Journal of Emerging Technologies in Web Intelligence

ISSN 1798-0461

Volume 2, Number 2, May 2010

Special Issue: Emerging Technologies in Internet Security and Networks

Guest Editors: Himanshu Aggarwal and Vishal Goyal

Contents

Special Issue on Emerging Technologies in Internet Security and Networks

Guest Editorial

Computer systems have been the core of information technology to provide signal processing and storage capabilities and form the basis for the Internet connectivity. Computer systems act as the fundamental core hardware for implementation of the physical layer of today's worldwide information super highways. Networking technologies provide end-users with the facilities for network connection, data access, and communications through wired and wireless IP-based networks. On the other hand, communication technologies provide means for signal transmission and reception via efficient multiple access capability through various media, including copper wires, optical fibers, radio frequency (RF) and other approaches (such as infrared, laser, etc.). Integration of computer systems, networks, and communications is an important trend for future information technologies (ITs).

With the advances in the communication and internet technologies the Information security, data integrity and safe exchange of data and Information are most important and contemporary issues in the today's world. With a mission to report emerging a state-of-the-art research in Technologies in Network security and Computer Networks, the most important areas of information technology, a special issue of JEWTI has been dedicated The research contributions have been collected through the National Conference –Cum- Workshop on Information Security and Network (ISAN). The mission of ISAN was to promote research, development and new inventions of Information Security and Networks including mobile and wireless technologies DCAS, intrusion detection etc.

This Issue of the Journal is designed for researchers, developers, practitioners, policy-makers, professional trainers, educators, and other specialists.

Guest Editors:

Dr. Himanshu Aggarwal, Punjabi University, India.

Dr. Vishal Goyal, Punjabi University, India.

Himanshu Aggarwal, Ph.D., is Reader in Computer Engineering at Punjabi University, Patiala. He has more than 15 years of teaching experience and served academic institutions such as Thapar Institute of Engineering & Technology, Patiala, Guru Nanak Dev Engineering College, Ludhiana and Technical Teacher's Training Institute, Chandigarh. He is an active researcher who has supervised many M.Tech. Dissertations and contributed 40 articles in various Research Journals and Conferences. He is on the editorial Board of several journals of repute. His areas of interest are Information Systems, ERP and Parallel Computing. Himanshu Aggarwal can be contacted at: himagrawal@rediffmail.com.

Vishal Goyal, MCA, M.Tech., is senior lecturer at Department of Computer Science, Punjabi University Patiala, Punjab, India. He has more than 10 years of teaching, research and software industry experience. He has about 50 publications in various journals and conferences of national and international repute. He is member of editorial board of various journals of computer science of national and international repute. He is Technical team member of E&R Department of Infosys Technologies. DST has funded him research project worth Rs 28Lakhs for development of Punjabi Grammar checker. He is very active researcher in the field of Natural Language Processing. He has guided approx. 25 M.tech. students and 15 M.Phil Students for their research work. He has been honored with Young Scientist award by Punjab Academy of Sciences in 2005. He is being regularly invited by colleges and universities for his expert talks. He can be reached at vishal.pup@gmail.com or http://www.punjabiuniversity.ac.in.

doi:10.4304/jetwi.2.2.79-79

Performance of Dominating and Adaptive Partial Dominating Sets in AODV Routing Protocol for MANETs

A. Nagaraju *, Dr. S. Ramachandram**, B. Eswar***
* Head & Associate Prof in I.T , Kamala Institute of Tech & Science, Huzurabad, Karimnagar, A.P
** Head & Prof in CSE , Osmania University , Hyderabad. A.P, INDIA
***Assistant Prof in I.T, Kamala Institute of Tech & Science, Huzurabad, Karimnagar, A.P. INDIA

***Abstract*— A mobile ad hoc network (MANET) is a wireless network that does not rely on any fixed infrastructure (i.e., routing facilities, such as wired networks and access points), and whose nodes must coordinate among themselves to determine connectivity and routing. The broadcast can target a portion of the network (e.g. gathering neighborhood information), or the entire network (e.g., discovering routes on demand). Broadcasting of signaling and data in MANETs raise redundant transmission of control packets to overcome these problems we applied dominating set and Adaptive partial Dominating (APDP) approach to existing routing protocols such as Ad-hoc On-demand Distance Vector (AODV). The focus of this paper is to apply the concept of DS and APDP to AODV and evaluate the performance of dominating sets in AODV that improve broadcasting, End-to-End Delay, Network load, Packet Latency, and also maintains secure packet transmission.**

***IndexTerms*—Adaptive partial dominating, Dominating sets, AODV**

I. INTRODUCTION

Mobile IP and wireless networks accessing the fixed networks have provided support for the mobility. But it is still restrictive in forcing the connectivity at least to the core network. It puts impediments on supporting the true mobility in the network. In this connection, one area which is getting much attention in last couple of years is Mobile Ad Hoc Networks (MANETs). A MANET (Mobile Ad-Hoc Network) is a type of ad-hoc network with rapidly changing topology. ***S***ince the nodes in a MANET are highly mobile, the topology changes frequently and the nodes are dynamically connected in an arbitrary manner.

As defined in [1] a MANET is an autonomous system of mobile nodes. MANET nodes are equipped with wireless transmitters and receivers using antenna which may be omni directional, highly-point-to-point, possibly steerable, etc. At a given point in time, depending on the positions of the nodes and their transmitters and receivers coverage patterns, transmission power levels and co-channel interference levels, a wireless connectivity in the form of a random, multi hop graph or ad hoc network exists among the nodes. This ad hoc topology may change with time as the nodes move or adjust their transmission and reception parameters. There is current and future need for dynamic ad hoc networking technology. The emerging field of mobile and nomadic computing, with its current emphasis on mobile IP operation, should gradually broaden and require highly-adaptive mobile networking technology to effectively manage multi hop, ad-hoc network clusters which can operate autonomously or, more than likely, be attached at some point(s) to the fixed Internet.

There are two broad categories of unicast routing protocols for MANETs, proactive and reactive. With *proactive routing* (e.g., OLSR [21]), nodes keep routing information to all nodes in the network, not subject to any existing data flow. OLSR is a link state protocol using an optimized broadcast mechanism for the dissemination of link state information. In *reactive routing* (e.g., AODV [3]), routes are found on demand and nodes find routes to their destinations as they are needed. Route discovery starts by broadcasting a *route request* (RREQ) message throughout the network. This message is relayed until it reaches a node with a valid route to the destination, or the destination itself. Once this happens, a *routereply* (RREP) message is sent back to the source by reversing the path traversed by the RREQ message. Only after receiving the corresponding RREP message can the source start sending packets to the destination. Reactive and proactive routing can be combined, resulting in *hybrid protocols* (e.g., ZRP 20]). In this case, routes to some nodes (usually the nearest ones) are kept proactively, while routes to the remaining nodes are found on-demand.

Neighbor-knowledge-based methods mainly depend on the following idea: To avoid flooding the whole network, a small set of forward nodes is selected such that the forward node set forms a connected dominating set (CDS). A node set is a connected dominating set if every node in the network is either in that set or the neighbor of a node in that set. Then the challenge is to select a small set of forward nodes in the absence of global network information

In our proposed paper we focused on applying APD to AODV and to evaluate the performance of Dominating sets in Ad-hoc On Demand Distance Vector Routing algorithm(AODV) that improve broadcasting, End-to-End Delay, Network load, Packet Latency, and also Maintains secure packet transmission . To do this, concepts from ***domination in graphs*** have been explored (i.e.

doi:10.4304/jetwi.2.2.80-85

Dominating sets), The rest of the paper is organized as follows. Section 2 is the related work. Section 3 comprises Dominating Sets (Domination in graph theory) as in [9]. Section 4 is Route Request Algorithm using Dominant Pruning, Section 5 presents simulation results of this method and Section 6 presents Conclusions and future scope.

II. Related Work

Several broadcasting techniques have been proposed, differing among each other on the heuristics applied to reduce the redundancy on broadcast transmissions. Broadcasting protocols can be categorized into the following four classes [1]:

Blind flooding [9]: Each node broadcasts a packet to its neighbors whenever it receives the first copy of a broadcast packet; therefore, all nodes in the network broadcast the packet exactly once.

Probability-based methods [12]: A node re-broadcasts a packet with a given probability p (if $p = 1$, we have blind flooding).

Area-based methods [12]: A node broadcasts a packet based on the information about its location and the location of its neighbors (e.g., if a node receives the packet from a neighbor really close to it, probably it will not reach other nodes other than the nodes reached by the first broadcast).

Neighbor information methods [15]: In these methods, a node has partial topology information, which typically consists of the topology within two hops from the node (two hop neighborhood). There are two main classes of methods in this category. In a *neighbor designated method* a node that transmits a packet to be flooded specifies which one-hop neighbors should forward the packet. In a *self-pruning method* a node simply broadcasts its packet, and each neighbor that receives the packet decides whether or not to forward the packet. Williams and Camp [1] have shown that *neighbor information* methods are preferred over other types of broadcast protocols. Between the two classes of neighbor information methods, Lim and Kim [6] show that the simplest form of neighbor-designated algorithm out performs the simplest form of self-pruning, and Wu and Dai [7] show that an improved self-pruning technique outperforms the most efficient neighbor-designated algorithm based on the two-hop neighborhood information

Dominating sets play a major role in deciding the forwarding list in neighbor designated algorithms. Extensive work has been done on finding good approximations for computing the *minimum cardinality* CDS (MCDS). An algorithm with a constant approximation of eight has been proposed by Wan et al. [13]. However, their approach requires that a spanning tree to be constructed first in order to select the dominating nodes (forwarding nodes), and only after the tree has been constructed a broadcast can be performed.

The forwarding nodes are selected using the *greedy set cover* (GSC) algorithm. GSC recursively chooses one-hop neighbors that cover the most two hop neighbors, repeating the process until all two-hop neighbors are covered. The identifiers (IDs) of the selected nodes are piggy-backed in the packet as the forwarding list. A receiving node that is requested to forward the packet again determines the forwarding list.

In our proposed method, we apply Dominating Set model to identify the best RREQ forwarding nodes among the existing neighbors. Here the RREQ are transferred using the forwarder list information. This kind of mechanism controls the overhead of Route Request Phase (RREQ) of AODV by eliminating the redundant RREQ forwarding towards the destination.

III. Dominating Sets (Domination in Graph Theory)

As our project involves the computation of dominating sets, we provide a brief introduction to domination in graph theory below.

In our notation, the undirected graph G = (V, E) consists of a set of vertices V represents a set of wireless mobile nodes and E represents a set of edges. A set $D \subseteq V$ of vertices in graph G called a dominating set (DS) if every $n_i \in V$ either an element of D or is adjacent to an element of D [15]. If the graph induced by the nodes in D is connected, we have a connected dominating set (CDS). The problem of computing the minimum cardinality DS or CDS of any arbitrary graph is known to be NP-complete [15].

In dominant pruning (DP) the sending node decides with adjacent nodes should relay the packet. The relaying nodes are selected using the distributed CDS algorithm, and the identifiers (IDs) of the selected nodes are piggybacked in the packet forward list. A receiving node that is requested to forward the packet again determines the forwarder list. The flooding ends when there is no more relaying nodes.

A. Dominant Pruning Algorithm

Selection Process:

Step 1 Let F(u,v) = [ ] (empty list)
Z = Ø (empty set) and K = U S_i
Where $S_i = N(v_i) \cap U(u,v)$ for $v_i \in B(u,v)$

Step 2 Find Set S_i whose size is maximum in K
(In case of a tie, the one with smallest identification I selected)

Step 3 $F(u,v) = F(u,v) \,||\, v_k$, $Z = Z \cup S_i$, $K = K - S_i$ and
$S_j = S_j - S_i$ for all $S_j \in K$

Step 4 If Z = U(u,v) exit ; Otherwise , go to Step 2

Nodes maintain the information about their two-hop neighborhood, which can be obtained by the nodes exchanging their adjacent node list with their neighbors. DP is the distributed algorithm that determines a set cover based

B. Route Request Algorithm Using DP Algorithm

On-demand route discovery is based on *route request* (RREQ) and *route reply* (RREP) messages (e.g., AODV [3] and DSR [4]). The way in which these

messages are handled may differ among different protocols, but their functionality remains the same: a request is relayed until it reaches a node with a valid route to the destination or the destination itself, which triggers a reply message sent back to the originator. Several parameters (such as how long to keep requests in a cache, timeouts for requests, timeouts for hellos) are subject to tuning, and the choices made may result in improvements in the protocol performance. However, RREQs are propagated using either an unrestricted broadcast or an expanding ring search. In either case, the resulting flooding operation causes considerable collisions of packets in wireless networks using contention-based channel access.

In addition to applying DP to reduce the number of nodes that need to propagate RREQs transmitted on broadcast mode, information regarding prior routes to a destination is used to unicast RREQs to a region close to the intended destination, so that broadcast RREQs are postponed as much as possible and occur (if necessary) only close to the destination, rather than on network-wide basis. This RREQ Algoritm presents the pseudo-code for the modified RREQ. A route request (RREQ) is handled as follows:

• If the source of a RREQ does not have any previous knowledge about the route to the destination or is retrying the RREQ, it calculates its forwarder list using DP, and broadcasts the packet (Lines 8, 9, and 14).

• On the other hand, if the source of a RREQ has knowledge about a recently expired route to the destination, and there is a valid route to the next hop towards the destination (Lines 2, 3, and 4), the node calculates the forwarder list using DP (Line 9), but instead of broadcasting the RREQ packet, the node unicasts the packet to the last known next hop towards the destination (Line 12).

Upon receiving a route request, a forwarder that cannot respond to this request calculates its own forwarder list using the information provided in the RREQ packet (i.e., forwarder list, second to previous forwarder list, and source node) and broadcasts or unicasts the packet (depending on which one of the two first cases apply) after updating it with its own forwarder list.

RREQ Algorithm

Data : n_i, destination D, B_i, U_i
Result : Unicast the RREQ, or Broadcast the RREQ
Begin

```
1   if recently expired route to D and not retrying
      then
2   NextHop ← previous_nextHop(D)
3      if validRoute(NextHop) then
4         result←Unicast
5      else
6         result←Broadcast
7   else
8      result←Broadcast
9   F_i←DP( n_i, B_i, U_i )
10  Update RREQ packet with F_i
11  if result== Unicast then
12     Unicast the RREQ packet to NextHop
13  else
14     broadcast the RREQ packet
    end
```

Eventually, the RREQ reaches a node with a route to the destination or the destination itself. Our approach attempts to reduce the delay of the route discovery by unicasting a RREQ towards the region where the destination was previously located. The success of this approach depends on how fresh the previous known route to the destination is, and how fast the destination node is moving out of the previous known location. If an intermediate node has completely removed any route to the destination, the RREQ is then broadcasted. The intended effect is to postpone the broadcast of a RREQ to the region closest to the destination. In the case that the unicast approach fails, or there is no previous route to the destination, the source broadcasts by default.

Because of topology changes, nodes may not have correct two-hop neighborhood information, which may result in forwarding lists that do not cover all nodes in the neighborhood. However, this is not a major problem when the request is broadcasted, because a node incorrectly excluded from the forwarder list may also receive the request and is able to respond in the case it has a route to the destination.

C. Enhanced Dominant Pruning Algorithm

In their paper [6], wei Lou and Jie Wu proposed two enhanced dominant pruning algorithms: the Total Dominant pruning (TDP) algorithm and Paratial Dominant Pruning (PDP).

In TDP algorithm, if node v can receive a packet piggybacked with N(N(v)) from node u , the 2-hop neighbour set that needs to be covered by v's forward node list F is reduced to U = N(N(v)) – N(N(u)). The main objective of the TDP algorithm is that 2-hop neighbourhood information of each sender is piggybacked in the broadcast packet which results in consumption of more bandwidth.

D. Partial Dominant Pruning Algorithm:

Just like the DP algorithm, in PDP, no neighbourhood information of the sender is piggybacked with the broadcast packet. Apart from excluding N (u) and N (v) from N (N (v)) as in the DP algorithm, we can here exclude some more nodes from neighbours of each node in N (u) ∩ N (v). Such a node set is denoted by P (u,v) (or simply P) = N(N(u) ∩ N(v)). Then 2-hop neighbour set U in the PDP algorithm can be given by U = N(N(v)) – N(u) – N(v) – P since P is a subset of N(N(u)), we can easily see P can be excluded from N(N(v)). Also it can be proved that U is subset of N(B), when P = N(N(u)∩N(v)), U = N(N(v)) – N(u) – N(v)-P and B = N(v) – N(u).

E Adaptive Partial Dominant Pruning Algorithm[19]

Adaptive dominant pruning algorithm (APDP) is similar to PDP. However, besides excluding N (u), N (v) and P from N (N (v)) as mentioned in PDP algorithm, adjacent nodes of U are eliminated from U.

APDP Algorithm

1 Node v uses N(N(v)) ,N(u), and N(v) to obtain P = N (N(u) ∩ N(v)), U^1 = U – E where U = N(N(v)) – N(u) – N(v) – P and E is the set of equivalent and adjacent nodes in U and B = N(v) – N(u)

2 Node v calls the selection process to determine F.

Now consider Fig 1 in the example of sample ad-hoc network with 12 nodes. Table 1 shows each node in Fig 1 1-hop and 2-hop neighbor nodes. Here we illustrate the difference between PDP and APDP algorithms.

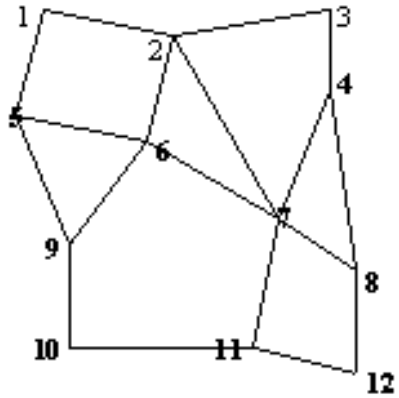


Fig 1 An Ad-hoc Network with 12 nodes.

Table 1: Two hop neighbors of each node

V	N(v)	N(N(v))
1	1,2,5	1,2,3,5,6,7,9
2	1,2,3,6,7	1,2,3,4,5,6,7,8,9,11
3	2,3,4	1,2,3,4,6,7,8
4	3,4,7,8	2,3,4,6,7,8,11,12
5	1,5,6,9	1,2,5,6,7,9,10
6	2,5,6,7,9	1,2,3,4,5,6,7,8,9,10,11
7	2,4,6,7,8,11	1,2,3,4,5,6,7,8,9,10,11,12
8	4,7,8,12	2,3,4,6,7,8,11,12
9	5,6,9,10	1,2,5,6,7,9,10,11
10	9,10,11	5,6,7,9,10,11,12
11	7,10,11,12	2,4,6,7,8,9,10,11,12
12	8,11,12	4,7,8,10,11,12

For PDP algorithm, node 6 again has same forward node list F (Ø,6) = [7,2,9]. From P(6,7) = {1,3,6,7 }, we have U(6,7) = N(N(7)) – N(6) – N(7) – P(6,7) = { 10,12 }. The forward node list for 7 is F(6,7) = {11 }, Similarly , from P(6,2)= { 2,4,6,8,11} , we have U(6,2) = N(N(2)) – N(6) –N(2) –P(2,6) = Ø and , then , F(6,2) = {10 }. Therefore, the total number of forward nodes is 1+3+2 = 6. The details of P, U, B and F are represented in the following Table 2.

The total number of forwarding nodes according to PDP is in the give example is 6 including source node i.e. { 6,2,7,9,11,10 }.

In our proposed model APDP, an enhanced version of PDP, the definition of existing U has been broadened to a new U to check and exclude if it contains any adjacent nodes. The results of the proposed model are presented in Table 3

Table 2: PDP algorithm

u	v	P	U	B	F
Ø	6	Ø	1,3,4,8,10,11	2,5,7,9	7,2,9
6	7	1,3,6,7	10,12	4,8,11	11
6	2	2,4,6,8,11	Ø	1,3	[ ]
6	9	1,6,9	9	10	10
7	11	Ø	9	10,12	10
9	10	Ø	7,12	11	11

Table 3: APDP algorithm

U	V	P	U	B	F
Ø	6	Ø	1,4,8,11	2,5,7,9	7,2
6	7	1,3,6,7	10,12	4,8,11	11
6	2	2,4,6,8,11	Ø	1,3	[ ]
6	9	1,6,9	9	10	10
7	11	Ø	9	10,12	10
9	10	Ø	7,12	11	11

The lower bound as per the AMCDS (Approximation Minimum Connected Dominating Set) reduces the minimum connected dominating set to {2, 6, 7, 11} minimizing the number of forward nodes to 4. According to our proposed model number of total forwarding nodes including source node is 5 i.e. {2, 6, 7, 10, 11} where as it is 6 in PDP.

IV. Simulations Results and Simulation Parameters

We used the simulator glomosim-2.03,[8] to run the simulation Table 3 summarizes the simulation parameters we used. The simulation time was 15 minutes according to simulator clock. A total of 45 nodes were randomly placed in field of 500 X 500 m^2 and in field of 2000 X 2000 m^2. Power range of each node is 250m. We performed the simulation for AODV with DP algorithm.

Table 3 Simulation Parameters

Parameter	Value	Description
Number of nodes	160 and 40	Simulation Nodes
Field range x	500m and 2000	X-Dimension
Field range y	500m and 2000	Y-Dimension
Power range	250m	Nodes power range
Mac protocol	IEEE 802.11	MAC layer protocol
Network Protocol	AODV &Rough AODV	Network Layer
Transport Layer Protocol	UDP	Transport Layer
Propagation function	FREE-Space	Propagation Function
Node placement	Random	Nodes are distributed in random manner
Simulation time	15M	According to simulation clock
Mobility Interval	10-30sec	Pause time of node

We run each simulation four times with different node pause time varying from 15 to 30sec with a step interval of 5 sec. The graphs are presented in results section.

Results

The performance of proposed protocol is evaluated using the following metrics:

Packet delivery ratio (Throughput): Packet delivery fraction is the ratio between the number of packets originated by the application layer CBR sources and the number of packets received by the CBR sinks at the final destinations. Packet delivery Ratio is higher for AODV with Dominating Sets than that of conventional AODV (figure 3 & 7)

Number of Control Packets: The total number of control packets occurred by different nodes is less in AODV with Dominating Sets than that of conventional AODV(figure 4 & 8).

Number of Route requests: It is the number of control packets generated by all the nodes in the simulation. The number of Route Request packets is less in AODV with Dominating Sets (figure 2 & 6) compare to conventional AODV.

End-to-End Delay: The End-to-End Delay of AODV with Dominating Sets is better when compared with conventional AODV (figure 5 & 9).

Performance of AODV with Dominatin sets and AODV in the Terrain Dimensions (500, 500)

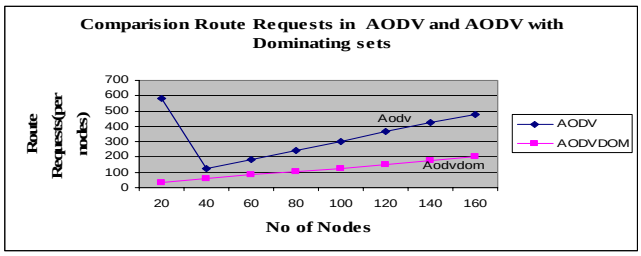


Fig 2 Route Request packet Comparison

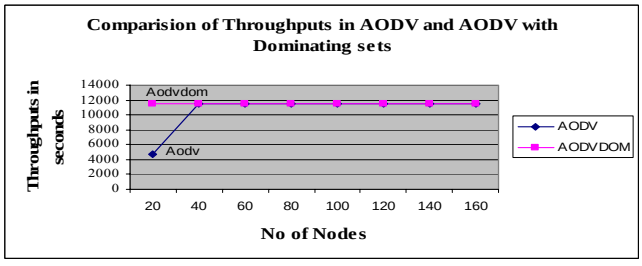


Fig 3 Throughput comparison

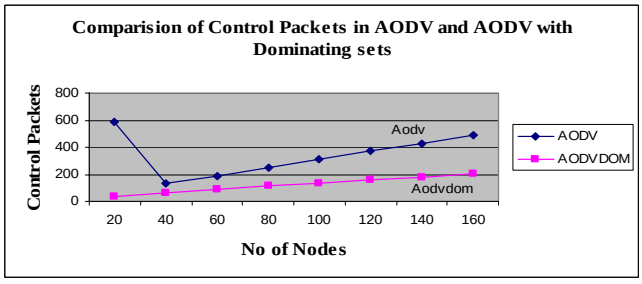


Fig4.Control Packets Comparison

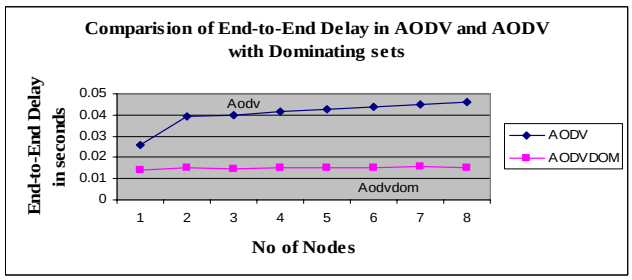


Fig 5 End-to-End Delay in Sec

Performance of AODV with dominating sets and AODV in the Terrain (2000, 2000) up to 40 nodes

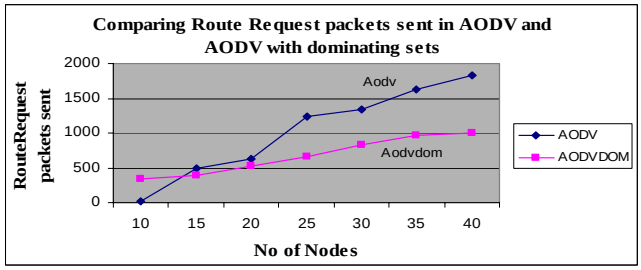


Fig 6 Route Request packet Comparision

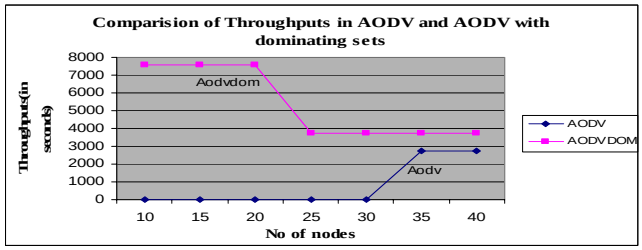


Fig 7 Throughput comparison

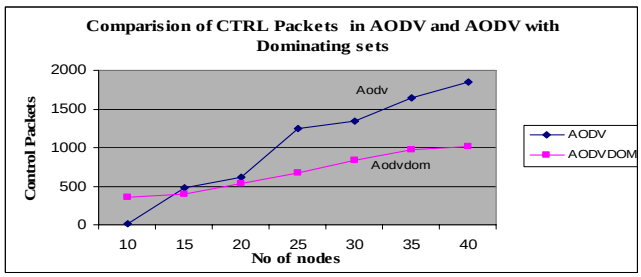


Fig 8.Control Packets Comparison

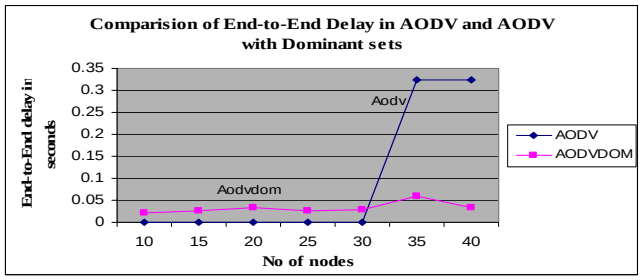


Fig 9 End-to-End Delay in Sec

V. Conclusions and Future Work

We presented an enhanced dominant pruning approach that allows pruning redundant broadcasts even more than the conventional dominant pruning heuristic. Redundant broadcasts increase the number of packet collisions, and consequently delay the response for RREQs in the route discovery process. DP is shown to reduce the number of broadcast transmissions when compared to standard DP. Because DP requires the two-hop neighborhood to determine the forwarder list, we built a neighbor protocol as part of AODV. By making the neighbor protocol part of AODV, the result is a more accurate view of the local topology, and therefore more accurate is the determination of the forwarder list. We also proposed a APDP algorithm which is an enhanced version of PDP. Here we have shown the performance comparison of AODV with DP algorithm. Future scope of the work is to compare to the performance of AODV with PDP algorithm and compare the performance of both.

REFERENCES

[1] Brad Williams and Tracy Camp. Comparison of broadcasting techniques for mobile adhoc networks. In *MobiHoc '02: Proceedings of the 3rd ACM international symposiumon mobile ad hoc networking & computing*, pages 194–205. ACM Press, 2002.

[2] C. Perkins. Ad-hoc on-demand distance vector routing. In *Proceedings of the IEEE workshop on mobile computing systems and applications*, pages 90–100, 1999.

[3] Charles E. Perkins and Elizabeth M.Royer. Ad-hoc on-demand distance vector routing. Technical report, Sun Micro Systems Laboratories, Advacnced Developmenet Group, USA.

[4] David B Johnson and David A Maltz. Dynamic source routing in ad hoc wireless networks.In *Mobile Computing*, volume 353, pages 153–179. Kluwer Academic Publishers,1996.

[5] Dipti Joshi, Sridhar Radhakrishnan, and Chandrasekheran Narayanan. A fast algorithm for generalized network location problems. In *Proceedings of the ACM/SIGAPP symposium on applied computing*, pages 701–8, 1993

[6] H. Lim and C. Kim. Flooding in wireless ad hoc networks. *Computer Communications*, 24(3–4):353–363, February 2001

[7] Jie Wu and Fei Dai. Broadcasting in ad hoc networks based on self-pruning. In *IEEE INFOCOM*, pages 2240–2250, 2003.

[8] Jorge Nuevo, "A comprensible glomosim tutorial", March4,2004 http://appsrv.cse.cuhk.edu.hk\~hndai\research.

[9] Katia Obraczka, Kumar Viswanath, and Gene Tsudik. Flooding for reliable multicast in multi-hop ad hoc networks. *Wireless Networks*, 7(6):627–634, 2001.

[10] Marco A. Spohn and J. J. Garcia-Luna-Aceves. Enhanced dominant pruning applied to the route discovery process of on-demand routing protocols. In *Proceedings of the 12th IEEE international conference on computer communications and networks*, pages 497–502, October 2003

[11] Mobile Ad Hoc Networking:Routing Protocol Performance Issues and Evaluation Considerations Request For Comments 2501 available at http://www.ietf.org

[12] Ning.P. and Sun.K. How to misuse aodv: a case study of insider attacks against mobile ad-hoc routing protocols. Technical report, Comput. Sci. Dept., North Carolina State Univ., Raleigh, NC, USA, 2003.

[13] Peng-Jun Wan, Khaled M. Alzoubi, and Ophir Frieder. Distributed construction of connected dominating set in wireless ad hoc networks. In *IEEE INFOCOM*, pages 1597–1604, June 2002.

[14] Rajive Bagrodia. README GloMoSim Software. University of California, Los Angeles - Department of Computer Science. Box 951596, 3532 Boelter Hall, Los Angeles-CA 90095-1596 / rajive@cs.ucla.edu.

[15] Teresa W. Haynes, Stephen T. Hedetniemi, and Peter J. Slater, editors. *Fundamentals of Domination in Graphs*. Marcel Dekker, Inc., 1998.

[16] Xiaoyan Hong, Kaixin Xu, and Mario Gerla. Scalable routing protocols for mobile ad hoc networks. 2002.

[17] Yu-Chee Tseng, Sze-Yao Ni, Yuh-Shyan Chen, and Jang-Ping Sheu. The broadcast storm problem in a mobile ad hoc network. *Wireless Networks*, 8:153–167, 2002. H. Lim and C. Kim. Flooding in wireless ad hoc networks. *Computer Communications*,24(3–4):353–363, February 2001-04-15).

[18] Z. Haas. A new routing protocol for the reconfigurable wireless network. In *Proceedingsof the IEEE international conference on universal personal communications*, pages 562–566, October 1997.

[19] A Nagaraju , Dr S.Ramachandram "Adaptive Partial Dominating Set Algorithm For Mobile Ad-Hoc Networks" , Presented in the international conference ACM ,COMPUTE-09, Jan 9th -11th,2009,Banglore.

[20] J.Hass and M.R.Pearlman, "The Zone Routing Protoco(ZRP) for Ad Hoc Networks," draft-ietf-manet-zone-zrp-02.txt,work in progress,June 1999

[21] T.Clausen, P.Jacquet,A.Laouiti, P.Muhlethaler,A. Q and L.Viennot. Optimized link state routing protocol for ad-hoc networks. In Proceedings of the IEEE internat multi topics conferences. 2001

Bee-Inspired Routing Protocols for Mobile Ad HOC Network (MANET)

Deepika Chaudhary
Chitkara Institute of Engg & Technology Jhansla
Teh. Rajpura,Distt Patiala-140401,Punjab
deepika.chaudhary@chitkara.edu.in

***Abstract* – Mobile AdHoc networks(MANETs) are receiving a significant amount of attention from researchers. This paper provides a high light on a new and very energy efficient algorithm for routing in Manets BeeAdHoc. This algorithm is a reactive source routing algorithm and consumes less energy as compared to other existing state of art routing algorithms because a fewer control packets for routing are sent as compared to other networks.**

I. Introduction

Mobile Ad HOC networks (MANET'S) are networks in which all nodes are mobile and communicate with each other via wireless connections. Nodes can join or leave the network at any time. There is no fixed infrastructure. All nodes are equal and there is no centralized control or overview. There no designated routers all nodes can serve as routers for each other and data packets are forwarded from node to node in multihop fashion.

Since a few years researcher's interest in MANETS have been growing and especially the design of MANET routing protocols has received a lot of attention. One of the reasons is that routing in MANETS is particularly challenging task due to the fact that the topology of the network changes constantly and paths, which were initially efficient, can quickly become inefficient or even infeasible. Moreover control information flow in the network is very restricted. This is because the bandwidth of the wireless medium is very limited and the medium is shared: nodes can only send or receive data if no other node is sending in their radio neighborhood. It is therefore important to design algorithm that are adaptive robust & self-healing. Nature self-organizing systems like insect societies show precisely these desirable properties. Making use of a number of relatively simple biological agents like ants, a variety of different organized behavior are generated at the system level from the local interaction among the agents and with the environment. The robustness and effectiveness of such collective behaviors with respect to variations of environment conditions are key aspects of their biological success. This kind of systems are often referred to with the term swarm intelligence. Swarm systems have recently become a source of inspiration for design of distributed & adaptive algorithms.

II. Design Issues of Routing Algorithms

The most important challenge in designing algorithms for MANETs are mobility and limited battery capacity of nodes. Mobility of nodes results in continuously evolving new topologies and consequently the routing algorithms have to discover or update the routes in real time but with small control overhead . The limited battery capacity requires that the packets if possible be distributed on multiple paths , which would result in the depletion of batteries of different nodes at an equal rate and hence as a result the life time of networks would increase[1]. Therefore an important challenge in Manets is to design a routing algorithm that is not only energy effieicent but also delivers performance same or better than existing state of art routing protocols.

III. Classification of Routing Algorithms in Manets

The routing algorithms for Manets can be broadly classified as proactive algorithms or reactive algorithms. Proactive algorithms periodically launch control packets which collect the new network state and update the routing tables accordingly. On the other hand, reactive algorithms find routes on demand only. Reactive algorithms looks more promising from the prespective of energy consumption in Manets. Each category of above mentioned algorithms can further be classified into host intelligent or router intelligent algorithms . A few reactive algorithms are DSR (Dynamic Source Routing) which is host intelligent algorithm while AODV(AdHoc On Demand Distance Vector Routing) which is a router –intelligent algorithm.However these algorithms are not designed for energy efficient routing. Here in this study an deep insight is provided on BeeAdHoc which delivers performance same as better than that of DSR,AODV but consumes less energy as compared to them. The algorithm achieves these objectives by transmitting fewer control packets and by distributing data packets on multiple paths.

IV. Existing Work on Nature –Inspired Manet Routing Protocols

The first algorithm which presents a detailed scheme for MANET routing based on ant colony principles is ARA [3]. The algorithm has its roots in ABC AND AntNet Routing algorithms for fixed networks and are inpired by the pheromone laying behavior of ant colonies . AntHocNet, which is hybrid algorithm having both reactive and proactive components have also been proposed.This algorithm tries to keep most of features of the original AntNet and shows promising results in simulation comared to AODV Termite is another MANET routing algorithm inspired by termite behaviour .Here no special agents are needed for updating the routing tables rather data packets are delegated this task.

V. Overview of BeeAdHoc Architecture

BeeAdHoc is an on-demand multi path routing algorithm for mobile adhoc networks inspired from the foraging principles of honey bees[4]. BeeAdHoc works with types of agents: packers, scouts foragers and swarms. The packers locate a forager and hand over the data packet to the discovered forager. Scouts discover new routes from the launching node to the destination

doi:10.4304/jetwi.2.2.86-88

node through broadcasting principle and an expanding time to live (TTL) timer. Foragers, the main workers.

The architecture of the hive is shown in Fig 1 where the entrance floor is an interface to the lower MAC layer, while the packing floor is an interface to the upper transport layer. The dance floor contains the foragers and the routing information to route locally generated packets. The functional characteristics of each floor composing the hive are explained in the following.

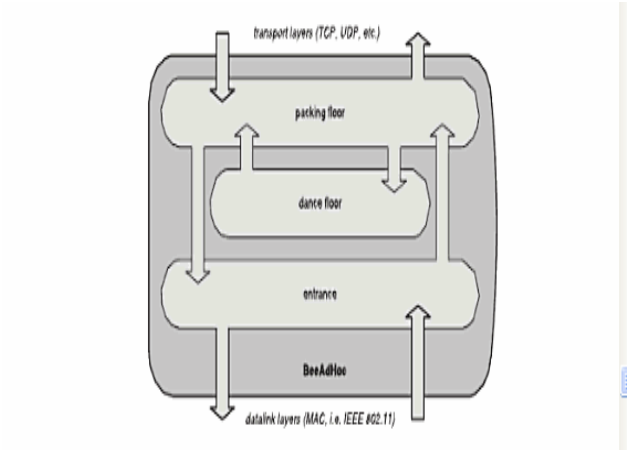


Fig 1. Overview of BeeAdHoc Architecture

VI. Packing Floor

The packing floor is an interface to the upper transport layer (e.g, TCP or UDP) . Once a data packet arrives from transport layer, a matching forager for it is looked up on the danced floor . If a forager is found then the data packet is encapsulated in its payload . Otherwise, the data packet is temporary buffered waiting for a returning forager. If no forager comes back within a certain predefined time, a scout is launched which is responsible for discovering new routes to the packet destination . Figure 2 explains the series of action performed at packing floor.

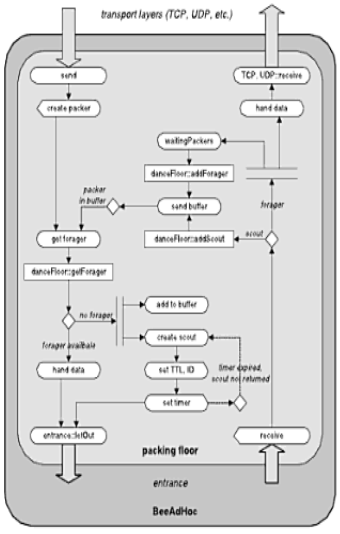


Fig 2 : The Packing Floor

VII. Entrance

The function performed in the entrance are shown in figure .3. The entrance is an interface to the lower level MAC layer. The entrance handles all incoming and out going packet. Action on the dance floor depends on the type of packet entered the floor from the MAC layer. If the packet is the forager and th e current node is its destination, then the forager is forwarded to the packing floor.; otherwise, it is directly routed to the MAC interface of the next hop node. If the packet is a scout, it is broadcast to the neighbor nodes if its TTL timer has not expired yet or if the current node is not its destination. The information about the ID of the scout and its source node is stored in a local list. If a replica of previously received scout arrives at the entrance floor , it is removed from the system. If a forager with the same destination as the scout already exists on the dance floor, then the forager's route to the destination is given to the scout by appending it to the route held so far by the scout.

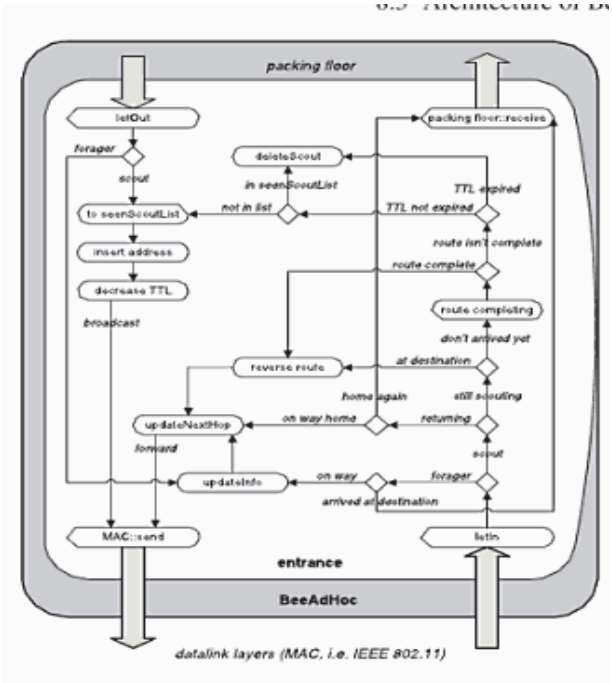


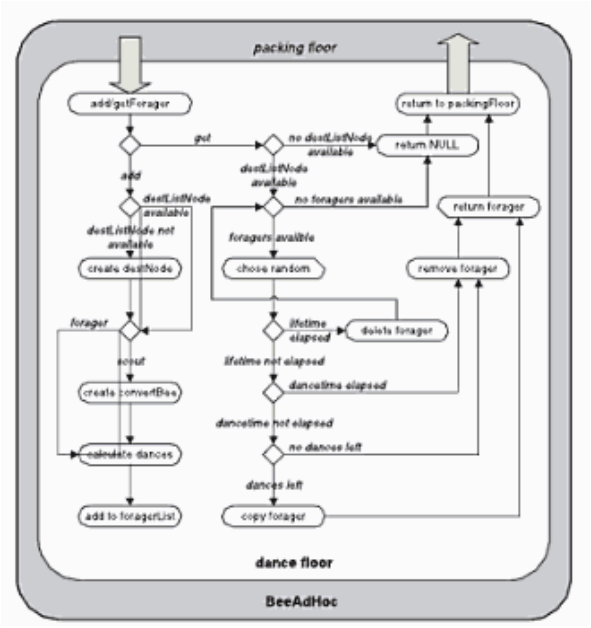
Fig 3: The Entrance

VIII. Dance Floor

The dance floor is the heart of the hive because it maintains the routing information in the form of foragers. The dance floor is populated with routing information by means of mechanism reminiscent of the waggle dance recruitment in natural bee hives once a forager returns after its journey, it recruits new forager by "dancing" according to the quality of the path it traversed.

A lifetime forager evaluates the quality of its route based on the average remaining battery capacity of the nodes along its route. The central activity of the dance floor module consists of sending a matching forager to the packing floor in response to a request from a packer. The foragers whose lifetime has expired are not considered for matching. If multiple path be identified for matching, then a forager is selected in a random way. This helps in distributing the packets over multiple paths, which in turn serves two purpose : avoiding congestion under high loads and depleting batteries of different nodes at comparable rate. A clone of the selected forager is sent to the packing floor and the number of permitted clones. If the dance number is 0 then the original forager is sent to the packing floor removing it in this way from the dance floor. This strategy aims at favoring young over old foragers.If the last forager for a destination leaves a hive then the hive does not have any more route to the destination. Nevertheless if a route to the desination still exists then soon a forager will be returning to the hive if no forager comes back within reasonable amount of time, then the node has

probably lost its connection to the destination node. In this way fewer control packets are transmitted resulting in less energy expenditure.



IX. Conclusion

The simplicity of BeeAdHoc which results from its simpler architecture and its using a smaller number of control packets pay off once we look at the energy consumption in transpoting the packets from their source to their destination. BeeAdHoc network employs a simple bee behavior to monitor the validity of the routes.When compared to other algorithms the battery level of BeeAdHoc is better because it tries to spread the data packets over different routes rather then always sending them on the best routes.

X. Future Aspects

Design and development of routing protocols for Mobile Ad Hoc Networks (MANETs) is an active area of research. The standard practice among researchers working in this emerging domain is to evaluate the performance of their routing protocols in a network simulator. In future a mathematical models of two key performance metrics, routing overhead and route optimality, of *BeeAdHoc* MANET routing protocol using a simulator can be studied . One of the key components of our *BeeAdHoc* model is the collision model at Medium Access Control (MAC) layer. To design the mathematical expressions of the performance metrics can also provide insight about the behavior of *BeeAdHoc* in particular, and a typical ad hoc routing protocol in general.

References

[1] Formal Modeling of *BeeAdHoc* : A Bio-inspired Mobile Ad Hoc Network Routing Protocol, Muhammad Saleem[1], Syed Ali Khayam[2] and Muddassar Farooq[3].

[2] The wisdom of the hive applied to mobile ad-hoc networks Wedde, H.R.; Farooq, M.

[3] BeeAdHoc: an Efficient, Secure and Scalable Routing Framework for Mobile AdHoc Networks , Prof. Dr. Horst F. Wedde (wedde@ls3.cs.uni-dortmund.de) ME Muddassar Farooq (muddassar.farooq@ls3.cs.uni-dortmund.de)

[4] M. Farooq, “From the Wisdom of the Hive to Intelligent Routing in Telecommunication Networks” PhD Thesis, Dortmund University, 2006.

[5] H. F. Wedde and M. Farooq. “The wisdom of the hive applied to mobile ad-hoc networks.” In Proceedings of the IEEE Swarm Intelligence Symposium, pages 341–348, 2005.

[6] P. K. Visscher, “Dance Language”, Encyclopedia of Insects, Academic Press, 2003

Distance and Frequency based Route Stability Estimation in Mobile Adhoc Networks

Ajay Koul
SMVD University/ School of computer Science and Engineering, Katra, India
Email:ajay.kaul@smvdu.ac.in

R. B. Patel and **V. K. Bhat**
M.M.University/Department of Computer Science and Engineering, Ambala, India
SMVD University/School of Applied Physics and Mathematics, Katra, India
Email :{ patel_r_b, vijaykumarbhat2000}@yahoo.com

Abstract-**This paper discusses the link stability estimation for Mobile Ad hoc Networks (MANETs). In this approach the total time for which the link remains connected with the neighboring nodes is estimated. This helps to predict the stability of the route which is required to forward the packets to the destination. This method does not use the selection of the next hop on the basis of shortest distance, but is based on the time period for which the next hop link remains connected. The parameters we use, are the distance and frequency and signal quality. These provide the way for a node to decide the best next hop neighbor and hence a perfect Quality of Service (QoS) is also obtained.**

Index Terms— **Route stability, frequency, distance, MANETs, QoS.**

I. Introduction

An ad hoc network is a dynamic multihop wireless network that is established by a set of mobile nodes. Such networks are, therefore, suitable for the environments where it is a difficult to create a fixed infrastructure. In this network, mobile nodes randomly move and communicate over radio channels. If two mobile nodes are in a radio transmission range, they can communicate with each other directly, otherwise, the source node sends/receives the packets via some intermediate nodes. Hence a proper routing algorithm is required to route the packets from source to destination. Most routing algorithms, like in [1] estimate the link stability based on the distance between the two neighboring nodes. In [2] the electric field as a parameter is used to find the stable route to the destination. The algorithm given in [3], finds the best route to the destination based on link perdurability. These, however, do not lead to the reliable solution in Mobile Ad hoc Networks (MANETs), as the environments are dynamic and hence the route estimation depends on the factors like mobility and direction as well. In this paper we have taken these factors into account and the next hop route is selected not on the basis of distance only but on mobility as well. The quality of strength is also taken into consideration. Our results show that our method can also be applied in the selection of the best neighboring node.

The rest of the paper is organized as follows.
In Section II we discuss the related works. In Section III, the node distance, position and mobility is estimated. The quality of signal along with the algorithms to be executed on intermediate and destination nodes are presented in Section IV. In Section V the practical scenario of estimating the parameters like distance, frequency, mobility, RSSI, etc. and their performance results are obtained. Finally the article is concluded in Sections VI.

II. Related work

Many routing protocols in the past have been proposed on route stability. In [4] the technique of signal stability is used for adaptive routing in Mobile Ad hoc networks. In this approach the on demand longer lived routes are discovered based on signal strength and location stability. The signal strength of the neighboring nodes are detected by sending beacons, and based on that, the classification of strong and weaker channels are established. This is further strengthened by choosing a channel which has existed for a longer period of time. This signal strength feature in combination with the location stability helps in selection of the next hop neighbor and hence the overall route is established. The protocol mentioned even though gives good results in establishing stable route, however, fails when the node density increases or the node mobility increases. The Global positioning system based reliable route discovery proposed in [5] used a different approach. The algorithm discovers routes based on two zones, the stable zone and the caution zone. The zones are decided based on the location and the mobility of the nodes using the Global positioning system (GPS). The stable zone and the caution zone change dynamically depending on the mobile nodes speed and direction information. The mobile nodes speed, direction and position is estimated based on GPS system. This method is the perfect method of determining the stability of the route as the critical parameters like direction and mobility is determined. The method mentioned however, has a serious drawback of additional cost involvement and more power requirement for the GPS based device. In [6] a model to find the best

doi:10.4304/jetwi.2.2.89-95

route based on link lifetime has been proposed. In this model the edge effect has been explored to find more stable route to the destination. The lifetime and the stability of the route is calculated by eliminating the edge effect to reduce route maintenance and route overheads. The disadvantage of this model however is that each method proposed under the model requires either pilot signal generation and monitoring of the pilot signal of the other nodes or monitoring of signal strength of the other nodes. With these weaknesses, the stable route still gets established and performance of the network also gets increased. Lifetime Prediction Routing (LPR) proposed in [7] also was suggested to find the stability of the route. In this the service life of the MANET is predicted and correspondingly the route is established. The service lifetime is predicted based on the past activity of the battery lifetime of the node. A simple moving average predictor is used to keep track of last N values of residual energies and the corresponding time instances for the last N packets received/ relayed by each mobile node. This way LPR not only captures the remaining battery capacity but also accounts for rate of energy discharge. This way the route stability and lifetime is estimated based on the battery life which is also an important factor. This method has some serious concerns like, when the node mobility increases it becomes difficult to predict the lifetime of the node and also the use of LPR involves certain overheads. One more model to predict the link stability was proposed in [8]. This model was named as the signal stability based routing protocol (SSA^+) model. SSA^+ is the enhanced version of SSA for finding the route stability in MANETs. In this method a route is maintained with the help of active neighbouring nodes. The neighbours are considered active if it relays or originates at least one packet within the most recent active timeout period. The SSA^+ solution was mainly offered to remove the problems of high node density and high mobility and the low node density and low mobility. In high node density and high mobility scenario the mobility is high, the probability of link failures remain also high. To cater to the problem the signal strength of the link is estimated and classified into the categories of weak, normal and strong signals. The nodes exchange information regarding the signal strength by sending the hello packets and based upon the signal strength value stored in the link state table, the life time of the route is determined. In low node density and low mobility the signals of the nodes remain weaker and the stability and the route lifetime is maintained based on two important conditions.

1) The distance between any two nodes gets shorter for the past few clicks.
2) Secondly the distance between any two nodes gets longer but the distance changes larger slowly for the past few clicks (the condition of low mobility).

The method above mentioned above, even though showed better results compared to SSA in terms of low node density and low mobility and high node density and high mobility has the drawback of predicting the lifetime of the route in a complicated way as the node has to maintain the Link stability table which adds lots of overheads in terms of route control and maintenance and which are difficult to maintain. In [9] [10] and [11], algorithms are proposed as DV-Hop, Hop-TERRAIN, and link stability with dynamic delay prediction, to determine the location of nodes based on hop counts and the appropriate route to the destination. The hop counts provided an estimate for the overall distance between the nodes and the dynamic delay ensured stability. These however, were mainly focusing on the Quality of service (QoS) based route establishment and hence were silent on the route lifetime parameters like mobility and battery power. A Novel route metric based on the fragility of the route was also proposed in [12]. In this approach the dynamic nature of the route is captured by studying the distance variation of the next hop neighbor in terms of expansion and contraction. A distributed algorithm is executed in every node to calculate the relative speed estimate of the neighbours. The destination node gets the information of every route and selects the best route based on the route fragility coefficient (RFC) which in turn depends on the cumulative expansion metrics (CEM) and cumulative contraction metrics (CCM). This approach finds the stability of the route and is good in terms of not requiring time bound measurements, like the global positioning system. This method however, fails to predict the node mobility and includes the destination latency. The estimation of link quality and residual time in vehicular Adhoc networks proposed in [13] is also the method to determine stability of the route in MANETs by predicting the active time of the link. In this method the signal processing along with empirical decomposition and robust regression is used to predict the link quality and the residual time. This method unlike the method proposed in this paper uses a three stage approach which involves lots of complexity in terms of practical deployment. This also involves additional burden of calculating the various essential parameters mentioned under signal processing, empirical decomposition and robust regression methods. The above mentioned literature even though provides excellent way of estimating the distance and the location of neighboring nodes through different techniques, however, are mostly silent on the mobility issue. This paper provides the route stability estimation based upon the neighboring node distance ,mobility and signal strength.

III. DESCRIPTION

Mobile Ad hoc Networks (MANETs) being dynamic in nature create challenges in terms of Quality of Service (QoS) at each level from application to physical. In physical layer level the mechanism to predict the lifetime of the neighboring node increases reliability from source to destination delivery. To achieve this we have identified the following three parameters for next hop path selection

1) Node Distance and position
2) Mobility
3) Signal quality

A. Node Distance and position

Let two circles with radii R_1 and R_2 intersect in a region as shown in Fig. 1 Let the circles be centered At A (0, 0) and B(d, 0). Then equations of the circles are

$$x^2 + y^2 = {R_1}^2 \tag{1}$$

$$(x-d)^2 + y^2 = {R_2}^2 \tag{2}$$

This implies that

$$x^2 - 2dx + d^2 - x^2 = {R_2}^2 - {R_1}^2 \tag{3}$$

$$x = \left(\frac{d^2 - {R_2}^2 + {R_1}^2}{2d} \right) \tag{4}$$

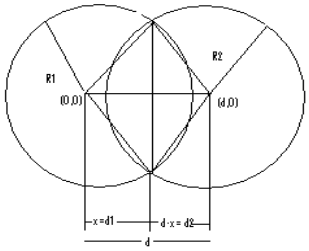


Figure 1. Showing the wireless range intersection.

Now y is half of the length of chord connecting the two circles, we have

$$y^2 = {R_1}^2 - x^2$$

$$y^2 = \frac{4d^2{R_1}^2 - (d^2 - {R_2}^2 + {R_1}^2)^2}{4d^2} \tag{5}$$

To find out d_1 and d_2 the distances from centre on nodes to the circle intersection point, we use the relation

$$d_1 = \frac{(d^2 - {R_2}^2 + {R_1}^2)}{2d} \tag{6}$$

$$d_2 = \frac{(d^2 + {R_2}^2 - {R_1}^2)}{2d} \tag{7}$$

To find the distance d, and position parameters x, y we assume the following:

1) That all the nodes are of equal strength and technical specifications
2) That the radii are known and are same for all the nodes because of same transmission power.

With these assumptions and from (1-7), the distance can be found using the method as given in [14], known as round trip time based method. This works on the concept that every data packet can be acknowledged. Under this work the time span from the moment at which a packet starts to occupy the wireless medium to the time at which the immediate acknowledgment is received is measured and denoted by T_R. The time duration between the reception of a data packet and issuing the corresponding immediate acknowledgment is also measured and denoted by T_L. The distance is computed based on the relation between the distance traveled and the speed of light as follows:

$$d = \frac{(TR - TL)}{2} \times C \tag{8}$$

Where C= 3×10^8 is the speed of light. Now solving for d_1, d_2, x, y from (7) and (8), the position P (α, β) of the neighboring node can be estimated. Now link stability L_s is, given by:

$$L_S = \begin{cases} 1, if\ d < \mu\ and\ |AP| \leq |AC| \\ 0, if\ d < \mu\ and\ |AP| \leq |AC| \end{cases}$$

$$\text{Where } AP = \sqrt{\alpha^2 + \beta^2}$$

$$AC = \sqrt{m^2 + n^2}$$

and μ and C(m, n) are respectively the maximum permissible distance and maximum co-ordinate position allowed to communicate between the nodes

B. Mobility

Measurement of distance and position as mentioned above are easy to compute and provides the route stability especially when the nodes are static. In case the nodes are dynamic the mobility plays an important factor. The distance and position in mobile environment will provide the feasibility of communication but will not provide the life time of the route with the neighboring node. To find mobility, we calculate the frequency of the neighboring node from the following equations using [10].

$$f_r = f_e \left[\frac{v}{v + v_{sr}} \right] \tag{9}$$

Where in (9), f_r is the received frequency, f_e is the emitted frequency, v is the speed of the waves in the medium and v_{sr} is the radial component of the velocity of the neighboring node with respect to the medium (positive if moving away from the observer, negative if moving towards the observer). A similar analysis for a moving observer and a stationary source yields the observed frequency from the following equation (the receiver's velocity being represented as v_r):

$$f_r = f_e \left[\frac{v}{v + v_r} \right] \tag{10}$$

where the same convention applies. We note that v_r is positive if the observer is moving away from the source and negative if the observer is moving towards the

source. These can be generalized into a single equation with both the source and receiver moving as given below:

$$f_r = f_e\left[1 - \frac{v_{sr}}{v + v_{sr}}\right] \quad (11)$$

Where v_{sr} is the source to receiver velocity radial component. Now since the source nodes have equal power so frequency of transmission will remain same for all the nodes within the vicinity. From (11) above we get the velocity of the neighboring nodes as

$$v_{sr} = v\left[\frac{f_e - f_r}{f_r}\right] \quad (12)$$

Now overall time period for which the link remains established will depend on the following relation

$$T = \frac{d}{v_{sr}} \quad (13)$$

where T in (13) is defined as the time period for which the link with the neighboring node will remain established. The time period depends on d because if the node covers the maximum distance away from the other node, then a small velocity of it in the opposite direction will disconnect the link. The time T will be more if it is negative as the velocity will be having +ve or –ve sign depending upon the direction of motion.

C. Signal Quality

The distance, and position plays an important role in finding mobility and hence the link stability. However, the stability of the link also depends on the quality of signal as well. In Mobile Adhoc Networks (MANETs) the signal quality plays an important role in selecting the route to the destination. The parameters like distance and mobility proposed above effects the signal quality, however in spite of the feasible distance and mobility threshold values, the node may not receive from the neighboring node the good quality signal due to noisy surroundings. When there is a signal transmission from a node with certain power in the noisy environment, the Bit error rate of the signal increases apart from the normal path loss in the medium. The receiver on the basis of the receiver sensitivity, and the threshold *Signal to Noise Ratio* (SNR), predicts the quality of signal. Overall it is the Receiver signal strength indicator (RSSI) which provides the details of the quality of signal received. RSSI which is basically a measurement of how well the radio is receiving or 'hearing' data to determine the quality of signal received. It's typically measured in -dBm, which is the power ratio in decibel (dB) of the measured power referenced to one milliwatt (mW). Normally in the real testing environments the RSSI above -60 dB is considered the threshold required to perform good networking functions and any value higher than that is the stable value. We denote this value by M_{RSSI}. Therefore overall link stability which is denoted by L_{SSO} should satisfy L_S, T *and* M_{RSSI} the three important parameters to determine the stability of the route.

IV. ROUTING

The above parameters calculated can be used to enhance the routing procedures already available. To inco-orporate the changes, the nodes needs to be modified. Below are the node level modifications suggested when the node behaves as, the intermediate and the destination node.

A. Intermediate node operation

When the intermediate node receives the RREQ packet from the source or from any of its neighbors it provides the additional information of the distance and the frequency in the RREQ packet and forwards it to the destination. The Algorithm 1 provides the operational details to be performed at each intermediate node.

Algorithm 1 Algorithm to be executed in the node

1 Set μ , (m, n), M_{RSSI}
  Set k stable link time threshold
2 Measure d
3 if $d < \mu$ $L_s = 1$ Goto x
  Else $d > \mu$ $L_s = 0$ stop.
3 x: calculate RSSI
4 if RSSI < M_{RSSI} goto 8 else
5 calculate v_{sr}
5 if vsr -ve calculate T
  Else if v_{sr} +ve calculate T
6 if $T < k$ reject
6 Else if $T > k$ select and modify RREQ
7 stop

B. Destination node operation

The destination node on receiving RREQ packets from the several nodes sends the reply to the source by selecting the best path on the basis of the T, M_{RSSI} value. Algorithm 2 provides the process details of the destination node when RREQ arrives from different nodes of the destination node.

Algorithm 2 Process execution when RREQ arrives

1 Analyze RREQ
2 Extract, d, v_{sr} M_{RSSI}^*
3 Compare
  d with d^* and M_{RSSI} *with* M_{RSSI}^* and v_{sr} with v_{sr}^*
4 if $d > d^*$ and $M_{RSSI} < M_{RSSI}^*$ and $v_{sr} > v_{sr}^*$
  Reject Route
  Else select Route

The source node when sending the RREQ packet, calculates before, the distance, mobility and SNR of the neighboring nodes. It then embeds the information in the RREQ packet and sends it to the destination via. Next hop neighbor.

V. RESULTS

The experiments were performed in the indoor environments. In the experiment, we used one wireless access point namely Wireless-G WAP54Gv.2 of Linksys make and a WLan laptop. The access point and the Laptop were supporting the wireless 802.11 b/g model protocols. For sending the ICMP packets, we used a Graphic tool namely Multiping Grapher to send the ping requests every second.The payloads in bytes were kept constant at 512bytes. The purpose of doing so was to see the graphic outcome of the response. Figures 2 and 3 show the details of the ping at a max of 6ms and 7ms by varying the indoor distance from maximum to minimum and vice versa.

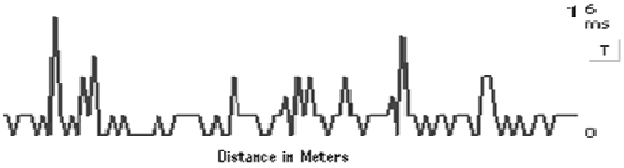


Figure. 2 Ping Response at 6ms Max

The physical and the operational mode rates have been fixed at 802.11g and 54mbps, 11mbps. The distance *d* as calculated in (8) is measured from the tool Snuffle. From this tool the T_R and T_L is computed from the MAC time stamps recorded on the transmitted request and received acknowledgment packets and by subtracting the MAC time stamps of the received reply and the issued acknowledgment packets.

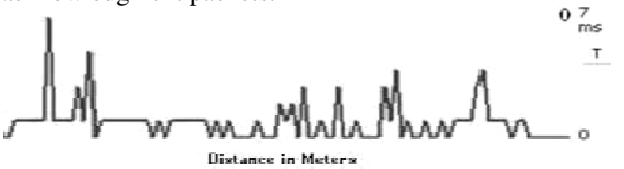


Figure. 3 Ping Response at 7 ms Max.

Figure 4. shows the distance calculated vs the difference in MAC delay. The calculated distance is however not the actual physical distance. The intension here is to check the distance between the nodes and to compare it with the threshold maximum allowable radial distance in this case we have fixed as μ . Now that the distance between the nodes is calculated, the mobility can be calculated by specifying the frequency of the source at rest which is 2.4 GHz and wave velocity 300,000 Kms/second. The experiments were carried in the indoors were the speed of the access point were varied and the frequency it was calculated by using the above equations and the Doppler's effect calculator.

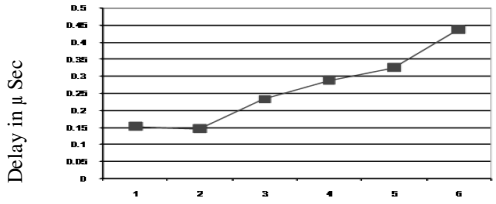


Figure. 4 Calculated distance vs Delay

TABLE I.
Velocity and Frequency Recorded Readings

S.No	velocity in m/s	Frequency in GHz
1	-2	2.4140
2	-3	2.4212
3	-4	2.4283
4	-5	2.4356
5	-6	2.4421
6	-7	2.4497
7	-8	2.4573
8	-9	2.4643
9	-10	2.4719
10	-11	2.4793
11	-12	2.4868

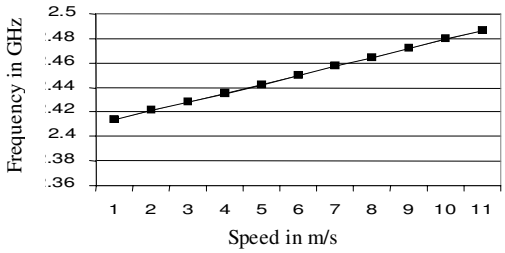


Figure. 5 Frequency vs Speed

The data collected are as shown in Figure 5. The velocities in the table 1 above are shown negative this indicates that the node is approaching the other node which is recording the frequency. By calculating the velocity we can find the link stability as per the equations given above in (13) and (14).

A. Received Signal Strength Indicator

To calculate the link stability based on RSSI parameter, we used a graphical tool called Wirelessmon. This tool provided the details of the RSSI Value recorded and the signal strength percentage which is proportional to the RSSI value. The more the percentage of signal strength the more the RSSI value closer to zero. Some of the Graphs plotted are shown from Fig. 6 to Fig. 8. The graphs clearly show that the percentage of signal received is good when the distance is only 1m and hence the RSSI value approaching –ve value towards zero. As soon as the node is moving away from the access point and some aluminum wall partition included in between, It shows the decline in the signal and hence in the RSSI value as well.

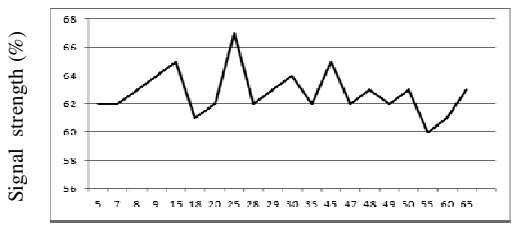


Figure. 6 Signal strength % vs time in seconds

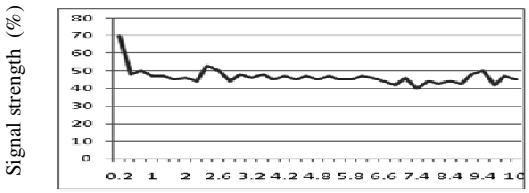


Figure. 7 Signal strength % vs Time in seconds distance =3m and one aluminium partition in between the nodes

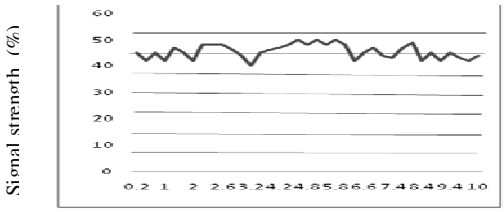


Figure. 8 Signal Strength % vs. time in sec with distance = 6m and two aluminum partitions between the nodes.

VI. CONCLUSIONS

We have presented an approach on how to measure the distance and position based on the propagation time of IEEE 802.11 packets. We used commercial WLAN card of linksys make supporting IEEE 802.11b and 802.11g. The distance and the position measure however is logical and has no relation with the physical distance or position in the co-ordinate system. The frequency measurement results reveal the mobility of the neighboring nodes. The distance parameter can be used to detect the location of the node especially when two nodes are static and is therefore, sufficient to determine next hop link stability. The frequency measurement to determine the logical mobility is however important to determine the next hop route stability if any one or both of the nodes are mobile. The signal quality also adds the QoS factor. This paper in spite of exploring three important factors like distance and mobility RSSI value however is silent on the other factors like S/N ratio, BER and overall signal quality. The distance factor can somehow presume the strength of signal but cannot provide the details of the quality parameters like S/N ratio, BER. These parameters S/N ratio and BER combined with the above mentioned parameters like distance and frequency may be used to find the best route life in near future. The algorithms depicted also will be incorporated in the nodes and a routing protocol will be used to see the performance enhancement of this method over others.

REFERENCES

[1] Nam T. Nguyen, Ady Wang, Peter Reiher, Geoff Kuenning " Electric Field Based Routing : A Reliable Framework for Routing in MANETs ", *ACM Mobile Computing and Communications Review,* USA, 8(2): 35-49, 2004.

[2] Sun Xuebin, Zhou Zheng, " Link Perdurability based Routing for Mobile Ad hoc Networks ", *Journal of Electronics,* China, 20(4): 299-304, 2004.

[3] M. Royer and C.-K. Toh, "A Review of current Routing Protocols for Ad Hoc Mobile Wireless Networks", *IEEE Personal Communication Magazine,* 6(2): 46-55, 1999.

[4] Rohit Dube, Cynthia D. Rais, Kuang-Yeh Wang, and Satish K. Tripathi, "Signal Stability-Based Adaptive Routing (SSA) for Ad Hoc Mobile Networks," *IEEE Personal Communications,* 4(1): 36-45 1997.

[5] Young Joo Soh, Won Kim, Dong Hee kwon, "GPS based Reliable Routing algorithms for Adhoc networks" *in Proceedings of IEEE International Conference on Communications,* Helsinki, Finland, June 11- 14, 2001, pp.3191-3195

[6] G. Lim, K. Shin, S. Lee, H. Yoon and J. S. Ma, "Link Stability and route lifetime in ad hoc wireless networks", in *Proceedings of International Conference on Parallel Processing workshop (ICPPW'02),* Vancouver, BC, Canada, August 18-21, 2002, pp. 116-123.

[7] M. Maleki, K. Dantu, and M. Pedram, "Lifetime prediction routing in mobile ad-hoc networks**,"** *IEEE Wireless Communications and Networking Conference,* USA, March 16-20, 2003, vol.2, pp.1185–1190.

[8] W.-F. Wang and P.-H. Shih, "Study on an enhanced link-stability based routing scheme for mobile ad hoc networks," in *Proceedings of 3rd Annual IEEE Communicati**on** Society Conference on Sensor and Ad Hoc Communications and Networks (SECON'06),* Reston, VA, USA, vol.3, September 25-28, 2006, pp. 797–802.

[9] F. Erbas, J. E.Garcia, K. Jobmann, "Position- based QoS routing in mobile ad hoc networks problem statement and a novel approach", in *Proceedings of 3rd IEEE International Conference on performance, Computing and communications,* Pheonix, AZ, USA, April 15- 17, 2004, pp. 619- 623.

[10] S.R. Medidi, K.H. Vik, "Quality of service- Aware source-Initiated ad-hoc routing ", in *Proceedings of IEEE SECON Conference* Pullman, USA, October 4-7, 2004, pp. 108-117.

[11] Peng Yang, Biao Huang, "QoS Routing Protocol Based on Link Stability with Dynamic Delay Prediction in MANET", *IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application (PACII),* Wuhan, China, vol. 1, December 19-20, 2008, pp.515-518..

[12] O. Tickoo, S. Raghunath, Kalyanaraman S, **"**Route fragility: a novel metric for route selection in mobile ad hoc networks", in Proccedings of the 11th IEEE international conference on networks (ICON), Sydney NSW, Austrailia, september 28- 1st october 2003, pp. 537-542.

[13] Nikoletta Sofra, K. K. Leung, " Estimation of Link quality and Residual Time in VANETs", *in Proceedings of IEEE wireless communications and Networking Conference (WCNC),* Sydney, Austrailia, 31 March- 3rd April, 2008, pp.2444-2449.

[14] Murad Abusubaih, Berthold Rathke, and Adam Wolisz, "A Dual Distance Measurement scheme for Indoor IEEE Wireless Local Area Networks", in *Proceedings of the 9th IFIP/IEEE International Conference on Mobile and wireless Communication Networks (MWCN'07),* Ireland, Cork, September 9-21, 2007, pp. 121-125.

[15] Branislav Kusy, Akos Ledeczi, Xenofon, Koutsoukos, "Tracking Mobile nodes using RF Doppler shifts", in Proceedings of the 5th International Conference on Embedded networked sensor systems, Sydney, NSW, Australia, November 6-9, 2007, pp. 29-42.

Ajay koul received his B.E. degree from Bangalore University, Bangalore, Karnataka, India in Electrical & Electronics in 1997 and MTech from Hyderabad University, Hyderabad, Andhra Pradesh, India in Computer Science in the year 2001.

He is currently a Ph.D. student in the School of Computer Science at SMVD University Katra, J&K, India. His research interests include wireless networks, Image processing and data storage.

Dr. R. B. Patel received his PhD from IIT Roorkee, Roorkee, Uttarakhand, India, in Computer Science & Engineering. PDF from Highest Institute of Education, Science & Technology (HIEST), Athens, Greece, MS (Software Systems) from BITS, Pilani, Rajasthan

He is in teaching and Research & Development since 1991. He has supervised several M. Tech, M. Phil and PhD . He has published more than 95 research papers in International/National Journals and Refereed International Conferences. His current research interests are in Mobile & Distributed Computing, Mobile Agent Security and Fault Tolerance, development infrastructure for mobile & Peer-To-Peer computing, Device and Computation Management, Cluster Computing, etc. He has written numbers books for engineering courses (These are "Fundamentals of Computing and Programming in C", "Theory of Automata and Formal Languages", "Expert Data Structures with C," out of these some are recommended by Indian Society for Technical Education (ISTE), India.

Dr. Patel received several research awards, like best research paper award at BIS 2003 Colorado, Spring, USA and is a member of various International Technical Societies such as IEEE-USA, Elsevier-USA, Technology, Knowledge & Society-Australia, WSEAS, Athens.

Dr. V.K. Bhat received his MscDegree from Kashmir University, Kashmir and Phd Degrees in Mathematics from Jammu University, Jammu, J&K, India.

He is currently serving in the School of Mathematics at SMVD university Katra India. He has supervised many research students in the area of Ring theory, Load balancing and distributed networks. He has published nearly 65 research papers mostly in international journals and conferences of repute. His current research interests include Graph theory and Ring theory.

Investigation of Blackhole Attack on AODV in MANET

Anu Bala
University Institute of Engg. & Tech Deptt, Panjab University, Chandigarh,India
Email: anubala22@gmail.com

Raj Kumari and Jagpreet Singh
University Institute of Engg. & Tech Deptt, Panjab University, Chandigarh,India
Guru Teg Bahadur Khalsa Institute of Engineering and Technology, Malout, India
Email:rajkumari_bhatia5@yahoo.com and drjagz@gmail.com

***Abstract*—Mobile Ad Hoc Network (MANET) consists of a collection of wireless mobile hosts without the required intervention of any existing infrastructure or centralized access point such as base station. The dynamic topology of MANET allows nodes to join and leave the network at any point of time. Wireless MANET is particularly vulnerable due to its fundamental characteristics such as open medium, dynamic topology, distributed cooperation and constrained capability. In this paper we simulate the blackhole attack which is one of the possible attacks on AODV routing protocol in mobile ad hoc networks by the help of network simulator (NS-2). The simulation results show the packet loss, throughput, and end-to-end delay with blackhole and without blackhole on AODV in MANET. We analyzed that the packet loss increases in the network with a blackhole node. We also observed that the throughput and end-to-end delay decreases in the network with a blackhole node.**

***Index Terms*—introduction, AODV routing protocol, blackhole attack in AODV, simulation environment and results, conclusion, acknowledgement, references**

I. Introduction

Wireless networks use some sort of radio frequencies in air to transmit and receive data instead of using some physical cables. Wireless networks are formed by routers and hosts. Ad-hoc networks are wireless networks where nodes communicate with each other using multi-hop links. Networks that support mobile wireless ad hoc architecture are typically called mobile ad hoc networks (MANET). A mobile ad hoc network is formed by mobile hosts. There is no stationary infrastructure or base station for communication. So the functioning of Ad-hoc networks is dependent on the trust and co-operation between nodes. Ref. [1] Nodes help each other in conveying information about the topology of the network and share the responsibility of managing the network. Ref. [2] Each mobile node acts as a host when requesting/providing information from/to other nodes in the network, and acts as router when discovering and maintaining routes for other nodes in the network.

Ref. [3] Based on the routing information update mechanism, routing protocols in ad hoc wireless networks can be classified into three broad categories: Proactive (or table-driven) protocols, Reactive (or on-demand) protocols, and Hybrid routing protocols. These are further divided into sub categories. Ref. [3] These are vulnerable to routing attacks. Routing attacks in ad hoc wireless networks can also be classified into five broad categories: Attacks using Impersonation, Modification, Fabrication, Replay, and Denial of Service (DoS). In this paper, we focus on blackhole attack that belongs to category of fabrication attacks.

Ref. [4] There are three main routing protocols proposed for MANET: Ad hoc On-demand Distance Vector (AODV) routing, Dynamic Source Routing (DSDV), and Destination Sequence Distance Vector routing protocols. AODV and DSR belong to on-demand routing protocols and DSDV is a table-driven routing protocol. These protocols are vulnerable to different security attacks. In this paper, we use AODV routing protocol because the AODV protocol is vulnerable to the blackhole attack. So we have simulated the behavior of blackhole attack on AODV in MANET.

II. AODV Routing Protocol

Ref. [5] Ad-Hoc On-Demand Distance Vector (AODV) is a reactive routing protocol in which the network generates routes at the start of communication. Ref. [7] The Ad Hoc On-Demand Distance Vector (AODV) routing protocol described in builds on the DSDV algorithm. AODV is an improvement on DSDV because it typically minimizes the number of required broadcasts by creating routes on a demand basis, as opposed to maintaining a complete list of routes as in the DSDV algorithm. Ref. [7] The authors of AODV classify it as a pure on-demand route acquisition system, since nodes that are not on a selected path do not maintain routing information or participate in routing table exchanges.

AODV builds routes using a route request / route reply query cycle. When a source node desires a route to a destination for which it does not already have a route, it broadcasts a route request (RREQ) packet across the network. Nodes receiving this packet update their information for the source node and set up backwards pointers to the source node in the route tables. In addition to the source node's IP address, current sequence number, and broadcast ID, the RREQ also contains the most recent sequence number for the destination of which the

doi:10.4304/jetwi.2.2.96-100

source node is aware. A node receiving the RREQ may send a route reply (RREP) if it is either the destination or if it has a route to the destination with corresponding sequence number greater than or equal to that contained in the RREQ. If this is the case, it unicasts a RREP back to the source. Otherwise, it rebroadcasts the RREQ. Nodes keep track of the RREQ's source IP address and broadcast ID. If they receive a RREQ which they have already processed, they discard the RREQ and do not forward it. As the RREP propagates back to the source, nodes set up forward pointers to the destination. Once the source node receives the RREP, it may begin to forward data packets to the destination. If the source later receives a RREP containing a greater sequence number or contains the same sequence number with a smaller hop count, it may update its routing information for that destination and begin using the better route. As long as the route remains active, it will continue to be maintained. If a link break occurs while the route is active, the node upstream of the break propagates a route error (RERR) message to the source node to inform it of the now unreachable destination(s). After receiving the RERR, if the source node still desires the route, it can reinitiate route discovery.

The rest of this paper is organized as follows. In section II, we discuss the AODV routing protocol in detail. Section III describes the characteristics of the blackhole attack on AODV. Section IV provides the simulation environment and results. Finally we conclude in section V.

III. BLACKHOLE ATTACK in AODV

Ref. [5] In a blackhole attack, a malicious node can impersonates a destination node by sending a spoofed route packet to a source node that initiates a route discovery. Ref. [2] A blackhole has two properties:

1. The node exploits the ad hoc routing protocol, such as AODV, to advertise itself as having a valid route to a destination, even though the route is spurious, with the intention of intercepting packets.
2. The node consumes the intercepted packets.

In an ad hoc network that uses the AODV protocol, a blackhole node absorbs the network traffic and drops all packets. To explain the blackhole attack we add a malicious node that exhibits blackhole behavior in the Fig. 1.

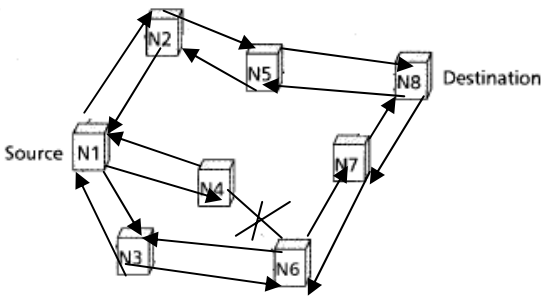


Figure 1: Blackhole attack in AODV

In Fig. 1, we assume that node N4 is the malicious node. Suppose node N1 wants to send data packets to node N8 in Fig. 1, and initiates the route discovery process. We assumed node N4 is a malicious node with no fresh enough route to destination node N8. However, node N4 claims that it has the route to the destination whenever it receives RREQ packets, and sends the response to source node N1. The destination node and any other normal intermediate nodes that have the fresh route to the destination may also give a reply. If the reply from a normal node reaches the source node of the RREQ. First, everything works well; but the reply from malicious node N4 could reach the source node first, if the malicious node is nearer to the source node. Moreover, a malicious node does not need to check its routing table when sending a false message; its response is more likely to reach the source node first. This makes the source node think that the route discovery process is complete, ignore all other reply messages, and begin to send data packets. As a result, all the packets through the malicious node are simply consumed or lost. The malicious node could be said to form a black hole in the network. In this way the malicious node can easily misroute a lot of network traffic to itself, and could cause an attack to the network with very little efforts on its part.

IV. SIMULATION ENVIRONMENT and RESULT

In this section we present a set of simulation experiments to evaluate the effect of blackhole attack on AODV protocol in MANET. First I have explained blackhole attack in detail via simulation in NS-2.We have generated a small size network with7 nodes in a flat grid of 670m x 670m including blackhole node. We have generated a connection between nodes 1 and node 2. We have also introduced some movements in our scenario. Duration of the scenario is 60 seconds. Node 1 is the source node, node 2 is the destination node and node 6 is the blackhole node. Fig. 2 shows the data flow from node 1 to node 2 via intermediate nodes 3 and 4. For some seconds, the link breaks and all data that is send from node 1 get lost as shown in Fig.3. Now Fig. 4 shows that node 1 again sends the RREQ to all nodes to find route. Nodes further rebroadcast the request if they are not the destination nodes. Node 6,that is blackhole node, claims that it has the route to destination whenever it receives RREQ packets and sends the response to source node 1. All other nodes that have the fresh route also send a reply. But the reply from node 6 reaches the source node first. Node 1 accepts it and ignores all other reply messages and begins to send data packets to node via node 3, 4, 5 and Node 6 being a blackhole node absorbs all the packets and traffic as shown in Fig. 5.

Secondly, to calculate network performance, we simulate blackhole node behavior in AODV in large number of nodes and connections with the help of Network Simulator 2 Ref. [6]. We set the parameters for our simulation as shown in Table 1.

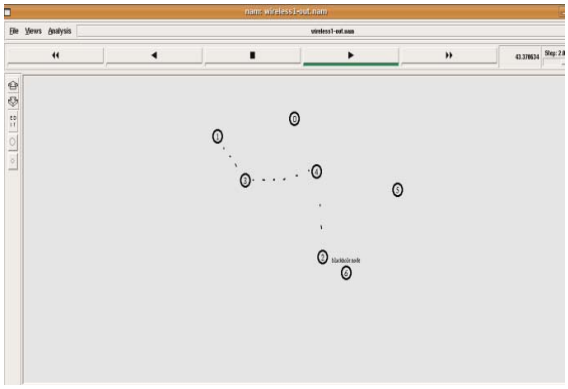


Figure 2 : Data flow between Node 1 to Node 2 via Node 3 and Node 4

Figure 3 : Link breakage and Data Loss

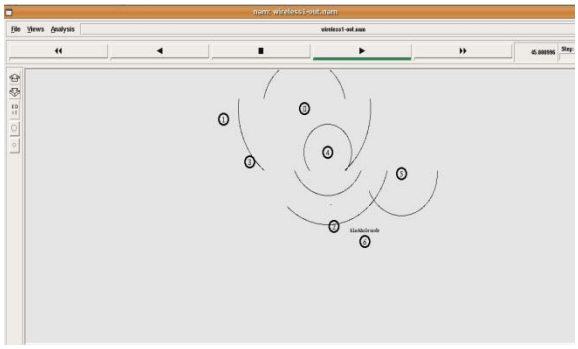


Figure 4 : Route Discovery Process

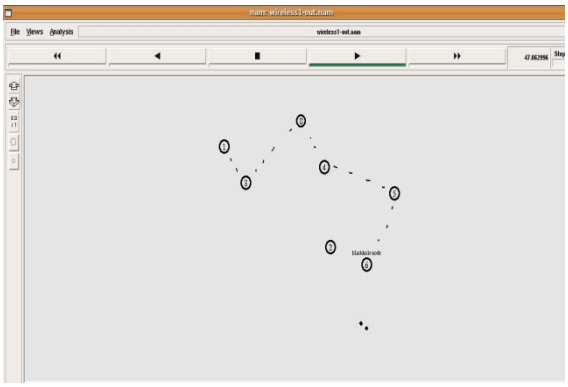


Figure 5 : Node 1 found new route and Node 6(Blackhole Node) absorbs the Data

Table 1: Simulation Parameters

Simulator	ns-2 (ver.2.31)
Simulation Time	500(s)
Number of Mobile Nodes	20
Number of Blackhole Nodes	1
Topology	750m x750m
Transmission Range	250m
Routing Protocol	AODV
Traffic	Constant Bit Rate (CBR)
Pause Time	10(s)
Maximum Connections	9
Packet Size	512 bytes
Data Rates	10 Kbits

We have taken four scenarios of defined parameters for our simulation with or without blackhole node. We have taken different positions and movements of nodes for each scenario. Then we have varied the blackhole nodes and simple nodes to evaluate the performance. We have also varied the mobility speed of mobile nodes.The metrics are used to evaluate the performance are packet loss percentage, throughput and end-to-end delay. We calculate data loss percentage with blackhole and without blackhole node. Then we compare the results of these two simulations to understand the network and node behaviors. The results of the simulation show that the packet loss in the network with a blackhole increases beyond that dropped by the blackhole node. This is due to increased congestion in the routes toward the blackhole node.

Table 2 shows the packet loss percentage on AODV with the presence and absence of blackhole nodes for four scenarios.

Table 2: Simulation Results with blackhole effect and without blackhole effect

Scenarios	**Total Loss of AODV**	**Total Loss of Black Hole AODV**	**Increase**
Scenario1	3.37	90.54	87.17
Scenario2	2.53	90.42	87.89
Scenario3	2.35	98.54	96.19
Scenario4	1.74	88.01	86.27

Our simulation results show that AODV network has normally 2.50 % data loss and if a blackhole node is introducing in this network data loss is increased to 89.38 %. As 2.50 % data loss already exists in this data traffic, blackhole node increases this data loss by 86.88 %.

We have also analyzed the throughput of received packets with the presence and absence of blackhole node with respect to the simulation time of 450(s).

Fig. 6 illustrates the graphic representation of packet loss percentage with and without blackhole node with respect to simulation time (1=100seconds).

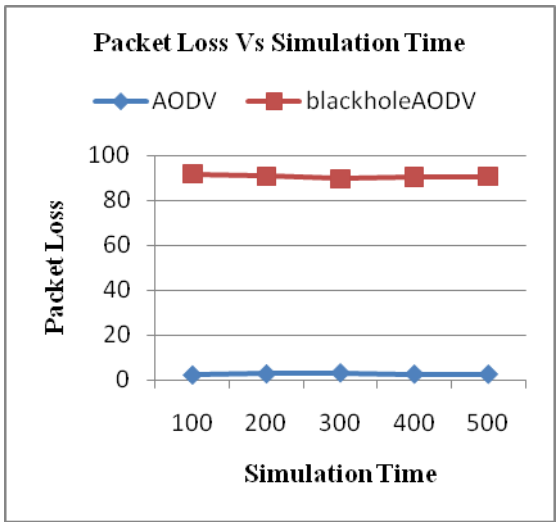


Figure 6 : Shows the Packet Loss of AODV and blackholeAODV

Fig. 7 shows the effect of blackhole attack on throughput of received packets of network. The result shows both the cases with blackhole and without blackhole attack. With our simulation, we analysed that the throughput of received packets in AODV is very high than the throughput of received packets in blackholeAODV. Because the packet loss in blackholeAODV is higher than the AODV protocol.

We studied the performance with varying number of Blackhole Nodes. Number of Blackhole Nodes varies from 1 to 4 with the increment of 1. Fig. 8 shows the impact of number of Blackhole Nodes on throughput in the network.. Simulation results show that the throughput decreases with the increase of number of Blackhole Nodes.

We also studied the performance with varying the number of nodes. Fig. 9 shows the impact of number of nodes on throughput without blackhole attack. The number of nodes is varying from 10 to 50 with the step of 10. Simulation results show that when the number of nodes increases, the throughput increases for AODV protocol.

We have evaluated the End-to-End Delay with varying the mobility speed of nodes without blackhole node and with blackhole node. The mobility speed varies from 10m/s to 50m/s with the increase of 10.

Fig. 10 illustrates the End-to-End Delay with blackhole attack and without blackhole attack. We observed that, there is increase in the average end-to-end delay without the effect of blackhole attack. This is due to the immediate reply from the malicious node because it doesn't check its routing table.

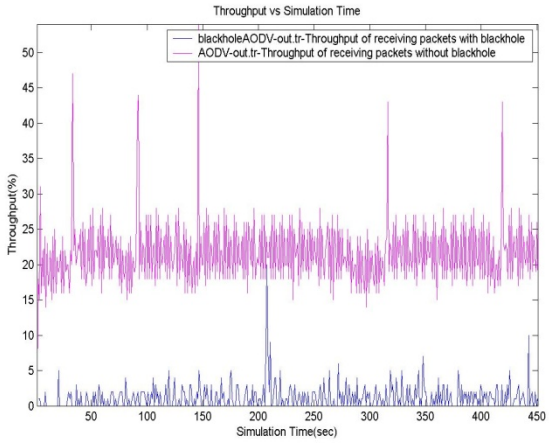


Figure 7: Impact of Blackhole Node on Throughput of Received Packets

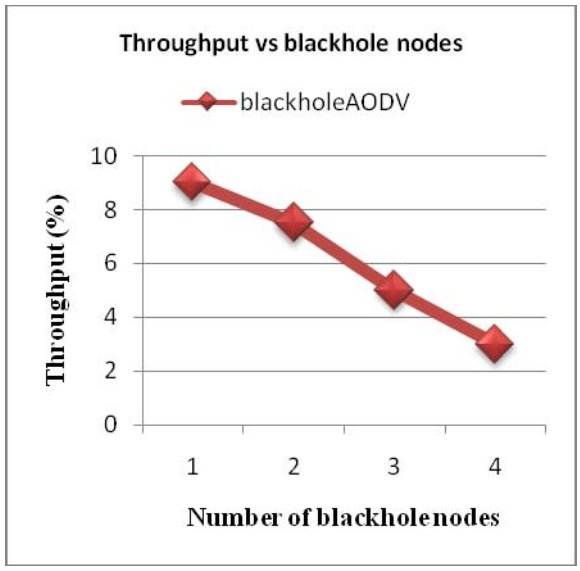


Figure 8: Impack of Number of Blackhole Nodes on Throughput

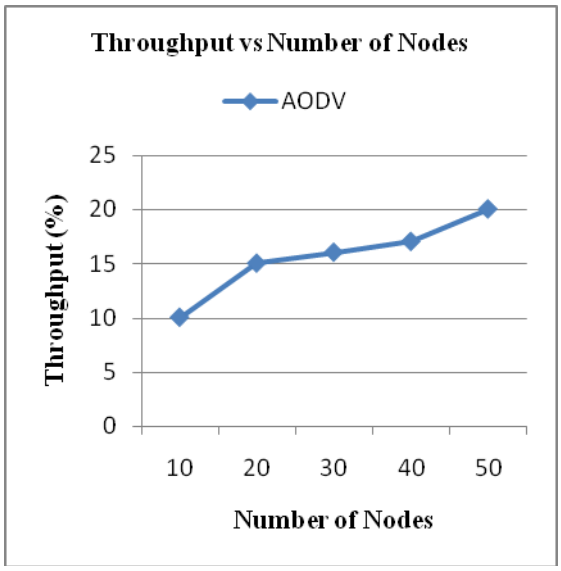


Figure 9: Impact of Number of Nodes on Throughput

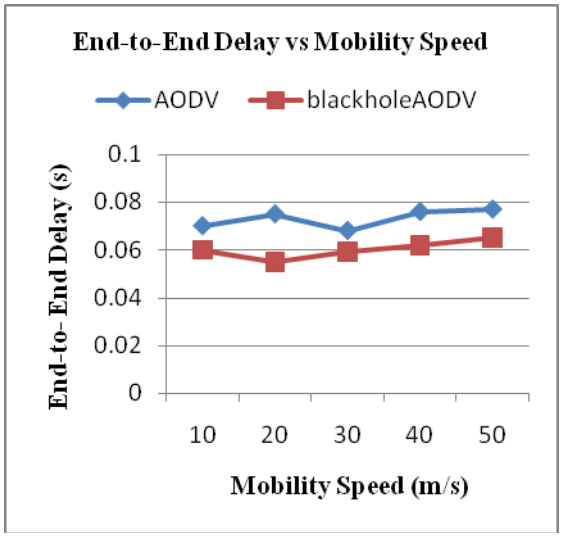


Figure 10: Impact of blackhole attack on End-to-End Delay

V. CONCLUSION

In this paper, we studied AODV in detail and the blackhole attack in AODV. We evaluate the effects of blackhole nodes on AODV in ad hoc networks. We simulate the blackhole behavior with the help of Network Simulator 2 and compared the performance of blackholeAODV with the original AODV in terms of packet loss percentage. The simulation results show that the packet loss increases in the network with a blackhole node. Simulation results also show that the throughput of the network is decreased with blackhole attack as compared to without blackhole attack.. When the number of blackhole nodes increases the throughput decreases. We observed that the End-to-end Delay without blackhole attack is slightly increased as compared to the effect of blackhole attack. The detection of blackhole in ad hoc networks is still to be a challenging task

ACKNOWLEDGEMENT

I would like to take the opportunity to thank people who guided and supported me during this process. Without their contributions, this project would not have been possible. I have a great pleasure in expressing my deep sense of gratitude and indebtedness to Mr. Jagpreet Singh, Lecturer, Teg Bahadur Khalsa Institute of Engineering and Technology, Malout and Ms. Rajkumari, lecturer, University Institute of Engineering and Technology, Panjab University, Chandigarh, my supervisor for their continuous guidance and invaluable suggestions at all the time during the research work. My special thanks to all my friends for being supportive in my hours of need. Finally, I would not forget to thank my parents for their love and blessings that support and encouraged me at every moment I need. They were the first ones that introduced the amazing world to me and encouraged me to explore the wonderful nature.

REFERENCES

[1] Latha Tamilselvan and Dr. V.Sankaranarayanan, "Solution to Prevent Rushing Attack in Wireless Mobile Ad hoc Networks".
[2] Latha Tamilselvan, Dr.V Sankaranarayanan, "Prevention of Blackhole Attack in MANET". The 2nd International Conference on Wireles Broadband and Ultra Wideband Communications (AusWireless 2007) India, 2007 IEEE.
[3] Mohammad O. Pervaiz, Mihaela Cardei, and Jie Wu, "Routing Security in Ad Hoc wireless Networks", Network Security, 2005 Springer.
[4] Elizabeth M. Royer, and Chai-Keong Toh, "A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks," IEEE Personal Communications, pp. 46-55, April 1999.
[5] Satoshi Kurosawa, Hidehisa Nakayama, Nei Kato, Abbas Jamalipour, and Yoshiaki Nemoto. "Detecting Blackhole Attack on AODV based Mobile Ad hoc networks by Dynamic Learning Method". International Journal of Network Security, Vol.5, No.3, PP.338–346, Nov 2007.
[6] ns-2 : http://www.isi.edu/nsnam/ns/
[7] C. E. Perkins and E. M. Royer, "Ad Hoc On-Demand Distance Vector Routing," *Proc. 2nd IEEE Wksp. Mobile Comp. Sys. and Apps.*, New Orleans, LA, Feb. 1999, pp. 90–100.

Anu Bala from Bhogpur (Jalandhar, India) and her date of birth is 22nd August 1985. She is a ME (IT) student in Information Technology at the University Institute of Engineering and Technology, Panjab University, Chandigarh, India. She received her B.Tech (CSE) degree from Panjab Technical University, jalandhar, India in 2006. Her research interests in the area of mobile ad hoc networks, especially ad hoc network security. Her recent work has focused on ad hoc routing protocol attacks.

She has worked as Lecturer for one year at IITT college of engg., Pojewal, Punjab, India.

Raj Kumari from chandigarh (India) and his date of birth is 6th June 1981. she received her M.Tech (IT) degree from GNDU, Amritsar, India in 2006 and her B.Tech from Panjab Technical University, Jalandhar, India in 2003.

She has works as Lecture for one year at college of Engg. Tangori, India. She has been working in UIET, Chandigarh, India since 2007

Jagpreet Singh from Malout (India) and his date of birth is 5th March 1983. He received his MS (Software System) degree from BITS Pilani, India and his B.Tech from Panjab Technical University, Jalandhar, India. He has been engaged in research on mobile adhoc network and he has enhanced algorithm of AODV protocol. He has been working in GTBIET, Malout, India since 2003. Two of his papers has been published in National Conferences.

1) Communication technology on UBI Quietest computing

2) Successful implementation of recruisite implementation of e-governance

Copyright Protection of Gray Scale Images by Watermarking Technique Using (N, N) Secret Sharing Scheme

Sushma Yalamanchili
Professor & Head, Department of CSE,V.R.Siddhartha Engineering College, Vijayawada.
Email: sushma_yalamanchili@yahoo.co.in

M.Kameswara Rao
Lecturer, P.B.Siddhartha College, P.G.Centre, Vijayawada.
Email: kamesh.manchiraju@gmail.com

***Abstract*— Since the rise of the Internet one of the most important factors of information technology and communication has been the security of information. A special case of information hiding is digital watermarking. Digital watermarking is the process of embedding information into digital multimedia content such that the information (the watermark) can later be extracted or detected for a variety of purposes including copy prevention and control. Among numerous cryptographic solutions proposed in the past few years, secret sharing schemes have been found sufficiently secure to facilitate distributed trust and shared control in various communication applications.**

In this paper, a new image watermarking algorithm is developed using (n,n) secret sharing scheme for copyright protection. The proposed method embeds the copyright image into original image and is to be shared among n participants. Then the copyright image could be recovered using simple XOR operations without any loss.

Experimental simulations are provided using MATLAB 7.1 to demonstrate the efficient performance of the developed technique in terms of reliability of watermark embedding and extraction. The scheme contains three phases: the copyright image embedding phase, embedded image secret sharing phase, copyright image extraction phase. The experimental results show that the proposed scheme can resist several attacks such as JPEG compression, resize and noise addition.

***Index Terms*— copyright protection, secret sharing, watermarking, steganography.**

I. INTRODUCTION

On the Internet today it is possible to duplicate digital information a million-fold and distribute it over the entire world in seconds. These issues worry creators of intellectual property to the point that they do not even consider to publish on the Internet. To solve the problem of publishing digital images, researchers have come up with digital image watermarking [1]. Digital watermarking is a method of embedding identifying information in an image, in such a manner that it cannot easily be removed .An application of watermarking is copyright control, in which an image owner seeks to prevent illegal copying of the image. A digital watermark is a code that is embedded inside an image. It acts as a digital signature, giving the image a sense of ownership or authenticity. Digital watermarking of an image has also been proposed for the prevention of copying of an image by unauthorized persons.

To a computer, an image is a collection of numbers that constitute different light intensities in different areas of the image [4]. This numeric representation forms a grid and the individual points are referred to as pixels. Most images on the Internet consists of a rectangular map of the image's pixels (represented as bits) where each pixel is located and its colour [5]. Digital colour images are typically stored in 24-bit files and use the RGB colour model, also known as true colour [6]. All colour variations for the pixels of a 24-bit image are derived from three primary colours: red, green and blue, and each primary colour is represented by 8 bits [4].

Secret sharing [2] refers to method for distributing a secret amongst a group of participants, each of which is allocated a share of the secret. The secret can be reconstructed only when the shares are combined together; individual shares are of no use on their own. Secret image sharing is a technique for protecting images that involves the dispersion of the secret image into many shadow images. This endows the method with a higher tolerance against data corruption or loss than other image-protection mechanisms, such as encryption or steganography.

A secret kept in a single information-carrier could be easily lost or damaged. Secret sharing (SS) schemes, called (n, n) schemes, have been proposed since late 1970s. To encode the secret into n pieces ("shadows" or "shares") that the pieces can be distributed to n participants at different locations. The secret can only be reconstructed from n pieces [2].

II. PROPOSED METHOD

Consider an image A of NR ×NC. Each pixel of A can take any one of c different colors or gray-levels. Image A is represented by an integer matrix A:
A = [aij]NR×NC , where i = 1, 2, . . . , NR,
j = 1, 2, . . . , Nc, and aij € {0, 1, . . . , c − 1}.We have c =2 for a binary image, and c =256 for a grayscale image

doi:10.4304/jetwi.2.2.101-105

with one byte per pixel. In a color image with one byte per pixel, the pixel value can be an index to a color table, thus c = 256. In a color image using an RGB model, each pixel has three integers: R (red), G (green) and B (blue). If each R, G or B takes value between 0 and 255, we have c = 256 [2]. The proposed method includes the following steps

Step 1: Consider two images - Original image and the copyright image - represented by integer matrices.

Step 2: Decide a value called weight_value between 0 and 1.For invisible watermark weight_value must be near to 0.

Watermark Embedding :

Step 3: Embed the copyright image inside the original image as follows

Embedded image=original image + (resized copyright image * weight_value).

(N,N) Secret Sharing Scheme For the embedded Image :

The output of a (n,n) scheme for embedded images is a set of n distinct NR×NC matrices A1, . . . , An, called shares or shadow images. Each share image has the same number of pixels as the original image A, but every pixel in a share may contain m sub pixels. A can be reconstructed from the set {Ai1, . . . , Aik } and even complete knowledge of k − 1 shares reveals no information about A. The first condition above is called precision, and the second condition is called security [2].

Step 4: Encode the embedded image into n shadows or secrets as follows and

generate n − 1 random matrices B1, . . . , Bn−1,

compute the shadow images as below:

$A1 = B1$,

$A2 = B1 \oplus B2$,

. . . . . .

$An-1 = Bn-2 \oplus Bn-1$,

$An = Bn-1 \oplus A$.

In the generation of the shadow images and in the reconstruction of the secret, Boolean operation XOR (“$\oplus$”) is used. For easy lookup, the truth-table of XOR for binary scalar inputs is given below.

	a=0	a=1
b=0	0	1
b=1	1	0

a $\oplus$ b

For example, when a =125 and b =18, the XOR between these two integers is

$a \oplus b = (125)_{10} \oplus (18)_{10} = (01111101)_2 \oplus (00010010)_2$

$= (01101111)_2 = (111)_{10}$.

Step 5 : Reveal the embedded image from the n shadow images as below:

$A' = A1 \oplus A2 \oplus \cdots \oplus An$.

Because the “$\oplus$” operation is associative and $Bi \oplus Bi$ is a zero matrix for any i, we have

$A = B1 \oplus (B1 \oplus B2) \oplus \cdots \oplus (Bn-2 \oplus Bn-1) \oplus (Bn-1 \oplus A)$

$= (B1 \oplus B1) \oplus \cdots \oplus (Bn-1 \oplus Bn-1) \oplus A = A.$

To demonstrate the computation steps in the revealing process, we give a trivial example for n = 3 and a single-pixel secret image A.

Given: $A = (231)_{10} = (11100111)_2$

Generate: $B1 = (46)_{10} = (00101110)_2$ and

$B2 = (188)_{10} = (10111100)_2$

Compute: $A1 = B1 = (46)_{10} = (00101110)_2$,

$A2 = B1 \oplus B2 = (10010010)_2$, and

$A3 = B2 \oplus A = (01011011)_2$

Reconstruct: $A1 \oplus A2 \oplus A3 = (11100111)_2 = (231)_{10}$.

An example for the proposed (n, n) scheme using a 3 × 3 secret image A is given below:

$$A = \begin{pmatrix} 209 & 214 & 225 \\ 233 & 227 & 228 \\ 222 & 221 & 226 \end{pmatrix}$$

$$B1 = \begin{pmatrix} 136 & 169 & 28 \\ 254 & 128 & 245 \\ 96 & 108 & 49 \end{pmatrix}$$

$$B2 = \begin{pmatrix} 8 & 94 & 65 \\ 218 & 137 & 228 \\ 46 & 71 & 222 \end{pmatrix}$$

,

$$A1 = B1 = \begin{pmatrix} 136 & 169 & 28 \\ 254 & 128 & 245 \\ 96 & 108 & 49 \end{pmatrix}$$

,

$$A2 = B1 \oplus B2 = \begin{pmatrix} 128 & 247 & 93 \\ 36 & 9 & 17 \\ 78 & 43 & 239 \end{pmatrix}$$

,

$$A3 = B2 \oplus A = \begin{pmatrix} 217 & 136 & 160 \\ 51 & 106 & 0 \\ 240 & 154 & 60 \end{pmatrix}$$

, and

$$A1 \oplus A2 \oplus A3 = \begin{pmatrix} 209 & 214 & 225 \\ 233 & 227 & 228 \\ 222 & 221 & 226 \end{pmatrix} = A$$

Watermark Extraction :

Step 6 : Extract the copyright image inside the embedded image as follows

copyright image =(Original image – embedded image)/weight_value.

III. Experimental Results On Gray Scale Images

The simulations were conducted on Intel machine with 2.4 GHz processor and 512 MB of RAM. MATLAB 7.1 was used for implementation of proposed scheme and

image processing operations respectively. Applying the proposed method on tree path image as original image and college logo image as copyright image under (n,n) scheme with $n = 3$.

Figure 1 (a) Original image

Figure 1 (b) Copyright Image

Figure 1(c) Embedded Image

Fig. 1 shows the original image in part (a), copyright image in part (b) and the embedded image (Original image + Copyright Image) in part(c) .

(n,n) Secret sharing scheme

A (n,n) secret sharing scheme encodes a secret image into n share images, which demonstrate randomly noisy patterns and hide the information about the secret image, are distributed to n recipients. The secret image can easily be decrypted by XOR operation of the n share images in an arbitrary order without complicated arithmetic.

Figure 2(a)Embedded Image

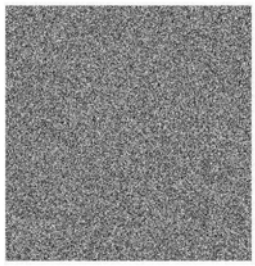


Figure 2(b)shadow image1

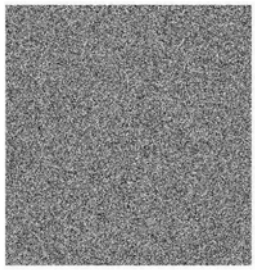


Figure 2(c) Shadow image 2

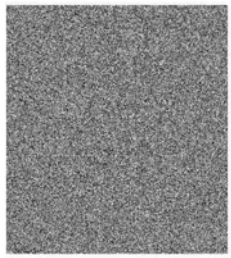


Figure 2(d) Shadow image 3

Figure 2 (e) Reconstructed Embedded image using 3 shadows

Fig. 2(a) shows the embedded image in part (a), shadow images in part(b-d)and the reconstructed embedded image using the 3 shadow images in part(e).

Strength of the (n,n) Scheme

The reconstruction phase of our algorithm computes n shadow images using XOR operation. the proposed (n, n) scheme reconstructs the secret image exactly, and it satisfies the security condition. That is, when fewer than n shadows are used, the original secret image A will not be revealed.

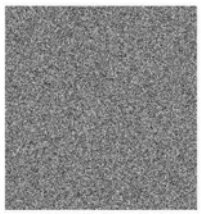


Figure 3(a)Reconstructed image using shadow image 1 and shadow image 2

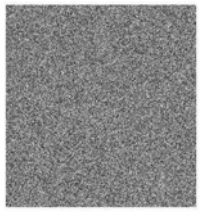


Figure 3(b)Reconstructed image using shadow image 1 and shadow image 3

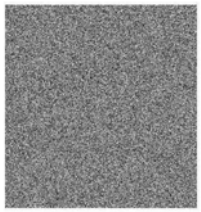


Figure 3(c)Reconstructed image using shadow image 2 and shadow image 3

Fig 3 shows the Reconstructed image using shadow images 2(a) and 2(b) in part(a) Reconstructed image using shadow images 2(a) and 2(c) in part(b) and Reconstructed image using shadow images 2(a) and 2(c) in part(c).

Watermark extraction process

Figure 4(a) Embedded image

Figure 4(b)Extracted copyright image

Fig 4 shows the embedded image in part (a) and the extracted copyright image in part (b)

Attacks on the Proposed Method

Our experiments show that for group of attacks like JPEG compression, resizing, adding noise the proposed method can be able to extract the copyright image with less loss in the quality.

Figure 5(a) Resized Embedded Image

Figure 5(b)Extracted Copyright image when the embedded image is resized

Fig 5 shows the Resized embedded image in part(a) and the extracted copyright image when the embedded image is resized in part(b)

Figure 6(a) Compressed Embedded Image

Figure 6(b)Extracted Copyright image when the embedded image is Compressed

Fig 6 shows the compressed embedded image in part (a) and the extracted copyright image when the embedded image is compressed in part (b)

Figure 7(a) Embedded Image with added noise

Figure 7(b)Extracted Copyright image when the embedded image is added with noise

Fig 7 shows the embedded image with added noise in part (a) and the extracted copyright image when the embedded image is added with noise in part (b)

IV. PSNR MEASUREMENT

One commonly used measure to evaluate the imperceptibility of the watermarked image is the peak signal to noise ratio (PSNR) which is given by

PSNR = 10 * log10((255)^2 / mean_square_error)

Where

OriginalImage_x_size = size(OriginalImage, 2);
OriginalImage_y_size = size(OriginalImage, 1);
CopyrightImage_x_size = size(CopyrightImage, 2);
CopyrightImage_y_size = size(CopyrightImage, 1);
mean_square_error = sum(sum(sum((CopyrightImage – Extracted CopyrightImage).^2))) / double(CopyrightImage_x_size * CopyrightImage_y_size * 3);

TABLE I.
PSNR VALUES

Type of Operation on Image	PSNR
Original Embedded Image	32.6737
After resizing embedded Image	29.0508
After compressing the Embedded image to a .jpg	28.4083
After adding random noise to the Embedded image	33.4338

Table I illustrates PSNR values between the Original Copyright Image and the Extracted Copyright Image taken from various operations

V. CONCLUSIONS

We have demonstrated a new watermarking technique that uses (n,n) secret sharing scheme to embed a copyright image into original image . This technique works well with images of all sizes. This technique provides two layers of security. In the first step, a copyright image is embedded into original image for copyright protection. Also, the embedded image is shared among n participants where all the n shares must be used to reconstruct the embedded image. This makes the system more secure. The method can with stand attacks like JPEG compression, resize and adding noise with less loss in quality of the image. Further this work can be extended by calculating the hash value of the image and encrypt the hash value using either symmetric key or public key and embed the hash value inside the image so that at the receivers end the authentication and integrity of the image can be verified by recalculating the hash and verifying it. Similarly digital signatures can be generated for images and can be verified. Also (k,n) threshold secret sharing schemes can be implemented for much security.

VI. REFERENCES

[1] Adrian Perrig Andrew Willmott, ”Digital Image Watermarking in the Real World" Extended Abstract, March 9, 1998.

[2] DaoshunWang, Lei Zhang, Ning Ma, Xiaobo Li,“Two secret sharing schemes based on Boolean operations”, Science Direct –Pattern Recognition 2007.

[3] A. Shamir, “How to share a secret”, Commun. Assoc. Comput. Mach. 22 (11) (1979) 612–613.

[4]Johnson, N.F. & Jajodia, S., “Exploring Steganography: Seeing the Unseen”, Computer Journal, February 1998 .

[5]“Reference guide: Graphics Technical Options and Decisions”, http://www.devx.com/projectcool/Article/1997

[6] Owens, M., “A discussion of covert channels and steganography”, SANS Institute, 2002.

Broadband Integrated Services over Proposed Open CPE Architecture

Kushal Roy
Principal Investigator to DST Government of India and faculty,
Department of Electronics and Communication Engineering
Haldia Institute of Technology, ICARE Complex, PO H.I.T Haldia
PIN- 721657, West Bengal, India
Email: profkushalroy@hotmail.com

***Abstract*— Over the decades, due to technology limitations and regulatory restrictions, information services have been delivered to subscribers through multiple service providers. In recent years, however, the advancement of digital communication technology, the passing of the 1996 Telecommunications Act in the U.S., and the emergence of the Internet are driving network convergence. As a result, the industry is calling for the consolidation of an integrated Customer Premises Equipment (CPE) in the customer's premises to provide integrated services, such as multi-media, voice, and data services. However, the ever-changing network standards and competing technologies not only confuse carriers, but also prevent them from committing to massive CPE deployment.**

In this paper, the problems carriers face in the deployment of integrated services over broadband are first described. In order to remove these deployment obstacles, and also to create new value-added services, service models are investigated and an open CPE architecture is proposed that will support the wide range of broadband access technologies and ever-changing network standards.

***Index Terms*—Integrated Services, Broadband, VoIP, PSN, SoftSwitch Model.**

I. INTRODUCTION

The convergence of voice and Internet data traffic also calls for new Customer Premises Equipment (CPE) such as Digital Subscriber Lines (DSLs), modems, cable modems, and wireless modems that provide users with broadband access to the Internet. However, instead of adding new CPE to the home, the industry is moving towards consolidating CPE into an Integrated CPE (I-CPE) to help lower the cost, reduce network management complexities, and improve the efficiency of network resources. The I-CPE is also called an Integrated Access Device (IAD) and a Residential Gateway (RG) when used in residential areas. In this paper, the term I-CPE is used, because I-CPE is intended to serve multiple market segments including residential, Small Office Home Office (SOHO), and small businesses to provide integrated services such as voice, multi-media, and Internet access.

While the upgrade of the network infrastructure is well underway, the deployment of I-CPE remains a big obstacle to the delivery of integrated services over broadband due to the volume and time that are required for the deployment. It was believed that the lack of auto-configuration was the main roadblock to massive I-CPE deployment. However, the largest obstacle today is not so much how to deploy I-CPE but rather, which type of I-CPE the carriers should be deploying, because this telecommunication world is filled with too many standards, e.g., Media Gateway Control Protocol (MGCP), H.248 [7], H.323, Session Initiation Protocol (SIP), Voice over Internet Protocol (VoIP), Voice over ATM (VoATM), etc.; and competing technologies, e.g., DSL, Cable, Wireless, Fiber, Ethernet, HomeRF∗, IEEE 802.11, Bluetooth, etc. As a result, carriers are facing great difficulties in choosing an I-CPE for deployment because they fear it being replaced in the near future.

II. THE NEXT-GENERATION NETWORK INFRASTRUCTURE

Unlike the dial-up modem, the introduction of I-CPE will require major upgrades in the infrastructure of the existing network. The new network is intended to support the convergence of the voice-centric Public Switched Telephone Network (PSTN) and the Internet, and it is commonly referred as the Next-Generation Network (NGN). The NGN is based on the distributed Softswitch architecture in which the call control is separated from the media transport. Figure 1 shows the role of I-CPE in the NGN infrastructure. It indicates that an I-CPE is the portal to a customer's premises and it acts as the gateway to interconnect the Local Area Network (LAN) with the Wide Area Network (WAN), which consist of the Access Network and the Core Network [6]. The Access Network, consisting of Access Nodes, Regional Broadband Networks, Transit Gateways, and Because of the passing of the Telecommunications Act in 1996 and the FCC's ongoing adventures in local loop deregulation, companies, including Inter-eXchange Carriers (IXC), Incumbent Local Exchange Carriers (ILEC), Competitive Local Exchange Carriers (CLEC), CATV service providers, and many other emerging service providers, are all eager to enter this lucrative broadband access business. Therefore, it is certain that I-CPE will support various broadband interfaces, such as wireless, cable, xDSL, and fiber optics, to target different market sectors depending on price, performance, environment, or the type of users. Then, the recent advancement of IEEE

doi:10.4304/jetwi.2.2.106-109

80.3ae 10 GigE standard [13] enables the Ethernet to work beyond the LAN. Thus, an I-CPE may also support Ethernet broadband interface to the WAN. Access Nodes terminate the broadband interfaces from I-CPE and aggregate multiple CPE traffic to the Regional Broadband Networks. Regional Broadband Networks provide transport and switching functions among Access Nodes and Transit Gateways. Transit Gateways may consist of Trunking Gateways and Signaling Gateways that perform signaling and bearer interworking functions between Regional Broadband Networks and PSN/CSN.

An MGC is also referred to as the Call Manager or Softswitch that typically runs call control software in a server to provide call-processing functions. In the distributed Softswitch architecture, the value-added services or applications can reside in an Application Server. A Media Server is controlled by the Softswitch to play tonal announcements or perform media streaming functions, such as Interactive Voice Response (IVR) and Conference Bridge. The media streaming is carried in bearer connections between the Access Node and the Transit Gateway for on-net calls, or between Access Nodes for off-net calls, while call control functions are provided by an MGC. GigaPoP (i.e., Point of Presence) aggregates the traffic from Ethernet CPE to PSN [14]. It should be noted that the NGN architecture shown above is meant to be logical, so the Access Nodes, Transit Gateways, Media Server, Application Server, and MGC are logical functional blocks which may be implemented as stand-alone devices or embedded in other network elements.

The Media Gateway Controller (MGC), provides the ramp to Packet-Switched Networks (PSN) and Circuit-Switched Networks (CSN).

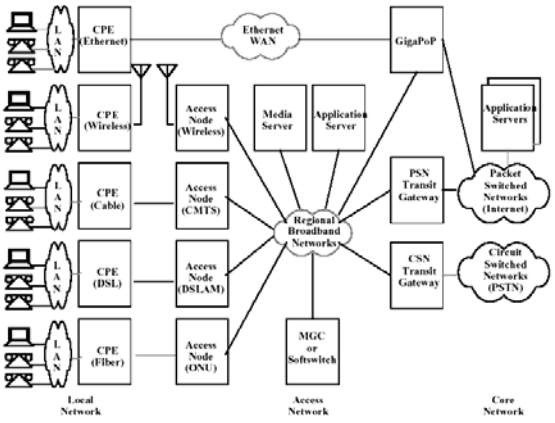


Figure 1: Next-Generation Network infrastructure

III OPEN CPE ARCHITECTURE

Broadband access is, for many, the solution to the "World Wide Wait" problem (slow Web performance), and it would also enable the creation of many new services that promise enormous benefits to subscribers and service providers. However, service providers are facing many challenges during the deployment of broadband access networks, not least of which is how to deploy the greatest number of Customer Premises Equipment (CPE) in the timeliest manner.

.

A. I-CPE Characteristics

The following characteristics of the Integrated Customer Premises Equipment (I-CPE) architecture are important to the resolution of these deployment problems.

- *Open Architecture*–In the past several decades, open architecture has become the trend in the industry. It started with the computer industry in the early 80s, and it gradually spread to the telecommunication industry, where the monolithic central office switch is under great pressure from the Internet telephony to open up. The broadband CPE industry will not buck this trend since openness will only foster new services and product competition, which, in turn, will lower CPE cost, reduce time-to-market, and inspire innovation [3].
- *Flexibility*–The ever-changing network standards and infrastructures, the wide range of access network technologies, and the uncertainty of which value-added service will emerge in the future make it very difficult to have a fit-it-all CPE architecture. Thus, the I-CPE should be flexible to contain the must-have core features needed to support today's services, yet it should still be able to accommodate future upgrades. Flexible architecture will lower CPE costs, something that is crucial to I-CPE being accepted in the cost-sensitive consumer market. For example, an I-CPE might be designed to provide home automation and security services, but the interface standards as well as the functionalities to control home appliances have yet to be defined. The flexible architecture needs to be able to accommodate a Bluetooth wireless modem plug-in and the necessary software that will need to be added to the CPE to support this feature when the standard and technology become available.
- *Value-Added Services* - Each time a new service or technology is introduced, it needs to be perceived by the public as adding value in order that it will be bought and used. Due to the competition from Internet telephony, the cost of toll telephony services is continually decreasing. Soon, it is believed, you will be charged a flat monthly rate for telephone services instead of being billed on a per-minute basis. Therefore, the driver for the Next-Generation Network (NGN) is not about inventing a new way to provide existing services, but focusing on value-added services such as Presence, Voice Web, Voice Portal, Unified Messaging, and Voice Virtual Private Network (VPN), which hold great potential for generating additional revenue for service providers.

B. Integrated CPE Functional Decomposition

Figure 2 shows the Open CPE architecture that is intended to fulfill the requirements as listed above and to support multiple broadband access technologies and telephony protocol standards. It is composed of a CPE Controller Module (CCM), LAN Interface Modules (LIM), and WAN Interface Modules (WIM). The CCM consists of a CPE Controller, Flash, ROM, and RAM that contain a Real-Time Operating System (RTOS) and software to implement telephony call control, network management, and service logic functions, as well as third-party applications. The CCM may also include a Digital Signaling Processor (DSP) to implement media streaming processing functions, such as vocoder, echo cancellation, tone generation and detection. The CCM can be re-configured to support different protocols or telephony standards through software downloading.

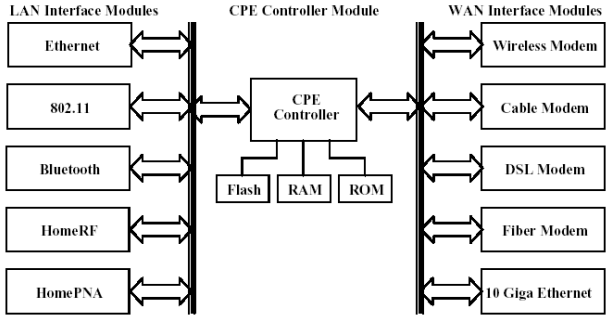


Figure 2: Open CPE architecture

WIM and LIM contain plug-ins to support various WAN and LAN interfaces. WIM may consist of a Wireless Modem , a Cable Modem, a DSL Modem, a Fiber Modem, and a 10 Gigabit Ethernet module that provide interfaces to the broadband access network. LIM may include interfaces to analog phones. There are two buses responsible for the distribution of user data, real-time voice streaming, control signaling, and the management data between the CCM Controller and the WIM/LIM. The interface specification should meet the requirements of transporting real-time and non-real-time data. The open CPE architecture as described above is derived from the abstraction of multiple CPE platforms that may use various broadband access technologies and standards. The goal of the abstraction is to find out what functions are common to all CPEs and therefore should be implemented in CCM, and what functions are interface dependent. and therefore are more appropriately implemented in plug-in modules. Figure 3 gives an example of the CPE functional decomposition. It depicts the protocol stacks of I-CPE supporting cable, xDSL, and 10 Gigabit Ethernet broadband interfaces to provide voice and data convergence services.

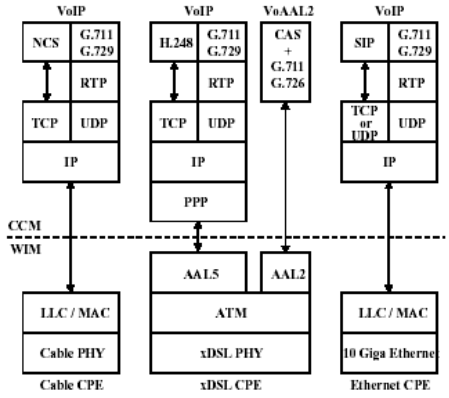


Figure 3: I-CPE protocol stack

The Cable CPE is intended to provide VoIP services. It uses Network-based Call Signaling (NCS) [4] as the signaling protocol to perform telephony call control functions, while encapsulating the voice streaming data in the Real-Time Protocol (RTP) packets to be sent to the Cable Modem Termination System (CMTS). xDSL CPE provides VoIP and Voice over AAL2 (ATM Adaptation Layer type 2) Loop Emulation Services (LES) [5]. It uses H.248 [7] and Channel Associated Signaling (CAS) protocols to implement telephony call control functions for VoIP and VoAAL2, respectively. The voice streaming data are encapsulated in either RTP or AAL2 packets to be sent to the DSL Access Multiplex (DSLAM). In this example, the Ethernet CPE acts as a SIP client to provide VoIP services. SIP is used to establish, modify, and terminate multi-media sessions and calls [10]. The voice streaming is encapsulated in the RTP packets transmitting to the terminating party via the User Datagram Protocol (UDP) connection.

In the Time Division Multiplex (TDM) networks, only call control and management functions are implemented by software in the processor, leaving the voice processing to be handled by the hardware, because of the latency concern.

However, the trend toward packet telephony, along with the advancement and increasing performance of processors in recent years, has made it possible for, and even demanded that, voice streaming be processed by software. Thus, as Figure 3 indicates, it makes good sense to locate the CCM-WIM interface between Layer 2 and Layer 3 of the protocol stack. The CCM contains software to implement voice processing, the signaling protocol, and management functions. WIM should implement Physical layer (PHY), Medium Access Control (MAC), Logical Link Control (LLC), Asynchronous Transfer Mode (ATM), and ATM. Adaptation Layer (AAL) functions that are closely coupled with each broadband access interface.

IV. CONCLUSION

The explosion of the Internet along with regulation and technology changes are reshaping telecommunication networks in many ways. The industry is moving toward the convergence of PSTN and the Internet. The

converged network, the NGN, will operate in a very similar way to the Internet topology in which central office switches are decomposed into distributed systems that are based on the Softswitch model. As a result, the NGN will no longer provide transport services between telephone equipment (as PSTN previously did), but will provide personalized multi-media services. This convergence of voice and Internet data traffic also calls for the deployment of I-CPE to provide integrated services. However, the ever-changing network standards and competing technologies not only confuse carriers, but also prevent them from committing to massive CPE deployment, because they are afraid that the CPE just deployed will have to be replaced in a few years. Hardware replacement is very common in the PC and cellular phone business when new standards or technologies are introduced, but it presents a big threat to the wireline broadband business because CPE deployment is such a daunting task that it cannot be completed easily. In this paper, I proposed an open CPE architecture to solve the CPE deployment dilemma. The architecture is very flexible and can support multiple WAN/LAN technologies and IP Telephony standards. It also includes a common API to allow users to customize or even create new services and third-party developers to create new applications and services. The open CPE architecture will allow CPE vendors to lower costs and reduce time-to-market, and most importantly enable service providers to provide many value-added services that promise great potential for generating additional revenue

REFERENCES

[1] C. A. Eldering, "Customer Premises Equipment for Residential Broadband Gateway," IEEE *Communication Magazine*, , pp. 114-121, June 1997.

[2] S. Moyer and A. Umar, "The Impact of Network Convergence on Telecommunications Software," IEEE *Communication Magazine*, pp. 78-84, January 2001

[3] C. Low, "Integrating Communication Services," IEEE *Communication Magazine*, pp. 164-169, June 1997.

[4] "PacketCable Network-based Call Signaling Protocol Specification," Pkt-SP-EC-MGCP-101-990312, Cable Lab, 3/12/1999.

[5] "Voice and Multimedia over ATM–Loop Emulation Service Using AAL2," af-vmoa-145.000, ATM Forum, July 2000.

[6] J. Chou, "The migration of LES to the Next Generation Network based on H.248," ATM Forum, atm00-164, San Francisco, May 2000

[7] F. Cuervo, N. Greene, A. Rayhan, C. Huitema, B. Rosen, and J. Segers. "Megaco Protocol Version 1.0," IETF RFC3015, November 2000.

[8] J. Lennox, H. Schulzrinne, "Call Processing Language Framework and Requirements," IETF RFC2824, May 2000.

[9] J. Rosenberg, J. Lennox, H. Schulzrinne, "Programming Internet Telephony Services," IEEE Internet Computing, May-June 1999, pp. 63-72.

[10] M. Handley, H. Schulzrinne, E. Schooler, and J. Rosenberg, "SIP: Session Initiation Protocol," IETF, RFC2543, March 1999.

[11] "OpenCable™," http://www.opencable.com, Cable Labs.

[12] J. Lennox, H. Schulzrinne, J. Rosenberg, "Common Gateway Interface for SIP," IETF RFC3050, January 2001.

[13] IEEE P802.3ae 10Gb/s Ethernet Task Force, http://grouper.ieee.org/groups/802/3/ae/index.html.

[14] "Lighting Internet in the WAN," Telecommunication Magazine, September, 2000, http://telecoms-mag.com/issues/200009/tcs/lighting_internet.html.

Kushal Roy, born in M.P Republic of India on December 13th 1979, received his Bachelors Degree with *Honors* in Electronics and Communication Engineering form Government Engineering College Ujjain, M.P India, in the year 2002, achieved his Master's Degree M,*Phil* in Instrumentation from Indian Institute of Technology Roorkee, U.A Republic of India in the year 2004 with nano scale instrumentation and opto instrumentation as major fields of study.

He is involved extensively involved in Research education and development since 2004 with more than 5 years of experience in teaching undergraduate as well as post graduate students in various renowned Institutes across the country. Presently he is serving as faculty in the Department of Electronics and communication Engineering, Haldia Institute of Technology ICARE (Indian Center for Advancement in Research and Education) Complex, Haldia, West Bengal Republic of India.

Mr. Roy has contributed to more than 6 international conferences papers across the globe in Italy, Singapore, Morocco, USA etc. He is also working as Principal Investigator to a sponsored project on "Study and Development of nano scale piezo electronic material Lead Zirconium Titanate, (popularly known as PZT) by sol-gel process." (conferred vide Honorable President of India Direct sanction order SR/FTP/ETA-49/07) worth 0.522 million INR, Science and Engineering Research Council, Department of Science and Technology, Ministry of Science and Technology Government of India.

Recovery Based Architecture To Protect Hids Log Files Using Time Stamps

Surinder S. Khurana
Punjab Engg. College, Chandigarh, India
surindersingh.cs07@pec.edu.in

Divya Bansal , Prof. Sanjeev Sofat
Punjab Engg. College, Chandigarh, India

***Abstract* – After the great revolution in the field of Information Technology, many applications made necessity to run computer systems (either servers or client machines) all the time. Along with improvements and new inventions in technology, the threat of attacks through computer networks becomes a large issue. Host Based Intrusion Detection is a part of security system that protects hosts from various kinds of attacks. It also provides a great degree of visibility (of system activities). It is quite widest that HIDS are vulnerable to attacks. An adversary, if successfully enters in a system can disable HIDS or modify HIDS rules to hide its existence. One can easily evade HIDS. In [7] we propose a new architecture that protects HIDS from such attacks. In this paper, we have proposed a new mechanism to check integrity of log files. We have discussed its affects on performance of system.**

I. INTRODUCTION

Intrusion Detection System [5] is an imperative ingredient in network computer security which plays a vital role in detecting the intrusive activities before they occur. Intrusion Detection is usually done through scanning network traffic and/or hosts data and activities. These intrusive activities can be defined as - activities performed by some adversary for gaining unlawful benefits. The adverse affects of such intrusion activities are in terms of loss of confidentiality, integrity and availability of resources or services.

IDSs can be classified under various categories. Figure-1 illustrates the various classifications of Intrusion Detection Systems.

Response Based	• Active IDS • Passive IDS
Domain Of Detection Based	• HIDS • NIDS
Underlying Detection Technique Based	• Anomaly based IDS • Signature based IDS

Figure-1: IDS Classification

An IDS can be active or passive. Passive IDS detect the attacks and logs information or raise alarms. Active IDS takes action in response to an already detected attack. Active IDSs are also known as Intrusion Prevention System.

Section 2 discusses Detection and Recovery based Architecture to protect HIDS. Section 3 describes time stamping based protocol to check integrity of log files. In section 4 and 5, we discuss implementation details. Affects of our architecture on system performance has described in section 6. In section 7, we discuss some related works. We present directions for future work in 6 and our conclusion in section 8.

II. OVERVIEW OF DETECTION AND RECOVERY BASED ARCHITECTURE

In [1], architecture has proposed to detect attacks on HIDS and recover HIDS to its previous healthy state. The architecture protects HIDS from two type of attacks: first is that it does not allow adversary to kill the HIDS process. And the other is it does not allow unauthorized modification of rule or signature database. The basic idea behind the architecture is to allow the adversary to perform attack on HIDS and then recover the HIDS to its previous healthy state. A new process called MonitorIDS has been introduced in the proposed architecture. MonitorIDS is a lightweight system process which detects the attacks that affects HIDS and takes required actions to recover HIDS from affects of that attack. MonitorIDS takes care of both HIDS process and integrity of HIDS related information. However, this architecture does not consider underlying technique used by HIDS to detect intrusions. It can be used with either signature based or anomaly based IDS.

Figure 2 depicts working of proposed architecture. As shown in figure 2, a backup of HIDS related files has been created immediately after the installation. MonitorIDS process is embedded with HIDS. This process regularly monitors the HIDS process and files after a small fraction of time gap. If MonitorIDS found HIDS process dead (detect unauthorized kill of HIDS) it immediately restarts the HIDS. It also monitors integrity of files related to HIDS. These files may include rule or signature database. If it found any unauthorized modification of HIDS files it replace the modified files

doi:10.4304/jetwi.2.2.110-114

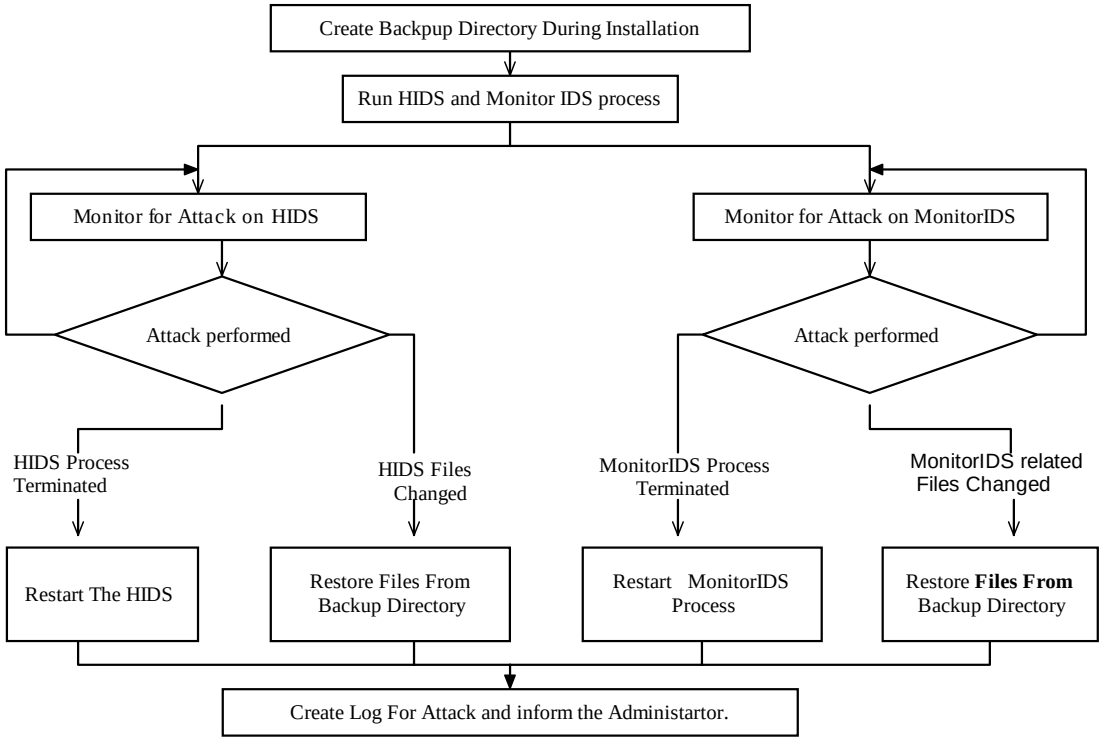


Figure 2 : Block Diagram of proposed architecture

with files from backup directory. It also logs the attack information and informs the administrator.

The architecture also takes care of backup directory and MonitorIDS process related information, so that these cannot be modified by some adversary. One important point to note about MonitorIDS is that it is a System process. To run this process as system process an entry labeled with respawn has made in /etc/inittab file. That cannot be stopped. If some adversary makes attempt to kill this process, operating system kernel automatically restarts it. MonitorIDS process also takes care of /etc/inittab file so that entry related to it cannot be removed.

III. TIME STAMPING BASED MECHANISM TO CHECK INTEGRITY OF LOG FILES

The main goal is to detect and recover unauthorized modification of log files by any process that is not related to HIDS. Such detection is made based on the last modification time (a file attribute, changed when the file was modified) of the log file.

As shown in Figure-3, when HIDS is running under this mechanism it should follow the below given sequence of steps to write an entry in a file.

1. Write log entry in log file and in backup copy of log file.
2. Encrypt Last Modification Time of log file and its backup copy.
3. Write encrypted version of last Modification time of log file and its backup copy into a file.

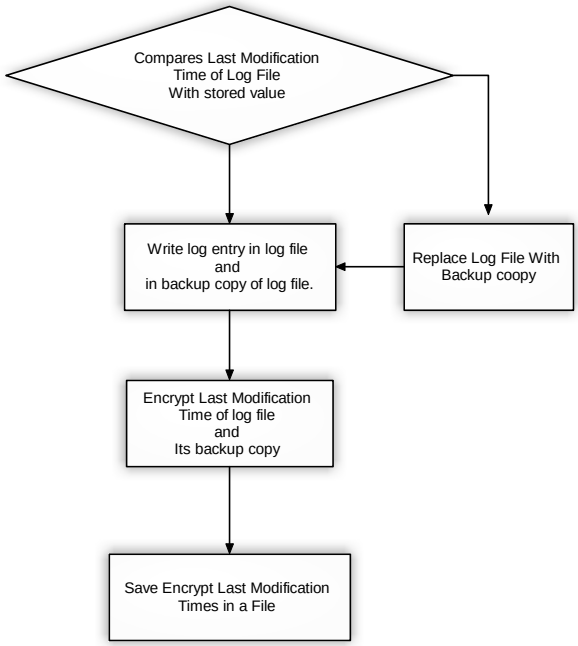


Figure - 3: Sequence of Steps to Write Entry in Log File

Before writing any entry IDS process compares the current last modification time of log file and last modification time that was stored in a file as specified in step 3 given above. A mismatch between these two values represents that any other process changed the log file means unauthorized modification of log file. In such a case log file is replaced with its backup copy or if unauthorized access to the backup file is detected, it will be replaced with log file.

IV. IMPLEMENTATION

The proposed architecture has been implemented in Red Hat Linux environment. Our architecture does not emphasis on underlying technology used by host based intrusion detection system to detect intrusion. Rather than construct a HIDS from scratch, we decided to leverage an HIDS within current implementation of our architecture.

The two hypotheses that underlie this dissertation are practical in nature. First, they intend to show that it is feasible to protect HIDS using proposed architecture. Second, proposed architecture does not affect system performance adversely. Therefore, an implementation was a center point for the development of this dissertation and was used both for practical verification of the intended features of the architecture and for aiding in reasoning about and experimenting with its characteristics.

V. NEW PROCESSES PROPOSED IN ARCHITECTURE

Practically MonitorIDS process (introduced in our proposed architecture) has implemented with two sub processes :

1. MonitorProcesses
2. MonitorFiles

To ensure that MoitorIDS processes live forever we run these processes as system processes. When the adversary tries to kill this process kernel automatically restarts it. A respawn labeled entry related to these processes has been written in system file '\etc\inittab' to run these processes as system processes.

MonitorProcesses process ensures that all processes related to OSSEC are always running. If it found any process dead it restarts the process corresponding HIDS process. MonitorProcesses also prevents modification related to these entries in '\etc\inittab' file. In case of deletion or modification of these entries from '\etc\inittab' file, this process writes new entries. The purpose of MonitorFiles sub-process is to protect HIDS files from unauthorized modification. If it detects any modification or deletion, it restores the victim file with genuine file from backup directory.

VI. AFFECTS ON PERFORMANCE OF THE SYSTEM

To evaluate the affects on the performance on system following steps are carried out :

1. To evaluate the affects on execution time of basic Linux commands, some time measurements were carried out.
2. System Monitoring Utility was used to check the changes in utilization of processor due to processes related to our proposed architecture.

A. Affects on execution of Linux commands

For this purpose, most of the time measurements were carried out regarding basic Linux commands (ps, find, who). Each Command was executed 100 times and all corresponding 100 execution times were recorded. The average of these timings has been considered as the Average Execution Time(ATE). Average Execution Time (AET) was calculated twice. In first run, we calculate execution time (ATE1) without running processes (such as MonitorIDS process) related to our architecture. And in second run we calculate execution time (ATE2) while processes related to our architecture were running.

Table-1 Average Execution Time in milliseconds

Command	ps –ef	find / >/dev/null
Number of system calls	536	10055
(a) Average Execution Time (AET 1) (time required to execute on machine while not running processes related to proposed architecture)	25.9	63.1
(b) Average Execution Time (AET 2) (time required to execute on machine while processes related to proposed architecture are running)	30.1	65.5
Overhead relative to (a)	0.16%	0.03%

Table 1 represents average execution time (in milliseconds) and corresponding overhead. Very small fraction of time was observed as overhead due to the MonitorIDS process.

B. Changes in processor utilization due MonitorIDS Process

MonitorIDS process is a very lightweight process that works in iterative manner. In each iteration checks the state (process terminated or running) of HIDS processes and integrity of HIDS related files. One iteration requires approximately .001 Second time for execution.

Figure 3 represents the CPU utilization graph when MonitorIDS process was not running. Because the processor in use is Dual Core two lines represents utilizations of processor. As shown in the graph CPU utilization is between 0% to 20%. Most of the time the CPU utilization is below 10%.

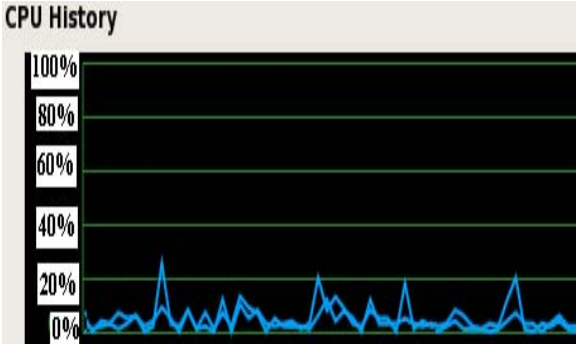


Figure 3: Graph showing CPU utilization while MonitorIDS not running

But as shown in graph given in figure 4, while MonitorIDS process has been running the CPU use is increased to some extent. Most of the time one of the CPU utilization lies between 15 to 20 % and other CPU utilization lies between 5 to 15%.

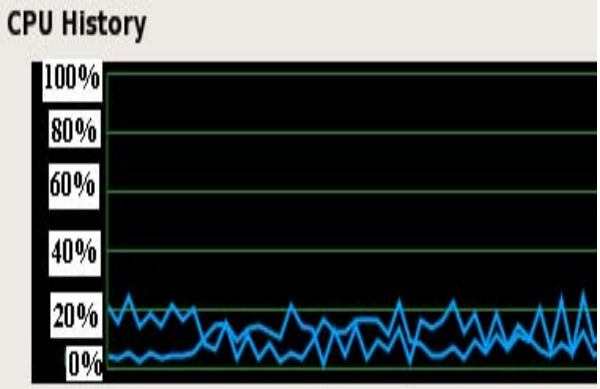


Figure 4 Graph showing CPU utilization while MonitorIDS running

.

As in comparison to graph in figure3 the CPU use is increased about approximately 8 to 10% of total CPU utilization while MonitorIDS process has been running.

VII. RELATED WORK

To protect the HIDS many researchers have given the proposals. In [2], Laureano and Jamhour define the use of virtual machine to protect the HIDS. Virtual machine is used because of its inherent properties like separation of execution space. Other benefits [3] of virtual machine, that are useful in system security are isolation (a process running outside the VM cannot be accessed internal process of VM), inspection (VM state can be accessed by VM Monitor), and interposition (any operation issued by VM can be modified by VM Monitor). The idea behind the architecture is to encapsulate the system activities (both user and guest activities) in virtual machine and place the Intrusion Detection System outside the scope of virtual machine. Now any process has not been able to access the IDS. IDS monitors the activities performed in the virtual machine and identify the intrusive activities. Their approach use type II virtual machine. The architecture protects the IDS from attacks by placing it out from the scope of other processes. However, this approach has one basic problem regarding with theperformance issue. The overhead due to virtual machine degrades the system performance to very worse status. As results given by them if we compare the execution timings of basic routines under an actual machine and virtual environment, there is a large variation. The time taken by routines under virtual machine is much more than time taken on an actual machine.

Table 2 represents a subset of results given by Laureano and Jamhour in [2]. The virtual machine overhead is so high that each routine requires double or even more time for execution as in comparison to time it requires to execute on a normal machine.

Table-2 Virtual Machine Overhead on execution time

Command	ps –ef	find / >/dev/null
Number of system calls	536	10055
(a) Host Time (time required to execute On real machine)	25	125
(b) Guest Time (time required to execute On real machine)	68	484
Overhead relative to (a)	172%	287%

The architecture also affects the network performance. The performance of applications such as FTP and HTML etc. is also turned down.

The work in [4] represents the use of virtual machine type-I. The basic idea is same as to run the HIDS in execution space that cannot be accessible by other processes. However, virtual machine overheads also affect this architecture.

The use of virtual machines for the security of systems has defined by G. Dunlap & et. al.[6]. The proposal defines an intermediate layer between the monitor and the host system, called Revirt. This layer captures the data sent through the syslog process (the standard UNIX logging daemon) of the virtual machine and sends it to the host system for storing and later analysis. However, if the virtual system is compromised, the guest syslog process can be terminated and/or the log messages can be manipulated. by the intruder, and consequently they are no longer reliable.

VIII. FUTURE WORK

Still there are many issues to be addressed about how the proposed architecture can be best implemented and used. MonitorIDS process checks HIDS continuously in iterative way. To check HIDS in an iteration MonitorIDS process requires .001 seconds.

On an average, the adversary has .0005 (.0001/2) seconds to disable our architecture and HIDS by executing following steps :

1. Terminate the MonitorIDS process
2. Remove the entry related to MonitorIDS process from /etc/inittab file to stop the Operating system to restart MonitorIDS process.
3. Terminate HIDS processes or change files related to HIDS.

A DFA (Deterministic Finite Automata) shown in Figure 4, represent such sequences of steps. State q1 is the initial healthy state and q4 is the final state. After reaching at q4 state attacker can tamper the HIDS. To avoid this case, there should be some mechanism that

prevents transitions of system states from q1 to q4. As discussed above in average case transition from q1 to q4 is only possible if made in .0005 seconds. Such mechanism can be based on detecting system call sequence required for this transition. If any such sequence is detected it should delay (approximately .0005 execution of these system calls so that before the system state will change from q2 or q3 to state q4, transition from q1 or q2 to state q1(initial healthy state) took place. This mechanism can be implemented by changing operating system kernel.

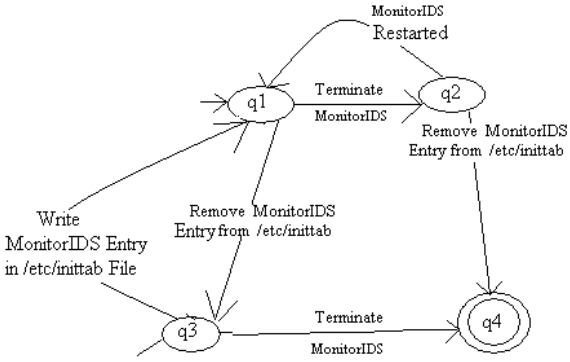


Figure-4: DFA representing sequence of steps to disable our architecture.

Issue of authorization and integrity of genuine updates of HIDS information is another important issue to be addressed. In this approach, we did not consider any technique to authorize of update. Because IDS's rules/signature database requires frequent updates, inclusion of a strong authorization mechanism would be required.

IX. CONCLUSION

In this paper, we propose a new time stamping based approach to check integrity of log files of HIDS. This approach can be combined with Detection and Auto Recovery based architecture to Protect Host Based IDS. As discussed in section 4, unlike other approaches our approach does not use virtual machine and hence does not affect system performance adversely.

Our mechanism ensures that HIDS process always live (cannot be killed by adversary) and the information related to HIDS can never be updated by any adversary. This architecture also ensures the integrity of frequently updated log files of HIDS.

REFERENCES

[1] A. Abraham, C. Grosan and C.M. Vide. Evolutionary Design of Intrusion Detection Programs. International Journal of Network Security, Vol. 4, No. 3, 2007.

[2] M Laureano, C Maziero, E Jamhour, Protecting host-based intrusion detectors through virtual machines- Computer Networks- Elsevier, 2007.

[3] P. Chen, B. Noble, When Virtual Is Better Than Real, Workshop on Hot Topics in Operating Systems, 2001.

[4] T. Garfinkel, M. Rosenblum, A virtual machine introspection based architecture for intrusion detection, ISOC Network and Distributed System Security Symposium (2003).

[5] S. Axelsson. Research in intrusion detection systems: A survey. Technical report, Chalmers University of Technology, 1999.

[6] G. Dunlap, S. King, S. Cinar, M. Basrai, P. Chen, ReVirt: Enabling Intrusion Analysis through Virtual-Machine Logging and Replay, USENIX Symposium on Operating Systems Design and Implementation, 2002.

[7] Surinder Singh khurana, Ms. Divya Bansal, Prof. Sanjeev Sofat "Detection and Auto Recovery Approach to Protect Host Based IDS" 2009 IEEE International Advance Computing Conference (IACC 2009)

Software Radio

Varun Sharma
Infosys Technologies Limited, Chandigarh, India
Email: Varun_sharma15@infosys.com

Yadvinder Singh Mann
Infosys Technologies Limited, Chandigarh, India
Email: YadvinderS_Mann@infosys.com

***Abstract*— This paper aims to provide an overview on rapidly growing technology in the radio domain which overcomes the drawbacks suffered by the conventional analog radio. This is the age of Software radio – the technology which tries to transform the hardware radio transceivers into smart programmable devices which can fit into various devices available in today's rapidly evolving wireless communication industry. This new technology has some or the entire physical layer functions software defined. All of the waveform processing, including the physical layer, of a wireless device moves into the software. An ideal Software Radio provides improved device flexibility, software portability, and reduced development costs. This paper tries to get into the details of all this. It takes one through a brief history of conventional radios, analyzes the drawbacks and then focuses on the Software radio in overcoming these short comings.**

I. HISTORY

Radio, as anyone would perceive, is a device that can wirelessly transmit or receive signals in the radio frequency (RF) part of the electromagnetic spectrum. J.C Maxwell postulated the theory of electromagnetic wave propagation which was confirmed by H Hertz. These electromagnetic waves travel through space either directly, or have their path altered by reflection, refraction or diffraction. When EM Waves come in contact with a conductor; they get converted to an electrical energy, radio or micro wave depending on its wavelength. Radio waves can carry information by varying a combination of the amplitude, frequency and phase of the wave within a frequency band. Nikola Tesla and Guglielmo Marconi invented devices that used radio waves for communication.

Earlier, radio systems relied entirely on the energy collected by an antenna to produce signals for the operator. Radios became more useful after the invention of electronic devices such as the vacuum tube and later the transistor, which made it possible to amplify weak signals.

Today most of our gadgets contain a radio system in it. From a cell phone to television , a walkie-talkie in a toy, space vehicles, cell phones or Satellite phones, Audio/Video broadcasting, Navigation Systems, RADAR and many other applications. Traditional Radios were built only for a particular frequency range, modulation type, and output power. "Figure 1. shows a typical traditional radio receiver.

The RF signal is converted to Intermediate Frequency (IF) using a programmable local oscillator. The IF signal is a fixed frequency and the IF signal is then amplified and filtered before feeding it to demodulator. Each kind of radio will have its own type of demodulator to extract the audio data for playing it on speaker. These radios would require hardware changes to modify these fundamental characteristics like frequency range, modulation type etc. Moreover, with the advancement of technology, new communication standards keep coming and are used to varied degree in different countries e.g. CDMA, GSM EDGE, 3G, etc. The Conventional radios cannot cope with these advances, due to compatibility issues. Hence, there was a need for dynamically configurable radios which could be used for various applications by simply re-configuring the software running in them. They can also be made to comply with the various communication standards, just by changing the implemented protocol. This is the world of Software Radio.

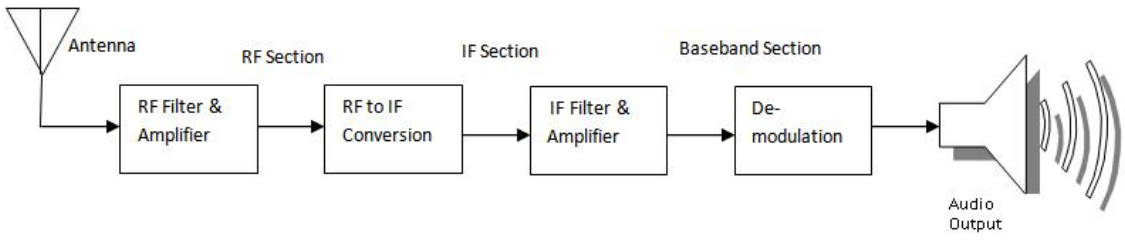


Figure 1. Block Diagram of a Traditional Radio Receiver

doi:10.4304/jetwi.2.2.115-121

II. What is a Software Radio?

Joe Mitola, who came up with the term Software Radio defines it as "A radio whose channel modulation waveforms are defined in software. That is, waveforms are generated as sampled digital signals, converted from digital to analog via a wideband DAC and then possibly up converted from IF to RF. The receiver, similarly, employs a wideband Analog to Digital Converter (ADC) that captures all of the channels of the software radio node. The receiver then extracts, down converts and demodulates the channel waveform using software on a general purpose processor." [2]

The software radio, by his definition, should have as little hardware as possible.

A Software Radio consists of a receiver and a transmitter. For the reception case as shown in “Figure 2. ” [1], when the Antenna receives the signal, it passes the signal to the receiver section for filtering and separating the signal from noise and interference. The signal is then converted to a desired frequency and amplitude which is compatible with the analog to digital converter (ADC), and finally to digital data. The digitized data is then processed using digital signal processing techniques for further use.

For the transmission process as shown in "Figure 3. [1], the software programmable digital signal processing techniques are used for generating the modulated signal(s) in the digital domain. The digital samples are then converted to analog signal using Digital to Analog Converter (DAC) and then amplified to the appropriate voltage, current or power before transmission through the antenna.

III. Components of a Software Radio

Software Radio provides a flexible radio architecture that makes it re-programmable for different usages with no or minimal change in the hardware. This is supported by features like interference rejection techniques, encryption, voice recognition and compression, software-enabled power minimization and control, different addressing protocols and advanced error recovery schemes.

The technology provides the flexibility to combine different grades of hardware and software to strike the right balance between cost and network resilience. The reconfigurable blocks in Software Radio allows easy changes to the radio’s fundamental characteristics such as modulation types, operating frequencies, bandwidth, multiple access schemes, source and channel encoding/decoding methods, frequency spreading/de-spreading techniques, and encryption/decryption algorithm.

The various components of a software radio are shown in “Figure 4. and described below:

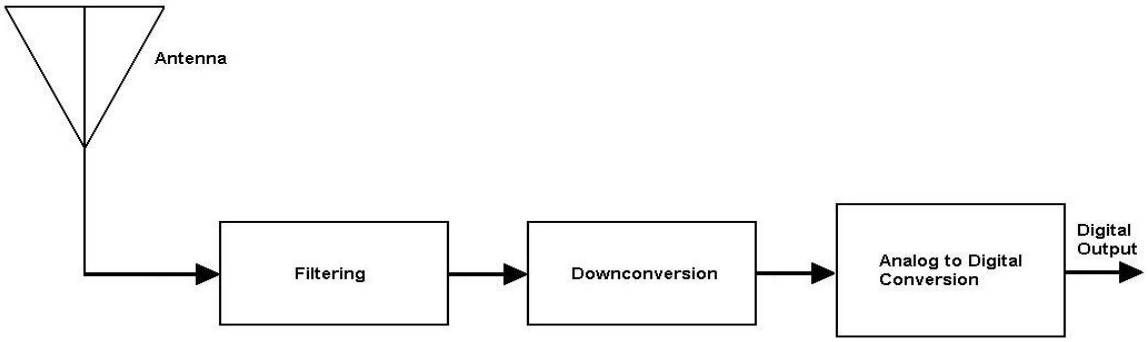


Figure 2. Radio Signal Reception Process

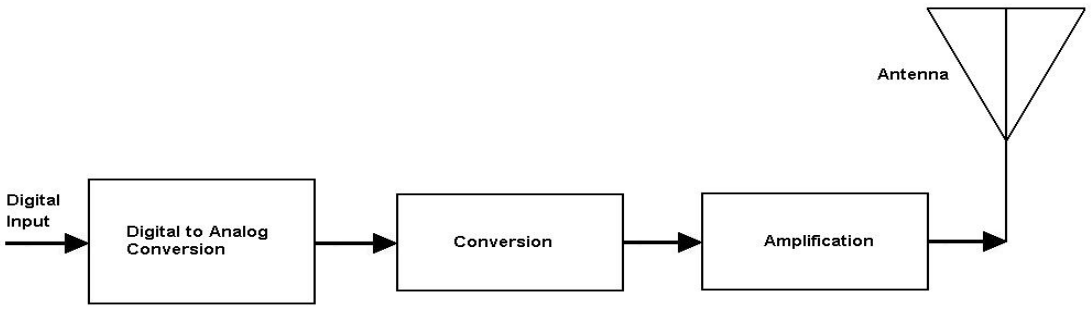


Figure 3. Radio Signal Transmission Process

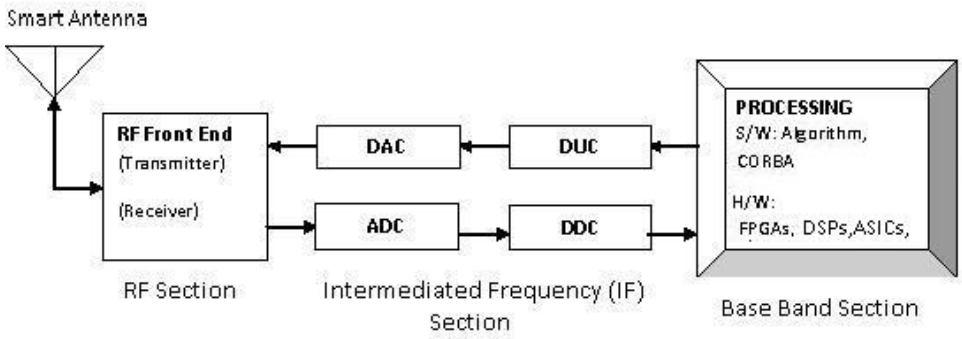


Figure 4. Block diagram of Software Radio

A. Smart Antenna

Antennas are conducting devices which transmit or receive electromagnetic radiations. An antenna array consists of distributed antenna elements whose outputs are combined and is a practical tool for enhancing wireless system performance. The choice of an antenna for a software radio is crucial as it is expected to support multiple bands [1].

It is an antenna array system with in-built signal processing "smart" algorithms to help adapt to different signal environments. It can also identify the direction of arrival of the signal; add phases to the signal to create a constructive radiation pattern to nullify interference. Smart antennas mitigate fading through diversity reception and beam-forming while minimizing interference through spatial filtering. Software radios provide the flexibility needed for effective smart antennas and smart antennas put help in effective implementation and utilization of software radio.

B. RF Front End

It appears before the Intermediate Frequency change state and after the signal is received from the antenna. It does the job of converting the incoming signal to IF frequency by using a tunable local oscillator.

RF front end design process for a software radio has the challenge of catering to a variety of waveforms with widely changing parameters such as amplitude, frequency, phase etc. Due to this wide spectrum, the presence of noise and interference is manifold. This makes the process of achieving a dynamic range even more difficult. The dynamic range is the measure of the highest and lowest level signals that can be simultaneously contained in a radio. In case of mobile communication, increase in dynamic range would mean more battery consumption, which becomes a major trade-off for mobile phones.

Low level signals suffer from the problem of noise. Noise enters at the bottom of dynamic range due to the thermal effects of the components, deviation in quantization or sampling aperture jitter in an ADC. High level signals are limited by interference. Interference is caused at the high end due to adjacent channel, co-channel or is self induced by transceiver. Traditional wireless communication receivers require single RF front end for each channel. Software radio, however, has only one wideband RF front end which can digitize desired signal into separated channels via software to provide a low cost solution.

C. Analog to Digital Conversion

Before all the processing can be done in the software, the analog signal received at IF stage needs to be converted to digital samples by use of Analog to Digital converter. Ideal Software Radio would require a wide bandwidth and good dynamic range from the ADC. As shown in "Figure 5. , the translation of the signal from analog to digital is performed via sampling and quantization. Sampling changes the signal that exists continuously in time to a signal that is non zero only at discrete intervals of time. Quantization changes the continuous valued signal to discrete valued signal.

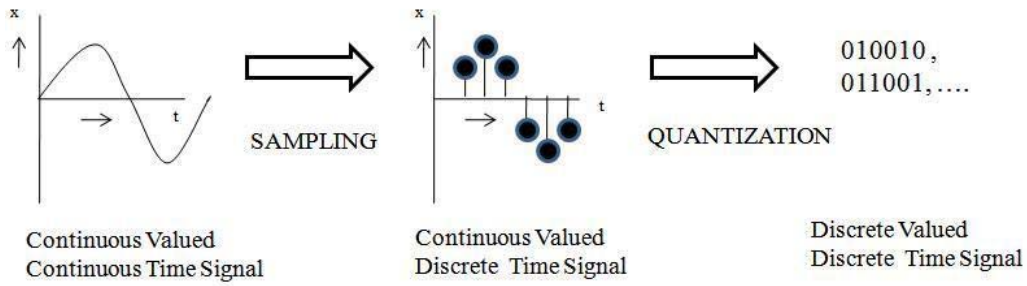


Figure 5. Analog to Digital Conversion

In an ideal software radio, Analog to Digital conversion should occur near the RF end. This would require the following criteria's to be met:

- A very high or variable sampling rate to support wide signals bandwidth.
- Greater number of quantization bits to support a high dynamic range.
- Operating bandwidth of several GHz to allow conversion of a signal over a wide range of frequencies.
- Lesser distortion and wide dynamic range to allow recovery of low level signals in the presence of strong interferers.
- Economical components to meet the above criteria's without consuming an excessive amount of power.

Due to the limitations of the fabrication technology, few of the above mentioned criteria's need to be traded off for an acceptable design solution for the software radio.

D. Digital to Analog Conversion

In Software Radio, entire waveform and modulation is generated in digital domain and is then fed to Digital to Analog Converter before final transmission. As shown in "Figure 6. , the translation of the signal from digital to analog is performed via voltage mapping and reconstruction, where the digital signal is changed to a continuous valued signal. The number of bits and the frequency range of the DAC are the important factors as they would determine the reconstruction of the continuous signal.

In a traditional analog radio, the frequency synthesis is performed via bulky devices like quartz crystal, inductor, capacitor, mechanical resonators etc. But different frequencies and waveforms can be generated in the digital domain by Direct Digital Synthesis (DDS) techniques and it is fast replacing the analog devices as they provide better accuracy, frequency can be changed with software and is simple to implement.

E. Digital Down Conversion

A digital down converter (DDC) provides the link between the analog RF front end and the digital baseband of a receiver. The received signal is usually sampled at much higher sampling rates than required to relax the specifications of the anti-aliasing filter required. But to reduce the computational power required from Digital Signal Processors (DSPs), the sampling rate is immediately reduced by down converter. DSP requirements are directly proportional to sampling rate and reducing sampling rates reduces CPU cycles required to perform the same digital signal processing algorithms and thus saving significantly on cost and DSP power consumption.

F. Digital Up Conversion

A digital up converter (DUC) provides the link between the digital base band and analog RF front end and is required on the transmitter end. The signal to be transmitted is generated by digital signal processing at lower sampling rates to reduce computations of DSP. But before it can be fed to DAC for converting to analog signal, it is up converted to higher sampling rate by Digital Up Converter to relax interpolation filter specifications.

G. Digital Signal Processing

DSP is the brain behind all digital radio technology. A DSP core consists of an arithmetic logic unit (ALU), accumulator(s), multiply and accumulate MAC unit(s), data and the address buses. A DSP is designed to support high performance, repetitive, numerically intensive tasks and very high I/O performance. Large accumulators in DSPs help reduce the precision problems. The digital signal processor performs all the functions of filtering, demodulating, generating the signal for transmission using various modulation techniques in software radio. DSPs also perform functions of data compression, encryption and special functions like speech recognition, image enhancement, neural networks for artificial intelligence etc. The three main categories of digital hardware are ASICs (Application Specific Integrated Circuit), FPGAs (Field Programmable Gate Array) and DSPs. A DSP represents the most generalized type of hardware that can be programmed repeatedly using high level programming language to perform various functions. An ASIC is the most specialized piece of hardware and can be used only in the specific application for which it has been designed. ASICs are generally used when DSP has insufficient processing power and function to be performed is fixed. The set-up of ASICs is implemented on fixed silicon, which optimizes the speed and power consumption of the circuit.

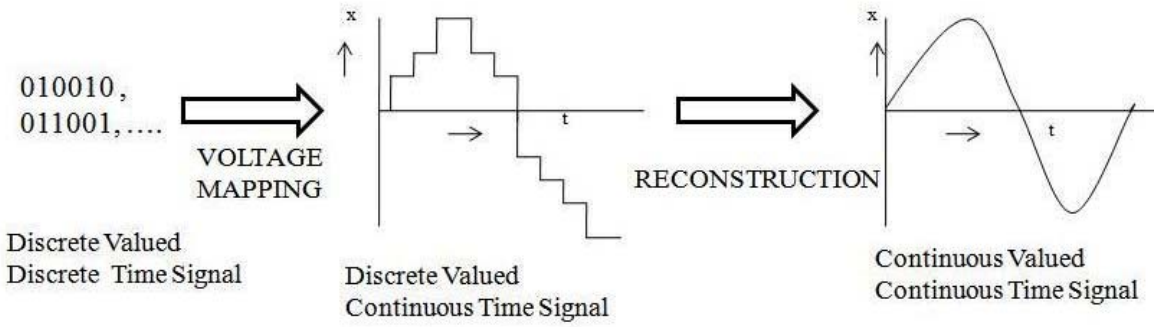


Figure 6. Digital to Analog Conversion

FPGAs help conserve silicon area since one chip can be configured to perform more than one function hence the configuration can be done at run time. Gate arrays offer higher degree of parallelism in comparison to DSP. A hardware description language (HDL) is used for designing highly complex circuitry. FPGAs are more flexible than ASICs but lesser than DSPs. FPGAs are being used because of their ability to perform high-speed parallel multiplication and accumulation functions and are especially well suited to handle algorithms like Finite Impulse Response (FIR) filters (for decimation or interpolation) and transforms like Fast Fourier Transformations (FFT) and Discrete Cosine Transforms (DCT). Also they have the advantage of providing flexibility to integrate logic design with signal processing. These three hardware components constitute a design space which trades flexibility, processing speed and power compensation.

In an ideal scenario, Digital signal processing (DSP) software should help to perform most of the radio functions at astonishing performance levels, enhance the performance using digital filtering and drastically reducing noise and interferences, all of it with less bulky equipments.

IV. Modulation and Demodulation

To understand the advantage and ease with which modulated signals can be generated and demodulation can be done in digital domain using digital signal processing, an example of **Amplitude Modulation** and **Demodulation** is discussed below.

A. Modulating the signal

The carrier signal as shown in "Figure 7. , can be generated using either DDS technique or using sine series library function. The modulating signal will be the output of microphone amplifier sampled at a particular rate by ADC. Here for illustration purposes, carrier frequency of 1000Hz and modulation signal of 20 Hz has been generated using the following equations with sampling rate of 10 KHz:

```
for ( i = 0; i < 1000, i++)
{
        Carrier = Sin(2 * PI * i * Fc / SR);
        ModS = 1 + Sin(2 * PI * i * Fm / SR);
}.
```

(1)

In (1) above,
SR is sampling Rate = 10KHz,
Fc = 1000Hz (Carrier),
Fm = 20Hz (Modulation signal).

Amplitude Modulation involves changing the amplitude of carrier as per the amplitude of the modulating signal. Since we have the samples of carrier and modulating signal in digital domain, the amplitude modulation in digital domain is simply multiplication of these samples. "**Error! Reference source not found.**", shows the generation of modulated signal.

$$ModSignal = Carrier * ModS. \quad (2)$$

B. Demodulating the received signal

At the receiving end when this signal is received, it is full wave rectified which in digital domain is nothing but taking the absolute value of all the samples. "**Error! Reference source not found.** shows the output of full wave rectifier.

$$RectSignal = abs(ModSignal). \quad (3)$$

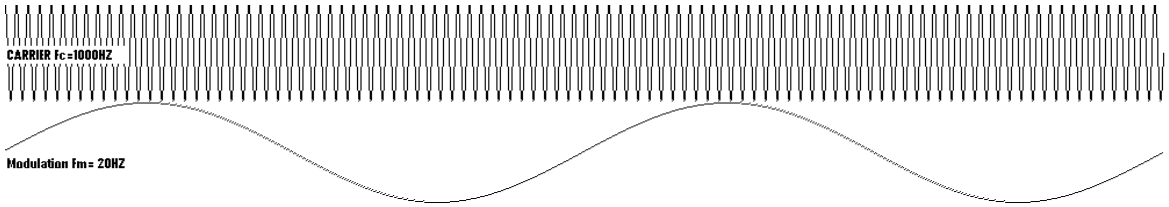


Figure 7. Generation of carrier and modulation signals

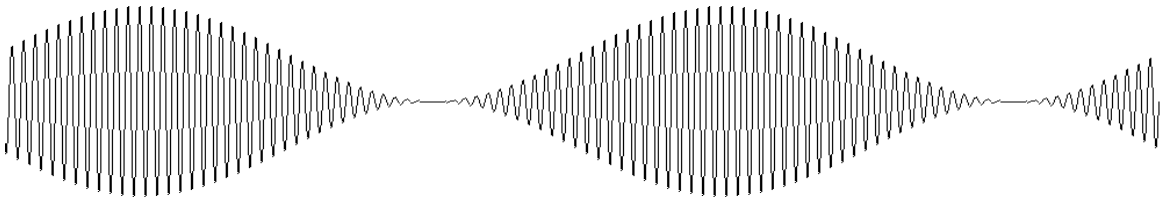


Figure 8. Amplitude modulated carrier

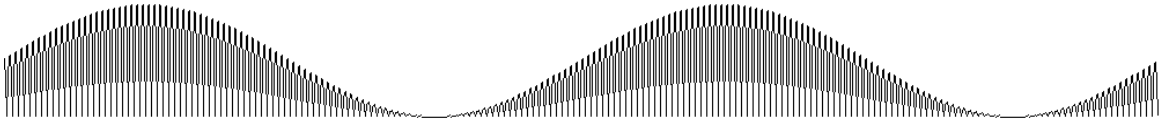


Figure 9. Full wave rectified AM waveform

Now the modulated signal is contained in the form of envelope of full wave rectified signal and can be recovered by applying a low pass filter to remove the carrier frequency. "**Error! Reference source not found.**shows the output of a first order Infinite Impulse Response (IIR) filter. The equation of filter used is:

$$Yn = 0.997*Yn-1 + 0.003*Xn. \quad (4)$$

In (4) above,
Xn is new input sample (RectSignal),
Yn-1 is previous output sample,
Yn is new output sample (demodulated signal).

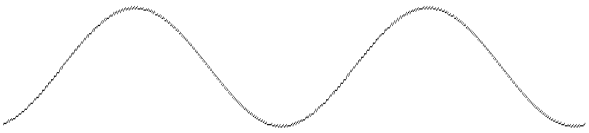


Figure 10. Demodulated signal after low pass filter

V. Software Radio as a solution to Conventional Analog Radios

A Software Radio in contrast to a traditional Radio runs on a generic hardware platform to perform all the modulation/demodulation, frequency selection, filtering etc in digital domain as also explained above in the example. It can perform role of a cordless phone, cell phone, internet access tool, a GPS receiver etc. all in one. This is possible by changing the software used for processing. Hence, there is no limitation of any particular frequency range, modulation type, and output power in a Software Radio. All these flexibilities of software radio offer advantages which can be broadly classified as:

A. Frequency-Agile

The software programmable processing allows it to operate in the entire radio waves frequency bands.

B. Easy to plug n play

It is a radio into which multiple contractors could plug parts and software.

C. Global Mobility

It caters to the model as it is flexible to support most of the communications standards like CDMA, GSM, IS-136 etc.

D. Inter-operability

It can be made to inter-operate with different wireless protocols, incorporate new services, and upgrade to new standards etc by simply downloading a new program (or a newer version of the software) or by change of selection from the user interface.

E. Compact and Power Efficient

Due to reduced number of components used in Software Radio, its size is reduced resulting in ease of manufacturing. Also less hardware results in less power consumption.

F. Quality of Service

As Software Radio is Software dependent; hence a greater quality of service is delivered consistently.

VI. Applications of Software Radio

The ultimate goal of Software Radio is to provide a single radio trans-receiver which can play the roles of cordless telephone, cell phone, wireless fax, wireless e-mail system, pager, wireless videoconferencing unit, wireless Web browser, Global Positioning System (GPS) unit, and other functions still in the realm of science fiction, operable from any location on the surface of the earth, and perhaps in space as well. Few of the applications of software Radio has been covered henceforth and shows the variety of functionality to which it can cater to.

A. Pacemakers and Implantable Cardiac Defibrillators

Wirelessly reprogrammable and implantable medical devices (IMDs) such as pacemakers, implantable cardioverter defibrillators (ICDs), neurostimulators, and implantable drug pumps use embedded computers and radios to monitor chronic disorders and treat patients with automatic therapies. For instance, an ICD that senses a rapid heartbeat can administer an electrical shock also known as the software radio attacks to restore a normal heart rhythm, then later reports this event to a health care practitioner who uses a commercial device programmer with wireless capabilities to extract data from the ICD or modify its settings without surgery. Clinical trials have shown that these devices significantly improve survival rates in certain populations. [4]

B. Smart sensors

Software Radios can be implemented as smart sensors to determine the amount of various particles present in atmosphere, hence useful for pollution check, weather forecast etc. When used as Bio-sensor with configurable software, it can be used in study of complex anatomy.

C. One Cell phone for many standards

It is highly inconvenient for a regular inter country traveler to change his mobile connection for communication due to network issues. For example, 3GPP UMTS is popular in European nations, while C/TDMA standards are widely used across North and South America. So, every time one travels through a country he has to have the country specific cellular phone. But with the coming of Software radio, it can be a thing of past as the Software Radio can adapt to change in the local air interface by a simple software change. All the base stations and terminals using Software Radio architecture can support multiple air interfaces during periods of transition and be easily upgraded.

D. Cost Effective Up gradation of a Cellular Base Station

The growth in radio technology is in-step with the demand of today's customer, but is the business ready? The customer may have asked for audio files in his

cellular services yesterday, but today they look for video streams in their phone and tomorrow they might ask for live data feed. The cellular industry cannot afford to lose business by not making such changes. Software Radio helps to save them millions by easy upgrades in software. There is no need for investment in the hardware, saving them both money and time. It is also a boon for customers as they don't need to purchase newer handsets for newer technology. Just as any individual or business can update a software program used on a PC, by using Software Radio technology, any cellular providers can upgrade software easily and quickly to make changes on their systems. As a result, the cost for cellular phone service can continue to be decreased for consumers; while the providers can see higher margins for the service they provide.

E. Usability in Emergency

As the software radio is compact and easily deployable, it can be useful for temporary coverage during emergencies and special events. A software radio can be programmed and mounted onto a vehicle and driven to the site of a natural disaster to bring in communications where they may have been damaged and to support multiple standards.

F. High Definition TV Applications

Software Radios are finding significant applications in the conversion of both analog and digital transmissions to HDTV, an upcoming broadcast technology.

VII. Challenges & Limitations

Software radio technology alters traditional radios in three ways namely moving A/D closer to antenna, substitute hardware with software processing and facilitating transition from dedicated to general-purpose hardware. This decreases the number of hardware components resulting in high degree of extensibility for a wide range of features. This will also eliminate the redundant hardware as found in dual band phones etc.

There are several technological challenges faced in Software Radio today, like the usage of faster A/D converters, requirement for less power consumption, implementing smart antennas and making the overall radio more cost effective. ADCs which can digitize the high range signal are very expensive. They also increase the level of SFDR (spurious free dynamic range). Hence, the signal received are down converted from RF signal and digitized at IF range via ADCs before sending the data to DSPs. To overcome this limitation, more research and development would be required to build less expensive ADCs covering more dynamic range and supporting greater sampling rates. For FPGAs, interval to reprogram FPGA's limits its usability.

The authenticity and security of downloaded software is another area of concern. There is a need for a globally recognized regulatory body which can control these processes. Certification from bodies like Federal Communications Commission (FCC) is another hurdle. FCC certification process presumes that any software that is bundled with the hardware be certified as a unit. But software radio being programmable at any time poses challenges in certification. A new certification procedure has been created. Software changes which can affect radio frequency, power and modulation are subject to more streamlined process. However, FCC does not certify any changes by third party software vendors. All modifications are the responsibility of the original manufacturer whose original hardware/software radio bundle was certified. This requires that all the independent vendors need to work via equipment manufacturers. Hence, certifying software radios pose a significant challenge because of cumbersome certification processes, especially for domains like aerospace and defense which require testing for all possible scenarios where software may misbehave.

VIII. Future of Software Radio

Ultimate aim of software radio would be to move ADC and DAC just next to Antenna and doing everything in software. Currently, the cost of high frequency ADCs and specialized hardware for digital signal processing are few of the factors limiting the application of software radios to high end use only. As the technology advances, the costs will come down and processing power of general purpose computing platforms will increase. Once the processing power crosses the threshold required for performing necessary radio functions, broader applications and scales of economy will come into place and we will see software radios being used in our day to day use gadgets. How soon it can happen is still an open question but software radio is likely to be an important technology in the years to come.

References

[1] Jeffrey H. Reed, "Software Radio".
[2] J Mitola, "The Software Radio", *IEEE National Telesystems Conference, 1992* - Digital Object Identifier 10.1109/NTC.1992.267870.
[3] Ulrich L. Rohde, "Digital HF Radio: A Sampling of Techniques", *Ham Radio Magazine*, April, 1985.
[4] http://www.secure-medicine.org/icd-study/icd-study.pdf
[5] "A Software-Defined Radio for the Masses, Part 1- 4", Gerald Youngblood, AC5OG
[6] http://www.sdrforum.org/
[7] http://rfdesign.com/mag/radio_fast_hot_data/index.html
[8] http://www.dspguide.com/

SDR and Error Correction using Convolution Encoding with Viterbi Decoding

Shriram K. Vasudevan
VIT/Embedded Systems, Vellore, India
Email: shriramkv@rocketmail.com

Siva Janakiraman and Subashri Vasudevan
SASTRA/Communication, Tanjore, India
SASTRA/Computer Science, Tanjore, India
Email: sivajanakiraman@gmail.com,vrsuba@gmail.com

Abstract— *The aim of the paper is to build a software based radio capable of demodulating Frequency Modulated signals with the ability to receive upto four stations simultaneously. In addition convolution encoding along with Viterbi decoding is also implemented to aid in error correction capability of digital transmission systems. The initial phase is aimed at software based demodulation of multiple channels of Frequency Modulated signals. The second phase is directed towards the Error correction using convolution encoding with Viterbi decoding for streaming data encountered in digital broadcast systems.*

Software Radio is a way of designing software very close to the antenna. The basic feature of software radio is that software will define the transmitted waveforms, and software will demodulate the received waveforms. This paper is developed with the help of free software based radio toolkit given by a community named GNU radio. Channel coding schemes need to be used extensively in the case of transmission of digital data over the air or through any other medium. These coding schemes also called FEC (Forward Error Correction) are used in cases where the re-transmission of the data is not feasible or possible. The channel coding schemes comes to the rescue in such cases where redundant data are sent over the transmission medium along with the message. In the receiver side, this channel coded signal is decoded to get back the original data even if the channel coded signal undergoes some interference from the noise in the transmission medium. Though the channel coding has a downside of requirement to transmit additional data over the transmission medium, the advantages offered by error detection and correction mechanisms to the receiver outweighs the disadvantage of additional data transfer rate and bandwidth requirements.

Index Terms—**SDR, Viterbi Decoder, Convolution Encoder**

I. Introduction

Software-Defined Radio (SDR) is an evolving technology that has received enormous recognition. In the past few years, analog radio systems are replaced by digital radio systems for various radio applications in civilian, military and commercial spaces. In addition to this, programmable hardware modules are playing a vital role in digital radio systems at different functional levels. SDR totally aims to take benefits of these programmable hardware modules to get open-architecture based radio system software.

SDR technology supports implementation of some of the functional modules in a radio system such as modulation/demodulation, signal generation, coding and link-layer protocols in software. This assists in getting reconfigurable software radio systems where dynamic selection of parameters for each of the above-mentioned functional modules is possible. A Complete hardware based system has many limitations.

Before getting into what software radio does, it is good to review the design of a traditional analog, hardware-based radio (Figure.1). In wireless communications, information is encoded into radio waves. These are collected (or transmitted) from (to) the air by the antenna. The received signal is then passed to a series of components that extract the useful information and convert it into the output of the radio. The basic design is the same whether the radio signal is destined for a cell phone, microwave repeater, or AM/FM car radio. Traditional radios are based on the super heterodyne (superhet) receiver circuit.

Manuscript received August 2, 2009; revised August 20, 2009;

doi:10.4304/jetwi.2.2.122-130

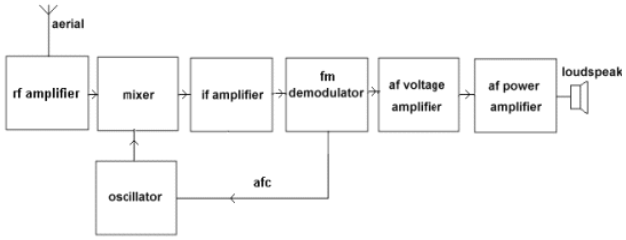


Figure 1. Generic FM receiver

In the superhet receiver, the incoming signal, which is Radio Frequency (RF), is first down-converted to a lower intermediate frequency (IF). The IF is then filtered for noise and amplified before being demodulated to produce the baseband signal that represents the desired information. This analog baseband signal may be passed directly to further downstage processing, or it may first be digitized and subjected to additional signal processing. There are several important reasons for down converting to a lower and standardized IF:

(a) It is easier and hence less expensive, to build filters and amplifiers – especially linear amplifiers – for a lower frequency signal.

(b) Use of a common IF enables standardization in the design of radio components.

Traditional designs for implementing the super heterodyne receiver architecture were optimized for specific frequencies and applications and each of the stages were implemented in hardware that was closely coupled. This was due largely to the difficulties inherent in and limitations in the state of the art in the design of analog signal processing components. Analog processing is much more complicated than digital processing, which is one of the key reasons why the transition to digital signals is so important.

In this paper chapter 2 deals about the performance and power limitations, Chapter 3 talks about architecture of SDR. Chapter 4 is on signal flow, Chapter 5 is dealing with how to derive instantaneous frequency. Chapter 6 talks about De Emphasizer. Chapter 7 explains about multi channel FM reception. Chapter 8 will give details about the Forward Error Correction. Chapter 9 and 10 speaks on implementation and results. Chapter 11 is concluding the paper followed by references.

II. Performance and Power Limitations

The change from hardware to software radios has faced lot of problems. First, performance generally comes down in the shift from dedicated to general-purpose hardware.

Secondly, the transition from hardware to software processing results in a substantial increase in computation which ultimately results in increased power requirements. This reduces battery life and is one of the key reasons why software radios will not be deployed first in end-user devices such as cell phones, but rather in base stations which can take advantage of external power sources.

III. Architecture

We now take a view at the overview of a basic conventional digital radio system and then look at how SDR technology can be used to implement radio functions in software. This will be followed by the software architecture of SDR. The definition of the software radio system is not absolute and depends on where the Analog to digital conversion is performed. The question of where the A/D conversion is performed determines what radio functions can be moved into software and what types of hardware are required. At one extreme, we consider calling it software radio if software is used at any stage within the radio. This case typically involves the digitization at the baseband for signal processing. In this case the actual demodulation process is performed before the conversion to digital format. This setup is more of a digital radio than software radio. What process can be done using software is very much limited. Since the actual demodulation itself takes place with hardware, it cannot be used for demodulation of other kinds of signal and hence is very limited in its functionality.

At the other extreme, we might choose to use software radio for cases in which A/D conversion is performed right at the antenna with all radio functionality implemented in software running on general-purpose hardware. In this case the Digitization of the signal takes place right at the antenna and thus gives us the complete flexibility. This setup requires the use of a very high speed and wideband ADC, not economically feasible at this point in time. Figure 2 shows the basic software radio architecture.

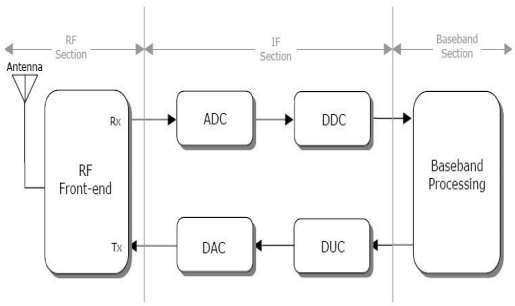


Figure.2 Basic Software Radio Architecture

In this paper, we take a stance that lies in between the two cases described above. The digitization of the signal is performed in the Intermediate frequency band, followed by the digital down conversion. The baseband processing is done in a general purpose computer. Though this setup does not give us a complete software control over the radio spectrum, it does offer us a great level of flexibility with the software since we can actually demodulate the signal using software. The architecture used in this project is indicated in Figure 3. This helps us to explore the part of radio spectrum which falls in the IF region that can be digitized.

The digital radio system consists of three main functional blocks: RF section, IF section and baseband section. The RF section consists of essentially analog hardware modules while IF and baseband sections contain digital hardware modules. The RF section (also called as RF front-end) is responsible for transmitting/receiving the radio frequency (RF) signal from the antenna via a coupler and converting the RF signal to an intermediate frequency (IF) signal. The RF front-end on the receive path performs RF amplification and analog down conversion from RF to IF. On the transmit path, RF front-end performs analog up conversion and RF power amplification.

The ADC/DAC blocks perform analog-to-digital conversion (on receive path) and digital-to analog conversion (on transmit path), respectively. DDC/DUC blocks perform digital-down conversion (on receive path) and digital-up-conversion (on transmit path), respectively. DUC/DDC blocks essentially perform modem operations, i.e., modulation of the signal on transmit path and demodulation of the signal on receive path.

For this paper work, the RF section and the IF section are handled by the hardware, followed by the baseband processing by Computer. The hardware used is known as the Universal Software Radio Peripheral (USRP).

III. Software based FM receiver

The real implementation of the project starts from here. The hardware used is a custom built board provided by the creators of the GNU Radio community, which is designed to run the free software toolkit provided by them. The main goal is to take the FM based demodulation one step further to receive up to four FM stations simultaneously.

What the USRP gives to the Computer is a Digital-Down converted, complex, quadrature signal in the Baseband. The remaining processing will be taken care by the software after getting the signal from the USRP board. Thus once the signal enters the computer, the demodulation of the FM signal consists of the following steps.

(a) Signal flow from the air to the computer (from real to complex)
(b) Getting the instantaneous frequency (from complex to real)
(c) De-emphasizer
(d) Audio FIR decimation filter
(e) Output to Soundcard/File

Thus each stage of this signal processing acts as a block with input and output ports. Each block receives the signal as input from the previous block, performs the required signal processing and gives the output to the next block in the chain.

The software architecture is implemented in two layers. The bottom layer is implemented in C++, which satisfies the performance requirement of the demodulation and signal processing. The top layer is implemented in python, which acts as a mask layer, interconnecting the bottom C++ layers which perform the bulk of the signal processing. This two layered architecture gives us the advantage or reusing the signal processing blocks and makes it easier to customize the signal processing flow necessary. Thus connecting or removing a processing block from the software demodulation chain is much easier. The entire software layer is built and run on a free UNIX port for windows called cygwin. Cygwin offers a UNIX platform on windows based system

IV. Signal Flow from the air to the computer (from real to complex).

Basically what the USRP does is to select the part of the spectrum we are interested in and decimate the digital sequence by some factor N. The resulting signal is complex with I/Q two channels. Thus USRP gives out a `complex' signal, with a data rate 256k samples per second, called `quadrature rate' – or quad rate because the complex signal has I/Q quadrature components (Figure. 3)

.

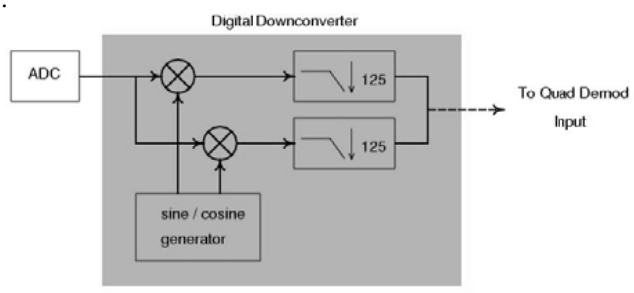


Figure.3 Digital down Converter

V. GETTING THE INSTANTANEOUS FREQUENCY (FROM COMPLEX TO REAL).

For FM, the instantaneous frequency of the transmitted waveform is varied as a function of the input signal. The instantaneous frequency at any time is given by the following formula:

$$f(t) = k * m(t) + f_c$$

m(t) is the input signal, k is a constant that controls the frequency sensitivity and f_c is the frequency of the carrier (for example, 100.1MHz). So to recover m(t), two steps are needed. First we need to remove the carrier f_c, then we're left with a baseband signal that has an instantaneous frequency proportional to the original message m(t). The second step is to compute the instantaneous frequency of the baseband signal.

Removing the carrier is taken care by the USRP, via the digital down converter (DDC). The resulting signal coming into the Computer has already become a baseband signal and the remaining task is to calculate its instantaneous frequency. If we integrate frequency, we get phase, or angle. Conversely, differentiating phase with respect to time gives frequency. These are the key insights we use to build the receiver. The angle between two subsequent samples can be determined by multiplying one by the complex conjugate of the other and then taking the arc tangent of the product.

Thus Arctangent of the product gives the phase difference between adjacent samples, if we divide it by the sample interval or multiply the data rate, we get the radian frequency ω, which gives the instantaneous frequency (f) if further divided by 2ω. This process of recovering the instantaneous frequency from the quadrature signal from the USRP is called the "Quadrature Demodulation" (Figure. 4).

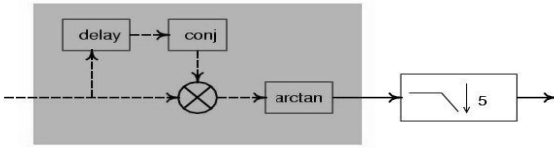


Figure.4 Quadrature Demodulation

VI. DE-EMPLASIZER

It has been theoretically proved that, in FM detector, the power of the output noise increases with the frequency quadratically. However, for most practical signals, such as human voice and music, the power of the signal decreases significantly as frequency increases. As a result, the "Signal to Noise Ratio" (SNR) at the high frequency end usually becomes unbearable. To circumvent this effect, `pre-emphasis' and `de-emphasis' were introduced into the FM system. At the transmitter, proper pre-emphasis circuits are used to manually amplify the high frequency components, and the converse operations are done at the receiver to recover the original power distribution of the signal. As a result, the SNR is improved effectively. In the analog world, a simple first order RLC circuit usually suffices for pre-emphasis and de-emphasis. In the case of the Software defined Radio, a first order IIR filter is implemented to get the magnitude of amplified, high frequency signal down to the magnitude of normal lower frequency signal, thereby improving the SNR.

Let us now see about Audio FIR decimation filter:

After passing the de-emphasizer, it is a real signal with a data rate of 256kHz.It is a baseband signal, containing all the frequency components of a FM station. The bandwidth of a FM station is usually around 2 * 100kHz. A sample rate of 256kHz is suitable for the 200kHz bandwidth, without losing any spectrum information

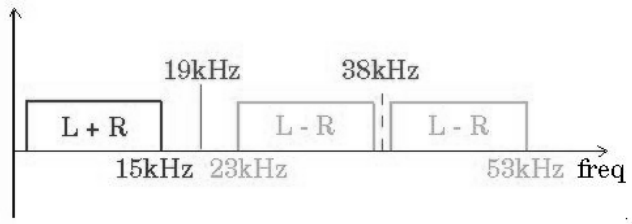


Figure.5 Typical FM Band

The FM signal band is spread out with various band of information (Figure. 5). From 0 to about 16 kHz is the left plus right (L + R) audio. The peak at 19 kHz is the stereo pilot tone. The left minus right (L - R) stereo information is centered at 2x the pilot (38kHz) and is AM-modulated on top of the FM. Additional sub carriers are sometimes found in the region of 57kHz - 96kHz. Thus for the reception of a mono FM station, a Low pass FIR filter with a pass band frequency of 15Khz and a transition band with 1Khz should suffice. Thus, once the signal comes out of the FIR decimation filter block, it is now ready to be played into the soundcard.

Output to Soundcard/File:

Once the signal passes through the FIR Decimation filter, it is ready to be played by the soundcard or alternatively stored in a File. The signal from the FIR decimation filter is in a raw bit format and can be converted to other formats using any suitable compression algorithms.

VII. Forward Error Correction

Forward error correction [FEC] is a system of error control for data transmission, whereby the sender adds redundant data to its messages, which allows the receiver to detect and correct errors (within some bound) without the need to ask the sender for additional data. The advantage of forward error correction is that retransmission of data can often be avoided, at the cost of higher bandwidth requirements on average, and is therefore applied in situations where retransmissions are relatively costly or impossible.

Such a typical situation where the receiver is not able to make a request for the re-transmission of message is FM reception. FEC devices are often located close to the receiver of an analog signal, in the first stage of digital processing after a signal has been received. That is, FEC circuits are often an integral part of the analog-to-digital conversion process.

Forward Error Correction Implemented:

The Forward error correction algorithm consists of two parts namely Encoding and Decoding. The Encoding is implemented using a Convolution algorithm and the Decoding implemented is the Viterbi algorithm. The convolution encoder and the corresponding Viterbi decoder are implemented in C programming language.

Motivation for Forward Error Correction

The situation in which the receiver is not able to make a request to the sender for re-transmission of message forms the basis for all forward error correction techniques. A typical radio receiver is a classic example in which the receiver has no provision to communicate with the sender. Under such circumstances, if an error occurred in the signal transmitted, there will be no way for the receiver to detect the error unless some error control mechanisms are incorporated into the signal before transmission itself. This process of adding error control to the signal before transmission is called Forward Error Correction.

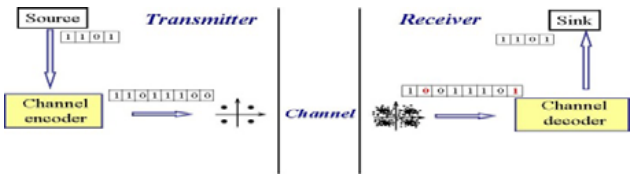


Figure 6. Typical wireless communication channel

A typical communication channel in which the data to be transmitted is encoded using some encoding algorithm by adding some redundant bits (Figure. 6). Thus the output of the encoder is called Channel Symbols. These channel symbols are then transmitted over the wireless channel.

Thus it can be seen that unless the error correction mechanisms are introduced there is no way for the receiver to recover the original information correctly. Hence Forward Error Correction mechanisms are essential if the receiver is to recover the original information. There are numerous forward error correction mechanisms each applicable to particular type of communication. The factors which govern which error correction algorithm is used for a particular communication type are

- Bandwidth of the spectrum
- Depth of error correction required
- Rate of encoding
- Processing capacity of Sender/Receiver

Introduction to Convolution Encoding

Convolution Encoding typically involves encoding of stream of data bits by using previous bits to perform a logical operation, the output of which is the encoded channel symbols. Thus depending upon the type of convolution encoding, a single data bit can have its influence on two or more adjacent bits thereby providing the required redundancy to the data. This influence which a bit has on other bits is what enables the receiver to identify any bit errors that occurred during the data transmission.

A convolution encoder is called so because it performs a convolution of the input stream with encoder's impulse responses. It can be represented mathematically as

$$y_i^j = \sum_{k=0}^{\infty} h_k^j x_{i-k},$$

where x is an input sequence, y^j is a sequence from output j and h^j is an impulse response for output j. A convolution encoder is a discrete linear time-invariant system. Every output of an encoder can be described by its own transfer function, which is closely related to a generator polynomial.

C.1. Convolution Encoding Algorithm in this paper

The convolution algorithm implemented in this project is a Half rate,(5,7) encoder with a constraint length of K=3, which is suitable for streaming data. .The convolution encoding typically involves the process of encoding data by using the bits of data to perform modulo 2 addition, thereby adding redundancy (Figure. 7).

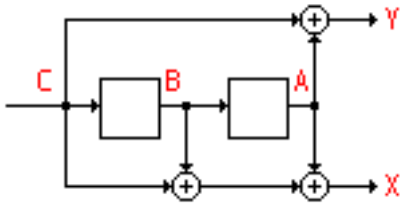


Figure 6. Basic convolution encoding

XIII. VITERBI ALGORITHM

The Viterbi algorithm implemented in this project is suitable for decoding data that has been encoded with a Half-rate,(7,5) convolution encoder with a constraint length of K=3.

IX. IMPLEMENTATION OF ALGORITHMS

The steps involved in simulating a communication channel using convolutional encoding and Viterbi decoding are as follows

1. Generate the data to be transmitted through the channel.
2. Convolutionally encode the data.
3. Introduce channel errors
4. Perform Viterbi decoding on the received channel symbols

X. RESULTS

The results of the project implemented are analyzed in the sections below.

Software Based Demodulation:

Forward Error Correction

The simulation setup consists of a channel encoder which generates random bit streams consisting of 0's and 1's which represents the message bits. This bit stream is then encoded. Then the encoded data is given to a channel error simulator to simulate bit errors. Then the message with errors is given as input to the Viterbi decoder and its error correction capability is analyzed.

Convolution Encoder

The length of the generated bit stream is taken as input from the user. Once the bit stream is generated, the convolution encoding is performed and the encoded data is written to a file. The encoder simulation is shown in figure.7

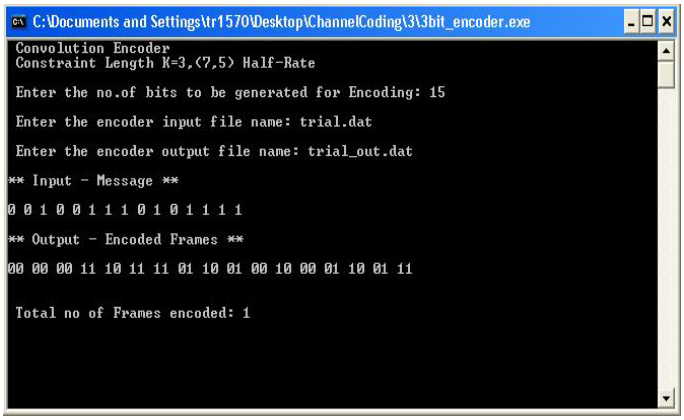


Figure.7 Simulation result of encoder

Since the encoding is typically performed on a stream of data, it must have a known length for the decoder to decode the message properly. Hence the encoder encodes the data in frames. The frame length used in this project is 15 message bits with 3 flushing bits for each frame. Hence the stream of data is broken up into frames of constant length and then encoded. Once the encoding process is completed, the encoded message is then written to a file.

Channel Error Simulation

In order to verify the error correction capability of the decoder, the encoded message must be included with random bit errors to simulate the message transmitted through a noisy channel. A model in which random bit errors occurs is created.

The number of bit errors to be introduced in each frame is taken as an input from the user and correspondingly random bits are inverted i.e. 0's changed to 1's and 1's changed to 0's which provides a reasonable model to simulate channel errors. The channel error simulation module is shown in figure 8

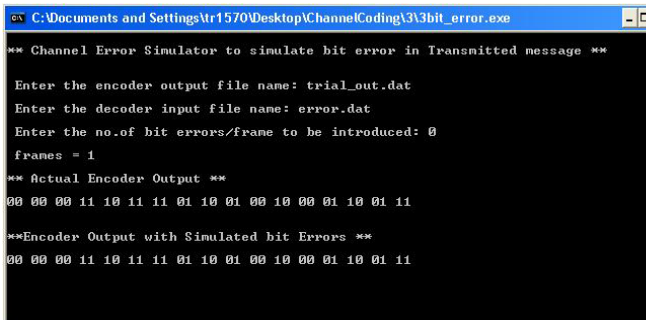


Figure.8 Channel error simulation

In this case, the number of bit errors introduced is zero and hence the encoder output and the received channel symbols are identical. The output of the channel error simulator is written to a file, which will then be taken as the input for the decoder.

Decoding received symbols

The input to the Viterbi decoder is the file containing channel errors, given as output by the channel error simulator. The decoder takes this file as input and then performs decoding and tries to recover the original message. The decoding simulation is shown in figure 9

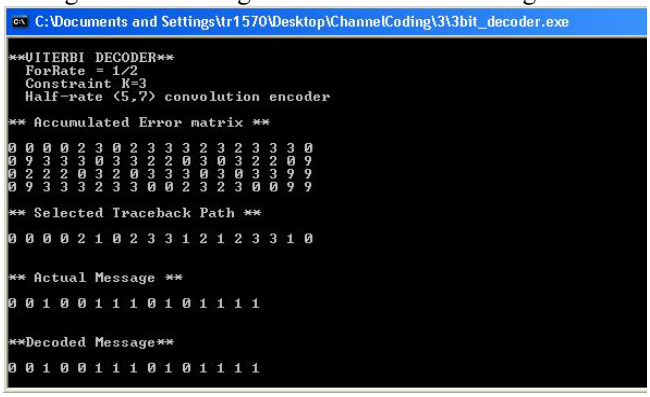


Figure.9 Decoding simulation result

While encoding messages are encoded in frame, so that decoder can start from a known state for each frame. Figure 9 shows decoding of channel symbols containing only one frame.

Performance of Viterbi Decoder

The Viterbi decoder implemented in this project is a Hard-decision decoder capable of decoding the channel symbols encoded by half-rate,(7,5) convolution encoder with a constraint length of 3. Various amounts of bit errors are simulated and the performance of the decoder is then analyzed.

The Test case taken up consists of a 60-bit message encoded into four frames by the encoder. This setup is shown in Figure

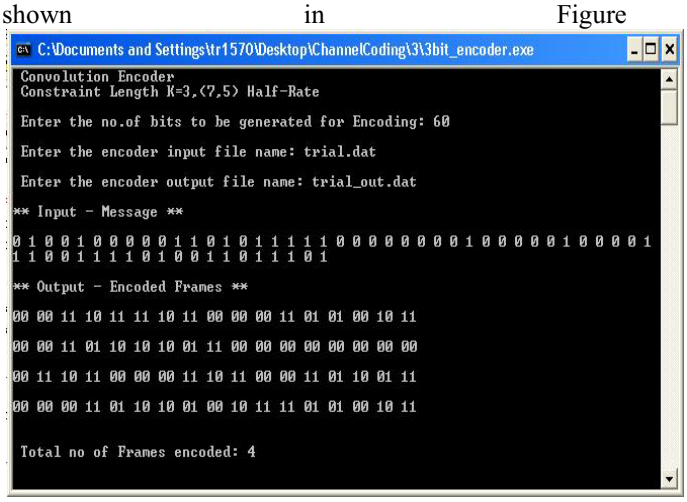
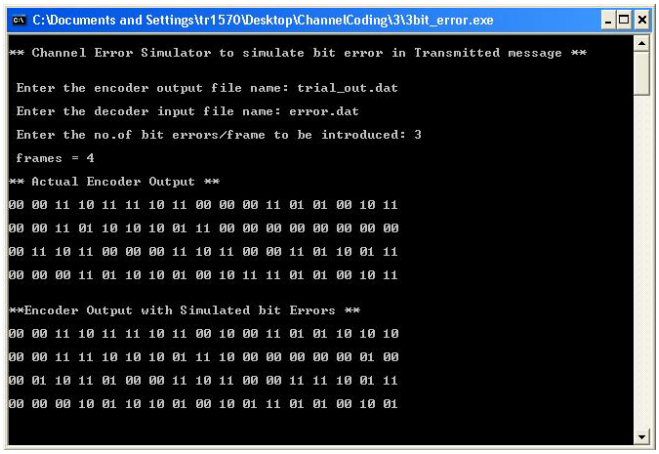
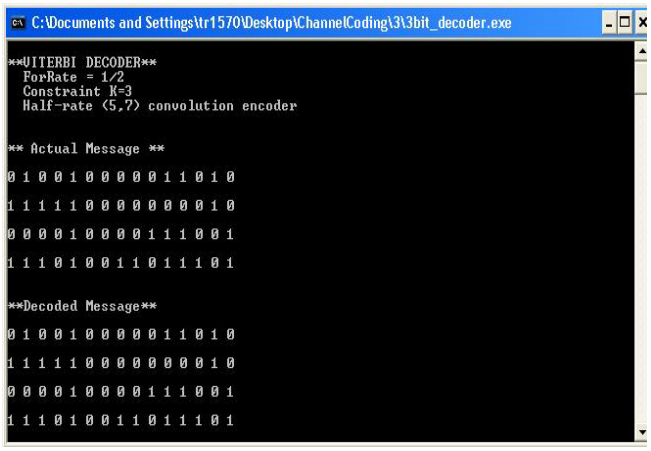


Figure 10. An example depiction

The message is encoded into four frames and is passed to the channel error simulator to introduce various amounts of bit errors and the decoder performance is analyzed.

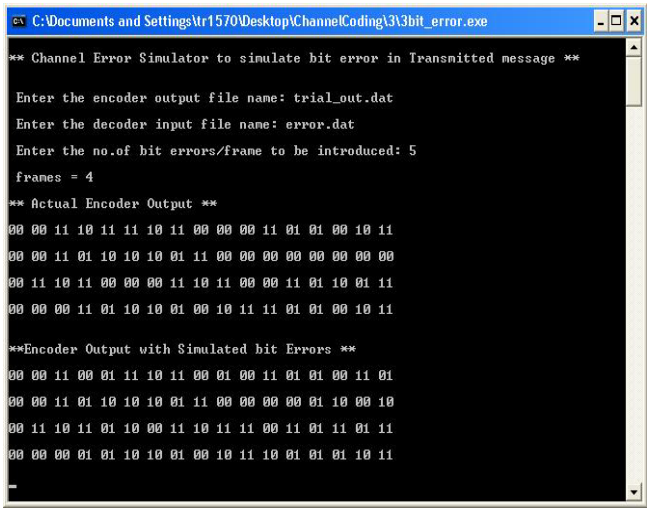


In the first case, three random bit errors are introduced in each frame. The performance of the decoder for this test case is shown if following Screen Shot.

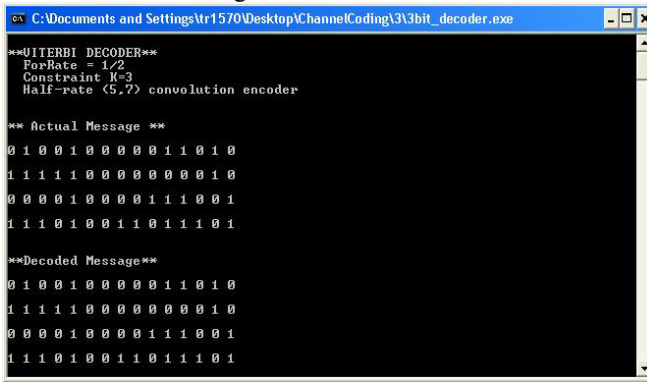


It is found that the decoder recovers the original message when three random bit errors are produced in each frame of the received channel symbols, offering 100% error correction capability for 3 random bit errors.

In the next test case, Five bit errors are introduced in each frame as shown in screen shot as follows.

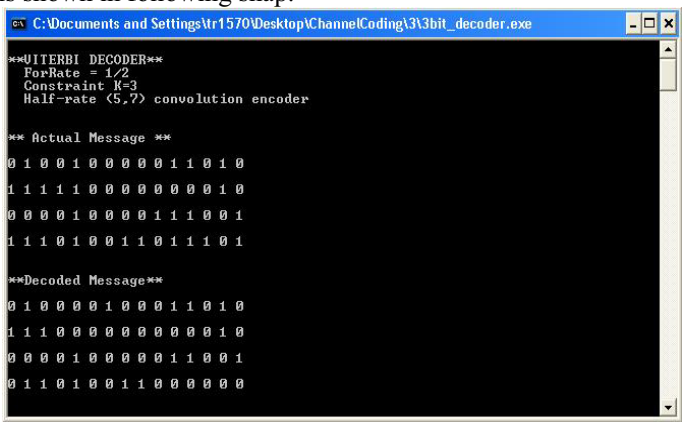


The performance of the decoder for this test case is shown in the following screen shot



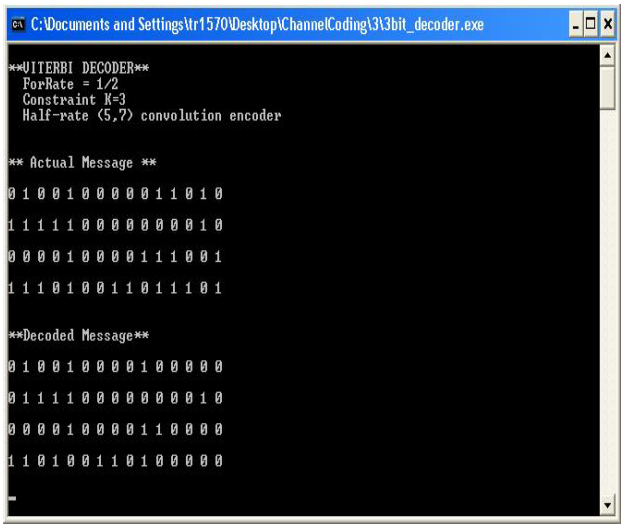
It is found that the Viterbi decoder is able to recover the original message in most of the cases when 5 random bit errors are introduced in each frame, thus offering almost 100% error correction when there is 33% error in the received channel symbols.

The next test case involves introducing seven bit errors per frame and the performance of the decoder in this case is shown in following snap.



It is found that the performance of the decoder deteriorates as the bit error increases to 7 errors per frame. The decoded messages differ from original message and on an average only 32% of the errors are corrected on a trial run with 10 samples.

When the number of bit errors per frame is further increased to 9 bit errors per frame, the performance of the decoder still goes down and the original message in not recovered by the decoder as indicated in next screen shot.



It is found that in case of nine bit errors per frame, the error recovery rate falls to as low as 8.2% on average run with 10 samples.

Thus overall, the performance of the Hard-decision Viterbi decoder falls with the increase in bit error rates and good level of error correction is available up to 5 bit errors per frame above which the message keeps getting deteriorated. This is indicated in the figure .11

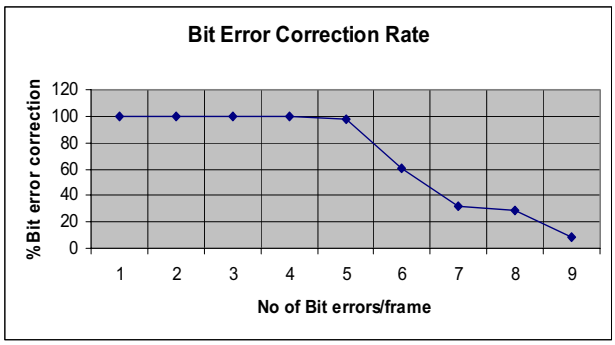


Figure 11. Analysis graph

XI. Conclusion

Thus the software based multi channel FM reception offers the capability to receive from one up to four FM stations at the same time. One station is streamed live to the speakers while the rest are stored in the disk in raw format. This offers the ability to listen to the recorded data at a later time thereby preventing the user from missing out on some important information.

Working with the FM band of the radio spectrum is the entrance to the Software based radio world which is capable of handling many parts of the radio spectrum. This project provided a good practical exposure to the problems involved in replacing a conventional radio with software based one and demonstrated the power and flexibility offered by using software in place of conventional hardware.

This project just touches the tip of the iceberg in the world of software defined radio, which has tremendous potential to be unleashed. With the aid of technology advancement, the software based radios will be able to explore the full radio spectrum very soon into the future.

With the move towards processing signal systems in digital format comes the problem of error detection and correction. Thankfully, this has been taken care by numerous error detection and correction algorithms, which were proposed long back, but were not implemented due to large computation requirements.

The convolution encoder and the Viterbi decoder implemented in this project offer a reasonable amount of error correction capability to a message signal.

These algorithms are the basic error correction algorithms on which many other complex error correction algorithms are built upon.

Acknowlegdement

We wish to thank all the co authors who supported this project for its success.

References

[1] John G Proakis, Dimitris G Manolakis, "***Digital Signal Processing: Principles, Algorithms and Applications***", 1996, Prentice Hall, ISBN: 0-13-373762-4

[2] Todd K. Moon, "***Error correction coding: Mathematical methods and algorithms***",2002, John Wiley and sons, ISBN: 0-471-64800-0

Shriram K. Vasudevan was born at TamilNadu on 29-07-1983. He holds an M.Tech in Embedded Systems and B.E., in Electronics and Instrumentation. He has done his B.E., from Annamalai University, Chidambaram, India. He proceeded with his M.Tech at TIFAC-CORE of SASTRA University, Tanjore, India.

He has got overall 4 years of industrial experience in the field of Embedded Systems and Optical Networking. Currently he is holding the responsibility of Assistant Professor in the field of Embedded Systems in VIT University, Vellore, India. He was employed with Wipro Technologies for 2 years. He has visited USA, Dallas for his telecom training.

He has presented papers in National and International conferences in the areas of VLSI, Embedded and Networking. His research area focuses on Embedded Networking.

Siva Janakiraman was born at TamilNadu on 05-06-1982. He holds an M.Tech in Embedded Systems and B.E., in Electronics and Communication. He has done his B.E., from SASTRA University, Tanjore, India. He proceeded with his M.Tech at SASTRA University, Tanjore, India. He has got overall 5 years of teaching experience in the field of Embedded Systems. Currently he is holding the responsibility of Assistant Professor in the school of Electrical and Electronics Engineering in SASTRA University, Tanjore, India. His research area focuses on Embedded Systems and Cryptography.

Subashri Vasudevan, born at TamilNadu on 21-01-1989. She is currently doing her B.Tech and she is in her final year of her studies. She is graduating at SASTRA University. She has been consistently working on many projects and she is one of the few toppers in her department. She is currently doing project on Software Radio.

She has presented papers in the area of Software Radio and Wireless communications. Her research interest includes Wireless communications and Embedded Systems.

Webinar – Education through Digital Collaboration

Anuradha Verma
Infosys Technologies Ltd, Chandigarh, India
Email: anuradha_thakur@infosys.com

Anoop Singh
Infosys Technologies Ltd, Bangalore, India
Email: anoop_singh@infosys.com

***Abstract*—Transition in the learning habits and trends of students is evident from the traditional text based to the current dynamic modes like world-wide web, CAI and simulation. Webinar as a learning technology offers a platform to overcome the gap in this digital divide and help the students get acquainted with the latest. A case describing the use of webinar by Infosys helps the reader comprehend the digital collaboration webinar offers discussing its shortcomings and benefits.**

***Index Terms*—webinar; learning habits, teaching methodology**

I. Learning Technologies

As per Gretchen Rhines Cheney et al. (2005) [1] India has the second largest education system in the world, falling only behind China. University Education Commission (UGC) [2] set up in 1948 provides guidelines for coordination, determination and maintenance of standards of university education in India. Studying the vast historical background of education in India, it can be seen that though the Indian education system is powerful, yet it needs to dynamically revamp various facets to emerge strong in today's ever changing scenario.

In today's fast pace world students are well informed and require a facilitator rather than an instructor. They have islands of information available but do not have the patience to understand its use or consult their faculty. The learning technologies have also seen a change in the learning habits of students from the past to the recent times. The current day students are well versed with the modern technologies and are avid internet users compared to their predecessors who were more comfortable in a text based instruction based approach.

However in the current education system, there exists a digital divide – a gap in the content and method of knowledge dissipation and what value it offers to its students in today's ever changing world of technology collaboration. Studying the trends in the past, a shift can be ascertained in learning habits of students from the traditional method of lecture based to the interactive and descriptive based on World Wide Web (WWW) or internet, e-mentoring, etc.

This paper talks about the shift in the learning habits and technologies and the impact webinar can play as a learning technology taking a case study from Campus Connect team of Infosys Technologies Ltd.

II. Change In Learning Habits From Traditional To Technology Based

Till late, the teaching methodology used in most of the educational institutes in India was following the traditional method of chalk and talk but in recent times this has seen a paradigm shift.

A. Traditional Methodology

Key features of the traditional teaching methodology can be collated as under:

- One way communication without any interaction
- No openness between the faculty and students as a forum to exchange ideas
- Less focus on analytical skills and more on memory based.
- Only lecture based class room approach followed for knowledge delivery

This approach faced a lot of challenges, some being:

- *Lack of Flexibility* – the students were not offered any flexibility in terms of their learning styles and habits or of implementing their learning. There was a fixed approach being followed.
- *Less or no access to Information* - Improving access of information to the students is difficult as there is was no mode for information sharing.
- *No Standardization* – Quality check at all levels of education was absent offering no standardisation.

B. Scientific Methodology

The advent of multi-media revolutionised training techniques and brought in greater diversity and interaction. The black boards have been replaced by LCD screens. Some of the ingenious methods using scientific or technology base are:

doi:10.4304/jetwi.2.2.131-136

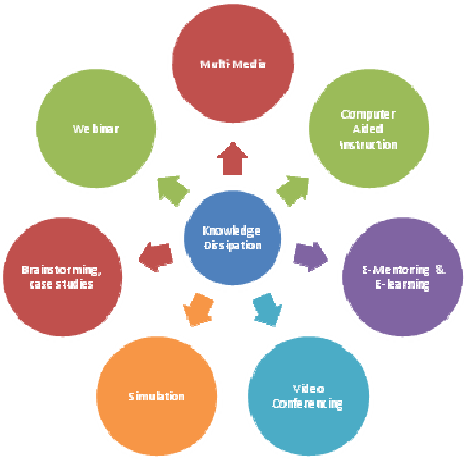


Figure 1. Various Knowledge Dissipation Methods

- *Multi-media Aids* – teaching using medium like power point presentations and slides, audio-visual clipping, etc enhancing the learning experience.
- *Computer aided instruction (CAI)* - use of computers in delivering training, supervising trainee progress, feedback and assessing results. A similar technique is Computer based trainings (CBT).
- *E-mentoring and E-learning* – as these techniques are not dependent on the physical presence of the instructor; they help students to study on their own comfort and speed.
- *Video Conferencing* – It incorporates voice, image and data during long distance transmission offering a real time and highly proficient means.
- *Brainstorming & Case studies* – is a group creativity technique designed to generate a large number of ideas for the solution of a problem.
- *Simulation* – reproduction of the real time environment for introducing students to real time challenges.
- *Webinar* – seminar or lecture over the internet [3].

This study attempts to establish the role of Webinars – an upcoming technology in the field of digital collaboration - in bridging the digital divide.

III. WHAT IS A WEBINAR

'Webinar' is a union of 'web + seminar' which simply means a seminar over the internet. This software is a remarkable innovation in the field of technology which offers a platform for people to interact and collaborate over vast geographical boundaries through WWW. This platform offers a two-way communication leading to higher effectiveness and involvement by the audience.

Typically a webinar consist of a presentation hosted by a service provider on a web server. The link of the webinar is shared with the attendees who can log on to the site and participate in it. The Webinar platform has already carved a niche for itself in the arena of business and has now started being use in education arena as well. After increasing the dynamism in the industry, webinars are all set to bring a revolution in the Education sector.

The focal feature of a webinar is its potential to discuss and share information. Offering a one stop shop for interacting with an array of experts, this platform provides a great appeal to its users.

Some more characteristics of webinars are discussed below:

A. *Characteristics*

- *Sharing Application* – It allows presenters to share their desktops, applications, etc to help the audience get a better understanding of the topic.

- *Chat window* – Attendee students can ask their queries during the session without disturbing the flow of the session through chat window. It helps in interaction with the presenters privately, interaction with the panellists privately, interaction with other participants privately or interaction with all the participants in one go.

- *Session Recording*– It may be beneficial to record the session delivered by the presenter for re-use. This is an out-of-box feature in webinars. The session can be recorded and shared with the students or participants in the form of CDs, etc. This also aids in archival of valuable information.

- *Survey* – The presenter can choose to conduct polls and surveys for the audience.

B. *Infrastructure Requirement*

For Organisers:

- One dedicated personal computer (PC)
- Online WEBINAR monitoring PC station (Preferred)
- Dedicated phone system and line (high quality preferably Polycom)
- Voice-Tap hardware for WEBINAR recording

For Participants:

- One dedicated personal computer (PC)
- Installation of Software of the vendor being used for Webinar (For example WebEx[4], GoToWebinar[5], etc)
- Dedicated direct phone line
- LCD projector
- High quality speaker phone with amplifier (if required for large audience) and mute button

C. *Steps to organise a Webinar*

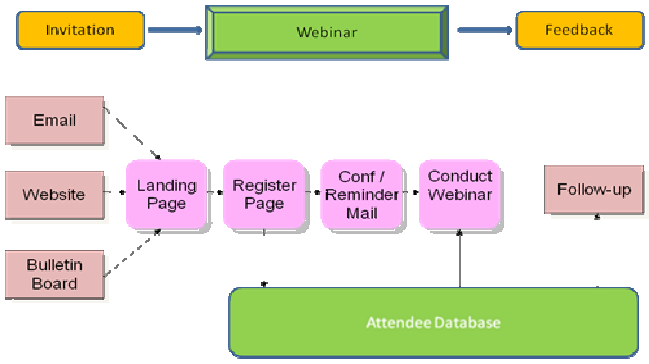


Figure 2. Webinar Event flow

The above diagram describes the event flow for a webinar. The host can send the invitation of the webinar event to intended audience (Student, faculty, etc) through an email, or post it on organization website or bulletin board sharing the link in the invitation mail. On the 'Landing Page' the logged in users can find detailed information about the webinar like objective, speaker profiles, reading materials links, etc. For attending the event, participants need to register on 'Registration Page'.

Once the registration is complete, participants receive an email detailing the next set of instructions for joining the webinar. A reminder mail can also be set for the event. For the audio of the webinar event, participants need to use their telephone and a computer for viewing the presentation. After the event the attendee details can be used for collating feedback and follow-up.

IV. IMPACT OF WEBINAR ON EDUCATION – A CASE STUDY OF INFOSYS CAMPUS CONNECT INITIATVE

Campus Connect Initiative was launched in 2004 By Infosys Technologies Ltd as an industry- academia partnership which with the aim of enhancing the quality and quantity of talent pool in India.

Campus Connect team had a target to reach 500+ partner colleges in an proficient and cost effective way for sharing technical, soft-skills, and domain knowledge. This required reaching the colleges across tier 2, 3 cities and deliver sessions which required approximately 5 hours of travel. Having Infosys Subject Matter Experts (SMEs) from various units spend such a large amount of time on travel wasn't a viable approach.

Hence, the team decided to use a technology platform called Webinar. To augment the planning and implementation of webinars a process was put in place. The team decided to work with WebEx which is a globally acclaimed online meeting applications and software services provider. Once this was zeroed in, the next challenge was to successfully deploy it across all 9 Development Centres (DCs) of Infosys in India. It also required each DC be equipped with the knowledge of Webinar – understanding the concept, its set up and conduction.

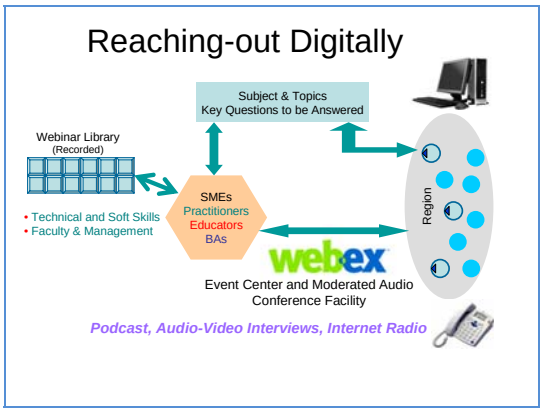


Figure 3. Webinar Platform

To provide clarification for all of this and to define a formal process, the team came up with a Field Guide [6] for webinars which served as a single point of reference for the above.

The following give an idea on the effectiveness and use of webinar by the team. [7]

A. *Stakeholder Feedback*

1. *Faculty feedback*

The webinar efficacy was calculated as part of the Campus Connect Perception Measurement survey in 306 colleges with 521 faculty member responses. The average rating turned out to be **3.53 on a scale of 1 to 4** (Scale: 1. Strongly disagree 2. Disagree 3. Agree 4. Strongly agree).

2. *Student feedback –*

Comprehensive feedback was collected for each webinar from the students. It covered feedback for content, flow, effective delivery, interaction through Q&A sessions, etc. Average feedback scores for major categories are given in the table below. (Scale: 1–Poor, 2-Satisfactory, 3- Good, 4-Very good, 5-Excellent). Overall Webinar Effectiveness was rated as 4.29 out of 5, above 'Very Good'.

TABLE I.
CONSOLIDATED WEBINAR FEEDBACK

Area of Feedback	Average Rating (on a scale of 1-5)
Webinar Content	3.93
Webinar Design	3.90
Presenter Effectiveness	4.05
Planning & Delivery	3.95
Overall Webinar Effectiveness	4.29

B. *Webinar Metrics*

Some consolidated metrics for the webinars conducted so far across the DCs are mentioned below. Table 2 gives the number of webinars conducted till date and their impact.

TABLE II.
WEBINAR COVERAGE DATA PERIOD - FEB 2006 - MAY 2009, 9 DCS

No.	Item	Number
1	Total number of Webinars conducted	59
2	Number of Colleges impacted across India	509
3	Number of Faculty impacted	1147
4	Number of Students impacted	22527

C. Rationalization of Travel Cost and Productivity

Total cost of 1 hour webinar session covering 10 partner colleges is approximately Rs. 12,000/-(Twelve Thousand) and to conduct similar seminar session in 1 partner college by SME through physical travel cost Rs. 39500 (Thirty Nine Thousand). So each webinar session results in saving of Rs. 383000/- (Three Lack Eighty Three Thousand).

TABLE III.
COST BREAKUP TO CONDUCT 1 SEMINAR SESSION

No.	Cost Saving Per Webinar Session	Rs.
1	Cost of Webinar covering 10 colleges	12000
2	Cost of Seminar covering 10 colleges	395000 (39500 * 10 Colleges)
	Total Savings per webinar session	383000

TABLE IV
COST SAVING PER WEBINAR SESSION

No.	Cost Breakup to conduct 1 Seminar	Rs.
1	Flight Tickets / Taxi Charges	12000
2	Accommodation – 2 Days	5000
3	Food /Misc. Expense	2500
4	Productivity loss due to travelling	10000
5	Banner / Marketing Material etc	10000
	Total	39500

Savings in terms of the SME travel time is another significant benefit. Taking a conservative estimate of two business days the SME would have spent in travelling to the college and delivering the seminar, a saving of 1018 business days has resulted because of using webinars.

D. Capability improvement – Scalability and Reach

Before the advent of Webinars in campus connect knowledge dissipation, SMEs were required to travel to the colleges posing a restraint on the number of colleges that can be covered. The reach to the colleges before and after introducing webinars gives a clear picture of the extent of change brought about. Table 5 shows the number of colleges the team was able to reach before and after introducing webinars.

Without Webinars, the coverage percentage would have come down from 22.18% to 18.53%. Because of the wide reach of Webinars, the coverage percentage has increased from 22.18% to 57.22%, an increase of 35.04% in coverage of colleges. This coverage has further increased to 40% covering almost each Campus Connect Partner Colleges.

TABLE V
IMPROVEMENT IN COLLEGE COVERAGE

No.	Period	# Partner Colleges	# Colleges Covered	% Coverage
1	Nov 2005 - Oct 2006 (Without Webinars)	275	61	22.18
2	Nov 2006 - Oct 2007 (Without Webinars)	367	68	18.53
3	Nov 2006 - Oct 2007 (With Webinars)	367	210 (142+68)	57.22
4	Nov 2007 – Till Date (With Webinars)	520	509	97.8

E. Reusability and Reproduction

These webinars covered a wide range of topics from Technical, Soft-Skills, Effective English, Quality and Domain areas. These webinar sessions delivered by SME are recorded and re-use later. These webinar sessions are reproduced and edited in such a manner that they can be used later on by partner colleges. Recording of these sessions are shared with participants in the form of CDs or by uploading the content on WebEx portal. This also aid in archival of Digital library.

V. BENEFITS WEBINAR PLATFORM USAGE IN INFOSYS CAMPUS CONNECT

Some benefits the team achieved were:

- *Collaboration* – It offered the presenters to share information, data, applications, desktops, etc along with allowing participants and presenters to interact amongst themselves publically or privately thus serving as a successful collaborative medium.
- *Multiple speakers* – The team could employ the advantage of multiple speakers for a single session without bearing the travelling cost for each. It has been seen that a panel of speakers creates more impact than an individual, especially for longer presentations.

- *Greater target access* – Webinar has an edge in terms of its greater reach to audience and overcome geographical boundaries. It is easy to access and use through WWW and is available anywhere anytime.
- *No location dependency* – As webinar are location independent, it can involve Subject Matter Experts (SME) from a variety of locations to talk about their respective areas of expertise. This helped the students leverage the knowledge base of these SMEs which may not have been possible otherwise.
- *Interactive platform* – It offered an interactive platform to the session audience (students, faculty or college management) to communicate with the facilitator hence making it an effective means of communication. Having a question answer round with all the participants in a set order like alphabetical helps solve the common queries of the participants and record them.
- *Learning offline* – Recording allows participants to re-run the session and learn the concepts in detail.
- *Engage senior speakers*– as the presenters of the sessions were practitioners and senior Infoscions, physical travel was not possible generally. Webinars helped the team overcome the issues and increase the SME base.
- *Download Material* – Participants could download the material of the webinar and read it offline. This increased the learning window for them. The SMEs could share applications during their sessions as well.
- *Feedback* –Instructor feedback could be recorded through webinars. It helped to analyse and improve further.

VI. Hurdles For The Team In Webinars

- *Need for high end Infrastructure*- requires quality of service and has dependency of electricity connection especially for the VOIP (Voice over internet protocol). Call drop-off rates during webinar session is typically 10 percent so live helpdesk support through phone and online chat is required in case of any technical issue.
- *Cost of hosting* – As a medium of instruction, webinar proves to be costly as compared to the others. Institutions or students may find it difficult to purchase licenses for webinar software and sustain the same.
- *Restricted Audience Involvement*–The two-way communication channel has a limited usage here. Hence, this may prove to a hurdle at times. It may not be able to live up to the same level as a physical interaction with the expert.
- *Limited modes of multi-media usage* –Power Point Presentation Animation along with video cannot be played smoothly if a participant doesn't have the required bandwidth. Webinar doesn't allow the instructor to use audio and video mode simultaneously.

VII. Conclusion

The traditional methods of training in educational system may have been in practice since a long time, the current time throws a challenge for students to update themselves on the technologies and their usage. With the advent of IT era, the learning habits of students have undergone a transformation from referring the text books to browsing the internet, having 3D images and animations as aids.

Webinar today has gained significant popularity. More and more institutions want to leverage technology in the field of education and percolate its maximum advantage to their students. It has helped overcome the issues of bandwidth and leverage expertise but has still to develop its potential as an interactive forum and cost effectiveness.

Going ahead, this technology can be deployed to harness a formal mechanism of measurement and an evaluation framework. Its usage can be extended to other areas of learning and its target audience base increased.

While the current study talks about the webinars conducted through WebEx, some common vendors that offer this platform can be listed as below:

- InterCall[8]
- GoToMeeting[9]
- IBM Lotus Sametime[10]
- MeetingBridge[11]
- ReadyTalk[12]
- WebMeetLive[13]

Acknowledgment

We would like to thank Dr Ramesh Babu, Mr. R.N. Prasad, Mr. Mukundh Nagarajan and Mr. Sarfaraz Abdul Aziz Jaitapkar from Infosys Technologies Ltd. for sharing their inputs and experiences for this paper.

References

[1] Gretchen Rhines Cheney, Betsy Brown Ruzzi and Karthic Muralidharan, "A Profile of Indian Education System, National Center on Education and the Economy," *November 2005*

[2] Home Page, http://www.ugc.ac.in/

[3] Webinar Definition, http://www.pcmag.com/encyclopedia_term/0,2542,t=Webinar&i=54380,00.asp

[4] Home Page, http://www.webex.com/

[5] Home Page, http://www.gotowebinar.com/

[6] Webinar Field Guide V1.0, Infosys Technologies Ltd., 2006

[7] Anuradha Verma, Anoop Singh, "Leveraging Webinar for Student Learning", *International Workshop on Technology for Education, 2009*

[8] http://www.intercall.com/

[9] https://www1.gotomeeting.com/?Portal=www.gotomeeting.com
[10] http://www.ibm.com/lotus/sametime
[11] http://www.meetingbridge.com/
[12] http://www.readytalk.com/
[13] http://www.webmeetlive.com/

Anuradha Verma is currently working as Technical Evangelist in the Education & Research Department at Infosys Technologies Ltd. She has over 6 years of IT Industry experience covering areas like Training & Development, Testing, Software Development and Industry-Academia Partnership.

She has Masters in Business Administration along with a Masters in Journalism and Mass Communication. She is currently pursuing her Doctor of Philosophy in Management with Human Resource Specialization. She has authored publications in International Conferences and Journal.

Anoop Singh. At Infosys, Anoop Singh is currently working as GROUP LEADER in the E&R Department. He has over 12 years of IT Industry experience including 4 years of exposure in global market, working in various counties of South East Asia region like Singapore, Hong Kong and China. Anoop has extensive IT Industry experience covering areas like Training & Development, Business Partner Development, Product Line management and Industry-Academia Partnership.

Anoop has earlier worked with IT organizations like NIIT, NIIT Shanghai Ltd. and have server large global customer like Microsoft, Singapore Airline, Beijing International Airport and various Chinese universities. He has worked on technologies areas like Database, Network Design and various programming languages. Anoop also holds a Master's Degree in Information Science.

Efficient Visual Cryptography

Er. Supriya Kinger
CSE Department, Chitkara Institute of Engineering and Technology,
Rajpura, Punjab, India
Email: ahujasupriya@gmail.com

Abstract **– Visual cryptography scheme (VCS) is a secret-sharing scheme which allows the encryption of a secret image into n shares that are distributed to n participants. The beauty of such a scheme is that, the decryption of the secret image requires neither the knowledge of cryptography nor complex computation. Colour visual cryptography becomes an interesting research topic after the formal introduction of visual cryptography by Naor and Shamir in 1995. It is a powerful technique which combines the notions of perfect ciphers and secret sharing in cryptography with that of raster graphics. A binary image can be divided into shares which can be stacked together to approximately recover the original image. Unfortunately, it has not been used much primarily because the decryption process entails a severe degradation in image quality in terms of loss of resolution and contrast. Its usage is also hampered by the lack of proper techniques for handling grayscale and color images. In this paper, I have developed a novel technique which enables visual cryptography of color as well as grayscale images. The physical transparency stacking type of decryption allows for the recovery of the traditional visual cryptography quality image. An enhanced stacking technique allows for the decryption into a halftone quality image. And finally, a computation based decryption scheme makes the perfect recovery of the original image possible. Based on this basic scheme, I have then established a progressive mechanism to share color images at multiple resolutions. I extracted shares from each resolution layer to construct a hierarchical structure; the images of different resolutions can then be restored by stacking the different shared images together. I have implemented our technique and present results.**

Index Terms **– Secret sharing, Color halftoning, image sharing, multiple resolutions, secret sharing, and visual cryptography**

I. Introduction

Visual cryptography was originally proposed for the problem of secret sharing. Secret sharing is one of the early problems to be considered in cryptography. The idea of the visual cryptography model proposed in [1] is to split an image into two random shares (printed on transparencies) which separately reveal no information about the original secret image other than the size of the secret image. The image is composed of black and white pixels. The original image can be recovered by superimposing the two shares. The underlying operation of this visual cryptography model is OR.

Visual cryptography is a unique technique in the sense that the encrypted messages can be decrypted directly by the human visual system. Therefore, a system employing visual cryptography can be used by anyone without any knowledge of cryptography. Another interesting thing about visual cryptography is that it is a perfectly secure cipher. There is a simple analogy of the one time-pad cipher to visual cryptography.

II. Background On Visual Cryptography

Besides introducing the new paradigm, Naor and Shamir also provided their constructions of visual cryptographic solutions for the general *k* out of *n* secret sharing problem. One can assume that every secret message can be represented as an image, and furthermore that the image is just a collection of black and white pixels i.e. it is assumed to be a binary image. Each original pixel appears in *n* modified versions (called shares) of the image, one for each transparency. Each share consists of *m* black and white sub-pixels. Each share of sub-pixels is printed on the transparency in close proximity (to best aid the human perception, they are typically arranged together to form a square with *m* selected as a square number). The resulting structure can be described by a Boolean matrix $M = (m_{ij})_{n*m}$ where $m_{ij} = 1$ if and only if the *j*-th sub-pixel of the *i*-th share (transparency) is black. The important parameters of the scheme are:

1. *m*, the number of pixels in a share. This parameter represents the loss in resolution from the original image to the recovered one.
2. α, the relative difference in the weight between the combined shares that come from a white pixel and a black pixel in the original image. This parameter represents the loss in contrast.
3. γ, the size of the collection of C_0 and C_1. C_0 refers to the sub-pixel patterns in the shares for a white pixel and black refers to the sub-pixel patterns in the shares for the 1 pixel.

The constructions can be clearly illustrated by a 2 out of 2 visual cryptographic scheme. Here we define the following collections of 2*4 matrices:

C_0 = all the matrices obtained by permuting the columns of

$$\begin{vmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{vmatrix}$$

C_1 = all the matrices obtained by permuting the columns of

$$\begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix}$$

doi:10.4304/jetwi.2.2.137-141

The six patterns of shares created based on the above matrices are shown in figure 1. Note that *one* pixel of the original image now corresponds to *four* pixels in each share.

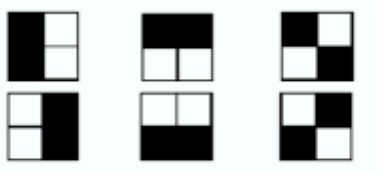


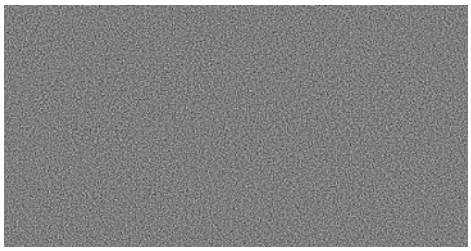
Fig 1: The six patterns of 4 pixel shares: vertical, horizontal and diagonal

A visual cryptography scheme can then be constructed by picking shares in the following manner:

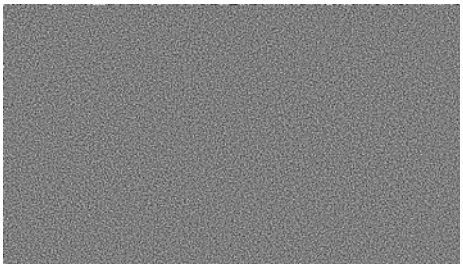
a) If the pixel of the original binary image is white, randomly pick the same pattern 0 of *f our* pixels for both shares. It is important to pick the patterns randomly in order to make the pattern random.
b) If the pixel of the original image is black, pick a complementary pair of patterns, i.e., the patterns from the same column in figure 1.

It can be easily verified that the resultant scheme has the parameters [m = 4; α= 12 ; γ= 6]: any two shares of C_0 cover *two* out *f our* of the pixels, while any pair of shares from C_1 covers all the *f our* pixels. An example of the above scheme is shown in figure 2. The first image is the original image, the next two are the shares and the last image is the recovered original image obtained by performing the equivalent of physically stacking two image shares on top of each other (assuming that they are printed on transparencies). It should be noted that the last three images in figure 2 are four times as large as the first one but I have scaled them to the same size as the original image.

(a) Sample of monochrome image



(b) The first Share



(c) The second share

(d) The stacked Image

Fig 2. Implementation of existing methodology

III. Related Work

There has been a steadily growing interest in visual cryptography. Despite its appearance of being a simple technique, visual cryptography is a secure and effective cryptographic scheme. Since the origin of this new paradigm, various extensions to the basic scheme have been developed to improve the contrast and the areas of application have also been greatly expanded.

In [1], the construction of (n,n)-VCS was extended for (k,n)-VCS. In 1996, the same authors introduced the idea of cover based semi-group to further improve the contrast [3]. Ateniese et al. [4] provided the first construction of (2, n)-VCS having the best possible contrast for any n·2. Blundo et al. [5] provided a contrast optimal (3,n)-VCS and gave a proof on the upper bound on the contrast of any (3,n)-VCS. [1] first considered the problem of concealing the existence of the secret image. [6] provided a general solution for that problem.

The random nature of secret shares makes shares unsuitable for transmission over an open channel. [6] used a modified scheme to embed some meaningful images into the shares. [7] used different moiré patterns to visualize the secret instead of different gray levels. As far as extending to color images goes, [8] provided a primitive scheme for images of 24 colors. Hou [9] then proposed a novel approach to share color images based on halftoning. Other interesting topics include visual authentication [10] and watermarking based on visual cryptography [11]. Recently, there has been an attempt to build a physical visual cryptographic system based on optical interferometry [12]. However, all of these earlier works result in a decrypted image of reduced quality.

IV. Our Contribution

The state of the art in visual cryptography leads to the degradation in the quality of the decoded images, which makes it unsuitable for digital media (image, video) sharing and protection. This is quite obvious in figure 2 where the white background of the original image becomes gray in the decrypted image.

Through this paper, I propose a visual cryptographic schemes that not only can support grayscale and color images, but also allow high quality images including that of perfect (original) quality to be reconstructed.

The nagging presence of the loss of contrast makes traditional visual cryptography scheme practical only when quality is not an issue which is quite rare. I have therefore focused our attention on specifically overcoming this problem by primarily devoting our efforts towards improving the quality of the reconstructed images. I first extend the basic scheme from [1] to allow visual cryptography to be directly applied on grayscale and color images. Image halftoning is employed in order to transform the original image from the grayscale/color space into the monochrome space which has proved to be quite effective.

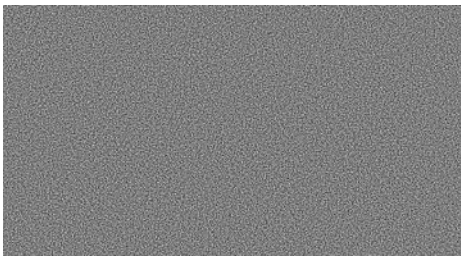
It is a well known fact that the digital halftoning is always a lossy process [2], which means that whenever a halftoning is used for the transformation, it is impossible to fully reconstruct the original secret image. A new encoding scheme has therefore been developed which allows for perfectly lossless transformation between monochrome, grayscale and color spaces. This new encoding scheme can be seamlessly incorporated into the proposed scheme for visual cryptography and it allows the original secret image to be perfectly restored. I believe this advancement in visual cryptography can be useful in secret sharing of images, in transmission of secret images over multiple untrustworthy channels, in e-commerce of digital media and in digital rights management of digital media.

V. Our Approach

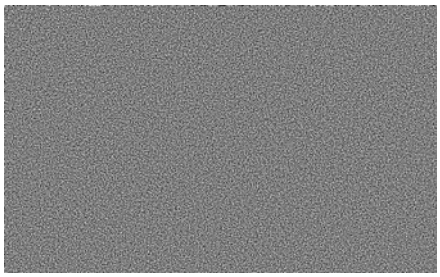
A. For Monochrome Images

Using XOR to Fully Restore Monochrome Secret Images :-I, first made the crucial observation that with just one additional computational operation; even traditional visual cryptography can allow full recovery of the secret binary image. Normally, when we superimpose the two shares printed on transparencies, this stacking operation is computationally modeled as the binary OR operation which causes the contrast level to be lowered. By simply substituting this OR operation with the XOR operation, the original binary image can be recovered without any loss in contrast. Thus, the produced image could have a more visually pleasant appearance with less storage space requirement. However, the XOR operation needs computation - the physical stacking process can only simulate the OR operation. Figure 3 recovers the same secret image as in figure 2 using the XOR operation and thus it is clearly evident that the contrast of the original image is restored.

(a) Sample of monochrome image



(b) The first Share



(c) The second share

(d) The stacked Image with XOR

Fig 3. Implementation of proposed methodology

As we have seen earlier, the application of digital halftoning techniques results in some downgrading of the original image quality due to its inherently lossy nature and it is not possible to recover the original image from its halftone version. We will refer to this proposed scheme as EVCS (Efficient Visual Cryptographic Scheme).

The novelty of my approach is that it not only allows the secret image to be just seen but allows the secret image to be reconstructed with perfect quality. The advantage of this approach is that it still retains the crucial advantages of traditional visual cryptography like simplicity, visual decoding and perfect security. The extra feature is that depending on whether additional computing resources are provided, images of different

quality can be decoded from the same set of shares. If only the stacking operation is allowed (i.e. no computations), then our scheme recovers the original visual cryptographic quality. If the XOR operation is provided (instead of the OR operation of stacking), then we can fully restore the original quality image.

B. For Colored Images (Halftone-based Grayscale and Color Visual Cryptography)

Digital halftoning has been extensively used in printing applications where it has been proved to be very effective. For visual cryptography, the use of digital half toning is for the purpose of converting the grayscale image into a monochrome image. Once we have a binary image, then the original visual cryptography technique can be applied. However, the concomitant loss in quality is unavoidable in this case.

For color images, there are two alternatives for applying digital halftoning. One is to split the color image into channels of cyan, magenta and yellow. Then each channel is treated as a grayscale image to which halftoning and visual cryptography are applied independently. After the monochrome shares are generated for each channel, channels are combined separately to create the color shares. This is the approach presented in [9]. The alternative approach would be to directly apply color halftoning, then perform the separation into color channels followed by the application of visual cryptography to each channel independently. Actually, these two approaches lead to the same results finally. There are many mature halftoning techniques available for selection like dispersed-dot dithering, clustered-dot dithering and error diffusion techniques.

Halftoning based visual cryptographic scheme can be summarized as follows:

a) **Encryption***:* This stage is for the creation of shares. This can be further divided into the following steps:

i. *Color halftoning:* Standard algorithms such as the ones described in [2], [13] and [14] can be used for this step. One could do the color channel splitting first and then do the grayscale halftoning for each channel:

$$I \xrightarrow{split\ \ CMY} [I^C, I^M, I^Y] \xrightarrow{halftoning} [I^C_{hft}, I^M_{hft}, I^Y_{hft}]$$

Or one could do color halftoning first followed by the splitting:

$$I \xrightarrow{color\ \ halftoning} I_{hft} \xrightarrow{split\ \ CMY} [I^C_{hft}, I^M_{hft}, I^Y_{hft}]$$

ii. *Creation of shares:* Considering the case of (2,2)-VCS, the steps are

$$I^C_{hft} \xrightarrow{(2,2)-VCS} [S^C_0, S^C_1]$$
$$I^M_{hft} \xrightarrow{(2,2)-VCS} [S^M_0, S^M_1]$$
$$I^Y_{hft} \xrightarrow{(2,2)-VCS} [S^Y_0, S^Y_1]$$

b) **Decryption:** This stage is for the reconstruction of the original secret image. This can be further divided into the following steps:

i. *Stacking of shares:* The following stacking (OR) operation needs to be performed:

$$[S^C_0, S^C_1] \xrightarrow{stacking} I^{mg}_C$$
$$[S^M_0, S^M_1] \xrightarrow{stacking} I^{mg}_M$$
$$[S^Y_0, S^Y_1] \xrightarrow{stacking} I^{mg}_Y$$

ii. *Subsampling for reconstruction:* These operations need to be performed where every block of *f our* pixels is sub-sampled into *one* pixel of the final image. This step is optional and should be used only with the XOR recovery described in Section III-B.1 to achieve better quality.

$$[I^{mg}_C, I^{mg}_M, I^{mg}_Y] \xrightarrow{combine\ CMY} I^{mg}$$

Then, for every 2*2 block $B(i; j)$ of I, where

$$B(i,j) = \begin{bmatrix} I^{mg}(2i,2j) & I^{mg}(2i,2j+1) \\ I^{mg}(2i+1,2j) & I^{mg}(2i+1,2j+1) \end{bmatrix}$$

$$I^{subsampled}(i,j) = I^{mg}(2i,2j)$$

It is clear that our technique, though independently developed, is quite similar in spirit to the one described in [9]. So both share the same drawback that digital halftoning always leads to permanent loss of information which means that the original image can never be perfectly restored. Inverse halftoning is a possible solution that can attempt to recover the image. The best results can obtain a restoration quality of 30 dB measured in PSNR, which is quite good. But this is not sufficient for applications which require that the original image be faithfully recovered. In fact, in all other cryptographic techniques, it is taken for granted that the decryption of a ciphertext perfectly recovers the plaintext. But visual cryptography has been a glaring exception so far.

VI. CONCLUSION

In this paper, I have extended traditional visual cryptography by employing new schemes which overcome its limitations. I propose a technique for grayscale and color visual cryptography. Our insight is that the OR operation in the traditional visual cryptography can be replaced by the XOR operation in order to allow for lossless decryption. However, there are some practical issues that need careful consideration. First, the transparencies should be precisely aligned in order to obtain a clear reconstruction. Secondly, there is usually some unavoidable noise introduced during the printing process. Thirdly, the stacking method can only simulate the OR operation which always leads to a loss in contrast. Proper alignment is absolutely essential when superimposing the shares. In real experiments, we have found that obtaining perfect alignment is always troublesome. As visual cryptographic schemes operate at the pixel levels, each pixel on one share must be matched correctly with the corresponding pixel on the other share. Superimposing the shares with even a slight shift in alignment results in a drastic degradation in the quality of the reconstructed image. In the worst case, even a single pixel shift can render the secret image totally invisible. This alignment problem can be resolved if the boundary of each share is clearly marked which can act as guides for the alignment.

REFERENCES

[1] M. Naor and A. Shamir, "Visual cryptography," in *Advances in Cryptology -EUROCRYPT'94*, A. D. Santis., Ed., vol. 950. Springer-Verlag, 1995, pp. 1–12.

[2] H. R. Kang, *Digital Color Halftoning*, ser. SPIE/IEE Series on Imaging Science and Engineering, E. R. Dougherty, Ed.Bellingham, Washington USA and New York: Copublished by SPIE Optical Engineering Press and IEEE Press, 1999.

[3] M. Naor and A. Shamir, "Visual cryptography 2: Improving the contrast via the cover base," 1996, a preliminary versionappears in "Security Protocols", M. Lomas ed. Vol. 1189 of Lecture Notes in Compute Science, Springer-Verlag, Berlin,pp.197-202, 1997.

[4] A. D. S. G. Ateniese, C. Blundo and D. R. Stinson, "Constructions and bounds for visual cryptography," in *23rd International Colloquium on Automata, Languages and Programming*, ser. Lecture Notes in Computer Science, F. M. auf der Heide and B. Monien, Eds., vol. 1099. Berlin: Springer-Verlag, 1996, pp. 416–428.

[5] C. Blundo, P. D'Arco, A. D. Snatis, and D. R. Stinson, "Contrast optimal threshold visual cryptography schemes," *SIAM Journal on Discrete Mathematics, available at: http://citeseer.nj.nec.com/blundo98contrast.html*, vol. 16, no. 2, pp. 224–261, April 1998.

[6] G. Ateniese, C. Blundo, A. D. Santis, and D. Stinson, "Extended schemes for visual cryptography," *Theoretical ComputerScience*, vol. 250, pp. 143–161, 2001.

[7] Y. Desmedt and T. V. Le, "Moire cryptography," in *the 7th ACM Conference on Computer and Communications Security'00*, Athens, Greece, 2000.

[8] V. Rijmen and B. Preneel, "Efficient color visual encryption for shared colors of benetton," 1996, EUCRYPTO'96 RumpSession. Availabe at http://www.iacr/org/conferences/ec96/rump/preneel.ps.

[9] Y. C. Hou, C. Y. Chang, and S. F. Tu, "Visual cryptography for color images based on halftone technology," in*International Conference on Information Systems, Analysis and Synthesis. World Multiconference on Systemics, Cyberneticsand Informatics. Image, Acoustic, Speech And Signal Processing: Part II*, 2001.

[10] M. Naor and B. Pinkas, "Visual authentication and identification," *Lecture Notes in Computer Science*, vol. 1294, pp. 322–336, 1997. [Online]. Available: citeseer.nj.nec.com/67294.html

[11] Q. B. Sun, P. R. Feng, and R. Deng, in *International Conference on Information Technology: Coding and Computing (ITCC '01), available at: http://dlib.computer.org/conferen/itcc/1062/pdf/10620065.pdf*, Las Vegas, April 2001.

[12] S.-S. Lee, J.-C. Na, S.-W. Sohn, C. Park, D.-H. Seo, and S.-J. Kim, "Visual cryptography based on an interferometric encryption technique," *ETRI Journal*, vol. 24, pp. 373–380, 2002, available at http://etrij.etri.re.kr/etrij/pdfdata/24-05-05.pdf.

Supriya A. Kinger was born in Haryana, India on 15th july 1982. She did her Master's in technology in field of computer science from YMCA, Haryana in year 2005, India and Bachelor's in technology in Computer Science from KU, Haryana, India. In year 2003.Currently she is doing research in field of Software Engineering (Component Based Software Engineering).

She has more than 5 Years of experience in teaching and research. She has attended and organized a number of workshops and conferences. She hosted a conference ASET-2006 at CIET and acted as convener in it. She has presented number of papers in national and international conferences of repute on the topics of web crawlers, Component based software Engineering, Information Security and Networks. Currently she is Working with Chitkara Institute of Enginnering and Technology, Punjab, India as Senior Lecturer. Earlier she has worked at Institute of Engineering and Technology. She is life member of ISTE

Multistage Interconnection Networks: A Transition from Electronic to Optical

Rinkle Rani Aggarwal
Department of Computer Science & Engineering,
Thapar University, Patiala –147004 (India)
raggarwal@thapar.edu

Dr. Lakhwinder Kaur, Dr. Himanshu Aggarwal
Department of Computer Engineering,
Punjabi University, Patiala–147002 (India)
mahal2k8@yahoo.com, himanshu@pbi.ac.in

***Abstract*—Optical communication are necessary for achieving reliable, fast and flexible communication. Advances in electro-optic technologies have made optical communication a reliable networking choice to meet the increasing demands for high bandwidth and low communication latency of high-performance computing/communication applications. So optical networks gives high performance as well as low latency. Although optical MINs hold great promise and have advantages over their electronic networks, they also hold their own challenges. This paper compares electronic and Optical MINs. The design issues and solution approaches available for optical MINs are also explained and analyzed.**

***Index Terms*— Multistage Interconnection Networks (MIN), Optical networks, Crosstalk, Window methods.**

I. INTRODUCTION

To meet the increasing demands of high performance computing applications for high channel bandwidth and low communication latency, traditional metal-based communication technology used in parallel computing systems is becoming a potential bottleneck. Now, the need arise either for some significant progress in the traditional interconnects or for some new interconnect technology be introduced in parallel computing systems. Electro-optic technologies have made optical communication a promising network choice to meet the increasing demands with its advancement in the technology. Fiber optic communications offer a combination of high bandwidth, low error probability and gigabit transmission capacity.

Multistage interconnection networks have been extensively accepted as an interconnecting scheme for parallel computing systems. As optical technology advances, there is considerable interest in using optical technology to implement interconnection network and switches. A multistage interconnection network is composed of several stages of switch elements by which any input port can be connected to any output port in the network. Optical MIN represents a very important class of interconnecting schemes used for constructing Optical interconnections for communication networks and multiprocessor systems. This network consists of N inputs, N outputs and n stages ($n=\log_2 N$). Each stage has N/2 switching elements each SE has two inputs and two outputs connected in a certain pattern. The most widely used MINs are the electronic MINs. In electronic MINs, electricity is used, since in optical MINs, light is used to transmit the messages. Although electronic MINs and optical MINs have many similarities but there are some fundamental differences between them. Available optical MINs were built mainly on banyan or its equivalent (e.g. *baseline*, *omega*) networks because they are fast in switch setting (self-routing) and also have a small number of switches between an input-output pair. Banyan networks have a unique path between an input-output pair, and this makes them blocking networks. Non-blocking networks can be constructed by either appending some extra stages to the back of a regular banyan network. Crosstalk in optical networks is one of the major shortcomings in optical switching networks, and avoiding crosstalk is an important for making optical communication properly. To avoid a crosstalk, many approaches have been used such as time domain and space domain approaches. Because the messages should be partitioned into several groups to send to the network, some methods are used to find conflicts between the messages.

II. MULTISTAGE INTERCONNECTION NETWORKS

Multistage interconnection networks (MINs) consist of more than one stage of small interconnection elements called switching elements and links interconnecting them. Multistage interconnection networks (MINs) are used in multiprocessing systems to provide cost-effective, high-bandwidth communication between processors and/or memory modules. A MIN normally connects N inputs to N outputs and is referred as an N × N MIN [9,10]. The parameter N is called the size of the network. There are several different multistage interconnection networks

doi:10.4304/jetwi.2.2.142-147

proposed and studied in the literature. Figure 1 illustrates a structure of multistage interconnection network, which are representatives of a general class of networks. This figure shows the connection between p inputs and b outputs, and connection between these is via number of stages. Multistage interconnection network is actually a compromise between crossbar and shared bus networks of various types of multiprocessor interconnections networks [1]. Multistage interconnection networks attempt to reduce cost and decrease the path length

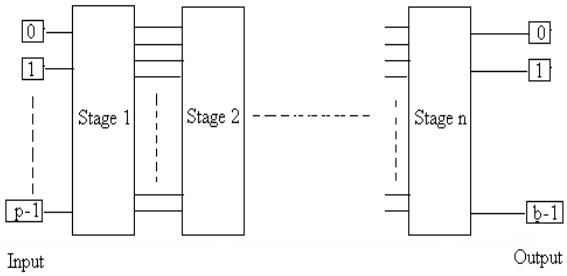


Figure 1: A MultistageNetwork

TABLE I
PROPERTIES OF DIFFERENT INTERCONNECTION TECHNIQUES

Property	***Bus***	***Crossbar***	***Multistage***
Speed	Low	High	High
Cost	Low	High	Moderate
Reliability	Low	High	High
Configurability	High	Low	Moderate
Complexity	Low	High	Moderate

III. OPTICAL MULTISTAGE INTERCONNECTION NETWORKS

An optical MIN can be implemented with either free-space optics or guided wave technology. It uses the Time Division Multiplexing. To exploit the huge optical bandwidth of fiber, the Wavelength Division Multiplexing (WDM) technique can also be used. With WDM, the optical spectrum is divided into many different logical channels, and each channel corresponds to a unique wavelength. Optical switching, involves the switching of optical signals, rather than electronic signals as in conventional electronic systems. Two types of guided wave optical switching systems can be used [5]. The first is a hybrid approach in which optical signals are switched, but the switches are electronically controlled. With this approach, the use of electronic control signals means that the routing will be carried out electronically. As such, the speed of the electronic switch control signals can be much less than the bit rate of the optical signals being switched. So, with this approach there is a big speed mismatch occur due to the high speed of optical signals. The second approach is all-optical switching. This has removed the problem that occurred with the hybrid approach but, such systems will not become practical in the future and hence only hybrid optical MINs are considered. In hybrid optical MINs, the electronically controlled optical switches, such as lithium niobate directional couplers, can have switching speeds from hundreds of picoseconds to tens of nanoseconds.

A. Switching in optical networks

In optical networks, circuit switching is used. Packet switching is not possible with Optical Multistage Interconnection Networks. If packet switching is used, the address information in each packet must be decoded in order to determine the switch state. In a hybrid MIN, it means it require conversions from optical signals to electronic ones, which could be very costly [4]. For this reason, circuit switching is usually preferred in optical MINs.

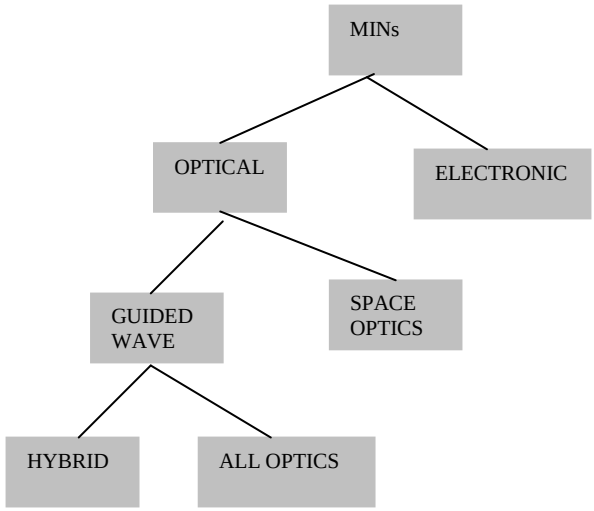


Figure 2: Types of Multistage Networks

B. Comparison of Electronic and Optical Networks

There are lots of benefits of optical networks over the electronic ones. The main benefit of the optical networks over the electronic network is the high speed of the Optical signals. In the optical networks light is transmitted which has a very good speed but in the electronic Multistage interconnection networks electricity is used which has very slow speed. The second advantage is the Bandwidth. Now a day's applications in communication require high bandwidth, the optical networks gives combination of very high bandwidth and low latency. Therefore, they have been used in the parallel processing applications. Optical MINs are also used in wide area networks, which require less error probability and very high bandwidth. Fiber optic

transmission distance is significantly greater than the electronic ones, signal need not to be regenerated in optical networks. Optical fiber has very less weight in comparison to electronic MINs. Thus Optical networks give the combination of high bandwidth and low latency.

TABLE II
COMARISON OF ELECTRONIC AND OPTICAL NETWORKS

Characteristics	***Electronic Multistage Netwotks***	***Optical Multistage Networks***
Speed	Less	High
Energy Transmitted	Electricity	Light
Bandwidth	Used for less bandwidth applications	Used for high bandwidth applications
Latency	High	Less
Error Probability	High	Less
Weight	More	Less
Cost	Less	More
Switching	Packet Switching	Circuit Switching
Path	Provide Multi path from source to destination.	Provide single path from source to destination
Complexity	More Complex	Less Complex
Structure considered	2-dimensional	3-dimensional

IV. PROBLEMS IN OPTICAL NETWORKS

Due to the difference in speeds of the electronic and optical switching elements and the nature of optical signals, optical MINs also hold their own challenges.

A. Path Dependent Loss

Path dependent loss means that optical signals become weak after passing through an optical path. In a large MIN, a big part of the path-dependent loss is directly proportional to the number of couplers that the optical path passes through [16]. Hence, it depends on the architecture used and its network size. Hence, if the optical signal has to pass through more no of stages or switches the path dependent loss will be more.

B. Optical Crosstalk

Optical crosstalk occurs when two signal channels interact with each other. There are two ways in which optical paths can interact in a switching network. The channels carrying the signals could cross each other. Alternatively; two paths sharing a switch could experience some undesired coupling from one path to another within a switch. For example, assume that the two inputs are y and z, respectively, the two outputs will have ly+lxz and lz+lxy, respectively, where l is path loss and x is signal crosstalk in a switch. Using the best device x=35 dB and l=0.25 dB. For more practically available devices, it is more likely that x=20 dB and l=1 dB [5]. Hence, when a signal passes many switches, the input signal will be distorted at the output due to the loss and crosstalk introduced on the path.

Crosstalk problem is more dangerous than the path-dependent loss problem with current optical technology. Thus, switch crosstalk is the most significant factor that reduces the signal-to-noise ratio and limits the size of a network. Luckily, ensuring that a switch is not used by two input signals simultaneously can eliminate first-order crosstalk. Once the major source of crosstalk disappears, crosstalk in an optical MIN will have a very small effect on the signal-to-noise ratio and thus a large optical MIN can be built and effectively used in parallel computing systems.

V. APPROACHES TO SOLVE CROSSTALK

A. Space Domain Approach

One way to solve crosstalk problem is a space domain approach, where a MIN is duplicated and combined to avoid crosstalk [8]. The number of switches required for the same connectivity in networks with space domain approach is slightly larger than twice that for the regular network. This approach uses more than double the original network hardware to achieve the same. Thus for the same permutation the hardware or we can say the no of switches will be double. Thus cost will be more with the networks using space domain approach. In all the four cases only one input and only one output is active at a given time so that no cross talk occurs. With the space domain approach, extra switching elements (SEs) and links are used to ensure that at most one input and one output of every SE will be used at any given time.

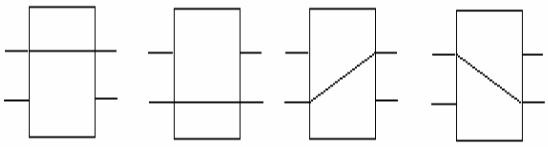


Figure 3. Crosstalk avoidance using space domain approach

B. Time Domain Approach

Another way to solve the problem of crosstalk is the time domain approach [3]. With the time domain approach, the same objective is achieved by treating crosstalk as a conflict; that is, two connections will be established at different times if they use the same SE. Whereas we want to distribute the messages to be sent to the network into several groups, a method is used to find

out which messages should not be in the same group because they will cause crosstalk in the network. A set of connections is partitioned into several subsets such that the connections in each subset can be established simultaneously in a network. There is no crosstalk in these subsections. This approach makes importance in optical MINs for various reasons. First, most of the multiprocessors use electronic processors and optical MINs. There is a big mismatch between the slow processing speed in processors and the high communication speed in networks carrying optical signals [15]. Second, there is a mismatch between the routing control and the fast signal transmission speed. To avoid crosstalk, the TDM approach is used, where the set of messages are partitioned into several groups such that the messages in each group can be sent simultaneously through the network without any crosstalk.

VI. METHODS FOR MESSAGE PARTITIONING IN TDM APPROACH

A. Window method

Window method is the method that is used to find the messages that are not in the same group because it causes crosstalk in the network. If we consider the network of size N x N, there are N source and N destination address. Combining source and its destination address forms combination matrix. From this, optical window size is M -1 where $M = \log_2 N$ and N is size of network. In window method, the number of windows is equal to the number of stages [11].

After finding conflicts using window method, conflict graph is generated shown in figure. The number of nodes is the size of the network. The nodes that are having conflict are connected through edge. Degree of each message is the number of conflicts to the other message. Conflict graph is shown in figure 4.

The conflict matrix is a square matrix with N x N entry, it consists of the output of the window method, as shown in figure 5. The definition of Conflict Matrix is the matrix M_{ij} with size N x N. N is the size of the network.

B. Improved window method

In this method the first window is eliminated for this we make the conflict matrix initialized to 0, here Number of windows is M-1. It takes less time to find conflicts than the windows method. Therefore, it is called improved windows method [11,12].

0 0 0 1 0 1 **message 000 and 100 have conflict**
0 0 1 0 0 1 **message 001 and 101 have conflict**
0 1 0 0 1 1 **message 010 and 110 have conflict**
0 1 1 1 1 0 **message 011 and 111 have conflict**
1 0 0 0 0 0
1 0 1 0 1 0
1 1 0 1 0 0
1 1 1 1 1 1
Step 1(w0)

0 0 0 1 0 1 **message 000 and 110 have conflict**
0 0 1 0 0 1 **message 001 and 101 have conflict**
0 1 0 0 1 1 **message 010 and 100 have conflict**
0 1 1 1 1 0 **message 011 and 111 have conflict**
1 0 0 0 0 0
1 0 1 0 1 0
1 1 0 1 0 0
1 1 1 1 1 1
Step 2(w1)

0 0 0 1 0 1 **message 000 and 110 have conflict**
0 0 1 0 0 1 **message 001 and 100 have conflict**
0 1 0 0 1 1 **message 010 and 101 have conflict**
0 1 1 1 1 0 **message 011 and 111 have conflict**
1 0 0 0 0 0
1 0 1 0 1 0
1 1 0 1 0 0
1 1 1 1 1 1

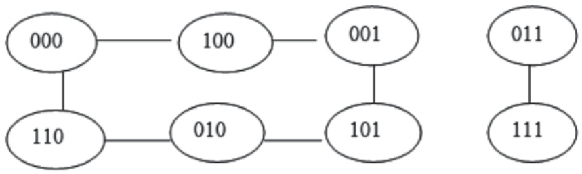


Figure 4. Conflict graph

msg	000	1	0	1	0	1	0	1
0	0	0	0	0	1	0	1	0
1	0	0	0	0	1	1	0	0
0	0	0	0	0	1	1	1	0
1	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0

Figure 5. Conflict matrix

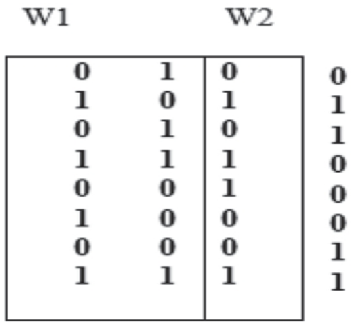
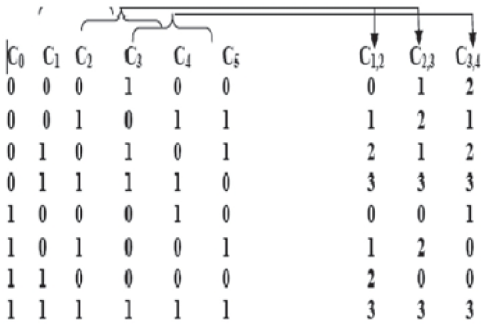


Figure 6. Improved window method

C. Bitwise combination matrix

For Bitwise combination matrix, all binary bits of single rows is each windows are converted to decimal, no. of window is reduced to *n*. By this method, time is reduced approximately by ten times. It is very effective method even when the network is very large.
For example the bitwise combination matrix for 8 8 network size is demonstrated in Fig. 5. The number of columns in WM is 6 (*Ci*, *i* = 2*n*) and for bitwise WM is 2 *(Ci, i = n)* [11,12].



Combination Matrix Bitwise Combination Matrix

D. Bitwise window method

In this method, source and destination address is in decimal format. Number of windows is $\log_2 N$. Thus, from combination matrix, the optical window size is only one for a different network size and the number of window is $\log_2 N$. In other words, there are only one decimal number in each row and each window for comparison and finding a conflict [11,12].

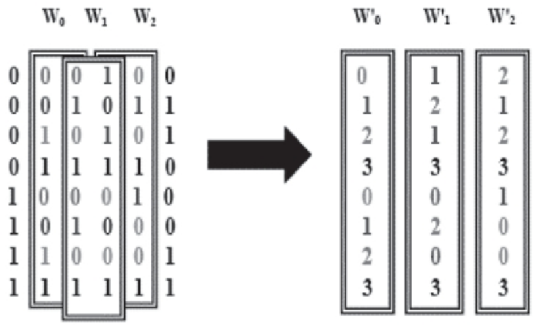


Figure 7. Bitwise window method

VII. CONCLUSION

In this paper properties of electronic and optical MINs have been explained and compared. It is concluded that for today's applications such as in wide area networks (WANs) optical networks is the promising choice to meet the high demand of Speed and Bandwidth. The paper also describes the problems and solutions of optical MINs. Various methods available in literature, which are used for crosstalk avoidance in TDM approach, are described in detail. It has been observed that the improved window method takes lesser time to find conflicts as compared to window method. The bitwise window method reduces the execution time ten times than the other algorithms even when the network size is large.

REFERENCES

[1] A. Verma and C.S. Raghvendra, "Interconnection Networks for Multiprocessors and Mul-ticomputers: Theory and Practice", IEEE Computer Society Press, Los Alamitos, California, 1994.
[2] A.K. Katangur, Y. Pan and M.D Fraser, "Message Routing and Scheduling in Optical Mul-tistage Networks Using Simulated Annealing", International Proceedings of the Parallel and Distributed Processing Symposium (IPDPS), 2002.
[3] C. Qiao, and R. Melhem, "A Time Domain Approach For Avoiding Crosstalk In Optical Blocking Multi-stage Interconnection Networks", Journal of Lightwave Technology, vol. 12 no. 10, October 1994, pp. 1854-1862.
[4] C. Siu and X. Tiehong , "New Algorithm for Message Routing and Scheduling in Optical Multistage Interconnection Network", Proceedings of International Conferwnce on Optical Communications Systems and Networks, 2004.
[5] D. K. Hunter and I. Andonovic, "Guided wave optical switch architectures," International Journal of Optoelectronics, vol. 9, no. 6, 1994, pp. 477-487.
[6] Hasan, "Rearrangeability of (2n − 1)-Stage Shuffle-Exchange Networks", Society for Industrial and Applied Mathematics, vol. 32, no. 3, 2003, pp. 557-585.
[7] J. T. Blaket and K.S. Trivedi, "Reliabilities of Two Fault-Tolerant Interconnection Networks", Proceeding of IEEE, 1988, pp. 300-305.
[8] K. Padmanabhan and A. Netravali, "Dilated Networks for Photonic Switching", IEEE Transactions on Communication, vol. 35, no. 12, 1987, pp. 1357-1365.
[9] L. N. Bhuyan and D.P. Aggarwal, "Design and performance of generalized interconnection networks",

IEEE Transactions on computers, vol. C-32, no. 2, 1983, pp. 1081–1090.
[10] L. N. Bhuyan, Q. Yang Qing and D.P. Aggarwal, "Performance of Multiprocessor Interconnection Networks", IEEE Computers, vol. 22, 1989, pp. 25-37.
[11] M. A, M. Othman and R. Johari, "An efficient approach to avoid crosstalk in optical Omega Network", International Journal of Computer, Internet and Management, vol. 14, no. 1, 2005, pp. 50-60.
[12] M. Ali, M. Othman, R. Johari and S. Subramaniam, "New Algorithm to Avoid Cros- stalk in Optical Multistage Interconnection Networks", Proceedings of IEEE International Conference on Network (MICC-ICON), 2005, pp. 501-504.
[13] M. Fang, Layout Optimization for Point to Multi Point, Wireless Optical Networks via Simulated Annealing & Genetic Algorithm", Master Project, University of Bridgeport, 2000.
[14] Regis Bates, "Optical Switching and Networking Handbook", McGraw-Hill, New York, 2001.
[15] X. Shen, F. Yang and Y. Pan, "Equivalent permutation capabilities between time division optical omega networks and non-optical extra-stage omega networks", IEEE Transactions on Networking, vol.9, no. 4, 2001, pp. 518-524.
[16] Y. Pan , X. Lin, and X. Jia, Evolutionary Approach For Message Scheduling In Optical Omega Networks, Fifth Intern. Conf. on Algorithms and Architectures for Parallel Processing (ICA3PP), 2002.

Biographies

Rinkle Rani Aggarwal, B.Tech (Computer Science & Engg.), M.S. (Software Systems), is Senior Lecturer in Computer Science & Engineering Department at Thapar University, Patiala. She has more than 12 years of teaching experience and served academic institutions such as Guru Nanak Dev Engineering College, Ludhiana and S.S.I.E.T 'Derabassi. She has supervised many M.Tech. Dissertations and contributed 23 articles in Conferences and 11 papers in research Journals. Her areas of interest are Parallel Computing and Algorithms.

Lakhwinder Kaur, Ph.D. is Reader in Computer Engineering at University College of Engineering Punjabi University, Patiala. She has 17 years of teaching experience. She has published 12 research papers in International Journals. Her areas of interest are Image processing, Parallel Computing and Computer Graphics.

Himanshu Aggarwal, Ph.D., is Reader in Computer Engineering at University College of Engineering, Punjabi University, Patiala. He has more than 16 years of teaching experience and served academic institutions such as Thapar Institute of Engineering & Technology, Patiala, Guru Nanak Dev Engineering College, Ludhiana and Technical Teacher's Training Institute, Chandigarh. He is an active researcher who has supervised many M.Tech. Dissertations and contributed 32 articles in Conferences and 13 papers in research Journals. His areas of interest are Information Systems, ERP and Parallel Computing.

Web Based Hindi to Punjabi Machine Translation System

Vishal Goyal and Gurpreet Singh Lehal
Department of Computer Science, Punjabi University, Patiala, Punjab, India
{vishal.pup,gslehal}@gmail.com

***Abstract* - Hindi and Punjabi are closely related languages with lots of similarities in syntax and vocabulary Both Punjabi and Hindi languages have originated from Sanskrit which is one of the oldest language. In terms of speakers, Hindi is third most widely spoken language and Punjabi is twelfth most widely spoken language. Punjabi language is mostly used in the Northern India and in some areas of Pakistan as well as in UK, Canada and USA. Hindi is the national language of India and is spoken and used by the people all over the country. In the present research, Basic Hindi to Punjabi machine translation system using direct translation approach has been developed. The results of this translation system are surprisingly good. The system includes lexicon based translation, transliteration and continuously improving the system through machine learning module. It also takes care of basic word sense disambiguation.**

***Index Terms* - Machine Translation System, Closely related Languages, Hindi, Punjabi, Natural Language Processing, Computational Linguistics, Transliteration.**

I. Introduction

MT is a field of research that has been around since the birth of electronic computers. Warren Weaver, a director of the Rockefeller Foundation received much credit for bringing the concept of MT to the public when he published an influential paper on using computers for translation in 1949. MT is the name for computerized methods that automate all or part of the process of translating from one human language to another. Fully-automatic general purpose high quality machine translation system (FGH-MT) is extremely difficult to build. In fact, there is no system in the world of any pair of languages which qualifies to be called FGH-MT. This paper explains the methodology followed for developing the machine translation system between closely related languages – Hindi and Punjabi. Closely related languages have lots of similarities in syntax and vocabulary. Machine Translation between closely related languages is easier than between language pairs that are not related with each other. Having many parts of their grammars and vocabularies in common reduces the amount of effort needed to develop a translation system between related languages. Closely related languages are the languages of people who have similar cultures and common historical roots.

II. History

The first attempt to verify the hypothesis that related languages are easier to translate started in mid 80s at Charles University in Prague [FEMTI; Hajic et al. 2000]. The project was called RUSLAN and aimed at the translation of documentation in the domain of operating systems for mainframe computers. From that date to till date so many examples are there in history which support the argument that with close languages, the quality of MT system, with simple techniques, is better. To name a few one are CESILKO (a system for translating Czech and Slovak), MT system for translating Turkish To Crimean Tatar etc. We are also trying to strengthen the same concept by experimenting with a word for word direct translation system for Hindi to Punjabi. These languages are very closely related and have many features in common.

III. System Description

The major task behind direct Machine translation system is developing an exhaustive lexicon consisting of source language words along with its corresponding translated version of the target language. There is no machine readable dictionary available for Hindi to Punjabi language. Two traditional dictionaries – one from Bhasha Vibhag, Patiala and another from National Book Trust has been published. We got it digitized and moulded required for machine translation purpose which is itself a big job. Now lexicon consists of approx. 1,00,000 words. This lexicon is used for word for word translation. Extending the lexicon demands large hindi corpus. Sometimes it is available in Unicode format and sometimes it is available in number of other non standard fonts like Susha, Agra, Krutidev etc. which needs to be converted into Unicode first. Conversion is being done through Font Converter that converts non standard fonts into Unicode Format. The beauty of the developed Font converter is that it is able to convert MS-Access files, MS Word files, HTML files and Text Files. Because the corpus can be in any file type, so it handles all the possible file types. Besides translation it performs the task of transliteration as well. Transliteration means to replace character by character of word from source language character to target language character like प्रेमचंद into ਪ੍ਰੇਮਚੰਦ. This system is basically an extension

doi:10.4304/jetwi.2.2.148-151

of previous system that does not handle any word sense disambiguation. The above said system just checks the words to be translated in the dictionary, if found it is replaced with the translated version stored in the dictionary, otherwise it is transliterated. But now in the extended system, the lexicon is divided into two parts – one table consists of words with no disambiguation and second consists of words that have multiple meanings depending upon the context of the word in which it has been used in the sentence.

IV. System Architecture

The architecture for the HPMTS (Hindi to Punjabi Machine Translation System) consists of number of modules that are listed below:

a. Training the system with training corpus
b. Input Text Font Conversion into Unicode Format
c. Hindi Text Normalization
d. Finding and Replacing Collocations
e. Finding and replacing named entities
f. Word to word translation using lexicons
g. Resolving Ambiguity among words
h. Transliteration of words
i. Post Processing
j. Improving the accuracy of the system through machine learning during every translation job.
k. Testing the system using test corpus other than train corpus

In the above architecture, the most important part and starting point is to train the system. Train the system means generating the lexicon using the already existing corpus. The second module is optional and is skipped if the inputted text is already in Unicode format. Unicode Font requirement arises due to internalization of the system and making the system free from specific font dependency. This font converter can be also used for converting the non-Unicode corpus into Unicode format corpus. Indian language words face spelling standardization issues, thereby resulting in multiple spelling variants for the same word. The main reason for this phenomenon can be attributed to the phonetic nature of Indian Languages and multiple dialects. To give an idea of this data problem, these words were found – मंजिल, मन्जिल, मंज़िल Third module is Hindi Text Normalization that solves this spelling variant problem. Hindi text is normalized into standard spellings before it goes for translation. Next Module of the system find and replaces all the collocations using the lexicon enteries. A Collocation is an expression consisting of two or more words that correspond to some conventional way of saying things. Or in the other words of Firth (1957:181) : "Collocations of a given word are statements of the habitual or customary places of that word". This module helps in increasing the accuracy of the translation. Generating Lexicon for Collocations is itself a challenging task. Then comes the turn of the heart of the system – word for word translation uses the lexicon. This search for the Hindi word in the lexicon and replaces it with the corresponding Punjabi translated version present in the lexicon. If this Hindi word is not found in the lexicon it searches that word in the database of ambiguous words, if found using tri-gram approach it resolves the ambiguity of word and replaces it with correct Punjabi meaning among multiple Punjabi meanings. For Example, the hindi word ਸਰੂਪ can be translated into either of the two Punjabi words - ਸਮਾਨ, ਸੁੰਦਰ. But how will the system decide which word to choose is basically to know the context in which the Hindi word ਸਰੂਪ has been used in the sentence. If the word is not found in both the tables it means it is not available in the database and need to be transliterated. For improving the accuracy of the system, this is must to know the system about which new words have been come across and if they have been transliterated accurately or not. If they were not present in the database and need to be present, it is added to lexicon for future translations. If it has been translated wrongly but required one, it is corrected first before adding to the lexicon. In this way this is the ongoing improvement of the system performance during every translation exercise through machine learning module. Post Processing Module takes into consideration some common grammatical mistakes that has been done during translation phases and based on the rules framed, it removes those mistakes and increases the accuracy to the system. Now system has been trained a lot by number of translation exercises, it is time to check the accuracy of the system by testing the system through test data other than the data used for training. Testing the system is also very tedious task. First step in it is to prepare the test cases that covers all the possibilities.

V. Web Based Tool

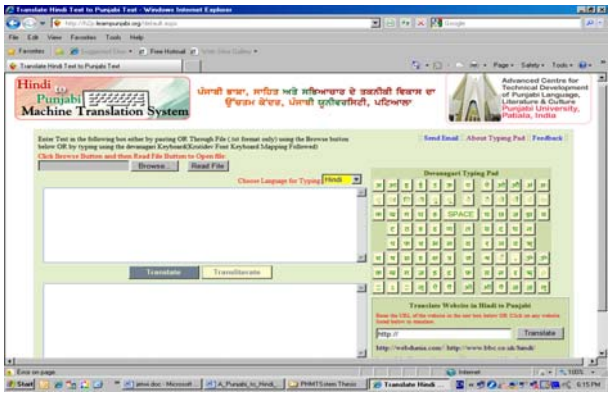
Research must not be restricted to papers, It must be propagated to public for use and test. Taking this aim, the whole system has been developed as a web tool and is online for use fre of cost. The website address is http://h2p.learnpunjabi.org/ . Following are the features of this web tool:

a. Hindi Text can be written in Unicode encoding by using the most popular Hindi Font Krutidev. This concept is very useful for those who are in habit of typing the text in Krutidev and later they find some source for converting them into Unicode encoding. Thus, this feature has solved their purpose in very easy manner. Now, they will type in their style and the typed matter will also be in Unicode.
b. The text can also be input to the system for translation through text file. File can be read using the Browse button provided.
c. Input text can be translated into Punjabi text by just clicking the Translate button. Within seconds, the text is translated into Punjabi.

d. Input Text can be transliterated also, if there is need by clicking on the Transliterate button.
e. Email facility has also been provided. Text can be typed in Hindi and English both. Subject can be written in both the languages. Then while sending email there is an option of sending the email in original or after translating in Punjabi. This is also very powerful feature.
f. The main feature of the webtool is user can translate the full Hindi website into Punjabi website by just providing the link of hindi website and press Translate. The hindi website is translated into Punjabi within seconds and Punjabi version of the website is displayed in the same format as it was originally in hindi website.

VI. Screen Shots

Following is the screen shot of web based machine translation system:



VII. Sample Output

A. Input Text:

भारत और पड़ोस

गुरुवार, **01** नवंबर, **2007** को **08:02 GMT** तक के समाचार

कर्नाटक को लेकर राजनीतिक सरगर्मी बढ़ी

कुमारस्वामी ने राजनीतिक चाल बदलते हुए येदियुरप्पा को मुख्यमंत्री के रुप में समर्थन देने का फ़ैसला किया है

कर्नाटक को लेकर राजनीतिक गहमागहमी तेज़ हो गई है और अब यह मामला दिल्ली आ गया है.

संभावना है कि जल्दी ही कर्नाटक के राजनीतिक भविष्य का कोई फ़ैसला हो जाएगा.

बुधवार को जहाँ भाजपा के उच्च स्तरीय प्रतिनिधि मंडल ने प्रधानमंत्री मनमोहन सिंह से मुलाक़ात की है प्रधानमंत्री और सोनिया गाँधी ने कांग्रेस कोरग्रुप की बैठक में कर्नाटक के राजनीतिक हालात पर चर्चा की है.

उधर भाजपा ने धमकी दी है कि यदि भाजपा-जनतादल (एस) को सरकार बनाने के लिए आमंत्रित नहीं किया गया तो दोनों दलों के **129** विधायक राष्ट्रपति के सामने परेड करेंगे.

उल्लेखनीय है कि जनता दल (एस) और भाजपा का गठबंधन सात अक्तूबर तक सत्ता में था. लेकिन समझौते के अनुसार मुख्यमंत्री बदले जाने के विवाद के चलते गठबंधन टूट गया और भाजपा ने सरकार से समर्थन वापस ले लिया.

B. Translated Output Text:

ਭਾਰਤ ਅਤੇ ਗੁਆਂਢ

ਵੀਰਵਾਰ , **01** ਨਵੰਬਰ , **2007** ਨੂੰ **08 : 02 GMT** ਤੱਕ ਦੀਆਂ ਖ਼ਬਰਾਂ

ਕਰਨਾਟਕ ਨੂੰ ਲੈਕੇ ਰਾਜਨੀਤਕ ਜੋਸ਼ ਵਧਿਆ

ਕੁਮਾਰਸਵਾਮੀ ਨੇ ਰਾਜਨੀਤਕ ਚਾਲ ਬਦਲਦੇ ਹੋਏ ਯੇਦਿਉਰੱਪਾ ਨੂੰ ਮੁੱਖਮੰਤਰੀ ਦੇ ਰੁਪ ਵਿੱਚ ਸਮਰਥਨ ਦੇਣ ਦਾ ਫੈਸਲਾ ਕੀਤਾ ਹੈ

ਕਰਨਾਟਕ ਨੂੰ ਲੈਕੇ ਰਾਜਨੀਤਕ ਗਹਿਮਾਗਹਿਮੀ ਤੇਜ਼ ਹੋ ਗਈ ਹੈ ਅਤੇ ਹੁਣ ਇਹ ਮਾਮਲਾ ਦਿੱਲੀ ਆ ਗਿਆ ਹੈ .

ਸੰਭਾਵਨਾ ਹੈ ਕਿ ਜੱਲਦੀ ਹੀ ਕਰਨਾਟਕ ਦੇ ਰਾਜਨੀਤਕ ਭਵਿੱਖ ਦਾ ਕੋਈ ਫੈਸਲਾ ਹੋ ਜਾਵੇਗਾ .

ਬੁੱਧਵਾਰ ਨੂੰ ਜਿੱਥੇ ਭਾਜਪਾ ਦੇ ਉੱਚ ਪੱਧਰ ਪ੍ਰਤਿਨਿੱਧੀ ਮੰਡਲ ਨੇ ਪ੍ਰਧਾਨਮੰਤਰੀ ਮਨਮੋਹਿਨ ਸਿੰਘ ਨਾਲ ਮੁਲਾਕ਼ਾਤ ਕੀਤੀ ਹੈ ਪ੍ਰਧਾਨਮੰਤਰੀ ਅਤੇ ਸੋਨਿਆ ਗਾਂਧੀ ਨੇ ਕਾਂਗਰਸ ਕੋਰਗਰੁਪ ਦੀ ਬੈਠਕ ਵਿੱਚ ਕਰਨਾਟਕ ਦੇ ਰਾਜਨੀਤਕ ਹਾਲਾਤ ਤੇ ਚਰਚਾ ਕੀਤੀ ਹੈ .

ਓਧਰ ਭਾਜਪਾ ਨੇ ਧਮਕੀ ਦਿੱਤੀ ਹੈ ਕਿ ਜੇਕਰ ਭਾਜਪਾ - ਜਨਤਾਦਲ (ਏਸ) ਨੂੰ ਸਰਕਾਰ ਬਣਾਉਣ ਲਈ ਸੱਦਿਆ ਨਹੀਂ ਕੀਤਾ ਗਿਆ ਤਾਂ ਦੋਨਾਂ ਦਲਾਂ ਦੇ **129** ਵਿਧਾਇਕ ਰਾਸ਼ਟਰਪਤੀ ਦੇ ਸਾਹਮਣੇ ਪਰੇਡ ਕਰਣਗੇ .

ਲਿਖਣ ਯੋਗ ਹੈ ਕਿ ਜਨਤਾ ਦਲ (ਏਸ) ਅਤੇ ਭਾਜਪਾ ਦਾ ਗੰਢ-ਜੋੜਾ ਸੱਤ ਅਕਤੂਬਰ ਤੱਕ ਸੱਤਾ ਵਿੱਚ ਸੀ . ਪਰ ਸਮੱਝੌਤੇ ਦੇ ਅਨੁਸਾਰ ਮੁੱਖਮੰਤਰੀ ਬਦਲੇ ਜਾਣ ਦੇ ਝੱਗੜਾ ਦੇ ਚੱਲ ਦੇ ਗੰਢ-ਜੋੜਾ ਟੁੱਟ ਗਿਆ ਅਤੇ ਭਾਜਪਾ ਨੇ ਸਰਕਾਰ ਨਾਲ ਸਮਰਥਨ ਵਾਪਸ ਲੈ ਲਿਆ .

The accuracy of the translation comes out be approx. 95%.

VIII. Conclusion

The present system is translating any complex sentence. The System accuracy is measured up to 95%. This web tool has number of applications in real world. This web tool can be used by Newspaper agencies, any website owner, using email facilty by community of different countries or regions i.e. Writing the email in Hindi and recipient will receive the email in Punjabi. Thus removing the language bars, communication becomes easy in one’s owm language.

REFERENCES

1. Kemal Altintas, Dept. of Computer Engineering, Bilkent University, Ankara, Turkey, “A Machine Translation System Between a Pair of Closely Related Languages”. Internet: http://www.cs.bilkent.edu.tr/~ilyas/PDF/iscis2002.pdf
2. Bharati A., Chaitanya V and Sangal R, "Natural Language processing: A Paninian Perspective", Prentice Hall of India, New Delhi, 1995.
3. Joseph Seasly, “Machine Translation: A Survey of Approaches”, University of Michigan, Ann Arbor, 2003.
4. Durgesh Rao, “Machine Translation in India: A brief survey”, National Centre for Software Technology, Mumbai, 2001. Internet: http://www.eldra.fr/en/rproj/scalla/SCALLA2001Rao.pdf
5. R.M.K. Sinha, R. Jain and A. Jain "Translation from English to Indian Languages: ANGLABHARTI Approach", Proceeding of STRANS-2002, pp. 69-85, 2002.
6. Christopher D. Manning and Hinrich Schutze, “Foundations of Statistical Natural Language Processing”, MIT Press,1999.
7. Manish Sinha, Mahesh Kumar Reddy, Pushpak Bhattacharya, “Hindi Word Sense Disambiguation”. Internet: http://www.cse.iiitb.ac.in/Pb/papers/HindiWSD.pdf
8. R. Canals-Marote, A. Esteve-Guillén, A. Garrido-Alenda, M.I. Guardiola-Savall, A. Iturraspe-Bellver, S. Montserrat-Buendia, S. Ortiz-Rojas, H. Pastor-Pina, P.M. Pérez-Antón, and M.L. Forcada, “The Spanish_ Catalan machine translation system interNOSTRUM”, Internet: http://internostrum.com/docum/iN-MTS.pdf
9. Kemal Altintas, Ilyas Cicekli,”A Machine Translation System Between a Pair of Closely Related Languages”, Internet: http://www.cs.bilkent.edu.tr/~ilyas/PDF/iscis2002.pdf
10. Kevin P. Scannell ,”Machine translation for closely related language pairs” Internet: http://borel.slu.edu/pub/ga2gd.pdf
11. Jan AJIC, Jan HRIC, Vladislav KUBON, “CESILKO – an MT system for closely related languages”, Internet: http://www.cs.ust.hk/acl2000/Demo/03_kubon.pdf

Protecting Data From the Cyber Theft – A Virulent Disease

[1]Dr. S.N. Panda and [2]Vikram Mangla
[1]Professor & Principal, [2]Assistant Professor
[1] RIMT-IMCT, Mandi Gobind Garh, Punjab.
[2]Chitkara Institue of Engineering & Technology, Rajpura, Punjab.
[1] panda.india@gmail.com, [2] mangla.vikram@gmail.com

***Abstract* - Network security policies are essential elements in Internet security. Network security perimeter devices such as firewalls, IPSec, and IDS/IPS devices operate based on locally configured policies. Malware-related data breaches have reached pandemic proportions as criminals discover that Internet crime is easy to commit, highly lucrative, and largely under-policed. With a few hundred dollars, a cyber criminal can begin a career of breaking into computers to steal identity and confidential data for sale to the highest bidder. This paper will cover current and emerging trends of stealth malware, such as moving primarily to the Web since most organizations allow Web traffic into the network. It will also cover new advances in network security technologies that use multi-phase heuristic and virtual machine analysis to detect and mitigate the damages that result from malware-related data thefts.**

***Index Terms* - Network Security, Web Threats, Malware, Phishing**

I. Introduction

With the global connectivity provided by the Internet, network security has gained significant attention in research and Industrial communities. Due to the increasing threats of network attacks, network security devices such like firewalls and IPSec gatewaye have become important integrated elements not only in enterprise networks but also in small size and home networks. Motivated by the lure of profits from the sale of stolen confidential information, cyber criminals today are shifting to the Web as their chosen attack vector, which provides an ideal environment for cyber crime. Malware-related data breaches have reached pandemic proportions as criminals discover that Internet crime is easy to commit, highly lucrative, and largely under-policed. With a few hundred dollars, a cyber criminal can begin a career of breaking into computers to steal identity and confidential data for sale to the highest bidder. Fraudsters who purchase the data have developed a variety of schemes to monetize that information ranging from transacting unauthorized stock trades to transferring funds to offshore bank accounts. The cyber crime economy is so robust that there is a vibrant market for professional malware toolkits available for $500 to $1,000 and come pre-configured with a range of attack modules, exploit 'maintenance' updates, and 24 x 7 online technical support.

Many Web threats can be deployed unbeknownst to the user, requiring no additional action than merely opening a Web page. Large numbers of users, an assortment of technologies, and a complex network structure provide criminals with the targets, exploitable weaknesses, and anonymity required for large-scale fraud. Web threats pose a broad range of risks, including financial damages, identity theft, and loss of confidential business information, theft of network resources, damaged brand or personal reputation, and erosion of consumer confidence in e-commerce. These high stakes, the pervasive use of the Web, and the complexity of protecting against Web threats combine to form perhaps the greatest challenge to protecting personal and business information in a decade.

In August 2007, a scene played out as cyber criminals infiltrated the monster.com job site through "Monster for Employers" accounts, compromising the personal information of 1.6 million users. Many of these users then received official-looking emails, claiming to be from monster.com and encouraging them to download a "helper application" that turned out to be yet more malware.

These attacks were well-researched, using familiar language and branding, and coded to transfer data slowly, under the radar of IT administrators looking for suspicious network traffic.[1] Web threats also include malware that is downloaded from an email attachment, but accesses the Web to convey information to the hacker. In 2007, fraudulent emails were sent purporting to be from the Federal Trade Commission. These emails claimed that a complaint had been filed against the company and contained an attachment. If the recipient opened the attachment, a keylogging Trojan was deployed that attempted to steal login information from the user's computer and send it back to the hacker. [2].

Phishing is a prevalent Web threat, spoofing legitimate companies to trick people into providing confidential information. Consumer phishing is wide-spread, sending emails that spoof organizations like banks and on-line retailers. These phishing emails often use links to take recipients to Web sites where confidential information is gathered. Employees can fall victim to these consumer threats, but phishing can also affect corporations more directly. In 2005, phishing emails targeted CEOs and other high-level executives of US credit unions in an attempt to gain control of millions of personal financial records. The email messages contained a link to a Web site where a Trojan was downloaded. Even one successful

doi:10.4304/jetwi.2.2.152-155

infection could have caused millions of dollars of damage and caused irreparable harm to hundreds of thousands of users through identity and asset theft. [3]

But Web threats don't just steal confidential information; they can also steal network resources. Variations of e-greeting card spam were sent throughout 2007. These simple spam messages told recipients that a friend had sent them an e-greeting card and to follow the link in the email to view the card. If recipients followed the link, it took them to a Web site that downloaded malicious code.

This code hijacked the computer, turning it into a "bot" and allowing the hackers to use the machine for their own purposes—sending spam, hosting malicious Web sites, and much more. Consumer and corporate computers were infected by the millions. Hackers network these infected computers to create botnets, stealing resources and further perpetuating their fraudulent activities.

II. Web Threats defined

Web threats are any threat that uses the Web to facilitate cyber crime. They are sophisticated in their methods, using multiple types of malware and fraud, all of which utilize HTTP or HTTPS protocols, but can also employ other protocols as components of the attack, such as links in email or IM, or malware in attachments or on servers that access the Web. The creators of such threats frequently update Web site content, variants, and malware types in order to evade detection and achieve greater success.

Web threats based on malware are hidden within Web pages and victims are infected when they visit the page. Fraudulent sites mimic legitimate business Web sites and use social engineering to request visitors to disclose confidential information. Individuals once characterized as hackers, virus writers, spammers, and spy ware makers are now simply known as cyber criminals with financial profit their primary aim.

Over the last 15 years, information security threats have evolved through a series of incarnations. In each case, malware writers and fraudsters sought out the medium that was most used and least protected (for example email). Today, a new wave of threats is emerging that uses the Web as a delivery vehicle. These Web threats are gaining traction at a time when the Web has become a major commerce engine as well as social networking vehicle, with usage continuing to grow.

At the same time, the Web is relatively unprotected, compared to messaging for example, as a medium to deliver malware and conduct fraud. According to IDC, "Up to 30% of companies with 500 or more staff have been infected as a result of Internet surfing, while only 20%-25% of the same companies experienced viruses and worms from emails." [4]

III. Web Threat Delivery Mechanisms

Web threats can be divided into two primary categories, based on delivery method – push and pull. Push based threats use spam, phishing, or other fraudulent means to lure a user to a malicious (often spoofed) Web site, which then collects information and/or injects malware. Push attacks use phishing, DNS poisoning (or pharming), and other means to appear to originate from a trusted source. Their creators have researched their target well enough to spoof corporate logos, official Web site copy, and other convincing evidence to increase the appearance of authenticity. Precisely-targeted push-based threats are often called "spear phishing" to reflect the focus of their data gathering ("phishing") attack.

Spear phishing typically targets specific individuals and groups for financial gain. In November 2006, a medical center fell victim to a spear phishing attack. Employees of the medical center received an email telling them they had been laid off. The email also contained a link that claimed to take the recipient to a career counseling site. Recipients that followed the link were infected by a keylogging Trojan. [5] In other push-based threats, malware authors use social engineering such as enticing email subject lines that reference holidays, popular personalities, sports, pornography, world events, and other popular topics to persuade recipients to open the email and follow links to malicious sites or open attachments with malware that accesses the Web.

Pull-based threats are often referred to as "drive-by" threats, since they can affect any visitor, regardless of precautions. Pull threat developers infect legitimate Web sites, which unknowingly transmit malware to visitors or alter search results to take users to malicious sites. Upon loading the page, the user's browser passively runs a malware downloader in a hidden HTML frame (IFRAME) without any user interaction. Both push- and pull-based Web threat variants target infection at a regional or local level (for example, via local language sites aimed at particular demographics), rather than using the mass infection technique of many earlier malware approaches. These threats typically take advantage of Internet port 80, which is almost always open to permit access to the information, communication, and productivity that the Web affords to employees.

IV. Today's Insider - Threat Is Stealth Malware

Law enforcement, computer crime experts, and even the military are playing catch up to the threat posed to consumers, businesses, and national security as cyber criminals cash in on stolen identity data, fraudulent online transactions, and cyber espionage. It is no surprise that the rise in cyber crime has coincided with the increased use of the Internet and especially "Web 2.0" technologies.

Web sites and applications now support user-contributed content, syndicated content, iframes, third-party widgets (or applets), and convoluted advertising distribution networks into which 'stealth' malware can easily be injected somewhere along the line. In a 2007 USENIX paper, Google researchers determined that approximately 9% of all suspicious web sites launched "drive-by" downloads of stealth malware binaries[12]. Government studies[13] estimate that 65% of all exploits

now enter via the Web and IBM Internet Security Systems (ISS) estimates that nearly 100% of Web attacks now utilize obfuscated JavaScript as a very effective technique to bypass antivirus and intrusion prevention.

Today, once a PC is infected with stealth malware, it typically opens two-way communications to a "command and control" (C&C) server to establish a channel back to the cyber criminal. This allows the "bot" (as in "robot computer") to report status as well as any valuable information that is immediately accessible. Groups of these remotely controlled, malware-infected computers are commonly called botnets, and serve as the foundation of most cybercrime on the Internet.

How do victims get infected? A user may be drawn by a phishing e-mail to a Web site hosted on a hijacked server, which serves up a browser exploit; this downloads and installs a bot on the user's PC. The bot then downloads more malware like "keyloggers" that silently record keyboard and mouse activities to execute further criminal activities, such as stealing user credentials and capturing other sensitive information. All of this takes place without the knowledge of the user or administrator. As their prevalence has increased, remote-control malware/botnets have become serious concerns for security administrators.

The recent January, 2009 malware-related data thefts at Heartland Payment Systems and earlier malware infiltrations at Hannaford Supermarkets, University of Florida Medical Center, and NASA underscore the escalating threat of malware-related data breaches. The Identity Theft Resource Center, a nonprofit group focused on understanding and preventing identity theft, reported that 656 known security breaches had taken place in 2008, reflecting a 47 percent increase over 2007's total. As of March 17, 2009 the resource center had already reported 110 breaches in 2009.

Recent Research[14] Has Found:
11 % of the world's computers are enmeshed in at least one botnet
23 % of home computers become infected despite having security enabled
72 % of corporate networks larger than 100 PC's have an infection

V. Stealth Malware Attacks Are Outmaneuvering Conventional Defenses

Defending corporate networks from today's malware-related data thefts requires modern protection that goes beyond current signature- and heuristic-based detection techniques. Modern threats exploit the inability of conventional network protection to provide a unified defense against a criminal who attacks on multiple fronts, from OS and browser vulnerabilities to social engineering. The anachronistic concept of detecting infections with a single technique, such as signatures, has left many businesses and consumers open to attack, despite their deployment of antivirus and IPS (intrusion prevention systems). The sheer volume and escalating danger of modern attacks are overwhelming limited IT resources and outmaneuvering conventional defenses that may already be in place. To enable a more efficient IT security process, accurate and timely identification of infected machines is the first step in preventing malware-related data breaches. And, the only viable solutions are those that provide thorough coverage across the many vectors that are used in attacks.

VI. Conventional Approaches Fail To Protect Against Web Threats

Web threat scanning has specific requirements that are not met by the traditional approach to virus scanning. Conventional antivirus software installed on client machines, for example, while crucial to the protection of these machines from a variety of threats, does not adequately protect against the evolving set of Web threats. One reason is that the conventional approach to virus protection involves collecting samples of viruses, developing patterns, and quickly distributing these patterns to users. Because many Web threats are targeted attacks and span many variants, collecting samples is almost impossible.

The large numbers of variants use multiple delivery vehicles (for example, spam, instant messaging, and Web sites), rendering the conventional sample collection, pattern creation, and deployment process insufficient. Another reason that conventional virus detection processes fall short involves a fundamental difference between these viruses and evolving Web threats. Conventional viruses were fundamentally designed to spread as quickly as possible, and were therefore often easy to spot. With the advent of Web threats, malware has evolved from this outbreak model to stealthy "sleeper" infections that are therefore difficult to detect via conventional antivirus techniques.

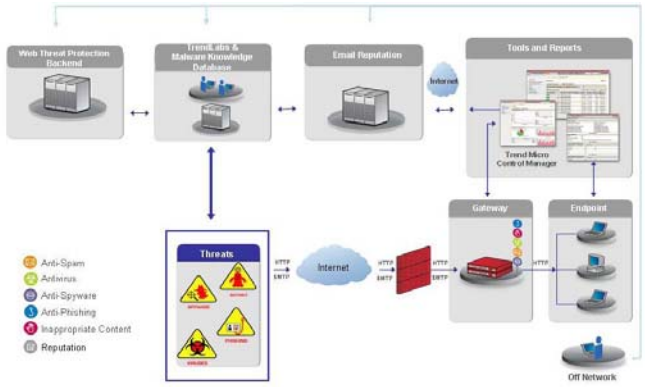
Recovering from infections also presents new challenges. In some cases, Web threats may result in a system infection that is so extensive (for example, via a rootkit in which the system file is replaced) that conventional uninstall or system cleaning approaches become useless. Infected systems often require a complete system recovery, in which the hard drive is wiped and the operating system, applications, and user data are reinstalled.

VII. Future work

A New Approach Is Needed: Integrated, Multi-Layered Protection - Clearly, users need a new approach to addressing Web threats that complements existing techniques.The most effective approach will employ multiple layers of protection and incorporate a range of protective measures. In addition, the evolving nature of the threat necessitates some form of information feedback and integration, in which information gathered in one portion of the protection network is used to update information in other layers. Any effective approach

should also address all relevant protocols, because Web threats leverage multiple protocols in their attacks, in particular email as the initial delivery mechanism and the Web as the threat host. However, other mechanisms can also help perpetrate attacks such as links in IM and infected files.

Coordinating measures requires efficient, centralized management of region-specific expertise to help address the regional, and even localized nature of many of the threats. The key to effectively addressing Web threats is a multi-layered approach. The network points are categorized in four different layers (see Figure 2): 1) "in-the-cloud" (i.e. before the traffic reaches the Internet gateway), 2) at the Internet gateway, 3) across the network servers, 4) and at the endpoint (for example, the client). In the below example, the description uses the points in the network for high level organization and describes the protocol protection and security technologies that can be deployed at these points. The subsections on protocol protection and security technologies describe email solutions first, which is often the first step in a Web threat attack, followed by Web solutions that directly protect Web usage.



A multi-layered approach is needed to protect against the broad range of Web threats

DNA of an Ideal Solution:

Dynamic, real-time detection of threat: Finds the latest stealth, 0-day attacks

Accurate detection: No false positives, and no false negatives

Return on security investment: Easy to install, manage, support and scale

VIII. Conclusion

Web threats are prevalent today and are growing in numbers and impact. Their complexity, large number of variants, and use of multiple vectors, combined with their exploitation of the most commonly used medium today - the Web - make Web threats the most challenging threat that consumers, businesses, and services providers, have faced in a long time.

Potential costs associated with these threats include confidential information leakage and theft of network resources, with the adverse impact of erosion of customers, trust, and brand reputation; regulatory and legal implications; negative public relations; and loss of competitive advantage. Because conventional approaches fail to protect against Web threats, the information security industry is at a crossroads. Businesses of all sizes, as well as service providers, need to deploy solutions via an integrated, multi-layered approach to provide real-time, comprehensive protection against these threats.

References

1. Gregg Keizer, Computerworld, August 19, 2007, "Identity attack spreads; 1.6M records stolen from Monster.com," http://computerworld.com/action/article.do?command=viewArticleBasic&articleId=9031418&pageNumber=1.
2. Dan Kaplan, SC Magazine, October 30, 2007, "FTC Spam Contains Keylogging Trojan", http://www.scmagazineus.com/FTC-spam-contains-keylogging-trojan/article/58273/
3. Paul F. Roberts, eWeek.com, December 16, 2005, "Spear Phishing Attack Targets Credit Unions," http://www.eweek.com/article2/0,1895,1902896,00.asp.
4. IDC, press release, July 18, 2006, "Private Internet Use by Staff Threatens IT Security in Danish Companies, Says IDC," http://www.idc.com/getdoc.jsp?containerId=pr2006_07_14_125434.
5. Cara Garretson, NetworkWorld.com, January 11, 2006, "Spam that Delivers a Pink Slip" http://www.networkworld.com/news/2006/110106-spam-spear-phishing.html
6. Gregg Keizer, TechWeb Technology News, January 24, 2006, "Botnet Creator Pleads Guilty, Faces 25 Years," http://www.techweb.com/wire/security/177103378.
7. Marius Oiaga, Softpedia, October 4, 2006, "Hacking Russian Trio Gets 24 Years in Prison," http://news.softpedia.com/news/Hacking-Russian-Trio-Gets-24-Years-in-Prison-37149.shtml.
8. Byron Acohido and Jon Swartz, USA TODAY "Cybercrime flourishes in online hacker forums," October 11, 2006, http://www.usatoday.com/tech/news/computersecurity/infotheft/2006-10-11-cybercrime-hackerforums_x.htm.
9. Police of the City of Munich, August 25, 2006, http://www.sueddeutsche.de/,tt3m3/muenchen/artikel/612/83529.
10. Avivah Litan, "Phishing Attacks Escalate, Morph, and Cause Considerable Damage," Gartner, December 12, 2007.
11. Tom Krazit, Cnet, "Two in three retail PCs are notebooks," December 20, 2006, http://news.com.com/Two+in+three+retail+PCs+are+notebooks/2100-1044_3-6144921.html.
12. Niels Provos, Dean McNamee, Panayiotis Mavrommatis, Ke Wang, and Nagendra Modadugu: The Ghost in the Browser Analysis of Web-based Malware, May 2007.
13 David Barroso, ENISA Position Paper No. 3: Botnets – The Silent Threat, November 2007, http://www.enisa.europa.eu/doc/pdf/deliverables/enisa_pp_botnets.pdf.
14 Panda Security, http://www.pandasecurity.com/homeusers/media/press-releases/viewnews?noticia=9077

Call for Papers and Special Issues

Aims and Scope

Journal of Emerging Technologies in Web Intelligence (JETWI, ISSN 1798-0461) is a peer reviewed and indexed international journal, aims at gathering the latest advances of various topics in web intelligence and reporting how organizations can gain competitive advantages by applying the different emergent techniques in the real-world scenarios. Papers and studies which couple the intelligence techniques and theories with specific web technology problems are mainly targeted. Survey and tutorial articles that emphasize the research and application of web intelligence in a particular domain are also welcomed. These areas include, but are not limited to, the following:

- Web 3.0
- Enterprise Mashup
- Ambient Intelligence (AmI)
- Situational Applications
- Emerging Web-based Systems
- Ambient Awareness
- Ambient and Ubiquitous Learning
- Ambient Assisted Living
- Telepresence
- Lifelong Integrated Learning
- Smart Environments
- Web 2.0 and Social intelligence
- Context Aware Ubiquitous Computing
- Intelligent Brokers and Mediators
- Web Mining and Farming
- Wisdom Web
- Web Security
- Web Information Filtering and Access Control Models
- Web Services and Semantic Web
- Human-Web Interaction
- Web Technologies and Protocols
- Web Agents and Agent-based Systems
- Agent Self-organization, Learning, and Adaptation
- Agent-based Knowledge Discovery
- Agent-mediated Markets
- Knowledge Grid and Grid intelligence
- Knowledge Management, Networks, and Communities
- Agent Infrastructure and Architecture
- Agent-mediated Markets
- Cooperative Problem Solving
- Distributed Intelligence and Emergent Behavior
- Information Ecology
- Mediators and Middlewares
- Granular Computing for the Web
- Ontology Engineering
- Personalization Techniques
- Semantic Web
- Web based Support Systems
- Web based Information Retrieval Support Systems
- Web Services, Services Discovery & Composition
- Ubiquitous Imaging and Multimedia
- Wearable, Wireless and Mobile e-interfacing
- E-Applications
- Cloud Computing
- Web-Oriented Architectrues

Special Issue Guidelines

Special issues feature specifically aimed and targeted topics of interest contributed by authors responding to a particular Call for Papers or by invitation, edited by guest editor(s). We encourage you to submit proposals for creating special issues in areas that are of interest to the Journal. Preference will be given to proposals that cover some unique aspect of the technology and ones that include subjects that are timely and useful to the readers of the Journal. A Special Issue is typically made of 10 to 15 papers, with each paper 8 to 12 pages of length.

The following information should be included as part of the proposal:

- Proposed title for the Special Issue
- Description of the topic area to be focused upon and justification
- Review process for the selection and rejection of papers.
- Name, contact, position, affiliation, and biography of the Guest Editor(s)
- List of potential reviewers
- Potential authors to the issue
- Tentative time-table for the call for papers and reviews

If a proposal is accepted, the guest editor will be responsible for:

- Preparing the “Call for Papers” to be included on the Journal’s Web site.
- Distribution of the Call for Papers broadly to various mailing lists and sites.
- Getting submissions, arranging review process, making decisions, and carrying out all correspondence with the authors. Authors should be informed the Instructions for Authors.
- Providing us the completed and approved final versions of the papers formatted in the Journal’s style, together with all authors’ contact information.
- Writing a one- or two-page introductory editorial to be published in the Special Issue.

Special Issue for a Conference/Workshop

A special issue for a Conference/Workshop is usually released in association with the committee members of the Conference/Workshop like general chairs and/or program chairs who are appointed as the Guest Editors of the Special Issue. Special Issue for a Conference/Workshop is typically made of 10 to 15 papers, with each paper 8 to 12 pages of length.

Guest Editors are involved in the following steps in guest-editing a Special Issue based on a Conference/Workshop:

- Selecting a Title for the Special Issue, e.g. “Special Issue: Selected Best Papers of XYZ Conference”.
- Sending us a formal “Letter of Intent” for the Special Issue.
- Creating a “Call for Papers” for the Special Issue, posting it on the conference web site, and publicizing it to the conference attendees. Information about the Journal and Academy Publisher can be included in the Call for Papers.
- Establishing criteria for paper selection/rejections. The papers can be nominated based on multiple criteria, e.g. rank in review process plus the evaluation from the Session Chairs and the feedback from the Conference attendees.
- Selecting and inviting submissions, arranging review process, making decisions, and carrying out all correspondence with the authors. Authors should be informed the Author Instructions. Usually, the Proceedings manuscripts should be expanded and enhanced.
- Providing us the completed and approved final versions of the papers formatted in the Journal’s style, together with all authors’ contact information.
- Writing a one- or two-page introductory editorial to be published in the Special Issue.

More information is available on the web site at http://www.academypublisher.com/jetwi/.

(Contents Continued from Back Cover)