.$$

This step 2 and 3 will later help in determining which disk has a better ranking in case of a situation where the sum of all the parameters gives the same sum

value, but each of them has different individual S.M.A.R.T parameters.

Step 4: Repeat Step 2 through Step 3 for all the S.M.A.R.T Parameters.

Step5: For the Disk group, sum the RDn<sub>pm</sub> (Deviation from the Average value) for all the disks.

$$SRdn = Dnr1 + Dnr2 + \dots + Dnrm.$$

Step 6: Depending on the sum SRdn for each drive, now assign rank accordingly for each disk

The disk with the lowest SRdn value is the best and ranked no. 1 (R1)

From the final ranking select the top most disks and using these disks for the logical unit creation. Figure 2 shows the flowchart for the entire process.

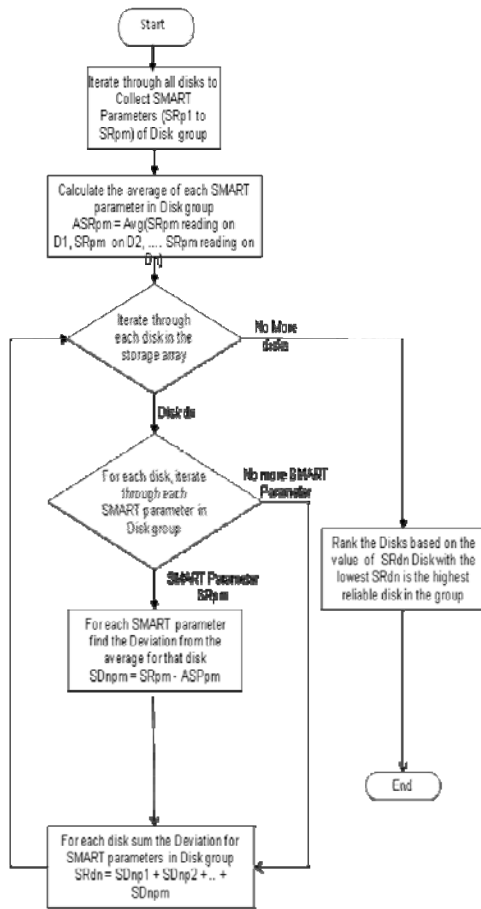


Figure 2: Algorithm for finding the reliable disk using SMART parameter

#### IV. RESULTS AND DISCUSSION

The sample data is collected using the HP Arrays, which can have a combination of SATA and SCSI drives, connected to a Windows 2003 host. Specific SCSI commands are used to obtain the Smart Parameter data.

Multiple disks were inserted in the JBOD enclosures. Using the SCSI pass-through collected the SMART value of the each disk before creating the storage pool. The above said algorithm was used to show the status/rank of the each disk.

The Table 1 shows different SMART parameters for Drives D1 – D10 and the average for each SMART parameter and the Table 2 shows the Deviation of Each parameter from its above calculated Average. Figure 3 showing the Smart Parameter values for Drives D1-D10 and Figure 4 shows the R1-R9 Indicates the Rank of each Disk assigned based on the sum SRdn for each drive, and we can see that Disk with the Rank R1 is the most Reliable Disk.

Disk/parameters	Unrecoverable Read Errors	ECC-Corrected Read Errors	Unrecoverable Write Errors	Seek Errors	Reallocated Sectors	Current Pending Sector Count	Not Ready Failures
D1	5	0	0	2	5	1	2
D2	3	0	0	0	3	0	0
D3	0	0	0	1	5	0	1
D5	0	1	1	3	1	0	1
D6	0	1	0	2	10	3	0
D6	1	1	4	5	0	0	0
D7	0	0	0	0	0	1	4
D8	0	0	0	0	0	0	0
D9	0	0	3	7	2	0	0
D10	0	0	4	2	3	0	0
Average	0.9	0.3	1.2	2.2	2.9	0.5	0.8

Table 1 Average for each smart parameters of the Drives D1-D10



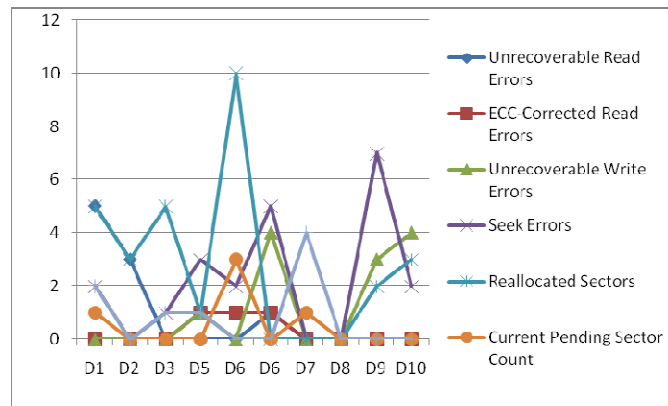
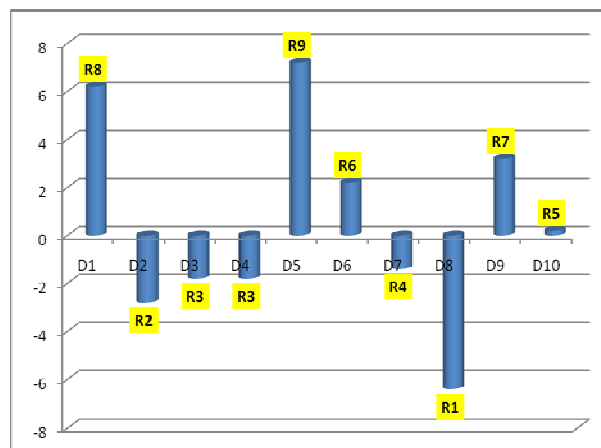


Figure 3 : Showing the Smart Parameter values for Drives D1- D10

Disk/p aramet ers	Unrecove rable Read Errors	ECC-Cor rected Read Errors	Unrecove rable Write Errors	Seek Errors	Reallocate d Sectors	Current Pending Sector Count	Not Ready Failures	Sum of the deviation from Average for each disk	RANK
D1	4.1	-0.3	-1.2	-0.2	2.1	0.5	1.2	6.2	8
D2	2.1	-0.3	-1.2	-2.2	0.1	-0.5	-0.8	-2.8	2
D3	-0.9	-0.3	-1.2	-1.2	2.1	-0.5	0.2	-1.8	3
D4	-0.9	0.7	-0.2	0.8	-1.9	-0.5	0.2	-1.8	3
D5	-0.9	0.7	-1.2	-0.2	7.1	2.5	-0.8	7.2	9
D6	0.1	0.7	2.8	2.8	-2.9	-0.5	-0.8	2.2	6
D7	-0.9	-0.3	1.2	-2.2	-2.9	0.5	3.2	-1.4	4
D8	-0.9	-0.3	1.2	-2.2	-2.9	-0.5	-0.8	-6.4	1
D9	-0.9	-0.3	1.8	4.8	-0.9	-0.5	-0.8	3.2	7
D10	-0.9	-0.3	2.8	-0.2	0.1	-0.5	-0.8	0.2	5

Table 2. Deviation of Each parameter from its above calculated Average

Figure 4: R1-R9 Indicates the Rank of each Disk assigned based on the sum SRdn for each drive.  
Disk with the Rank R1 is the most Reliable Disk

#### .V. CONCLUSION

This paper corroborates an innovative approach for creating the storage pool using the different SMART values, which play a major role to influence the reliability of the disk drives. The use cases are primarily before creating the different raid levels on the DAS to choose the most for reliable disks. This will be helpful for the Storage System that has a combination of old and relatively new disks.

#### REFERENCES

- [1] [http:// en.wikipedia.org/ wiki/S.M.A.R.T](http://en.wikipedia.org/wiki/S.M.A.R.T), the different SMART parameters.
- [2] M. H. Hassoun. Fundamentals of Artificial Neural Networks. MIT Press, Cambridge, 1995.
- [3] E. Anderson. Simple table-based modeling of storage devices. Technical Report HPL-SSP-2001-04, HP Laboratories, Jul 2001. <http://www.hpl.hp.com/SSP/papers/>.
- [4] Mengzhi Wang, Kinman Au, Anastassia Ailamaki, Anthony Brockwell, Christos Faloutsos, Gregory R. Ganger: Storage Device Performance Prediction with CART Models. MASCOTS 2004: 588-595
- [5] D. W. Marquardt. An Algorithms for the Least-Squares Estimation of Nonlinear Parameters. SIAM Journal of Applied Mathematics, 11(2):431–441, June 1963.
- [6] C. Ruemmler and J. Wilkes. An introduction to disk drive modeling. IEEE Computer, 27(3):17–28, 1994.
- [7] J. Schindler and G. R. Ganger. Automated Disk Drive Characterization. In Proceedings of the Sigmetrics 2000, page 109–126. ACM Press, June 2000.
- [8] Elizabeth Shriver, Arif Merchant, and John Wilkes. An analytic behavior model for disk drives with readahead caches and request reordering. International Conference on Measurement and Modeling of Computer Systems (Madison, WI., 22-26 June 1998). Published as Perform. Eval. Rev., 26(1):182–191. ACM, June 1998.
- [9] Mohanta Taranisen, Ananthamurthy Srikanth . A method to predict Hard disk Failures using SMART monitored parameters. Recent Developments in National Seminar on Devices, Circuits and Communication (NASDEC2 - 06). Page 243-246, November-2006
- [10] Jon G. Elerath and Sandeep Shah. Disk drive reliability case study: Dependence upon fly-height and quantity of heads. In Proceedings of the Annual Symposium on Reliability and Maintainability, January 2003.

□□□

# High Speed Parallel Architecture For Cyclic Convolution Based On FNT

P. Janardhan<sup>1</sup>, G. Mamatha<sup>2</sup> & P. Pushpalatha<sup>3</sup>

<sup>1&2</sup> Department ECE, JNTU, ANANTAPUR, India

<sup>3</sup> Department ECE, AITAM, TEKKALI, India

S

---

**Abstract** - This paper presents a high speed parallel architecture for cyclic convolution based on Fermat Number Transform (FNT) in the diminished-1 number system. A code conversion method without addition (CCWA) and a butterfly operation method without addition (BOWA) are proposed to perform the FNT and its inverse (IFNT) except their final stages in the convolution. The point wise multiplication in the convolution is accomplished by modulo  $2^n+1$  partial product multipliers (MPPM) and output partial products which are inputs to the IFNT. Thus modulo  $2^n+1$  carry propagation additions are avoided in the FNT and the IFNT except their final stages and the modulo  $2^n+1$  multiplier. The execution delay of the parallel architecture is reduced evidently due to the decrease of modulo  $2^n+1$  carry- propagation addition. Compared with the existing cyclic convolution architecture, the proposed one has better throughput performance and involves less hardware complexity. Synthesis results using 130nm CMOS technology demonstrate the superiority of the proposed architecture

**General Terms** - This paper studying parallel architecture design for various applications in VLSI Design, for high speed data transformation, using Fermat number transform, different designing method to provide high speed data transmission.

**Key words** - VLSI, DSP, FPGA, Implementation using Verilog and VHDL coding method, Tools ate Xilinx new version.

---

## I. INTRODUCTION

The cyclic convolution based on FFT is a widely used operation in signal processing, which needs to be performed in a complex domain even if both of the sequences to be performed would be real. Additionally, the dynamic range of the numbers varies widely so that one need to use floating point numbers to avoid scaling and quantization problems. Some architecture for efficient cyclic convolution has been developed to overcome the problems based on Number Theory Transform (NTT). They replace the complex domain with a finite field or a finite residue ring and can be defined by the FFT-like formula. All arithmetic operations are performed modulo  $m$  and the convolution results are exact without rounding errors.

When the modulus in NTT is a Fermat number ( $F_t=2^{2^t}+1$ ,  $t^{\text{th}}$  Fermat;  $t$  is an integer) the NTT. Turns into the Fermat Number Transform(FNT). The multiplication in the FNT and its inverse (IFNT) can be converted into bit shifts when the transform kernel is 2 or its integer power.

Though the modulus of the FNT has a strict relationship with its maximum transform, the cyclic convolution based on FNT is more attractive than the conventional method in some areas.

Most cyclic convolution architectures based on FNT are implemented for the operands in the diminished-1 representation. Thus the code conversion (CC) stage which converts the normal binary numbers into their diminished-1 representation is compulsory.

Other arithmetic operations described originally by Leibowitz includes modulo  $2^n+1$  negation, addition, subtraction, multiplication operations in the diminished-1 number system. These operations constitute the butterfly operation (BO) which is the most important element in the FNT. The CC and the BO are both mainly composed of modulo  $2^n+1$  adders of which the fastest one in the diminished-1 number system is proposed by Vergos so far. The fast modulo  $2^n+1$  adder involves the carry-propagation addition computation and is used in the recent FNT implementations.

In this paper, a code conversion method without addition (CCWA) and a butterfly operation method without addition (BOWA) which take full advantage of the carry-save adder are proposed to accomplish the cyclic convolution with the unity root 2 or its integer power. The modulo  $2^n+1$  partial product multiplier (MPPM) is used to accomplished the point wise multiplication so that the final carry-propagation addition of two partial product in the multiplier is avoided.

Thus the execution delay of the architecture is reduced evidently. Model estimations and experiment results show that the proposed architecture is faster than the existing one when the modulus of the FNT is no less than  $2^8+1$ . For wider modulus, the proposed parallel architecture leads to considerably faster hardware implementations than those presented in.

The FNT of a sequence of length  $N$

$$X(K) = \sum_{i=0}^{N-1} x(i)^{ik}, k = 0, 1, \dots, N-1$$

where  $F_t=2^{2^t}+1$ , the  $t^{\text{th}}$  Fermat;  $N$  is a power of 2 and  $\alpha$  is an  $N^{\text{th}}$  root unit (i.e.  $\alpha^N \bmod F_t = 1$  and  $\alpha^M \bmod F_t \neq 1, 1 \leq M < N$ ). The notation  $\langle ik \rangle$  means  $ik$  modulo  $N$ .

The inverse FNT is given by

$$x_i = \frac{1}{N} \sum_{k=0}^{N-1} x_k \alpha_N^{-ik} \bmod F_t (i = 0, 1, \dots, N-1)$$

where  $1/N$  is an element in the finite field or ring of integer and satisfies the following condition:

$$(N \cdot 1/N) \bmod F_t = 1$$

Parameters  $\alpha$ ,  $F_t$ ,  $N$  must be chosen carefully and some conditions must be satisfied so that the FNT possesses the cyclic convolution property. In this project,  $\alpha=2$ ,  $F_t=2^{2^t}+1$  and  $N=2 \cdot 2^t$  where  $t$  is an integer.

## II. IMPORTANT OPERATIONS IN CYCLIC CONVOLUTION BASED ON FNT

Important operations of the cyclic convolution based on FNT with the unity root 2 include the CCWA, the BOWA and the MPPM. The CCWA and the BOWA both consist of novel modulo  $2n+1$  4-2 compressors mainly which are composed of the 4-2compressor introduced by Nagamatsu. The 4-2 compressor, the novel modulo  $2n+1$  4-2 compressor and the BOWA are shown in Fig. 1. In the figure, “ $X^*$ ” denotes the diminished-1 representation of  $X$ ,  $ix^n = x-1$ .

### 2.1 Code conversion without addition

The CC converts normal binary numbers (NBCs) into their diminished-1 representation. It is the first stage in the FNT. Delay and area of CC of a  $2n$ -bit NBC are no less than the ones of two  $n$ -bit carry propagation adders. To reduce the cost, we propose the CCWA that is performed by the modulo  $2n+1$  4-2 compressor.

Let  $A$  and  $B$  represent two operands whose widths are no more than  $2n$  bits. We define two new variables:

Let  $A$  and  $B$  represent two operands whose widths are no more than  $2n$  bits. We define two new variables:

$$A = 2^n A_H + A_L$$

$$B = 2^n B_H + B_L$$

And

$$M_0 = (2^n - 1) - A_H = \bar{A}_H$$

$$M_1 = (2^n - 1) - B_H = \bar{B}_H$$

$$M_2 = (2^n - 1) - B_L = \bar{B}_L$$

If the subsequent operation of CC is modulo  $2^n+1$  addition, assign  $A_L$ ,  $M_0$ ,  $B_L$  and  $M_1$  to  $I_0$ ,  $I_1$ ,  $I_2$ ,  $I_3$  in the modulo  $2^n+1$  4-2 compressor respectively.  $I_0$ ,  $I_1$ ,  $I_2$ ,  $I_3$  are defined as follows:

$$I_0 = I_{0(n-1)} I_{0(n-2)} \dots I_{01} I_{00}$$

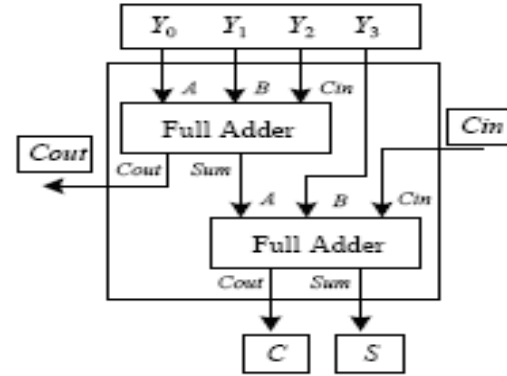
$$I_1 = I_{1(n-1)} I_{1(n-2)} \dots I_{11} I_{10}$$

$$I_2 = I_{2(n-1)} I_{2(n-2)} \dots I_{21} I_{20}$$

$$I_3 = I_{3(n-1)} I_{3(n-2)} \dots I_{31} I_{30}$$

We obtain the sum vector  $H_O^*$  and carry vector.

$H_1^*$  In the diminished-1 number system



4-2 Compressor

We obtain the sum vector  $H_O^*$  and carry vector  $H_1^*$  in the diminished-1 number system. The most significant bit of  $H_1^*$  is complemented and connected back to its least significant bit. That is to say

$$H_O^* = H_{0(n-1)} H_{0(n-2)} \dots H_{01} H_{00}$$

$$H_1^* = H_{1(n-2)} \dots H_{11} H_{10} H_{1(n-1)}$$

The result of modulo  $2^n+1$  addition of  $A^*$  and  $B^*$  is equal to the result of modulo  $2^n+1$  addition of  $H_O^*$   $H_1^*$  in this way,  $A$  and  $B$  are converted into their equivalent diminished-1 representations  $H_O^*$  and  $H_1^*$

Let  $|A^* + B^*|_{2^n+1}$ ,  $|\bar{A}^*|_{2^n+1}$ ,  $|A^* - B^*|_{2^n+1}$ , and  $+2^i|_{2^n+1}$  denote modulo  $2^n+1$  addition, negation, subtraction and multiplication by the power of 2

respectively which are proposed by Leibowitz originally. The CCWA for subsequent modulo  $2^{n+1}$  addition can be described as follows.

$$|A^* + B^*|_{2^{n+1}} = |A_L + M_0 + B_L + M_1|_{2^{n+1}} = |H_O^* + H_1^*|_{2^{n+1}}$$

If the subsequent operation is modulo  $2^{n+1}$  subtraction, we assign  $A_L$ ,  $M_0$ ,  $M_2$  and  $B_H$  to  $I_0$ ,  $I_1$ ,  $I_2$ ,  $I_3$  respectively. Then  $H_O^*$  and  $H_1^*$  in the modulo  $2^{n+1}$  4-2 compressor constitute the result of the CCWA. The conversion is described as follows:

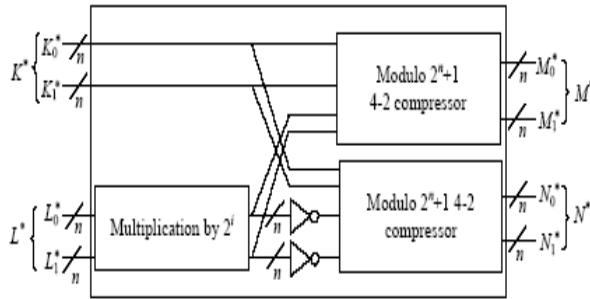
$$\begin{aligned} |A^* - B^*|_{2^{n+1}} &= |A - B|_{2^{n+1}} = \\ A + B|_{2^{n+1}} &= |A_L + M_0 + B_L + M_1|_{2^{n+1}} \\ &= |H_O^* + H_1^*|_{2^{n+1}} \end{aligned}$$

After CCWA, we obtain the result consisting of two diminished-1 numbers. The result also includes the information of modulo  $2^{n+1}$  addition or subtraction in the first stage of previous BO.

## 2.2 Butterfly operation without addition

After the CCWA, we obtain the results of modulo  $2n+1$  addition and subtraction in the diminished-1 representation. Each result consists of two diminished-1 values. The subsequent butterfly operation involves four operands. The proposed BOWA involves two Modulo  $2n+1$  4-2 compressors, a multiplier and some Inverters as shown in Fig. 1(c). The multiplication by an integer power of 2 in the diminished-1 number system in the BOWA is trivial and can be performed by left shifting the low-order  $n-i$  bits of the number by  $i$  bit positions then inverting and circulating the high-order  $i$  bits into the  $i$  least significant bit positions.

Thus the BOWA can be performed without the carry-propagation chain so as to reduce the delay and the area obviously.  $K^*$ ,  $L^*$ ,  $M^*$ ,  $N^*$  are corresponding to two inputs and two outputs of previous BO in the diminished-1 number system.

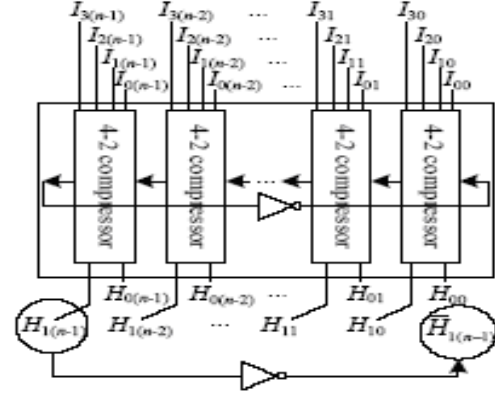


Butterfly operation without addition

## 2.3 Modulo $2^{n+1}$ partial product multiplier

For the modulo  $2^{n+1}$  multiplier proposed by

Efstathiou, there are  $n+3$  partial products that are derived by simple AND and NAND gates [11]. An FA-based Dadda tree that reduces the  $n+3$  partial products into two summands is followed. Then a modulo  $2^{n+1}$  adder for diminished -1 operands is employed to accept these two summands and produce the required product.



Modulo  $2^{n+1}$  4-2 Compressor

In the proposed parallel architecture for cyclic convolution based on FNT, the BOWA can accept four operands in the diminished-1 number system. Every point wise multiplication only needs to produce two partial products rather than one product. The operation can be accomplished by taking away the final modulo  $2^{n+1}$  adder of two partial products in the multiplier. Thus the final modulo  $2^{n+1}$  adder is omitted and the modulo  $2^{n+1}$  partial product multiplier is employed to save the delay and the area.

## III. PARALLEL ARCHITECTURE FOR CYCLIC CONVOLUTION

Based on the CCWA, the BOWA and the MPPM, We design the whole parallel architecture for the cyclic Convolution based on FNT as shown in Fig. 2. Includes the FNTs, the point wise multiplication and the IFNT mainly. FNTs of two input sequences  $\{a_i\}$  and  $\{b_i\}$  produce two sequences  $\{A_i\}$  and  $\{B_i\}$  ( $i=1,2,\dots,N-1$ ). Sequences  $\{A_i\}$  and  $\{B_i\}$  are sent to  $N$  MPPMs to accomplish the point wise multiplication and produce  $N$  pairs of partial products. Then the IFNT of the partial products are performed to produce the resulting sequence  $\{p_i\}$  of the cyclic convolution.

IFNT. Illustrative examples of the FNT and the IFNT are shown in Fig. 3 in the case the transform length is 16 and the modulus is  $2^8+1$ . Commutators in Fig. 3 are used to adjust the operand order of every stage of FNT and IFNT according to the radix-2 DIT algorithm.

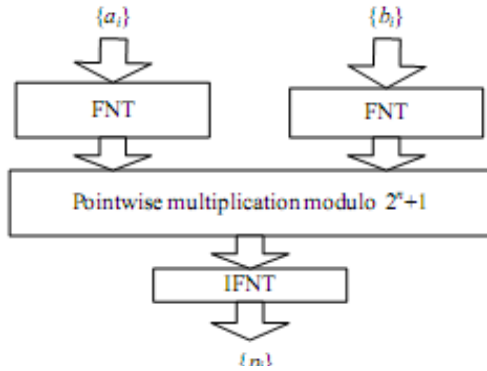


Fig. 2 : Parallel architecture for the cyclic convolution based on FNT

The efficient FNT structure involves  $\log_2 N + 1$  stages of operations. The original operands are converted into the diminished-1 representation in the CCWA stage, containing the information of modulo  $2^n + 1$  addition or subtraction in the first butterfly operation stage of the previous FNT structure. Then the results are sent to the next stage of BOWA. After  $\log_2 N - 1$  stages of BOWAs, the results composed of two diminished-1 operands are obtained. The final stage of FNT consists of modulo  $2^n + 1$  carry-propagation adders which are used to evaluate the final results in the diminished-1 representation.

The CCWA stage, the BOWA stage and the modulo  $2^n + 1$  addition stage in the FNT involves  $N/2$  couples of code conversions including the information of modulo  $2^n + 1$  addition and subtraction,  $N/2$  butterfly operations and  $N/2$  couple of modulo  $2^n + 1$  additions respectively.

From the definition of FNT and IFNT in section 2, the only difference between the FNT and the IFNT is the normalization factor  $1/N$  and the sign of the phase factor  $\alpha N$ . If ignoring the normalization factor  $1/N$ , the above formula is the same as that given in the FNT except that all transform coefficients  $\alpha_N^{(k)}$  used for the

FNT need to be replaced by  $\alpha_N^{-ik}$  for the IFNT Computation. The proposed FNT structure can be used to complete the IFNT as well with little modification as shown in Fig. 3(b). After the IFNT of  $N$ -point bit-reversed input data, the interim results are multiplied by  $1/N$  in the finite field or ring. Then  $x[j]$  and  $x[j + N/2]$  ( $j = 1, 2, \dots, N/2 - 1$ ) exchange their positions to produce the final results of the IFNT in natural order. Our architecture for the cyclic convolution gives a good speed performance without requiring a complicated control. Furthermore, it is very suitable for implementation of the overlap-save and overlap-add

techniques which are used to reduce a long linear convolution to a series of short cyclic convolutions.

#### IV. COMPARISON AND RESULTS

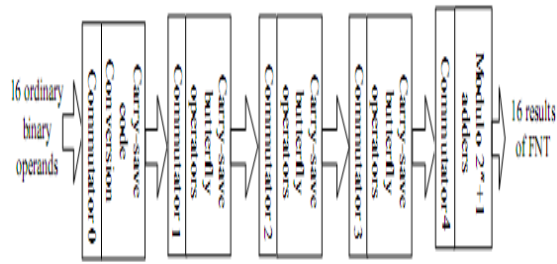
In this section, we compare the proposed parallel architecture for the cyclic convolution against that introduced by Conway [3]. The modulo  $2^n + 1$  addition for the diminished-1 number system is the crucial operation which contains a standard  $n$ -bit carry-propagation computation such as a parallel-prefix adder with a carry-logic block and a zero indicator of the diminished-1 operand to determine whether to perform subsequent operations. It produces the longest execution delay and requires large area in the previous solution. The proposed CCWA and BOWA overcome the disadvantage of the carry-propagation adder and don't require a zero indicator. Thus our architecture is faster and more efficient than the existing one.

For quantitative comparison, we employ the "unit gate model" presented in, which was also used in this model assumes that each two-input gate excluding XOR is equivalent to one elementary gate for both area and delay. An XOR gate counts for two gates for both area and delay.

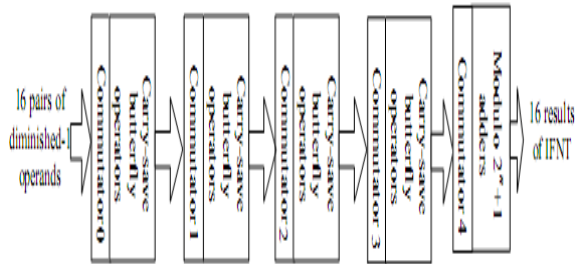
Thus, a full adder has an area of seven gates and a delay of four gates. This model does not involve the cost of buffering and routing, but achieve a reasonable accuracy for the purpose of comparison. The delay and the area estimations of modulo  $2^n + 1$  adder and modulo  $2^n + 1$  multiplier in the cyclic convolution are given in table.

To obtain more accurate results, we describe the proposed parallel cyclic convolution in verilog for  $F_t = 2^8 + 1, 2^{16} + 1, 2^{32} + 1$ . The validated Verilog code is synthesized using a  $0.13\text{-}\mu\text{m}$  CMOS standard cells library in the worst operating condition by the Synopsis Design Compiler. The units of area and delay are  $\mu\text{m}^2$  and  $\text{ns}$  respectively. Each design was recursively optimized for speed until the EDA software can't provide a faster design. The results for the fastest derived implementation are listed.

Table. Indicate that for values of  $F_t \geq 2^8 + 1$ , the proposed architecture comprising the CCWA and the BOWA require less delay and area than the previous one. The former results in a 12.6% reduction in area and a 26% reduction in delay respectively compared with the latter in the case  $F_t$  is  $2^{32} + 1$  and the transform length is 64. Moreover, our algorithm will be more and more advantageous with the growth of modulus width.



(a) Parallel FNT structure



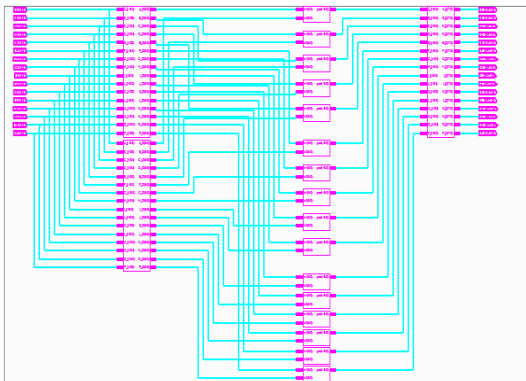
(b) Parallel IFNT structure

 Fig. 3 : Structures for FNT and IFNT ( $F_t=2^8+1$ )

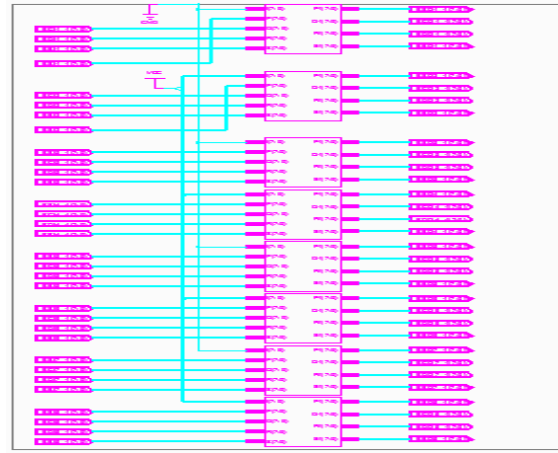
Table: Area and delay results of cyclic convolution based on FNT

$F_t$	Area ( $\mu m^2$ )		Delay (ns)	
	This paper	[3]	This paper	[3]
$2^8+1$	$3.5 \times 10^5$	$3.9 \times 10^5$	8.9	9.9
$2^{16}+1$	$1.86 \times 10^6$	$2.05 \times 10^6$	11.6	14.4
$2^{32}+1$	$1.08 \times 10^7$	$1.24 \times 10^7$	15.1	20.4

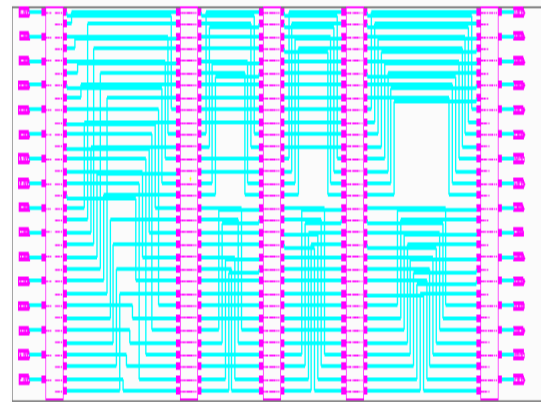
## RESULTS



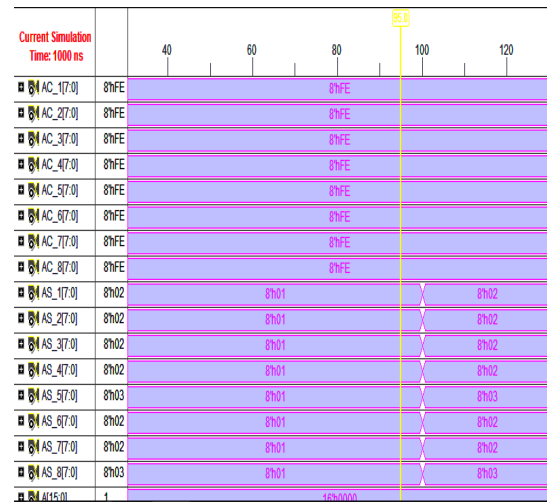
FNT parallel Architecture



RTL Schematic of Butterfly Operation



FNT parallel Architecture over all Schematic



Code conversion simulation result



Current Simulation Time: 1000 ns		200	400	600	750.0	800	1000
int op							
Q[A_Q(8)]	9..	136	1	2	3	4	5
Q[B_Q(8)]	9..	39	1	226	234	41	35
Q[C_Q(8)]	9..	90	1	254	81	45	45
Q[D_Q(8)]	9..	166	1	129	127	131	123
Q[E_Q(8)]	9..	120	1	17	1	0	16
Q[F_Q(8)]	9..	225	1	3	132	88	36
Q[G_Q(8)]	9..	91	1	194	198	214	21
Q[H_Q(8)]	9..	100	1	250	25	21	15
Q[I_Q(8)]	9..	249	1	0	256	0	25
Q[J_Q(8)]	9..	141	1	33	25	89	81
Q[K_Q(8)]	9..	150	1	5	198	182	17
Q[L_Q(8)]	9..	16	1	130	132	136	14
Q[M_Q(8)]	9..	121	1	242	1	0	24
Q[N_Q(8)]	9..	75	1	256	127	63	95
Q[O_Q(8)]	9..	151	1	65	81	77	13
Q[P_Q(8)]	9..	202	1	9	234	230	10

FNT Top module

Current Simulation Time: 1000 ns		200	400	600	750.0	800	1000
int op							
Q[A_I(15)]	1..	18496	1	4	9	16	25
Q[B_I(15)]	1..	62001	1	0	1	0	1
Q[C_I(15)]	1..	14400	1	289	1	0	256
Q[D_I(15)]	1..	14641	1	58564	1	0	58041
Q[E_I(15)]	1..	8100	1	64516	3721	2025	2401
Q[F_I(15)]	1..	22500	1	25	38204	33124	31844
Q[G_I(15)]	1..	8281	1	37636	38204	45796	44
Q[H_I(15)]	1..	22801	1	4225	3721	5929	16
Q[I_I(15)]	1..	1521	1	51076	54756	1681	1521
Q[J_I(15)]	1..	19801	1	1089	825	7921	8281
Q[K_I(15)]	1..	50625	1	9	17424	4624	1296
Q[L_I(15)]	1..	5625	1	1	16129	3969	9025
Q[M_I(15)]	1..	27556	1	16641	16129	17161	15129
Q[N_I(15)]	1..	256	1	16900	17424	18496	20736
Q[O_I(15)]	1..	10000	1	62500	825	441	22940
Q[P_I(15)]	1..	40804	1	81	54756	52900	10201

IFNT Top module

## V. CONCLUSIONS

A novel parallel architecture for the cyclic convolution based on FNT is proposed in the case the principle root of unity is equal to 2 or its integer power. The FNT and the IFNT are accomplished by the CCWA and the BOWA mainly. The point wise multiplication is performed by the modulo  $2^n+1$  partial product multiplier. Thus there are very little modulo  $2^n+1$  carry-propagation addition compared to the existing cyclic convolution architecture. A theoretical model was applied to access the efficiency independently of the target technology. VLSI implementations using a 0.13 um standard cell library show the proposed parallel architecture can attain lower area and delay than that of the existing solution when the modulus is no less than  $2^8+1$ .



## ACKNOWLEDGMENT

It is a great pleasure to express my gratitude and sincere thanks to my external guide Mr.G.SANATH, Senior Designer Engineer, CITD for his valuable guidance, suggestions and constant encouragement through out the project work.

## REFERENCES

- [1] C. Cheng, K.K. Parhi, "Hardware efficient fast DCT based on novel cyclic convolution structures", IEEE Trans. Signal processing, 2006, 54(11), pp. 4419-4434
- [2] H.C. Chen, J.I. Guo, T.S. Chang, et al., "A memory-efficient realization of cyclic convolution and its application to discrete cosine transform", IEEE Trans.
- [3] R. Conway, "Modified Overlap Technique Using Fermat and Mersenne Transforms", IEEE Trans. Circuits and Systems II: Express Briefs, 2006, 53(8), pp.632 – 636
- [4] A. B. O'Donnell, C. J. Bleakley, "Area efficient fault tolerant convolution using RRNS with NTTs and WSCA", Electronics Letters, 2008, 44(10), pp.648-649
- [5] H. H. Alaeddine, E. H. Baghious and G. Madre et al., "Realization of multi-delay filter using Fermat number transforms", IEICE Trans. Fundamentals, 2008, E91A(9), pp. 2571-2577
- [6] N. S. Rubanov, E. I. Bovbel, P. D. Kukharchik, V. J. Bodrov, "Modified number theoretic transform over the direct sum of finite fields to compute the linear convolution", IEEE Trans. Signal Processing, 1998, 46(3), pp. 813-817



# Development of Laboratory Scale Self-Balancing Prototype Robot

M.K.A. Ahamed Khan, I. Elamvazuthi, P. Vasant, T. Ganesan and R.A. Amin

Universiti Industri Selangor, Kampus Bestari Jaya, 45600 Berjuntai Bestari, Malaysia  
Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak Darul Ridzuan, Malaysia

**Abstract** - This paper presents the development of a laboratory scale self-balancing prototype robot. It addresses the balancing mechanism issue based on the theories of inverted pendulum and rotational inertia. It comprises of two major sub-systems, i.e., mechanical and control sub-systems. The implementation of the mechanical and control sub-systems is described in detail.

**Keywords**- self-balancing robot, mechanical, control, inverted pendulum

## I. INTRODUCTION

Over the last decade, the research on self-balancing robot has gained interest amongst the robotics enthusiasts as well as the robotics laboratory around the world [1]. There are considerable research efforts towards building a robust self-balancing robot [2-7]. The application of self-balancing robot is unlimited where it can serve the community and encounter the problems that many have raised in the industrial and service sectors. For example, a motorized wheelchair utilizing this technology would give the operator greater maneuverability and thus access to place most able-bodied people take for granted [8].

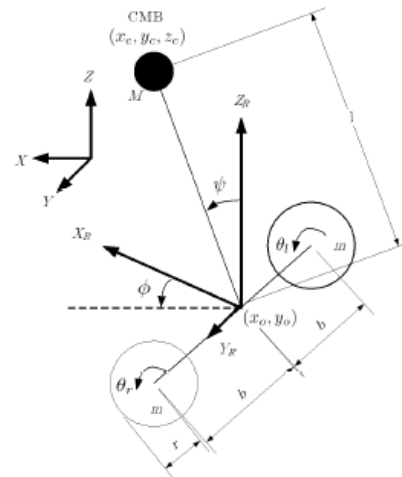
As compared to four-wheels trolley, having a platform with only two wheels that is able to balance itself and move around is advantageous as it possibly maneuvers better, move on either flat or sloppy surfaces, and to further enhances, apply for carrying limitless tasks, hence benefit the community. Yet, the main challenges to this possibility is how to have a two wheels robot that is able to stand upright and balance itself, furthermore carry object from one point to other point efficiently. This is where the significant of this project comes from which is to build a self-balancing robot [9].

The main objective for this project is to build a self-balancing robot. The robot is comprised platforms with two wheels and equipped with successful balancing mechanism. The self-balancing robot should be able to stand upright and balance itself, move from one point to other point as pre-programmed and carry an object from one point to the other.

This paper is organized as follows: Section II presents the self-balancing robot dynamics. Section III provides the system design. Section IV discusses the implementation details and finally, section V is the conclusions.

## II. SELF-BALANCING ROBOT DYNAMICS

The derivation of dynamics of self-balancing robot is given in the following based on Figure 1. For detail derivation, see [10].



The system parameters based on Figure 1 is given in Table 1.

Table 1 System Parameters

$r$	Radius of wheel
$m$	Mass of a wheel
$J_c$	Moment of inertia of body
$J_y$	Moment of inertia of the motor and the wheel about the y axis
$J_d$	Moment of inertia of the motor and the wheels about their diameter
$b_r, b_b, b_v, b_w$	Damping coefficients of the robot
$\theta$	Rotational angle of the wheels
$\psi$	Tilt angle of CMB
$\omega$	The angular velocity of the origin
$\phi$	Steering angle about the origin
$M$	Mass of the CMB
$l$	Distance between CMB and origin
$2b$	Distance between wheels
$b_m, J_m, R_a, L_a, k_m$	Motor parameters

Given the kinetic energy of the wheels and Centre Mass of Body (CMB), the damping energy, and potential energy, as  $K_\omega$ ,  $K_c$ ,  $\beta$ ,  $U_c$  respectively, we obtain,

$$K_\omega = \frac{1}{2}(mr^2 + J_y)(\dot{\theta}_r^2 + \dot{\theta}_l^2) + J_d \dot{\phi}^2 \quad (1)$$

$$K_c = \frac{1}{2}v_c^T M v_c + \frac{1}{2}\omega_c^T J_c \omega_c \quad (2)$$

$$\beta = \frac{1}{2}b_r(\dot{\theta}_r - \dot{\psi}_l)^2 + \frac{1}{2}b_l(\dot{\theta}_l - \dot{\psi}_l)^2 + \frac{1}{2}v_c^T b_v v_c + \frac{1}{2}\omega_c^T b_\omega \omega_c \quad (3)$$

$$U_c = Mgl \cos \psi \quad (4)$$

The robot is powered by two motors. The controlled variables of the model are the position and orientation of the mobile robot, while the control variables are the angular velocities of the left wheel and the right wheel.

### III. SYSTEM DESIGN

The laboratory prototype self-balancing robot is divided into two major sub-systems, i.e., mechanical and control sub-systems.

#### A. Mechanical Sub-system

The mechanical sub-system comprises of three round platforms arranged vertically, two wheels at the bottom of each side, with circuitry, servo motors and batteries attached to the structure as illustrated in Fig 2.

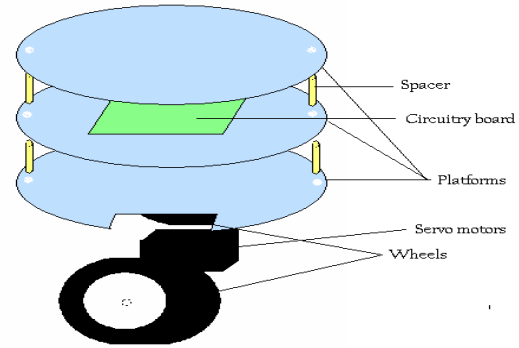


Figure 2. Mechanical sub-system

The base area for the mechanical sub-system is designed to be reasonably wide. Theoretically, the wider the base area is the more stable the structure would be. Even so, the structure for this is not required to be perfectly stable without the balancing mechanism. The structure needed to wobble or tilt at significant for the accelerometer, Memsic2125 to sense it easily, or else the balancing mechanism may not work as expected. At the same time, the structure should not be in the state of not very stable or else it would difficult for it to balance. Hence, the base is determined as 150mm in diameter in accordance to the dimension of the platforms, as shown in Figure 3. Note that, the size of the wheels, which is 50mm in diameter, is taken into consideration in determining the base area. Also note that, the material of the platforms is 2.5mm Perspex. The purpose of this choice is to keep the total weight of the structure at a range that the servo motors can operate at, which is 8 kg at maximum.

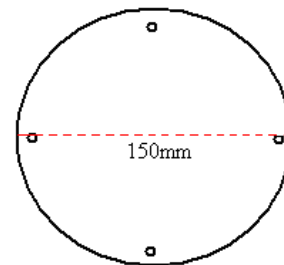


Figure 3. Dimension of the platform

Three round platforms are used in this project. The first platform which is the base platform is a round shape with two rectangle shapes cut out at each side, as shown in Figure 4. Those rectangles are meant for the wheels.

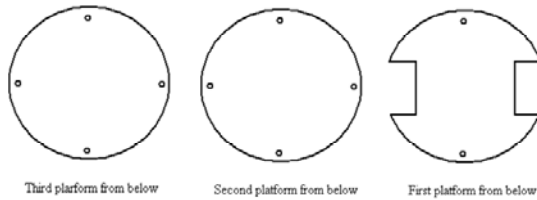


Figure 4. Platforms for structure

The dimension of the wheels for this project is chosen to be reasonably large also depends on its availability in the market. Hence the dimension of the wheels is 50mm diameter and 25mm wide. The wheels are Tamiya manufactured.

#### B. Control Sub-system

The microcontroller chip used is AtMega16, the developed circuit schematic is as shown in Figure 5. This circuitry comprises power, reset, clock, input, output connections and in-circuit programming.

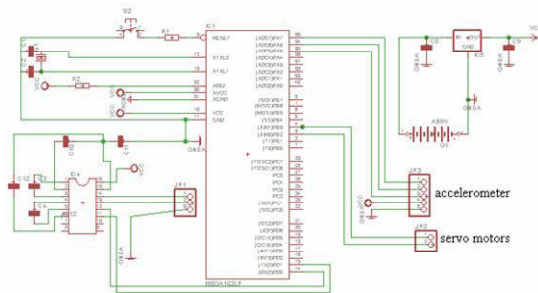


Figure 5. Schematic for circuitry

The purpose of power circuit is to regulate the battery's voltage from 12V to 5V as appropriate for this circuit. The reset circuit is to reset the operation of AtMega16, whilst, the clock circuit is to provide clock cycle for the AtMega16 so that it could operate faster. It should be noted that connecting to high frequency of crystal clock may not assure that the microcontroller chip works as desired, in a worse case may damage the chip. Therefore, a proper frequency of crystal clock has to be ascertained depending on the type of microcontroller chip. Accelerometer pin-outs are for the Memsic2125 connections and servo motors pin-outs are for servo motors connections. The purpose to have in-circuit programming is to eliminate the need to detach and reattach AtMega16 for programming. This would

also eliminate the possibility of damaging the pins of AtMega16.

The circuit design is transferred to PCB (printed circuit board) design using Eagle software. The design for the PCB is as shown in Figure 6.

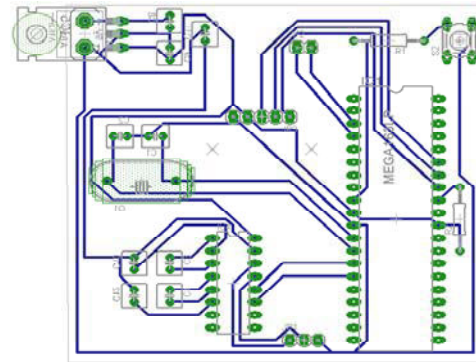


Figure 6. PCB layout for circuitry

#### IV. IMPLEMENTATION DETAILS

The mechanical structure of the robot is made of a combination of Aluminium and Perspex. The control subsystem of all the robots is comprised of a PIC Microcontroller Board, DC Brushless motor, brush motor, IR sensors, solenoids, and limit switches. The developed laboratory scale self-balancing prototype robot is shown in Figure 7.

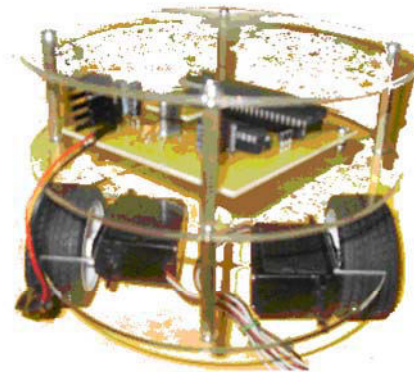


Figure 7. Self-Balancing Robot

Figure 8 shows the control sub-system. The main function of the control system is to generate electrical signals to effect motion based on input given by the control software. Considering the requirements of the self-balancing robot, it was decided that the simplest and most cost effective way of achieving the sequence

control was to use a AtMega16 microcontroller board which acts as the heart of the robot.

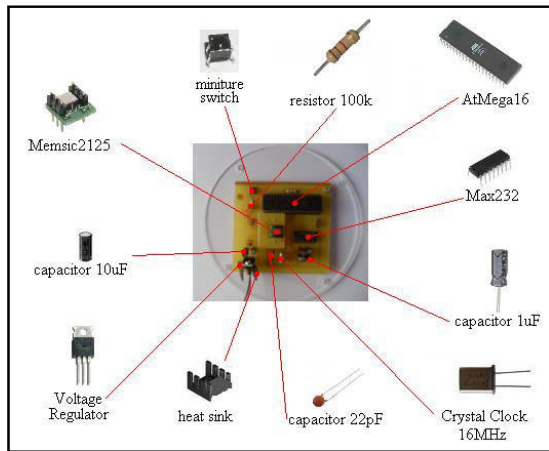


Figure 8. Control Sub-system

The rest of the components are carefully soldered unto the PCB board.

Programming the microcontroller is necessary to perform the desired application of the robot. These codes enable the robot to do its explicit function, designed by the programmer depending on the tasks that are needed to be completed. The functional diagram of the robot is shown in Figure 9.

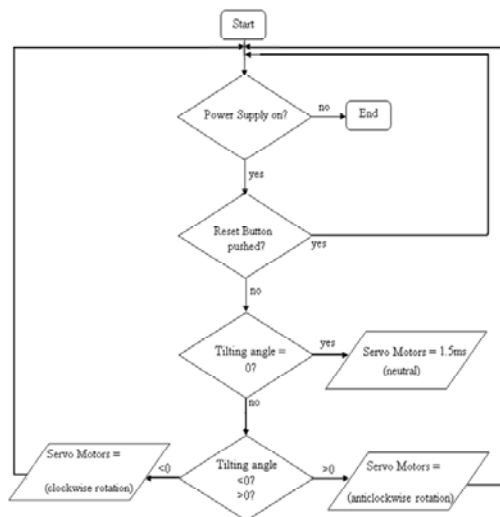


Figure 9. Functional diagram

From Figure 9, it can be seen that initially, it ensures whether the power supply is turned on or off. Then, it senses for reset button. If the reset is pressed, the flow continues to the input data from Memsic2125. This would then continue in a loop. A series of tests

were carried out to determine the robustness of the algorithms and it performed well the repeatability reaching 73%.

## V. CONCLUSION

A laboratory scale self-balancing robot was developed and tested. The robot was able to perform the desired tasks effectively. Further research should study the robustness of the control strategy with additional incorporation of IR sensors or gyroscope for better maneuverability and control.

## ACKNOWLEDGEMENTS

The authors are grateful to UNISEL and UTP for supporting this work.

## REFERENCES

- [1] Hu, J. S, 2009, "Self-balancing Control and Manipulation of a Glove Puppet Robot on a Two-wheel Mobile Platform", International Conference on Intelligent Robots and Systems, St. Louis, USA
- [2] Ha, Y. and Yuta, S. Trajectory tracking control for navigation of the inverse pendulum type self-contained mobile robot. Robot. Autonomous Systems, 1996, 17, 65 – 80
- [3] Kumagai, M. and Ochiai, T. "Development of a Robot Balancing on a Ball," Proc. ICCAS 2008, 2008, pp 433 – 438
- [4] <http://www.geology.smu.edu/~dpa-www/robo/nbot/>
- [5] Segway, LLC, [www.segway.com](http://www.segway.com)
- [6] Thomas Braunl, Embedded Robotics. Berlin Heidelberg: Springer – Verlag, 2006, ch.9
- [7] Ong Yin Chee, Mohamad Shukri b. Zainal Abidin, "Design And Development of Two Wheeled Autonomous Balancing Robot", 4<sup>th</sup> Student Conference on Research and Development (SCORed 2006), Shah Alam, Selangor, MALAYSIA, 27-28 June 2006.
- [8] N. M. Abdul Ghani, N. I. Mat Yatim, N. A. Azmi, Comparative Assessment for Two Wheels Inverted Pendulum Mobile Robot Using Robust Control", International Conference on Control, Automation and System 2010, KINTEX, Gyeonggi-do, Korea, Oct 27-30, 2010
- [9] Hua Sun, Haixu Zhou, Xiang Li, Yanhui Wei and Xiao Li, "Design of Two-Wheel Self-Balanced Electric Vehicle based on MEMS", Proceedings of the 2009 4th IEEE International Conference on Nano/Micro Engineered and Molecular Systems, Shenzhen, China, January 5-8, 2009
- [10] Araghi, M.H. and Kermani, M.R., Computer-Aided System Design for Educational Purposes: An Autonomous Self-Balancing Two-Wheeled Inverted Pendulum Robot, IEEE, CCECE 2011.

# Simulink Based FFT-IFFT Application On Image Processing

Divya S Das<sup>1</sup> & Raji C Nair<sup>2</sup>

<sup>1</sup>REVA Institute of Technology & Management, Bengaluru-560064, India

<sup>2</sup>Dept. of ECE, RITM, Bengaluru-560064, India

---

**Abstract** - A VLSI architecture realizing the FFT-IFFT has to perform in real time. This paper presents an application of FFT in image processing using simulink. The design involves radix-2 and radix-4 processing elements with complex multipliers and can be configured in real time to accommodate FFT computations of length 16, 32, 64, 128 and 256. For further speeding up the entire algorithm, the design can include a technique for parallel accessing the memory banks of each processing element.

**Key words** - FFT, IFFT, Radix-2, Radix-4, Image processing.

---

## I. INTRODUCTION

The spectral analysis of discrete signals plays an important role both in Digital Signal Processing and in Telecommunication Applications. The Fast Fourier Transform (FFT) Algorithm, as introduced by Cooley and Tukey, is an important improvement on the Discrete Fourier Transform (DFT) Algorithm, due to the achievement of significantly lower Computational complexity.

Here 2-D image is transferred through FFT and IFFT process. This procedure requires real time Inverse FFT and FFT processing during the modulation and demodulation of the signal respectively. The performance required by the FFT processing demands either a single processor driven to a very high clock frequency or alternatively the implementation of an Application Specific Integrated Circuit (ASIC), solution utilizing parallel processing and bit-pipelining techniques. Both solutions have to meet requirements that are extremely demanding in terms of throughput and low power consumption.

## II. PROBLEM DEFINITION

This Section presents the FFT computation and its application on image processing and related results with respect to the Process.

*The FFT in Image Processing Systems:*

As shown in the Decimation in Time FFT (DIT FFT)  $X[k]$  of a sequence  $x[n]$  is computed by successively decomposing the input signal in odd and even samples and recursively applying the DFT

algorithm to the resulting sequences. This can be illustrated as follows:

$$X[K] = \sum_{i=1}^N X[i] W_N^{Ki}$$
$$X[K] = \sum_{i=1}^{N/2} X[2i] W_N^{K2i} + \sum_{i=1}^{N/2} X[2i+1] W_N^{K(2i+1)}$$
$$X[K] = \sum_{i=1}^{N/2} X[2i] W_N^{K2i} + W_N^K \sum_{i=1}^{N/2} X[2i+1] W_N^{K2i}$$

Where:  $W_N = e^{(-j2\pi/N)}$

Exploiting the inherent symmetry of the twiddle factors and recursively applying the above formula, we obtain the elementary “butterfly” operation used in the calculation of the FFT.

*Related work:*

In [1] presents how the perfect shuffle interconnection pattern can be used in order to perform FFT transforms. Several methods for performing FFT computations are presented in [2]. A variety of algorithms for pipeline and parallel pipeline processors are examined, with respect to VLSI implementation. [3] Presents among others a set of algebraic tools that can be used to describe processor networks in terms of their patterns of connections. A radix-22 algorithm is presented in [4], which combines the radix-2 butterfly structure and the radix-4 multiplicative complexity.

### III. IMAGE PROCESSING

Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or, a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it.

Image processing usually refers to digital image processing, but optical and analog image processing also are possible. This article is about general techniques that apply to all of them. The acquisition of images (producing the input image in the first place) is referred to as imaging.

#### 3.1 ARCHITECTURE

##### *FFT:*

This Section describes the overall architecture and the details of the individual blocks, namely the processing elements, the deserialiser. The FFT-architecture performs FFT or IFFT computation of  $2n$  points,  $4 < n < 8$ .

The FFT architecture consists of Processing Elements (PEs): radix-4 PEs and radix-2 PE, as shown in Figure. Each PE performs a single stage of the FFT computation within the time required to input one data frame. The FFT architecture can be configured in real time, in order to perform FFT computations of 16, 32, 64, 128 or 256 points.

The FFT core computes an  $N$ -point forward DFT or inverse DFT (IDFT) where  $N$  can be  $2^m$ ,  $m = 3-16$ . The input data is a vector of  $N$  complex values represented as dual  $bx$ -bit two's complement numbers, that is,  $bx$  bits for each of the real and imaginary components of the data sample, where  $bx$  is in the range 8 to 24 bits inclusive. Similarly, the phase factors  $bw$  can be 8 to 24 bits wide.

All memory is on-chip using either block RAM or distributed RAM. The  $N$  element output vector is represented using  $by$  bits for each of the real and imaginary components of the output data. Input data is presented in natural order, and the output data can be in either natural or bit/digit reversed order. The complex nature of data input and output is intrinsic to the FFT algorithm, not the implementation.

##### *Processing Element Description:*

This section describes the functionality of the blocks that support each processing element (PE). The

following paragraphs define the interfaces of the blocks within the PEs, as well as the interfaces among the PEs.

##### *Radix 4:*

A radix-4 processing element performs one stage of radix-4 butterfly computations to the data. Each radix-4 PE consists of the following blocks:

##### *RAM:*

This block is used to store the input data to the processing element. The preceding butterfly stage supplies the write addresses to the radix-4 PE. The read addresses are generated within the radix-4 PE (as described below). The memory block is organized internally with two memory banks. Each bank can store one (1) frame of data. The first bank can be considered as the working bank for the FFT core. The second Bank is used to store the incoming input data. The two memory banks switch roles at the beginning of each incoming frame.

##### *Address Generators:*

There are two address generators. The first is used for supplying read addresses to the RAM block internal to the PE. The second is used for supplying write addresses for the data that exit the PE. The addressing scheme is the same for both Address Generators. Each radix-4 PE uses a distinct addressing scheme depending on the FFT stage realized by the PE.

##### *Twiddle (W) Generator:*

A Look-Up-Table (LUT) contains the max (N) roots of unity, where max (N) is the maximum size of FFT that is supported by the architecture (256). The Twiddle Address Generator is a simple N-counter-based architecture. At each data cycle the Twiddle Generator fetches the appropriate twiddle factor.

##### *Butterfly Core:*

This block performs the radix-4 FFT butterfly computation. The  $N$  input data are read sequentially from the RAM block. Each set of four consecutive input data forms the input to each radix-4 calculation. The corresponding twiddle factors are also fetched from the Twiddle Generator. Four (4) complex accumulators are used to process the input data in parallel. Each accumulator-process involves the add-subtract of the four data, as these operations are defined by the radix-4 data flow. A single complex multiplier unit operates on the four (4) accumulated results and the twiddle factors in a pipeline fashion. The resulting data are written sequentially to the RAM block of the following (FFT) PE.

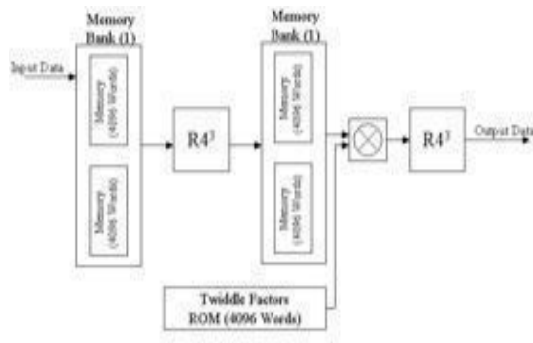


Fig. 3.1 : Overall FFT Architecture

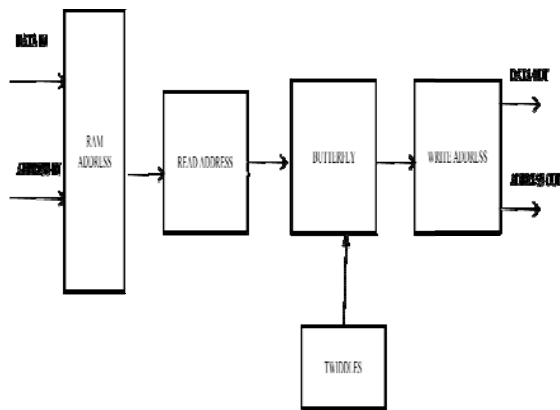


Fig. 3.2 : Radix-4 &amp; Radix-2 processing element internal Units

Each radix-4 processing element in the reconfigurable FFT pipeline has to be provided the length of the data frame it is processing by the Overall Control Unit (OCU). With this information the addressing scheme used by the address generator block and the twiddle factor generator block is changed. This is accomplished by selecting the proper permutation of the binary counter that all three generator blocks realize.

#### Radix 2:

The radix-2 PE applies one stage of radix-2 butterfly computations to its data. It is used when the size of the frame to be processed is 32 or 128 points. The radix-2 PE is realized as a simplified radix-4 PE. The Butterfly Core is replaced with the simpler radix-2 butterfly network, consisting of two (2) complex adders/subtractions and one (1) complex multiplier.

This circuit though is optimized further. In the split radix 128 and 32 point FFT computation, the twiddle factors for all radix-2 butterflies have the constant value of  $(1+0j)$ . Plugging into the radix-2 butterfly equations we obtain:

$$\begin{cases} A^1 = A & B.W \\ B^1 = A & B.W \end{cases} \quad \begin{cases} A^1 = A & B \\ B^1 = A & B. \end{cases}$$

Consequently, the complex multiplier (in the butterfly core) and the twiddle generator blocks are omitted.

#### IV. BLOCK DIAGRAM:

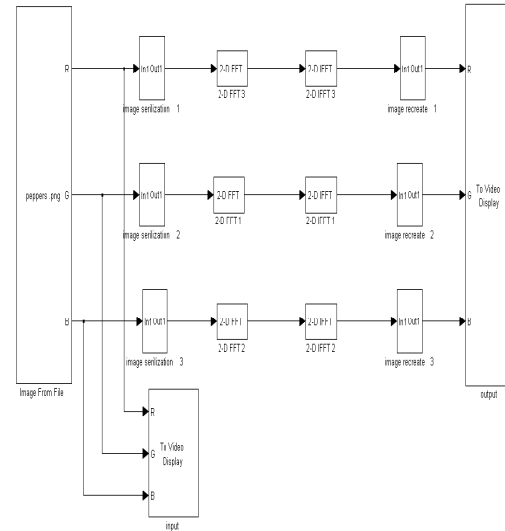


Fig. 4 : Simulation of image using image processing

The above block shows the simulation of image using FFT and IFFT, this image gives a clarity in processing as the size increases.

For the current process the size of the image given is 512, so the overall calculation is carried out for three stages that is  $512 \times 512 \times 512$ . The resizing block is used to reduce the overall image size; the total image cannot be given into the process because it is time consuming and lengthy process. Serialization of the resized image is carried, data of which is transferred to the FFT for the processing. The IFFT is carried out along with deserialization and the output is given to the video image block for viewing.



## V. RESULTS



Fig. 5 : Input image given for processing



Fig. 5.1 : Output image after IFFT

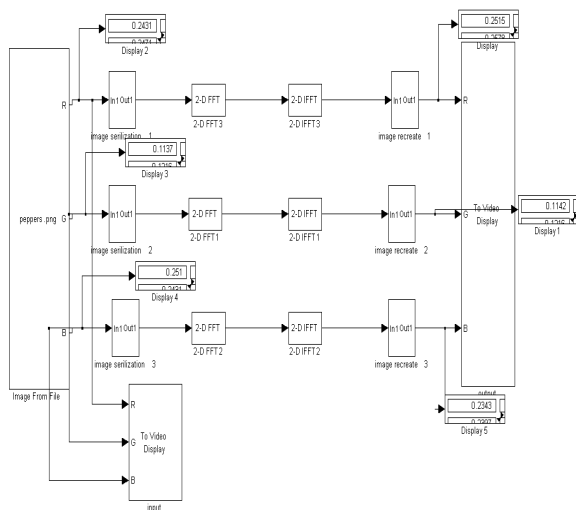


Fig. 5.2 : Displaying the values of input and output after processing

## VI. CONCLUSION

A VLSI architecture realizing the FFT has to perform in real time. FFT and IFFT process is carried out for the given image and the result is displayed.

This application of FFT-IFFT in image processing is done using MATLAB simulink. The design involves radix-2 and radix-4 processing elements with complex multipliers and can be configured in real time to accommodate FFT computations of length 16, 32, 64, 128 and 256. For further speeding up the entire algorithm, the design can include a technique for parallel accessing the memory banks of each processing element.

## REFERENCES

- [1] H. D. Stone 'Parallel Processing with the Perfect Shuffle.', IEEE Transactions on Computers, vol. C-20, no.2, 1971.
- [2] E.H. Wold and A.M. Despain. 'Pipeline and Parallel FFT Processors for VLSI Implementations', IEEE Transactions on Computers, vol. C-33, 1984.
- [3] D. Stott Parker 'Notes on Shuffle/Exchange-Type Switching Networks.' IEEE Transactions on Computers, 1980
- [4] S. He and M. Torkelson 'A New Approach to Pipeline FFT Processor.' Proceedings of the IPPS, 1996.
- [5] Reconfigurable Fast Fourier Transform Architecture for Orthogonal Frequency Division Multiplexing Systems  
Konstantinos E. MANOLOPOULOS,  
Konstantinos G. NAKOS, Dionysios I. REISIS  
and Nikolaos G. VLASSOPOULOS  
Electronics Laboratory, Department of Physics  
National and Capodistrian University of Athens  
Physics Buildings IV & V, Panepistimiopolis,  
Athens, Greece – 15784, Email:  
dreisis@cc.uoa.gr
- [6] A HIGH PERFORMANCE VLSI FFT ARCHITECTURE, K. Babionitakis, K. Manolopoulos, K. Nakos, D. Reisis, N. Vlassopoulos, Electronics Laboratory, Department of Physics, National and Kapodistrian University of Athens, Athens, Greece dreisis@phys.uoa.gr V.A. Chouliaras  
Department of Electronic and Electrical Engineering, Loughborough University, Loughborough, LEICS, LE11 3TU, UK  
V.A.Chouliaras@lboro.ac.uk



# A Novel Approach For The Enhancement Of Medical Images Using Thresholding

Nirpjeet Kaur<sup>1</sup> & Rajpreet Kaur<sup>2</sup>

<sup>1</sup>Deptt. Of Computer Science and Engineering, IET BHADDAL Ropar, India

<sup>2</sup>Deptt. Of Computer Science and Engineering, BBSBEC Fatehgarh Sahib, India

**Abstract** - The main objective of medical image segmentation is to extract and characterize anatomical structures with respect to some input features or expert knowledge. This paper describes a way of medical image enhancement using optimized approach. The purposed technique is compared with the popular algorithms i.e improved Otsu method ,improved moment preserving thresholding. Our algorithm outperformed over other algorithms in terms of speed ,threshold value and visual output.

**Key words** - Thresholding, Moment, Image segmentation, Threshold value, Image enhancement , Algorithms.

## I. INTRODUCTION

Medical image processing has revolutionized the field of medicine by providing novel methods for extracting and visualizing anatomical regions of interest. In particular, medical image segmentation plays a vital role in many medical imaging applications such as the quantification of tissue volumes, the extraction of anatomical structures, and computer-aided surgery[1,4]. Segmentation is the process of labeling each pixel in a medical image data set to indicate its tissue type or anatomical structure. Because of its intuitive properties and simplicity of implementation, image thresholding enjoys a central position in applications of image segmentation. Thresholding is the simplest method of image segmentation. Individual pixels in a grayscale image are marked as “object” pixels if their value is greater than some threshold value (assuming an object to be brighter than the background) and as “background” pixels otherwise[2]. Typically, an object pixel is given a value of “1” while a background pixel is given a value of “0.” In this paper, a novel image enhancement operation is proposed and applied to medical images prior to object/background binarization[3]. The proposed method adaptively determines the global threshold to enhance the grey scale image based on triangle approach and finally give the bianrized image. The purposed technique gives better output than the other existing techniques.

## II. RELATED WORK

### A) Otsu method [6,17]

otsu is a global thresholding method [which is based on discriminate analysis. The threshold operation is regarded as the partitioning of the pixels of an image into two classes C0 and C1 (e.g., objects and background) at grey-level t, i.e., C0 = {0, 1,2,..., t} and C1 = {t + 1,t+2,...L-1}. As stated in [11], let  $\sigma_w^2$ ,  $\sigma_B^2$  and  $\sigma_T^2$  be the within-class variance[7], between-class variance, and the total variance, respectively. An optimal threshold can be determined by minimizing one of the following (equivalent) criterion functions with respect to :

$$\lambda = \frac{\sigma_B^2}{\sigma_w^2}, \quad \eta = \frac{\sigma_B^2}{\sigma_T^2}, \quad \kappa = \frac{\sigma_T^2}{\sigma_w^2}$$

The above three criterion functions,  $\eta$  are the simplest. Thus, the optimal threshold t[19] is defined

as  $t = \text{ArgMin } \eta$

$$\text{Where } \sigma_T^2 = \sum_{i=0}^{L-1} [1 - \mu_T]^2 p_i, \quad \mu_T = \sum_{i=0}^{L-1} [i p_i]$$

$$\sigma_B^2 = w_0 w_1 (\mu_0 \mu_1)^2$$

$$w_0 = \sum p_i, \quad w_1 = 1 - w_0$$

$$\mu_1 = \frac{\mu_T - \mu_0}{1 - \mu_0}, \quad \mu_0 = \frac{\mu_1}{w_0}, \quad \mu_t = \sum_{i=0}^t (i p_i)$$

$$p_i = n_i / n$$

Where  $n_i$  is the number of pixels with grey-level  $i$  and  $n$  is the total number of pixels in a given image defined as  $n = \sum n_i$ . Moreover,  $p_i$  is the probability of occurrence of grey-level  $i$ . Otsu's method as proposed affords further means to analyze further aspects other than the selection of the optimal threshold for a given image [12,13]. For a selected threshold  $t$  of a given image, the class probabilities  $w_0$  and  $w_1$  indicate the portions of the areas occupied by the classes  $c_0$  and  $c_1$ . The class means  $\mu_0$  and  $\mu_1$  serve as estimates of the mean levels of the classes in the original greylevel image [17].

#### B) Moment preserving thresholding [15,16]

Given an image  $f$  with  $n$  pixels whose gray value at pixel  $(x, y)$  is denoted by  $f(x, y)$ , we want to threshold  $f$  into two pixel classes, the below-threshold pixels [18] and the above-threshold ones. The  $i$ th moment  $m_i$  of  $f$  is defined as

$$m_i = (1/n) \sum_{x,y} f_i(x,y) \quad i = 1, 2, 3, \dots \quad (1)$$

Moments can also be computed from the histogram of  $f$  in the following way:

$$m_i = (1/n) \sum n_j (z_j)^i \quad (2)$$

$$= \sum p_j (z_j)^i$$

where  $n_j$  is the total number of the pixels in  $f$  with gray value  $z_j$  and  $p_j = n_j / n$ . We also define  $m_0 = 1$ . Image  $f$  can be considered as a blurred version of an ideal bi-level image [17] which consists of pixels with only two gray values  $z_0$  and  $z_1$ , where  $z_0 < z_1$ . The moment-preserving thresholding is to select a threshold value such that if all below-threshold gray values in  $f$  are replaced by  $z_0$  and all above-threshold gray values replaced by  $z_1$  then the first three moments of image  $f$  are preserved in the resulting bi-level image  $g$ . [9] Image  $g$  so obtained may be regarded as an ideal unblurred version of  $f$ . Let  $p_0$  and  $p_1$  denote the fractions of the below-threshold pixels and the above-threshold pixels in  $f$ , respectively, then the first three moments of  $g$  are just

$$m'_i = \sum_{j=0}^1 p_j (z_j)^i \quad i = 1, 2, 3. \quad (3)$$

And preserving the first three moments in  $g$  means the following equalities:

$$m'_i = m_i, \quad i = 1, 2, 3. \quad (4)$$

Note that

$$p_0 + p_1 = 1. \quad (5)$$

The four equalities described by (4) and (5) above are equivalent to

$$\begin{aligned} p_0 z_0^0 + p_1 z_1^0 &= m_0, \\ p_0 z_0^1 + p_1 z_1^1 &= m_1, \\ p_0 z_0^2 + p_1 z_1^2 &= m_2, \\ p_0 z_0^3 + p_1 z_1^3 &= m_3, \end{aligned} \quad (6)$$

where  $m_i$ , with  $i = 1, 2, 3$  are computed by (1) or (2) and  $m_0 = 1$ . To find the desired threshold value  $t$ , we can first solve the four equations described by (6) above to obtain  $p_0$  and  $p_1$ , and then choose  $t$  as the  $p_0$ -tile of the histogram of  $f$ , i.e., choose  $t$  such that

$$p_0 = (1/n) \sum_{z_j \leq t} n_j$$

In practice, there may exist no discrete gray value which is exactly the  $P_0$ -tile of the histogram. Then, the threshold  $t$  should be chosen the gray value closest to the  $P_0$ -tile.

### III. PURPOSED TECHNIQUE

The purposed technique calculate the global threshold and that threshold segment the image from object to background. This technique is histogram based in which first the histogram is generated for the grey scale image then we mark the grey levels into two categories one which has minimum no of pixels and the other which has maximum no of pixels. Then we draw a line from minimum to the maximum grey value which forms a triangle. We compare the actual histogram curve and the line which is earlier drawn, the grey level which has the maximum distance from the actual histogram curve and line, is chosen as optimal threshold. This threshold segment the image as object and background. Algorithm for the techniques is given below.

1. First convert the grey scale image into bmp format, we have chosen bmp format because it is compression less.
2. Now generate the histogram of the grey scale image.
3. Determine the highest grey level position  $G_{\max}$ , of the highest peak from the histogram.
4. Determine the lowest position,  $G_{\min}$ , of the first peak from histogram.
5. Construct a line between the peak of the histogram and the lowest value, which is represented as

$$Ag + Bh(g) + C = 0$$

where  $A, B, C$  are coefficients given by

$$A = \frac{-H(g) G_{\max} - H(g) G_{\min}}{G_{\max} - G_{\min}}$$

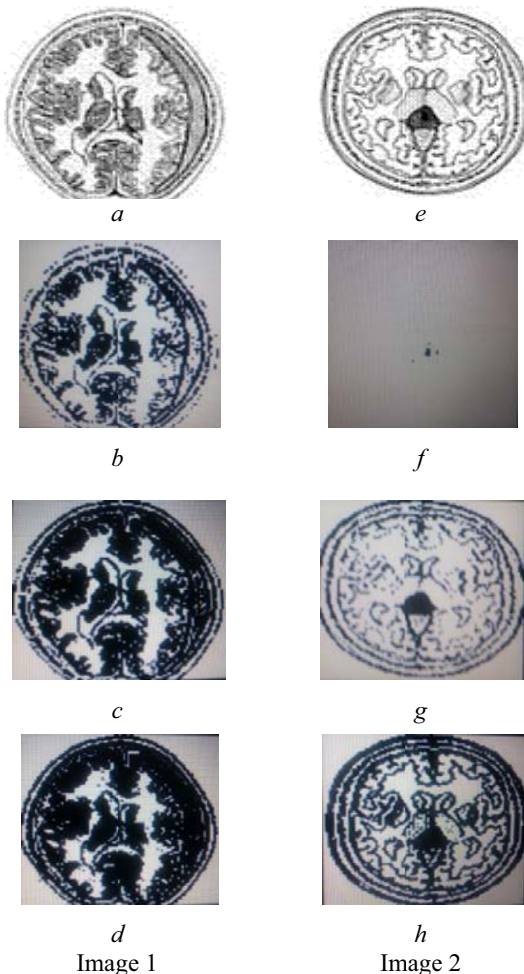
$$B=1, C= A(G_{\min})-H(g) G_{\min}.$$

6. If the histogram  $H(g) G_{\max}$  is greater than ,the  $H(g) G_{\min}$  then the distance,  $d$  between the line and the histogram  $Hg(g)$  is computed for all values  $g$  of from  $LOW G_{\max}$  to  $G_{\max-1}$  otherwise  $g$  is computed from  $G_{\min+1}$  to  $UPG_{\min}$  . where  $LOW G_{\max}$  is the left valley of the highest unsaturated peak and  $UP G_{\min}$  is the right valley of the first saturated peak. The distance,  $d$ , between the line and a point is computed by

$$d= \frac{Ag_0+BHg(g_0)+C}{\sqrt{A^2+B^2}}$$

The best threshold value occurs where the distance between and the line is maximal

7. After calculating the value, segment the image into object and background.



(a),(e)Original image (b),(f),image by otsu method  
(c),(g)image by moment preserving thresholding

(d),(h)image by purposed algorithm

#### IV. RESULTS AND CONCLUSION

The purposed algorithm is compared with the otsu and moment preserving thresholding. Our algorithm gives enhanced image as comparative to other in terms of visual output or in terms of time and threshold value .The purposed algorithm is tested on 10 different grey level medical images and proved to be better than the improved otsu and improved moment preserving thresholding . In future work ,the same can be tested on color images. The table 1 shows threshold value and time required to perform thresholding by otsu method , moment preserving method and purposed method .

Table 1

S. No	Method	Threshold value	Time
1	Otsu	Image 1	17
		Image 2	7
2	Moment preserving	Image 1	169
		Image 2	72
3	Purposed Algorithm	Image 1	196
		Image2	141

#### REFERENCES

- [1] Bülent Sankur<sup>a,\*</sup>, Mehmet Sezgin<sup>b</sup> “Survey over image thresholding techniques and quantitative performance evaluation” Journal of Electronic Imaging 13(1), 146–165 (January 2004).
- [2]. O. D. Trier and A. K. Jain, “Goal-directed evaluation of binarization methods,” IEEE Trans. Pattern Anal. Mach. Intell. PAMI-17, 1191–1201 ~1995!.
- [3]. M. Sezgin and B. Sankur, “Comparison of thresholding methods for non-destructive testing applications,” IEEE ICIP’2001, Intl. Conf. Image Process., pp. 764–767 ~2001!.
- [4]. P. Bock, R. Klinnert, R. Kober, R. M. Rovner, and H. Schmidt, “Gray-scale ALIAS,” IEEE Trans. Knowl. Data Eng. 4, 109–122~1992!.
- [5] WANG Hongzhi<sup>1</sup>, DONG Ying “An Improved Image Segmentation Algorithm Based on Otsu Method” Photo electronic Detection and Imaging

- : Related Technologies and Applications, Proc. of SPIE Vol. 6625, 66250I, (2008)
- [6]. N. Otsu, A threshold selection method from gray-level histogram, IEEE Transactions on Systems Man Cybernet, SMC-8 pp. 62-66, 1978.
- [7] P. K. Sahoo, S.Soltani, A.K. Wong, and Y. C. Chan, A survey of thresholding techniques, Computer Vision Graphics, and Image Processing, vol.41, pp. 233-260, 1998
- [8] J. Sauvola and M .Pietikainen, Adaptive document image binarization, Pattern Recognition, vol.33, pp.225-236,2000
- [9] Chun Che Fung "A Review of Evaluation of Optimal Binarization Technique for Character Segmentation in Historical Manuscripts", 2010 Third International Conference on Knowledge Discovery and Data Mining .
- [10] Tengku Mohd Afendi Zulcaffle," A Thresholding Algorithm for Text/Background Segmentation in Degraded Handwritten Jawi Documents" ,2010 Second International Conference on Advances in Computing, Control, and Telecommunication Technologies
- [11] Min-Ki Kim, "Adaptive Thresholding Technique for Binarization of License Plate Images," J. Opt. Soc. Korea 14, 368-375 (2010)
- [12] Leedham, G., S. Varma, A. Patankar, V.Govindaraju "Separating Text and Background in Degraded Document Images" Proceedings Eighth International Workshop on Frontiers of Handwriting Recognition, pp. 244-249, September, 2002
- [13] J.He, Q.D.M.Do, A.C.Downton, J.H.Kim, "A Comparison of Binarization Methods for Historical Archive Documents", ICDAR'05, p.538-542, 2005.
- [14] Niblack, W. "An Introduction to Digital image processing", pp 115-116, Prentice Hall, 2005.
- [15] S.D. Yanowitz and A.M. Bruckstein, "A new method for image segmentation," Computer Vision, Graphics and Image Process., vol. 46, no.1, pp. 82-95, Apr. 2003.
- [16] M.K. Yanni and E. Horne: "A new approach to dynamic thresholding," Proc. of the 9th European Conf. on Signal Process., Island of Rhodes,Greece, pp. 34-44, Sept 1998
- [17] Gatos B., Pratikakis I. and Perantonis S.J. "An adaptive binarisation technique for low quality historical documents". IAPR Workshop on Document Analysis systems (DAS'2006), Lecture Notes in Computer Science (3163), Florence, Italy, pp. 102-113
- [18] R. Gonzalez and R. E. Woods, Digital Image Processing, Prentice Hall, 2002
- [19] Ch.hima bindu,qiscet, ongole"An improved medical image segmentation algorithm using otsu method" International Journal of Recent Trends in Engineering, Vol 2, No. 3, November 2009.



# Potential of Electricity Generation through Solar Radiation at YCCE

Prasad B. Joshi, V. K. Deshpande & R. M. Moharil

Department of Electrical Engineering , YCCE, Nagpur, India

---

**Abstract** -The project presents estimation of steam generation potential at YCCE.Solar technology offers great potential in supply the world energy need. However it's current contribution to the world is still limited. There are different types of devices i.e. Collector, concentrator and receiver.These delivers different output quantitatively. Qualitatively the output of these devices is hot water. With some metrological input we would like to see which delivers proper desired output. In this project different types of devices will be simulated and their performance will be compared.

**Key words** - Solar collectors, Optical and thermal analysis, Water heating, Cooling, Solar power Generation.

---

## I. INTRODUCTION

Solar energy is a vital energy source that can make environment friendly energy more flexible, cost effective and commercially widespread. Photovoltaic cell is not use for steam generation. For maximum efficiency and more steam generation. Concentrator, collector and receiver are developed to solve this problem. Solar concentrator is a device that allows the collection of sunlight from a large area and focusing it on a smaller receiver or exit. A conceptual representation of a solar concentrator used in harnessing the power from the sun to generate electricity. A solar collector is used to concentrate solar radiation onto a receiver where heat transfer to a fluid takes place. There are different types of solar collector used.

The sun is a sphere of intensely hot gaseous matter with a diameter of  $1.39 \times 10^9$  m. The solar energy strikes our planet a mere 8 min and 20 s after leaving the giant furnace, the sun which is  $1.5 \times 10^{11}$  m away. The sun has an effective blackbody Temperature of 5762 K. The temperature in the central region is much higher and it is estimated at  $8 \times 10^6$  to  $40 \times 10^6$  K. The solar energy strikes our planet a mere 8 min and 20 s after leaving the giant furnace, the sun which is  $1.5 \times 10^{11}$  m away. The sun's total energy output is  $3.8 \times 10^{20}$  MW. Which is equal to 63 MW/m<sup>2</sup> of the sun's surface? This energy radiates outwards in all directions. This energy radiates outwards in all directions. Only a tiny fraction,  $1.7 \times 10^{14}$  kw, of the total radiation emitted is intercepted by the earth. The greatest advantage of solar energy as compared with Other forms of energy is that it is clean and can be supplied Without any environmental pollution. The benefits arising from the installation and operation of renewable energy systems can be

distinguished into three categories; energy saving, generation of new working posts and the decrease of environmental pollution. A worldwide research and development in the field of renewable energy resources and systems is carried out.

During the last two decades. Energy conversion systems that are based on renewable energy technologies appeared to be cost effective compared to the projected high cost of oil. Furthermore, renewable energy systems can have a beneficial impact on the environmental, economic, issues of the world. At the end of 2001 the total installed capacity of renewable energy systems was equivalent to 9% of the total electricity generation By applying a renewable energy intensive scenario the global consumption of renewable sources by 2050 would reach 318 exa joules.

Presently data regarding electricity consumption at YCCE collected. Conventionally for this much of average electricity consumption per month the machine rating is being calculated .To run this machine we will need a boiler the present work .Aim at replacing this boiler by harnessing solar energy available at this site the work compares of simulating and comparing various solar device, so that best out of them can be decide.

## II. SOLAR COLLECTORS

Solar energy collectors are special kind of heat exchangers that transform solar radiation energy to internal energy of the transport medium.The major component of any solar system is the solar collector.This is a device which absorbs the incoming solar radiation, converts it into heat, and transfers this

heat to a fluid (usually air, water, or oil) flowing through the collector.

### 2.1. Types of Solar Collectors

1. Non-concentrating or stationary.
2. Concentrating.

#### 2.1.1. Non-concentrating or stationary

A non-concentrating collector has the same area for intercepting and for absorbing solar radiation, whereas a sun-tracking.

#### 2.1.2. Concentrating

Concentrating solar collector usually has concave reflecting surfaces to intercept and focus the sun's beam radiation to a smaller receiving area, thereby increasing the radiation flux.

## III. STATIONARY COLLECTORS

Solar energy collectors are basically distinguished by their motion, i.e. stationary, single axis tracking and two axis tracking, and the operating temperature. These collectors are permanently fixed in position and do not track the sun .

### 3.1. Types of Stationary collectors

Following types of flat plat collector

1. Flat plate collectors (FPC).
2. Stationary compound parabolic collectors (CPC).
3. Evacuated tube collectors (ETC).

#### 3.1.1. Flat plate collectors (FPC)

Following components use in flat plat collector

1. Glazing
2. Tubes, fins, or passages
3. Absorber plates
4. Headers or manifolds
5. Insulation
6. Container or casing

#### 1. Glazing

One or more sheets of glass or other diathermanous (radiation-transmitting) material

#### 2. Tubes, fins, or passages

To conduct or direct the heat transfer fluid from the inlet to the outlet.

### 3. Absorber plates

Flat, corrugated, or grooved plates, to which the tubes, fins, or passages are attached. The plate may be integral with the tubes

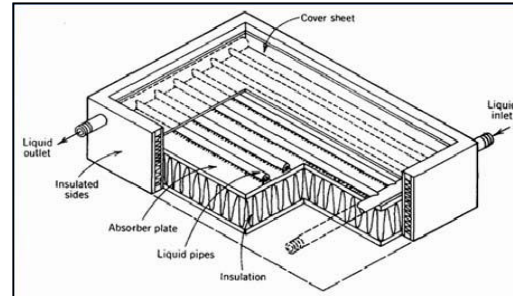


Fig. 1 : Flat plat collectors

### 4. Headers or manifolds

To admit and discharge the fluid.

### 5. Insulation.

To minimize the heat loss from the back and sides of the collector.

### 6. Container or casing.

To surround the a aforementioned components and keep them free from dust, moisture, etc

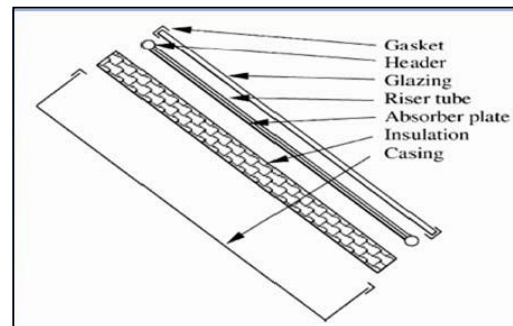


Fig. 2 : Exploded view of a flat plat collector

### 3.1.2. Working Principle

Flat-plate collectors are the most common solar collectors for use in solar water-heating systems in homes and in solar space heating. A flat-plate collector consists basically of an insulated metal box with a glass or plastic cover (the glazing) and a dark-colored absorber plate. Solar radiation is absorbed by the absorber plate and transferred to a fluid that circulates through the collector in tubes. In an air-based collector the circulating fluid is air, whereas in a liquid-based collector it is usually water . Flat collectors can be mounted in a variety of ways, depending on the type of

building, application, and size of collector. Options include mounting on a roof, in the roof itself, or free .

#### Advantages

1. Flat-plate collectors will absorb energy coming from all directions above the absorber.
2. flat-plate collectors do not need to track the sun
3. Simple in design
4. No moving part
5. Requires little maintenance
6. Less cost

#### 3.2. Various types of flat-plate solar collectors

Following fig shows various type of flat plat collector.

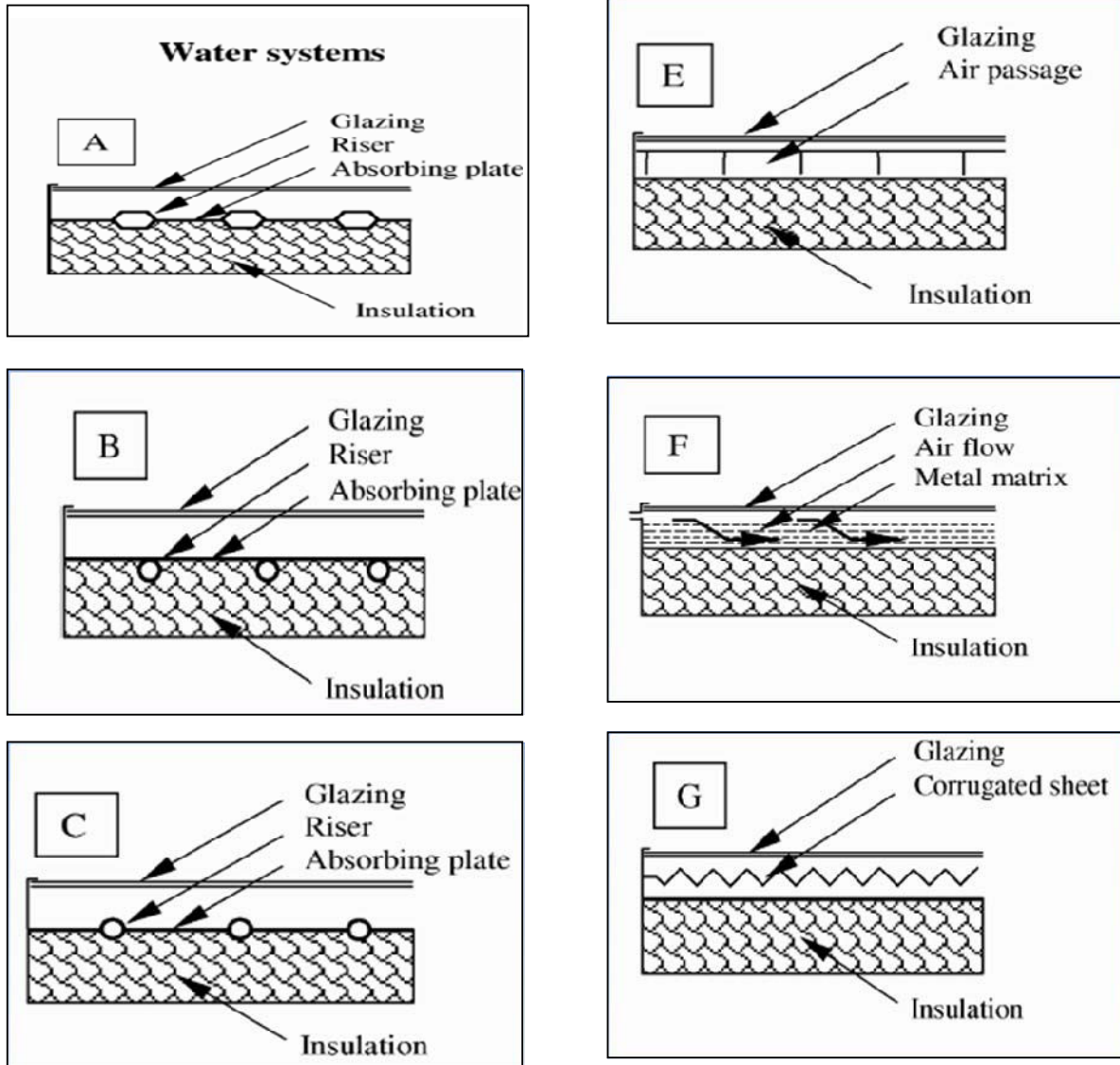


Fig. 3 : Various type of flat plat collector



### 3.2.1. Description

(Fig A) Bonded sheet design Fig shows a bonded sheet design, in which the fluid passages are integral with the plate to ensure good thermal conduct between the metal and the fluid. Fig B) & C) shows fluid heaters with tubes soldered, brazed, or otherwise fastened to upper or lower surfaces of sheets or strips of copper. Copper tubes are used most often because of their superior resistance to corrosion. The Fig D) shows the use of extruded rectangular tubing to obtain a larger heat transfer area between tube and plate. Mechanical pressure, thermal cement, or brazing may be used to make the assembly. Soft solder must be avoided because of the high plate temperature encountered at stagnation conditions. Fig E) Shows It is used to counteract the low heat transfer coefficients between metal and air .Metal or fabric matrices. Metal or fabric matrices (Fig. F) or thin corrugated metal sheets (Fig. G) may be used, with selective surfaces applied to the latter when a high level of performance is required. The principal requirement is a large contact area between the absorbing surface and the air. Various applications of solar air collectors are reported. A design procedure for solar air heating systems is presented whereas the optimization of the flow passage geometry is presented.

## IV. CONCENTRATING TYPE

Following types of concentrating types

1. Parabolic trough collector.
2. Linear Fresnel reflector (LFR).
3. Parabolic dish.
4. Central receiver

### 4.1. Parabolic trough collector

Concentrating, or focusing, collectors intercept direct radiation over a large area and focus it onto a small absorber area. These collectors can provide high temperatures more efficiently than flat-plate collectors, since the absorption surface area is much smaller. However, diffused sky radiation cannot be focused onto the absorber. Most concentrating collectors require mechanical equipment that constantly orients the collectors toward the sun and keeps the absorber at the point of focus. Therefore; there are many types of concentrating collectors.

- Parabolic troughs are devices that are shaped like the letter “u”.
- The troughs concentrate sunlight onto a receiver tube that is positioned along the focal line of the trough.

- Sometimes a transparent glass tube envelops the receiver tube to reduce heat loss.

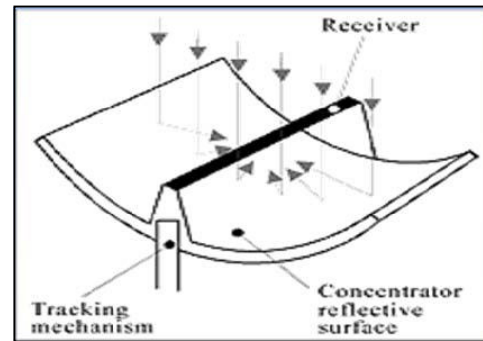


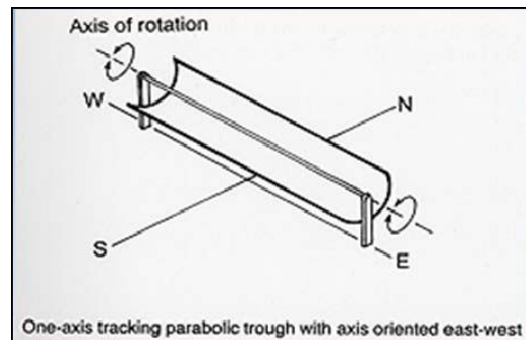
Fig. 4 : Parabolic trough system



Fig. 5 : Parabolic trough system

Temperatures at the receiver can reach 400 °C and produce steam for generating electricity. However with long operating experience, continued technology improvements, and operating and maintenance cost reductions, troughs are the least expensive, most reliable solar thermal power production technology for near-term

### 4.2. Parabolic trough often use single-axis or dual-axis tracking





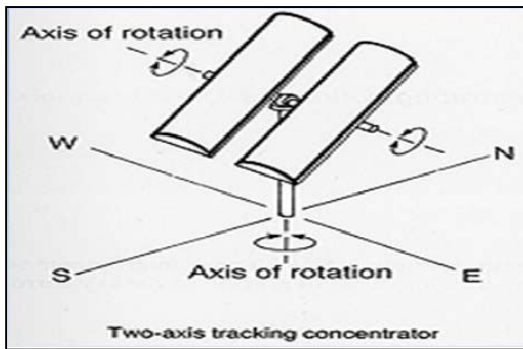


Fig. 6 : Single axis and two axis tracking concentrator

## V. THERMAL ANALYSIS OF COLLECTORS

The basic parameter to consider is the collector thermal efficiency. This is defined as the ratio of the useful energy delivered to the energy incident on the collector aperture. The incident solar flux consists of direct and diffuse radiation. While FPC can collect both, concentrating collectors can only utilize direct radiation if the concentration ratio is greater than 10.

### 5.1. Flat-plate collector's performance

Following equation show the thermal analysis of Flat plat collector.

$$\eta = F_R \left[ \tau\alpha - \frac{U_L(T_f - T_a)}{G_o} \right]$$

$F_R$  = Heat Removal Factor

$\tau\alpha$  = Transmittance–Absorptance Product

$U_L$  – Solar Collector Heat Transfer Loss Coefficient (w/m2)

$T_f$  = Temperatures of the Fluid Entering the Collector(c)

$T_a$  = Ambient Temperature (c)

$G_o$  = Total (Direct Plus Diffuse) Solar Energy Incident On The Collector Aperture (w/m2)

### 5.2. Concentrating collector's performance

Following equation show the thermal analysis of concentrating type collector

$$\eta = F_R \left[ \eta_o - U_L \left( \frac{T_f - T_a}{G_b C} \right) \right]$$

$F_R$  = Heat Removal Factor

$\eta_o$  = Collector Optical Efficiency

$C$  = Collector concentration ratio

$G_b$  = Beam (Or Direct) Irradiation (w/m2)

$T_f$  =Temperatures of the Fluid Entering the Collector(c)

$T_a$  = Ambient Temperature (c)

## VI. FLAT PLAT COLLECTOR AND PARABOLIC TROUGH COLLECTOR CONNECTED IN SERIES, PARALLEL AND SERIES-PARALLEL

Solar flat plat collector under normal circumstances does not have any loss of water either room temperature in hot condition there for when there are two or more flat plat collector connected in series the output of the first collector as an input of second naturally we say that when two collected are connected in series discharge remain constant .when two or more collected are connected parallel and if they adjacent to each other radiation received by both of them will be ideal speaking same other design parameter being head constant .Temperature accurse will be identical.

### 6.1. Flat plat collector and Parabolic through collector

- Series
  - Q : Constant
  - T : Increase
- Parallel
  - Q : Increase
  - T : Constant
- Series-Parallel
  - Q : Increase
  - T : Increase

### 6.2. YCCE College Data

Table No.(1) : YCCE Collage Data

S.N.	MONTH	KWH	KVAH	P.F
1	JAN .10	52500	52500	1
2	FEB.10	64435	64439	1
3	MAR.10	78025	78040	1
4	APRIL.10	12764	12860	1
5	MAY.10	77420	77485	1
6	JUNE.10	74265	74295	1
7	JULY.10	93410	93585	1
8	AUG.10	101020	101125	1
9	SEP.10	91805	91120	1

10	OCT.10	91815	91970	1
11	NOV.10	63475	63600	0.998
12	DEC.10	63590	63590	0.999
13	JAN.11	57210	57230	1
14	FEB.11	68470	68495	0.99
15	MAR.11	80185	80265	1
16	APRIL.11	93525	93695	1
17	MAY.11	84275	84370	1
18	JUNE.11	76980	77015	1
	<b>TOTAL</b>	<b>1325244</b>	<b>1325675</b>	

### 6.2.1 Calculation

Following are the calculation

$$\text{KWH} = 1325244 / 18$$

$$= 73624.666116 \text{ (18 Month)}$$

$$\text{KWAH} = 1325675 / 18$$

$$= 73648.611(18 \text{ Month})$$

$$= 73600 \text{ per month}$$

➤ Approximately  $P = 75000 \text{ KWH}$

➤ Convert into KW for one month

$$75000 / 24 \times 30$$

$$= 104.16 \text{ KW}$$

➤ 125 KW convert into HP

$$= 167.16 \text{ HP}$$

$$P = 167.16 \text{ HP}$$

## VII. MODELING OF SOLAR SYSTEMS

The proper sizing of the components of a solar system is a complex problem which includes both predictable (collector and other components performance characteristics) and unpredictable (weather data) components. In this section an overview of the simulation techniques and programs suitable for solar heating and cooling systems is presented.

Computer modeling of thermal systems presents many advantages the most important of which are the following.

1. Eliminate the expense of building prototypes.
2. Complex systems are organized in an understandable Format.

3. Provide thorough understanding of system operation and Component interactions.
4. It is possible to optimize the system components.
5. Estimate the amount of energy delivery from the system.
6. Provide temperature variations of the system.
7. Estimate the design variable changes on system Performance by using the same weather conditions.

The following sections describe briefly five Software programs.

- a) TRNSYS simulation program.
- b) WATSUN simulation program.
- c) Polysun simulation program.
- d) Chart method and program.
- e) Mat lab and Artificial neural networks in solar energy
- f) Systems modeling and prediction

## VIII. SOLAR COLLECTOR APPLICATIONS

Solar collectors have been used in a variety of applications. These are described in this section. In the most important technologies in use are listed together with the type of collector that can be used in each Case.

- a) Solar water heating systems
- b) Space heating and cooling
- c) Solar refrigeration
- d) Industrial process heat.
- e) Solar desalination.
- f) Solar thermal power systems

## IX. CONCLUSION

Form this paper it can be seen that any machine (turbine or it equivalent) which can given above  $P = 170 \text{ HP}$ . May be presently adequate for the electricity requirement at YCCE taking into consideration 200 HP machine may be sufficient for this purpose. In project phase two simulation of various devices to generate stem which can run this machine will be carried out further discussion taken.

## REFERENCES

- [1] Duffie and Backman 'Solar Engineering Of Thermal Processes' second edition, a Wiley – interscience publication, New York, June 1991.
- [2] D. P. Kothari, K. C. Singal and Rakesh Ranjan 'Renewable energy sources and emerging technologies', PHI Publication, 2008.

- [3] S P Sukhatme 'Solar energy', second edition, second edition Tata McGraw-Publication, 2005.
- [4] Soteris A. Kalogirou 'Solar thermal collectors and applications' Progress in Energy and Combustion Science 30 (2004) PP-231–295
- [5] Soteris A.Kalogirou and Sofia 'Thermo siphon solar domestic water heating system: Long term performance prediction using artificial neural networks' Solar Energy, Vol 69 pp 163-174 (2000)
- [6] Kalogirou S. (1996) artificial neural networks for predicting the local concentration ratio of parabolic through collector. In proceeding of the international conference EUROSUN (96). Freiburg, Germany pp. 470-473

